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 ordinary differential equations.
6. Rescaled Fourier transform.
7. The Poisson summation formula.
8. The sampling theorem.
9. The uncertainty principle.

Symbols. • indicates important results. $*$ indicates complements.
[Ex] means that the proof is left as exercise. [] means that the 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 $\mathcal{A}(\Omega)$ (which must be defined):

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

One is thus induced to consider all the functionals of the topological dual $\mathcal{A}'(\Omega)$ of $\mathcal{A}(\Omega)$. In this way several classes of distributions are generated. The space $\mathcal{A}(\Omega)$ must be so large that the functional T_f determines a unique f. On the other hand, the smaller is the space $\mathcal{A}(\Omega)$, the larger is its topological dual $\mathcal{A}'(\Omega)$. It happens that there exists a very small space $\mathcal{A}(\Omega)$, so that $\mathcal{A}'(\Omega)$ is very large (actually, this is the largest space used in analysis); 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 $\mathbb{R}^N$. Let us denote by $\mathcal{D}(\Omega)$ the space of infinitely differentiable functions $\Omega \to \mathbb{C}$ whose support is a compact subset of Ω. These are called *test functions.*

By the theorem of analytic continuation, the null function is the only analytic function in $\mathcal{D}(\Omega)$, since any element of this space vanishes in some open set. By a classical example of Cauchy, the function

$$f(x) = e^{-1/|x|^2} \qquad \forall x \neq 0, \qquad f(0) = 0$$

is infinitely differentiable but not analytic. It is then easy to check that 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. In particular, for any open neighbourhood of any $x_0 \in \Omega$ contains the support of a bell-shaped functions.

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. These are linear subspaces of $C^\infty(\Omega)$, and $\mathcal{D}(\Omega) = \bigcup_{K \subset\subset \Omega} \mathcal{D}_K(\Omega)$. The space $\mathcal{D}(\Omega)$ will be equipped with the corresponding *inductive-limit topology.* [1] Here we shall not study the subtle properties of 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 $\alpha \in \mathbb{N}^N$.

A sequence $\{u_n\}$ in $\mathcal{D}(\Omega)$ converges to $u \in \mathcal{D}(\Omega)$ in the inductive topology iff, for some $K \subset\subset \Omega$,

$$u_n, u \in \mathcal{D}_K(\Omega) \qquad \forall n, \qquad D^\alpha u_n \to D^\alpha u \qquad \text{uniformly in } K, \forall \alpha \in \mathbb{N}^N.$$

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) and $\{a_n\}$ is a sequence in $\mathbb{R}^N$, 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 sequentially continuous, i.e., $T(v_n) \to 0$ whenever $v_n \to 0$ in $\mathcal{D}(\Omega)$;

(iii)

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

[1] This is the finest topology among those that make all injections $\mathcal{D}_K(\Omega) \to \mathcal{D}(\Omega)$ continuous. This topology makes $\mathcal{D}(\Omega)$ a nonmetrizable locally convex Hausdorff space.

A set $A \subset \mathcal{D}(\Omega)$ is open in this topology iff $A \cap \mathcal{D}_K(\Omega)$ is open for any $K \subset\subset \Omega$.

If $m = m_{T,K}$ is the smallest integer that fulfills this condition, one says that T has order m on the compact set K.

* *Proof.* Obviously (i) entails (ii). Let us prove that (ii) entails (iii). By contradiction, let us assume that

$$\exists K \subset\subset \Omega : \forall m \in \mathbb{N}, \exists v_m \in \mathcal{D}_K(\Omega) : \quad |T(v_m)| > m \max_{|\alpha| \leq m} \sup_K |D^\alpha v_m|.$$

(we have taken $C = m$). Possibly dividing v_m by $T(v_m)$, we can assume that $T(v_m) = 1$ for any m. Hence $\sup_K |D^\alpha v_m| < 1/m$ for any $\alpha \in \mathbb{N}^N$ and any $m \geq |\alpha|$. Thus $v_m \to 0$ in $\mathcal{D}_K(\Omega)$, whence in $\mathcal{D}(\Omega)$, although $T(v_m)$ does not vanish. This contradicts (ii).

Finally, (iii) entails (ii), since if $v_n \to 0$ then $\sup_K |D^\alpha v|$. This simple argument can be extended to nets, so that actually (iii) yields (i). □

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

$$T_n \to T \quad \text{in } \mathcal{D}'(\Omega) \quad \Leftrightarrow \quad T_n(v) \to T(v) \quad \forall v \in \mathcal{D}(\Omega). \tag{1.4}$$

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

We can then define a series of distributions as the limit of a partial sum of distributions.

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

Examples of distributions. (i) For any measurable function $f : \Omega \to \mathbb{C}$, the integral functional

$$T_f : v \mapsto \int_\Omega f(x)\, v(x)\, dx \tag{1.5}$$

is a distribution iff $f \in L^1_{\text{loc}}(\Omega)$. [Ex] [2] The mapping $f \mapsto T_f$ is injective, so that we can identify $L^1_{\text{loc}}(\Omega)$ with a closed subspace of $\mathcal{D}'(\Omega)$. These distributions are called *regular;* the other distributions 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

$$T_\mu : v \mapsto \int_\Omega v(x)\, d\mu(x) \tag{1.6}$$

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

(iii) The function $x \mapsto 1/x$ is not locally integrable in $\mathbb{R}$, and so it is not a distribution. Next we show that, on the other hand, its *principal value* (p.v.),

$$\langle p.v.\ \frac{1}{x}, v\rangle := \lim_{\varepsilon \to 0} \int_{|x|>\varepsilon} \frac{v(x)}{x}\, dx \qquad \forall v \in \mathcal{D}(\mathbb{R}) \tag{1.7}$$

[2] $L^1_{\text{loc}}(\Omega)$ is a Fréchet space.

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

$$\begin{aligned}
\langle p.v.\ \frac{1}{x}, v\rangle &= \lim_{\varepsilon\to 0^+}\left(\int_{\varepsilon<|x|<a}\frac{v(x)-v(0)}{x}\,dx + \int_{\varepsilon<|x|<a}\frac{v(0)}{x}\,dx\right)\\
&\left(\int_{\varepsilon<|x|<a}\frac{v(0)}{x}\,dx = 0 \text{ as the integrand is an odd function}\right)\\
&= \lim_{\varepsilon\to 0^+}\int_{\varepsilon<|x|<a}\frac{v(x)-v(0)}{x}\,dx = \int_{-a}^{a}\frac{v(x)-v(0)}{x}\,dx.
\end{aligned} \tag{1.8}$$

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

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

By the characterization of Theorem 1.1, we conclude that p.v. $1/x$ is a distribution.

Note that the principal value is different from other notions of (Remann-type) *generalized integral.*

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

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

(vi) More generally, if μ is a *Radon measure* (i.e. a Borel measure that is finite on compact subsets of Ω), then $v \mapsto \int v\,d\mu$ is a distribution.

(vii) The series $T = \sum_{n=1}^{\infty}\delta_{x_n}$ is a distribution of $\mathcal{D}'(\Omega)$ iff any $K \subset\subset \Omega$ contains at most a finite number of points of the sequence $\{x_n\}$. [Ex] Indeed, if this condition is fulfilled, then for any $v \in \mathcal{D}(\Omega)$, $T(v)$ is a finite sum; moreover, the condition (1.3) is promptly checked. On the other hand, if that condition fails, then there exists $v \in \mathcal{D}(\Omega)$ such that $v(x_n) = 1$ for any n, so that $T(v)$ diverges. So for instance (for $N = 1$)

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

Differentiation of distributions. We define the multiplication of a distribution by a C^∞-function and the differentiation [3] of a distribution via **transposition**:

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

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

Via the characterization (1.3), it may be checked that $\widetilde{D}^\alpha T$ is a distribution, and that the operator $\widetilde{D}^\alpha$ is continuous in $\mathcal{D}'(\Omega)$. [Ex] Indeed, defining $m_{T,K}$ as above,

$$m_{fT,K} = m_{T,K}, \qquad m_{\widetilde{D}^\alpha T,K} \le m_{T,K} + |\alpha| \qquad \forall K \subset\subset \Omega.$$

Thus any distribution has derivatives of any order. More specifically, for any $f \in C^\infty(\Omega)$, the operators $T \mapsto fT$ and $\widetilde{D}^\alpha$ are linear and continuous in $\mathcal{D}'(\Omega)$. [Ex]

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

The definition (1.10) is consistent with partial integration: for any $f \in L^1_{\rm loc}(\Omega)$ such that $\widetilde{D}^\alpha f \in L^1_{\rm loc}(\Omega)$, (1.10) indeed reads

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

(Here no boundary term occurs as the support of v is compact.) Similarly, multiplication of a distribution by a C^∞-function is consistent with the multiplication of functions of $L^1_{\rm loc}(\Omega)$, namely

$$T_{\varphi f} = \varphi T_f \qquad \forall f \in L^1_{\rm loc}(\Omega), \forall \varphi \in C^\infty(\Omega).$$

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

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

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

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

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

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

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

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

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

NOTE: Any derivative of distributions is the limit of the corresponding incremental ratios.

Comparison with classical derivatives. We just state two results that characterize the coincidence between distributional and ordinary derivatives.

Theorem 1.3 *Du-Bois Reymond] For any $f \in C^0(\Omega)$ and any $i \in \{1, \ldots, N\}$, the two following conditions are equivalent:*
(i) $\widetilde{D}_i f \in C^0(\Omega)$, [4]
(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. []

[4] that is, $\widetilde{D}_i f$ is a regular distribution that can be identified with a function $h \in C^0(\Omega) \cap L^1_{\rm loc}(\Omega)$. Using the notation (1.5), 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 $\Omega :=]a,b[$. First we remind the reader that

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

This entails that f is continuous and $f' = g$ a.e. in $]a,b[$. (Here by f' we denote the ordinary derivative: $f'(x) = \lim_{h\to 0}[f(x+h) - f(x)]/h$, whenever this limit exists and is finite.) Thus if $f \in L^1_{\rm loc}(a,b)$ is absolutely continuous, then it is everywhere continuous, a.e. differentiable (in the classical sense), and $f' \in L^1_{\rm loc}(a,b)$.

The converse does not hold: $f \in L^1_{\rm loc}(a,b)$ need not be absolutely continuous, even if f is everywhere continuous and a.e. differentiable, and $f' \in L^1_{\rm loc}(a,b)$; moreover in this case $\widetilde{D}_i f$ need not be a regular distribution. The classical Cantor function is a counterexample. [Ex] If we drop the condition of continuity, a simpler counterexample is provided by the *Heaviside function* H*:*

$$H(x) := 0 \quad \forall x < 0 \qquad H(x) := 1 \quad \forall x \geq 0. \tag{1.14}$$

Indeed $DH = 0$ a.e. in $\mathbb{R}$, but of course H is not absolutely continuous. Note that $\widetilde{D}H = \delta_0$ since

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

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 not a distribution. We claim that

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

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

$$\begin{aligned}
&\langle D\log|x|, v\rangle = -\langle \log|x|, v'\rangle = -\int_{\mathbb{R}} (\log|x|)\,v'(x)\,dx \\
&= -\lim_{\varepsilon\to 0^+} \int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} (\log|x|)\,v'(x)\,dx \\
&= \lim_{\varepsilon\to 0^+} \left\{ \int_{[-a,a]\setminus[-\varepsilon,\varepsilon]} \frac{1}{x}\,v(x)\,dx + (\log|\varepsilon|)\,[v(\varepsilon) - v(-\varepsilon)] \right\} \\
&\Big(\text{as } \int_{[-a,a]\setminus[-\varepsilon,\varepsilon]} \frac{v(0)}{x}\,dx = 0 \ \text{ and } \lim_{\varepsilon\to 0^+} (\log|\varepsilon|)\,[v(\varepsilon) - v(-\varepsilon)] = 0\Big) \\
&= \lim_{\varepsilon\to 0^+} \int_{[-a,a]\setminus[-\varepsilon,\varepsilon]} \frac{v(x) - v(0)}{x}\,dx \overset{(1.8)}{=} \langle \text{p.v.}\frac{1}{x}, v\rangle.
\end{aligned} \tag{1.16}$$

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

$$
\begin{aligned}
\langle D(\text{p.v.}\frac{1}{x}), v\rangle &= -\langle \text{p.v.}\frac{1}{x}, v'\rangle \overset{(1.8)}{=} -\int_{-a}^{a} \frac{v'(x) - v'(0)}{x}\,dx \\
&= -\int_{-a}^{a} \frac{[v(x) - v(0) - xv'(0)]'}{x}\,dx = \int_{-a}^{a} \frac{v(x) - v(0) - xv'(0)}{x^2}\,dx.
\end{aligned}
\tag{1.17}
$$

(This equality is obtained by partial integration, as the boundary term vanishes.) The final integrand equals $v''(\xi_x)/2$, for some ξ_x between 0 and x. Hence this distribution has order two. Notice that this integral cannot be decomposed into a sum of integrals, not even using principal values.

** (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 any distribution. On the other hand, it is easily seen that the next two limits exist

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

Let us also set $g(0) = 0$. Thus $g(x) = \int_0^x f(t)\,dt$, 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, it is 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 $\text{supp}(v) \subset [-a, a]$, still understanding integrals as generalized Riemann integrals,

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

for a suitable mapping $x \mapsto \xi_x$. If is $v'(0) \neq 0$ then the latter integral is a Lebesgue integral. This completes our representation of Dg.

* (iv) The modifications of (iii) 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 can 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 (**??**) 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 obvious. Let us see the case of $N = 1$, along the lines of [Gilardi: Analisi 3].

Proposition 1.5 *For any $T \in \mathcal{D}'$ and $m \in \mathbb{N}$,*

$$x^m T = 0 \quad \Rightarrow \quad \exists c_0, ..., c_{m-1} \in \mathbb{C} : T = \sum_{n=0}^{m-1} c_n D^n \delta_0. \tag{1.22}$$

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

$$\langle T\big|_{\widetilde{\Omega}}, v\rangle := \langle T, v\rangle \qquad \forall v \in \mathcal{D}(\Omega) \text{ such that } \operatorname{supp}(v) \subset \widetilde{\Omega}. \tag{1.23}$$

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 $\Omega_1, \Omega_2, \Omega_3$,

$$\Omega_1 \subset \Omega_2 \subset \Omega_3 \quad \Rightarrow \quad \big(T\big|_{\Omega_2}\big)\big|_{\Omega_1} = T\big|_{\Omega_1} \qquad \forall T \in \mathcal{D}'(\Omega_3). \tag{1.24}$$

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 $K \subset\subset \Omega$, we defined $m_{T,K}$ as the smallest integer m that fulfills the estimate (1.3), and called it the order of T in K. Here we set $m_T := \sup\{m_{T,K} : K \subset\subset \Omega\}$, and call it the **order** of T; each distribution has thus either finite or infinite order. For instance,

(i) regular distributions and distributions associated to Radon measures (e.g., 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 $\mathcal{D}'(\mathbb{R})$.

For any open set $\widehat{\Omega} \subset \Omega$ and any $T \in \mathcal{D}(\Omega)$, the restriction $T|_{\widehat{\Omega}}$ is defined by transposing the trivial extension $\mathcal{D}(\widehat{\Omega}) \to \mathcal{D}(\Omega)$ (that is, the extension with null value):

$${}_{\mathcal{D}(\widehat{\Omega})}\langle T|_{\widehat{\Omega}}, \varphi\rangle_{\mathcal{D}(\widehat{\Omega})} = {}_{\mathcal{D}(\Omega)}\langle T, \widetilde{\varphi}\rangle_{\mathcal{D}(\widehat{\Omega})} \qquad \forall \varphi \in \mathcal{D}(\Omega).$$

It is easy to check that $T|_{\widehat{\Omega}} \in \mathcal{D}'(\widehat{\Omega})$.

The next statement easily follows from (1.3).

Remarks 1.6 (i) In general distributions cannot be extended to a larger open set: this may fail even for regular distributions.

(ii) For any p, $L^p(\Omega)$ can be trivially identified with $L^p(\widetilde{\Omega})$ whenever Ω and $\widetilde{\Omega}$ are open sets, $\widetilde{\Omega} \subset \Omega$ and $|\Omega \setminus \widetilde{\Omega}| = 0$. The analogous identification fails for L^p_{loc} (if $p \neq \infty$), and a fortiori for distributions: $T \in \mathcal{D}'(\Omega)$ entails $T|_{\widetilde{\Omega}} \in \mathcal{D}'(\widetilde{\Omega})$, but the converse fails, even for regular distributions. Counterexamples are provided by δ_0 and $f(x) = x^{-2}$ in $\Omega = \mathbb{R}$ and in $\widetilde{\Omega} = \mathbb{R} \setminus \{0\}$. □

Theorem 1.7 *Any compactly supported distribution is of finite order.*

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

Theorem 1.8 *Any distribution whose support is the origin is a finite combination of derivatives of the Dirac mass. []*

Exercise. May any distribution of infinite order be approximated by a sequence of distributions of finite order?

The space $\mathcal{E}(\Omega)$ 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 $u \in \mathcal{E}$,

$$\begin{aligned} &u_n \to u \quad \text{in } \mathcal{E}(\Omega) \quad \Leftrightarrow \\ &\sup_{x\in K} |D^\alpha(u_n - u)(x)| \to 0 \quad \forall K \subset\subset \Omega, \quad \forall \alpha \in \mathbb{N}^N. \end{aligned} \tag{1.25}$$

Notice that

$$\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega) \quad \text{with continuous and sequentially dense injection,} \tag{1.26}$$

namely, any element of $\mathcal{E}(\Omega)$ can be approximated by a sequence of $\mathcal{D}(\Omega)$. This can be checked via multiplication by a suitable sequence of compactly supported smooth functions. [Ex] By (1.26)

$$\mathcal{E}'(\Omega) \subset \mathcal{D}'(\Omega) \quad \text{with continuous and sequentially dense injection,} \tag{1.27}$$

so that we can identify $\mathcal{E}'(\Omega)$ with a closed 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 $\{T_n\}$ and any $T \in \mathcal{E}'(\Omega)$,

$$T_n \to T \quad \text{in } \mathcal{E}'(\Omega) \quad \Leftrightarrow \quad T_n(v) \to T(v) \quad \forall v \in \mathcal{E}(\Omega). \tag{1.28}$$

[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 $\{T_n\}$ and T in $\mathcal{E}'(\Omega)$ and any $T \in \mathcal{E}'(\Omega)$,

$$T_n \to T \quad \text{in } \mathcal{E}'(\Omega) \quad \overset{\not\Leftarrow}{\Rightarrow} \quad T_n \to T \quad \text{in } \mathcal{D}'(\Omega).\text{[Ex]} \tag{1.29}$$

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 1.9 $\mathcal{E}'(\Omega)$ *can be identified with the closed 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 can 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 can 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 $f \in L^1_{\rm loc}$ 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) $T = \sum_{n=1}^\infty n^{-2} D^{\alpha_n}\delta_{a_n} \in \mathcal{E}'(\Omega)$, for any sequence $\{a_n\}$ contained in a compact subset of Ω, and any bounded sequence of multi-indices $\{\alpha_n\}$. (If $\{a_n\}$ is not contained in a compact subset of Ω, then $T \in \mathcal{D}'(\Omega)$.)

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

The space $\mathcal{S}$ of rapidly decreasing smooth functions. As we shall see, in order to extend the Fourier transform to distributions, Laurent Schwartz introduced the space of *(infinitely differentiable) rapidly decreasing functions* (at ∞): [5]

$$\begin{aligned}\mathcal{S}\big(\mathbb{R}^N\big) :=& \big\{v \in C^\infty : \forall \alpha, \beta \in \mathbb{N}^N, x^\beta D^\alpha v \in L^\infty\big\} \\ =& \big\{v \in C^\infty : \forall \alpha \in \mathbb{N}^N, \forall m \in \mathbb{N}, \\ & |x|^m D^\alpha v(x) \to 0 \text{ as } |x| \to +\infty\big\}.\end{aligned} \tag{1.30}$$

(It is easy to check that these two sets coincide.) [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 []

$$|v|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^N} |x^\beta D^\alpha v(x)| \qquad \alpha, \beta \in \mathbb{N}^N, \tag{1.31}$$

$$|v|_{m,\alpha} := \sup_{x\in\mathbb{R}^N} (1 + |x|^2)^m |D^\alpha v(x)| \qquad m \in \mathbb{N}, \alpha \in \mathbb{N}^N. \tag{1.32}$$

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

$$u \mapsto P(x)Q(D)u, \qquad u \mapsto P(D)[Q(x)u] \tag{1.33}$$

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

$$\mathcal{D} \subset \mathcal{S} \subset \mathcal{E} \quad \text{with continuous and sequentially dense injections.} \tag{1.34}$$

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

$$\{v_n\} \subset \mathcal{S}, \quad v_n \to 0 \quad \text{in } \mathcal{S} \quad \Rightarrow \quad \langle T, v_n\rangle \to 0. \tag{1.35}$$

[5] 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. $\mathcal{S}(\mathbb{R}^N)$ is also called the *Schwartz space.*

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 $w \in L^1$. These are called *slowly increasing functions,* and include polynomials and L^p for any $p \in [1, +\infty]$.

