Distributions and Fourier Transform

This chapter includes the following sections:
1. Distributions.
2. Convolution.
3. Fourier transform of functions.
4. Extensions of the Fourier transform.
5. Fourier transform and differential equations.
6. Uncertainty principle.

The symbol [Ex] means that the proof is left as exercise. [] means that a proof is missing.

1 Distributions

The theory of distributions was introduced in the 1940s by Laurent Schwartz, who provided a thorough functional formulation to previous ideas of Heaviside, Dirac and others, and forged a powerful tool of calculus. Distributions also offer a solid basis for the construction of Sobolev spaces, that had been introduced by Sobolev in the 1930s using the notion of *weak derivative.* These spaces play a fundamental role in the modern analysis of linear and nonlinear partial differential equations.

We shall denote by Ω a nonempty domain of $\mathbb{R}^N$. The notion of distribution rests upon the idea of regarding any locally integrable function $f : \Omega \to \mathbb{C}$ as a continuous linear functional acting on a topological vector space $\mathcal{T}(\Omega)$ (which must be defined):

$$T_f(v) := \int_\Omega f(x)v(x)\,dx \qquad \forall v \in \mathcal{T}(\Omega). \tag{1.1}$$

One is thus induced to consider all the functionals of the topological dual $\mathcal{T}'(\Omega)$ of $\mathcal{T}(\Omega)$. In this way several classes of distributions are generated. The space $\mathcal{T}(\Omega)$ must be so large that the functional T_f determines a unique f. On the other hand, the smaller is the space $\mathcal{T}(\Omega)$, the larger is its topological dual $\mathcal{T}'(\Omega)$. It happens that there exists a smallest space $\mathcal{T}(\Omega)$, so that $\mathcal{T}'(\Omega)$ is the largest one; the elements of this dual space are what we shall name *distributions.*

In this section we outline some basic tenets of this theory, and provide some tools that we will use ahead.

Test functions. Let Ω be a domain of $\mathbb{R}^N$. By $\mathcal{D}(\Omega)$ we denote the space of infinitely differentiable functions $\Omega \to \mathbb{C}$ whose support is a compact subset of Ω; these are called **test functions.**

The null function is the only analytic function in $\mathcal{D}(\Omega)$, since any element of this space vanishes in some open set. The *bell-shaped* function

$$\rho(x) := \begin{cases} \exp\left[(|x|^2-1)^{-1}\right] & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geq 1 \end{cases} \tag{1.2}$$

belongs to $\mathcal{D}(\mathbb{R}^N)$. By suitably translating ρ and by rescaling w.r.t. x, nontrivial elements of $\mathcal{D}(\Omega)$ are easily constructed for any Ω.

For any $K \subset\subset \Omega$ (i.e., any compact subset K of Ω), let us denote by $\mathcal{D}_K(\Omega)$ the space of the infinitely differentiable functions $\Omega \to \mathbb{C}$ whose support is contained in K. This is a linear subspace of $C^\infty(\Omega)$, and $\mathcal{D}(\Omega) = \bigcup_{K\subset\subset\Omega} \mathcal{D}_K(\Omega)$. The space $\mathcal{D}(\Omega)$ is equipped with the finest topology among those that make all injections $\mathcal{D}_K(\Omega) \to \mathcal{D}(\Omega)$ continuous (so-called *inductive-limit topology).* This topology makes $\mathcal{D}(\Omega)$ a nonmetrizable locally convex Hausdorff space. []

By definition of the inductive-limit topology, a set $A \subset \mathcal{D}(\Omega)$ is open in this topology iff $A \cap \mathcal{D}_K(\Omega)$ is open for any $K \subset\subset \Omega$. Here we shall not study this topology: for our purposes, it will suffice to characterize the corresponding notions of bounded subsets and of convergent sequences.

A subset $B \subset \mathcal{D}(\Omega)$ is bounded in the inductive topology iff it is contained and is bounded in $\mathcal{D}_K(\Omega)$ for some $K \subset\subset \Omega$. [] This means that

(i) there exists a $K \subset\subset \Omega$ that contains the support of all the functions of B, and

(ii) $\sup_{v\in B} \sup_{x\in\Omega} |D^\alpha v(x)| < +\infty$ for any $\alpha \in \mathbb{N}^N$.

As any convergent sequence is necessarily bounded, the following characterization of convergent sequences of $\mathcal{D}(\Omega)$ should be easily understood. A sequence $\{u_n\}$ in $\mathcal{D}(\Omega)$ converges to $u \in \mathcal{D}(\Omega)$ in the inductive topology iff, for some $K \subset\subset \Omega$, $u_n, u \in \mathcal{D}_K(\Omega)$ for any n, and $u_n \to u$ in $\mathcal{D}_K(\Omega)$. [] This means that

(i) there exists $K \subset\subset \Omega$ that contains the support of any u_n and of u, and

(ii) $\sup_{x\in\Omega} |D^\alpha (u_n - u)(x)| \to 0$ for any $\alpha \in \mathbb{N}^N$. [Ex]

For instance, if ρ is defined as in (1.2), then the sequence $\{\rho(\cdot - a_n)\}$ is bounded in $\mathcal{D}(\mathbb{R})$ iff the sequence $\{a_n\}$ is bounded. Moreover $\rho(\cdot - a_n) \to \rho(\cdot - a)$ in $\mathcal{D}(\mathbb{R}^N)$ iff $a_n \to a$. [Ex]

Distributions. All linear and continuous functionals $\mathcal{D}(\Omega) \to \mathbb{C}$ are called **distributions**; these functionals form the (topological) dual space $\mathcal{D}'(\Omega)$. For any $T \in \mathcal{D}'(\Omega)$ and any $v \in \mathcal{D}(\Omega)$ we also write $\langle T, v\rangle$ in place of $T(v)$.

Theorem 1.1 (Characterization of Distributions)
For any linear functional $T : \mathcal{D}(\Omega) \to \mathbb{C}$, the following properties are mutually equivalent:

(i) T is continuous, i.e., $T \in \mathcal{D}'(\Omega)$;

(ii) T is bounded, i.e., it maps bounded subsets of $\mathcal{D}(\Omega)$ to bounded subsets of $\mathbb{C}$;

(iii) T is sequentially continuous, i.e., $T(v_n) \to 0$ whenever $v_n \to 0$ in $\mathcal{D}(\Omega)$;

(iv)

$$\begin{aligned} &\forall K \subset\subset \Omega, \exists m \in \mathbb{N}, \exists C > 0 : \forall v \in \mathcal{D}(\Omega), \\ &\mathrm{supp}(v) \subset K \quad \Rightarrow \quad |T(v)| \le C \max_{|\alpha|\le m} \sup_K |D^\alpha v|. \qquad [] \end{aligned} \tag{1.3}$$

(If m is the smallest integer integer such that the latter condition is fulfilled, one says that T has order m on the compact set K; m may actually depend on K.)

Here are some examples of distributions:

(i) For any $f \in L^1_{\mathrm{loc}}(\Omega)$, the integral functional

$$T_f : v \mapsto \int_\Omega f(x)\, v(x)\, dx \tag{1.4}$$

is a distribution. The mapping $f \mapsto T_f$ is injective, so that we may identify $L^1_{\mathrm{loc}}(\Omega)$ with a subspace of $\mathcal{D}'(\Omega)$. These distributions are called *regular;* the others are called *singular.*

(ii) Let μ be either a complex-valued Borel measure on Ω, or a positive measure on Ω that is finite on any $K \subset\subset \Omega$. In either case the functional

$$T_\mu : v \mapsto \int_\Omega v(x)\, d\mu(x) \tag{1.5}$$

is a distribution, that is usually identified with μ itself. in particular this applies to continuous functions.

(iii) Although the function $x \mapsto 1/x$ is not locally integrable in $\mathbb{R}$, its **principal value** (p.v.),

$$\langle p.v.\ \frac{1}{x}, v\rangle := \lim_{\varepsilon\to 0} \int_{|x|>\varepsilon} \frac{v(x)}{x}\, dx \qquad \forall v \in \mathcal{D}(\mathbb{R}) \tag{1.6}$$

is a distribution. For any $v \in \mathcal{D}(\mathbb{R})$ and for any $a > 0$ such that $\mathrm{supp}(v) \subset [-a, a]$, by the oddness of the function $x \mapsto 1/x$ we have

$$\begin{aligned} \langle p.v.\ \frac{1}{x}, v\rangle &= \lim_{\varepsilon\to 0^+} \left(\int_{\varepsilon<|x|<a} \frac{v(x)-v(0)}{x}\, dx + \int_{\varepsilon<|x|<a} \frac{v(0)}{x}\, dx \right) \\ &= \lim_{\varepsilon\to 0^+} \int_{\varepsilon<|x|<a} \frac{v(x)-v(0)}{x}\, dx = \int_{-a}^{a} \frac{v(x)-v(0)}{x}\, dx. \end{aligned} \tag{1.7}$$

This limit exists and is finite, since by the mean value theorem

$$\left| \int_{\varepsilon<|x|<a} \frac{v(x)-v(0)}{x}\, dx \right| \leq 2a \max_{\mathbb{R}} |v'| \qquad \forall \varepsilon > 0.$$

Notice that the principal value is quite different from other notions of *generalized integral.*

(iv) For any $x_0 \in \Omega$ ($\subset \mathbb{N}^N$) the **Dirac mass** $\delta_{x_0} : v \mapsto v(x_0)$ is a distribution. [Ex] In particular $\delta_0 \in \mathcal{D}'(\mathbb{R})$.

(v) The series of Dirac masses $\sum_{n=1}^{\infty} \delta_{x_n}/n^2$ is a distribution for any sequence $\{x_n\}$ in Ω. [Ex]

(vi) The series $\sum_{n=1}^{\infty} \delta_{x_n}$ is a distribution iff any $K \subset\subset \Omega$ contains at most a finite number of points of the sequence $\{x_n\}$ (i.e., iff $|x_n| \to +\infty$). [Ex] (Indeed, if this condition is fulfilled, whenever any test function is applied to the series this is reduced to a finite sum.) So for instance

$$\sum_{n=1}^{\infty} \delta_n \in \mathcal{D}'(\mathbb{R}), \quad \sum_{n=1}^{\infty} \delta_{1/n} \in \mathcal{D}'(\mathbb{R} \setminus \{0\}), \quad \text{but} \quad \sum_{n=1}^{\infty} \delta_{1/n} \notin \mathcal{D}'(\mathbb{R}).$$

We equip the space $\mathcal{D}'(\Omega)$ with the sequential (weak star) convergence: for any sequence $\{T_n\}$ and any T in $\mathcal{D}'(\Omega)$,

$$T_n \to T \quad \text{in } \mathcal{D}'(\Omega) \quad \Leftrightarrow \quad T_n(v) \to T(v) \quad \forall v \in \mathcal{D}(\Omega). \tag{1.8}$$

This makes $\mathcal{D}'(\Omega)$ a nonmetrizable locally convex Hausdorff space. []

Proposition 1.2 *If $T_n \to T$ in $\mathcal{D}'(\Omega)$ and $v_n \to v$ in $\mathcal{D}(\Omega)$, then $T_n(v_n) \to T(v)$. []*

— Esercizio: di possono calcolare i coefficienti di Fourier della funzione tangente?

Differentiation of distributions. We define the multiplication of a distribution by a C^∞-function and the differentiation [1] of a distribution via **transposition**:

$$\langle fT, v\rangle := \langle T, fv\rangle \qquad \forall T\in\mathcal{D}'(\Omega), \forall f\in C^\infty(\Omega), \forall v\in\mathcal{D}(\Omega)\,, \tag{1.9}$$

$$\langle \tilde{D}^\alpha T, v\rangle := (-1)^{|\alpha|}\langle T, D^\alpha v\rangle \qquad \forall T\in\mathcal{D}'(\Omega), \forall v\in\mathcal{D}(\Omega), \forall\alpha\in\mathbb{N}^N. \tag{1.10}$$

Via the characterization (1.3), it may be checked that $\tilde{D}^\alpha T$ is a distribution, and that the operator $\tilde{D}^\alpha$ is continuous in $\mathcal{D}'(\Omega)$. [Ex] (Actually, by (1.3), the operator $\tilde{D}^\alpha$ may just increase the order of T at most of $|\alpha|$ on any $K\subset\subset\Omega$; see ahead.) Thus any distribution has derivatives of any order. More specifically, for any $f\in C^\infty(\Omega)$, the operators $T\mapsto fT$ and $\tilde{D}^\alpha$ are linear and continuous in $\mathcal{D}'(\Omega)$. [Ex]

The definition (1.9) is consistent with the properties of $L^1_{\rm loc}(\Omega)$. For any $f\in L^1_{\rm loc}(\Omega)$, the definition (1.10) is also consistent with partial integration: if $T=T_f$, (1.10) indeed reads

$$\int_\Omega [D^\alpha f(x)]v(x)\,dx = (-1)^{|\alpha|}\int_\Omega f(x)D^\alpha v(x)\,dx \qquad \forall v\in\mathcal{D}(\Omega), \forall\alpha\in\mathbb{N}^N.$$

(No boundary terms appears as the support of v is compact.)

By (1.10) and as derivatives commute in $\mathcal{D}(\Omega)$, the same applies to $\mathcal{D}'(\Omega)$, that is,

$$\tilde{D}^\alpha\circ\tilde{D}^\beta T = \tilde{D}^{\alpha+\beta}T = \tilde{D}^\beta\circ\tilde{D}^\alpha T \qquad \forall T\in\mathcal{D}'(\Omega), \forall\alpha,\beta\in\mathbb{N}^N. \tag{1.11}$$

The formula of differentiation of the product is extended as follows:

$$\begin{aligned}&\tilde{D}_i(fT) = (D_i f)T + f\tilde{D}_i T\\ &\forall f\in C^\infty(\Omega), \forall T\in\mathcal{D}'(\Omega),\ \text{for } i=1,..,N;\end{aligned} \tag{1.12}$$

in fact, for any $v\in\mathcal{D}(\Omega)$,

$$\begin{aligned}\langle\tilde{D}_i(fT), v\rangle &= -\langle fT, D_i v\rangle = -\langle T, fD_i v\rangle = \langle T, (D_i f)v\rangle - \langle T, D_i(fv)\rangle\\ &= \langle (D_i f)T, v\rangle + \langle\tilde{D}_i T, fv\rangle = \langle (D_i f)T + f\tilde{D}_i T, v\rangle.\end{aligned}$$

A recursive procedure then yields the extension of the classical Leibniz rule:

$$\begin{aligned}&\tilde{D}^\alpha(fT) = \sum_{\beta\le\alpha}\binom{\alpha}{\beta}(D^{\alpha-\beta}f)\tilde{D}^\beta T\\ &\forall f\in C^\infty(\Omega), \forall T\in\mathcal{D}'(\Omega), \forall\alpha\in\mathbb{N}^N. \quad \text{[Ex]}\end{aligned} \tag{1.13}$$

The translation (for $\Omega=\mathbb{R}^N$), the conjugation and other linear operations on functions are also easily extended to distributions via transposition. [Ex]

Comparison with classical derivatives.

Theorem 1.3 (Du-Bois Reymond)
For any $f\in C^0(\Omega)$ and any $i\in\{1,\dots,N\}$, the two following conditions are equivalent:
(i) $\widetilde{D}_i f\in C^0(\Omega)$, [2]
(ii) f is classically differentiable w.r.t. x_i at each point of Ω, and $D_i f\in C^0(\Omega)$.
In either case $\widetilde{D}_i f = D_i f$ in Ω. []

[1] In this section we denote the distributional derivative by $\tilde{D}^\alpha$, and the classical derivative, i.e. the pointwise limit of the difference quotient, by D^α, whenever the latter exists.

[2] that is, $\widetilde{D}_i f$ is a regular distribution that may be identified with a function $h\in C^0(\Omega)\cap L^1_{\rm loc}(\Omega)$. Using the notation (**??**), this condition and the final assertion read $\widetilde{D}_i T_f = T_h$ and $\widetilde{D}_i T_f = T_{D_i f}$ in Ω, respectively.

The next theorem applies to $\Omega :=]a, b[$, for $-\infty \le a < b \le +\infty$. First we remind the reader that

$$\begin{aligned}
&\text{a function } f \in L^1(a,b) \text{ is absolutely continuous iff}\\
&\exists g \in L^1(a,b): \ f(x) = f(y) + \int_y^x g(\xi)\, d\xi \qquad \forall x, y \in]a,b[.
\end{aligned}$$

This entails that $f' = g$ a.e. in $]a, b[$. Thus if $f \in L^1(a, b)$ is absolutely continuous, then it is a.e. differentiable (in the classical sense) and $f' \in L^1(a, b)$.

The converse does not hold: even if f is a.e. differentiable and $f' \in L^1(a, b)$, $f \in L^1(a, b)$ need not be absolutely continuous and $\tilde{D}_i f$ need not be a regular distribution. A counterexample is provided by the Heaviside function H:

$$H(x) := 0 \quad \forall x < 0 \qquad H(x) := 1 \quad \forall x \ge 0. \qquad \text{[Ex]} \tag{1.14}$$

$DH = 0$ a.e. in $\mathbb{R}$, but of course H is not (a.e. equal to) an absolutely continuous function. Notice that $\tilde{D}H = \delta_0$ since

$$\langle \tilde{D}H, v\rangle = -\int_{\mathbb{R}} H(x)\, Dv(x)\, dx = -\int_{\mathbb{R}^+} Dv(x)\, dx = v(0) = \langle \delta_0, v\rangle \quad \forall v \in \mathcal{D}(\mathbb{R}).$$

Theorem 1.4 *For any $f \in L^1(a, b)$, the two following conditions are equivalent:*
(i) $\widetilde{D}f \in L^1(a, b)$,
(ii) f is a.e. equal to an absolutely continuous function.
In either case $\widetilde{D}f = Df$ in Ω. []

Thus, for complex functions of a single variable:
(i) f is of class C^1 iff f and $\widetilde{D}f$ are both continuous,
(ii) f is absolutely continuous iff f and $\widetilde{D}f$ are both locally integrable.

