This chapter includes the following sections:
1. Distributions.
2. Convolution.
3. Fourier transform of functions.
4. Extensions of the Fourier transform.
5. Laplace transform.
6. P.D.E.s with constant coefficients.
7. A glance at system theory.

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

1 Distributions

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

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

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

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

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

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

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

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

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

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

A set $A \subset \mathcal{D}(\Omega)$ is open in this topology iff $A \cap \mathcal{D}_K(\Omega)$ is open for any $K \subset\subset \Omega$. Here we shall not study this topology, and just characterize the corresponding notions of bounded subsets and of convergent sequences, which suffice for our purposes.

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

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

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

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

(i) there exists a $K \subset\subset \Omega$ 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 $\alpha \in \mathbb{N}^N$.

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

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

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

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

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

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

(iv)

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

(If the latter condition is fulfilled, one says that T has order m on the compact set K. It should be noticed that m may depend on K.)

Here are some examples of distributions:

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

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

is a distribution. The mapping $f \mapsto T_f$ is injective, so that we can identify $L^1_{\text{loc}}(\Omega)$ with a subspace of $\mathcal{D}'(\Omega)$.

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

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

is a distribution, that is usually identified with μ itself.

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

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

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

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

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

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

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

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

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

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

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

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

$$T_n \to T \quad \text{in } \mathcal{D}'(\Omega) \quad \Leftrightarrow \quad T_n(v) \to T \quad \forall v \in \mathcal{D}(\Omega)\,. \tag{1.7}$$

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

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

Differentiation of Distributions.

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

$$\langle fT, v\rangle := \langle T, fv\rangle \qquad \forall T \in \mathcal{D}'(\Omega), \forall f \in C^\infty(\Omega), \forall v \in \mathcal{D}(\Omega)\,, \tag{1.8}$$

$$\langle \tilde{D}^\alpha T, v\rangle := (-1)^{|\alpha|}\langle T, D^\alpha v\rangle \qquad \forall T \in \mathcal{D}'(\Omega), \forall v \in \mathcal{D}(\Omega), \forall \alpha \in \mathbb{N}^N\,. \tag{1.9}$$

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

The definition (1.8) is consistent with the properties of $L^1_{\mathrm{loc}}(\Omega)$. For any $f \in L^1_{\mathrm{loc}}(\Omega)$, the definition (1.9) is also consistent with partial integration: if $T = T_f$, (1.9) indeed reads

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

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

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

By (1.9) and as derivatives commute in $\mathcal{D}(\Omega)$, the same applies to $\mathcal{D}'(\Omega)$, that is,

$$\tilde{D}^\alpha \circ \tilde{D}^\beta T = \tilde{D}^{\alpha+\beta} T = \tilde{D}^\beta \circ \tilde{D}^\alpha T \qquad \forall T \in \mathcal{D}'(\Omega), \forall \alpha, \beta \in \mathbb{N}^N. \tag{1.10}$$

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

$$\begin{aligned} &\tilde{D}_i(fT) = (D_i f)T + f\tilde{D}_i T \\ &\forall f \in C^\infty(\Omega), \forall T \in \mathcal{D}'(\Omega), \text{ for } i = 1, .., N; \end{aligned} \tag{1.11}$$

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

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

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

$$\begin{aligned} &\tilde{D}^\alpha(fT) = \sum_{\beta\le\alpha} \binom{\alpha}{\beta} (D^{\alpha-\beta} f)\tilde{D}^\beta T \\ &\forall f \in C^\infty(\Omega), \forall T \in \mathcal{D}'(\Omega), \forall \alpha \in \mathbb{N}^N \,. \quad [Ex] \end{aligned} \tag{1.12}$$

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

Comparison with Classical Derivatives.

Theorem 1.3 (Du-Bois Reymond)
Let $f \in C^0(\Omega)$ and $i \in \{1, \dots, N\}$. Then $\tilde{D}_i f \in C^0(\Omega)$ (possibly after modification in a set of vanishing measure) iff f is classically differentiable w.r.t. x_i in the whole Ω, and $D_i f \in C^0(\Omega)$. In this case $\tilde{D}_i f = D_i f$ in Ω. []

The next statement applies to $\Omega :=]a, b[$, for $-\infty \le a < b \le +\infty$.

First we remind the reader that

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

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

$$H(x) := 0 \quad \forall x < 0 \qquad H(x) := 1 \quad \forall x \ge 0. \qquad [Ex] \tag{1.13}$$

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

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

Theorem 1.4 *A function $f \in L^1(a,b)$ is a.e. equal to an absolutely continuous function $\hat{f}$, iff $\tilde{D}f \in L^1(a,b)$. In this case $\tilde{D}f = D\hat{f}$ a.e. in $]a,b[$. []*

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.

Support and Order of Distributions. 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 $\Omega_1, \Omega_2, \Omega_3$,

$$\Omega_1 \subset \Omega_2 \subset \Omega_3 \quad \Rightarrow \quad \big(T\big|_{\Omega_2}\big)\big|_{\Omega_1} = T\big|_{\Omega_1} \qquad \forall T \in \mathcal{D}'(\Omega_3). \tag{1.14}$$

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 $\mathrm{supp}(T)$.

For any $K \subset\subset \Omega$, the smallest integer m that fulfills the estimate (1.3) is called the order of T in K. The supremum of these orders is called the **order** of T; each distribution is thus of either finite or infinite order. For instance, the functions in $L^1_{\mathrm{loc}}(\Omega)$ and the Dirac mass are distributions of order zero. [Ex] The distribution $D^\alpha\delta_0$ is of order $|\alpha|$ for any $\alpha \in \mathbb{N}^N$, and the distribution p.v. $(1/x)$ is of order one in $\mathcal{D}'(\mathbb{R})$. [Ex] On the other hand, $\sum_{n=1}^\infty D^n\delta_n$ is of infinite order in $\mathcal{D}'(\mathbb{R})$.

The next statement easily follows from (1.3), and establishes a strict relation between support and order.

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

The converse statement fails: a counterexample is provided by $\sum_{n=1}^\infty a_n D^n \delta_{x_n}$, for any sequence $\{a_n\}$ in $\mathbb{C}$ and for any sequence $\{x_n\}$ that belongs to any $K \subset\subset \Omega$ at most for a finite number of points. (E.g., $\{x_n\} = \mathbb{N}$ if $\Omega = \mathbb{R}$, $\{x_n\} = \{1/n\}$ if $\Omega = \mathbb{R} \setminus \{0\}$ but not if $\Omega = \mathbb{R}$.)

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 $u \in \mathcal{E}$,

$$\begin{aligned} &u_n \to u \quad \text{ in } \mathcal{E}(\Omega) \quad \Leftrightarrow \\ &\sup_{x\in K} |D^\alpha(u_n - u)(x)| \to 0 \quad \forall K \subset\subset \Omega, \quad \forall \alpha \in \mathbb{N}^N. \end{aligned} \tag{1.15}$$

Notice that

$$\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega) \quad \text{with continuous and dense injection} \tag{1.16}$$

(namely, $\mathcal{D}(\Omega)$ is a dense subset of $\mathcal{E}(\Omega)$). This may be checked via convolution with a sequence of regularizing kernels

$$\rho_n(x) = C_n^{-1}\rho(nx) \quad \text{ with } C_n = \textstyle\int_{\mathbb{R}^N} \rho(nx)\,dx \text{ and } \rho \text{ as in (1.2)}, \forall n. \tag{1.17}$$

[Ex] We may then identify the (topological) dual space $\mathcal{E}'(\Omega)$ with a dense subspace of $\mathcal{D}'(\Omega)$, and equip it with the induced convergence. For any sequence $\{T_n\}$ and any T in $\mathcal{E}'(\Omega)$, thus $T_n \to T$ in $\mathcal{E}'(\Omega)$ iff $T_n \to T$ in $\mathcal{D}'(\Omega)$. By (1.16) it follows that

$$T_n \to T \quad \text{ in } \mathcal{E}'(\Omega) \quad \Leftrightarrow \quad T_n(v) \to T(v) \quad \forall v \in \mathcal{E}(\Omega)\,.[Ex] \tag{1.18}$$

Theorem 1.6 $\mathcal{E}'(\Omega)$ *may be identified with the subspace of distributions having compact support.*

We just outline a part of the argument. Let $T \in \mathcal{D}'(\Omega)$ have support $K \subset\subset \Omega$. For any $v \in \mathcal{E}(\Omega)$, multiplying it by χ_K and then applying the convolution with a regularizing kernel ρ (see (1.17)), one may construct $v_0 \in \mathcal{D}(\Omega)$ such that $v_0 = v$ in K. [Ex] One may thus define $\tilde{T}(v)$ by setting $\tilde{T}(v) = T(v_0)$, and this determines a unique $\tilde{T} \in \mathcal{E}'(\Omega)$. Compactly supported distributions may thus be identified with some elements of $\mathcal{E}'(\Omega)$. The surjectivity of the mapping $T \mapsto \tilde{T}$ is less straightforward.