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

One can show that $\mathcal{S}$ is the smallest subspace of that is stable under differentiation and multiplication by a polynomial.

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 $T \in \mathcal{S}'$,

$$T_n \to T \quad \text{in } \mathcal{S}' \quad \Leftrightarrow \quad T_n(v) \to T(v) \quad \forall v \in \mathcal{S}. \tag{1.36}$$

[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

$$\mathcal{S}' \subset \mathcal{D}' \quad \text{with continuous and sequentially dense injection; [Ex]} \tag{1.37}$$

namely, any element of $\mathcal{D}'$ can 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 $T \in \mathcal{S}'$,

$$T_n \to T \quad \text{in } \mathcal{S}' \quad \overset{\not\Leftarrow}{\Rightarrow} \quad T_n \to T \quad \text{in } \mathcal{D}'.\text{[Ex]} \tag{1.38}$$

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

$$e^{|x|}\chi_{[n,n+1]} \to 0 \qquad \text{in } \mathcal{D}' \text{ but not in } \mathcal{S}'. \tag{1.39}$$

On the other hand $L^1_{\rm loc}$ 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

$$\mathcal{E}' \subset \mathcal{S}' \quad \text{with continuous and sequentially dense injection; [Ex]} \tag{1.40}$$

Because of (1.40), we can apply to $\mathcal{S}'$ the operations that we defined for distributions. It is straightforward to check that this space is continuously transformed to itself by differentiation and by multiplication by a smooth function.

For any $\alpha \in \mathbb{N}^N$, the linear and continuous operators $u \mapsto x^\alpha u$ and $u \mapsto D^\alpha u$ are extended from $\mathcal{S}$ to $\mathcal{S}'$ by transposition:

$$\begin{aligned} \langle x^\alpha T, v\rangle := \langle T, x^\alpha v\rangle, \qquad & \langle D^\alpha T, v\rangle := (-1)^{|\alpha|}\langle T, D^\alpha v\rangle \\ & \forall v \in \mathcal{S}, \forall T \in \mathcal{S}', \forall \alpha \in \mathbb{N}^N. \end{aligned} \tag{1.41}$$

These operators are linear and continuous in $\mathcal{S}'$. The same then holds for the operators $T \mapsto P(x)T$ and $T \mapsto P(D)T$, for any polynomial P of N variables.

For instance, for $N = 1$, the function $x \mapsto \sin(e^x)$ is slowly increasing, hence tempered. Therefore its derivative $x \mapsto e^x \cos(e^x)$ is also tempered, at variance with its absolute value.

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

1.1 Overview and Commentaries

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

$$\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega) \quad \text{with continuous and dense injection.} \tag{1.42}$$

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

$$\mathcal{D} \subset \mathcal{S} \subset \mathcal{E} \quad \text{with continuous and dense injection.} \tag{1.43}$$

We equipped the respective dual spaces with the sequential weak star convergence. (1.42) and (1.43) respectively yield

$$\mathcal{E}'(\Omega) \subset \mathcal{D}'(\Omega) \quad \text{with continuous and sequentially dense injection,} \tag{1.44}$$

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

$$\mathcal{E}' \subset \mathcal{S}' \subset \mathcal{D}' \quad \text{with continuous and sequentially dense injection.} \tag{1.45}$$

Moreover,

$$\mathcal{D} \subset \mathcal{D}', \qquad \mathcal{S} \subset \mathcal{S}', \quad \text{with continuous and sequentially dense injections.} \tag{1.46}$$

On the other hand, $\mathcal{E}$ is not included in $\mathcal{E}'$.

These density properties can be proved via regularization, by a procedure that we shall introduce in the next section.

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

$$(f * g)(x) := \int f(x-y)g(y)\,dy \qquad \text{for a.e. } x \in \mathbb{R}^N, \tag{2.1}$$

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

$$\mathrm{supp}(f * g) \subset \overline{\mathrm{supp}(f) + \mathrm{supp}(g)}. \quad \text{[Ex]} \tag{2.2}$$

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.

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

$$\|f * g\|_{L^1} \le \|f\|_{L^1} \|g\|_{L^1} \qquad \forall f, g \in L^1. \tag{2.3}$$

(ii) $L^1_{\rm loc} * L^1_{\rm comp} \subset L^1_{\rm loc}$, and [6]

$$\begin{aligned} &\|f * g\|_{L^1(K)} \le \|f\|_{L^1(K - {\rm supp}(g))} \|g\|_{L^1} \\ &\forall K \subset\subset \mathbb{R}^N, \forall f \in L^1_{\rm loc}, \forall g \in L^1_{\rm comp}. \end{aligned} \tag{2.4}$$

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

(iii) For $N = 1$, $L^1_{\rm loc}(\mathbb{R}^+) * L^1_{\rm loc}(\mathbb{R}^+) \subset L^1_{\rm loc}(\mathbb{R}^+)$. [7] *For any $f, g \in L^1_{\rm loc}(\mathbb{R}^+)$,*

$$(f * g)(x) = \begin{cases} \int_0^x f(x-y) g(y)\, dy & \textit{for a.e. } x \ge 0 \\ 0 & \textit{for a.e. } x < 0, \end{cases} \tag{2.5}$$

$$\|f * g\|_{L^1(0,M)} \le \|f\|_{L^1(0,M)} \|g\|_{L^1(0,M)} \qquad \forall M > 0. \tag{2.6}$$

(iv) In each of these cases the convolution is commutative and associative, and the mapping $(u, v) \mapsto u * v$ is continuous (sequentially continuous if it involves the nonmetrizable space $L^1_{\rm comp}$).

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

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

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

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

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 $L^1_{\rm comp} * L^1_{\rm comp} \subset L^1_{\rm comp}$ is based on (2.2), and is left to the Reader.

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

[6] By $L^1_{\rm comp}$ 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..

[7] Any function or distribution defined on $\mathbb{R}^+$ will be automatically extended to the whole $\mathbb{R}$ with value 0. In *Signal Analysis* (also called Signal Processing in enginering), the functions of time that vanish for any $t < 0$ are said *causal.* One can also define the space $\mathcal{D}'(\mathbb{R}^+)$ of *causal distributions,* namely distributions with support in $[0, +\infty[$.

Proposition 2.2 *L^1, $L^1_{\rm comp}$ and $L^1_{\rm loc}(\mathbb{R}^+)$, equipped with the convolution product, are commutative algebras (without unit).* [8] *In particular,*

$$\begin{aligned}&f*g=g*f, \qquad (f*g)*h=f*(g*h) \qquad \text{a.e. in } \mathbb{R}^N\\ &\forall(f,g,h)\in(L^1)^3\cup\big(L^1_{\rm loc}\times L^1_{\rm comp}\times L^1_{\rm comp}\big).\end{aligned} \tag{2.7}$$

(We may then write $f*g*h$ without risk of ambiguity.)

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

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

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

The cases of $(f,g,h)\in\big(L^1_{\rm loc}\times L^1_{\rm comp}\times L^1_{\rm comp}\big)$ and $(f,g,h)\in L^1_{\rm loc}(\mathbb{R}^+)^3$ are analogously checked. The rest of the proof is left to the reader. □

By means of these results is easily seen that $(L^1,*)$ (here "$\cdot$" stands for the pointwise product) is a commutative Banach algebra. The same clearly holds for $(L^\infty,\cdot)$, which also has the unit element $e\equiv 1$.

Convolution of L^p-functions. The following result generalizes Theorem 2.1. [9]

---

[8] ***** 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 $u\in X$.

For any $N\in\mathbb{N}$, the space of square matrices $\mathbb{C}^{N\times N}$ equipped with the usual multiplication is a noncommutative Banach algebra with unit. The same applies for the space of bounded complex-valued functions defined on any nonempty set, equipped with pointwise multiplication.

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

$$uv\in L^r(\Omega), \quad \|uv\|_r\le\|u\|_p\|v\|_q \qquad \forall u\in L^p(\Omega),\forall v\in L^q(\Omega). \tag{2.8}$$

• **Theorem 2.3 (Young)** *Let*

$$p, q, r \in [1, +\infty], \qquad p^{-1} + q^{-1} = 1 + r^{-1}.\ ^{10} \tag{2.9}$$

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

$$\|f * g\|_{L^r} \le \|f\|_{L^p} \|g\|_{L^q} \qquad \forall f \in L^p, \forall g \in L^q. \tag{2.10}$$

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

$$\begin{aligned} &\|f * g\|_{L^r(K)} \le \|f\|_{L^p(K - {\rm supp}(g))} \|g\|_{L^q} \\ &\forall K \subset\subset \mathbb{R}^N, \forall f \in L^p_{\rm loc}, \forall g \in L^q_{\rm comp}. \end{aligned} \tag{2.11}$$

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*

$$\begin{aligned} &\|f * g\|_{L^r(0,M)} \le \|f\|_{L^p(0,M)} \|g\|_{L^q(0,M)} \\ &\forall M > 0, \forall f \in L^p_{\rm loc}(\mathbb{R}^+), \forall g \in L^q_{\rm loc}(\mathbb{R}^+). \end{aligned} \tag{2.12}$$

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 (2.10) 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

$$\|f * g\|_{L^p} = \Big\| \int f(x-y) g(y)\, dy \Big\|_{L^p} \le \|f\|_{L^p} \|g\|_{L^1} \qquad \forall g \in L^1,$$
$$\|f * g\|_{L^\infty} = \operatorname*{ess\,sup}_{x \in \mathbb{R}^N} \Big| \int f(x-y) g(y)\, dy \Big| \le \|f\|_{L^p} \|g\|_{L^{p'}} \qquad \forall g \in L^{p'}$$

$(p^{-1} + (p')^{-1} = 1)$. 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 (2.10) holds, provided that

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

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 $S_g := {\rm supp}(g)$,

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

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

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

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

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

(iii) Part (iii) can 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)** *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* [11]

$$A : L^{p_1}(\Omega) + L^{p_2}(\Omega) \to L^{q_1}(\Omega') + L^{q_2}(\Omega') \tag{2.13}$$

is a linear operator such that

$$A : L^{p_i}(\Omega) \to L^{q_i}(\Omega') \text{ is continuous.} \tag{2.14}$$

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

$$\frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}, \qquad \frac{1}{q} = \frac{\theta}{q_1} + \frac{1-\theta}{q_2}. \tag{2.15}$$

Then A maps $L^p(\Omega)$ to $L^q(\Omega')$, is linear and continuous. Moreover, if M_1 and M_2 are two constants such that

$$\|Af\|_{L^{q_i}(\Omega')} \leq M_i \|f\|_{L^{p_i}(\Omega)} \qquad \forall f \in L^{p_i}(\Omega)\ (i = 1, 2), \tag{2.16}$$

then

$$\|Af\|_{L^q(\Omega')} \leq M_1^\theta M_2^{1-\theta} \|f\|_{L^p(\Omega)} \qquad \forall f \in L^p(\Omega). \qquad [] \tag{2.17}$$

By this result, we may regard $L^{p(\theta)}(\Omega)$ as an *interpolate space* between $L^{p_1}(\Omega)$ and $L^{p_2}(\Omega)$. (2.17) is accordingly called the *interpolation inequality.* [12]

Let us set

$$\check{f}(x) = f(-x) \qquad \forall x \in \mathbb{R}^N, \forall f \in L^1_{\text{loc}}. \tag{2.18}$$

*** Corollary 2.5** *Let*

$$p, q, s \in [1, +\infty], \qquad p^{-1} + q^{-1} + s^{-1} = 2. \tag{2.19}$$

Then

$$\begin{aligned} &\forall (f, g, h) \in L^p \times L^q \times L^s, \\ &(f * g) \cdot h,\ g \cdot (\check{f} * h),\ f \cdot (\check{g} * h) \in L^1, \quad \textit{and} \\ &\int (f * g) \cdot h = \int g \cdot (\check{f} * h) = \int f \cdot (\check{g} * h). \end{aligned} \tag{2.20}$$

The same holds also

$$\begin{aligned} &\forall (f, g, h) \in \big(L^p_{\text{comp}} \times L^q_{\text{loc}} \times L^s_{\text{comp}}\big), \\ &\forall (f, g, h) \in L^p_{\text{loc}}(\mathbb{R}^+) \times L^q_{\text{loc}}(\mathbb{R}^+) \times L^s_{\text{comp}}(\mathbb{R}^+). \end{aligned} \tag{2.21}$$

---

[11] $L^{p_1}(\Omega) + L^{p_2}(\Omega)$ and $L^{q_1}(\widetilde{\Omega}) + L^{q_2}(\widetilde{\Omega})$ are *topological linear spaces,* that is, linear spaces equipped with a topology for which the linear operations are continuous.

[12] This theorem is actually a prototype of the theory of Banach spaces interpolation.

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

Proof. As $r^{-1}+s^{-1}=1$ by (2.10) and (2.19), the Hölder inequality yields the inclusions of (2.20). [Ex] The first equality in (2.20) follows from the computation

$$\begin{aligned}\int (f*g)h\,dx &= \int\Big(\int f(x-y)g(y)\,dy\Big)h(x)\,dx\\ &= \int g(y)\Big(\int \check{f}(y-x)h(x)\,dx\Big)\,dy = \int g(\check{f}*h)\,dy;\end{aligned}$$

As $f*g=g*f$, the second equality in (2.20) also holds. The assertions (2.21) are similarly checked; this is left to the reader. □

Remark 2.6 * Let $f:\mathbb{R}^N\times\mathbb{R}^N\to\mathbb{C}$ and $g:\mathbb{R}^N\to\mathbb{C}$ be two measurable functions such that the integral

$$I(x)=\int f(x,y)g(y)\,dy \tag{2.22}$$

(absolutely) converges for a.e. x. This is sometimes referred to as a *Volterra convolution,* and extends the ordinary convolution (2.1), which is retrieved if the function u is of the form $f(x,y)=\widetilde{f}(x-y)$.

Some of the foregoing results can be extended to this more general set-up. Let us denote by $L^\infty_y(L^1_x)$ the space of functions f such that $f(\cdot,y)\in L^1$ for a.e. y, and the function $y\mapsto\|f(\cdot,y)\|_{L^1}$ belongs to L^∞. For instance, part (i) of Theorem 2.1 is easily extended as follows:

$$\begin{gathered}\int f(x,y)g(y)\,dy\in L^1_x \qquad \forall f\in L^\infty_y(L^1_x),\forall g\in L^1,\\ \Big\|\int f(x,y)g(y)\,dy\Big\|_{L^1_x}\le\|f\|_{L^\infty_y(L^1_x)}\,\|g\|_{L^1}.\end{gathered} \tag{2.23}$$

Several variants of this statement can be derived along the lines above, including an extension of Young's Theorem 2.3; but we shall not delve on this. □

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

Lemma 2.7 *As $h\to 0$,*

$$\tau_h f\to f \qquad \text{in } C^0,\ \forall f\in C^0, \tag{2.24}$$

$$\tau_h f\to f \qquad \text{in } L^p,\ \forall f\in L^p,\ \forall p\in[1,+\infty[. \tag{2.25}$$

Proof. As any $f\in C^0$ is locally uniformly continuous, $\tau_h f\to f$ uniformly in any $K\subset\subset\mathbb{R}^N$; (2.24) thus holds. This yields (2.25), 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^∞ can be replaced by $L^\infty\cap C^0$, and in part (i) also by $L^\infty\cap C^0_0$.

* **Proposition 2.8** *Let $p, q \in [1,+\infty]$ be such that $p^{-1}+q^{-1}=1$. Then:*

$$f*g \in C^0 \qquad \forall (f,g) \in (L^p \times L^q) \cup (L^p_{\rm loc} \times L^q_{\rm comp}), \tag{2.26}$$

$$f*g \in C^0 \qquad \forall (f,g) \in L^p_{\rm loc}(\mathbb{R}^+) \times L^q_{\rm loc}(\mathbb{R}^+) \quad \text{if } N=1, \tag{2.27}$$

$$(f*g)(x) \to 0 \quad \text{as} \quad |x| \to +\infty \quad \forall (f,g) \in L^p \times L^q, \forall p,q \in [1,+\infty[. \tag{2.28}$$

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 2.7,

$$\begin{aligned} \|\tau_h(f*g)-(f*g)\|_{L^\infty} &= \left\| \int [f(x+h-y)-f(x-y)]g(y)]\,dy \right\|_{L^\infty} \\ &\le \|\tau_h f - f\|_{L^p}\|g\|_{L^q} \to 0 \qquad \text{as } h \to 0; \end{aligned} \tag{2.29}$$

the function $f*g$ can then be identified with a uniformly continuous function. (2.26) is thus established. (2.27) can be similarly checked.

Let $\{f_n\} \subset L^p_{\rm comp}$ and $\{g_n\} \subset L^q_{\rm comp}$ 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. □

(2.28) fails if either p or $q=+\infty$. E.g., for any $u \in L^1$, $u*1 = \int u(y)\,dy$ is constant.

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

$$\rho \in C^\infty(\mathbb{R}^N), \quad \rho \ge 0, \quad \rho(x)=0 \text{ if } |x| \ge 1, \quad \int_{\mathbb{R}^N} \rho(x)\,dx = 1. \tag{2.30}$$

A standard construction provides an example. Let us set

$$\begin{aligned} &v(x) := \exp\left[\left(|x|^2-1\right)^{-1}\right] \quad \text{if } |x|<1, \qquad v(x) := 0 \quad \text{if } |x| \ge 1, \\ &\rho(x) := \frac{v(x)}{\int_{\mathbb{R}^N} v(y)\,dy}, \qquad \rho_\varepsilon(x) := \varepsilon^{-N}\rho\Big(\frac{x}{\varepsilon}\Big) \qquad \forall x \in \mathbb{R}^N, \forall \varepsilon > 0. \end{aligned} \tag{2.31}$$

It is easily checked that ρ_ε is a mollifier for any $\varepsilon > 0$.

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 $R_\varepsilon u$ by

$$R_\varepsilon u(x) := (\rho_\varepsilon * \widetilde{u})(x) = \int_\Omega \rho_\varepsilon(x-y)u(y)\,dy \qquad \forall x \in \mathbb{R}^N. \tag{2.32}$$

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

$$R_\varepsilon u(x) = \varepsilon^{-N} \int_{\mathbb{R}^N} \rho\Big(\frac{y}{\varepsilon}\Big)\widetilde{u}(x-y)\,dy = \int_{\mathbb{R}^N} \rho(t)\widetilde{u}(x-\varepsilon t)\,dt. \tag{2.33}$$

The following theorem summarizes some properties of the operator R_ε.

Proposition 2.9 *Let $u \in L^1_{\rm loc}$ and define R_ε as above. Then:*

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

$$D^\alpha R_\varepsilon u(x) = \varepsilon^{-N-|\alpha|} \int_\Omega \Big[(D^\alpha \rho)\Big(\frac{x-y}{\varepsilon}\Big)\Big] u(y)\, dy = \varepsilon^{-|\alpha|} \int_{\mathbb{R}^N} \big[D^\alpha \rho(y)\big]\ \widetilde{u}(x - \varepsilon y)\, dy \quad (2.34)$$

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

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

for all $x \in \mathbb{R}^N$ and all $\alpha \in \mathbb{N}^N$. Next let us differentiate both sides of (2.33), and interchange derivation and integration by Lebesgue's theorem of dominated convergence. This yields (2.34).

(ii) The stated property on the support of $R_\varepsilon f$ stems from (2.2) and (2.33).

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

$$\begin{aligned} |R_\varepsilon v(x) - v(x)| &\overset{(2.33)}{=} \int_{B(0,1)} \rho(y)[v(x - \varepsilon y) - v(x)]\, dy \\ &\le \int_{B(0,1)} \rho(y)|v(x-\varepsilon y) - v(x)|\, dy \le \max_{|z-x|\le\varepsilon} |v(z) - v(x)|. \end{aligned}$$

Therefore $R_\varepsilon v \to v$ uniformly as $\varepsilon \to 0$.

(iv) As $\|\rho_\varepsilon\|_{L^1(\mathbb{R}^N)} = 1$, for any $p \in [1, +\infty[$ the Young Theorem 2.3 yields

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

If $u \in L^\infty(\Omega)$ the same holds for $p = \infty$, [13]

(v) Let $p \ne \infty$. As $C^0_c(\Omega)$ is dense in $L^p(\Omega)$, for any $u \in L^p(\Omega)$ and any $\eta > 0$, there exists $v_\eta \in C^0_c(\Omega)$ with $\|u - v_\eta\|_{L^p(\Omega)} \le \eta$. By using the linearity of R_ε and part (iii), we get

$$\begin{aligned} &\|R_\varepsilon u - u\|_{L^p(\Omega)} \\ &\le \|R_\varepsilon u - R_\varepsilon v_\eta\|_{L^p(\Omega)} + \|R_\varepsilon v_\eta - v_\eta\|_{L^p(\Omega)} + \|v_\eta - u\|_{L^p(\Omega)} \\ &\le \|R_\varepsilon v_\eta - v_\eta\|_{L^p(\Omega)} + 2\eta. \end{aligned} \quad (2.35)$$

For η fixed, by part (iii) $\|R_\varepsilon v_\eta - v_\eta\|_{L^p(\Omega)} \to 0$ as $\varepsilon \to 0$. Therefore $\limsup_{\varepsilon\to 0} \|R_\varepsilon u - u\|_{L^p(\Omega)} \le 2\eta$. Since $\eta > 0$ was arbitrary, the assertion follows. □

[13] Here we apply the following result. If $u \in \bigcap_{q<p<+\infty} L^p(\Omega)$ and $C = \sup_{q<p<+\infty} \|u\|_p < +\infty$, then $u \in L^\infty(\Omega)$ and $\|u\|_\infty \le C$.