Henceforth all derivatives will be meant in the sense of distributions, if not otherwise stated. We shall denote them by D^α, dropping the tilde.

Examples. (i) $D \log|x| = 1/x$ (in $\mathbb{R}$) in standard calculus, but not in the theory of distributions, as $1/x$ is not locally integrable in any neighbourhood of $x = 0$, and thus it is no distribution. We claim that, for any $v \in \mathcal{D}(\mathbb{R})$ and any $a > 0$ such that $\mathrm{supp}(v) \subset [-a, a]$,

$$D \log|x| = \text{p.v.}\frac{1}{x} \qquad \text{in } \mathcal{D}'(\mathbf{R}). \tag{1.15}$$

Indeed, as the support of any $v \in \mathcal{D}(\mathbb{R})$ is contained in some symmetric interval $[-a, a]$, we have

$$\begin{aligned}
&\langle D\log|x|, v\rangle = -\langle \log|x|, v'\rangle = -\lim_{\varepsilon\to 0^+} \int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} (\log|x|)\, v'(x)\, dx\\
&= \lim_{\varepsilon\to 0^+} \left\{ \int_{[-a,a]\setminus[-\varepsilon,\varepsilon]} \frac{1}{x} v(x)\, dx + (\log|\varepsilon|)\, [v(\varepsilon) - v(-\varepsilon)] \right\}\\
&\left(\text{as} \int_{[-a,a]\setminus[-\varepsilon,\varepsilon]} \frac{v(0)}{x}\, dx = 0 \ \text{ and } \lim_{\varepsilon\to 0^+} (\log|\varepsilon|)\, [v(\varepsilon) - v(-\varepsilon)] = 0\right)\\
&= \lim_{\varepsilon\to 0^+} \int_{[-a,a]\setminus[-\varepsilon,\varepsilon]} \frac{v(x) - v(0)}{x}\, dx = \langle \text{p.v.}\frac{1}{x}, v\rangle.
\end{aligned} \tag{1.16}$$

* (ii) $D[\text{p.v.}(1/x)] \neq -1/x^2$ as the latter is no distribution. Instead, for any $v \in \mathcal{D}(\mathbb{R})$ and any $a > 0$ such that $\text{supp}(v) \subset [-a, a]$, we have

$$
\begin{aligned}
&\langle \widetilde{D}(\text{p.v.}\frac{1}{x}), v\rangle = -\langle \text{p.v.}\frac{1}{x}, v'\rangle \overset{(1.7)}{=} -\int_{-a}^{a} \frac{v'(x) - v'(0)}{x}\, dx \\
&= -\lim_{a\to+\infty} \int_{-a}^{a} \frac{[v(x) - v(0) - xv'(0)]'}{x}\, dx \\
&= (\text{by partial integration}) \;\; -\lim_{a\to+\infty} \int_{-a}^{a} \frac{v(x) - v(0) - xv'(0)}{x^2}\, dx.
\end{aligned}
\tag{1.17}
$$

The latter integral converges, since v has compact support and (by the mean-value theorem) the integrand equals $v''(\xi_x)$, for some ξ_x between 0 and x. (In passing notice that the condition (1.3) is fulfilled.)

* (iii) The even function

$$
f(x) = \frac{\sin(1/|x|)}{|x|} \qquad \text{for a.e. } x \in \mathbb{R} \tag{1.18}
$$

is not locally (Lebesgue)-integrable in $\mathbb{R}$; hence it cannot be identified with a distribution. On the other hand, it is easily seen that the next two limits exist

$$
\begin{aligned}
g(x) &:= \lim_{\varepsilon\to 0+} \int_{\varepsilon}^{x} f(t)\, dt \qquad \forall x > 0, \\
g(x) &:= \lim_{\varepsilon\to 0-} \int_{\varepsilon}^{x} f(t)\, dt \qquad \forall x < 0.
\end{aligned}
\tag{1.19}
$$

That is, $g(x) := \int_0^x f(t)\, dt$, if this is understood as a generalized Riemann integral. Moreover, $g \in L^1_{\text{loc}}(\mathbb{R}) \subset \mathcal{D}'(\mathbb{R})$, so that $Dg \in \mathcal{D}'(\mathbb{R})$; however, Dg cannot be identified with f ($\notin \mathcal{D}'(\mathbb{R})$). Actually, the distribution Dg is a *regularization* of the function f (namely, a distribution T whose restriction to $\mathbb{R} \setminus \{0\}$ coincides with f).

As g is odd and has a finite limit (denoted $g(+\infty)$) at $+\infty$, for any $v \in \mathcal{D}(\mathbb{R})$ and any $a > 0$ such that $\text{supp}(v) \subset [-a, a]$,

$$
\begin{aligned}
\langle Dg, v\rangle &= -\langle g, v'\rangle = -\lim_{b\to+\infty} \int_{-b}^{b} g(x)[v(x) - v(0)]'\, dx \\
&= \lim_{b\to+\infty} \int_{-b}^{b} f(x)[v(x) - v(0)]\, dx + \lim_{b\to+\infty} [g(b) - g(-b)]v(0) \\
&= \int_{-a}^{a} f(x)[v(x) - v(0)]\, dx + 2g(+\infty)v(0) \qquad \forall v \in \mathcal{D}(\mathbb{R}).
\end{aligned}
\tag{1.20}
$$

* (iv) The modifications for the odd function $\tilde{f}(x) = [\sin(1/|x|)]/x$ are left to the reader. □

* **Problems of division.** For any $f \in C^\infty(\mathbb{R}^N)$ and $S \in \mathcal{D}'(\mathbb{R}^N)$, let us consider the problem

$$
\text{find } \; T \in \mathcal{D}'(\mathbb{R}^N) \; \text{ such that } \; fT = S. \tag{1.21}
$$

(This is named a *problem of division,* since formally $T = S/f$.) The general solution may be represented as the sum of a particular solution of the nonhomogeneous equation and the general solution of the homogeneous equation $fT_0 = 0$. The latter may depend on a number of arbitrary constants.

If f does not vanish in $\mathbb{R}^N$, then $1/f \in C^\infty(\mathbb{R}^N)$ and (1.21) has one and only one solution: $T = (1/f)S$. On the other hand, if f vanishes at some points of $\mathbb{R}^N$, the solution is less trivial. Let us see the case of $N = 1$, along the lines of [Gilardi: Analisi 3]. For instance, if $f(x) = x^m$ (with $m \in \mathbb{N}$), then the homogeneous equation $x^m T = 0$ has the general solution $T_0 = \sum_{n=0}^{m-1} c_n D^n \delta_0$, with $c_n \in \mathbb{C}$ for any n. [Ex] On the other hand, even the simple-looking equation $x^m T = 1$ is more demanding: notice that $x^{-m} \notin \mathcal{D}'(\mathbb{R})$ for any integer $m \geq 1$.

Support and order of distributions. For any open set $\widetilde{\Omega} \subset \Omega$ and any $T \in \mathcal{D}'(\Omega)$, we define the restriction of T to $\widetilde{\Omega}$, denoted $T\big|_{\widetilde{\Omega}}$, by

$$\langle T\big|_{\widetilde{\Omega}}, v\rangle := \langle T, v\rangle \qquad \forall v \in \mathcal{D}(\Omega) \text{ such that } \operatorname{supp}(v) \subset \widetilde{\Omega}.$$

Because of Theorem 1.1, $T\big|_{\widetilde{\Omega}} \in \mathcal{D}'(\widetilde{\Omega})$.

A distribution $T \in \mathcal{D}'(\Omega)$ is said to vanish in an open subset $\tilde{\Omega}$ of Ω iff it vanishes on any function of $\mathcal{D}(\Omega)$ supported in $\tilde{\Omega}$. Notice that, for any triplet of Euclidean domains $\Omega_1, \Omega_2, \Omega_3$,

$$\Omega_1 \subset \Omega_2 \subset \Omega_3 \quad \Rightarrow \quad \big(T\big|_{\Omega_2}\big)\big|_{\Omega_1} = T\big|_{\Omega_1} \qquad \forall T \in \mathcal{D}'(\Omega_3). \tag{1.22}$$

There exists then a (possibly empty) largest open set $A \subset \Omega$ in which T vanishes. [Ex] Its complement in Ω is called the **support** of T, and will be denoted by $\operatorname{supp}(T)$.

For any $K \subset\subset \Omega$, the smallest integer m that fulfills the estimate (1.3) is called the order of T in K. The supremum of these orders is called the **order** of T; each distribution is thus of either finite or infinite order. For instance,

(i) regular distributions and the Dirac mass are of order zero; [Ex]

(ii) $D^\alpha \delta_0$ is of order $|\alpha|$ for any $\alpha \in \mathbb{N}^N$;

(iii) p.v. $(1/x)$ is of order one in $\mathcal{D}'(\mathbb{R})$. [Ex]

On the other hand, $\sum_{n=1}^\infty D^n \delta_n$ is of infinite order in $\mathcal{D}'(\mathbb{R})$.

The next statement directly follows from (1.3).

Theorem 1.5 *Any compactly supported distribution is of finite order.*

The next theorem is also relevant, and will be applied ahead.

Theorem 1.6 *Any distribution whose support is the origin is a finite combination of derivatives of the Dirac mass. []*

The space $\boldsymbol{\mathcal{E}(\Omega)}$ and its dual. In his theory of distributions, Laurent Schwartz denoted by $\mathcal{E}(\Omega)$ the space $C^\infty(\Omega)$, equipped with the family of seminorms

$$|v|_{K,\alpha} := \sup_{x \in K} |D^\alpha v(x)| \qquad \forall K \subset\subset \Omega, \forall \alpha \in \mathbb{N}^N.$$

This renders $\mathcal{E}(\Omega)$ a locally convex Frèchet space, and induces the topology of uniform convergence of all derivatives on any compact subset of Ω: for any sequence $\{u_n\}$ in $\mathcal{E}(\Omega)$ and any $u \in \mathcal{E}$,

$$\begin{aligned} &u_n \to u \quad \text{ in } \mathcal{E}(\Omega) \quad \Leftrightarrow \\ &\sup_{x \in K} |D^\alpha (u_n - u)(x)| \to 0 \quad \forall K \subset\subset \Omega, \quad \forall \alpha \in \mathbb{N}^N. \end{aligned} \tag{1.23}$$

Notice that

$$\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega) \quad \text{with continuous and sequentially dense injection,} \tag{1.24}$$

namely, any element of $\mathcal{E}(\Omega)$ may be approximated by a sequence of $\mathcal{D}(\Omega)$. This may be checked via multiplication by a suitable sequence of compactly supported smooth functions. [Ex] By (1.24)

$$\mathcal{E}'(\Omega) \subset \mathcal{D}'(\Omega) \quad \text{with continuous and sequentially dense injection,} \tag{1.25}$$

so that we may identify $\mathcal{E}'(\Omega)$ with a subspace of $\mathcal{D}'(\Omega)$.

As we did for $\mathcal{D}'(\Omega)$, we shall equip the space $\mathcal{E}'(\Omega)$ with the sequential weak star convergence: for any sequence $\{T_n\}$ in $\mathcal{E}'(\Omega)$ and any $T \in \mathcal{E}'(\Omega)$,

$$T_n \to T \quad \text{ in } \mathcal{E}'(\Omega) \quad \Leftrightarrow \quad T_n(v) \to T(v) \quad \forall v \in \mathcal{E}(\Omega). \tag{1.26}$$

[This makes $\mathcal{E}'(\Omega)$ a nonmetrizable locally convex Hausdorff space.]

The sequential weak star convergence of $\mathcal{E}'(\Omega)$ is strictly stronger than that induced by $\mathcal{D}'(\Omega)$: for any sequence $\{T_n\}$ in $\mathcal{E}'(\Omega)$ and any $T \in \mathcal{E}'(\Omega)$,

$$T_n \to T \quad \text{ in } \mathcal{E}'(\Omega) \quad \overset{\not\Leftarrow}{\Rightarrow} \quad T_n \to T \quad \text{ in } \mathcal{D}'(\Omega).\text{[Ex]} \tag{1.27}$$

If $\Omega = \mathbb{R}$, the sequence $\{\chi_{[n,n+1]}\}$ (the characteristic functions of the intervals $[n, n+1]$) is a counterexample to the converse implication:

$$\chi_{[n,n+1]} \to 0 \qquad \text{in } \mathcal{D}'(\mathbb{R}^N) \text{ but not in } \mathcal{E}'(\mathbb{R}^N).$$

Theorem 1.7 $\mathcal{E}'(\Omega)$ *may be identified with the subspace of distributions having compact support.*

We just outline a part of the argument. Let $T \in \mathcal{D}'(\Omega)$ have support $K \subset\subset \Omega$. For any $v \in \mathcal{E}(\Omega)$, multiplying it by χ_K and then convoluting with a regularizing kernel ρ (see (1.2)), one may construct $v_0 \in \mathcal{D}(\Omega)$ such that $v_0 = v$ in K. [Ex] One may thus define $\tilde{T}(v)$ by setting $\tilde{T}(v) = T(v_0)$. It is easily checked that this determines a unique $\tilde{T} \in \mathcal{E}'(\Omega)$. Compactly supported distributions may thus be identified with certain elements of $\mathcal{E}'(\Omega)$.

The proof of the surjectivity of the mapping $T \mapsto \tilde{T}$ is less straightforward, and is here omitted.

On the basis of the latter theorem, examples of elements of $\mathcal{E}'(\Omega)$ are easily provided. E.g.:

(i) any compactly supported $f \in L^1_{\text{loc}}$ belongs to $\mathcal{E}'(\Omega)$,

(ii) $\sum_{n=1}^{m} D^{\alpha_n}\delta_{a_n} \in \mathcal{E}'(\Omega)$, for any finite families $a_1, ..., a_m \in \Omega$ and $\alpha_1, ..., \alpha_m \in \mathbb{N}^N$,

(iii) $\sum_{n=1}^{\infty} n^{-2} D^{\alpha_n}\delta_{a_n} \in \mathcal{E}'(\Omega)$, for any sequence $\{a_n\}$ contained in a compact subset of Ω, and any bounded sequence of multi-indices $\{\alpha_n\}$. (This is no element of $\mathcal{E}'(\Omega)$ if either the coefficients n^{-2} are dropped, or the sequence $\{a_n\}$ is not confined to a compact subset of Ω, or the sequence of multi-indices $\{\alpha_n\}$ is unbounded.)

On the basis of the latter theorem, we may apply to $\mathcal{E}'(\Omega)$ the operations that we defined for distributions. It is straightforward to check that this space is stable by differentiation, multiplication by a smooth function, and so on.

The space $\mathcal{S}$ of rapidly decreasing functions. In order to extend the Fourier transform to distributions, Laurent Schwartz introduced the space of **(infinitely differentiable) rapidly decreasing functions** (at ∞): [3]

$$\begin{aligned} \mathcal{S}\big(\mathbb{R}^N\big) :=& \big\{v \in C^\infty : \forall \alpha, \beta \in \mathbb{N}^N, x^\beta D^\alpha v \in L^\infty\big\} \\ =& \big\{v \in C^\infty : \forall \alpha \in \mathbb{N}^N, \forall m \in \mathbb{N}, \\ & \ |x|^m D^\alpha v(x) \to 0 \text{ as } |x| \to +\infty\big\}. \end{aligned} \tag{1.28}$$

(The latter equality is easily checked.) [Ex] We shall write $\mathcal{S}$ in place of $\mathcal{S}\big(\mathbb{R}^N\big)$. This is a locally convex Fréchet space equipped with either of the following equivalent families of seminorms []

$$|v|_{\alpha,\beta} := \sup_{x \in \mathbb{R}^N} |x^\beta D^\alpha v(x)| \qquad \alpha, \beta \in \mathbb{N}^N, \tag{1.29}$$

$$|v|_{m,\alpha} := \sup_{x \in \mathbb{R}^N} (1 + |x|^2)^m |D^\alpha v(x)| \qquad m \in \mathbb{N}, \alpha \in \mathbb{N}^N. \tag{1.30}$$

For instance, for any $\theta \in C^\infty$ such that $\theta(x)/|x|^a \to +\infty$ as $|x| \to +\infty$ for some $a > 0$, $e^{-\theta(x)} \in \mathcal{S}$. By the Leibniz rule, for any polynomials P and Q, the operators

$$u \mapsto P(x)Q(D)u, \qquad u \mapsto P(D)[Q(x)u] \tag{1.31}$$

map $\mathcal{S}$ to $\mathcal{S}$ and are continuous. [Ex] It is easily checked that

$$\mathcal{D} \subset \mathcal{S} \subset \mathcal{E} \quad \text{with continuous and sequentially dense injections.} \tag{1.32}$$

The space $\mathcal{S}'$ of tempered distributions. We shall denote the (topological) dual space of $\mathcal{S}$ by $\mathcal{S}'$. As $\mathcal{S}$ is a metric space, this is the space of the linear functionals $T : \mathcal{S} \to \mathbb{C}$ such that

$$\{v_n\} \subset \mathcal{S}, \quad v_n \to 0 \quad \text{in } \mathcal{S} \quad \Rightarrow \quad \langle T, v_n \rangle \to 0. \tag{1.33}$$

The elements of this space are named **tempered distributions:** we shall see that actually $\mathcal{S}' \subset \mathcal{D}'$ (up to identifications) with continuous injection. Here are some examples:
(i) any compactly supported $T \in \mathcal{D}'(\Omega)$,
(ii) any $f \in L^p$ with $p \in [1, +\infty]$ (since, by the Hölder inequality, $(1 + |x|)^{-a} f \in L^1$ for any $a > 1/p'$, p' being the conjugate index of p),
(iii) any function f such that $|f(x)| \le C(1 + |x|)^m$ for some $C > 0$ and $m \in \mathbb{N}$,
(iv) $f(x) = p(x)w(x)$, for any polynomial p and any $w \in L^1$. [Ex]
On the other hand L^1_{loc} is not included in $\mathcal{S}'$. E.g., $e^{|x|} \notin \mathcal{S}'$. Nevertheless ahead we shall see that $e^x \cos(e^x) \in \mathcal{S}'$ for $N = 1$, at variance with what might be expected.