The Space $\mathcal{S}$ of Rapidly Decreasing Functions. In order to extend the Fourier transform to distributions, Laurent Schwartz introduced the space of **(infinitely differentiable) rapidly decreasing functions** (at ∞): [2]

$$\begin{aligned}\mathcal{S}\big(\mathbb{R}^N\big) :=& \big\{v \in C^\infty : \forall\alpha, \beta \in \mathbb{N}^N, x^\beta D^\alpha v \in L^\infty\big\}\\ =& \big\{v \in C^\infty : \forall\alpha \in \mathbb{N}^N, \forall m \in \mathbb{N},\\ & |x|^m D^\alpha v(x) \to 0 \text{ as } |x| \to +\infty\big\}.\end{aligned} \tag{1.19}$$

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 []

$$|v|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^N} |x^\beta D^\alpha v(x)| \qquad \alpha, \beta \in \mathbb{N}^N\,, \tag{1.20}$$

$$|v|_{m,\alpha} := \sup_{x\in\mathbb{R}^N} (1 + |x|^2)^m |D^\alpha v(x)| \qquad m \in \mathbb{N}, \alpha \in \mathbb{N}^N\,. \tag{1.21}$$

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

$$u \mapsto P(x)Q(D)u, \qquad u \mapsto P(D)[Q(x)u] \tag{1.22}$$

map $\mathcal{S}$ to $\mathcal{S}$ and are continuous.

The Space $\mathcal{S}'$ of Tempered Distributions. We shall denote the (topological) dual of $\mathcal{S}$ by $\mathcal{S}'$. Similarly to what we saw for $\mathcal{E}(\Omega)$, it is easily checked that

$$\mathcal{D} \subset \mathcal{S} \quad \text{with continuous and dense injection.}[Ex] \tag{1.23}$$

We may then identify the (topological) dual space $\mathcal{S}'$ with a dense subspace of $\mathcal{D}'$, and equip it with the induced convergence. For any sequence $\{T_n\}$ and any T in $\mathcal{S}'(\Omega)$, thus

$$T_n \to T \quad \text{in } \mathcal{S}' \quad \Leftrightarrow \quad T_n(v) \to T(v) \quad \forall v \in \mathcal{S}\,. \tag{1.24}$$

Here are some examples of tempered distributions:
— Any compactly supported distribution.
— Any $f \in L^p$ with $p \in [1, +\infty]$.
— Any function f such that $|f(x)| \le C(1 + |x|)^m$ for some $C > 0$ and $m \in \mathbb{N}$.
— Any function of the form $f(x) = p(x)w(x)$, p being a polynomial and $w \in L^1$. [Ex]
On the other hand $L^1_{\rm loc}$ is not included in $\mathcal{S}'$, not even for $N = 1$. E.g., $e^{-|x|} \notin \mathcal{S}'$.

Theorem 1.7 *The following applies for X equal to any of the spaces $\mathcal{D}(\Omega)$, $\mathcal{D}'(\Omega)$, $\mathcal{E}(\Omega)$, $\mathcal{E}'(\Omega)$, $\mathcal{S}$, $\mathcal{S}'$:*
(i) The space X is sequentially complete.
(ii) A subset of X is sequentially relatively compact iff it is bounded. []

[2] 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 ones are Fréchet space, at variance with the first one; the same holds for the respective (topological) duals.

Notice that this does not subsume any monotonicity property; e.g., the nonmonotone function $e^{-|x|^2} \sin x$ is an element of $\mathcal{S}(\mathbb{R})$.

For instance for $X = \mathcal{D}'(\Omega)$, this means that for any sequence $\{T_n\}$ in $\mathcal{D}'(\Omega)$:
(i) if $T_n(v)$ is a Cauchy sequence for any $v \in \mathcal{D}(\Omega)$, then $\{T_n\}$ converges.
(ii) $\{T_n\}$ has a convergent subsequence iff the sequence $\{T_n(v)\}$ is bounded for any $v \in \mathcal{D}(\Omega)$.

The next statement conveys information on the structure of distributions.

Theorem 1.8 *For any $T \in \mathcal{D}'(\Omega)$, and for any $\alpha \in \mathbb{N}^N$ there exists $f_\alpha \in C^0(\Omega)$ such that any $K \subset\subset \Omega$ intersects the support of just a finite number of f_αs, and $T = \sum_\alpha D^\alpha f_\alpha$.*

If T has finite order, then the functions f_α can be selected so that only a finite number of them does not vanish identically. []

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

$$(f * g)(x) := \int f(x-y)g(y)\,dy \qquad \text{for a.e. } x \in \mathbb{R}^N, \tag{2.25}$$

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

$$\text{supp}(f * g) \subset \text{supp}(f) + \text{supp}(g). \quad [Ex] \tag{2.26}$$

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

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

$$\|f * g\|_{L^1} \le \|f\|_{L^1}\|g\|_{L^1} \qquad \forall f, g \in L^1. \tag{2.27}$$

(ii) $L^1_{loc} * L^1_{comp} \subset L^1_{loc}$, and [3]

$$\begin{aligned} &\|f * g\|_{L^1(K)} \le \|f\|_{L^1(K - supp(g))}\|g\|_{L^1} \\ &\forall K \subset\subset \mathbb{R}^N, \forall f \in L^1_{loc}, \forall g \in L^1_{comp}. \end{aligned} \tag{2.28}$$

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

(iii) For $N = 1$, $L^1_{loc}(\mathbb{R}^+) * L^1_{loc}(\mathbb{R}^+) \subset L^1_{loc}(\mathbb{R}^+)$. [4] *For any $f, g \in L^1_{loc}(\mathbb{R}^+)$,*

$$(f * g)(x) = \begin{cases} \int_0^x f(x-y)g(y)\,dy & \textit{for a.e. } x \ge 0 \\ 0 & \textit{for a.e. } x < 0, \end{cases} \tag{2.29}$$

$$\|f * g\|_{L^1(0,M)} \le \|f\|_{L^1(0,M)}\|g\|_{L^1(0,M)} \qquad \forall M > 0. \tag{2.30}$$

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

Proof. (i) For any $f, g \in L^1$, the function $(\mathbb{R}^N)^2 \to \mathbb{C} : (z, y) \mapsto f(z)g(y)$ is (absolutely) integrable, and by a change of 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 in L^1. Moreover

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

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

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

[3] By L^1_{comp} we denote the space of the functions $v \in L^1$ that have compact support.

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

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

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

The proof of the inclusion $L^1_{\rm comp} * L^1_{\rm comp} \subset L^1_{\rm comp}$ is based on (2.26), and is left to the Reader.

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

Proposition 2.2 *L^1, L^1_{comp} and $L^1_{loc}(\mathbb{R}^+)$, equipped with the convolution product, are commutative algebras (without unit).* [5] *In particular,*

$$\begin{aligned}&f * g = g * f, \qquad (f * g) * h = f * (g * h) \qquad a.e.\ in\ \mathbb{R}^N \\ &\forall (f,g,h) \in (L^1)^3 \cup \big(L^1_{loc} \times L^1_{comp} \times L^1_{comp}\big).\end{aligned} \tag{2.31}$$

If $N = 1$, the same holds for any $(f,g,h) \in L^1_{loc}(\mathbb{R}^+)^3$, too.