*** Convolution of distributions.** The convolution between a distribution and a smooth function $\mathbb{R}^N \to \mathbb{C}$ is defined by extending (2.1):

$$(T * v)(x) := \langle T, v(x - \cdot)\rangle \qquad \forall x \in \mathbb{R}^N, \tag{2.36}$$

whenever T and v belong to two spaces in duality, like $\mathcal{D}' \times \mathcal{D}$ or $\mathcal{E}' \times \mathcal{E}$ or $\mathcal{S}' \times \mathcal{S}$. By this formula we mean that, defining the function $v_x : y \mapsto v(x - y)$ for any fixed $x \in \mathbb{R}^N$, $(T * v)(x) := \langle T, v_x\rangle$. Bilinearity, separate (sequential) continuity and the support property (2.2) can be routinely checked.

Next we deal with the convolution of two distributions, starting by regular distributions. By part (ii) of Theorem 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

$$\int (f * g)(x)\varphi(x)\, dx = \iint f(x - y)g(y)\varphi(x)\, dxdy = \iint f(z)g(y)\varphi(z + y)\, dzdy. \tag{2.37}$$

By Fubini's theorem, each of these double integrals equals each of the corresponding iterated integrals. As L^1_{loc} and L^1_{comp} are respectively dense in $\mathcal{D}'$ and $\mathcal{E}'$, one can then extend the convolution to these spaces by continuity. More explicitly, for any $(T, S) \in (\mathcal{D}' \times \mathcal{E}') \cup (\mathcal{E}' \times \mathcal{D}')$, we define $T * S \in \mathcal{D}'$ by

$$\langle T * S, \varphi\rangle := \langle T_x, \langle S_y, \varphi(x + y)\rangle\rangle \qquad \forall \varphi \in \mathcal{D}. \tag{2.38}$$

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 definition is meaningful, since

$$\begin{aligned} &\langle S_y, \varphi(x + y)\rangle \in \mathcal{D} \qquad \forall \varphi \in \mathcal{D}, \forall S \in \mathcal{E}', \\ &\langle S_y, \varphi(x + y)\rangle \in \mathcal{E} \qquad \forall \varphi \in \mathcal{D}, \forall S \in \mathcal{D}'.\text{[Ex]} \end{aligned} \tag{2.39}$$

Let us recall the definition (2.18), and define $\check{S}$ for any distribution S by transposition:

$$\langle \check{S}, v\rangle = \langle S, \check{v}\rangle \qquad \forall v \in \mathcal{D}, \tag{2.40}$$

extending the analogous definition for functions. By (2.38), we can then extend the statement (2.20):

$$\langle T * S, \varphi\rangle = \langle T, \check{S} * \varphi\rangle \qquad \forall \varphi \in \mathcal{D}, \forall (T, S) \in (\mathcal{D}' \times \mathcal{E}') \cup (\mathcal{E}' \times \mathcal{D}'). \tag{2.41}$$

By approximating with regular distributions, it is not difficult to extend the properties that we saw for the convolution of functions. More specifically, in $(\mathcal{D}' \times \mathcal{E}') \cup (\mathcal{E}' \times \mathcal{D}')$ the convolution is symmetric, linear, separately (sequential) continuous and fulfills the support property (2.2).

On the other hand in general for $(T, S) \in \mathcal{D}' \times \mathcal{S}'$, and a fortiori for $(T, S) \in \mathcal{D}' \times \mathcal{D}'$, one cannot define $T * S$; this may fail even for $T, S \in L^1_{\text{loc}}$. Actually, 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 $\varphi \equiv 0$.

Anyway, analogously to what we saw for functions, in some cases the convolution of two distributions can be defined also if both factors have noncompact support. In particular for $N = 1$ this applies to causal distributions: for any $T, S \in \mathcal{D}'(\mathbb{R}^+)$ one can define $T * S \in \mathcal{D}'(\mathbb{R}^+)$ as in (2.38). As $L^1_{\text{loc}}(\mathbb{R}^+)$ is dense in $\mathcal{D}'(\mathbb{R}^+)$, this can be justified by regularizing the distribution, then using part (iii) of Theorem 2.1, and finally passing to the limit.

Theorem 2.10 *(i)*

$$\mathcal{E}' * \mathcal{E}' \subset \mathcal{E}', \qquad \mathcal{D}'(\mathbb{R}^+)' * \mathcal{D}'(\mathbb{R}^+)' \subset \mathcal{D}'(\mathbb{R}^+)', \qquad \mathcal{S}'(\mathbb{R}^+)' * \mathcal{S}'(\mathbb{R}^+)' \subset \mathcal{S}'(\mathbb{R}^+)'. \tag{2.42}$$

(ii) $(\mathcal{E}', *)$, $(\mathcal{D}'(\mathbb{R}^+), *)$ *and* $(\mathcal{S}'(\mathbb{R}^+), *)$ *are associative and commutative algebras, with unit element* δ_0.

(ii) Moreover,

$$\mathcal{D}' * \mathcal{E}' \subset \mathcal{D}', \qquad \mathcal{S}' * \mathcal{E}' \subset \mathcal{S}', \qquad \mathcal{S} * \mathcal{S}' \subset \mathcal{E} \cap \mathcal{S}', \qquad \mathcal{E} * \mathcal{E}' \subset \mathcal{E}, \tag{2.43}$$

and in each of these cases the convolution product is linear, symmetric and separately continuous.

(iv) Finally,

$$D^\alpha(T * S) = (D^\alpha T) * S = T * D^\alpha S \qquad \forall (T, S) \in (\mathcal{D}' \times \mathcal{E}') \cup (\mathcal{E}' \times \mathcal{D}'), \forall \alpha \in \mathbb{N}^N, \tag{2.44}$$

and the same holds for the pairs of parts (ii) and (iii). []

Outline of the proof. The argument uses transposition and the previous results of function convolution. For instance, the inclusion $\mathcal{D}' * \mathcal{E}' \subset \mathcal{D}'$ is an extension of $L^1_{\rm loc} * L^1_{\rm comp} \subset L^1_{\rm loc}$. In order to prove parts (i) and (ii), one can approximate the distributions by functions of $L^1_{\rm comp}$ use the convolution properties of functions, and then pass to the limit. Parts (iii) and (iv) can also be proved via approximation. □

The convolution provides a procedure for regularizing distributions. For any $\varepsilon > 0$, let us define a mollifier ρ_ε as in (2.31), and set

$$u_\varepsilon(x) := u * \rho_\varepsilon \qquad \forall x \in \mathbb{R}^N, \forall \varepsilon > 0. \tag{2.45}$$

By (2.44), then $u_\varepsilon \in C^\infty$ for any $\varepsilon > 0$, and $u_\varepsilon \to u$ in $\mathcal{D}'$.

Remarks 2.11 (i) In $\mathcal{E}'$, in $\mathcal{D}'(\mathbb{R}^+)$ and in $\mathcal{S}'(\mathbb{R}^+)$ the convolution is associative. In general, assuming that the convolutions $(u * v) * w$ and $u * (v * w)$ are admissible, they coincide if at least two distributions have compact support. Otherwise this equality may fail: for instance,

$$\begin{aligned} &(1 * \delta_0') * H = (1 * \delta_0)' * H = (1' * \delta_0) * H = (0 * \delta_0) * H = 0 * H = 0, \\ &1 * (\delta_0' * H) = 1 * (\delta_0 * H)' = 1 * (\delta_0 * H') = 1 * (\delta_0 * \delta_0) = 1 * \delta_0 = 1. \end{aligned} \tag{2.46}$$

(ii) The shift operator coincides with the convolution with the translated δ_0, that is,

$$\tau_h \phi = \phi * \delta_{-h} \qquad \forall \phi \in C^0, \forall h \in \mathbb{R}^N. \tag{2.47}$$

In particular $\delta_a * \delta_b = \delta_{a+b}$, for any $a, b \in \mathbb{R}^N$.

(iii) By means of (2.44) we can represent any linear partial differential operator with constant coefficients $P(D) = \sum_{|\alpha| \le m} a_\alpha D^\alpha$ as a convolution: as $\delta_0 \in \mathcal{E}'$,

$$P(D)u = P(D)(u * \delta_0) = u * P(D\delta_0) \qquad \forall u \in \mathcal{D}'. \qquad \square \tag{2.48}$$

The next result has important applications in the theory of filters.

*** Theorem 2.12** *If* $\Phi : \mathcal{D} \to C^\infty$ *is linear and continuous and if* $\tau_h \Phi = \Phi \tau_h$ *for any* $h \in \mathbb{R}^N$, *then there exists one and only one* $u \in \mathcal{D}'$ *such that* $\Phi(v) = u * v$ *for any* $v \in \mathcal{D}$. []

3 The Fourier Transform in L^1

Integral transforms. Next we outline some features that characterize a large number of integral transforms. These are linear integral operators $\mathcal{T}$ that typically act on functions $\mathbb{R}^N \to \mathbb{C}$, and have the form

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

for a prescribed kernel $K : \mathbb{R}^{2N} \to \mathbb{C}$, and for any transformable function u. We postpone the specification of regularity properties, and here briefly illustrate some of the main properties of these transforms.

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

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

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, that is, $\widetilde{\mathcal{T}} = \mathcal{T}^{-1}$ (once the domains of these operators are suitably specified).

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

$$\mathcal{T}Lu = \widetilde{L} \cdot \mathcal{T}u \tag{3.3}$$

for any transformable function u. That is, $\mathcal{T}L\mathcal{T}^{-1} = \widetilde{L}\cdot$.

For any prescribed transformable function f, the equation $Lu = f$ is then transformed to the algebraic equation

$$\widetilde{L}(\xi) \cdot \widehat{u}(\xi) = \widehat{f}(\xi) \qquad \forall \xi \ (\widehat{u} := \mathcal{T}u,\ \widehat{f} := \mathcal{T}f).$$

If $\widetilde{L}(\xi) \neq 0$ for any ξ, then $\widehat{u} = \widehat{f}/\widetilde{L}$ pointwise, whence $u = \widetilde{\mathcal{T}}(\widehat{f}/\widetilde{L})$. This procedure is at the basis of what is called *symbolic (or operational) calculus.*

An example: Fourier series. The operator that maps any locally integrable periodic function $u : \mathbb{R} \to \mathbb{C}$ (e.g., of period 2π) to the sequence of its Fourier coefficients

$$c_k := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx \qquad \forall k \in \mathbb{Z} \tag{3.4}$$

is an integral transform, that here we denote by $\mathcal{T}$. Formally, the Fourier series

$$S(x) = \sum_{k\in\mathbb{Z}} c_k e^{ikx} \qquad \forall x \in \mathbb{R} \tag{3.5}$$

is its inverse. This series is an integral w.r.t. the counting measure (the measure that associates to any subset of $\mathbb{Z}$ the number of its elements). Several functional setups can be provided. As we saw, the Hilbert formulation in $L^2(-\pi, \pi)$ is especially convenient.

For any 2π-period function f, let us consider the differential equation

$$\sum_{n=0}^{m} a_n D^n u = f \qquad \forall t \in \mathbb{R}\ (a_0, ..., a_m \in \mathbb{C}) \tag{3.6}$$

Note that, under suitable assumptions of regularity,

$$D^n u = D^n \sum_{k\in\mathbb{Z}} c_k e^{ikx} = \sum_{k\in\mathbb{Z}} (ik)^n c_k e^{ikx} \qquad \forall n \in \mathbb{N}. \tag{3.7}$$

Let us assume that the function f can also be represented as a Fourier series: $f(x) = \sum_{k\in\mathbb{Z}} b_k e^{ikx}$ for any x. Let us also define the *characteristic polynomial* $P(s) = \sum_{n=0}^{m} a_n s^n$ for any $\in \mathbb{R}$. The equation (3.6) thus reads

$$\sum_{k\in\mathbb{Z}} P(ik) c_k e^{ikx} = \sum_{k\in\mathbb{Z}} b_k e^{ikx} \qquad \forall x \in \mathbb{R}. \tag{3.8}$$

If $P(ik) \neq 0$ for any $k \in \mathbb{Z}$, then (3.8) is equivalent to

$$P(ik)c_k = b_k \quad \text{that is,} \quad c_k = b_k/P(ik) \qquad \forall k \in \mathbb{Z}. \tag{3.9}$$

For instance, for the differential equation $u - D^2 u = f$ the characteristic polynomial reads $P(ik) = 1 + k^2$. As $P(ik) = 1 + k^2 \neq 0$ for any $k \in \mathbb{Z}$, (3.9) defines the solution operator $f \mapsto u$.

The Fourier transform in L^1. We shall systematically deal with spaces of functions $\mathbb{R}^N \to \mathbb{C}$, and 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

$$\widehat{u}(\xi) := (2\pi)^{-N/2} \int_{\mathbb{R}^N} e^{-i\xi\cdot x} u(x)\, dx \qquad \forall \xi \in \mathbb{R}^N; \tag{3.10}$$

here $\xi \cdot x := \sum_{i=1}^{N} \xi_i x_i$. [14] This is a Lebesgue integral. [15] The variable x may represent space or (for $N = 1$) time. The dual variable ξ is interpreted as the vector of frequencies (more precisely, the frequency is $\xi/2\pi$, but we shall omit this distinction).

Different scalings may be used to define this transform. Some authors introduce a factor 2π in the exponent under the integral, others amend the factor in front of the integral, and so on. All these transforms are of the form $u \mapsto A \int_{\mathbb{R}^N} e^{-iB\xi\cdot x} u(x)\, dx$, for some real constants $A, B > 0$. Each of these definitions may simplify some formulas, but none is able to simplify all formulas. Our choice is maybe the most frequently used in analysis.

In engineering applications one often uses $u \mapsto \int_{\mathbb{R}^N} e^{-i2\pi\xi\cdot x} u(x)\, dx$. For $N = 1$, in the latter case ξ represents *frequency,* namely the number of cycles in the unit time. In (3.9) ξ represents *angular frequency* (also called *radian frequency),* namely the angle in the unit time. Here we shall use the term *frequency* in all cases.

Proposition 3.1 *The formula (3.10) defines a linear and continuous operator*

$$\begin{aligned} &\mathcal{F} : L^1 \to C^0_b : u \mapsto \widehat{u};\\ &\|\widehat{u}\|_{L^\infty} \leq (2\pi)^{-N/2} \|u\|_{L^1} \quad \forall u \in L^1. \end{aligned} \tag{3.11}$$

---

[14] We denote the multiple integral of a function $f : \mathbb{R}^N \to \mathbb{C}$ by $\int_{\mathbb{R}^N} f(x)dx$ or just $\int f(x)dx$. Alternative notations are e.g. $\iint_{\mathbb{R}^N} f(x)\, dx$ or $\iint_{\mathbb{R}^N} f(x)\, dx_1...dx_N$.

[15] This integral thus converges absolutely, and also coincides with its principle value: $\int_{\mathbb{R}^N} e^{-i\xi\cdot x} u(x)\, dx = \lim_{A\to+\infty} \int_{]-A,A[^N} e^{-i\xi\cdot x} u(x)\, dx$.

(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, for any $u \in L^1$, (denoting by H the Heaviside function)

$$\widehat{u}(0) = (2\pi)^{-N/2} \int_{\mathbb{R}^N} u(x)\, dx = (2\pi)^{-N/2} \lim_{t\to+\infty} (u * H)(t),$$

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

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

Remark 3.2 For any $u \in L^1$, (3.10) also reads

$$\widehat{u}(\xi) = (2\pi)^{-N/2}\int_{\mathbb{R}^N} \cos(\xi\cdot x) u(x)\, dx - i(2\pi)^{-N/2}\int_{\mathbb{R}^N} \sin(\xi\cdot x) u(x)\, dx \tag{3.12}$$

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

$$C_u(\xi) = (2/\pi)^{-N/2}\int_{\mathbb{R}^N} \cos(\xi\cdot x) u(x)\, dx \qquad \forall \xi \in \mathbb{R}^N, \forall u \in L^1, \tag{3.13}$$

$$S_u(\xi) = (2/\pi)^{-N/2}\int_{\mathbb{R}^N} \sin(\xi\cdot x) u(x)\, dx \qquad \forall \xi \in \mathbb{R}^N, \forall u \in L^1, \tag{3.14}$$

we thus have

$$\widehat{u} = C_u - iS_u \qquad \forall u \in L^1. \tag{3.15}$$

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

$$u \text{ is even} \quad \Leftrightarrow \quad S_u = 0 \quad \Leftrightarrow \quad \widehat{u}(\xi) = C_u(\xi) \qquad \forall \xi \in \mathbb{R}^N, \tag{3.16}$$

$$u \text{ is odd} \quad \Leftrightarrow \quad C_u = 0 \quad \Leftrightarrow \quad \widehat{u}(\xi) = -iS_u(\xi) \qquad \forall \xi \in \mathbb{R}^N. \tag{3.17}$$

The functions C_u and S_u are real valued iff $u(x) \in \mathbb{R}$ for any x. Above the Fourier series of periodic functions were similarly decomposed into the sum of a *cosine series* and a *sine series.* □

The following formulas mimic those we derived for the Fourier series, and have the same basis: the properties of the exponential functiona $x \mapsto e^{-i\xi\cdot x}$. [16]

[16] For any $\xi \in \mathbb{R}^N$, let us set $e_\xi(x) := e^{-i\xi\cdot x}$ for all $x \in \mathbb{R}^N$. These functions map the additive group $(\mathbb{R}^N, +)$ to the multiplicative group $(\mathbb{C}, \cdot)$, and are called *characters.* Their properties are the basis of those of Fourier series and Fourier transform.

Proposition 3.3 *For any $u \in L^1$,* [17]

$$v(x) = u(x-y) \quad \Rightarrow \quad \widehat{v}(\xi) = e^{-i\xi\cdot y}\widehat{u}(\xi) \qquad \forall y \in \mathbb{R}^N, \tag{3.18}$$

$$v(x) = e^{ix\cdot\eta}u(x) \quad \Rightarrow \quad \widehat{v}(\xi) = \widehat{u}(\xi-\eta) \qquad \forall \eta \in \mathbb{R}^N, \tag{3.19}$$

$$v(x) = \overline{u(x)} \quad \Rightarrow \quad \widehat{v}(\xi) = \overline{\widehat{u}(-\xi)}, \tag{3.20}$$

$$u \text{ is real} \quad \Rightarrow \quad \widehat{u}(-\xi) = \overline{\widehat{u}(\xi)}, \tag{3.21}$$

$$u \text{ is imaginary} \quad \Rightarrow \quad \widehat{u}(-\xi) = -\overline{\widehat{u}(\xi)}, \tag{3.22}$$

$$u \text{ is even} \quad \Rightarrow \quad \widehat{u} \text{ is even}, \tag{3.23}$$

$$u \text{ is odd} \quad \Rightarrow \quad \widehat{u} \text{ is odd}, \tag{3.24}$$

$$u \text{ is radial} \quad \Rightarrow \quad \widehat{u} \text{ is radial}, \tag{3.25}$$

$$v(x) = u\big(A^{-1}x\big) \;\Rightarrow\; \widehat{v}(\xi) = |\det A|\,\widehat{u}(A^*\xi) \quad \forall A \in \mathbb{R}^{N^2}, \det A \neq 0. \tag{3.26}$$

[Ex]

Remark 3.4 (3.18) and (3.19) establish a duality between the operations of shift and *modulation* (namely, multiplication by an oscillating function):

(i) if the argument of the input function is shifted, then the transformed function is modulated;

(ii) conversely, if the input function is modulated, then the transformed function is shifted.

Defining the translation and modulation operators

$$(T_y u)(x) = u(x-y), \qquad (M_y u)(x) = e^{i\xi\cdot y}u(x) \qquad \forall x, y \in \mathbb{R}^N, \forall u \in L^1, \tag{3.27}$$

(3.18) and (3.19) read

$$\widehat{T_y u} = M_{-y}\widehat{u}, \qquad \widehat{M_\eta u} = T_\eta \widehat{u} \qquad \forall y \in \mathbb{R}^N, \forall u \in L^1. \tag{3.28}$$

This entails

$$\widehat{T_y M_\eta u} = M_{-y}T_\eta\widehat{u}, \qquad \widehat{M_\eta T_y u} = T_\eta M_{-y}\widehat{u} \qquad \forall y, \eta \in \mathbb{R}^N, \forall u \in L^1. \qquad \square \tag{3.29}$$

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

Lemma 3.5 *Let $j \in \{1, ..., N\}$. If $\varphi, D_j\varphi \in L^1$ then $\int_{\mathbb{R}^N} D_j\varphi(x)\,dx = 0$.*

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

$$\theta_n(x) := \rho\big(x/n\big) \qquad \forall x \in \mathbb{R}^N, \forall n \in \mathbb{N}. \tag{3.30}$$

As each function θ_n has compact support, by partial integration

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

as $n \to \infty$. Since $\theta_n(x) \to 1$ pointwise in $\mathbb{R}^N$, by dominated convergence [18] we then conclude that

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

In passing note that shift and modulation operators are isometries in L^p for any $p \in [1, +\infty]$. $\square$

[17] 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 $A^* = A^{-1}$).