Convergence in $\mathcal{S}'$. As we did for $\mathcal{D}'(\Omega)$ and $\mathcal{E}'(\Omega)$, we shall equip the space $\mathcal{S}'$ with the sequential weak star convergence: for any sequence $\{T_n\}$ in $\mathcal{S}'$ and any $T \in \mathcal{S}'$,

$$T_n \to T \quad \text{in } \mathcal{S}' \quad \Leftrightarrow \quad T_n(v) \to T(v) \quad \forall v \in \mathcal{S}. \tag{1.34}$$

[3] Laurent Schwartz founded the theory of distributions upon the dual of three main function spaces: $\mathcal{D}(\Omega)$, $\mathcal{E}(\Omega)$ and $\mathcal{S}(\mathbb{R}^N)$. The two latter are Fréchet space, at variance with the first one and with the respective (topological) duals.
Notice that this does not subsume any monotonicity property; e.g., the nonmonotone function $e^{-|x|^2} \sin x$ is an element of $\mathcal{S}(\mathbb{R})$.

[This makes $\mathcal{S}'$ a nonmetrizable locally convex Hausdorff space.]

As $\mathcal{D} \subset \mathcal{S}' \subset \mathcal{D}'$ and $\mathcal{D}$ is a sequentially dense subset of $\mathcal{D}'$, it follows that

$$\mathcal{S}' \subset \mathcal{D}' \quad \text{with continuous and sequentially dense injection; [Ex]} \tag{1.35}$$

namely, any element of $\mathcal{D}'$ may be approximated by a sequence of $\mathcal{S}'$. The sequential weak star convergence of $\mathcal{S}'$ is strictly stronger than that induced by $\mathcal{D}'$: for any sequence $\{T_n\}$ in $\mathcal{S}'$ and any $T \in \mathcal{S}'$,

$$T_n \to T \quad \text{in } \mathcal{S}' \quad \overset{\not\Leftarrow}{\Rightarrow} \quad T_n \to T \quad \text{in } \mathcal{D}'.\text{[Ex]} \tag{1.36}$$

In $\mathbb{R}$, $\{e^{|x|}\chi_{[n,n+1]}\}$ is a counterexample to the converse implication:

$$e^{|x|}\chi_{[n,n+1]} \to 0 \qquad \text{in } \mathcal{D}' \text{ but not in } \mathcal{S}'. \tag{1.37}$$

On the other hand L^1_{loc} is not included in $\mathcal{S}'$, not even for $N = 1$. E.g., $e^{|x|} \notin \mathcal{S}'$.

As $\mathcal{S} \subset \mathcal{E}$ with sequentially dense inclusion, it follows that

$$\mathcal{E}' \subset \mathcal{S}' \quad \text{with continuous and sequentially dense injection; [Ex]} \tag{1.38}$$

Because of (1.38), we may apply to $\mathcal{S}'$ the operations that we defined for distributions. It is straightforward to check that this space is stable by differentiation, multiplication by smooth functions, and so on.

Overview of distribution spaces. We introduced the spaces $\mathcal{D}(\Omega), \mathcal{E}(\Omega)$, with (up to identifications)

$$\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega) \quad \text{with continuous and dense injection.} \tag{1.39}$$

For $\Omega = \mathbb{R}^N$ (which is not displayed), we also defined $\mathcal{S}$, which is such that

$$\mathcal{D} \subset \mathcal{S} \subset \mathcal{E} \quad \text{with continuous and dense injection.} \tag{1.40}$$

We equipped the respective dual spaces with the weak star convergence. (1.39) and (1.40) respectively yield

$$\mathcal{E}'(\Omega) \subset \mathcal{D}'(\Omega) \quad \text{with continuous and sequentially dense injection,} \tag{1.41}$$

and, for $\Omega = \mathbb{R}^N$,

$$\mathcal{E}' \subset \mathcal{S}' \subset \mathcal{D}' \quad \text{with continuous and sequentially dense injection.} \tag{1.42}$$

L. Schwartz also introduced spaces of slowly increasing functions and rapidly decreasing distributions. But we shall not delve on them.

Exercises.

2 Convolution

Convolution of L^1-functions. For any measurable functions $f, g : \mathbb{R}^N \to \mathbb{C}$, we call **convolution product** (or just **convolution)** of f and g the function

$$(f * g)(x) := \int f(x-y)g(y)\,dy \qquad \text{for a.e. } x \in \mathbb{R}^N, \tag{2.43}$$

whenever this integral converges (absolutely) for a.e. x. (We write $\int ...dy$ in place of $\int ... \int_{\mathbb{R}^N} ... dy_1...dy_N$, and omit to display the domain $\mathbb{R}^N$.) Note that

$$\mathrm{supp}(f * g) \subset \overline{\mathrm{supp}(f) + \mathrm{supp}(g)}. \quad \text{[Ex]} \tag{2.44}$$

Henceforth, whenever A and B are two topological vector spaces of functions for which the convolution makes sense, we set $A * B := \{f * g : f \in A, g \in B\}$, and define $A \cdot B$ similarly.

Proposition 2.1 *(i) $L^1 * L^1 \subset L^1$, and*

$$\|f * g\|_{L^1} \le \|f\|_{L^1}\|g\|_{L^1} \qquad \forall f, g \in L^1\,. \tag{2.45}$$

(ii) $L^1_{\mathrm{loc}} * L^1_{\mathrm{comp}} \subset L^1_{\mathrm{loc}}$, and [4]

$$\begin{aligned} &\|f * g\|_{L^1(K)} \le \|f\|_{L^1(K-\mathrm{supp}(g))}\|g\|_{L^1} \\ &\forall K \subset\subset \mathbb{R}^N, \forall f \in L^1_{\mathrm{loc}}, \forall g \in L^1_{\mathrm{comp}}. \end{aligned} \tag{2.46}$$

Moreover $L^1_{\mathrm{comp}} * L^1_{\mathrm{comp}} \subset L^1_{\mathrm{comp}}$.

(iii) For $N = 1$, $L^1_{\mathrm{loc}}(\mathbb{R}^+) * L^1_{\mathrm{loc}}(\mathbb{R}^+) \subset L^1_{\mathrm{loc}}(\mathbb{R}^+)$. [5] *For any $f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^+)$,*

$$(f * g)(x) = \begin{cases} \int_0^x f(x-y)g(y)\,dy & \text{for a.e. } x \ge 0 \\ 0 & \text{for a.e. } x < 0, \end{cases} \tag{2.47}$$

$$\|f * g\|_{L^1(0,M)} \le \|f\|_{L^1(0,M)}\|g\|_{L^1(0,M)} \qquad \forall M > 0. \tag{2.48}$$

The mapping $(f, g) \mapsto f * g$ is thus continuous in each of these three cases.

Proof. (i) For any $f, g \in L^1$, the function $(\mathbb{R}^N)^2 \to \mathbb{C} : (z, y) \mapsto f(z)g(y)$ is (absolutely) integrable, and by changing integration variable we get

$$\iint f(z)g(y)\,dz\,dy = \iint f(x-y)g(y)\,dy\,dx.$$

By Fubini's theorem the function $f * g : x \mapsto \int f(x-y)g(y)\,dy$ is then integrable. Moreover

$$\begin{aligned} \|f * g\|_{L^1} &= \int dx \left|\int f(x-y)g(y)\,dy\right| \\ &\le \iint |f(x-y)||g(y)|\,dx\,dy = \iint |f(z)||g(y)|\,dz\,dy = \|f\|_{L^1}\|g\|_{L^1}. \end{aligned}$$

[4] By L^1_{comp} we denote the space of the functions $v \in L^1$ that have compact support. The support of a measurable function $v : \Omega \to \mathbb{R}$ is the complement in Ω of the set of the points that have a neighborhood in which v vanishes a.e..

[5] Any function or distribution defined on $\mathbb{R}^+$ will be automatically extended to the whole $\mathbb{R}$ with value 0. (In signal theory, the functions of time that vanish for any $t < 0$ are said *causal).*

(ii) For any $f \in L^1_{\text{loc}}$ and $g \in L^1_{\text{comp}}$, setting $S_g := \text{supp}(g)$,

$$(f * g)(x) = \int_{S_g} f(x-y)g(y)\,dy \qquad \text{for a.e. } x \in \mathbb{R}^N.$$

Moreover, for any $K \subset\subset \mathbb{R}^N$,

$$\begin{aligned}\|f * g\|_{L^1(K)} &\le \int_K dx \int_{S_g} |f(x-y)g(y)|\,dy = \int_{S_g} dy \int_K |f(x-y)g(y)|\,dx \\ &= \int_{S_g} dy \int_{K-S_g} |f(z)g(y)|\,dz \le \|f\|_{L^1(K-S_g)}\|g\|_{L^1}.\end{aligned}$$

The proof of the inclusion $L^1_{\text{comp}} * L^1_{\text{comp}} \subset L^1_{\text{comp}}$ is based on (2.44), and is left to the Reader.

(iii) Part (iii) may be proved by means of an argument similar to that of part (ii), that we also leave to the reader. □

Proposition 2.2 *L^1, L^1_{comp} and $L^1_{\text{loc}}(\mathbb{R}^+)$, equipped with the convolution product, are commutative algebras (without unit).* [6] *In particular,*

$$\begin{aligned}&f * g = g * f, \qquad (f * g) * h = f * (g * h) \qquad \text{a.e. in } \mathbb{R}^N \\ &\forall (f,g,h) \in (L^1)^3 \cup \big(L^1_{\text{loc}} \times L^1_{\text{comp}} \times L^1_{\text{comp}}\big).\end{aligned} \tag{2.49}$$

If $N = 1$, the same holds for any $(f,g,h) \in L^1_{\text{loc}}(\mathbb{R}^+)^3$, too.
The mapping $(f,g,h) \mapsto f * g * h$ is continuous for any choice of the above spaces.

Proof. For any $(f,g,h) \in (L^1)^3$ and a.e. $x \in \mathbb{R}^N$,

$$\begin{aligned}(f * g)(x) &= \int f(x-y)g(y)\,dy = \int f(z)g(x-z)dz = (g * f)(x), \\ [(f * g) * h](x) &= \int [(f * g)](z)\,h(x-z)\,dz = \int dz \int f(y)g(z-y)\,dy\ h(x-z) \\ &= \iint f(y)g(t)h((x-y)-t)\,dt\,dy \\ &= \int dy\, f(y) \int g(t)h(x-y-t)\,dt \\ &= \int f(y)[(g * h)](x-y)\,dy = [f * (g * h)](x).\end{aligned}$$

[6] ***** Let a linear space X over a field $\mathbb{K}$ ($= \mathbb{C}$ or $\mathbb{R}$) be equipped with a product $* : X \times X \to X$. This is called an **algebra** iff, for any $u, v, z \in X$ and any $\lambda \in \mathbb{K}$:

(i) $u * (v * z) = (u * v) * z$,

(ii) $(u + v) * z = u * z + v * z, \quad z * (u + v) = z * u + z * v$,

(iii) $\lambda(u * v) = (\lambda u) * v = u * (\lambda v)$.

The algebra is said **commutative** iff the product $*$ is commutative.

X is called a **Banach algebra** iff it is both an algebra and a Banach space (over the same field), and, denoting the norm by $\|\cdot\|$,

(iv) $\|u * v\| \le \|u\|\|v\|$ for any $u, v \in X$.

X is called a **Banach algebra with unit** iff

(v) there exists (a necessarily unique) $e \in X$ such that $\|e\| = 1$, and $e * u = u * e = u$ for any $u \in X$.

(If the unit is missing, it may be constructed in a canonical way...)

The cases of $(f,g,h) \in \left(L^1_{\rm loc} \times L^1_{\rm comp} \times L^1_{\rm comp}\right)$ and $(f,g,h) \in L^1_{\rm loc}(\mathbb{R}^+)^3$ are analogously checked. [Ex] $\square$

It is easily seen that $(L^1, *)$ and $(L^\infty, \cdot)$ (here "$\cdot$" stands for the pointwise product) are commutative Banach algebras; $(L^\infty, \cdot)$ has the unit element $e \equiv 1$.

Convolution of L^p-functions. The following result generalizes Proposition 2.1. [7]

• **Theorem 2.3 (Young)** *Let*

$$p, q, r \in [1, +\infty], \qquad p^{-1} + q^{-1} = 1 + r^{-1}.\,^{8} \tag{2.51}$$

Then: (i) $L^p * L^q \subset L^r$ and

$$\|f * g\|_{L^r} \leq \|f\|_{L^p} \|g\|_{L^q} \qquad \forall f \in L^p, \forall g \in L^q. \tag{2.52}$$

(ii) $L^p_{\rm loc} * L^q_{\rm comp} \subset L^r_{\rm loc}$ and

$$\begin{aligned} &\|f * g\|_{L^r(K)} \leq \|f\|_{L^p(K - {\rm supp}(g))} \|g\|_{L^q} \\ &\forall K \subset\subset \mathbb{R}^N, \forall f \in L^p_{\rm loc}, \forall g \in L^q_{\rm comp}. \end{aligned} \tag{2.53}$$

Moreover $L^p_{\rm comp} * L^q_{\rm comp} \subset L^r_{\rm comp}$.

(iii) For $N = 1$, $L^p_{\rm loc}(\mathbb{R}^+) * L^q_{\rm loc}(\mathbb{R}^+) \subset L^r_{\rm loc}(\mathbb{R}^+)$, and

$$\begin{aligned} &\|f * g\|_{L^r(0,M)} \leq \|f\|_{L^p(0,M)} \|g\|_{L^q(0,M)} \\ &\forall M > 0, \forall f \in L^p_{\rm loc}(\mathbb{R}^+), \forall g \in L^q_{\rm loc}(\mathbb{R}^+). \end{aligned} \tag{2.54}$$

The mapping $(f, g) \mapsto f * g$ is thus continuous in each of these three cases.

* *Proof.* (i) If $p = +\infty$, then by (2.8) $q = 1$ and $r = +\infty$, and (2.52) obviously holds; let us then assume that $p < +\infty$. For any fixed $f \in L^p$, the generalized (integral) Minkowski inequality and the Hölder inequality respectively yield

$$\|f * g\|_{L^p} = \left\| \int f(x - y) g(y) \, dy \right\|_{L^p} \leq \|f\|_{L^p} \|g\|_{L^1} \qquad \forall g \in L^1\,,$$

$$\|f * g\|_{L^\infty} = \operatorname*{ess\,sup}_{x \in \mathbb{R}^N} \int f(x - y) g(y) \, dy \leq \|f\|_{L^p} \|g\|_{L^{p'}} \qquad \forall g \in L^{p'}$$

$(p^{-1} + (p')^{-1} = 1)$. Thus the mapping $g \mapsto f * g$ is (linear and) continuous from L^1 to L^p and from $L^{p'}$ to L^∞. By the Riesz-Thorin Theorem (see below), this mapping is then continuous from L^q to L^r and inequality (2.52) holds, provided that

$$\exists \theta \in \,]0, 1[: \quad \frac{1}{q} = \frac{\theta}{1} + \frac{1 - \theta}{p'}, \quad \frac{1}{r} = \frac{\theta}{p} + \frac{1 - \theta}{\infty}.$$

[7] This theorem may be compared with the following result, that easily follows from the Hölder inequality: If $p, q, r \in [1, +\infty[$ are such that $p^{-1} + q^{-1} = r^{-1}$, then

$$uv \in L^r(\Omega), \quad \|uv\|_r \leq \|u\|_p \|v\|_q \qquad \forall u \in L^p(\Omega), \forall v \in L^q(\Omega). \text{[Ex]} \tag{2.50}$$

[8] Here we set $(+\infty)^{-1} := 0$.

As the latter equality yields $\theta = p/r$, by the first one we get $p^{-1} + q^{-1} = 1 + r^{-1}$.

(ii) For any $f \in L^p_{\rm loc}$ and $g \in L^q_{\rm comp}$, setting $S_g := {\rm supp}(g)$,

$$(f * g)(x) = \int_{S_g} f(x-y)g(y)\,dy \qquad \text{converges for a.e. } x \in \mathbb{R}^N.$$

If $r = +\infty$ then $p = q = 1$, and we are in the situation of part (ii) of Proposition 2.1; let us then assume that $r \neq +\infty$. For any $K \subset\subset \mathbb{R}^N$, denoting by $\chi_{K,g}$ the characteristic function of $K - S_g$, we have

$$\begin{aligned}\|f * g\|^r_{L^r(K)} &= \int_K \left| \int_{S_g} f(x-y)g(y)\,dy \right|^r dx \\ &\leq \int \left| \int (\chi_{K,g} f)(x-y)g(y)\,dy \right|^r dx.\end{aligned}$$

As $\chi_{K,g} f \in L^p$, by part (i) the latter integral is finite.