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

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

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 similarly checked. □

It is easily checked that $(L^1, *)$ and $(L^\infty, \cdot)$ (here "$\cdot$" stands for the pointwise product) are commutative Banach algebras; $(L^\infty, \cdot)$ has the unit element $e \equiv 1$. On the other hand, $(\mathcal{S}, *)$ and $(\mathcal{S}, \cdot)$ are commutative algebras without unit.

Convolution of L^p-Functions. The following result generalizes Proposition 2.1.

---

[5] Let a vector 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) $\|u * v\| \leq \|u\|\|v\|$ 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 $u \in X$.

• **Theorem 2.3 (Young)** *Let*

$$p, q, r \in [1, +\infty], \qquad p^{-1} + q^{-1} = 1 + r^{-1}.\,^{6} \tag{2.32}$$

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

$$\|f * g\|_{L^r} \le \|f\|_{L^p} \|g\|_{L^q} \qquad \forall f \in L^p, \forall g \in L^q. \tag{2.33}$$

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

$$\begin{aligned} &\|f * g\|_{L^r(K)} \le \|f\|_{L^p(K - supp(g))} \|g\|_{L^q} \\ &\forall K \subset\subset \mathbb{R}^N, \forall f \in L^p_{loc}, \forall g \in L^q_{loc}. \end{aligned} \tag{2.34}$$

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

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

$$\begin{aligned} &\|f * g\|_{L^r(0,M)} \le \|f\|_{L^p(0,M)} \|g\|_{L^q(0,M)} \\ &\forall M > 0, \forall f \in L^p_{loc}(\mathbb{R}^+), \forall g \in L^q_{loc}(\mathbb{R}^+). \end{aligned} \tag{2.35}$$

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

Proof. (i) If $p = +\infty$, then by (2.8) $q = 1$ and $r = +\infty$, and (2.33) obviously holds; let us then assume that $p < +\infty$. For any fixed $f \in L^p$, the generalized Minkowski inequality V.3.4' and the Hölder inequality V.3.2 respectively yield

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

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

($p^{-1} + (p')^{-1} = 1$). 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 (2.33) 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 $f \in L^1_{\rm loc}$ and $g \in L^1_{\rm comp}$, setting $S_g := {\rm supp}(g)$,

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

If $r = +\infty$ then $p = q = 1$, and we are in the situation of part (ii) of Proposition 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

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

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. □

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

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

$$T : L^{p_1}(\Omega) + L^{p_2}(\Omega) \to L^{q_1}(\Omega') + L^{q_2}(\Omega') \tag{2.36}$$

is a linear operator such that

$$T : L^{p_i}(\Omega) \to L^{q_i}(\Omega') \text{ is continuous.} \tag{2.37}$$

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

$$\frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}, \qquad \frac{1}{q} = \frac{\theta}{q_1} + \frac{1-\theta}{q_2}. \tag{2.38}$$

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

$$\|Tf\|_{L^{q_i}(\Omega')} \le M_i \|f\|_{L^{p_i}(\Omega)} \qquad \forall f \in L^{p_i}(\Omega) \ (i = 1, 2), \tag{2.39}$$

then

$$\|Tf\|_{L^q(\Omega')} \le M_1^{\theta} M_2^{1-\theta} \|f\|_{L^p(\Omega)} \qquad \forall f \in L^p(\Omega). \qquad [] \tag{2.40}$$

By this result, we one may regard $L^{p(\theta)}(\Omega)$ as an *interpolate space* between $L^{p_1}(\Omega)$ and $L^{p_2}(\Omega)$. This is actually the prototypical example of the theory of Banach spaces interpolation.

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

Corollary 2.5 *Let*

$$p, q, s \in [1, +\infty], \qquad p^{-1} + q^{-1} + s^{-1} = 2. \tag{2.41}$$

Then:

$$\begin{aligned} &\forall (f, g, h) \in L^p \times L^q \times L^s, \\ &(f * g) \cdot h, \ g \cdot (\check{f} * h), \ f \cdot (\check{g} * h) \in L^1, \quad \textit{and} \\ &\int (f * g) \cdot h = \int g \cdot (\check{f} * h) = \int f \cdot (\check{g} * h). \end{aligned} \tag{2.42}$$

The same holds also

$$\begin{aligned} &\forall (f, g, h) \in \left(L^p_{comp} \times L^q_{loc} \times L^s_{comp}\right), \\ &\forall (f, g, h) \in L^p_{loc}(\mathbb{R}^+) \times L^q_{loc}(\mathbb{R}^+) \times L^s_{comp}(\mathbb{R}^+). \end{aligned} \tag{2.43}$$

Proof. For any $(f, g, h) \in L^p \times L^q \times L^s$, by the Young theorem $f * g \in L^r$ for r as in (2.32). By (2.41) then $r^{-1} + s^{-1} = 1$, and (2.42) follows.

The remainder is similarly checked. □

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

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

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

$$\tau_h f \to f \qquad \text{in } C^0, \ \forall f \in C^0, \tag{2.44}$$

$$\tau_h f \to f \qquad \text{in } L^p, \ \forall f \in L^p, \ \forall p \in [1, +\infty[. \tag{2.45}$$

Proof. As any $f \in C^0$ is locally uniformly continuous, $\tau_h f \to f$ uniformly in any $K \subset\subset \mathbb{R}^N$; (2.44) thus holds. This yields (2.45), as C^0 is dense in L^p for any $p \in [1, +\infty[$, cf. Theorem V.3.8. □

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

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

$$f * g \in C^0 \qquad \forall (f, g) \in (L^p \times L^q) \cup (L^p_{loc} \times L^q_{comp}), \tag{2.46}$$

$$f * g \in C^0 \qquad \forall (f, g) \in L^p_{loc}(\mathbb{R}^+) \times L^q_{loc}(\mathbb{R}^+) \quad \text{if } N = 1, \tag{2.47}$$

$$(f * g)(x) \to 0 \quad \text{as} \quad |x| \to +\infty \quad \forall (f, g) \in L^p \times L^q, \forall p, q \in [1, +\infty[. \tag{2.48}$$

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

$$\begin{aligned} \|\tau_h(f * g) - (f * g)\|_{L^\infty} &= \Big\| \int [f(x + h - y) - f(x - y)] g(y)\, dy \Big\|_{L^\infty} \\ &\leq \|\tau_h f - f\|_{L^p} \|g\|_{L^q} \to 0 \quad \text{ as } h \to 0; \end{aligned} \tag{2.49}$$

the function $f * g$ can then be identified to a uniformly continuous function.

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. □

It is easily seen that if either p or $q = +\infty$ then (2.48) fails.

Convolution of Distributions. By part (i) of Proposition 2.1, for any $(f, g) \in (L^1_{\rm loc} \times L^1) \cup (L^1 \times L^1_{\rm loc})$, $f * g \in L^1_{\rm loc}$. For any $\varphi \in \mathcal{D}$ then

$$\int (f * g)(x) \varphi(x)\, dx = \iint f(x - y) g(y) \varphi(x)\, dx dy = \iint f(z) g(y) \varphi(z + y)\, dx dy, \tag{2.50}$$

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

$$\langle T * S, v \rangle := \langle T_x, \big(\langle S_y, \varphi(x + y) \rangle \big) \rangle. \tag{2.51}$$

(By this notation we mean that in the bracketed term x is just a parameter, and the duality pairing just involves the variable y.) This is meaningful, since

$$S \in \mathcal{E}' \ \ (S \in \mathcal{D}', \text{ resp.}) \qquad \Rightarrow \quad \langle S_y, \varphi(x + y) \rangle \in \mathcal{D} \ \ (\in \mathcal{E}, \text{ resp.}). \ [Ex] \tag{2.52}$$

For $N = 1$, if $T \in \mathcal{D}'(\mathbb{R}^+)$, then (2.51) also makes sense.

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

The following properties hold: The convolution commutes and is associative; thus $(\mathcal{E}', *)$ is a convolution algebra, with unit element δ_0.