[18] We shall often apply this theorem to justify loose calculations via regularization.

• **Theorem 3.6 (Differentiation formulas)** *For any* $m \in \mathbb{N}$,

$$D_x^\alpha u \in L^1 \;\; \forall |\alpha| \le m \quad \Rightarrow \quad (i\xi)^\alpha \widehat{u} = (D_x^\alpha u)\widehat{\ } \in C_b^0 \;\; \forall |\alpha| \le m, \tag{3.31}$$

$$x^\alpha u \in L^1 \;\; \forall |\alpha| \le m \quad \Rightarrow \quad D_\xi^\alpha \widehat{u} = [(-ix)^\alpha u]\widehat{\ } \in C_b^0 \;\; \forall |\alpha| \le m. \tag{3.32}$$

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 3.5. 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. □

*** Remarks 3.7** (i) One can transform differential operators with constant coefficients of any integer order m. For any family $\{c_\alpha : |\alpha| \le m\} \subset \mathbb{C}$, let us define the operator $P(D) := \sum_{|\alpha|\le m} c_\alpha D^\alpha$. By (3.31) and (3.32),

$$\mathcal{F}(P(D)u) = [P(i\xi)\mathcal{F}(u)] \qquad \forall u \in \mathcal{S}. \tag{3.33}$$

(ii) Defining the multiplicative operators

$$(Xu)(x) = ixu(x), \qquad (\Xi u)(\xi) = i\xi u(\xi), \qquad \forall x, \xi \in \mathbb{R}^N, \forall u \in L^1, \tag{3.34}$$

(3.31) and (3.32) respectively read

$$u, D_x^\alpha u \in L^1 \quad \Rightarrow \quad \Xi^\alpha \widehat{u} = (D_x^\alpha u)\widehat{\ } \in C_b^0, \tag{3.35}$$

$$u, x^\alpha u \in L^1 \quad \Rightarrow \quad D_\xi^\alpha \widehat{u} = [(-X)^\alpha u]\widehat{\ } \in C_b^0. \tag{3.36}$$

This shows that the Fourier transform establishes a duality between differentiation and the above multiplicative operators.

(iii) Differentiation w.r.t. either x or ξ and the multiplicative operators Ξ and X are the generators respectively of the shift and modulation semigroups. This duality is then at the basis of the duality between those semigroups, that we pointed out in Remark 3.4, and conversely.

(iv) Because of (3.31), differential operators amplify high frequency components and attenuate low frequency ones. The inverse operation of integration therefore attenuates high frequencies and amplifies low frequencies. □

• **Corollary 3.8 (Smoothness vs. decay principle)** *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* $(1+|\xi|)^m \widehat{u}(\xi) \in C_b^0$.
(ii) If $(1+|x|)^m u \in L^1$, *then* $\widehat{u} \in C^m$ *and* $D^\alpha \widehat{u} \in C_b^0$ *for any* α *with* $|\alpha| \le m$. [Ex]

Remark 3.9 This result entails that
(i) the more u is regular, the faster $|\widehat{u}|$ decays at infinity, and conversely;
(ii) the faster $|u|$ decays at infinity, the more $\widehat{u}$ is regular, and conversely. □

Examples. (i) For $N = 1$, [19]

$$u = \chi_{[-A,A]} \quad \Rightarrow \quad \widehat{u}(\xi) = \sqrt{2/\pi}\,[\sin(A\xi)]/\xi \qquad \forall \xi \in \mathbb{R}^N. \tag{3.37}$$

(ii) We claim that for any $a > 0$

$$\begin{aligned} &u(x) = \exp(-a|x|^2) \qquad \forall x \in \mathbb{R}^N \quad \Rightarrow \\ &\widehat{u}(\xi) = (2a)^{-N/2} \exp(-|\xi|^2/(4a)) \qquad \forall \xi \in \mathbb{R}^N. \end{aligned} \tag{3.38}$$

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

$$i\xi\widehat{u}(\xi) \overset{(3.31)}{=} \widehat{D_x u}(\xi) = \widehat{-xu}(\xi) \overset{(3.32)}{=} -iD_\xi \widehat{u}(\xi),$$

that is, $D_\xi \widehat{u} = -\xi\widehat{u}$ for any $\xi \in \mathbb{R}^N$. The functions u and $\widehat{u}$ thus solve the same first-order differential equation. On the other hand, by the classical Poisson formula $\int_{\mathbb{R}} \exp(-y^2)\,dy = \sqrt{\pi}$,

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

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

$$u(x) = \exp(-x^2/2) \quad \forall x \in \mathbb{R} \quad \Rightarrow \quad \widehat{u}(\xi) = \exp(-\xi^2/2) \quad \forall \xi \in \mathbb{R}. \tag{3.39}$$

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

$$\begin{aligned} \widehat{u}(\xi) &= (2\pi)^{-N/2} \int e^{-i\xi\cdot x} e^{-|x|^2/2}\,dx = (2\pi)^{-N/2} \int \dots \int e^{\sum_{j=1}^N (-i\xi_j x_j + x_j^2/2)}\,dx_1 \dots dx_N \\ &= \prod_{j=1}^N \Big\{ (2\pi)^{-1/2} \int e^{-i\xi_j x_j} e^{-x_j^2/2}\,dx_j \Big\} \overset{(3.39)}{=} \prod_{j=1}^N e^{-\xi_j^2/2} = e^{-|\xi|^2/2} \quad \forall \xi \in \mathbb{R}^N. \end{aligned}$$

This concludes the proof of (6.25) for $a = 1/2$. (6.25) then follows from (3.26).

The next theorem mimics a property that we saw for Fourier series (the same applies to several other results of Fourier theory).

Theorem 3.10 (Riemann-Lebesgue) *For any* $u \in L^1$, $\widehat{u}$ *is uniformly continuous in* $\mathbb{R}^N$, *and* $\widehat{u}(\xi) \to 0$ *as* $|\xi| \to +\infty$.

[19] The *cardinal sinus* function $\operatorname{sinc} v := \sin(\pi v)/\pi v$ for any $v \in \mathbb{R}$ plays an important role in Signal Analysis, as we shall see.

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 3.8, for any n then $\widehat{u}_n(\xi) \to 0$ 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. [20] As $\widehat{u}$ is continuous, it is then uniformly continuous. □

Remarks 3.11 (i) This theorem entails that, for any interval $]a, b[\subset \mathbb{R}$,

$$\int_a^b e^{inx} u(x)\, dx \to 0 \qquad \text{as } n \to \pm\infty, \forall u \in L^1(a,b). \tag{3.40}$$

This is easily checked by extending u to $\mathbb{R}$ with vanishing value outside the interval$]a, b[$, and noticing that this integral is proportional to $\widehat{u}(-n)$. By the Riemann-Lebesgue Theorem then $\widehat{u}(-n) \to 0$ as $n \to \infty$. Thus $e^{ix} \to 0$ weakly star in $L^\infty(a, b)$, and of course the same holds for the real and imaginary parts of e^{ix}: $\cos(nx) \to 0$, $\sin(nx) \to 0$ weakly star in $L^\infty(a, b)$.

(ii) For any $T > 0$, any $u \in L^1(0, T)$ can be developed in Fourier series:

$$\begin{aligned} u(x) &= \sum_{k\in\mathbb{Z}} c_k e^{i2\pi kx/T} \qquad \text{for a.e. } x \in\,]0, T[, \\ c_k &= \frac{1}{T}\int_0^T u(x) e^{-i2\pi kx/T}\, dx \qquad \forall k \in \mathbb{Z}. \end{aligned} \tag{3.41}$$

By the Riemann-Lebesgue Theorem, $c_k \to 0$ as $k \to \pm\infty$. □

The space L^1 can be identified with a normed subspace of the dual space of C_b^0 via the duality pairing $(u, v) \mapsto \int_{\mathbb{R}^N} u\, v\, dx$. As by Proposition 3.1 $\mathcal{F}$ a linear and continuous operator $L^1 \to C_b^0$, one can define the formal adjoint $\mathcal{F}^\tau : L^1 \to L^\infty$ by the formula

$$\int_{\mathbb{R}^N} \mathcal{F}(u)\, v\, dx = \int_{\mathbb{R}^N} u\, \mathcal{F}^\tau(v)\, dx \qquad \forall u, v \in L^1. \tag{3.42}$$

• **Theorem 3.12 (Parseval)** *The formal adjoint of $\mathcal{F}$ coincides with $\mathcal{F}$ itself,* [21] *that is,*

$$\int_{\mathbb{R}^N} \widehat{u}\, v\, dx = \int_{\mathbb{R}^N} u\, \widehat{v}\, dx \qquad \forall u, v \in L^1. \tag{3.43}$$

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

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

• **Theorem 3.13 (First convolution formula)** [22]

$$u{*}v \in L^1, \quad \textit{and} \quad \widehat{u{*}v} = (2\pi)^{N/2} \widehat{u}\, \widehat{v} \qquad \forall u, v \in L^1. \tag{3.44}$$

[20] Here is an alternative argument. By direct evaluation of the integral one can 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.

[21] One cannot say that $\mathcal{F}$ is self-adjoint, since this terminology is used just for Hilbert spaces.

[22] A second convolution formula reads $\widehat{uv} = (2\pi)^{-N/2} \widehat{u} * \widehat{v}$, but it requires rather restrictive assumptions for $\mathcal{F} : L^1 \to C_b^0$. Dealing with the extensions of the Fourier transform, we shall meet a set-up in which this second formula holds without restrictions.

Proof. By the change of integration variable $z = x - y$, for any $u, v \in L^1$,

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

Next we present the inversion formula for the Fourier transform. First, we introduce what is called the *conjugate Fourier transform:* [23]

$$\widetilde{\mathcal{F}}(v)(x) := (2\pi)^{-N/2} \int_{\mathbb{R}^N} e^{i\xi\cdot x} v(\xi) d\xi \qquad \forall v \in L^1, \forall x \in \mathbb{R}^N. \tag{3.45}$$

This operator differs from $\mathcal{F}$ just in the sign of the imaginary unit. Obviously, $\widetilde{\mathcal{F}}v = \overline{\mathcal{F}\bar{v}}$ for any $v \in L^1$. [24] Clearly the properties of $\widetilde{\mathcal{F}}$ mimic those of $\mathcal{F}$, up to a change in the sign of the imaginary unit i. For instance,

$$u, D_x^\alpha u \in L^1 \quad \Rightarrow \quad (i\xi)^\alpha \widetilde{u} = (-D_x^\alpha u)\widetilde{\ } \in C_b^0, \tag{3.46}$$

$$u, x^\alpha u \in L^1 \quad \Rightarrow \quad D_\xi^\alpha \widetilde{u} = [(ix)^\alpha u]\widetilde{\ } \in C_b^0 \tag{3.47}$$

can be compared with (3.31) and (3.32).

Theorem 3.14 (Inversion) *For any $u \in L^1 \cap C_b^0$, if $\mathcal{F}(u) \in L^1$ then*

$$u(x) = \widetilde{\mathcal{F}}[\mathcal{F}(u)](x) = \mathcal{F}[\widetilde{\mathcal{F}}(u)](x) \qquad \forall x \in \mathbb{R}^N. \tag{3.48}$$

Proof. Let us set $v(x) := \exp(-|x|^2/2)$ for any $x \in \mathbb{R}^N$. As $\mathcal{F}(u) \in L^1$, by the Tonelli and Fubini theorems we have

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

Let us now replace $v(\xi)$ by $v_\varepsilon(\xi) := v(\varepsilon\xi)$, for any $\varepsilon > 0$. By (6.25), $\widehat{v_\varepsilon}(z) = \varepsilon^{-N}\widehat{v}(\varepsilon^{-1}z)$; 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 can pass to the limit under integral as $\varepsilon \to 0$, obtaining

$$v(0) \int_{\mathbb{R}^N} \widehat{u}(\xi)e^{i\xi\cdot x}\,d\xi = u(x) \int_{\mathbb{R}^N} \widehat{v}(y)\,dy. \tag{3.49}$$

[23] By this definition, in general the conjugate Fourier transform of a function is not the conjugate of the Fourier transform of that function. Actually, $\widetilde{\mathcal{F}}(v) = \overline{\mathcal{F}(v)}$ only if v is real-valued.

[24] By $\bar{z}$ we denote the complex conjugate of any $z \in \mathbb{C}$.

On the other hand, by (6.25)

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

As $v(0) = 1$, (3.49) then yields $\int_{\mathbb{R}^N} \widehat{u}(\xi)e^{i\xi\cdot x}\,d\xi = u(x)$, that is, $\widetilde{\mathcal{F}}[\mathcal{F}(u)] = u$. The proof of $\mathcal{F}[\widetilde{\mathcal{F}}(u)] = u$ is similar. □

Remarks 3.15 (i) By a more refined argument one might show that (3.48) holds under the only hypotheses that $u, \widehat{u} \in L^1$, see Rudin [Ru1 186]. (Of course, a posteriori one then gets that $u, \widehat{u} \in C^0_b$.)

(ii) By Theorem 3.14, $\mathcal{F}(u) \equiv 0$ 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

$$\overline{\widehat{\widehat{u}}(x)} = \bar{u}(-x) \qquad \forall x \in \mathbb{R}^N. \tag{3.50}$$

(iii) However the formula (3.48) does not yield an authentic inversion theorem, since we have not characterized the image set $\mathcal{F}(L^1)$. □

*** The Fourier-Laplace transform.** Under suitable assumptions, Fourier transformed functions can be extended to holomorphic functions of several complex variables: $\mathbb{C}^N \to \mathbb{C}$, that is, functions that are separately holomorphic with respect to each variable. We denote the space of these functions by $C^\omega(\mathbb{C}^N)$. This extended operator is called *Fourier-Laplace transform,* for reasons that will be clear after introducing the Laplace transform in the next chapter.

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

Theorem 3.16 (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 $\widehat{u} : (\mathbb{R} \times iB(0,\lambda))^N \to \mathbb{C}$:*

$$\widehat{u}(z) := (2\pi)^{-N/2} \int_{\mathbb{R}^N} e^{-iz\cdot x} u(x)\,dx \qquad \forall z \in \mathbb{R}^N; \tag{3.51}$$

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

$$M(z) := \sup_{x\in\mathbb{R}} \big\{|x|e^{|x|[\lambda+\mathrm{Im}(z)]}\big\}\ (< +\infty) \qquad \forall z \in \mathbb{R} \times iB(0,\lambda), \tag{3.52}$$

and applying the mean value theorem to the complex function $h \mapsto (e^{-ih\cdot x} - 1)/h$, we have

$$\begin{aligned}&\Big|\frac{e^{-i(z+h)\cdot x} - e^{-iz\cdot x}}{h}\Big| = \Big|e^{-iz\cdot x}\Big|\ \Big|\frac{e^{-ih\cdot x} - 1}{h}\Big| = ...\\ &\le e^{|\,\mathrm{Im}(z)|\,|x|}|x| \sup_{t\in[0,1]} e^{t\,\mathrm{Im}(h)|x|} \le M(z)e^{\lambda|x|} \qquad \forall z \in \mathbb{R} \times iB(0,\lambda). \quad [\mathrm{Ex}]\end{aligned} \tag{3.53}$$

By dominated convergence, for any $z \in \mathbb{R} \times iB(0,\lambda)$ one then has

$$\begin{aligned}\widehat{u}'(z) &= \lim_{\mathbb{C}\ni h\to 0} (2\pi)^{-N/2} \int \frac{e^{-i(z+h)\cdot x} - e^{-iz\cdot x}}{h} u(x)\,dx\\ &= -(2\pi)^{-N/2} ix \int e^{-iz\cdot x} u(x)\,dx. \qquad \square\end{aligned} \tag{3.54}$$

Remarks 3.17 (i) As we already remarked, the support of any holomorphic function different from the null function coincides with the whole $\mathbb{R}$.

(ii) Theorem 3.6 and other results also hold for the Fourier-Laplace transform $\widehat{u}$ in its domain of holomorphy, A. In particular, by analytic extension one gets

$$u, D_x^\alpha u \in L^1 \quad \Rightarrow \quad (iz)^\alpha \widehat{u}(z) = (D_x^\alpha u)\widehat{\ }(z) \quad \forall z \in A, \tag{3.55}$$

$$u, x^\alpha u \in L^1 \quad \Rightarrow \quad D_z^\alpha \widehat{u}(z) = [(-ix)^\alpha u]\widehat{\ }(z) \quad \forall z \in A. \qquad \square \tag{3.56}$$

Next we assume that the support of u is bounded, and provide an estimate on the growth at infinity of $\widehat{u}$ and of its derivatives.

Proposition 3.18 *If $u \in L^1$ and $\operatorname{supp} u \subset B(0,R)$ for some $R > 0$, then $\mathcal{F}(u)$ can be extended to a (necessarily unique) holomorphic function $\widehat{u} : \mathbb{C}^N \to \mathbb{C}$ such that*

$$|D^\alpha \widehat{u}(z)| \le (2\pi)^{-N/2} R^{|\alpha|} e^{R|\operatorname{Im}(z)|} \|u\|_{L^1} \qquad \forall z \in \mathbb{C}^N, \forall \alpha \in \mathbb{N}^N. \tag{3.57}$$

Proof. By Theorem 3.16, $\widehat{u}$ can be extended to a unique holomorphic function (which we still denote by $\widehat{u}$) defined on the whole $\mathbb{C}^N$. By (3.56),

$$\begin{aligned} |D^\alpha \widehat{u}(z)| &= (2\pi)^{-N/2} \Big| \int (-ix)^\alpha e^{-iz\cdot x} u(x)\, dx \Big| \\ &\le (2\pi)^{-N/2} \int_{B(0,R)} |(-ix)^\alpha| |e^{-iz\cdot x}| |u(x)|\, dx \\ &\le (2\pi)^{-N/2} R^{|\alpha|} e^{R|\operatorname{Im}(z)|} \|u\|_{L^1} \qquad \forall z \in \mathbb{C}^N. \qquad \square \end{aligned} \tag{3.58}$$

Next we state a deeper classical result, which provides a necessary and sufficient condition for existence of the holomorphic extension.

Theorem 3.19 (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

$$\forall m \in \mathbb{N}, \exists C > 0 : \forall z \in \mathbb{C}^N \qquad |\widehat{u}(z)| \le C \frac{e^{R|\operatorname{Im}(z)|}}{(1+|z|)^m}. \tag{3.59}$$

(The constant C may depend on u and m.)

Proof of "(i) $\Rightarrow$ (ii)". By condition (i), for any $\alpha \in \mathbb{N}^N$, $u, D_x^\alpha u \in L^1$. By (3.55), then

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

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

$$|\widehat{u}(z)| \le C e^{R|\operatorname{Im}(z)|} (1+|z|)^{-m} \qquad \forall z \in \mathbb{C}^N. \square$$

Remark 3.20 Let us consider the Gaussian function $u(x) = \exp(-a|x|^2)$ for $x \in \mathbb{R}^N$, and recall that $\widehat{u}(\xi)$ is also a Gaussian, see (6.25). By Theorem 3.16, $\widehat{u}$ can be extended to a holomorphic function $\mathbb{C}^N \to \mathbb{C}$, but by the Paley-Wiener Theorem the extension does not fulfill the estimate (3.59). (Notice that $\widehat{u}(z)$ is obtained from $\widehat{u}(\xi)$ replacing $|\xi|^2$ by $z\cdot z$, not by $|z|^2 = z\cdot\bar{z}$.) $\square$

3.1 Overview and Commentaries

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, under suitable restrictions, we saw the following. Some of these restrictions will be removed in the extensions that we introduce in the next section.

(i) *Differentiation formulas.* $\mathcal{F}$ 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) *Smoothness vs. decay principle.* $\mathcal{F}$ 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 $\widehat{u}$. [25] If u decays exponentially then the Fourier transform can be extended to an entire holomorphic function $\mathbb{C}^N \to \mathbb{C}$. Further asymptotic information is provided by the Riemann-Lebesgue theorem.

(iii) *Parseval's theorem.* The formal adjoint of $\mathcal{F}$ coincides with $\mathcal{F}$.

(iv) *Convolution formula.* $\mathcal{F}$ transforms the convolution of two functions to the product of their transforms, up to a multiplicative constant.

(v) *Inversion.* Under suitable regularity assumptions, the conjugate Fourier transform plays the role of inverse of Fourier transform. However, it does not seem easy to determine the image set $\mathcal{F}(L^1) \subset C_b^0$.