(iii) Part (iii) may be proved by means of an argument similar to that of part (ii), that we leave to the reader. □

*** Lemma 2.4 (Riesz-Thorin's theorem)** *Let Ω, Ω' be nonempty open subsets of $\mathbb{R}^N$. For $i = 1, 2$, let $p_i, q_i \in [1, +\infty]$ and assume that*

$$T : L^{p_1}(\Omega) + L^{p_2}(\Omega) \to L^{q_1}(\Omega') + L^{q_2}(\Omega') \tag{2.55}$$

is a linear operator such that

$$T : L^{p_i}(\Omega) \to L^{q_i}(\Omega') \text{ is continuous.} \tag{2.56}$$

Let $\theta \in]0, 1[$, and $p := p(\theta)$, $q := q(\theta)$ be such that

$$\frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}, \qquad \frac{1}{q} = \frac{\theta}{q_1} + \frac{1-\theta}{q_2}. \tag{2.57}$$

Then T maps $L^p(\Omega)$ to $L^q(\Omega')$, is linear and continuous. Moreover, if M_1 and M_2 are two constants such that

$$\|Tf\|_{L^{q_i}(\Omega')} \leq M_i \|f\|_{L^{p_i}(\Omega)} \qquad \forall f \in L^{p_i}(\Omega)\ (i = 1, 2), \tag{2.58}$$

then

$$\|Tf\|_{L^q(\Omega')} \leq M_1^\theta M_2^{1-\theta} \|f\|_{L^p(\Omega)} \qquad \forall f \in L^p(\Omega). \qquad [] \tag{2.59}$$

By this result, we may regard $L^{p(\theta)}(\Omega)$ as an *interpolate space* between $L^{p_1}(\Omega)$ and $L^{p_2}(\Omega)$. ((2.59) is accordingly called the *interpolate inequality.)* This theorem is actually a prototype of the theory of Banach spaces interpolation.

For any $f : \mathbb{R}^N \to \mathbb{C}$, let us set $\check{f}(x) = f(-x)$.

Corollary 2.5 *Let*

$$p, q, s \in [1, +\infty], \qquad p^{-1} + q^{-1} + s^{-1} = 2. \tag{2.60}$$

Then

$$
\begin{aligned}
&\forall (f,g,h) \in L^p \times L^q \times L^s, \\
&(f * g)\cdot h,\ g\cdot(\check{f} * h),\ f\cdot(\check{g} * h) \in L^1, \quad \textit{and} \\
&\int (f*g)\cdot h = \int g\cdot(\check{f} * h) = \int f\cdot(\check{g} * h).
\end{aligned}
\tag{2.61}
$$

The same holds also

$$
\begin{aligned}
&\forall (f,g,h) \in \big(L^p_{\rm comp} \times L^q_{\rm loc} \times L^s_{\rm comp}\big), \\
&\forall (f,g,h) \in L^p_{\rm loc}(\mathbb{R}^+) \times L^q_{\rm loc}(\mathbb{R}^+) \times L^s_{\rm comp}(\mathbb{R}^+).
\end{aligned}
\tag{2.62}
$$

The same holds also

$$
\begin{aligned}
&\forall (u,v,w) \in \big(L^p_{\rm comp} \times L^q_{\rm loc} \times L^s_{\rm comp}\big), \\
&\forall (u,v,w) \in L^p_{\rm loc}(\mathbb{R}^+) \times L^q_{\rm loc}(\mathbb{R}^+) \times L^s_{\rm comp}(\mathbb{R}^+).
\end{aligned}
\tag{2.63}
$$

(In the language of operator theory, $\check{f}*$ is the adjoint of the operator $f*$.)

Proof. As $r^{-1}+s^{-1}=1$ by (2.52) and (2.60), the Hölder inequality yields the inclusions of (2.61). [Ex] The first equality in (2.61) follows from the computation

$$
\begin{aligned}
\int (u*v)w\,dx &= \int \Big(\int u(x-y)v(y)\,dy\Big) w(x)\,dx \\
&= \int v(y)\Big(\int \check{u}(y-x)w(x)\,dx\Big)\,dy = \int v(\check{u} * w)\,dy;
\end{aligned}
$$

the second holds since $u*v=v*u$. The assertions (2.63) are similarly checked; this is left to the reader. □

Convolution and translation. Let us next set $\tau_h f(x) := f(x+h)$ for any $f:\mathbb{R}^N \to \mathbb{C}$ and any $x,h \in \mathbb{R}^N$.

Let us denote by $C^0(\mathbb{R}^N)$ the space of continuous functions $\mathbb{R}^N \to \mathbb{C}$ (which is a Fréchet space equipped with the family of sup-norms on compact subsets of $\mathbb{R}^N$), and by $C^0_0(\mathbb{R}^N)$ the subspace of $C^0(\mathbb{R}^N)$ of functions that vanish at infinity (this is a Banach space equipped with the sup-norm).

Lemma 2.6 *As $h \to 0$,*

$$
\tau_h f \to f \qquad \textit{in } C^0,\ \forall f \in C^0, \tag{2.64}
$$

$$
\tau_h f \to f \qquad \textit{in } L^p,\ \forall f \in L^p,\ \forall p \in [1,+\infty[. \tag{2.65}
$$

Proof. As any $f \in C^0$ is locally uniformly continuous, $\tau_h f \to f$ uniformly in any $K \subset\subset \mathbb{R}^N$; (2.64) thus holds. This yields (2.65), as C^0 is dense in L^p for any $p \in [1,+\infty[$. □

By the next result, in the Young theorem the space L^∞ may be replaced by $L^\infty \cap C^0$, and in part (i) also by $L^\infty \cap C^0_0$.

*** Proposition 2.7** *Let $p,q \in [1,+\infty]$ be such that $p^{-1}+q^{-1}=1$. Then:*

$$
f*g \in C^0 \qquad \forall (f,g) \in (L^p \times L^q) \cup (L^p_{\rm loc} \times L^q_{\rm comp}), \tag{2.66}
$$

$$
f*g \in C^0 \qquad \forall (f,g) \in L^p_{\rm loc}(\mathbb{R}^+) \times L^q_{\rm loc}(\mathbb{R}^+) \quad \textit{if } N=1, \tag{2.67}
$$

$$
(f*g)(x) \to 0 \quad \textit{as} \quad |x| \to +\infty \quad \forall (f,g) \in L^p \times L^q, \forall p,q \in [1,+\infty[. \tag{2.68}
$$

Proof. For instance, let $p \neq +\infty$ and $(f, g) \in L^p \times L^q$; the other cases may be dealt with analogously. By Lemma 2.6,

$$
\begin{aligned}
\|\tau_h(f * g) - (f * g)\|_{L^\infty} &= \Big\| \int [f(x+h-y) - f(x-y)] g(y)\, dy \Big\|_{L^\infty} \\
&\leq \|\tau_h f - f\|_{L^p} \|g\|_{L^q} \to 0 \qquad \text{as } h \to 0;
\end{aligned}
\tag{2.69}
$$

the function $f * g$ may then be identified with a uniformly continuous function.

Let $\{f_n\} \subset L^p_{\text{comp}}$ and $\{g_n\} \subset L^q_{\text{comp}}$ be such that $f_n \to f$ in L^p and $g_n \to g$ in L^q. Hence $f_n * g_n$ has compact support, and $f_n * g_n \to f * g$ uniformly. This yields the final statement of the theorem. □

It is easily seen that (2.68) fails if either p or $q = +\infty$.

Regularization by convolution. A function $\rho : \mathbb{R}^N \to \mathbb{R}$ is called a **mollifier** iff

$$
\rho \in C^\infty(\mathbb{R}^N), \quad \rho \geq 0, \quad \rho(x) = 0 \ \text{ if } \ |x| \geq 1, \quad \int_{\mathbb{R}^N} \rho(x)\, dx = 1. \tag{2.70}
$$

A standard construction provides an example:

$$
\begin{aligned}
&v(x) := \exp\left[\left(|x|^2 - 1\right)^{-1}\right] \quad \text{if } |x| < 1, \qquad v(x) := 0 \quad \text{if } |x| \geq 1, \\
&\rho(x) := \frac{v(x)}{\int_{\mathbb{R}^N} v(y)\, dy}\,, \qquad \rho_\varepsilon(x) := \varepsilon^{-N} \rho\Big(\frac{x}{\varepsilon}\Big) \qquad \forall x \in \mathbb{R}^N, \forall \varepsilon > 0.
\end{aligned}
\tag{2.71}
$$

For any $u \in L^1(\Omega)$, let us denote by $\widetilde{u} \in L^1(\mathbb{R}^N)$ the extension of u with zero value on $\mathbb{R}^N \setminus \Omega$. For any $\varepsilon > 0$, we then define the regularization $R_\varepsilon u$ of u by

$$
R_\varepsilon u(x) := (\rho_\varepsilon * \widetilde{u})(x) = \int_\Omega \rho_\varepsilon(x-y) u(y)\, dy \qquad \forall x \in \mathbb{R}^N. \tag{2.72}
$$

Notice that, since $\rho_\varepsilon * \widetilde{u} = \widetilde{u} * \rho_\varepsilon$,

$$
R_\varepsilon u(x) = \varepsilon^{-N} \int_{\mathbb{R}^N} \rho\Big(\frac{y}{\varepsilon}\Big) \widetilde{u}(x-y)\, dy = \int_{\mathbb{R}^N} \rho(t) \widetilde{u}(x - \varepsilon t)\, dt. \tag{2.73}
$$

The following theorem summarizes some properties of the operator R_ε.

Proposition 2.8 *Let $u \in L^1_{\text{loc}}$ and define R_ε as above. Then:*

(i) For any $\varepsilon > 0$, $R_\varepsilon u \in C^\infty(\mathbb{R}^N)$ and

$$
D^\alpha R_\varepsilon u(x) = \varepsilon^{-N-|\alpha|} \int_\Omega \Big[(D^\alpha \rho)\Big(\frac{x-y}{\varepsilon}\Big)\Big] u(y)\, dy = \varepsilon^{-|\alpha|} \int_{\mathbb{R}^N} \big[D^\alpha \rho(y)\big]\, \widetilde{u}(x - \varepsilon y)\, dy \tag{2.74}
$$

for any $x \in \mathbb{R}^N$ and any $\alpha \in \mathbb{N}^N$.

(ii) For any $\varepsilon > 0$, the support of $R_\varepsilon u$ is contained within the ε-neighbourhood of the support of u. (Mollification thus preserves the compactness of the support.)

(iii) For any $v \in C^0_c(\mathbb{R}^N)$, $R_\varepsilon v \to v$ uniformly as $\varepsilon \to 0$.

Let us next assume that Ω is an open subset of $\mathbb{R}^N$, and $u \in L^p(\Omega)$.

(iv) For any $p \in [1,+\infty]$ and $u \in L^p(\Omega)$, $R_\varepsilon u \in L^p(\Omega)$ and $\|R_\varepsilon u\|_{L^p(\Omega)} \le \|u\|_{L^p(\Omega)}$.

(v) For any $p \in [1,+\infty[$ and $u \in L^p(\Omega)$, $\|R_\varepsilon u - u\|_{L^p(\Omega)} \to 0$ as $\varepsilon \to 0$.

Proof. (i) All the derivatives of ρ are bounded, and

$$D_x^\alpha[\rho_\varepsilon(x-y)] = \varepsilon^{-N} D_x^\alpha\Big[\rho\Big(\frac{x-y}{\varepsilon}\Big)\Big] = \varepsilon^{-N-|\alpha|}(D^\alpha\rho)\Big(\frac{x-y}{\varepsilon}\Big),$$

for all $x \in \mathbb{R}^N$ and all $\alpha \in \mathbb{N}^N$. Next let us differentiate both sides of (2.73), and interchange derivation and integration (this is easily justified via dominated convergence). This yields (2.74).

(ii) The stated property on the support of $R_\varepsilon f$ stems from (2.44) and (2.73).

(iii) Let $v \in C_c^0(\mathbb{R}^N)$. As $\|\rho\|_{L^1(\mathbb{R}^N)} = 1$ and $\rho \ge 0$, for any $x \in \mathbb{R}^N$ and any $\varepsilon > 0$ we have

$$\begin{aligned}|R_\varepsilon v(x) - v(x)| &\overset{(2.73)}{=} \int_{B(0,1)} \rho(y)[v(x-\varepsilon y) - v(x)]\,dy \\ &\le \int_{B(0,1)} \rho(y)|v(x-\varepsilon y) - v(x)|\,dy \le \max_{|z-x|\le\varepsilon} |v(z) - v(x)|.\end{aligned}$$

As v is assumed to be uniformly continuous on $\mathbb{R}^N$, it follows that $R_\varepsilon v \to v$ uniformly as $\varepsilon \to 0$.

(iv) As $\|\rho_\varepsilon\|_{L^1(\mathbb{R}^N)} = 1$, by the Young Theorem we get

$$\|R_\varepsilon u\|_{L^p(\Omega)} = \|\rho_\varepsilon * \widetilde{u}\|_{L^p(\Omega)} \le \|\rho_\varepsilon\|_{L^1(\mathbb{R}^N)}\, \|\widetilde{u}\|_{L^p(\mathbb{R}^N)} = \|u\|_{L^p(\Omega)}.$$

If $u \in L^\infty(\Omega)$ the same holds for $p = \infty$.

(v) By virtue of the density of $C_c^0(\Omega)$ in $L^p(\Omega)$, for any $\eta > 0$ there exists $v_\eta \in C_c^0(\Omega)$ with $\|u - v_\eta\|_{L^p(\Omega)} \le \eta$. By using the linearity of R_ε and part (iii), we get

$$\begin{aligned}&\|R_\varepsilon u - u\|_{L^p(\Omega)} \\ &\le \|R_\varepsilon u - R_\varepsilon v_\eta\|_{L^p(\Omega)} + \|R_\varepsilon v_\eta - v_\eta\|_{L^p(\Omega)} + \|v_\eta - u\|_{L^p(\Omega)} \\ &\le \|R_\varepsilon v_\eta - v_\eta\|_{L^p(\Omega)} + 2\eta.\end{aligned} \tag{2.75}$$

Keeping $\eta > 0$ fixed, for $\varepsilon \to 0$ we have $\|R_\varepsilon v_\eta - v_\eta\|_{L^p(\Omega)} \to 0$ as a consequence of part (iii). From (2.75) we thus conclude that $\limsup_{\varepsilon\to 0} \|R_\varepsilon u - u\|_{L^p(\Omega)} \le 2\eta$. Since $\eta > 0$ was arbitrary, the assertion follows. □

*** Convolution of distributions.** By part (ii) of Proposition 2.1,

$$f * g \in L^1_{\text{loc}} \qquad \forall (f,g) \in (L^1_{\text{loc}} \times L^1_{\text{comp}}) \cup (L^1_{\text{comp}} \times L^1_{\text{loc}}).$$

For any $\varphi \in \mathcal{D}$, then

$$\int (f*g)(x)\varphi(x)\,dx = \iint f(x-y)g(y)\varphi(x)\,dxdy = \iint f(z)g(y)\varphi(z+y)\,dzdy, \tag{2.76}$$

and of course each of these double integrals equals the corresponding iterated integral, by Fubini's theorem. This formula allows one to extend the operation of convolution to distributions, under analogous restrictions on the supports. Let either $(T,S) \in (\mathcal{D}' \times \mathcal{E}') \cup (\mathcal{E}' \times \mathcal{D}')$, and define

$$\langle T * S, \varphi\rangle := \langle T_x, \langle S_y, \varphi(x+y)\rangle\rangle \qquad \forall \varphi \in \mathcal{D}. \tag{2.77}$$

(In $\langle S_y, \varphi(x+y)\rangle$ the variable x is just a parameter; if this pairing is reduced to an integration, then y is the integration variable.) This is meaningful, since

$$S \in \mathcal{E}' \ \ (S \in \mathcal{D}', \text{ resp.}) \qquad \Rightarrow \qquad \langle S_y, \varphi(x+y)\rangle \in \mathcal{D} \ \ (\in \mathcal{E}, \text{ resp.}) \quad \forall \varphi \in \mathcal{D}.[\text{Ex}] \tag{2.78}$$

For $N = 1$, if $T \in \mathcal{D}'(\mathbb{R}^+)$, then (2.77) still makes sense.

On the other hand, one cannot write $\langle T_x S_y, \varphi(x+y)\rangle$ in the duality between $\mathcal{D}'(\mathbb{R}^N \times \mathbb{R}^N)$ and $\mathcal{D}(\mathbb{R}^N \times \mathbb{R}^N)$, since the support of the mapping $(x,y) \mapsto \varphi(x+y)$ is compact only if $\varphi \equiv 0$.

In $\mathcal{E}'$ the convolution commutes and is associative. Thus $(\mathcal{E}', *)$ is a convolution algebra, with unit element δ_0. Here are some further properties:

$$\mathcal{D}' * \mathcal{E}' \subset \mathcal{D}', \qquad \mathcal{E}' * \mathcal{E}' \subset \mathcal{E}', \tag{2.79}$$

$$\mathcal{S}' * \mathcal{E}' \subset \mathcal{S}', \qquad \mathcal{S} * \mathcal{S}' \subset \mathcal{E} \cap \mathcal{S}', \qquad \mathcal{S} * \mathcal{E}' \subset \mathcal{S}, \tag{2.80}$$

and in all of these cases the convolution is separately continuous w.r.t. either factor.