Here are some further properties:

$$\mathcal{D}' * \mathcal{E}' \subset \mathcal{D}', \qquad \mathcal{E}' * \mathcal{E}' \subset \mathcal{E}', \tag{2.53}$$

$$\mathcal{S}' * \mathcal{E}' \subset \mathcal{S}', \qquad \mathcal{S} * \mathcal{S}' \subset \mathcal{E} \cap \mathcal{S}', \tag{2.54}$$

$$\mathcal{S} * \mathcal{E}' \subset \mathcal{S}, \qquad \mathcal{S} * \mathcal{S}' \subset \mathcal{E}, \tag{2.55}$$

and in all of these cases the convolution is separately continuous w.r.t. each of the two factors.

3 The Fourier Transform in L^1

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

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

for a prescribed kernel $K : \mathbb{R}^2 \to \mathbb{C}$, and for any transformable function u. [7] The main properties of this class of transforms include the following:

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

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

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

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

$$TLT^{-1} = \widetilde{L}(s) \quad \text{(this is a multiplicative operator).} \tag{3.3}$$

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

The first of the transforms that we illustrate is named after J. Fourier, who introduced it at the beginning of the 19th century, and is the keystone of all integral transforms. In the 1950s Laurent Schwartz introduced the space of *tempered distributions,* and extended the transform to this class. This transform allows one to reduce linear ordinary differential equations with constant coefficients to algebraic equations, and this has found many uses in the study of stationary problems. This is useful for applications, and is also an important tool in functional analysis, as we shall see in Chapter X.

The Fourier Transform in L^1. We shall systematically deal with spaces of functions from the whole $\mathbb{R}^N$ to $\mathbb{C}$. We shall then write L^1 in place of $L^1(\mathbb{R}^N)$, C^0 in place of $C^0(\mathbb{R}^N)$, and so on. For any $u \in L^1$, we define the **Fourier transform** (also called **Fourier integral**) $\widehat{u}$ of u by [8]

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

here $\xi \cdot x := \sum_{i=1}^N \xi_i x_i$.

Proposition 3.1 *Formula (3.5) defines a linear and continuous operator*

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

(By C^0_b we denote the Banach space $C^0_b\mathcal{L}^\infty$.)

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

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

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

---

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

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

Proposition 3.2 *For any $u \in L^1$,* [9]

$$\begin{aligned}
v(x) = u(x-y) \quad &\Rightarrow \quad \widehat{v}(\xi) = e^{-i\xi\cdot y}\widehat{u}(\xi) \qquad \forall y \in \mathbb{R}^N, && (3.6)\\
v(x) = e^{ix\cdot\eta}u(x) \quad &\Rightarrow \quad \widehat{v}(\xi) = \widehat{u}(\xi-\eta) \qquad \forall \eta \in \mathbb{R}^N, && (3.7)\\
v(x) = u\big(A^{-1}x\big) \quad &\Rightarrow \quad \widehat{v}(\xi) = |\det A|\widehat{u}(A^*\xi) \quad \forall A \in \mathbb{R}^{N^2}, \det A \neq 0\,, && (3.8)\\
v(x) = \overline{u(x)} \quad &\Rightarrow \quad \widehat{v}(\xi) = \overline{\widehat{u}(-\xi)}, && (3.9)\\
u \text{ is even (odd, resp.)} \quad &\Rightarrow \quad \widehat{u} \text{ is even (odd, resp.)}, && (3.10)\\
u \text{ is real and even} \quad &\Rightarrow \quad \widehat{u} \text{ is real (and even)}, && (3.11)\\
u \text{ is real and odd} \quad &\Rightarrow \quad \widehat{u} \text{ is imaginary (and odd)}, && (3.12)\\
u \text{ is radial} \quad &\Rightarrow \quad \widehat{u} \text{ is radial.} && (3.13)
\end{aligned}$$

[Ex]

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

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

Proof. Let us set

$$\rho(x) := \exp\Big(-\frac{|x|^2}{1-|x|^2}\Big) \quad \text{if } |x| < 1, \quad \rho(x) := 0 \quad \text{if } |x| \geq 1,$$
$$\rho_n(x) := \rho\big(\frac{x}{n}\big) \quad \forall x \in \mathbb{R}^N, \forall n \in \mathbb{N}.$$

Hence $\rho_n(x) \to 1$ pointwise in $\mathbb{R}^N$ as $n \to \infty$, and

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

Therefore, by the dominated convergence theorem,

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

□

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

$$\begin{aligned}
u, D_x^\alpha u \in L^1 \quad &\Rightarrow \quad i^{|\alpha|}\xi^\alpha\widehat{u} = (D_x^\alpha u)\widehat{\ } \in C_b^0, && (3.14)\\
u, x^\alpha u \in L^1 \quad &\Rightarrow \quad D_\xi^\alpha\widehat{u} = (-i)^{|\alpha|}(x^\alpha u)\widehat{\ } \in C_b^0. && (3.15)
\end{aligned}$$

Proof. In both cases it suffices to prove the equality for any first-order derivative $D_j = \partial/\partial x_j$; the general case follows by induction.
(i) As
$$D_j[e^{-i\xi\cdot x}u(x)] = -i\xi_j e^{-i\xi\cdot x}u(x) + e^{-i\xi\cdot x}D_ju(x),$$
the integrability assumptions entail that $D_j[e^{-i\xi\cdot x}u(x)] \in L^1$. It then suffices to integrate the latter equality over $\mathbb{R}^N$, and to notice that $\int_{\mathbb{R}^N} D_j[e^{-i\xi\cdot x}u(x)]\,dx = 0$ by Lemma 3.3. Finally $(D_x^\alpha u)\widehat{\ } \in C_b^0$,

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

by Proposition 3.1.
(ii) Denoting by e_j the unit vector in the jth direction, we have

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

Passing to the limit as $t \to 0$, by the dominated convergence theorem we then get $D_j\widehat{u}(\xi) = -i(x_j u)\widehat{\ }(\xi)$ for any ξ. By Proposition 3.1, this is an element of C^0_b. □

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

In other terms:
(i) the faster u decreases at infinity, the greater is the regularity of $\widehat{u}$;
(ii) the greater is the regularity of u, the faster $\widehat{u}$ decreases at infinity.

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

Proof. For any $u \in L^1$, there exists a sequence $\{u_n\}$ in $\mathcal{D}$ such that $u_n \to u$ in L^1. By part (i) of Corollary 3.5, $\widehat{u}_n(\xi) \to 0$ as $|\xi| \to +\infty$. This holds also for $\widehat{u}$, as $\widehat{u}_n \to \widehat{u}$ uniformly in $\mathbb{R}^N$ by Proposition 3.1.

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

As $\widehat{u} \in C^0_b$, the uniform continuity follows. □

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

$$\int_{\mathbb{R}^N} \widehat{u}\, v\,dx = \int_{\mathbb{R}^N} u\,\widehat{v}\,dx \qquad \forall u, v \in L^1. \tag{3.16}$$

Moreover,

$$u{*}v \in L^1, \quad \text{and} \quad (u{*}v)\widehat{\ } = (2\pi)^{N/2}\widehat{u}\,\widehat{v} \qquad \forall u, v \in L^1. \tag{3.17}$$

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

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

as well as, via the change of integration variable $z = x - y$,

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

$\square$

We now present the inversion formula for the Fourier transform. First let us introduce the so-called *conjugate Fourier transform:*

$$\widetilde{\mathcal{F}}(v)(x) := (2\pi)^{-N/2} \int_{\mathbb{R}^N} e^{2\pi i\xi\cdot x} v(\xi) d\xi \qquad \forall v \in L^1. \tag{3.18}$$

Clearly this has analogous properties to the operator $\mathcal{F}$.