The properties of the conjugate transform are analogous to those of the transform, because of the similarity between the two definitions.

The inversion formula (3.48) also provides an interpretation of the Fourier transform. (3.48) represents u as a weighted average of the *harmonic components* $x \mapsto e^{i\xi\cdot x}$ *(plane waves).* [26] 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). [27] Therefore any function which fulfills (3.48) can 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.

However, the Fourier transform cannot be interpreted as a spectral decomposition in the Hilbert space L^2 in the customary sense, at variance with the Fourier series. Indeed the functions $e_\xi(x) := e^{-i\xi\cdot x}$ do not form an orthogonal family; actually, they are not even elements of L^2. (They are elements of L^2 and of $\mathcal{S}$, but these are not Hilbert space.)

Symmetries of the Fourier transform. (i) The transform of the shift is the modulation of the transform, and dually the transform of the modulation is the shift of the transform; see (3.28).

(ii) The transform of the derivative is the transform of the original function multiplied by the variable, and dually the transform of the function multiplied by the variable is the derivative of the transform (changed of sign); see (3.31), (3.32).

[25] The Reader will notice that this principle establishes a strict relation between the qualitative property of smoothness and the quantitive property property of decay at infinity.

[26] This is especially clear in L^2. However these functions do not form an orthogonal basis, for the simple reason that they are not elements of that space.

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

(iii) The more the function is regular, the faster the transform decays at infinity; and dually the faster the function decays at infinity, the more the transform is regular; see Remark 3.7.

(iv) The transform coincides with its formal adjoint, see Parseval Theorem 3.12.

(v) The transform of a convolution coincides with the pointwise product of the transforms; see (3.44).

(vi) The conjugate transform is the inverse of the transform, under suitable restrictions; see the Inversion Theorem 3.14.

(vii) The transform $\mathcal{F}$ and the conjugate transform $\widetilde{\mathcal{F}}$ fulfills analogous properties, just with $-i$ in place of i.

4 Extensions of the Fourier Transform

Fourier transform of measures. The Fourier transform can be extended to any finite complex Borel measure μ on $\mathbb{R}^N$. This is a σ-additive measure defined on the σ-algebra $\mathcal{B}$ of Borel subsets of $\mathbb{R}^N$, having finite total mass $\sup_{A\in\mathcal{B}} |\mu(A)|$.

The Fourier transform can be defined on $L^1(\mathbb{R}^N, \mu)$, the space of complex functions that are μ-integrable in the sense of Lebesgue. Formally, this simply corresponds to replacing $f(x)\,dx$ by $d\mu(x)$ in (3.10). 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^0_b. They fulfill the same properties of transformation of derivatives and multiplication by a power of x, the Parseval theorem and the convolution formulas. The Riemann-Lebesgue Theorem 3.10) instead fails: e.g.,

$$\widehat{\delta_y}(\xi) = (2\pi)^{-N/2} e^{-i\xi\cdot y} \qquad \forall y \in \mathbb{R}^N, \tag{4.1}$$

which does not vanish as $|\xi| \to +\infty$. In particular, $\widehat{\delta_0}(\xi) = (2\pi)^{-N/2}$.

We do not go on along this line, since this is a particular case of a more general extension of the Fourier transform that we are going to introduce.

Fourier transform in $\mathcal{S}$. For any $u \in \mathcal{D}$, by Theorem 3.19 $\widehat{u}$ is holomorphic. If $\widehat{u} \in \mathcal{D}$ then it has compact support, so that $\widehat{u} \equiv 0$, by the theorem of analytic continuation. By the inversion formula (3.48) then $u \equiv 0$. In conclusion,

$$u, \widehat{u} \in \mathcal{D} \quad \Rightarrow \quad u \equiv 0.$$

The Fourier transform thus does not map $\mathcal{D}$ to itself. This means that the set of frequencies of the harmonic components of any non-identically vanishing $u \in \mathcal{D}$ is unbounded. In other terms, any nontrivial $u \in \mathcal{D}$ 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, and omit to display restrictions.

In order to construct an operator that acts and is invertible in the same space, we shall use a procedure based on transposition. First, we shall restrict $\mathcal{F}$ to the space of rapidly decreasing functions, $\mathcal{S}$; $\mathcal{F}$ operates in this space, and is linear and continuous. By Proposition XXX, we can identify $\mathcal{S}'$ with a dense subset of $\mathcal{D}'$. By transposition, we shall then extend $\mathcal{F}$ to an operator acting in $\mathcal{S}'$. Finally, we shall see that $\widetilde{\mathcal{F}}$ operates in L^2.

In $\mathcal{S}$ the Fourier transform fulfills all the properties that we saw in L^1. Here we also have a second convolution formula, (4.7), which in general is meaningless for $u, v \in L^1$.

Proposition 4.1 *(The restriction of) $\mathcal{F}$ maps $\mathcal{S}$ to $\mathcal{S}$, is continuous, and is invertible:*

$$\mathcal{F}^{-1} = \widetilde{\mathcal{F}} \qquad \text{in } \mathcal{S}. \tag{4.2}$$

Moreover, for any $u, v \in \mathcal{S}$,

$$(i\xi)^\alpha \widehat{u} = \widehat{D_x^\alpha u}, \tag{4.3}$$

$$D_\xi^\alpha \widehat{u} = [(-ix)^\alpha u]\widehat{\ }, \tag{4.4}$$

$$\int_{\mathbb{R}^N} \widehat{u}\, v\, dx = \int_{\mathbb{R}^N} u\, \widehat{v}\, dx, \tag{4.5}$$

$$u{*}v \in \mathcal{S}, \qquad \widehat{u{*}v} = (2\pi)^{N/2} \widehat{u}\, \widehat{v} \quad \text{ in } \mathcal{S}, \tag{4.6}$$

$$uv \in \mathcal{S}, \qquad \widehat{uv} = (2\pi)^{-N/2} \widehat{u}{*}\widehat{v} \quad \text{ in } \mathcal{S}. \tag{4.7}$$

The conjugate Fourier transform $\widetilde{\mathcal{F}}$ fulfills analogous properties, with $-i$ in place of i in (4.3) *and* (4.4).

Proof. The boundedness of $\mathcal{F}$ in $\mathcal{S}$ is easily checked by repeated use of the Leibniz rule, because the space $\mathcal{S}$ is stable by multiplication by any polynomial and by application of any differential operator (with constant coefficients). [Ex] [28] For $u, v \in \mathcal{S}$, the formulas (4.3)—(4.5) are just particular cases of (3.31), (3.32), (3.43), since $\mathcal{S} \subset L^1$. As it is easily checked that $uv, u{*}v, (u{*}v)\widehat{\ }, \widehat{u}{*}\widehat{v} \in \mathcal{S}$, the *first convolution formula* (4.6) follows from (3.44). By writing (3.44) with $\widehat{u}$ and $\widehat{v}$ in place of u and v, and with $\widetilde{\mathcal{F}}$ 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 *second convolution formula* (4.7) follows.

The final claim of the theorem is obvious. □

Fourier transform in $\mathcal{S}'$. Next we generalize the Fourier transform to $\mathcal{S}'$ by extending the Parseval Theorem 3.12. This is reminiscent of how we generalized differentiation to $\mathcal{D}'$ by extending the formula of partial integration. Both procedures rest upon transposition of the operator, namely, here the Fourier transform, and in Section 1 the differentiation.

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:

$$\langle \mathcal{F}^\tau u, v \rangle := \langle u, \mathcal{F} v \rangle \qquad \forall v \in \mathcal{S}, \forall u \in \mathcal{S}'. \tag{4.8}$$

The conjugate transform $\widetilde{\mathcal{F}}^\tau$ is similarly extended to $\mathcal{S}'$ by transposition.

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

(ii) For any $u \in \mathcal{S}'$, denoting $\mathcal{F}^\tau$ by the hat ($\widehat{\ }$),

$$(i\xi)^\alpha \widehat{u} = \widehat{D_x^\alpha u}, \tag{4.9}$$

$$D_\xi^\alpha \widehat{u} = [(-ix)^\alpha u]\widehat{\ }. \tag{4.10}$$

[28] $\mathcal{S}$ is the smallest space that contains $\mathcal{D}$ and has these properties. Indeed L. Schwartz designed $\mathcal{S}$ for use in the Fourier theory.

(iii) For any $u \in \mathcal{S}'$ and any $v \in \mathcal{S}$,

$$\langle \widehat{u}, v\rangle = \langle u, \widehat{v}\rangle \qquad \textit{(extended Parseval formula)}, \tag{4.11}$$

$$u{*}v \in \mathcal{S}', \qquad \widehat{u{*}v} = (2\pi)^{N/2}\widehat{u}\,\widehat{v} \quad \textit{in } \mathcal{S}', \tag{4.12}$$

$$uv \in \mathcal{S}', \qquad \widehat{uv} = (2\pi)^{-N/2}\widehat{u}{*}\widehat{v} \quad \textit{in } \mathcal{S}'. \tag{4.13}$$

(iv) The operator $\mathcal{F}^\tau$ is invertible in $\mathcal{S}'$ and $(\mathcal{F}^\tau)^{-1} = \widetilde{\mathcal{F}}^\tau$.
(v) The operator $\widetilde{\mathcal{F}}^\tau$ fulfills properties analogous to those of $\widetilde{\mathcal{F}}$.

Proof. [29] (i) By (4.5),

$$\langle \mathcal{F}^\tau u, v\rangle = \langle u, \mathcal{F}v\rangle = \int u(x)\,[\mathcal{F}v](x)\,dx = \int [\mathcal{F}u](x)\,v(x)\,dx \qquad \forall u, v \in \mathcal{S}. \tag{4.14}$$

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

(ii) (4.9) and (4.10) are easily derived from the analogous statements for $\mathcal{S}$ via transposition.

For instance, we retrieve (3.31) in $\mathcal{S}'$ as follows, exchanging the role of the dual variables x and ξ. Let us select any $v = v(\xi) \in \mathcal{S}$, and set $\widehat{v}(x) := (2\pi)^{-N/2}\int_{\mathbb{R}^N} e^{-i\xi\cdot x}v(\xi)\,d\xi$. (3.32) is then replaced by $[(i\xi)^\alpha v]\widehat{\ } = (-D_x)^\alpha\widehat{v}$. Via this formula, denoting $\mathcal{F}^\tau(u)$ by $\widehat{u}$, we get

$$\begin{aligned} {}_{\mathcal{S}'}\langle (i\xi)^\alpha \widehat{u}, v\rangle_{\mathcal{S}} &= {}_{\mathcal{S}'}\langle \widehat{u}, (i\xi)^\alpha v\rangle_{\mathcal{S}} = {}_{\mathcal{S}'}\langle u, [(i\xi)^\alpha v]\widehat{\ }\rangle_{\mathcal{S}} \\ &= {}_{\mathcal{S}'}\langle u, (-D)^\alpha \widehat{v}\rangle_{\mathcal{S}} = {}_{\mathcal{S}'}\langle D^\alpha u, \widehat{v}\rangle_{\mathcal{S}} = {}_{\mathcal{S}'}\langle \widehat{D^\alpha_x u}, v\rangle_{\mathcal{S}} \qquad \forall u \in \mathcal{S}', \forall v \in \mathcal{S}. \end{aligned}$$

(3.31) is thus established.

(iii) As $\mathcal{S}$ is dense in $\mathcal{S}'$, (4.11)–(4.13) are also easily checked.

(iv) For any $u \in \mathcal{S}'$, $\widetilde{\mathcal{F}}^\tau[\mathcal{F}^\tau(u)] = u$ since

$$\langle \widetilde{\mathcal{F}}^\tau[\mathcal{F}^\tau(u)], v\rangle = \langle \mathcal{F}^\tau(u), \widetilde{\mathcal{F}}(v)\rangle = \langle u, \mathcal{F}[\widetilde{\mathcal{F}}(v)]\rangle \overset{(4.2)}{=} \langle u, v\rangle \qquad \forall v \in \mathcal{S}. \tag{4.15}$$

Similarly one can check that $\mathcal{F}^\tau[\widetilde{\mathcal{F}}^\tau(u)] = u$. Thus $(\mathcal{F}|_{\mathcal{S}})^{-1} = \widetilde{\mathcal{F}}|_{\mathcal{S}}$.

(v) As $\mathcal{S}$ is sequentially dense in $\mathcal{S}'$, part (v) follows from the analogous statement for $\widetilde{\mathcal{F}}$. □

On the basis of the latter theorem, henceforth we shall identify $\mathcal{F}^\tau$ with $\mathcal{F}$, and $\widetilde{\mathcal{F}}^\tau$ with $\widetilde{\mathcal{F}}$.

For instance, setting $e_\xi(x) = e^{-i\xi\cdot x}$ for any $\xi, x \in \mathbb{R}^N$,

$$\widehat{\delta_y} = (2\pi)^{-N/2}e_y \qquad \widehat{e_y}(\xi) = (2\pi)^{N/2}\delta_y \qquad \forall y \in \mathbb{R}^N. \tag{4.16}$$

So we retrieve (4.1). In particular, $\widehat{\delta_0} = (2\pi)^{-N/2}$ and $\widehat{1} = (2\pi)^{N/2}\widehat{\delta_0}$. The Riemann-Lebesgue theorem thus fails in $\mathcal{S}'$.

Next we provide a representation of $\mathcal{F}$ which is reminiscent of the principal value of Cauchy.

Proposition 4.3 *For any $u \in L^1_{\rm loc} \cap \mathcal{S}'$,*

$$(2\pi)^{-N/2}\int_{]-R,R[^N} e^{-i\xi\cdot x}u(x)\,dx \;\to\; \widehat{u}(\xi) \qquad \textit{in } \mathcal{S}',\ \textit{as } R \to +\infty. \tag{4.17}$$

[29] As $\mathcal{S}$ is a dense subspace of $\mathcal{S}'$, one can also prove this result via regularization.

Proof. Let us denote by χ_R the characteristic function of the N-dimensional interval $]-R,R[^N$. Note that $u\chi_R \in L^1$ and

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

It is easily checked that $u\chi_R \to u$ in $\mathcal{S}'$ s $R \to +\infty$. The continuity of $\mathcal{F}$ in $\mathcal{S}'$ then yields (4.17). □

Fourier-Laplace transform in $\mathcal{E}'$. Next we extend the Fourier-Laplace transform to suitable distributions. As for $\xi \in \mathbb{C}^N \setminus \mathbb{R}^N$ the function $e^{-ix\cdot\xi}$ is unbounded, here we just transform compactly supported distributions, namely elements of $\mathcal{E}'$. For any $T \in \mathcal{E}'$, first we define this transformed function on $\mathbb{R}^N$; afterwards in Theorem 4.5 we extend it to a holomorphic function defined on the whole $\mathbb{C}^N$.

Theorem 4.4 *For any $T \in \mathcal{E}'$,*

$$\widehat{T}(\xi) = {}_{\mathcal{E}'}\langle T, e^{-ix\cdot\xi}\rangle_{\mathcal{E}} \qquad \forall \xi \in \mathbb{R}^N. \tag{4.18}$$

* *Partial proof.* For any $\varepsilon > 0$, let us define the mollifier ρ_ε as in (2.31), set

$$(T * \rho_\varepsilon)(x) := \langle T, \rho_\varepsilon(x - \cdot)\rangle \qquad \forall x \in \mathbb{R}^N,$$

and note that $T * \rho_\varepsilon \in \mathcal{E}$ by (2.43). It is easily checked that $T * \rho_\varepsilon \to T$ in $\mathcal{E}'$, hence also in $\mathcal{S}'$. Therefore

$$(T * \rho_\varepsilon)\widehat{\ } \to \widehat{T} \qquad \text{in } \mathcal{S}'. \tag{4.19}$$

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

$$\begin{aligned}
&\int_{\mathbb{R}^N} e^{-i\xi\cdot x} (T * \rho_\varepsilon)(x)\, dx = \int_{\mathbb{R}^N} e^{-i\xi\cdot x} \langle T_y, \rho_\varepsilon(x-y)\rangle\, dx \\
&= \int_{\mathbb{R}^N} \langle T_y, e^{-i\xi\cdot y} e^{-i\xi\cdot(x-y)} \rho_\varepsilon(x-y)\rangle\, dx \\
&= \langle T_y, e^{-i\xi\cdot y} \int_{\mathbb{R}^N} e^{-i\xi\cdot(x-y)} \rho_\varepsilon(x-y)\, dx\rangle \qquad \forall \xi \in \mathbb{R}^N.
\end{aligned}$$

The last equality can be justified by approximation and use of the Fubini theorem. Thus

$$(T * \rho_\varepsilon)\widehat{\ }(\xi) = \langle T_y, e^{-i\xi\cdot y}\rangle \widehat{\rho}_\varepsilon(\xi) \qquad \forall \xi \in \mathbb{R}^N. \tag{4.20}$$

This is a holomorphic function of ξ. Since, as $\varepsilon \to 0$, $\rho_\varepsilon(\xi) \to \delta_0$ in the space of Borel measures on $\mathbb{R}^N$, $\widehat{\rho}_\varepsilon(\xi) \to 1$ uniformly on any compact subset of $\mathbb{R}^N$. Therefore

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

By (4.19) 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. □

By the Paley-Wiener Theorem 3.19, the transform of compactly supported C^∞-functions has a holomorphic extension which decays algebraically as $|\operatorname{Re}(z)| \to +\infty$, and conversely. By the next result the transform of any $T \in \mathcal{E}'$ has at most algebraic growth, and conversely.

*** Theorem 4.5 (Paley-Wiener-Schwartz)** *For any $T \in \mathcal{S}'$ and any $R \geq 0$, $\operatorname{supp} T \subset \overline{B(0,R)}$ iff $\mathcal{F}(T)$ can be extended to a holomorphic function $\widehat{T} : \mathbb{C}^N \to \mathbb{C}$ such that*

$$\exists m \in \mathbb{N}_0, \exists C > 0 : \forall z \in \mathbb{C}^N \qquad |\widehat{T}(z)| \leq C(1+|z|)^m e^{R|\operatorname{Im}(z)|}. \qquad [] \tag{4.21}$$

Corollary 4.6 *The null function is the only tempered distribution such that its support and that of its Fourier transform are both limited (in $\mathbb{R}$).*

Proof. By the Paley-Wiener-Schwartz Theorem and by analytic continuation, if $\operatorname{supp} T$ is bounded then $\operatorname{supp} \widehat{T}$ is a proper subset of $\mathbb{R}$ only if $\widehat{T}$. The dual statement is proved by the same argument. □

(The constant C may depend on u and m.)

Fourier transform in L^2. The operator $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ is the point of arrival of our extension of the Fourier transform. Next we deal with the Fourier transform of Lebesgue spaces, and start from L^2. We show that (the restriction of) $\mathcal{F}$ maps L^2 to itself, and is an isometric isomorphism in this space.

• **Theorem 4.7 (Plancherel)** *For any $u \in \mathcal{S}'$, $u \in L^2$ iff $\widehat{u} \in L^2$.*
More precisely, the restriction of $\mathcal{F}$ is an isometry in L^2, with inverse $\widetilde{\mathcal{F}}$:

$$\int_{\mathbb{R}^N} \widehat{u}\bar{\widehat{v}}\, dx = \int_{\mathbb{R}^N} u\bar{v}\, dx, \qquad \|\widehat{u}\|_{L^2} = \|u\|_{L^2} \qquad \forall u, v \in L^2. \tag{4.22}$$

Moreover, [30]

$$\mathcal{F}(u\chi_R) = (2\pi)^{-N/2} \int_{]-R,R[^N} e^{-i\xi\cdot x} u(x)\, dx \to \mathcal{F}(u) \qquad \text{in } L^2, \forall u \in L^2. \tag{4.23}$$

Proof. Denoting the inverse Fourier transform by the tilde, it is easily checked that

$$\mathcal{F}\big(\,\overline{\mathcal{F}(v)}\,\big) = \mathcal{F}(\widetilde{\mathcal{F}}(\bar{v})) = \bar{v} \qquad \forall v \in \mathcal{S}. \tag{4.24}$$

By the Parseval theorem, then

$$\int_{\mathbb{R}^N} \widehat{u}\bar{\widehat{v}}\, dx \overset{(3.43)}{=} \int_{\mathbb{R}^N} u\widehat{\bar{\widehat{v}}}\, dx = \int_{\mathbb{R}^N} u\bar{v}\, dx \qquad \forall u, v \in \mathcal{S}.$$

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

Remarks 4.8 (i) By (4.23), as $R \to +\infty$ $\mathcal{F}(u\chi_R) \to \mathcal{F}(u)$ in measure on any bounded subset of $\mathbb{R}^N$, hence also a.e. along a suitable sequence which may depend on u. But in general $\mathcal{F}(u)$ cannot be represented as an integral.

(ii) We derived the Plancherel theorem from the Parseval theorem. Conversely the Parseval theorem easily follows from the Plancherel theorem. □

[30] This is reminiscent of the principal value of Cauchy.