For instance, the inclusion $\mathcal{D}' * \mathcal{E}' \subset \mathcal{D}'$ is an extension of $L^1_{\text{loc}} * L^1_{\text{comp}} \subset L^1_{\text{loc}}$, and actually may be proved by approximating distributions by L^1_{loc}- or L^1_{comp}-functions, by using the foregoing inclusion, and then passing to the limit. This procedure may also be used to prove $\mathcal{E}' * \mathcal{E}' \subset \mathcal{E}'$. The other inclusions may analogously be justified by approximation and passage to the limit.

In general the convolution of distributions is not associative. For instance,

$$\begin{aligned} &(1 * \delta') * H = (1' * \delta) * H = (0 * \delta) * H = 0 * H = 0, \\ &1 * (\delta' * H) = 1 * (\delta * H') = 1 * \delta = 1. \end{aligned} \tag{2.81}$$

3 The Fourier Transform in L^1

Integral transforms. These are linear integral operators $\mathcal{T}$ that typically act on functions $\mathbb{R} \to \mathbb{C}$, and have the form

$$\big(\widehat{u}(\xi) := \big)\ (\mathcal{T}u)(\xi) = \int_{\mathbb{R}} K(\xi, x) u(x)\, dx \qquad \forall \xi \in \mathbb{R}, \tag{3.1}$$

for a prescribed kernel $K : \mathbb{R}^2 \to \mathbb{C}$, and for any transformable function u. [9] The main properties of this class of transforms include the following:

(i) Inverse transform. Under appropriate restrictions, there exists another kernel $\widetilde{K} : \mathbb{R}^2 \to \mathbb{C}$ such that (formally)

$$\int_{\mathbb{R}} \widetilde{K}(x, \xi) K(\xi, y)\, d\xi = \delta_0(x - y) \qquad \forall x, y \in \mathbb{R}. \tag{3.2}$$

Denoting by $\widetilde{\mathcal{T}}$ the integral operator associated to $\widetilde{K}$, we thus have $\widetilde{\mathcal{T}}\mathcal{T}u = \mathcal{T}\widetilde{\mathcal{T}}u = u$ for any transformable u.

(ii) Commutation Formula. Any integral transform is associated to a class of linear operators (typically of differential type), that act on functions of time. For any such operator, L, there exists a function, $\widetilde{L} = \widetilde{L}(\xi)$, such that

$$\mathcal{T} L \mathcal{T}^{-1} = \widetilde{L} \qquad (\widetilde{L} \text{ is a function}). \tag{3.3}$$

[9] To devise hypotheses that encompass a large number of integral transforms is not easy and may not be convenient. In this brief overview we then refer to the Fourier transform. We are intentionally sloppy and drop regularity properties, that however are specified ahead.

For a prescribed function $f = f(x)$, by applying $\mathcal{T}$ an equation of the form $Lu = f$ is then transformed to $\widetilde{L}(\xi)\widehat{u}(\xi) = \widehat{f}(\xi)$. Thus $\widehat{u} = \widehat{f}/\widetilde{L}$, whence $u = \widetilde{\mathcal{T}}(\widehat{f}/\widetilde{L})$, provided that $\widetilde{L}(\xi) \neq 0$ for a.e. ξ. This procedure is at the basis of so-called *symbolic (or operational) calculus,* that was introduced by O. Heaviside at the end of the 19th century.

The first of the transforms that we illustrate is named after J. Fourier, who introduced it at the beginning of the 19th century, and is the keystone of all integral transforms. In the 1950s Laurent Schwartz introduced the space of *tempered distributions,* and extended the Fourier transform to this class. Because of the commutation formula, this transform allows one to reduce linear ordinary differential equations with constant coefficients to algebraic equations; this found many uses in the study of stationary problems.

The Fourier transform in L^1. We shall systematically deal with spaces of functions from the whole $\mathbb{R}^N$ to $\mathbb{C}$. We shall then write L^1 in place of $L^1(\mathbb{R}^N)$, C^0 in place of $C^0(\mathbb{R}^N)$, and so on. For any $u \in L^1$, we define the **Fourier transform** (also called **Fourier integral**) $\widehat{u}$ of u by [10]

$$\widehat{u}(\xi) := (2\pi)^{-N/2} \int_{\mathbb{R}^N} e^{-i\xi\cdot x} u(x)\, dx \qquad \forall \xi \in \mathbb{R}^N; \tag{3.4}$$

here $\xi \cdot x := \sum_{i=1}^N \xi_i x_i$. This is a Lebesgue integral.

Proposition 3.1 *The formula (3.4) defines a linear and continuous operator*

$$\begin{aligned} &\mathcal{F} : L^1 \to C^0_b : u \mapsto \widehat{u}; \\ &\|\widehat{u}\|_{L^\infty} \le (2\pi)^{-N/2} \|u\|_{L^1} \quad \forall u \in L^1. \text{[Ex]} \end{aligned} \tag{3.5}$$

(By C^0_b we denote the Banach space $C^0 \cap L^\infty$, equipped with the sup-norm.)

Thus $\widehat{u}_n \to \widehat{u}$ uniformly in $\mathbb{R}^N$ whenever $u_n \to u$ in L^1. In passing notice that $\|\widehat{u}\|_{L^\infty} = (2\pi)^{-N/2}\|u\|_{L^1} = \widehat{u}(0)$ for any nonnegative $u \in L^1$, as in this case

$$\|\widehat{u}\|_{C^0_b} \le (2\pi)^{-N/2}\|u\|_{L^1} = \widehat{u}(0) \le \|\widehat{u}\|_{C^0_b}.$$

Apparently, no simple condition characterizes the image set $\mathcal{F}(L^1)$.

Cosine and sine transforms. For any $u \in L^1$, (3.4) also reads

$$\widehat{u}(\xi) = (2\pi)^{-N/2} \int_{\mathbb{R}^N} \cos(\xi\cdot x) u(x)\, dx - i(2\pi)^{-N/2} \int_{\mathbb{R}^N} \sin(\xi\cdot x) u(x)\, dx \tag{3.6}$$

for any $\xi \in \mathbb{R}^N$. Defining the so-called **cosine transform** and **sine transform** respectively by

$$C_u(\xi) = (2/\pi)^{-N/2} \int_{\mathbb{R}^N} \cos(\xi\cdot x) u(x)\, dx \qquad \forall \xi \in \mathbb{R}^N, \forall u \in L^1, \tag{3.7}$$

$$S_u(\xi) = (2/\pi)^{-N/2} \int_{\mathbb{R}^N} \sin(\xi\cdot x) u(x)\, dx \qquad \forall \xi \in \mathbb{R}^N, \forall u \in L^1; \tag{3.8}$$

we thus have

$$\widehat{u} = C_u - iS_u \qquad \forall u \in L^1. \tag{3.9}$$

[10] Some authors introduce a factor 2π in the exponent under the integral, others omit the factor in front of the integral. Our definition is maybe the most frequently used. Each of these modifications simplifies some formulas, but none is able to simplify all of them.

Therefore, for any $u \in L^1$,

$$u \text{ is even} \quad \Leftrightarrow \quad \widehat{u}(\xi) = C_u(\xi) \qquad \forall \xi \in \mathbb{R}^N, \tag{3.10}$$
$$u \text{ is odd} \quad \Leftrightarrow \quad \widehat{u}(\xi) = -iS_u(\xi) \qquad \forall \xi \in \mathbb{R}^N. \tag{3.11}$$

The functions C_u and S_u are real valued iff so is u itself. Above the Fourier series of periodic functions were similarly decomposed into the sum of a *cosine series* and a *sine series.*

The following properties mimic those of the Fourier series and have the same basis: the properties of the exponential function $e^{-i\xi\cdot x}$. In particular, if the argument of the input function is shifted, then the transformed function is multiplied by an exponential function; [11] conversely, if the input function is modulated, then the argument of the transformed function is shifted.

Proposition 3.2 *For any* $u \in L^1$, [12]

$$v(x) = u(x-y) \quad \Rightarrow \quad \widehat{v}(\xi) = e^{-i\xi\cdot y}\widehat{u}(\xi) \qquad \forall y \in \mathbb{R}^N, \tag{3.12}$$
$$v(x) = e^{ix\cdot\eta}u(x) \quad \Rightarrow \quad \widehat{v}(\xi) = \widehat{u}(\xi-\eta) \qquad \forall \eta \in \mathbb{R}^N, \tag{3.13}$$
$$v(x) = \overline{u(x)} \quad \Rightarrow \quad \widehat{v}(\xi) = \overline{\widehat{u}(-\xi)}, \tag{3.14}$$
$$u \text{ is real} \quad \Rightarrow \quad \widehat{u}(-\xi) = \overline{\widehat{u}(\xi)}, \tag{3.15}$$
$$u \text{ is imaginary} \quad \Rightarrow \quad \widehat{u}(-\xi) = -\overline{\widehat{u}(\xi)}, \tag{3.16}$$
$$u \text{ is even} \quad \Rightarrow \quad \widehat{u} \text{ is even}, \tag{3.17}$$
$$u \text{ is odd} \quad \Rightarrow \quad \widehat{u} \text{ is odd}, \tag{3.18}$$
$$u \text{ is radial} \quad \Rightarrow \quad \widehat{u} \text{ is radial}, \tag{3.19}$$
$$v(x) = u\big(A^{-1}x\big) \;\Rightarrow\; \widehat{v}(\xi) = |\det A|\,\widehat{u}(A^*\xi) \quad \forall A \in \mathbb{R}^{N^2}, \det A \neq 0. \tag{3.20}$$

[Ex]

Let us define the operators

$$(T_y u)(x) = u(x-y), \qquad (M_y u)(x) = e^{i\xi\cdot y}u(x) \qquad \forall x, y \in \mathbb{R}^N, \forall u \in L^1. \tag{3.21}$$

(3.12) and (3.12) then read

$$v = T_y u \quad \Rightarrow \quad \widehat{v}(\xi) = M_{-y}\widehat{u} \qquad \forall y \in \mathbb{R}^N, \tag{3.22}$$
$$v = M_\eta u \quad \Rightarrow \quad \widehat{v}(\xi) = T_\eta \widehat{u} \qquad \forall \eta \in \mathbb{R}^N. \tag{3.23}$$

This entails that

$$v = T_y M_\eta u \quad \Rightarrow \quad \widehat{v} = M_{-y}T_\eta \qquad \forall x, \eta \in \mathbb{R}^N, \forall u \in L^1. \tag{3.24}$$

Henceforth by D (or D_j or D^α) we shall denote the operation of derivation in the sense of distributions.

Lemma 3.3 *Let* $j \in \{1, ..., N\}$*. If* $\varphi, D_j\varphi \in L^1$ *then* $\int_{\mathbb{R}^N} D_j\varphi(x)\,dx = 0$.

[11] In the engineering literature, one says that the function is *modulated.*

[12] For any $A \in \mathbb{R}^{N^2}$, we set $(A^*)_{ij} := A_{ji}$ for any i, j. For any $z \in \mathbb{C}$, we denote its complex conjugate by $\overline{z}$. We say that u is **radial** iff $u(Ax) = u(x)$ for any x and any orthonormal matrix $A \in \mathbb{R}^{N^2}$ (i.e., with $A^* = A^{-1}$).

Proof. Let us recall the definition of the bell-shaped function ρ of (1.2), and set

$$\rho_n(x) := \rho\big(x/n\big) \qquad \forall x \in \mathbb{R}^N, \forall n \in \mathbb{N}. \tag{3.25}$$

As ρ_n has compact support, by partial integration

$$\Big| \int_{\mathbb{R}^N} \big[D_j\varphi(x)\big]\rho_n(x)\,dx \Big| = \Big| \int_{\mathbb{R}^N} \varphi(x) D_j\rho_n(x)\,dx \Big| \le \frac{1}{n}\|\varphi\|_{L^1} \cdot \|D_j\rho\|_\infty \to 0$$

as $n \to \infty$. Since $\rho_n(x) \to 1$ pointwise in $\mathbb{R}^N$, by the dominated convergence theorem we then conclude that

$$\int_{\mathbb{R}^N} D_j\varphi(x)\,dx = \lim_{n\to\infty} \int_{\mathbb{R}^N} \big[D_j\varphi(x)\big]\rho_n(x)\,dx = 0. \qquad \square$$

• **Theorem 3.4** *For any multi-index* $\alpha \in \mathbb{N}^N$,

$$u, D_x^\alpha u \in L^1 \quad \Rightarrow \quad (i\xi)^\alpha \widehat{u} = (D_x^\alpha u)\widehat{} \in C_b^0, \tag{3.26}$$

$$u, x^\alpha u \in L^1 \quad \Rightarrow \quad D_\xi^\alpha \widehat{u} = [(-ix)^\alpha u]\widehat{} \in C_b^0. \tag{3.27}$$

Proof. In both cases it suffices to prove the equality for any first-order derivative; the general case then follows by induction.

(i) Let us fix any $j \in \{1, ..., N\}$. As

$$D_j[e^{-i\xi\cdot x}u(x)] = -i\xi_j e^{-i\xi\cdot x}u(x) + e^{-i\xi\cdot x}D_j u(x) \qquad \forall x \in \mathbb{R}^N,$$

the integrability of u and $D_j u$ entails that $D_j[e^{-i\xi\cdot x}u(x)] \in L^1$. It then suffices to integrate the latter equality over $\mathbb{R}^N$, and to notice that $\int_{\mathbb{R}^N} D_j[e^{-i\xi\cdot x}u(x)]\,dx = 0$ by Lemma 3.3. Finally $(D_x^\alpha u)\widehat{} \in C_b^0$, by Proposition 3.1.

(ii) Let us denote by e_j the unit vector in the jth direction. By applying the classical formula $\frac{e^{is}-e^{-is}}{2i} = \sin s$ with $s = tx_j/2$, we have

$$\begin{aligned}\frac{\widehat{u}(\xi + te_j) - \widehat{u}(\xi)}{t} &= \int_{\mathbb{R}^N} \frac{e^{-i(\xi+te_j)\cdot x} - e^{-i\xi\cdot x}}{t} u(x)\,dx \\ &= -i\int_{\mathbb{R}^N} e^{-i(\xi\cdot x + tx_j/2)}\,\frac{\sin(tx_j/2)}{t/2} u(x)\,dx.\end{aligned}$$

Passing to the limit as $t \to 0$, by the dominated convergence theorem we then get

$$\frac{\widehat{u}(\xi + te_j) - \widehat{u}(\xi)}{t} \to -i\int_{\mathbb{R}^N} e^{-i\xi\cdot x}x_j u(x)\,dx = -i(x_j u)\widehat{}(\xi) \qquad \forall \xi.$$

By Proposition 3.1, this is an element of C_b^0. $\square$

Let us define the operators

$$(Xu)(x) = ixu(x), \qquad (\Xi u)(\xi) = i\xi u(\xi), \qquad \forall x, \xi \in \mathbb{R}^N, \forall u \in L^1. \tag{3.28}$$

(3.26) and (3.27) then read

$$u, D_x^\alpha u \in L^1 \quad \Rightarrow \quad \Xi^\alpha \widehat{u} = (D_x^\alpha u)\widehat{} \in C_b^0, \tag{3.29}$$

$$u, x^\alpha u \in L^1 \quad \Rightarrow \quad D_\xi^\alpha \widehat{u} = [(-X)^\alpha u]\widehat{} \in C_b^0. \tag{3.30}$$

Corollary 3.5 *Let $m \in \mathbb{N}_0$.*
(i) If $D_x^\alpha u \in L^1$ for any $\alpha \in \mathbb{N}_0^N$ with $|\alpha| \le m$, then $(1+|\xi|)^m \widehat{u}(\xi) \in L^\infty$.
(ii) If $(1+|x|)^m u \in L^1$, then $\widehat{u} \in C^m$ and $D^\alpha \widehat{u} \in L^\infty$ for any α. [Ex]

In other terms:
(i) the more u is regular, the faster $|\widehat{u}|$ decreases at infinity;
(ii) the faster $|u|$ decreases at infinity, the more $\widehat{u}$ is regular.

Examples. (i) For any $A > 0$, if $u = \chi_{[-A,A]}$, then $\widehat{u}(\xi) = \sqrt{2/\pi}\,\frac{\sin(A\xi)}{\xi}$. [Ex] [13]
(ii) We claim that

$$\begin{aligned} &u(x) = \exp(-a|x|^2) \qquad \forall x \in \mathbb{R}^N \quad \Rightarrow \\ &\widehat{u}(\xi) = (2a)^{-N/2} \exp(-|\xi|^2/(4a)) \qquad \forall \xi \in \mathbb{R}^N. \end{aligned} \tag{3.31}$$

Let us first prove this statement in the case of $a = 1/2$ and $N = 1$. [14] As $D_x u = -xu$ for any $x \in \mathbb{R}^N$,

$$i\xi\widehat{u}(\xi) \overset{(3.26)}{=} \widehat{D_x u}(\xi) = \widehat{-xu}(\xi) \overset{(3.27)}{=} -iD_\xi \widehat{u}(\xi),$$

that is, $D_\xi \widehat{u} = -\xi\widehat{u}$ for any $\xi \in \mathbb{R}^N$. On the other hand, by the classical Poisson formula $\int_{\mathbb{R}} \exp(-y^2)\,dy = \sqrt{\pi}$, for $N = 1$

$$\widehat{u}(0) = (2\pi)^{-1/2} \int e^0 e^{-x^2/2}\,dx = 1.$$

As $u(0) = 1$, we see that $\widehat{u}$ solves the same Cauchy problem as u. Therefore for $N = 1$

$$u(x) = \exp(-x^2/2) \quad \forall x \in \mathbb{R} \quad \Rightarrow \quad \widehat{u}(\xi) = \exp(-\xi^2/2) \quad \forall \xi \in \mathbb{R}. \tag{3.32}$$

For $N > 1$ and still for $a = 1/2$, $u(x) = \exp(-|x|^2/2) = \prod_{j=1}^N \exp(-x_j^2/2)$. Therefore

$$\begin{aligned} \widehat{u}(\xi) &= (2\pi)^{-N/2} \int e^{-i\xi\cdot x} e^{-|x|^2/2}\,dx = (2\pi)^{-N/2} \int \dots \int e^{\sum_{j=1}^N (-i\xi_j x_j + x_j^2/2)}\,dx_1 \dots dx_N \\ &= \prod_{j=1}^N \Big\{ (2\pi)^{-1/2} \int e^{-i\xi_j x_j} e^{-x_j^2/2}\,dx_j \Big\} \overset{(3.32)}{=} \prod_{j=1}^N e^{-\xi_j^2/2} = e^{-|\xi|^2/2} \quad \forall \xi \in \mathbb{R}^N. \end{aligned}$$

This concludes the proof of (3.31) for $a = 1/2$. (3.31) then follows from (3.20).