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

$$u(x) = \widetilde{\mathcal{F}}(\widehat{u})(x) \qquad \forall x \in \mathbb{R}^N. \tag{3.19}$$

Proof. Let us set $v(x) := \exp(-|x|^2/2)$ for any $x \in \mathbb{R}^N$. A calculation based on integration along paths in the complex plane shows that $\widehat{v}(\xi) := \exp(-|\xi|^2/2)$ for any $\xi \in \mathbb{R}^N$. [] By the Tonelli and Fubini theorems, we have

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

Let us now replace $v(\xi)$ by $v_\varepsilon(\xi) := v(\varepsilon\xi)$, for any $\varepsilon > 0$. By (3.8), $\widehat{v_\varepsilon}(z) = \varepsilon^{-N}\widehat{v}(\varepsilon^{-1}z)$, and via a further change of variable of integration we get

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

As u is continuous and bounded, by the dominated convergence theorem, as $\varepsilon \to 0$ we get

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

As $v(0) = 1$ and $\int_{\mathbb{R}^N} \widehat{v}(y)\, dy = (2\pi)^{N/2}$, (3.19) follows. $\square$

By Proposition 3.1, the regularity assumptions of Theorem 3.8 are actually needed, as $\bar{u} = \mathcal{F}(\overline{\widehat{u}})$. However, by a more refined argument one might show that (3.19) holds under the only hypotheses that $u, \widehat{u} \in L^1$. (Of course, a posteriori one then gets that $u, \widehat{u} \in C^0_b$.)

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

$$\overline{\widehat{\widehat{u}}(x)} = \bar{u}(-x) \qquad \forall x \in \mathbb{R}^N. \tag{3.20}$$

We recall that $B(0,R)$ denotes the closed ball in $\mathbb{R}^N$ with center at the origin and radius R.

Theorem 3.9 (Paley-Wiener)
For any $u \in C^\infty(\mathbb{R}^N)$ and any $R > 0$, $\operatorname{supp} u \subset B(0,R)$ iff $\mathcal{F}(u)$ can be extended to an analytic function $\mathbb{C}^N \to \mathbb{C}$ (also denoted by $\mathcal{F}(u)$) [10] *such that*

$$\forall m \in \mathbb{N}, \exists C \geq 0 : \forall z \in \mathbb{C}^N \qquad [\mathcal{F}(u)](z)| \leq C \frac{e^{R|\mathcal{I}(z)|}}{(1+|z|)^m}.[] \tag{3.21}$$

[10] A function $\mathbb{C}^N \to \mathbb{C}$ is called analytic iff it is separately analytic with respect to each variable. For any $z \in \mathbb{C}^N$, we set $|z| = \left(\sum_{i=1}^N |z_i|^2\right)^{1/2}$ and $\mathcal{I}(z) = (\mathcal{I}(z_1), ..., \mathcal{I}(z_N))$ (the vector of the imaginary parts).

This extended function $\mathcal{F}(u) : \mathbb{C}^N \to \mathbb{C}$ is called the **Fourier-Laplace transform** of u.

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

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

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

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

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

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

The inversion formula (3.19) also provides an interpretation of the Fourier transform. (3.19) represents u as a weighted average of the *harmonic components* $x \mapsto e^{i\xi\cdot x}$. For any $\xi \in \mathbb{R}^N$, $\widehat{u}(\xi)$ is the *amplitude* of the component having *vector frequency* ξ (that is, frequency ξ_i in each direction x_i). Therefore any function which fulfills (3.19) can equivalently be represented by specifying either the value $u(x)$ at a.a. points $x \in \mathbb{R}^N$, or the amplitude $\widehat{u}(\xi)$ for a.a. frequencies $\xi \in \mathbb{R}^N$. (Loosely speaking, any non-identically vanishing $u \in \mathcal{D}$ has harmonic components of arbitrarly large frequencies.)

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

4 Extensions of the Fourier Transform

Fourier Transform of Measures. The Fourier transform can easily be extended to any finite complex Borel measure μ on $\mathbb{R}^N$, simply by replacing $f(x)\,dx$ with $d\mu(x)$ in (3.5). In this case one speaks of the **Fourier-Stieltjes transform.** Most of the previously established properties holds also in this case. For instance, transformed functions are still elements of C^0_b and fulfill the properties of transformation of derivatives and multiplication by a power of x. The Riemann-Lebesgue theorem (Proposition 3.6) however fails; e.g., $\widehat{\delta_0}(\xi)$ does not vanish as $|\xi| \to +\infty$. For instance, the transform of the Dirac measure at the origin is the function identically equal to $(2\pi)^{-N/2}$, that is, $\widehat{\delta_0} = (2\pi)^{-N/2}$.

Fourier Transform in $\mathcal{S}$. For any $u \in \mathcal{D}$, by Theorem 3.9 $\widehat{u}$ is analytic; hence $\widehat{u} \in \mathcal{D}$ only if $\widehat{u} \equiv 0$, namely $u \equiv 0$. Thus $\mathcal{D}$ is not stable by Fourier transform. This means that the set of the frequencies of the harmonic components of any non-identically vanishing $u \in \mathcal{D}$ is unbounded. This situation induced L. Schwartz to introduce the space of rapidly decreasing functions $\mathcal{S}$, cf. Sect. 1, and to extend the Fourier transform to this space and to its (topological) dual. Next we review the tenets of that theory.

Proposition 4.1 *(The restriction of) $\mathcal{F}$ operates in $\mathcal{S}$ and is continuous. Moreover, the formulae of Proposition 3.4 and Theorem 3.7 hold in $\mathcal{S}$ without any restriction, $\mathcal{F}$ is invertible in $\mathcal{S}$, and $\mathcal{F}^{-1} = \widetilde{\widetilde{\mathcal{F}}}$ (cf. (3.19)). [Ex]*

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

The next statement extends and also completes (3.17).

Proposition 4.2 *For any $u, v \in \mathcal{S}$,*

$$u{*}v \in \mathcal{S}, \qquad (u{*}v)\widehat{\ } = (2\pi)^{N/2}\widehat{u}\,\widehat{v} \quad in\ \mathcal{S}, \tag{4.1}$$

$$uv \in \mathcal{S}\,, \qquad (uv)\widehat{\ } = (2\pi)^{-N/2}\widehat{u}{*}\widehat{v} \quad in\ \mathcal{S}. \tag{4.2}$$

Proof. The first statement is a direct extension of (3.17). Let us prove the second one.

It is easily checked that $uv \in \mathcal{S}$. By writing (3.17) with $\widehat{u}$ and $\widehat{v}$ in place of u and v, and $\widetilde{\mathcal{F}}$ in place of $\mathcal{F}$, we have

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

By applying $\mathcal{F}$ to both members of this equality, (4.2) follows. □

Fourier Transform in $\mathcal{S}'$. Denoting by $\mathcal{F}^\tau$ the transposed of $\mathcal{F}$, we set

$$\bar{\mathcal{F}} := [\mathcal{F}^\tau]^* : \mathcal{S}' \to \mathcal{S}'. \tag{4.3}$$

By Theorem 1.6, $\mathcal{F}^\tau = \mathcal{F}$; hence $\bar{\mathcal{F}} = \mathcal{F}^*$, that is,

$$\langle \bar{\mathcal{F}}(T), v\rangle := \langle T, \mathcal{F}(v)\rangle \qquad \forall v \in \mathcal{S}, \forall T \in \mathcal{S}'. \tag{4.4}$$

As $\mathcal{S}$ is dense in $\mathcal{S}'$, cf. Proposition VIII.6.2, $\bar{\mathcal{F}}$ is the unique continuous extension of the Fourier transform from $\mathcal{S}$ to $\mathcal{S}'$.

Henceforth we shall use the same symbols $\mathcal{F}$ or $\widehat{\ }$ for the many restrictions and extensions of the Fourier transform. We shall thus write $\mathcal{F}(T)$, or $\widehat{T}$, in place of $\bar{\mathcal{F}}(T)$.