The formulas of Proposition 3.3, the differentiation formulas (3.31) and (3.31), the Parseval formula (3.43) and the convolutions formulas (4.6) and (4.7) hold for $\mathcal{F}$ in L^2 without any restriction. The Riemann-Lebesgue theorem instead fails in L^2; as a counterexample it suffices to consider any $u \in L^2$ which does not vanish at infinity, and not that $u \in \mathcal{F}(L^2)$.

Remark 4.9 * If an electric current of density $u = u(t) \in L^2$ flows through a resistance R, it dissipates the *instantaneous power* $R|u(t)|^2$ and the (total) *energy* $R \int_{\mathbb{R}} |u(t)|^2 \, dt$. Let us assume that $R = 1$. If u $(\neq 0)$ is periodic of period T then its energy is infinite, and one defines the *mean power*

$$P(t) = \frac{1}{T} \int_{-T/2}^{T/2} |u(t)|^2 \, dt. \tag{4.25}$$

In Signal Theory one calls *energy signal (power signal,* resp.) a signal of finite energy (finite mean power, resp.). Note that the mean power of a periodic signal vanishes whenever the energy is finite, so it is of interest to deal with the mean power only if $u \notin L^2$.

The energy of an energy signal is thus distributed in time with *energy density* $|u(t)|^2$. As by the Plancherel theorem $\int_{\mathbb{R}} |u(t)|^2 \, dt = \int_{\mathbb{R}} |\widehat{u}(\xi)|^2 \, d\xi$, one can also regard the energy as distributed in frequency with *spectral density of energy* $|\widehat{u}(\xi)|^2$. Similarly, the power of a power signal is distributed in time with *power density* $|u(t)|^2/T$, and is distributed in frequency with *power spectral density* $|\widehat{u}(\xi)|^2/T$. This applies for instance to electromagnetic and mechanical (in particular acoustic) signals.

By the Convolution Theorem, setting $v_-(t) = v(-t)$ for any t, one can easily see that

$$\mathcal{F}^{-1}(|\widehat{u}|^2) = \sqrt{2\pi} \int u(t')\overline{u(t'-t)} \, dt' =: \sqrt{2\pi} R_{uu}(t) \qquad \forall u \in L^2. \tag{4.26}$$

R_{uu} is called the *time autocorrelation* function of the signal. □

*** Fourier transform of the triple product L^2.** For any $f, g, h \in L^2$, let us define two types of *triple product:*

$$\begin{aligned} [f \cdot (g * h)](x) &= f(x) \int g(y)h(x-y) \, dy \qquad q\forall x \in \mathbb{R}, \\ [(f \cdot g) * h](x) &= \int f(y)g(y)h(x-y) \, dy \qquad q\forall x \in \mathbb{R}. \end{aligned} \tag{4.27}$$

Both are elements of L^2, and are different in general; thus this triple product is not associative.

The following formulas show that $\mathcal{F}$ transforms the first type into the second type, and vice versa:

$$\begin{aligned} \mathcal{F}[f \cdot (g * h)] &= (2\pi)^{-N/2} \widehat{f} * \widehat{g * h} = \widehat{f} * (\widehat{g} \cdot \widehat{h}), \\ \mathcal{F}[(f \cdot g) * h] &= (2\pi)^{N/2} \widehat{f \cdot g} \cdot \widehat{h} = (\widehat{f} * \widehat{g}) \cdot \widehat{h} \qquad \forall f, g, h \in L^2. \end{aligned} \tag{4.28}$$

Fourier transform in L^p for $p \in [1, 2]$.

Lemma 4.10

$$L^q \subset L^p + L^r \qquad \forall p, q, r \in [1, +\infty], \text{ with } p < q < r. \tag{4.29}$$

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

$$u\chi_u \in L^p, \qquad u(1-\chi_u) \in L^r, \qquad u = u\chi_u + u(1-\chi_u) \qquad \forall u \in L^q. \qquad \square \tag{4.30}$$

Theorem 4.11 *(Hausdorff-Young) Let $p \in [1,2]$, and set $p' := p/(p-1)$ if $p > 1$, $p' = \infty$ if $p = 1$. Then (the restriction of) $\mathcal{F}$ is a linear and continuous operator $L^p \to L^{p'}$, and*

$$\|\widehat{u}\|_{L^{p'}} \leq \|u\|_{L^p} \qquad \forall u \in L^p. \tag{4.31}$$

Moreover, setting $\chi_u := 1$ where $|u| \geq 1$ and $\chi_u := 0$ elsewhere,

$$\mathcal{F}(u) = \mathcal{F}(u\chi_u) + \mathcal{F}(u(1-\chi_u)) \in L^\infty + L^2 \qquad \forall u \in L^p. \tag{4.32}$$

In this case $\mathcal{F}(u)$ thus is a regular distribution. Note that $\mathcal{F}(u)$ is nonexpansive, and for $p > 1$ need admit an integral representation. Anyway it can be written as a limit in L^p, similarly to (4.23).

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. Finally, (4.30) yields (4.32). $\square$

Remarks 4.12 (i) This extended Fourier transform fulfills analogous properties to those of the original transform, including the formulas of derivatives and multiplication by a power of x, the Parseval theorem and the first convolution formula, But the Riemann-Lebesgue Theorem, and (for $p < 2$) the second convolution formula fail in general.

(ii) For any $p \in \,]1,2]$, $\mathcal{F} : L^p \to L^{p'}$ is injective, because it is the restriction of $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$. But $\mathcal{F}$ is not surjective. (We know that this fails for $p = 1$.)

(iii) For any $p \in \,]1,2]$, $\mathcal{F} : L^p \to L^{p'}$ can be represented as a principal value.

(iv) For $p \in [1,2]$, the following scaling argument shows that $\mathcal{F}$ maps L^p to L^q with continuity only if $q = p'$. Therefore the Fourier transform cannot be further extended to a continuous operator between Lebsgue spaces. Anyway, for any $p \in \,]2,+\infty]$, $\mathcal{F} : L^p \to \mathcal{S}'$ is continuous.

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 (3.26) we have $\mathcal{F}(u_\lambda) = \lambda^{-N}\mathcal{F}(u)_{1/\lambda}$. Hence, for any $u \in \mathcal{S} \setminus \{0\}$,

$$\frac{\|\mathcal{F}(u_\lambda)\|_{L^q}}{\|u_\lambda\|_{L^p}} = \lambda^{-N}\frac{\|\mathcal{F}(u)_{1/\lambda}\|_{L^q}}{\|u_\lambda\|_{L^p}} = \lambda^{N(-1+1/q+1/p)}\frac{\|\mathcal{F}(u)\|_{L^q}}{\|u\|_{L^p}}, \tag{4.33}$$

and this is uniformly bounded w.r.t. λ iff $q = p'$. $\square$

Proposition 4.13 *For any $p \in [1,2]$ and any $u \in L^p$,*

$$(2\pi)^{-N/2}\int_{]-R,R[^N} e^{-i\xi\cdot x}u(x)\,dx \;\to\; \widehat{u}(\xi) \qquad \textit{in } L^p, \textit{ as } R \to +\infty.\ ^{31} \tag{4.34}$$

Therefore this integral also converges in measure on all bounded subsets of $\mathbb{R}^N$. As $R \to +\infty$ along a suitable sequence which may depend on u, it then also converges a.e..

[31] This is reminiscent of the principal value of Cauchy.

Proof. By (4.32) it suffices to prove (4.34) for $p = 2$. 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 (4.22) then

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

*** Fourier transform of periodic function.**

(This part follows the lines of [Gilardi pp. 474-484])

We assume that $N = 1$, but this discussion could easily be extended to any dimensioon.

First we define the space of T*-periodic distributions* for any $T > 0$:

$$\mathcal{D}'_T(\mathbb{R}) = \big\{u \in \mathcal{D}'(\mathbb{R}) : \langle u, \varphi(\cdot + T)\rangle = \langle u, \varphi\rangle, \forall \varphi \in \mathcal{D}'(\mathbb{R})\big\}. \tag{4.35}$$

The shift of a distribution is defined by transposition, so that $u(\cdot - T) = u$ iff $\langle u, \varphi(\cdot + T)\rangle = \langle u, \varphi\rangle$ for any $\varphi \in \mathcal{D}'(\mathbb{R})$. This definition of periodic distributions thus extends that of periodic functions.

$\mathcal{D}'_T(\mathbb{R})$ is a (sequentially) closed subspace of $\mathcal{D}'$. By the Characterization Theorem 1.1 it is easy to check that any periodic distribution has finite order.

At variance with the nonperiodic setting, here we have the following result.

Proposition 4.14 *Let $T > 0$. Any T-periodic distribution is tempered, i.e., $\mathcal{D}'_T(\mathbb{R}) \subset \mathcal{S}'$. Moreover, for any sequence $\{v_n\}$ in $\mathcal{D}'_T(\mathbb{R})$,*

$$v_n \to v \quad \text{in } \mathcal{D}' \quad \Leftrightarrow \quad v_n \to v \quad \text{in } \mathcal{S}'. \qquad [] \tag{4.36}$$

See e.g. [Gilardi 474] for the proof.

The relation between Fourier transform and Fourier series is illustrated by the next result, by which the frequencies of the harmonic components of a periodic distribution are confined to a discrete set, which consists of multiples of a fundamental frequency. One then says that the distribution has *discrete spectrum.*

*** Theorem 4.15 (Fourier series in $\mathcal{D}'$)** *Let $u \in \mathcal{S}'$ and $T > 0$. Then:*

(i) The following three statements are equivalent:

$$u \in \mathcal{D}'_T, \tag{4.37}$$

$$\exists \{a_k\}_{k\in\mathbb{Z}} \subset \mathbb{C} : u = \sum_{k\in\mathbb{Z}} a_k e^{2k\pi i x/T} \qquad \text{in } \mathcal{D}', \tag{4.38}$$

$$\exists \{a_k\}_{k\in\mathbb{Z}} \subset \mathbb{C} : \widehat{u} = \sqrt{2\pi} \sum_{k\in\mathbb{Z}} a_k \delta_{2k\pi/T} \qquad \text{in } \mathcal{D}'. \tag{4.39}$$

(ii) The two-sided sequence $\{a_k\}$ is uniquely determined by $u \in \mathcal{D}'_T$. If moreover $u \in L^1_{\text{loc}}$, then

$$a_k = \frac{1}{T} \int_0^T e^{-2k\pi i x/T} u(x)\, dx \left(= \frac{\sqrt{2\pi}}{T} \widehat{u\chi_{]0,T[}}(2k\pi/T) \right) \qquad \forall k \in \mathbb{Z}. \tag{4.40}$$

Proof. We just check the final equality: for any $u \in \mathcal{D}'_T \cap L^1_{\text{loc}}$,

$$\frac{1}{T} \int_0^T e^{-2h\pi i x/T} u(x)\, dx \overset{(4.38)}{=} \sum_{k\in\mathbb{Z}} \frac{a_k}{T} \int_0^T e^{-2h\pi i x/T} e^{2k\pi i x/T}\, dx = a_h \qquad \forall h \in \mathbb{Z}. \tag{4.41}$$

(This series can be integrated term by term since $u \in \mathcal{S}'$.) □

The series in (4.38) and the a_ks are respectively called the *Fourier series* and the *Fourier coefficients* of u. Consistently with what we saw in the chapter devoted to Fourier series, the spectrum $\widehat{u}$ is often identified with the two-sided sequence $\{\sqrt{2\pi} a_k\}$.

Proposition 4.16 *Let $T > 0$ and $\{a_k\}$ be a two-sided sequence in $\mathbb{C}$.*
The series $\sum_{k \in \mathbb{Z}} a_k e^{2k\pi ix/T}$ converges in $\mathcal{D}'$ (equivalently, in $\mathcal{S}'$) iff

$$\exists M, m > 0 : \forall k \in \mathbb{Z}, \qquad |a_k| \leq M(1 + |k|^m). \qquad [] \tag{4.42}$$

* **Theorem 4.17 (Fourier series in L^2)** *Let u be any T-periodic distribution, and the two-sided sequence $\{a_k\}$ be as in (4.38) or (4.39). Then $u|_{]0,T[} \in L^2(0,T)$ iff $\{a_k\} \in \ell^2$. In this case*

$$\|u|_{]0,T[}\|^2_{L^2(0,T)} = T \sum_{k \in \mathbb{Z}} |a_k|^2 \ \big(= T\|\{a_k\}\|^2_{\ell^2}\big). \tag{4.43}$$

Proof. For any $u \in \mathcal{D}'_T \cap L^1_{\rm loc}$, by (4.38)

$$\|u|_{]0,T[}\|^2_{L^2(0,T)} = \sum_{h,k \in \mathbb{Z}} a_h \overline{a_k} \int_0^T e^{2(k-h)\pi ix/T}\, dx = T \sum_{k \in \mathbb{Z}} |a_k|^2. \tag{4.44}$$

(This series can be integrated term by term since $u \in \mathcal{S}'$.) □

Remark 4.18 Any periodic distribution has discrete spectrum, but the converse may fail [Gilardi 484]. □

Overview of the extensions of the Fourier transform. The Fourier transform (3.11) has a natural extension for any complex Borel measure μ. Loosely speaking, this is just defined by replacing $u(x)dx$ by $d\mu$ in the definition (3.10).

By the Paley-Wiener theorem, $\mathcal{D}$ is not transformed to itself by 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 continuously, so that $\mathcal{F}$ can be extended to $\mathcal{S}'$ by transposition. $\mathcal{F}$ maps L^2 to itself isometrically (Plancherel theorem). Moreover, by the Riesz-Thorin Theorem, $\mathcal{F}$ is also linear and continuous from L^p to $L^{p/(p-1)}$ for any $p \in]1, 2]$; but this fails for $p > 2$.

Note: The Fourier transform is a nonsurjective homomorphism between the commutative algebras $(L^1, *)$ and $(C^0_b, \cdot)$ (here "$\cdot$" stands for the product a.e.). By the Riemann-Lebesgue theorem, it is easily checked that $\mathcal{F}(L^1)$ is a dense subalgebra of C^0_b. The Fourier transform is also an isomorphism between the commutative algebras $(\mathcal{S}, *)$ and $(\mathcal{S}, \cdot)$; cf. (4.6), (4.7).

5 Fourier Transform and Ordinary Differential Equations

Let m be a positive integer, $a_0, ..., a_m \in \mathbb{C}$ ($a_m \neq 0$), $f : \mathbb{R} \to \mathbb{C}$ be a given function, and consider the ODE [32]

$$P(D)u(t) := \sum_{n=0}^{m} a_n D^n u(t) = f(t) \qquad t \in \mathbb{R}. \tag{5.1}$$

[32] On account of typical applications of this theory, here it is natural to interpret the independent variable x as time.

Solutions in L^2. If we confine ourselves to functions, it is convenient to assume that $f \in L^2$. We then search for a solution $u \in L^2$ having distributional derivatives up to order m also in L^2. For the moment we assume this regularity for u, and shall check it a posteriori. By applying the Fourier transform to both members of (5.1), we see that the ODE is equivalent to the algebraic equation

$$\sum_{n=0}^{m} a_n (i\xi)^n \widehat{u} = \widehat{f} \qquad \forall \xi \in \mathbb{R}. \tag{5.2}$$

Here it is crucial to assume that no root of the polynomial is purely imaginary, i.e.,

$$\big(P(i\xi) =\big) \sum_{n=0}^{m} a_n (i\xi)^n \neq 0 \qquad \forall \xi \in \mathbb{R}, \tag{5.3}$$

so that the ODE can equivalently we rewritten as

$$\widehat{u}(\xi) = \frac{\widehat{f}(\xi)}{P(i\xi)} \qquad \text{for a.e. } \xi \in \mathbb{R}. \tag{5.4}$$

By (5.3), $1/P(i\xi)$ is asymptotic to $1/(a_m|\xi|^m)$ as $|\xi| \to +\infty$. As here $N = 1$, it is easy to see that

$$1/P(i\xi) \in L^p \qquad \forall p \in \,]1, +\infty] \ (p = 1 \text{ included if } m > 1). \tag{5.5}$$

As $\widehat{f} \in L^2$, we infer that $\widehat{f}(\xi)/P(i\xi) \in L^2$. By applying the inverse Fourier transform $\widetilde{\mathcal{F}} = \mathcal{F}^{-1}$, (5.1) is then equivalent to

$$u = \widetilde{\mathcal{F}}(\widehat{u}) = \widetilde{\mathcal{F}}\bigg(\frac{\widehat{f}(\xi)}{P(i\xi)}\bigg) = \widetilde{\mathcal{F}}\bigg(\frac{1}{P(i\xi)} \cdot \widehat{f}(\xi)\bigg) \qquad \text{a.e. in } \mathbb{R}. \tag{5.6}$$

Note that $u \in L^2$, which is the condition that we assumed at the beginning and had to check. Because of the second convolution formula (4.7), that we know to hold also in L^2, (5.6) then also reads

$$u = \frac{1}{\sqrt{2\pi}} \widetilde{\mathcal{F}}\bigg(\frac{1}{P(i\xi)}\bigg) * \widetilde{\mathcal{F}}(\widehat{f}) = \frac{1}{\sqrt{2\pi}} \widetilde{\mathcal{F}}\bigg(\frac{1}{P(i\xi)}\bigg) * f \qquad \text{in } \mathbb{R}. \tag{5.7}$$

As $\widetilde{\mathcal{F}}(1/P(i\xi)) \in L^1$, we infer that $u \in L^2$,

We have thus proved the following statement.

Proposition 5.1 *Let $a_0, ..., a_m \in \mathbb{C}$ ($a_m \neq 0$, $m \geq 1$), and $f \in L^2$. If* (5.3) *holds, then* (5.7) *is the only solution of the ODE* (5.1) *in L^2.*

Remarks 5.2 (i) Here we set the equation on the whole $\mathbb{R}$, without boundary conditions. In this case the L^2-integrability of u replaces the boundary conditions, so that the solution does not include any constant. [33]

(ii) If $P(i\widetilde{\xi}) = 0$ for some $\widetilde{\xi} \in \mathbb{R}^N$, then the above procedure does not apply. In this case the homogeneous equation $P(D)w = 0$ can be studied in $\mathcal{S}'$, instead of L^2.

As $P(i\widetilde{\xi}) = 0$, there exist an integer $n \leq m$ and a polynomial Q such that

$$Q(i\widetilde{\xi}) \neq 0, \qquad P(i\widetilde{\xi}) = (\xi - \widetilde{\xi})^n Q(i\xi) \qquad \forall \xi \in \mathbb{R}^N. \tag{5.8}$$

[33] This may be understood by recalling that $\mathcal{D}$ is dense in L^2...

The homogeneous ODE (5.1) then has the solutions $w_j(t) = t^{j-1}e^{-i\widetilde{\xi}t}$ for $j = 1, ..., n$. For instance, the equation $D^2u + a^2u = 0$ is transformed to $P(i\xi)\widehat{u}(\xi) = (-\xi^2 + a^2)\widehat{u}(\xi) = 0$. As $P(i\widetilde{\xi}) = 0$ iff $\widetilde{\xi} = \pm a$, this homogenous equation has then a two dimensional space of solutions in $\mathcal{S}'$, with basis $u_\pm(x) = e^{\pm iax}$.

In order to find the general solution of the nonhomogeneous ODE $P(D)w = 0$, it is then necessary to find a particular solution of this equation. This can be obtained via the classical method of *variation of parameters,* but we do not address this issue here.

(iii) In the argument above we transformed the source term f, but the final formula (5.7) is expressed just in terms of the original function f. Anyway, in some cases it may be convenient to antitransform $\widehat{f}(\xi)/P(i\xi)$, as in (5.6).

(iv) The above procedure can be extended to PDEs of the form

$$P(D)u := \sum_{|\alpha|\le m} a_\alpha D^\alpha u = f \qquad x \in \mathbb{R}^N \tag{5.9}$$

with constant complex coefficients a_α, provided that $P(i\xi) \neq 0$ for all $\xi \in \mathbb{R}^N$ and $m > N/2$, so that $1/P(i\xi) \in L^2$. (This is easily checked, since this function is asymptotic to $|\xi|^{-m}$ as $|\xi| \to +\infty$).

(v) Here we set the equation in the whole $\mathbb{R}$. Dealing with time evolution, Cauchy problems are especially relevant for applications, and will be studied via Laplace transform in the next chapter. □

Time-invariant systems. The differential equation (5.1) can be addressed from the point of view of *system theory.* Let us interpret the solution u (that here we assume unique) as the *response* of a linear system characterized by the mapping $L : f \mapsto u$. As the coefficients do not depend on x, this operator is invariant for time shifts: that is, for any $h \in \mathbb{R}$, setting $\tau_h v = v(\cdot + h)$, if $u = Lf$ then $\tau_h u = L(\tau_h f)$. One then says that the system is *time-invariant.*

Next let us go beyond functions, assume that $f, u \in \mathcal{S}'$, and set $h = L\delta_0$, which is tantamount to $\sum_{n=0}^m a_n D^n h = \delta_0$. If (5.3) is fulfilled, as δ_0 is compactly supported we can mimic the derivation of (5.7). This yields

$$L : \delta_0 \mapsto \frac{1}{\sqrt{2\pi}}\widetilde{\mathcal{F}}\left(\frac{1}{P(i\xi)}\right) * \delta_0 = \frac{1}{\sqrt{2\pi}}\widetilde{\mathcal{F}}\left(\frac{1}{P(i\xi)}\right) =: h. \tag{5.10}$$

The function $h = L\delta_0$ of (5.7) is called the *impulsive response in time* (or in space, according to the meaning of the independent variable x). The function $\mathcal{F}(h) = [\sqrt{2\pi}P(i\xi)]^{-1}$ is accordingly called the *impulsive response in frequency* (or just the *response function),*

For any $f \in \mathcal{S}'$, we then rewrite the solution $u \in \mathcal{S}'$ of the equation (5.1) in the form

$$u = h * f \quad \text{ or, in frequency, } \quad \widehat{u} = \sqrt{2\pi}\widehat{h}\widehat{f} = \frac{\widehat{f}}{P(i\xi)}. \tag{5.11}$$

$\widehat{h}$ is then also called the *transfer function* of the system (factor $\sqrt{2\pi}$ apart).