The next theorem mimics a property that we saw for Fourier series.

Theorem 3.6 (Riemann-Lebesgue theorem) *For any $u \in L^1$, $\widehat{u}$ is uniformly continuous in $\mathbb{R}^N$, and $\widehat{u}(\xi) \to 0$ as $|\xi| \to +\infty$.*

---

[13] Defining the *cardinal sinus* function

$$\operatorname{sinc} v := \frac{\sin v}{v} \qquad \forall v \in \mathbb{R},$$

this also reads $\widehat{u}(\xi) = A\sqrt{2/\pi}\,\operatorname{sinc}(A\xi)$. This function is often used in applications.

[14] A different proof of this result is based on integration along paths in the complex plane.

Proof. For any $u \in L^1$, there exists a sequence $\{u_n\}$ in $\mathcal{D}$ such that $u_n \to u$ in L^1. By part (i) of Corollary 3.5, $\widehat{u}_n(\xi) \to 0$ as $|\xi| \to +\infty$. The same holds also for $\widehat{u}$, since $\widehat{u}_n \to \widehat{u}$ uniformly in $\mathbb{R}^N$ by Proposition 3.1. [15] As $\widehat{u}$ is continuous, it is then uniformly continuous. $\square$

Remark. This theorem entails that, for any measurable Ω,

$$\int_{\mathbb{R}^N} e^{in\cdot x} u(x)\, dx \to 0 \qquad \text{as } n \to \infty, \forall u \in L^1(\Omega). \tag{3.33}$$

This is easily checked by extending u to $\mathbb{R}^N$ with vanishing value outside Ω. Thus $e^{in\cdot x} \to 0$ weakly star in $L^\infty(\Omega)$.

• **Theorem 3.7 (Parseval)** *The formal adjoint of $\mathcal{F}$ coincides with $\mathcal{F}$ itself, that is,*

$$\int_{\mathbb{R}^N} \widehat{u}\, v\, dx = \int_{\mathbb{R}^N} u\, \widehat{v}\, dx \qquad \forall u, v \in L^1. \tag{3.34}$$

Moreover,

$$u{*}v \in L^1\,, \quad \text{and} \quad (u{*}v)\widehat{} = (2\pi)^{N/2}\widehat{u}\,\widehat{v} \qquad \forall u, v \in L^1. \tag{3.35}$$

Proof. By the theorems of Tonelli and Fubini, for any $u, v \in L^1$

$$\int_{\mathbb{R}^N} \widehat{u}(y)v(y)\, dy = (2\pi)^{-N/2} \iint_{\mathbb{R}^N\times\mathbb{R}^N} e^{-iy\cdot x} u(x)v(y)\, dx\, dy = \int_{\mathbb{R}^N} u(y)\widehat{v}(y)\, dy.$$

On the other, by the change of integration variable $z = x - y$,

$$\begin{aligned}(u{*}v)\widehat{}(\xi) &= (2\pi)^{-N/2} \iint_{\mathbb{R}^N\times\mathbb{R}^N} e^{-i\xi\cdot x} u(x-y)v(y)\, dx\, dy \\ &= (2\pi)^{-N/2} \int_{\mathbb{R}^N} e^{-i\xi\cdot z} u(z)\, dz \int_{\mathbb{R}^N} e^{-i\xi\cdot y} v(y)\, dy \\ &= (2\pi)^{N/2}\widehat{u}(\xi)\widehat{v}(\xi). \qquad \square\end{aligned}$$

Next we present the inversion formula for the Fourier transform. First, we introduce the so-called *conjugate Fourier transform:*

$$\widetilde{\mathcal{F}}(v)(x) := (2\pi)^{-N/2} \int_{\mathbb{R}^N} e^{i\xi\cdot x} v(\xi) d\xi \qquad \forall v \in L^1, \forall x \in \mathbb{R}^N. \tag{3.36}$$

This operator differs from $\mathcal{F}$ just in the sign of the imaginary unit. Obviously, $\widetilde{\mathcal{F}}v = \overline{\mathcal{F}\overline{v}}$ for any $v \in L^1$. Clearly the properties of $\widetilde{\mathcal{F}}$ are mimic those of $\mathcal{F}$.

Theorem 3.8 *For any $u \in L^1 \cap C^0_b$, if $\widehat{u} \in L^1$ then*

$$u(x) = [\widetilde{\mathcal{F}}(\widehat{u})](x) \qquad \forall x \in \mathbb{R}^N. \tag{3.37}$$

[15] Here is an alternative argument. By direct evaluation of the integral one may check that the assertion holds for the characteristic function of any N-dimensional interval $[a_1, b_1] \times \cdots \times [a_N, b_N]$. It then suffices to approximate u in L^1 by a sequence of finite linear combinations of characteristic functions of N-dimensional intervals.

Proof. Let us set $v(x) := \exp(-|x|^2/2)$ for any $x \in \mathbb{R}^N$. By the Tonelli and Fubini theorems, we have

$$\int_{\mathbb{R}^N} \widehat{u}(\xi)v(\xi)e^{i\xi\cdot x}\,d\xi = (2\pi)^{-N/2}\iint_{\mathbb{R}^N\times\mathbb{R}^N} u(y)e^{-i\xi\cdot y}v(\xi)e^{i\xi\cdot x}\,dy\,d\xi$$
$$= \int_{\mathbb{R}^N} u(y)\widehat{v}(y-x)\,dy = \int_{\mathbb{R}^N} u(x+z)\widehat{v}(z)\,dz \qquad \forall x \in \mathbb{R}^N.$$

Let us now replace $v(\xi)$ by $v_\varepsilon(\xi) := v(\varepsilon\xi)$, for any $\varepsilon > 0$. By (3.20), $\widehat{v_\varepsilon}(z) = \varepsilon^{-N}\widehat{v}(\varepsilon^{-1}z)$; by a further change of variable of integration, we then get

$$\int_{\mathbb{R}^N} \widehat{u}(\xi)v(\varepsilon\xi)e^{i\xi\cdot x}\,d\xi = \int_{\mathbb{R}^N} u(x+\varepsilon y)\widehat{v}(y)\,dy \qquad \forall x \in \mathbb{R}^N.$$

As u and v are continuous and bounded, by the dominated convergence theorem we may pass to the limit under integral as $\varepsilon \to 0$. We thus get

$$v(0)\int_{\mathbb{R}^N} \widehat{u}(\xi)e^{i\xi\cdot x}\,d\xi = u(x)\int_{\mathbb{R}^N} \widehat{v}(y)\,dy. \tag{3.38}$$

On the other hand, by (3.31)

$$\int_{\mathbb{R}^N} \widehat{v}(y)\,dy = \int_{\mathbb{R}^N} \exp(-|y|^2/2)\,dy = \left(\int_{\mathbb{R}} \exp(-s^2/2)\,ds\right)^N = (2\pi)^{N/2}.$$

As $v(0) = 1$, (3.38) then yields (3.37). □

Remarks. (i) By Proposition 3.1, for the above argument the regularity assumptions of Theorem 3.8 are actually needed, as $\bar{u} = \mathcal{F}\big(\overline{\widehat{u}}\big)$. However, by a more refined argument one might show that (3.37) holds under the only hypotheses that $u, \widehat{u} \in L^1$. (Of course, a posteriori one then gets that $u, \widehat{u} \in C^0_b$.)

(ii) By Theorem 3.8, $\mathcal{F}(u) \equiv 0$ only if $u \equiv 0$; hence the Fourier transform $L^1 \to C^0_b$ is injective. Under the assumptions of this theorem, we also have

$$\overline{\widehat{\widehat{u}}(x)} = \bar{u}(-x) \qquad \forall x \in \mathbb{R}^N. \tag{3.39}$$

The Fourier-Laplace transform. This is the extension of the Fourier transform to holomorphic functions of several complex variables. [16] Its existence requires strong restrictions on the behaviour of $u(x)$ as $|x| \to +\infty$.

For any $z \in \mathbb{C}^N$, let us set $|z| = \big(\sum_{i=1}^N |z_i|^2\big)^{1/2}$ and $\mathrm{Im}(z) = (\mathrm{Im}(z_1), ..., \mathrm{Im}(z_N)) \in \mathbb{R}^N$. By $B(0,R)$ we still denote the ball of $\mathbb{R}^N$ with center the origin and radius R.

Theorem 3.9 (Holomorphy) *If $u \in L^1$ and $e^{\lambda|x|}u \in L^1$ for some $\lambda > 0$, then $\mathcal{F}(u)$ can be extended to a (necessarily unique) holomorphic function $\widehat{u} : (\mathbb{R} \times iB(0,\lambda))^N \to \mathbb{C}$.* [17]

[16] The reason of the denomination of *Fourier-Laplace transform* will be clear after introducing the Laplace transform in the next chapter.

[17] This means that $\widehat{u}$ is separately holomorphic with respect to each variable.

Proof. Here we assume that $N = 1$; however the argument is easily extended to any N. It suffices to prove that

$$\frac{\widehat{u}(z+h) - \widehat{u}(z)}{h} = \int \frac{e^{-i(z+h)\cdot x} - e^{-iz\cdot x}}{h} u(x)\,dx \qquad \forall z \in \mathbb{R} \times iB(0,\lambda) \tag{3.40}$$

converges as $h \to 0$ in $\mathbb{C}$. Defining $M(z) := \sup_{x\in\mathbb{R}} |x| e^{-|x|[\lambda - \mathrm{Im}(z)]/2}$ $(< +\infty)$, a direct computation shows that

$$\begin{aligned} &\Big|\frac{e^{-i(z+h)\cdot x} - e^{-iz\cdot x}}{h}\Big| = \Big|e^{-iz\cdot x}\Big|\,\Big|\frac{e^{-ih\cdot x} - 1}{h}\Big| \\ &\le e^{|\mathrm{Im}(z)|\,|x|}|x| \sup_{t\in[0,1]} e^{t\mathrm{Im}(h)|x|} \le M(z)e^{\lambda|x|} \qquad \forall z \in \mathbb{R} \times iB(0,\lambda). \quad [\text{Ex}] \end{aligned} \tag{3.41}$$

The modulus of the integrand of (3.40) is thus dominated by $M(z)e^{\lambda|x|}|u(x)| \in L^1$, so that one may pass to the limit in (3.40). □

Remark. Theorem 3.4 and other results also hold for the Fourier-Laplace transform $\widehat{u}$ in its domain of holomorphy, A. In particular,

$$u, D_x^\alpha u \in L^1 \quad \Rightarrow \quad (iz)^\alpha \widehat{u}(z) = (D_x^\alpha u)\widehat{\ }(z) \quad \forall z \in A, \tag{3.42}$$

$$u, x^\alpha u \in L^1 \quad \Rightarrow \quad D_z^\alpha \widehat{u}(z) = [(-ix)^\alpha u]\widehat{\ }(z) \quad \forall z \in A. \tag{3.43}$$

The next result assumes that the support of u is bounded, and provides an estimate on the growth at infinity of $\widehat{u}$ and of its derivatives.

Proposition 3.10 *Let $u \in L^1$ and $\operatorname{supp} u \subset B(0,R)$ for some $R > 0$. $\mathcal{F}(u)$ can then be extended to a (necessarily unique) holomorphic function $\widehat{u} : \mathbb{C}^N \to \mathbb{C}$ such that*

$$|D^\alpha \widehat{u}(z)| \le (2\pi)^{-N/2} R^{|\alpha|} e^{R|\mathrm{Im}(z)|} \|u\|_{L^1} \qquad \forall z \in \mathbb{C}^N, \forall \alpha \in \mathbb{N}^N. \tag{3.44}$$

Proof. By Theorem 3.10, $\widehat{u}$ is holomorphic on the whole $\mathbb{C}^N$. By (3.43),

$$\begin{aligned} |D^\alpha \widehat{u}(z)| &= (2\pi)^{-N/2}\Big|\int (-ix)^\alpha e^{-iz\cdot x} u(x)\,dx\Big| \\ &\le (2\pi)^{-N/2} \int_{B(0,R)} |(-ix)^\alpha||e^{-iz\cdot x}||u(x)|\,dx \\ &\le (2\pi)^{-N/2} R^{|\alpha|} e^{R|\mathrm{Im}(z)|} \|u\|_{L^1} \qquad \forall z \in \mathbb{C}^N. \square \end{aligned} \tag{3.45}$$

The next classical theorem provides a necessary and sufficient condition for the existence of the holomorphic extension of the Fourier transformed function.

Theorem 3.11 (Paley-Wiener) *Let $u \in L^1$ and $R > 0$. Then the following two conditions are equivalent:*

(i) $u \in C^\infty(\mathbb{R}^N)$ and $\operatorname{supp} u \subset B(0,R)$;

(ii) $\mathcal{F}(u)$ can be extended to a (necessarily unique) holomorphic function $\widehat{u} : \mathbb{C}^N \to \mathbb{C}$ such that

$$\forall m \in \mathbb{N}, \exists C > 0 : \forall z \in \mathbb{C}^N \qquad |\widehat{u}(z)| \le C\frac{e^{R|\mathrm{Im}(z)|}}{(1+|z|)^m}. \tag{3.46}$$

(The constant C depends on u and m.)

Proof of "(i) ⇒ (ii)". By condition (i), for any $\alpha \in \mathbb{N}^N$, $u, D_x^\alpha u \in L^1$. By (3.42), then

$$|z^\alpha||\widehat{u}(z)| = (2\pi)^{-N/2}\Big|\int_{B(0,R)} e^{-iz\cdot x} D^\alpha u(x)\,dx\Big| \le (2\pi)^{-N/2} e^{R|\mathrm{Im}(z)|}\|D^\alpha u\|_{L^1} \qquad \forall z \in \mathbb{C}^N.$$

Therefore, for any $m \in \mathbb{N}$ there exists a constant $C > 0$ (which will depend on u and m) such that

$$(1+|z|)^m|\widehat{u}(z)| \le Ce^{R|\mathrm{Im}(z)|} \qquad \forall z \in \mathbb{C}^N.\text{[Ex]}\square$$

Overview of the Fourier transform in L^1. We defined the classic Fourier transform $\mathcal{F} : L^1 \to C_b^0$ and derived its basic properties. In particular we saw the following:

(i) $\mathcal{F}$ transforms partial derivatives to multiplication by powers of the independent variable (up to a multiplicative constant) and conversely. This is at the basis of the application of the Fourier transform to the study of linear partial differential equations with constant coefficients, that we shall outline ahead.

(ii) $\mathcal{F}$ establishes a correspondence between the regularity of u and the order of decay of $\widehat{u}$ at ∞, and conversely between the order of decay of u at ∞ and the regularity of $\widehat{u}$. In the limit case of a compactly supported function, the Fourier transform may be extended to an entire holomorphic function $\mathbb{C}^N \to \mathbb{C}$.

(iii) $\mathcal{F}$ transforms the convolution of two functions to the product of their transforms (the converse statement may fail, because of summability restrictions).

(iv) Under suitable regularity restrictions, the inverse transform exists, and has an integral representation analogous to that of the direct transform.

The properties of the two transforms are then similar; this accounts for the duality of the statements (i) and (ii). However the assumptions are not perfectly symmetric; in the next section we shall see a different functional framework where this is remedied.

The inversion formula (3.37) also provides an interpretation of the Fourier transform. (3.37) represents u as a weighted average of the *harmonic components* $x \mapsto e^{i\xi\cdot x}$. For any $\xi \in \mathbb{R}^N$, $\widehat{u}(\xi)$ is the *amplitude* of the component having *vector frequency* ξ (that is, frequency ξ_i in each direction x_i). [18] Therefore any function which fulfills (3.37) may equivalently be represented by specifying either the value $u(x)$ at a.e. points $x \in \mathbb{R}^N$, or the amplitude $\widehat{u}(\xi)$ for a.e. frequencies $\xi \in \mathbb{R}^N$.

The analogy between the Fourier transform and the Fourier series is obvious, and will be briefly discussed at the end of the next section.

4 Extensions of the Fourier Transform

Fourier transform of measures. The Fourier transform may be extended to any finite complex Borel measure μ on $\mathbb{R}^N$. Formally, this simply corresponds to replacing $f(x)\,dx$ by $d\mu(x)$ in (3.4). This is called the **Fourier-Stieltjes transform.** Most of the previously established properties hold also in this more general set-up. For instance, transformed functions are still elements of C_b^0 and fulfill the properties of transformation of derivatives and multiplication by a power of x. The Riemann-Lebesgue Theorem 3.6) instead fails: e.g., $\widehat{\delta_0}(\xi) = (2\pi)^{N/2}$, which does not vanish as $|\xi| \to +\infty$.