Proposition 4.3 *$\mathcal{F}$ can be uniquely extended to an operator which acts in $\mathcal{S}'$ and is continuous. Moreover, the formulae of Proposition 3.4 and Theorem 3.7 hold in $\mathcal{S}$ without any restriction, $\mathcal{F}$ is invertible in $\mathcal{S}$, and $\mathcal{F}^{-1} = \widetilde{\mathcal{F}}$. [Ex]*

For instance, for any $v \in \mathcal{D}$ we have

$$\begin{aligned}
{}_{\mathcal{D}'}\langle i^{|\alpha|}\xi^\alpha \widehat{T}, v\rangle_{\mathcal{D}} &= {}_{\mathcal{D}'}\langle \widehat{T}, i^{|\alpha|}\xi^\alpha v\rangle_{\mathcal{D}} = {}_{\mathcal{D}'}\langle T, [i^{|\alpha|}\xi^\alpha v]\widehat{\ }\rangle_{\mathcal{D}} \\
&\overset{(3.15)}{=} {}_{\mathcal{D}'}\langle T, (-D)^\alpha \widehat{v}\rangle_{\mathcal{D}} = {}_{\mathcal{D}'}\langle D^\alpha T, \widehat{v}\rangle_{\mathcal{D}} = {}_{\mathcal{D}'}\langle (D^\alpha T)\widehat{\ }, v\rangle_{\mathcal{D}}.
\end{aligned}$$

As $\mathcal{D}$ is a dense subspace of $\mathcal{S}$, we conclude that

$$i^{|\alpha|}\xi^\alpha \widehat{T} = (D_x^\alpha T)\widehat{\ } \in \mathcal{S}' \qquad \forall T \in \mathcal{S}', \forall \alpha \in \mathbb{N}^N. \tag{4.5}$$

Proposition 4.4 *For any $u \in \mathcal{S}$ and any $T \in \mathcal{S}'$,*

$$u{*}T \in \mathcal{S}', \qquad (u{*}T)\widehat{\ } = (2\pi)^{N/2}\widehat{u}\,\widehat{T} \quad in\ \mathcal{S}', \tag{4.6}$$

$$uT \in \mathcal{S}', \qquad (uT)\widehat{\ } = (2\pi)^{-N/2}\widehat{u}{*}\widehat{T} \quad in\ \mathcal{S}'. \tag{4.7}$$

Theorem 4.5 *For any $T \in \mathcal{E}'$ and any $\xi \in \mathbb{R}^N$, $\widehat{T}(\xi) = {}_{\mathcal{E}'}\langle T, e^{-ix\cdot\xi}\rangle_{\mathcal{E}}$. This expression can be extended to any $\xi \in \mathbb{C}^N$, and is an analytic function.*

(This extends the Fourier-Laplace transform of the previous section.)

Proof. For any $\varepsilon > 0$, let us define the mollifier ρ_n as in (1.17), and set $(T{*}\rho_n)(x) := \langle T_y, \rho_n(x-y)\rangle$ for any $x \in \mathbb{R}^N$. (The index y indicates that T acts on the variable y; here x is just a parameter.) This yields $T * \rho_n \to T$ in $\mathcal{E}'$, hence also in $\mathcal{S}'$ as $\mathcal{E}' \subset \mathcal{S}'$ with continuous and dense injections.

Therefore

$$(T * \rho_\varepsilon)\widehat{\ } \to \widehat{T} \qquad \text{in } \mathcal{S}'. \tag{4.8}$$

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

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

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

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

By (4.8) we then conclude that $\widehat{T}(\xi) = \langle T_y, e^{-2\pi i\xi\cdot y}\rangle$ for any $\xi \in \mathbb{R}^N$, and this function is analytic. □

Theorem 4.6 (Paley-Wiener-Schwartz)
For any $T \in \mathcal{S}'$ and any $R > 0$, $\operatorname{supp} T \subset B(0,R)$ iff $\mathcal{F}(T)$ can be extended to an analytic function $\mathbb{C}^N \to \mathbb{C}$ (also denoted by $\mathcal{F}(T)$) [11] *such that*

$$\exists m \in \mathbb{N}_0, \exists C \geq 0 : \forall z \in \mathbb{C}^N \qquad |[\mathcal{F}(T)](z)| \leq C(1+|z|)^m e^{R|\mathcal{I}(z)|}. \qquad [] \tag{4.9}$$

Fourier Transform in L^2. As $L^2 \subset \mathcal{S}'$, any function of L^2 has a Fourier transform. Next we study the restriction of $\mathcal{F}$ to L^2 and show that it is an isometry.

- **Theorem 4.7 (Plancherel)**

$$u \in L^2 \quad \Leftrightarrow \quad \widehat{u} \in L^2 \qquad \forall u \in \mathcal{S}'. \tag{4.10}$$

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

$$\|\widehat{u}\|_{L^2} = \|u\|_{L^2} \qquad \forall u \in L^2. \tag{4.11}$$

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

$$(2\pi)^{-N/2} \int_{]-R,R[^N} e^{-i\xi\cdot x} u(x)\, dx \;\to\; \widehat{u}(\xi) \qquad \textit{in } L^2, \; \textit{as } R \to +\infty. \tag{4.12}$$

Therefore this sequence also converges in measure on any bounded subset of $\mathbb{R}^N$; it also converges a.e. in $\mathbb{R}^N$, as $R \to +\infty$ along some sequence which may depend on u.

Proof. For any $u \in \mathcal{S}$, we know that $\widehat{u} \in \mathcal{S}$. Moreover, by (3.16) and (3.20),

$$\begin{aligned}\int_{\mathbb{R}^N} |\widehat{u}|^2\, dx = \int_{\mathbb{R}^N} \widehat{u}\overline{\widehat{u}}\, dx = \int_{\mathbb{R}^N} u\widehat{\overline{\widehat{u}}}\, dx = \int_{\mathbb{R}^N} u(x)\overline{\widehat{\widehat{u}}(-x)}\, dx = \int_{\mathbb{R}^N} u\bar{u}\, dx \\ = \int_{\mathbb{R}^N} |u|^2\, dx.\end{aligned}$$

Therefore, as $\mathcal{S} \subset L^2$ with continuity and density, the restriction of $\mathcal{F}$ to L^2 is an isometry with respect to the L^2-metric. Hence $\mathcal{F}$ maps L^2 to itself.

[11] A function $\mathbb{C}^N \to \mathbb{C}$ is called analytic iff it is separately analytic with respect to each variable. For any $z \in \mathbb{C}^N$, we set $|z| = \big(\sum_{i=1}^N |z_i|^2\big)^{1/2}$ and $\mathcal{I}(z) = (\mathcal{I}(z_1), ..., \mathcal{I}(z_N))$ (the vector of the imaginary parts).

In order to prove (4.12), for any $R > 0$ and any $x \in \mathbb{R}$, let us set $\chi_R(x) := 1$ if $|x_i| \le R$ for $i = 1, ..., N$, and $\chi_R(x) := 0$ otherwise. Then $u\chi_R \in L^1 \cap L^2$ and $u\chi_R \to u$ in L^2. Hence, by (4.11),

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

□

Remarks. (i) We saw that in any Hilbert space H the scalar product is determined by the norm: $2(u, v) = \|u + v\|^2 - \|u\|^2 - \|v\|^2$ for any $u, v \in H$. (4.11) then entails that

$$\int_{\mathbb{R}^N} u(x)\bar{v}(x)\, dx = \int_{\mathbb{R}^N} \widehat{u}(\xi)\overline{\widehat{v}(\xi)}\, d\xi \qquad \forall u, v \in L^2, \tag{4.13}$$

(ii) The representation (4.12) is more general:

$$\int_{]-R,R[^N} e^{-2\pi i\xi\cdot x} u(x)\, dx \ \to\ \widehat{u}(\xi) \qquad \text{in } \mathcal{S}',\ \text{as } R \to +\infty\,, \forall u \in \mathcal{S}' \cap L^1_{\text{loc}}\,. \tag{4.14}$$

The above argument also allow one to generalize the inversion Theorem 3.8.
(iii) The Lebesgue-integral representation (3.5) is meaningful only if $u \in L^1$. Anyway it may be useful to know of cases in which the (extended) Fourier transform maps functions to functions. We claim that, for any $p \in [1, 2]$, any function $u \in L^p$ can be written as the sum of a function of L^1 and one of L^2, i.e.,

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

Indeed, setting $\chi := 1$ where $|u| \ge 1$ and $\chi := 0$ elsewhere, we have $u\chi \in L^1$, $u(1 - \chi) \in L^2$ and $u = u\chi + u(1 - \chi)$. [12] Therefore, as $\mathcal{F} : L^1 \to L^\infty$ and $\mathcal{F} : L^2 \to L^2$,

$$\mathcal{F}(u) = \mathcal{F}(u\chi) + \mathcal{F}(u(1 - \chi)) \in L^\infty + L^2 \qquad \forall u \in L^p, \forall p \in [1, 2]\,. \tag{4.16}$$

Thus $\mathcal{F}(u)$ is a regular distribution, although it may admit no integral representation.