Thus it suffices to know h (or equivalently $\widehat{h}$) to evaluate the response to any admissible input. In other terms, the impulsive response characterizes the system.

Remarks 5.3 (i) In the theory of linear differential equations a distribution $E \in \mathcal{D}'$ such that $P(D)E = \delta_0$ is called a *fundamental solution* of that equation. This entails that

$$E * [P(D)u] = u \quad \forall u \in \mathcal{E}', \qquad P(D)[E * f] = f \quad \forall f \in \mathcal{E}'. \tag{5.12}$$

A classical result states that any differential operator with constant coefficients $P(D) \neq 0$ has a fundamental solution.

This is not the unique solution in $\mathcal{D}'$, since this space prescribes no (either explicit or implicit) restriction on the behaviour at infinity.

(ii) The polynomial $P(\eta)$ (with $\eta \in \mathbb{C}^N$) is an element of $\mathcal{S}'$, and is called the *characteristic polynomial* of the operator $P(D)$. For $\xi \in \mathbb{R}^N$ the polynomial $P(i\xi)$ is also called the *symbol* of that operator.

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

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

has exactly m (possibly repeated) complex roots. Let us denote by $\{\xi_j : j = 1, ..., \ell\}$ the distinct complex roots, and by r_j the multiplicity of ξ_j for any j; thus $\ell \leq m$ and $r_1 + ... + r_\ell = m$. Notice these ξ_j's need not be real. Therefore

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

whence

$$P(D) = a_m \prod_{j=1}^{\ell} (D - i\xi_j)^{r_j}.$$ [34]

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 $\sum_{n=0}^{m} a_n D^n u(t) = 0$. Thus

$$\begin{cases} \xi_j \in \mathbb{R} & \Rightarrow \quad u_{j,1}, ..., u_{j,r_j} \in \mathcal{S}' \\ \xi_j \notin \mathbb{R} & \Rightarrow \quad u_{j,1}, ..., u_{j,r_j} \in \mathcal{D}' \setminus \mathcal{S}' \end{cases} \qquad \text{for } j = 1, ..., \ell. \tag{5.13}$$

We shall distinguish two cases:

(i) If $P(i\xi) \neq 0$ for any $\xi \in \mathbb{R}$, then defining the function h as in (5.10) we conclude that

$$\begin{aligned} &h = \widetilde{\mathcal{F}}(1/P(i\xi))/\sqrt{2\pi} \text{ is the unique fundamental solution of (5.1) in } \mathcal{S}'; \\ &\text{hence } u = h * f \text{ of (5.1) is the unique solution in } \mathcal{S}', \forall f \in \mathcal{S}'. \end{aligned} \tag{5.14}$$

(ii) If instead $P(i\widetilde{\xi}) = 0$ for some $\widetilde{\xi} \in \mathbb{R}$, then each of these roots corresponds to a solution $u(x) = e^{i\widetilde{\xi}x} \in \mathcal{S}'$ of the homogeneous equation (5.1). In this case we are not able to reproduce the above procedure, but one can show that there exists a solution in $u \in \mathcal{S}'$, [35] and

$$\begin{aligned} &\text{the fundamental solution is not unique in } \mathcal{S}'; \\ &\text{hence the solution of (5.1) also fails to be unique in } \mathcal{S}', \forall f \in \mathcal{S}'. \end{aligned} \tag{5.15}$$

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

[35] This is a deep theorem of Hörmander.

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

$$u - k^2 u'' = f(t), \qquad u + k^2 u'' = f(t). \tag{5.16}$$

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 (5.3) 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 $\mathcal{S}'$. For the second equation it is more appropriate to address the initial-value problem rather than the problem on the whole $\mathbb{R}$, and to use the Laplace transform, as we shall see in the next chapter. This mathematical aspect also reflects problems that typically arise in engineering applications.

This discussion can be extended to systems of linear ODEs. The extension to linear evolutionary PDEs with constant coefficients is more complex.

6 Rescaled Fourier Transform

Change in notation. Let us reconsider the definition of Fourier transform:

$$\widehat{u}(\xi) := (2\pi)^{-N/2} \int_{\mathbb{R}^N} u(x) e^{-i\xi\cdot x}\, dx \qquad \forall \xi \in \mathbb{R}^N, \forall u \in L^1, \tag{6.1}$$

with inverse transform (under suitable restrictions)

$$u(x) := (2\pi)^{-N/2} \int_{\mathbb{R}^N} \widehat{u}(\xi) e^{i\xi\cdot x}\, d\xi \qquad \forall x \in \mathbb{R}^N. \tag{6.2}$$

Dealing with Time-Frequency Analysis, Signal Processing, and several other applications of the Fourier theory, it is customary to rescale this transform as follows:

$$\widehat{u}(\xi) := \int_{\mathbb{R}^N} e^{-2\pi i\xi\cdot x} u(x)\, dx \qquad \forall \xi \in \mathbb{R}^N, \forall u \in L^1, \tag{6.3}$$

with inverse transform (under the known restrictions)

$$\widetilde{u}(x) := \int_{\mathbb{R}^N} e^{2\pi i\xi\cdot x} u(\xi)\, d\xi \qquad \forall x \in \mathbb{R}^N. \tag{6.4}$$

Henceforth we shall use this definition. For $N = 1$, here x is time, and ξ is number of cycles for unit time.

Rescaled formulas. Some relevant formulas are then rescaled as follows.

$$v(x) = u(x - y) \quad \Longrightarrow \quad \widehat{v}(\xi) = e^{-2\pi i\xi\cdot y}\widehat{u}(\xi) \qquad \forall y \in \mathbb{R}^N, \tag{6.5}$$

$$v(x) = e^{2\pi i x\cdot\eta} u(x) \quad \Longrightarrow \quad \widehat{v}(\xi) = \widehat{u}(\xi - \eta) \qquad \forall \eta \in \mathbb{R}^N, \tag{6.6}$$

$$u, D_x^\alpha u \in L^1 \quad \Rightarrow \quad (2\pi i)^{|\alpha|}\xi^\alpha \widehat{u} = \widehat{D_x^\alpha u} \in L^\infty \cap C^0, \tag{6.7}$$
$$u, x^\alpha u \in L^1 \quad \Rightarrow \quad D_\xi^\alpha \widehat{u} = (-2\pi i)^{|\alpha|}\widehat{x^\alpha u} \in L^\infty \cap C^0. \tag{6.8}$$

$$u*v \in \mathcal{S}', \qquad \widehat{u*v} = \widehat{u}\,\widehat{v}\ (\in \mathcal{S}'), \tag{6.9}$$
$$uv \in \mathcal{S}', \qquad \widehat{uv} = \widehat{u}*\widehat{v}\ (\in \mathcal{S}'). \tag{6.10}$$

The Parseval and Plancherel formula are unchanged.

For instance,

$$\widehat{\delta_y}(\xi) = e^{2\pi i \xi\cdot y} \qquad \widehat{e^{-2\pi i \xi\cdot y}} = \delta_y(\xi) \qquad \forall y \in \mathbb{R}^N. \tag{6.11}$$

In particular, $\widehat{\delta_0} = 1$ and $\widehat{1} = \widehat{\delta_0}$. Moreover,

$$v(x) = e^{-\pi|x|^2} \quad \Longrightarrow \quad \widehat{v}(\xi) = e^{-\pi|\xi|^2}. \tag{6.12}$$

*** Shift and modulation.** We redefine the shift and modulation operators T_y, M_η of (3.34) as follows:

$$(T_y u)(x) = u(x-y), \qquad (M_\eta u)(x) = e^{2\pi i \eta\cdot x}u(x) \qquad \forall x, y, \eta \in \mathbb{R}^N, \forall u : \mathbb{R} \to \mathbb{C}. \tag{6.13}$$

This entails the following commutation relation:

$$T_y M_\eta = e^{-2\pi i \eta\cdot y} M_\eta T_y \qquad \forall y, \eta \in \mathbb{R}^N. \tag{6.14}$$

Thus $T_y M_\eta = M_\eta T_y$ iff $\eta \cdot y \in \mathbb{Z}$.

The Fourier transform exchanges the roles of the operators T_y, M_η: (6.7) and (6.8) yield

$$\widehat{T_y u} = M_{-y}\widehat{u} \qquad \forall y \in \mathbb{R}^N, \forall u \in L^1, \tag{6.15}$$
$$\widehat{M_\eta u} = T_\eta \widehat{u} \qquad \forall \eta \in \mathbb{R}^N, \forall u \in L^1. \tag{6.16}$$

A modulation in time thus corresponds to a shift in frequency. Hence

$$(T_y M_\eta u)\widehat{\ } = M_{-y}T_\eta \widehat{u}\ \big(= e^{-2\pi i \eta\cdot y} T_\eta M_{-y}\widehat{u}\big) \qquad \forall y, \eta \in \mathbb{R}^N. \tag{6.17}$$

That is,

$$\mathcal{F}T_y = M_{-y}\mathcal{F}, \quad \mathcal{F}M_\eta = T_\eta \mathcal{F}, \quad \mathcal{F}T_y M_\eta = M_{-y}T_\eta \mathcal{F}. \tag{6.18}$$

*** Differentiation and multiplication.** Theorem 3.6 is here rescaled as follows: for any multi-index $\alpha \in \mathbb{N}^N$,

$$u, D_x^\alpha u \in L^1 \quad \Rightarrow \quad (2\pi i\xi)^\alpha \widehat{u} = (D_x^\alpha u)\widehat{\ } \in C_b^0, \tag{6.19}$$
$$u, x^\alpha u \in L^1 \quad \Rightarrow \quad D_\xi^\alpha \widehat{u} = [(-2\pi i x)^\alpha u]\widehat{\ } \in C_b^0. \tag{6.20}$$

Here we redefine the multiplication operators (3.34),

$$(Xu)(x) = 2\pi i x u(x), \qquad (\Xi u)(\xi) = 2\pi i \xi u(\xi), \qquad \forall x, \xi \in \mathbb{R}^N, \forall u \in L^1, \tag{6.21}$$

so that (6.19) and (6.20) read

$$u, D_x^\alpha u \in L^1 \quad \Rightarrow \quad \Xi^\alpha \widehat{u} = (D_x^\alpha u)\widehat{\ } \in C_b^0, \tag{6.22}$$
$$u, x^\alpha u \in L^1 \quad \Rightarrow \quad D_\xi^\alpha \widehat{u} = [(-X)^\alpha u]\widehat{\ } \in C_b^0. \tag{6.23}$$

Remark 6.1 * X and $-D$ are proportional to the infinitesimal generators of the modulation and shift semigroups, respectively:

$$D_\xi (M_\xi u)\big|_{\xi=0} = 2\pi i X u, \qquad D_y (T_y u)\big|_{y=0} = -Du \qquad \forall u \in L^1. \tag{6.24}$$

These semigroups are therefore mutually dual. □

The transform of the *Gaussian function* is here rescaled as follows, for any $a > 0$:

$$\begin{aligned} &u_a(x) = a^{-N/2} e^{-\pi |x/a|^2} \qquad \forall x \in \mathbb{R}^N \quad \Rightarrow \\ &\widehat{u_a}(\xi) = a^{N/2} e^{-\pi |a\xi|^2} \qquad \forall \xi \in \mathbb{R}^N. \end{aligned} \tag{6.25}$$

Thus $u_a = \widehat{u_a}$ iff $a = 1$.

* **Proposition 6.2 (Gaussians)** *For any $a > 0$, the set $\{T_x M_\xi u_a : x, \xi \in \mathbb{R}\}$ is dense in L^2. []*

7 The Poisson summation formula

(In this section we define the Fourier transform as in (6.3).)

Periodization. For any $T > 0$, let us set

$$u_T(x) = \sum_{k\in\mathbb{Z}^N} u(x + kT) \left(= \sum_{k\in\mathbb{Z}^N} u(x - kT) \right) \qquad \forall x \in \mathbb{R}^N, \forall u \in L^1. \tag{7.1}$$

This function is T-periodic in each coordinate direction. We shall call it the *T-periodized function* of u.

Remark 7.1 Given a function $v : [0, T[^N \to \mathbb{C}$, let us set $\widetilde{v} = v$ in $[0, T[^N$ and $\widetilde{v} = 0$ in $\mathbb{R} \setminus [0, T[^N$. Then $\widetilde{v}_T$ is the periodic extension of v, which we may identify with v itself. Notice however that the respective domains are $\mathbb{R}^N$ and $[0, T[^N$. □

Lemma 7.2 *For any $u \in L^1$ and any function $g \in L^\infty$ that is T-periodic in each coordinate direction,*

$$\int u(x)g(x)\, dx = \int_{]0,T[^N} u_T(x)g(x)\, dx, \tag{7.2}$$

$$\|u_T\|_{L^1(0,T)^N} \le \|u\|_{L^1} \qquad \textit{with equality if } u(\cdot) \in \mathbb{R} \textit{ and has constant sign.} \tag{7.3}$$

By setting $\mathcal{P}_T u = u_T|_{]0,T[^N}$, a linear and continuous operator $\mathcal{P}_T : L^1 \mapsto L^1(0, T)^N$ is thus defined.

If u is real and has constant sign, then loosely speaking the restriction of the periodization compresses the total mass of u to a period. Otherwise, some compensation occurs and the total mass is reduced.

Proof. By a simple change of integration variable and as $g(\cdot + kT) = g$, we have

$$\begin{aligned} \int u(x)g(x)\,dx &= \sum_{k\in\mathbb{Z}^N} \int_{]0,T[^N} [u(x+kT)g(x+kT)]\,dx \\ &= \text{(as } ug \in L^1) \int_{]0,T[^N} \sum_{k\in\mathbb{Z}^N} [u(x+kT)g(x+kT)]\,dx \\ &= \int_{]0,T[^N} \Big(\sum_{k\in\mathbb{Z}^N} u(x+kT)\Big) g(x)\,dx = \int_{]0,T[^N} u_T(x)g(x)\,dx. \end{aligned} \tag{7.4}$$

By selecting $g = \text{sign}(u_T)$, we have

$$\|u_T\|_{L^1(0,T)^N} = \int_{]0,T[^N} u_T(x)g(x)\,dx = \int u(x)g(x)\,dx \le \|u\|_{L^1},$$

with equality if u has constant sign. □

Theorem 7.3 (Poisson's summation formula) *Let $T > 0$. If $u \in L^1$, then*

$$([\mathcal{P}_T u](x) =) \sum_{k\in\mathbb{Z}^N} u(x+kT) = \frac{1}{T^N} \sum_{k\in\mathbb{Z}^N} \widehat{u}(k/T) e^{2\pi i k\cdot x/T} \qquad \textit{for a.e. } x \in\,]0,T[^N. \tag{7.5}$$

Dually, if $\widehat{u} \in L^1$, then

$$([\mathcal{P}_{1/T}\widehat{u}](\xi) =) \sum_{k\in\mathbb{Z}^N} \widehat{u}(\xi + k/T) = T^N \sum_{k\in\mathbb{Z}^N} u(kT) e^{-2\pi i k\cdot\xi T} \qquad \textit{for a.e. } \xi \in\,]0,1/T[^N. \tag{7.6}$$

Both series converge in $L^1(0,T)^N$ and $L^1(0,1/T)^N$, respectively. Moreover,

$$\sum_{k\in\mathbb{Z}^N} \delta_{kT} = \frac{1}{T^N} \sum_{k\in\mathbb{Z}^N} \widehat{\delta}_{k/T} \qquad \textit{in } \mathcal{S}'. \tag{7.7}$$

In particular, (7.5) and (7.6) hold for any $u \in \mathcal{S}$.

By (7.5), the Fourier coefficients of the periodized function are proportional to the values of the Fourier transform at the integer multiples of $1/T$.

Proof. (i) By Lemma 8.1, $u_T := \mathcal{P}_T u \in L^1(0,T)^N$, so that we can develop this function in Fourier series:

$$\begin{aligned} u_T(x) &= \sum_{k\in\mathbb{Z}^N} c_k e^{2\pi i k\cdot x/T} \qquad \text{for a.e. } x \in\,]0,T[^N, \\ c_k &= \frac{1}{T^N} \int_{]0,T[^N} u_T(x) e^{-2\pi i k\cdot x/T}\,dx \\ &\overset{(7.2)}{=} \frac{1}{T^N} \int_{\mathbb{R}^N} u(x) e^{-2\pi i k\cdot x/T}\,dx = \frac{1}{T^N}\widehat{u}(k/T) \qquad \forall k \in \mathbb{Z}^N. \end{aligned} \tag{7.8}$$

(Note that $\widehat{u}(k/T)$ is meaningful since $\widehat{u} \in C^0_b$.) This yields (7.5).

(7.6) is similarly proved via the dual argument.

For any $u \in \mathcal{S}$, selecting $x = 0$ in (7.5), we have

$$\sum_{k\in\mathbb{Z}^N} u(kT) = \frac{1}{T^N} \sum_{k\in\mathbb{Z}^N} \widehat{u}(k/T). \tag{7.9}$$

Hence, recalling the Parseval theorem,

$$\langle \sum_{k\in\mathbb{Z}^N} \delta_{kT}, u\rangle = \sum_{k\in\mathbb{Z}^N} u(kT) = \frac{1}{T^N} \sum_{k\in\mathbb{Z}^N} \widehat{u}(k/T) = \frac{1}{T^N}\langle \sum_{k\in\mathbb{Z}^N} \delta_{k/T}, \widehat{u}\rangle = \frac{1}{T^N}\langle \sum_{k\in\mathbb{Z}^N} \widehat{\delta_{k/T}}, u\rangle. \tag{7.10}$$

(7.7) is thus established. □

Remark 7.4 If $u \in L^1$ and $\limsup_{|x|\to+\infty} |u(x)||x|^d < +\infty$ for some $d > N$, then $\sum_{k\in\mathbb{Z}^N} |\widehat{u}(k/T)| < +\infty$ Thus the series (7.5) converges everywhere absolutely. A dual statements holds for the dual series. □

Dirac calculus. For the sake of simplicity, here we assume $T = 1$. However, this can easily be extended to any period $T > 0$. We define the *Dirac comb* (or *pulse train):*

$$s = \sum_{k\in\mathbb{Z}^N} \delta_k \ \left(= \sum_{k\in\mathbb{Z}^N} \delta_{-k}\right) \in \mathcal{S}'. \tag{7.11}$$

Thus (7.9) reads

$$s(t) = \widehat{s}(t) = \sum_{k\in\mathbb{Z}^N} \widehat{\delta_k} = \sum_{k\in\mathbb{Z}^N} e^{2\pi ikt} \qquad \text{in } \mathcal{S}'. \tag{7.12}$$

The periodization can be represented as follows in terms of s: for any $u \in \mathcal{S}$,

$$(u*s)(x) = \sum_{k\in\mathbb{Z}^N} \langle u(x-\cdot), \delta_{-k}\rangle = \sum_{k\in\mathbb{Z}^N} u(x+k) = [\mathcal{P}_1 u](x) \qquad \text{for a.e. } x \in \mathbb{R}. \tag{7.13}$$

We also introduce the *sample function* (with step 1)

$$u(x)\cdot s(x) = \sum_{k\in\mathbb{Z}} \langle u, \delta_k\rangle = \sum_{k\in\mathbb{Z}^N} \langle u(k), \delta_k\rangle \qquad \forall x \in \mathbb{R}, \forall u \in C^0. \tag{7.14}$$

This definition can easily be extended to any step $T > 0$.