[18] It is usual to use the term *frequency,* but ξ is rather the *angular frequency* (expressed in radians).

Fourier transform in $\mathcal{S}$. For any $u \in \mathcal{D}$, $\widehat{u}$ is holomorphic by Theorem 3.11. Hence $\widehat{u} \in \mathcal{D}$ only if $\widehat{u} \equiv 0$, namely $u \equiv 0$ by Theorem 3.8. Fourier transform thus does not map $\mathcal{D}$ to itself. This means that the set of the frequencies of the harmonic components of any non-identically vanishing $u \in \mathcal{D}$ is unbounded. In other terms, any non-identically vanishing $u \in \mathcal{D}$ has harmonic components of arbitrarily large frequencies. This induced L. Schwartz to introduce the space of rapidly decreasing functions $\mathcal{S}$, and to extend the Fourier transform to this space and to its dual. Next we review the tenets of that theory. We shall operate several identifications, omitting to display restrictions.

Proposition 4.1 *(The restriction of) $\mathcal{F}$ maps $\mathcal{S}$ to $\mathcal{S}$, is continuous, and is invertible: $\mathcal{F}^{-1} = \widetilde{\mathcal{F}}$. Moreover, for any $u, v \in \mathcal{S}$,*

$$(i\xi)^\alpha \widehat{u} = (D_x^\alpha u)\widehat{\ }, \tag{4.1}$$

$$D_\xi^\alpha \widehat{u} = [(-ix)^\alpha u]\widehat{\ }, \tag{4.2}$$

$$\int_{\mathbb{R}^N} \widehat{u}\, v\, dx = \int_{\mathbb{R}^N} u\, \widehat{v}\, dx, \tag{4.3}$$

$$u*v \in \mathcal{S}, \qquad (u*v)\widehat{\ } = (2\pi)^{N/2} \widehat{u}\, \widehat{v} \quad \text{ in } \mathcal{S}, \tag{4.4}$$

$$uv \in \mathcal{S}\,, \qquad (uv)\widehat{\ } = (2\pi)^{-N/2} \widehat{u} * \widehat{v} \quad \text{ in } \mathcal{S}. \tag{4.5}$$

The conjugate Fourier transform $\widetilde{\mathcal{F}}$ fulfills analogous properties, with $-i$ in place of i.

Proof. The first part is easily checked by repeated use of the Leibniz rule, because of the stability of the space $\mathcal{S}$ w.r.t. multiplication by any polynomial and w.r.t. application of any differential operator (with constant coefficients). [Ex] Actually, $\mathcal{S}$ is the smallest space that contains L^1 and has these properties.

The formulas (4.1)—(4.1) are just particular cases of (3.26), (3.27), (3.34), since $\mathcal{S} \subset L^1$.

It is easily checked that $uv, u*v, (u*v)\widehat{\ }, \widehat{u}*\widehat{v} \in \mathcal{S}$. The equality of (4.4) is a direct extension of (3.35). By writing (3.35) with $\widehat{u}$ and $\widehat{v}$ in place of u and v, and $\widetilde{\mathcal{F}}$ in place of $\mathcal{F}$, we have

$$\widetilde{\mathcal{F}}(\widehat{u}*\widehat{v}) = (2\pi)^{N/2} \widetilde{\mathcal{F}}(\widehat{u})\, \widetilde{\mathcal{F}}(\widehat{v}) = (2\pi)^{N/2} u\, v.$$

By applying $\mathcal{F}$ to both members of this equality, the equality of (4.5) follows.

The final statement is obvious. □

Remark. The identity of (4.5) extends (3.35), that we proved in L^1; on the other hand the identity of (4.5) is meaningless for $u, v \in L^1$.

Fourier transform in $\mathcal{S}'$. Let us first define $\mathcal{F}^\tau : \mathcal{S}' \to \mathcal{S}'$ to be the transpose operator of $\mathcal{F}|_{\mathcal{S}} : \mathcal{S} \to \mathcal{S}$, that is:

$$\langle \mathcal{F}^\tau T, v \rangle := \langle T, \mathcal{F} v \rangle \qquad \forall v \in \mathcal{S}, \forall T \in \mathcal{S}'. \tag{4.6}$$

The conjugate transform $\widetilde{\mathcal{F}}^\tau$ is similarly extended to $\mathcal{S}'$ by transposition. Actually, any results that holds for $\mathcal{F}$ may be reproduced in terms of $\widetilde{\mathcal{F}}^\tau$.

• **Theorem 4.2** *(i) The operator $\mathcal{F}^\tau : \mathcal{S}' \to \mathcal{S}'$ is linear, sequentially continuous, and is the unique continuous extension of $\mathcal{F}$ to $\mathcal{S}'$.*

(ii) The formulas of Proposition 3.2 and Theorem 3.4 hold for $\mathcal{F}^\tau$ in $\mathcal{S}'$ without any restriction.

(iii) The operator $\mathcal{F}^\tau$ is invertible in $\mathcal{S}'$ and $(\mathcal{F}^\tau)^{-1} = \widetilde{\mathcal{F}}^\tau$.

Proof. By 4.3,

$$\langle \mathcal{F}^\tau u, v\rangle = \langle u, \mathcal{F}v\rangle = \int u(x)\, \mathcal{F}v(x)\, dx = \int \mathcal{F}u(x)\, v(x)\, dx \qquad \forall u, v \in \mathcal{S}. \tag{4.7}$$

Thus $\mathcal{F}^\tau|_{\mathcal{S}} = \mathcal{F}|_{\mathcal{S}}$. As $\mathcal{S}$ is sequentially dense in $\mathcal{S}'$, we infer part (i).

Part (ii) is easily derived from the analogous statements for $\mathcal{S}$ via transposition. For instance, we retrieve (3.26) in $\mathcal{S}'$ as follows:

$$\begin{aligned} {}_{\mathcal{S}'}\langle (i\xi)^\alpha \widehat{T}, v\rangle_{\mathcal{S}} &= {}_{\mathcal{S}'}\langle \widehat{T}, (i\xi)^\alpha v\rangle_{\mathcal{S}} = {}_{\mathcal{S}'}\langle T, [(i\xi)^\alpha v]\widehat{\ }\rangle_{\mathcal{S}} \\ &\overset{(3.27)}{=} {}_{\mathcal{S}'}\langle T, (-D)^\alpha \widehat{v}\rangle_{\mathcal{S}} = {}_{\mathcal{S}'}\langle D^\alpha T, \widehat{v}\rangle_{\mathcal{S}} = {}_{\mathcal{S}'}\langle (D^\alpha T)\widehat{\ }, v\rangle_{\mathcal{S}} \qquad \forall T \in \mathcal{S}', \forall v \in \mathcal{S}. \end{aligned}$$

(Here we applied (3.27) exchanging the roles of x and ξ: $[(i\xi)^\alpha v(x)]\widehat{\ }(\xi) = (-D_x)^\alpha \widehat{v}(x)$). Thus, $(i\xi)^\alpha \widehat{T} = (D^\alpha T)\widehat{\ }$.

As we already pointed out, $(\mathcal{F}|_{\mathcal{S}})^{-1} = \widetilde{\mathcal{F}}|_{\mathcal{S}}$. As $\mathcal{S}$ is sequentially dense in $\mathcal{S}'$, part (iii) then follows. □

On the basis of the foregoing result, henceforth we shall write $\mathcal{F}$ in place of $\mathcal{F}^\tau$ (omitting the transposition) also in $\mathcal{S}'$.

Next we extend to $\mathcal{E}'$ the Fourier-Laplace transform of the previous section. First, we define this transformed function on $\mathbb{R}^N$; afterwards in Theorem 4.4 we extend it to a holomorphic function defined on the whole $\mathbb{C}^N$.

Theorem 4.3 *For any $T \in \mathcal{E}'$,*

$$\widehat{T}(\xi) = {}_{\mathcal{E}'}\langle T, e^{-ix\cdot\xi}\rangle_{\mathcal{E}} \qquad \forall \xi \in \mathbb{R}^N. \tag{4.8}$$

* *Proof.* For any $n \in \mathbb{N}$, let us define the mollifier ρ_n as in (3.25), and set $(T * \rho_n)(x) := \langle T_y, \rho_n(x-y)\rangle$ for any $x \in \mathbb{R}^N$. (The index y indicates that T acts on the variable y; on the other hand here x is just a parameter.) By testing on a function of $\mathcal{E}$, it is easily checked that $T * \rho_n \to T$ in $\mathcal{E}'$, hence also in $\mathcal{S}'$ as $\mathcal{E}' \subset \mathcal{S}'$ with continuous and sequentially dense injections. Therefore

$$(T * \rho_\varepsilon)\widehat{\ } \to \widehat{T} \qquad \text{in } \mathcal{S}'. \tag{4.9}$$

On the other hand, as $T * \rho_n \in \mathcal{E}$ and $\int_{\mathbb{R}^N} \rho_n(x)\, dx = 1$, we have

$$\begin{aligned} (T * \rho_n)\widehat{\ }(\xi) &= (2\pi)^{-N/2} \int_{\mathbb{R}^N} e^{-i\xi\cdot x} (T * \rho_n)(x)\, dx \\ &= (2\pi)^{-N/2} \int_{\mathbb{R}^N \times \mathbb{R}^N} e^{-i\xi\cdot x} \langle T_y, \rho_n(x-y)\rangle\, dx\, dy \\ &= (2\pi)^{-N/2} \langle T_y, e^{-i\xi\cdot y} \int_{\mathbb{R}^N} e^{-i\xi\cdot(x-y)} \rho_n(x-y)\, dx\rangle = \langle T_y, e^{-i\xi\cdot y}\rangle \widehat{\rho}_n(\xi), \end{aligned}$$

and this is a holomorphic function of ξ. As $\varepsilon \to 0$, $\widehat{\rho}_n(\xi) \to 1$ uniformly on any compact subset of $\mathbb{R}^N$. Therefore

$$(T * \rho_n)\widehat{\ }(\xi) = \langle T_y, e^{-i\xi\cdot y}\rangle \widehat{\rho}_n(\xi) \to \langle T_y, e^{-i\xi\cdot y}\rangle \qquad \text{in } \mathcal{S}'.$$

By (4.9) we then conclude that $\widehat{T}(\xi) = \langle T_y, e^{-i\xi\cdot y}\rangle$ for any $\xi \in \mathbb{R}^N$, and this function is holomorphic. □

The next result concerns tempered distributions that have a holomorphic Fourier-Laplace transform (after extension to the whole $\mathbb{C}^N$).

*** Theorem 4.4 (Paley-Wiener-Schwartz)** *Let $T \in \mathcal{S}'$ and $R > 0$. Then* $\operatorname{supp}(T) \subset B(0,R)$ *iff $\mathcal{F}(T)$ may be extended to a holomorphic function $\widehat{T} : \mathbb{C}^N \to \mathbb{C}$ such that*

$$\exists m \in \mathbb{N}_0, \exists C > 0 : \forall z \in \mathbb{C}^N \qquad |\widehat{T}(z)| \le C(1+|z|)^m e^{R|\mathrm{Im}(z)|}. \qquad [] \tag{4.10}$$

(Here C depends on m and T.)

Fourier transform in L^2. The operator $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ is the point of arrival of our extensions of the Fourier transform, as $\mathcal{S}'$ includes all other classes of functions on which we define this transform. Nevertheless it will be useful to know when the transformed function of a function is a function.

Next we show that (the restriction of) $\mathcal{F}$ maps L^2 to itself and is an isometric isomorphism.

• Theorem 4.5 (Plancherel) *The Fourier transform may be extended to L^2:*

$$u \in L^2 \quad \Leftrightarrow \quad \widehat{u} \in L^2 \qquad \forall u \in \mathcal{S}'. \tag{4.11}$$

The restriction of $\mathcal{F}$ to L^2 is an isometry, with inverse $\widetilde{\mathcal{F}}$:

$$\int_{\mathbb{R}^N} \widehat{u}\bar{\widehat{v}}\, dx = \int_{\mathbb{R}^N} u\bar{v}\, dx, \qquad \|\widehat{u}\|_{L^2} = \|u\|_{L^2} \qquad \forall u, v \in L^2. \tag{4.12}$$

Moreover, for any $u \in L^2$,

$$(2\pi)^{-N/2} \int_{]-R,R[^N} e^{-i\xi\cdot x} u(x)\, dx \;\to\; \widehat{u}(\xi) \qquad \textit{in } L^2, \;\textit{ as } R \to +\infty. \tag{4.13}$$

Therefore this integral also converges in measure on all bounded subsets of $\mathbb{R}^N$; it also converges a.e., as $R \to +\infty$ along a suitable sequence which depends on u. [19]

Proof. Denoting the inverse Fourier transform by the tilde, it is easily checked that $\widehat{\bar{\widehat{v}}} = \bar{\widetilde{\widetilde{v}}} = \bar{v}$ for any $v \in \mathcal{S}$. Hence

$$\int_{\mathbb{R}^N} \widehat{u}\bar{\widehat{v}}\, dx \overset{(3.34)}{=} \int_{\mathbb{R}^N} u\widehat{\bar{\widehat{v}}}\, dx = \int_{\mathbb{R}^N} u\bar{v}\, dx \qquad \forall u, v \in \mathcal{S}.$$

$\mathcal{F}\big|_{\mathcal{S}}$ is thus a surjective isometry with respect to the L^2-metric. As $\mathcal{S} \subset L^2$ with continuous and dense injection, the same holds for $\mathcal{F}\big|_{L^2}$.

In order to prove (4.13), let us set $\chi_R = \chi_{]-R,R[^N}$ for any $R > 0$, and notice that $u\chi_R \in L^1 \cap L^2$ and $u\chi_R \to u$ in L^2 as $R \to +\infty$. By (4.12) then

$$(2\pi)^{-N/2} \int_{]-R,R[^N} e^{-i\xi\cdot x} u(x)\, dx = (u\chi_R)\widehat{\ } \to \widehat{u} \qquad \text{in } L^2. \qquad \square$$

Proposition 4.6 *For any $u \in L^1_{\mathrm{loc}} \cap \mathcal{S}'$,*

$$(2\pi)^{-N/2} \int_{]-R,R[^N} e^{-i\xi\cdot x} u(x)\, dx \;\to\; \widehat{u}(\xi) \qquad \textit{in } \mathcal{S}', \;\textit{ as } R \to +\infty. \tag{4.14}$$

[19] This is reminiscent of the principal value of Cauchy, but does not coincide with it.

Proof. Let us denote by χ_R the characteristic function of the N-dimensional interval $]-R,R[^N$, and notice that $(u\chi_R)\widehat{\ }(\xi) = (2\pi)^{-N/2} \int e^{-i\xi\cdot x} u(x)\chi_R \, dx$, for a.e. ξ. As $u\chi_R \to u$ in $\mathcal{S}'$ as $R \to +\infty$ the continuity of $\mathcal{F}$ in $\mathcal{S}'$ yields (4.14). $\square$

Remark. One may thus generalize the inversion Theorem 3.8 to the whole L^2.

Lemma 4.7 *For any $p \in [1,2]$, any function $u \in L^p$ may be written as the sum of a function of L^1 and one of L^2, i.e.,*

$$L^p \subset L^1 + L^2 \qquad \forall p \in [1,2]. \tag{4.15}$$

Proof. Setting $\chi := 1$ where $|u| \geq 1$ and $\chi := 0$ elsewhere,

$$u\chi \in L^1, \qquad u(1-\chi) \in L^2, \qquad u = u\chi + u(1-\chi) \qquad \forall u \in L^p. \square \tag{4.16}$$

(Similarly one can show that $L^p \subset L^q + L^r$ whenever $1 \leq q < p < r \leq \infty$.)

As $\mathcal{F} : L^1 \to L^\infty$ and $\mathcal{F} : L^2 \to L^2$, because of this lemma

$$\mathcal{F}(u) = \mathcal{F}(u\chi) + \mathcal{F}(u(1-\chi)) \in L^\infty + L^2 \qquad \forall u \in L^p, \forall p \in [1,2]\,. \tag{4.17}$$

In this case $\mathcal{F}(u)$ thus is a regular distribution, although it may admit no integral representation. This is made more precise by the next result.

Theorem 4.8 (Hausdorff-Young) *Let $p \in [1,2]$, $p' := p/(p-1)$ if $p > 1$, and $p' = \infty$ if $p = 1$. Then (the restriction of) $\mathcal{F}$ is a linear and continuous operator $L^p \to L^{p'}$, and*

$$\|\widehat{u}\|_{L^{p'}} \leq \|u\|_{L^p} \qquad \forall u \in L^p. \tag{4.18}$$

Proof. The restriction of $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ maps L^1 to L^∞ and L^2 to L^2, and is continuous in these spaces. We may thus regard $\mathcal{F}$ as an operator $L^1 + L^2 \to L^\infty + L^2$. The Riesz-Thorin Theorem 2.4 then entails the thesis. $\square$

Remarks. (i) Because of the symmetry between the definition of the direct and inverse Fourier transform, see formulas (3.4) and (3.37), the results that we established for $\mathcal{F}$ in $\mathcal{S}$ and in $\mathcal{S}'$ hold also for $\mathcal{F}^{-1}$, which is defined on the whole $\mathcal{S}'$. In particular this applies to Theorems 4.1 and 4.5.