This is made more precise by the following result, which is a direct consequence of the classic Riesz-Thorin Theorem 2.4.

Theorem 4.8 (Hausdorff-Young) [13]
Let $p \in [1, 2]$, $p' := p/(p-1)$ if $p > 1$, and $p' = \infty$ if $p = 1$. Then (the restriction of) $\mathcal{F}$ is a linear and continuous operator $L^p \to L^{p'}$. More precisely, for any $u \in L^p$ we have $\widehat{u} \in L^{p'}$ and $\|\widehat{u}\|_{L^{p'}} \le \|u\|_{L^p}$.

Proof. As $\mathcal{F}$ is linear and continuous as an operators $L^1 \to L^\infty$ and $L^2 \to L^2$, it suffices to apply the classical Riesz-Thorin Theorem 2.4. □

Because of the symmetry between the definitions of the direct and inverse Fourier transform, see formulas (3.5) and (3.19), the results that we established for $\mathcal{F}$ in $\mathcal{S}$ and in $\mathcal{S}'$, in particular Theorems 4.1 and 4.7, hold also for $\mathcal{F}^{-1}$.

We claim that $\mathcal{F}$ does not map L^p to L^q for any $q \ne p'$. Let $u \in L^p$ be such that $\mathcal{F}(u) \in L^q$. For any $\lambda > 0$, setting $u_\lambda(x) := u(\lambda x)$ for any x, by (3.8) we have $\mathcal{F}(u_\lambda) = \lambda^{-N}\mathcal{F}(u)_{1/\lambda}$; hence

$$\frac{\|\mathcal{F}(u_\lambda)\|_{L^q}}{\|u_\lambda\|_{L^p}} = \lambda^{-N} \frac{\|\mathcal{F}(u)_{1/\lambda}\|_{L^q}}{\|u_\lambda\|_{L^p}} = \lambda^{N(-1+1/q+1/p)} \frac{\|\mathcal{F}(u)\|_{L^q}}{\|u\|_{L^p}} \tag{4.17}$$

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

[12] Similarly one can show that $L^p \subset L^q + L^r$ whenever $1 \le q < p < r \le \infty$.
[13] We set $\infty^{-1} := 0$.

Overview of the Extensions of the Fourier Transform. At first we noticed that the Fourier transform (3.5) has an obvious extension for any complex Borel measure μ; loosely speaking, this is just defined by replacing $u(x)dx$ by $d\mu$ in the definition (3.5). By the Paley-Wiener theorem, $\mathcal{D}$ is not stable under application of the Fourier transform. However, $\mathcal{F}$ maps the Schwartz space $\mathcal{S}$ to itself; this allowed us to extend $\mathcal{F}$ to $\mathcal{S}'$ by transposition. We also saw that $\mathcal{F}$ is an isometry in L^2 (Plancherel theorem), that in this space $\mathcal{F}$ admits an integral representation as a principal value, and that $\mathcal{F}$ is also linear and continuous from L^p to $L^{p/(p-1)}$, for any $p \in]1,2[$.

Finally, we saw that the Fourier series arise as Fourier transforms of periodic functions.

Note: The Fourier transform is a homomorphism from the algebra $(L^1, *)$ to the algebra $(L^\infty, \cdot)$ (here "$\cdot$" stands for the product a.e.), and is an isomorphism between the algebras $(\mathcal{S}, *)$ and $(\mathcal{S}, \cdot)$; cf. (3.17).

5 The Laplace Transform

At variance with the Fourier transform, this transform is especially relevant for the study of evolution problems. In particular the Laplace transform allows one to reduce initial-value problems for ordinary differential equations with constant coefficients to algebraic equations. This includes equations as simple as $u'(t) = au(t)$ ($a \in \mathbb{R}$), whose solution indeed reads $u(t) = Ce^{-at}$ for any $C \in \mathbb{C}$; this is not even a tempered distribution, so that it cannot be treated via Fourier transform.

In the applications, this theory is also known as *operational calculus.*

In this section we just deal with spaces of functions defined on $\mathbb{R}$, rather than $\mathbb{R}^N$ as for the Fourier transform; as above, we then write L^1 in place of $L^1(\mathbb{R})$, and so on. The Laplace transform acts on complex-valued functions defined on positive reals, (or equivalently, defined on the whole $\mathbb{R}$ and assumed to vanish in $\mathbb{R}^-$). We then extend this definition to a class of distributions. At variance with the Fourier transform, in either case this transform yields functions of a single *complex* variable.

Laplace Transform of Functions. We define:
the linear space of the **(Laplace) transformable functions**

$$D_{\mathcal{L}} := \big\{u \in L^1_{\mathrm{loc}} : u = 0 \text{ a.e. in } \mathbb{R}^-, \exists x \in \mathbb{R} : \{e^{-xt}u(t)\} \in L^1\big\}, \tag{5.1}$$

the **abscissa of (absolute) convergence**

$$\lambda_a(u) := \inf\big\{s \in \mathbb{R} : \{e^{-st}u(t)\} \in L^1\big\} \quad (\in [-\infty, +\infty[) \qquad \forall u \in D_{\mathcal{L}}, \tag{5.2}$$

the **half-plane of (absolute) convergence**

$$\mathbb{C}_{\lambda_a(u)} := \{s \in \mathbb{C} : \mathrm{Re}(s) > \lambda_a(u)\} \qquad \forall u \in D_{\mathcal{L}}, \tag{5.3}$$

and the **(unilateral) Laplace transform** (or *Laplace integral)*

$$[\mathcal{L}(u)](s) := \int_{\mathbb{R}} e^{-st}u(t)\, dt \qquad \forall s \in \mathbb{C}_{\lambda_a(u)}, \forall u \in D_{\mathcal{L}}. \tag{5.4}$$

In this section we shall often write $\widehat{u} := \mathcal{L}(u)$. If $\lambda_a(u) = -\infty$ then the *halph-plane* $\mathbb{C}_{\lambda_a(u)}$ coincides with the whole plane $\mathbb{C}$; this is the case e.g. for $u(t) = e^{-t^2}H(t)$. [14] On the other hand $u(t) = e^{t^2}H(t)$ is no elements of $D_{\mathcal{L}}$. Moreover,

$$\begin{aligned} &\forall u \in D_{\mathcal{L}}, \text{ if } \exists M \in \mathbb{R} \text{ such that } u(t) \neq 0 \ \ \forall t > M, \text{ then} \\ &\lambda_a(u) = \limsup_{t \to +\infty} \frac{\log|u(t)|}{t}. \qquad [Ex] \end{aligned} \tag{5.5}$$

[14] We still denote by H the Heaviside function.

The *Laplace integral* (5.4) clearly converges absolutely for any $s \in \mathbb{C}_{\lambda_a(u)}$ (namely, the modulus of the integrand is integrable). Indeed it is promptly seen that if $\mathrm{Re}(s) > \lambda_a(u)$ then $e^{-st}u(t)$ decays exponentially as $t \to +\infty$, so it is integrable on $\mathbb{R}$. On the other hand, if $\mathrm{Re}(s) < \lambda_a(u)$ then $e^{-st}u(t)$ is unbounded for t large enough, thus it is not integrable.

Other formulations allow for the simple convergence of the integral. There exists also a *bilateral* Laplace transform, in which the input function is not assumed to vanish in $\mathbb{R}^-$; in this case the region of convergence is a (possibly degenerate) strip parallel to the imaginary axis.