Thus, denoting the sample operator by $\mathcal{Q}_1$,

$$\mathcal{P}_1 : \mathcal{S} \to \mathcal{E} \cap L^1 : u \mapsto u*s, \qquad \mathcal{Q}_1 : \mathcal{S} \to \mathcal{S}' : u \mapsto us. \tag{7.15}$$

By the next two formulas the Fourier transform establishes a duality relation between periodization and sampling. In the engineering literature these formulas are known as *exchange relations.*

Theorem 7.5 *(Exchange)*

$$\widehat{u*s} = \widehat{u}s \qquad \forall u \in \mathcal{S}, \tag{7.16}$$

$$\widehat{us} = \widehat{u}*s \qquad \forall u \in \mathcal{S}. \tag{7.17}$$

Proof. The convolution formulas and (7.12) yield

$$\widehat{u*s} = \widehat{u}\widehat{s} = \widehat{u}s, \qquad \widehat{us} = \widehat{u}*\widehat{s} = \widehat{u}*s. \qquad \square$$

Remark 7.6 By this result $\mathcal{F}$ transforms the periodized function of u to the sample function of $\widehat{u}$, and vice versa transforms the sample function of u to the periodized function of $\widehat{u}$. In other terms,

$$\mathcal{F} \circ \mathcal{P}_1 = \mathcal{Q}_1 \circ \mathcal{F} : \mathcal{S} \to \mathcal{S}', \qquad \mathcal{F} \circ \mathcal{Q}_1 = \mathcal{P}_1 \circ \mathcal{F} : \mathcal{S} \to \mathcal{E} \cap L^1. \tag{7.18}$$

The Dirac comb, periodization, sampling and the exchange relations provide a *calculus.*

8 The Sampling theorem

(In this section we define the Fourier transform as in (6.3).)

Let us define the space of L^2-functions with spectrum of bounded support, i.e., band-limited signals of finite energy:

$$BL^2(-F_0, F_0) := \big\{u \in L^2(\mathbb{R}) : \widehat{u}(\xi) = 0,\ \forall \xi \notin [-F_0, F_0]\big\} \qquad \forall F_0 > 0. \tag{8.19}$$

This a Hilbert subspace of $L^2(\mathbb{R})$. By the Paley-Wiener Theorem, any function $u \in BL^2(-F_0, F_0)$ is holomorphic, hence it is supported on the whole $\mathbb{R}$ (if $u \not\equiv 0$).

For the sake of simplicity, here we assume $N = 1$, although this theory can be extended to any finite dimension.

The famous Sampling Theorem provides a formula for reconstructing a signal $u \in BL^2(-F_0, F_0)$ from its sample values, whenever the sampling frequency is larger than a certain threshold. This result was due to Whittaker, and is often ascribed to Kotelnikov, Nyquist and Shannon (see e.g. [Higgins 1984]).

As a first step towards that result, we show how $\widehat{u}$ can be reconstructed by a sequence of sample values of u, provided that the sampling frequency is not too small. We shall then complete the proof by antitransforming this formula.

Proposition 8.1 *Let $F_0 > 0$. For any $u \in BL^2(-F_0, F_0)$ and any $T \leq 1/(2F_0)$,*

$$\widehat{u}(\xi) = T \sum_{k\in\mathbb{Z}} u(kT) e^{-2\pi i k\cdot\xi T} \qquad \textit{for a.e. } \xi \in \,]0, 1/T[. \tag{8.20}$$

Proof. Let us set $F = 1/(2T)$ $(\geq F_0)$. As $u \in L^2$ and $\widehat{u}$ is supported in $[-F_0, F_0]$,

$$\mathcal{P}_{2F}\widehat{u} = \widehat{u}|_{]-F,F[} \in L^2(-F, F).$$

This function can thus be developed in Fourier series: by (7.8),

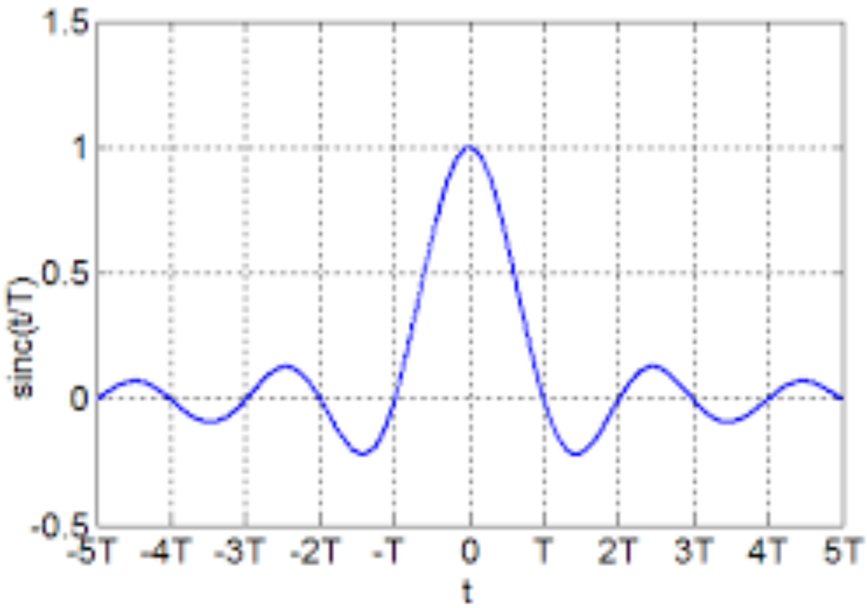
$$\begin{aligned} \widehat{u}(\xi) &= \sum_{k\in\mathbb{Z}} a_k e^{-2\pi i k\xi T} \qquad \text{for a.e. } \xi \in \,]0, 1/T[, \\ a_k &= T \int_{]0,1/T[} \mathcal{P}_{2F}\widehat{u}(\xi) e^{2\pi i k\xi T}\, d\xi \\ &\overset{(7.2)}{=} T \int_{\mathbb{R}} \widehat{u}(\xi) e^{2\pi i k\xi T}\, d\xi = T u(kT) \qquad \forall k \in \mathbb{Z}. \qquad \Box \end{aligned} \tag{8.21}$$

For any $T > 0$, let us define the *rectangle function* of width T, rec_T, and the *cardinal sinus* function, sinc:

$$\mathrm{rec}_T := \chi_{[-T/2,T/2]} \qquad \text{in } \mathbb{R}, \tag{8.22}$$

$$\mathrm{sinc}\ t := \frac{\sin(\pi t)}{\pi t} \qquad \forall t \in \mathbb{R}\setminus\{0\}, \qquad \mathrm{sinc}\ 0 := 1. \tag{8.23}$$

Thus $\mathrm{rec}_T(t) = H(t - T/2) - H(t + T/2)$ (H= Heaviside function).



Graph of the function sinc (cardinal sinus).

The function sinc is continuous, bounded and even. Moreover,

$$\mathrm{sinc}\ 0 = 1, \qquad \mathrm{sinc}\ n = 0 \quad \forall n \in \mathbb{Z} \setminus \{0\}, \tag{8.24}$$

$$[\mathcal{F}(\mathrm{rec}_T)](\xi) = \int_{-T/2}^{T/2} e^{-2\pi i \xi t}\, dt = T\,\mathrm{sinc}\,(T\xi) \qquad \forall \xi \in \mathbb{R}. \tag{8.25}$$

Note that $\mathrm{sinc} \notin L^1$, since it is the Fourier transform of a discontinuous function. As $T \to +\infty$, $\mathrm{rec}_T \to 1$ in $\mathcal{S}'$. Hence

$$\mathcal{F}(\mathrm{rec}_T) = T\,\mathrm{sinc}\,(T\cdot) \to \mathcal{F}(1) = \delta_0 \qquad \text{in } \mathcal{S}', \text{ as } T \to +\infty.$$

Theorem 8.2 *(Sampling theorem) Let $F_0 > 0$. For any $u \in BL^2(-F_0, F_0)$,*

$$u(t) = \sum_{k\in\mathbb{Z}} u(kT)\,\mathrm{sinc}\,(t/T - k) \qquad \forall t \in \mathbb{R}, \forall T \le 1/(2F_0). \tag{8.26}$$

This series converges in $L^2(\mathbb{R})$ and pointwise everywhere.

The series of (8.26) is called the *cardinal series,* and is the discrete-time Fourier transform (DTFT) of the sample sequence $\{u(kT)\}_{k\in\mathbb{Z}}$.

In passing note that (8.26) clearly holds if $t/T \in \mathbb{Z}$, because of (8.24). Moreover, for instance for $T = 1$, (8.26)) can be rewritten as $u = (u\ s) * \mathrm{sinc}$, by (7.14) and (8.24).

* **Proof.** Let us fix any $F \ge F_0$. By Proposition 8.1,

$$\widehat{u}(\xi) = \sum_{k\in\mathbb{Z}} a_k e^{-2\pi i k\xi/(2F)} \in L^2(-F, F), \tag{8.27}$$

$$a_k = \frac{1}{2F} u(k/(2F)) \qquad \forall k \in \mathbb{Z}. \tag{8.28}$$

As $\widehat{u} = \widehat{u}\text{rec}_{2F}$, denoting the conjugate Fourier transform by $\widetilde{\mathcal{F}}$, we then have

$$\begin{aligned} u(t) = [\widetilde{\mathcal{F}}(\widehat{u}\text{rec}_{2F})](t) &\overset{(8.27)}{=} \sum_{k\in\mathbb{Z}} a_k[\widetilde{\mathcal{F}}(e^{-2\pi ik\xi/(2F)}\text{rec}_{2F}(\xi))](t) \\ = \sum_{k\in\mathbb{Z}} a_k[\widetilde{\mathcal{F}}(\text{rec}_{2F}(\xi))](t-k/(2F)) &\overset{(8.25)}{=} \sum_{k\in\mathbb{Z}} a_k 2F\,\text{sinc}\,(-2Ft+k) \\ \overset{(8.28)}{=} \sum_{k\in\mathbb{Z}} u(k/(2F))\,\text{sinc}\,(-2Ft+k) &\qquad \forall t\in\mathbb{R}. \end{aligned} \tag{8.29}$$

As $\text{sinc}\,(-2Ft+k) = \text{sinc}\,(2Ft-k)$, setting $T = 1/(2F)$ we get (8.26).

As $u \in L^2$ and is holomorphic, the series (8.26) converges in $L^2(\mathbb{R})$ and pointwise everywhere. □

Corollary 8.3 *For any $u \in BL^2(-F_0, F_0)$ and any $T \le 1/(2F_0)$,*

$$\int_{\mathbb{R}} u(t)\,dt = T\sum_{k\in\mathbb{Z}} u(kT), \tag{8.30}$$

$$\int_{\mathbb{R}} |u(t)|^2\,dt = \int_{\mathbb{R}} |\widehat{u}(\xi)|^2\,d\xi = T\sum_{k\in\mathbb{Z}} |u(kT)|^2. \tag{8.31}$$

Proof. By (8.25),

$$\int_{\mathbb{R}} \text{sinc}\,(t/T-k)\,dt = T\int_{\mathbb{R}} \text{sinc}\; r\,dr = T[\widetilde{\mathcal{F}}(\text{sinc}\,)](0) = T\text{rec}_1(0) = T. \tag{8.32}$$

(8.30) then follows by integrating (8.26).

As we saw in the proof of the Sampling Theorem, $\widehat{u}$ can be developed in Fourier series, and its coefficients are as in (8.28). By the Plancherel Theorem we then get

$$\int_{\mathbb{R}} |u(t)|^2\,dt = \int_{\mathbb{R}} |\widehat{u}(\xi)|^2\,d\xi = \frac{1}{T}\sum_{k\in\mathbb{Z}} |a_k|^2 = T\sum_{k\in\mathbb{Z}} |u(kT)|^2. \qquad \square$$

Corollary 8.4 *For any $T > 0$, $\{\text{sinc}\,(t/T-k)\}_{k\in\mathbb{Z}}$ is an orthogonal family of functions of L^2, with*

$$\int_{\mathbb{R}} |\,\text{sinc}\,(t/T-k)|^2\,dt = T \qquad \forall k\in\mathbb{Z}. \tag{8.33}$$

This family generates the space $BL^2(-F_0, F_0)$ for any $F_0 \ge 1/(2T)$.

Proof. For any $t \in \mathbb{R}$ and any $k \in \mathbb{Z}$, let us set $v_k(t) = \text{sinc}\,(t/T-k)$, and note that $\widehat{v_k} = (1/T)\,\text{rec}_T(\xi)e^{-2\pi ik\xi/T}$. Therefore

$$\begin{aligned} \int_{\mathbb{R}} v_h(t)\overline{v_k(t)}\,dt &= \int_{\mathbb{R}} \widehat{v_h}(\xi)\overline{\widehat{v_k}(\xi)}\,d\xi = \int_{-T/2}^{T/2} e^{2\pi ih\xi/T}e^{-2\pi ik\xi/T}\,d\xi \\ &= \begin{cases} T & \text{if } h = k \\ 0 & \text{if } h \ne k. \end{cases} \end{aligned} \tag{8.34}$$

The final statement of the thesis stems from the Sampling Theorem. □

Remarks 8.5 (i) For a sampling frequency $1/T < 1/(2F_0)$ (so-called *undersampling),* the formula (8.26) does not reconstruct accurately the signal u. In Signal Processing this distortion is known as *aliasing.* A typical example occurs in movies. If the frequency of rotation of the wheels of a car is larger than the sampling rate of the camera, in the movie the rotating wheels may appear as steady or even turning backward.

On the other hand, a sampling frequency $1/T > 1/(2F_0)$ (so-called *oversampling)* does not improve the reconstruction (8.26), which is already perfect.

(ii) The results of this section can be extended to any number of dimensions. □

9 The Uncertainty Principle

(In this section we define the Fourier transform as in (6.3).)

Fourier analysis describes quantities as functions either of $x \in \mathbb{R}^N$ or of the dual variable $\xi \in \mathbb{R}^N$, establishing a net distinction between space- and spectral-representations. For $N = 1$, in Signal Analysis these variables typically represent time and frequency. Whenever the inversion theorem applies, $u(x)$ and $\widehat{u}(\xi)$ are equivalent representations of the same entity. Nevertheless, without operating the transform, it is may be hard to derive qualitative properties of $u(x)$ from those of $\widehat{u}(\xi)$, and conversely.

Time-Frequency Analysis instead combines information on $u(x)$ with that on $\widehat{u}(\xi)$. The *Heisenberg uncertainty principle* is a fundamental result of this theory. Its several variants put severe limitations: it is impossible to provide a precise representation of any phenomenon both in time and frequency. Instantaneous frequency and time concentration cannot be derived from a real signal.

Doppler theorem. We start by an example. (3.26) yields the following scaling formula, which is also called the *Doppler theorem:*

$$v_a(x) = a^{-N/2}u(x/a) \quad \Rightarrow \quad \widehat{v}_a(\xi) = a^{N/2}\widehat{u}(a\xi) \qquad \forall a > 0, \forall u \in L^1. \tag{9.1}$$

By this scaling and because $\mathcal{F}$ is an isometry in L^2, we get

$$\|v_a\|_{L^2} = \|u\|_{L^2} = \|\widehat{u}\|_{L^2} = \|\widehat{v}_a\|_{L^2} \qquad \forall a > 0.$$

In particular, by taking the Gaussians functions (6.12) as u by $(2\pi)^{-N/2}$, we have

$$v_a(x) = a^{-N/2}e^{-|x/a|^2} \quad \Rightarrow \quad \widehat{v}_a(\xi) = a^{N/2}e^{-|a\xi|^2} \qquad \forall a > 0. \tag{9.2}$$

As $\int e^{-|y|^2}\,dy = \pi^{N/2}$, we infer that $\|v_a\|_{L^2} = \|\widehat{v}_a\|_{L^2} = \pi^{N/2}$ for any $a > 0$.

In either case, by varying a the more the graph of v_a is spread the more that of $\widehat{v}_a$ is concentrated, and conversely. (This stems from the occurrence of the product $\xi \cdot x$ in the definition of the Fourier transform.) Next we state a more general result. Although here we take $N = 1$, this is easily extended to any N.

Theorem 9.1 *(Heisenberg's uncertainty principle) Let $u \in L^2(\mathbb{R})$, and set $E := \|u\|^2_{L^2} (= \|\widehat{u}\|^2_{L^2})$. Let us assume that the densities of probability $|u|^2/E$ and $|\widehat{u}|^2/E$ respectively have finite mean m_1, m_2 and finite variance σ_1^2, σ_2^2:*

$$m_1 = \frac{1}{E}\int x|u|^2(x)\,dx, \qquad \sigma_1^2 = \frac{1}{E}\int (x - m_1)^2|u|^2(x)\,dx, \tag{9.3}$$

$$m_2 = \frac{1}{E}\int \xi|\widehat{u}|^2(\xi)\,d\xi, \qquad \sigma_2^2 = \frac{1}{E}\int (\xi - m_2)^2|\widehat{u}|^2(\xi)\,d\xi. \tag{9.4}$$

Then

$$\sigma_1\sigma_2 \geq (4\pi)^{-1}. \tag{9.5}$$

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

Proof. Without loss of generality, we can assume that $m_1 = m_2 = 0$; thus

$$\sqrt{E}\sigma_1 = \|xu\|_{L^2}, \qquad \sqrt{E}\sigma_2 = \|\xi\widehat{u}\|_{L^2}. \tag{9.6}$$

Let us first assume that $u \in \mathcal{D}$. By the differentiation formulas and by L^2-isometry,

$$4\pi^2 \int \xi^2|\widehat{u}(\xi)|^2\,d\xi = \int |\widehat{u'}(\xi)|^2\,d\xi = \int |u'(t)|^2\,dt < +\infty. \tag{9.7}$$

Notice that

$$\begin{aligned} 2\,\mathrm{Re}\int xu(x)\overline{u'(x)}\,dx &= \int x\Big\{\overline{u(x)}u'(x) + u(x)\overline{u'(x)}\Big\}\,dx = \int x\frac{d}{dx}|u(x)|^2\,dx \\ &= \lim_{A\to+\infty} x|u(x)|^2\Big|_{x=-A}^{x=A} - \|u\|_{L^2}^2 = -\|u\|_{L^2}^2 = -E, \end{aligned}$$

and

$$\left|\mathrm{Re}\int xu(x)\overline{u'(x)}\,dx\right| \leq \|xu\|_{L^2}\|u'\|_{L^2} \overset{(9.7)}{=} 2\pi\|xu(t)\|_{L^2}\|\xi\widehat{u}(\xi)\|_{L^2} \overset{(9.6)}{=} 2\pi E\sigma_1\sigma_2.$$

These two formulas entail that $4\pi\sigma_1\sigma_2 \geq 1$ for all $u \in \mathcal{D}$. As $\mathcal{D}$ is dense in L^2, (9.5) follows. The final statement about Gaussian functions can be checked by direct computation. □

Remark 9.2 The standard deviations σ_1 and σ_2 can be regarded as crude measures respectively of the duration and of the bandwidth of the signal $u = u(x)$. This somehow generalizes the concept of support of a signal in time- and frequency, respectively. [36] □

Heisenberg boxes. The Uncertainty Principle may loosely be interpreted as follows in the time-frequency (t, ξ)-plane. It essentially prevents u and $\widehat{u}$ from being both small outside small sets of the respective time and frequency domains. In other terms, the mass of both u and $\widehat{u}$ cannot be concentrated in sets of small N-dimensional volume.

Let us represent the time-frequency resolution of a signal $u \in L^2$ by a rectangular box centered at the point $(x, \xi) = (m_1, m_2)$, having width σ_1 (σ_2, resp.) along the x-axis (the ξ-axis, resp.). This is called a *Heisenberg box.* By the Heisenberg inequality $\sigma_1\sigma_2 \geq (4\pi)^{-1}$. Thus the sharper is the information we have on time localization of a phenomenon (i.e., the smaller is σ_1), the rougher is the information we may get on its frequency (i.e., the larger is σ_2), and conversely. In this sense, time resolution and frequency resolution mutually compete.

The family of Gaussian functions $u_a(x) = a^{-N/2}e^{-a|x|^2}$, parameterized by $a > 0$, is an L^2-example. By inspecting (9.2), we already pointed out that the spreading of the graph of u_a is inversely proportional to that of $\widehat{u}_a$.

Passing to the limit in $\mathcal{S}'$ as $a \to 0$, $a^{-N/2}e^{-a|x|^2} \to \sqrt{\pi}\delta_0$ in $\mathcal{S}'$, which exactly specifies the time location of the signal. Its Fourier transform is the constant $\sqrt{\pi}$, which is supported over the whole $\mathbb{R}$,

[36] Consistently with the applicative literature, here by *support* of a function f we mean to the region in which most of the mass of the function is concentrated. This rather loose concept is different from that used in analysis, and can be formulated in several ways; for instance, the support if f may be defined as the region where $|f|$ is larger than some prescribed (small) constant. Of course this is not related to the definition of support that we used so far.

and thus is completely unlocalized in frequency. Similarly, for any $x_0 \in \mathbb{R}$, $a^{-N/2}e^{-a|x-x_0|^2} \to \sqrt{\pi}\delta_{x_0}$ in $\mathcal{S}'$ as $a \to 0$. Its Fourier transform is the sinusoid $\widehat{u}(\xi) = \sqrt{\pi}e^{-2\pi i\xi x_0}$, which is completely unlocalized in frequency.

Dually, for any x_0, the Dirac measure $\widehat{u} = \delta_{\xi_0}$ is an exact frequency specification. Its Fourier antitransform is the sinusoid $u(t) = e^{2\pi i\xi_0 t}$, which is completely unlocalized in time. [37]

Note that this discussion of Gaussians applies to any $a > 0$, but not to this limit case $a = 0$, since the Dirac measures and the sinusoids are not elements of L^2.

[37] This theory also has musical applications; for instance, the sinusoid may represent a pure musical tone.