(ii) Next we show that $\mathcal{F}$ maps L^p to L^q only if $q = p'$. Let $u \in L^p$ be such that $\mathcal{F}(u) \in L^q$. For any $\lambda > 0$, setting $u_\lambda(x) := u(\lambda x)$ for any x, by (3.20) we have $\mathcal{F}(u_\lambda) = \lambda^{-N}\mathcal{F}(u)_{1/\lambda}$. Hence, for any nonidentically vanishing $u \in \mathcal{S}$,

$$\frac{\|\mathcal{F}(u_\lambda)\|_{L^q}}{\|u_\lambda\|_{L^p}} = \lambda^{-N} \frac{\|\mathcal{F}(u)_{1/\lambda}\|_{L^q}}{\|u_\lambda\|_{L^p}} = \lambda^{N(-1+1/q+1/p)} \frac{\|\mathcal{F}(u)\|_{L^q}}{\|u\|_{L^p}}, \tag{4.19}$$

and the left side is uniformly bounded w.r.t. λ iff $q = p'$.

Overview of the extensions of the Fourier transform. At first we noticed that the Fourier transform (3.5) has an obvious extension for any complex Borel measure μ; loosely speaking, this is just defined by replacing $u(x)dx$ by $d\mu$ in the definition (3.4). By the Paley-Wiener theorem, $\mathcal{D}$ is not stable under application of the Fourier transform, so that $\mathcal{F}$ cannot be extended by continuity to the whole $\mathcal{D}'$. However, $\mathcal{F}$ maps the Schwartz space $\mathcal{S}$ to itself, and this allowed us to extend

$\mathcal{F}$ to $\mathcal{S}'$ by transposition. We also saw that $\mathcal{F}$ is an isometry in L^2 (Plancherel theorem), that in this space $\mathcal{F}$ admits an integral representation as a principal value, and that $\mathcal{F}$ is also linear and continuous from L^p to $L^{p/(p-1)}$, for any $p \in \,]1, 2[$.

Note: The Fourier transform is a homomorphism from the algebra $(L^1, *)$ to the algebra $(L^\infty, \cdot)$ (here "$\cdot$" stands for the product a.e.), and is an isomorphism between the algebras $(\mathcal{S}, *)$ and $(\mathcal{S}, \cdot)$; cf. (3.35).

ARed: Finally, ... the Fourier series arise as Fourier transforms of periodic functions.

4.1 Exercises

— Evaluate the Fourier transform of the Heaviside function.

5 Fourier Transform and Differential Equations

Let $a_0, ..., a_m \in \mathbb{C}$ ($a_m \neq 0$), $f : \mathbb{R} \to \mathbb{C}$ be a given function, and consider the linear ODE

$$P(D)u(t) := \sum_{n=0}^{m} a_n D^n u(t) = f(t) \qquad t \in \mathbb{R}. \tag{5.20}$$

If we confine ourselves to functions, it is especially convenient to assume that f, u and its (distributional) derivatives up to order m are elements of L^2. By applying the Fourier transform to both members of (5.20), we get

$$P(D)u = f \qquad \Leftrightarrow \qquad P(i\xi)\widehat{u} = \widehat{f}; \tag{5.21}$$

the differential equation is thus equivalent to the algebraic equation

$$\big(P(i\xi)\widehat{u}(\xi) =\big) \sum_{n=0}^{m} a_n (i\xi)^n \widehat{u}(\xi) = \widehat{f}(\xi) \qquad \forall \xi \in \mathbb{R}.$$

If

$$\big(P(i\xi) =\big) \sum_{n=0}^{m} a_n (i\xi)^n \neq 0 \qquad \forall \xi \in \mathbb{R}, \tag{5.22}$$

then the equation (5.20) is equivalent to

$$\widehat{u}(\xi) = \frac{\widehat{f}(\xi)}{P(i\xi)} \qquad \forall \xi \in \mathbb{R}. \tag{5.23}$$

The second member of (5.23) is Fourier-antitransformable, since it is an element of $\mathcal{S}'$ because of (5.22). At this point it suffices to invert the Fourier transform in L^2; (5.23) is then equivalent to

$$u = \mathcal{F}^{-1}(\widehat{u}) = \mathcal{F}^{-1}\left(\frac{\widehat{f}(\xi)}{P(i\xi)}\right). \tag{5.24}$$

Because of the Parseval theorem, this also reads

$$u = (2\pi)^{1/2} \mathcal{F}^{-1}\left(\frac{1}{P(i\xi)}\right) * \mathcal{F}^{-1}(\widehat{f}) = (2\pi)^{1/2} \mathcal{F}^{-1}\left(\frac{1}{P(i\xi)}\right) * f. \tag{5.25}$$

This is the unique solution of the equation (5.20) in L^2, under the hypothesis (5.22). For instance, for any $\eta \in \mathbb{R}\setminus\{0\}$, the equation $Du - \eta u = 0$ is transformed to $(i\xi - \eta)\hat{u} = 0$. As $P(i\xi) = i\xi - \eta \neq 0$ for any $\xi \in \mathbb{R}$, this homogenous equation has the unique solution $u = e^{\eta t} \in \mathcal{D}' \setminus \mathcal{S}'$.

If the condition (5.22) is not fulfilled, the solution of the equation $P(D)u = f$ need not be unique in $\mathcal{D}'$, and is determined up to the sum of any solution $u \in \mathcal{D}'$ of the homogeneous equation $P(D)u = 0$. E.g., this occurs for $P(D) = D^2$.

Remarks. (i) Notice that here we set the equation on the whole $\mathbb{R}$, without boundary conditions. In this case the solution does not include any additive constant, as the boundary conditions are actually surrogated by the L^2-integrability. [20]

(ii) The above procedure is easily extended to any PDE of the form

$$P(D)u(x) := \sum_{|\alpha|\leq m} a_\alpha D^\alpha u(x) = f(x) \qquad x \in \mathbb{R}^N \tag{5.26}$$

(where by α we denote a multi-index of $\mathbb{N}^N$), provided that $\sum_{|\alpha|\leq m} a_\alpha \xi^\alpha \neq 0$ for any $\xi \in \mathbb{R}^N$.

(iii) In the next chapter the Cauchy problem will be studied on the whole line via the Laplace transform. □

The fundamental solution. [21] The differential equation (5.20) may be addressed from the point of view of *system theory.* The (here assumed unique) solution is interpreted as the *response* of a linear system that is defined by the differential operator. The system is thus characterized by the inverse operator $L : f \mapsto u$. As the coefficients $\{a_n\}$ do not depend on t, this operator is invariant by time shifts.

Next let us go beyond functions, and assume that $f, u \in \mathcal{S}'$. If $f = \delta_0$ in (5.25), then the solution $h = Lu$ represents the response of a system to δ_0 (the unit pulse at the origin). In the theory of linear differential equations (ODEs or PDEs) this is called a *fundamental solution,* whereas in the theory of linear systems the term *impulsive response in time* is used:

$$h := (2\pi)^{N/2}\mathcal{F}^{-1}\left(\frac{1}{P(i\xi)}\right) * \delta_0 = (2\pi)^{N/2}\mathcal{F}^{-1}\left(\frac{1}{P(i\xi)}\right). \tag{5.27}$$

$[\mathcal{F}(h)](\xi) = (2\pi)^{N/2}/P(i\xi)$ is then called the *impulsive response in frequency* (or just the *response function).* [22] We may actually rewrite the solution of the equation (5.20) in the form

$$u = h * f \quad \text{ with } \quad \sum_{n=0}^{m} a_n D^n h = \delta_0. \tag{5.28}$$

Distributional solutions. Because of the fundamental theorem of algebra, the *characteristic equation*

$$\big(P(i\xi) =\big) \sum_{n=0}^{m} a_n (i\xi)^n = 0$$

[20] This may be somehow understood by recalling that the linear space of continuous functions $\mathbb{R} \to \mathbb{C}$ with compact support is dense in L^2...

[21] These notions will be revisited in the chapter devoted to filters.

[22] This terminology is typical of the *theory of linear systems,* that we shall briefly illustrate ahead. The term *fundamental solution* is used in mathematical analysis; in this set-up whenever the null value is assumed either at the boundary or at infinity, one may also speak of a *Green function.*

has exactly m (possibly repeated) complex roots. Let us denote by $\{\xi_j : j = 1, ..., \ell\}$ the distinct roots, and by r_j the multiplicity of ξ_j for any j; thus $\ell \leq m$ and $r_1 + ... + r_\ell = m$. Therefore

$$P(i\xi) = a_m \prod_{j=1}^{\ell} (i\xi - i\xi_j)^{r_j} \qquad \forall \xi \in \mathbb{C},$$

whence

$$P(D) = a_m \prod_{j=1}^{\ell} (D - i\xi_j)^{r_j}.[23]$$

As

$$(D - i\xi_j)^{r_j}(t^{k-1}e^{i\xi_j t}) = 0 \qquad \text{for } k = 1, ..., r_j,\ j = 1, ..., \ell,$$

the roots $\{\xi_j : j = 1, ..., \ell\}$ are associated to a linearly independent family of m solutions:

$$u_{j,k}(t) = t^{k-1}e^{i\xi_j t} \qquad (k = 1, ..., r_j,\ j = 1, ..., \ell)$$

of the homogeneous differential equation $\sum_{n=0}^{m} a_n D^n u(t) = 0$. For each j,

$$\begin{cases} \xi_j \in \mathbb{R} & \Rightarrow \quad u_{j,1}, ..., u_{j,r_j} \in \mathcal{S}' \\ \xi_j \notin \mathbb{R} & \Rightarrow \quad u_{j,1}, ..., u_{j,r_j} \in \mathcal{D}' \setminus \mathcal{S}' \end{cases} \qquad \text{for } j = 1, ..., \ell. \tag{5.29}$$

We shall distinguish two cases:

(i) If the condition (5.22) is fulfilled, defining the function h as in (5.27) we conclude that

$$\begin{aligned} &h \text{ is the unique fundamental solution of the equation (5.20) in } \mathcal{S}'; \\ &\text{hence } u = h * f \text{ is the unique solution of the equation (5.20) in } \mathcal{S}', \forall f \in \mathcal{S}'. \end{aligned} \tag{5.30}$$

(ii) If instead $\big(P(i\widetilde{\xi}) = \big) \sum_{n=0}^{m} a_n (i\widetilde{\xi})^n = 0$ for some $\widetilde{\xi} \in \mathbb{R}$, then this root corresponds to the solution $u(t) = e^{i\widetilde{\xi}t} \in \mathcal{S}'$ of the homogeneous equation (5.20). In this case

$$\begin{aligned} &\text{the fundamental solution is not unique in } \mathcal{S}'; \\ &\text{hence the solution of the nonhomogenous (5.20) also fails to be unique in } \mathcal{S}'. \end{aligned} \tag{5.31}$$

Examples. Let us fix any $k > 0$ and consider two differential equations

$$u - k^2 u'' = f(t), \qquad u + k^2 u'' = f(t). \tag{5.32}$$

These equations are respectively associated to the operators

$$P_1(D) := I - k^2 D^2, \qquad P_2(D) := I + k^2 D^2 \qquad (I\text{: operatore identità}),$$

which in turn correspond to the characteristic polynomials

$$P_1(i\xi) = 1 + k^2\xi^2, \qquad P_2(i\xi) = 1 - k^2\xi^2 \qquad (\xi \in \mathbb{R}).$$

The hypothesis (5.22) is satisfied by $P_1(i\xi)$, but not by $P_2(i\xi)$. The previous analysis may thus be applied just to the first equation, which therefore has a unique solution in $\mathcal{S}'$; moreover, the

[23] Here $i\xi_j$ stands for $i\xi_j I$, where I is the identity operator, and the product represents the composition product.

second equation has more than one solution in $\mathcal{S}'$. For the latter equation, mathematically it is more appropriate to address the initial-value problem rather than the problem on the whole $\mathbb{R}$, and use the Laplace transform, as we shall see in the next section. This mathematical aspect also reflects typical applications.

This discussion can be extended to systems of linear ODEs, and also to some nonlinear problems. The extension to PDEs is also viable, although more complex.

5.1 Exercises

1. Generalize the analysis of the ODEs (5.32) to the PDEs

$$u - \Delta u = f, \qquad u + \Delta u = f \qquad \text{in } \mathbb{R}^N. \tag{5.33}$$

6 Uncertainty Principle

Doppler theorem and uncertainty principle. (3.20) yields the following scaling formula (which is also referred to as the *Doppler theorem):*

$$v_a(x) = a^{-N/2}u(x/a) \quad \Rightarrow \quad \widehat{v}_a(\xi) = a^{N/2}\widehat{u}(a\xi) \qquad \forall a > 0, \forall u \in L^1. \tag{6.1}$$

(This scaling is such that $\|v_a\|_{L^2}$ and $\|\widehat{v}_a\|_{L^2}$ are independent of a.)

For instance, we shall see that the Gaussians function $u(x) = e^{-|x|^2/2}$ corresponds to $\widehat{u}(\xi) = e^{-|\xi|^2/2}$ (see (3.31)). By (6.1), $v_a(x) = a^{-N/2}e^{-|x/a|^2/2}$ then corresponds to $\widehat{v}_a(\xi) = a^{N/2}e^{-|a\xi|^2/2}$. All of these functions have the L^2-norm equal to 1. By varying a, we see that the more v_a is spread, the less $\widehat{v}_a$ is spread; and conversely, the less v_a is spread, the more $\widehat{v}_a$ is spread.

This behaviour is representative of a general result, which now we briefly illustrate.

Theorem 6.1 (Heisenberg Uncertainty Principle) *Let $u \in L^2$. By the Plancherel Theorem, $E := \|u\|_{L^2}^2 = \|\widehat{u}\|_{L^2}^2$, so that $|u|^2/E$ and $|\widehat{u}|^2/E$ are densities of probability. Let us assume that the respective means m_1, m_2 and variances σ_1^2, σ_2^2 exist and are finite:*

$$m_1 = \frac{1}{E}\int t|u|^2(t)\,dt \in \mathbb{R}, \qquad \sigma_1^2 = \frac{1}{E}\int (t-m_1)^2|u|^2(t)\,dt < +\infty, \tag{6.2}$$

$$m_2 = \frac{1}{E}\int \xi|\widehat{u}|^2(\xi)\,d\xi \in \mathbb{R}, \qquad \sigma_2^2 = \frac{1}{E}\int (\xi-m_2)^2|\widehat{u}|^2(\xi)\,d\xi < +\infty. \tag{6.3}$$

Then

$$\sigma_1\sigma_2 \geq 1/2. \tag{6.4}$$

Equality holds if $u = v_a(x) = a^{-N/2}e^{-|x/a|^2/2}$, for any $a > 0$.

Proof. Without loss of generality, we may assume that $m_1 = m_2 = 0$. Let us first assume that $u \in C^\infty$ with compact support. We known that then

$$\int \xi^2|\widehat{u}(\xi)|^2\,d\xi = \int |\widehat{u'}(\xi)|^2\,d\xi = \int |u'(t)|^2\,dt. \tag{6.5}$$

Notice that

$$\begin{aligned} 2\,\mathrm{Re}\int tu(t)\overline{u'(t)}\,dt &= \int t\Big\{\overline{u(t)}u'(t) + u(t)\overline{u'(t)}\Big\}\,dt = \int t\frac{d}{dt}|u(t)|^2\,dt \\ &= t|u(t)|^2\Big|_{t=-\infty}^{t=+\infty} - \|u\|_{L^2}^2 = -\|u\|_{L^2}^2 \end{aligned}$$

and

$$\left|\mathrm{Re}\int tu(t)\overline{u'(t)}\,dt\right| \leq \|tu\|_{L^2}\|u'\|_{L^2} = \|tu(t)\|_{L^2}\|\xi\widehat{u}(\xi)\|_{L^2}(= \sigma_1\sigma_2). \tag{6.6}$$

The two latter formulae yield

$$2\|tu(t)\|_{L^2}\|\xi\widehat{u}(\xi)\|_{L^2} \geq \|u\|_{L^2}^2. \tag{6.7}$$

(6.4) thus holds for any $u \in C^\infty$ with compact support. As this space is dense in L^2 and the Fourier transform is isometric w.r.t. the L^2 metric, the thesis follows. □

If $|u|^2$ is the density of probability of time localization of an event, then $|\widehat{u}|^2$ is the density of probability of its frequency. The time-frequency resolution of u is represented in the time-frequency

(t,ξ)-plane (t,ξ) by a rectangular *Heisenberg box* centered at $(t,\xi)=(m_1,m_2)$, and with width σ_{a1} (σ_{a2}, resp.) along the t-axis (ξ-axis, resp.).

By the Heisenberg inequality $\sigma_1\sigma_2 \geq 1/2$, the sharper is the information we have on the time localization, the rougher is the information we may get on the frequency: time resolution and frequency resolution thus conflict each other.

For instance, a Dirac measure $u = \delta_{t_0}$ is the most precise specification of time location; its Fourier transform is the sinusoid $\widehat{u}(\xi) = e^{-i\xi t_0}/\sqrt{2\pi}$, which is supported over the whole $\mathbb{R}$, and thus is completely unlocalized. Dually, a Dirac measure $\widehat{u} = \delta_{\xi_0}$ is the most precise specification of frequency location; its Fourier antitransform is the sinusoid $u(t) = e^{i\xi_0 t}/\sqrt{2\pi}$, which is supported over the whole $\mathbb{R}$, and thus is completely unlocalized.

Remark. For any $u \in L^1$ and any $a > 0$, let us define v_a as in (6.1) and the standard deviation σ_{a1} (σ_{a2}, resp.) of v_a ($\widehat{v_a}$, resp.) as above. Notice that the product $\sigma_{a1}\sigma_{a2}$ is independent of a. Loosely speaking, this scaling is thus consistent with the uncertainty principle.