Here are some simple examples. For any $\omega > 0$ and any $s_0 \in \mathbb{C}$,

$$\begin{aligned} &u(t) = H(t) &&\Rightarrow \quad \lambda_a(u) = 0, \quad \widehat{u}(s) = 1/s, \\ &u(t) = t^n H(t) &&\Rightarrow \quad \lambda_a(u) = 0, \quad \widehat{u}(s) = n!s^{-n-1} \ (\forall n \in \mathbb{N}), \\ &u(t) = (\sin \omega t)H(t) &&\Rightarrow \quad \lambda_a(u) = 0, \quad \widehat{u}(s) = \omega/(s^2+\omega), \\ &u(t) = e^{s_0 t}H(t) &&\Rightarrow \quad \lambda_a(u) = \mathrm{Re}(s_0), \quad \widehat{u}(s) = 1/(s-s_0). [Ex] \end{aligned} \tag{5.6}$$

The next result is a direct consequence of the definition of the Fourier and Laplace transforms, and illustrates how they are strictly related.

Proposition 5.1 *For any $u \in D_{\mathcal{L}}$,*

$$\begin{aligned} &\{e^{-xt}u(t)\} \in L^1, \\ &[\mathcal{L}(u)](x+iy) = \sqrt{2\pi}\big[\mathcal{F}(\{e^{-xt}u(t)\})\big](y) \end{aligned} \qquad \forall x+iy \in \mathbb{C}_{\lambda_a(u)}. \tag{5.7}$$

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

$$\lambda_a(u) < 0 \quad \Rightarrow \quad [\mathcal{F}(u)](\xi) = \frac{1}{\sqrt{2\pi}}[\mathcal{L}(u)](i\xi) \quad \forall \xi \in \mathbb{R}. \qquad [Ex] \tag{5.8}$$

The case of $\lambda(u) = 0$ is somehow delicate: as $u \in L^1$, $[\mathcal{F}(u)](\xi)$ exists for any $\xi \in \mathbb{R}$, but one cannot write $[\mathcal{L}(u)](i\xi)$. Thus $[\mathcal{L}(u)](x + 2\pi iy)$ need not converge to $[\mathcal{F}(u)](y)$ as $x \to 0^+$.

By (5.7), the Laplace transform may be regarded as the Fourier transform of a function which has preliminarily been damped down by means of a suitable real exponential factor. On the other hand, in (5.8) the one-dimensional Fourier transform is represented as the restriction of the Laplace transform to the imaginary axis (times the conventional factor $(2\pi)^{-1/2}$).

Next we display some basic properties of the Laplace transform of functions. They easily follow from the definition; because of Proposition 5.1, most of them may also be derived from the analogous properties of the Fourier transform.

Proposition 5.2 *For any $u \in D_{\mathcal{L}}$,*
(i) $\widehat{u}(s) \to 0$ as $\mathrm{Re}(s) \to +\infty$;
(ii) $\widehat{u}$ is bounded in the half-plane $\{s \in \mathbb{C} : \mathrm{Re}(s) \geq \lambda\}$, for any $\lambda > \lambda_a(u)$. [Ex]

The statement (i) means that $\widehat{u}(s_n) \to 0$ for any sequence $\{s_n\}$ in $\mathbb{C}_{\lambda_a(u)}$ such that $\mathrm{Re}(s_n) \to +\infty$, thus independently of $\mathrm{Im}(s_n)$.

The function $\widehat{u}$ need not be bounded in the whole half-plane $\{s \in \mathbb{C} : \mathrm{Re}(s) > \lambda(u)\}$. For instance, $\lambda(H) = 0$ and $\widehat{H}(s) = 1/s$ is unbounded in $\mathbb{C}_0 = \{s \in \mathbb{C} : \mathrm{Re}(s) > 0\}$.

Proposition 5.3 *For any $u \in D_{\mathcal{L}}$,*

$$v(t) = u(t-t_0) \quad \Rightarrow \quad \begin{aligned} &\widehat{v}(s) = e^{-t_0 s}\widehat{u}(s) \\ &\forall s \in \mathbb{C}_{\lambda_a(u)}, \ \forall t_0 > 0, \end{aligned} \tag{5.9}$$

$$v(t) = e^{s_0 t}u(t) \quad \Rightarrow \quad \begin{aligned} &\widehat{v}(s) = \widehat{u}(s-s_0) \\ &\forall s \in \mathbb{C}_{\lambda_a(u)+\mathrm{Re}(s_0)}, \ \forall s_0 \in \mathbb{C}, \end{aligned} \tag{5.10}$$

$$v(t) = u(\omega t) \quad \Rightarrow \quad \begin{aligned} &\widehat{v}(s) = \frac{1}{\omega}\widehat{u}\Big(\frac{s}{\omega}\Big) \\ &\forall s \in \mathbb{C}_{\omega\lambda_a(u)}, \ \forall \omega > 0. \end{aligned} \tag{5.11}$$

For any $u, v \in D_{\mathcal{L}}$,

$$\begin{aligned} &u * v \in D_{\mathcal{L}}, \quad \lambda_a(u * v) \leq \max\{\lambda_a(u), \lambda_a(v)\}, \\ &\widehat{u * v} = \widehat{u}\,\widehat{v} \qquad \text{in } \mathbb{C}_{\max\{\lambda_a(u),\lambda_a(v)\}}\,.[Ex] \end{aligned} \tag{5.12}$$

Proposition 5.4 *(i) For any $u \in D_{\mathcal{L}}$,*

$$\begin{aligned} &tu(t) \in D_{\mathcal{L}}, \quad \lambda_a(\{tu(t)\}) = \lambda_a(u), \\ &\widehat{u} \text{ is holomorphic in } \mathbb{C}_{\lambda_a(u)}, \\ &\widehat{tu(t)}(s) = -\widehat{u}'(s) \qquad \forall s \in \mathbb{C}_{\lambda_a(u)}. \end{aligned} \tag{5.13}$$

(ii) For any $u \in D_{\mathcal{L}}$, if $Du \in D_{\mathcal{L}}$ then

$$\widehat{Du}(s) = s\,\widehat{u}(s) \qquad \forall s \in \mathbb{C}_{\lambda_a(u)} \cap \mathbb{C}_{\lambda_a(Du)}. \qquad [] \tag{5.14}$$

By Du we denoted the derivative of u in the sense of distributions, and by $\widehat{u}'(s)$ the classical derivative of $\widehat{u}$. Note that $\lim_{t\to 0^+} u(t) = 0$ whenever $Du \in D_{\mathcal{L}}$. We shall see that the case where the latter condition fails is satisfactorily covered by the Laplace transform of distributions.

In part (ii) the assumption $Du \in D_{\mathcal{L}}$ cannot be dropped. For instance, for both $u(t) = t^{-1/2}H(t)$ and $u(t) = \sin(e^{t^2} - 1)H(t)$, $u \in D_{\mathcal{L}}$ but $Du \notin D_{\mathcal{L}}$.

In several cases the transformed function may be extended by analyticity to a larger domain than $\mathbb{C}_{\lambda_a(u)}$. Nevertheless it represents the Laplace transform, and thus benefits from the corresponding properties, just in that half-plane.

Corollary 5.5 *(i) For any $u \in D_{\mathcal{L}}$,*

$$\begin{aligned} &\{\textstyle\int_0^t u(\tau)\,d\tau\} \in D_{\mathcal{L}}, \quad \lambda_a(\{\int_0^t u(\tau)d\tau\}) = \lambda_a(u), \\ &\big[\mathcal{L}\big(\{\textstyle\int_0^t u(\tau)\,d\tau\}\big)\big](s) = \dfrac{\widehat{u}(s)}{s} \qquad \forall s \in \mathbb{C}_{\lambda_a(u)} \setminus \{0\}. \end{aligned} \tag{5.15}$$

(ii) For any $u \in D_{\mathcal{L}}$,

$$\begin{aligned} &\text{if } \{u(t)/t\} \in D_{\mathcal{L}}, \text{ then } \lambda_a(\{u(t)/t\}) = \lambda_a(u), \text{ and} \\ &[\mathcal{L}(\{u(t)/t\})](s) = \lim_{\sigma\to+\infty} \int_s^\sigma \widehat{u}(\tau)\,d\tau \qquad \forall s \in \mathbb{C}_{\lambda_a(u)}\,. \end{aligned} \tag{5.16}$$

The latter limit coincides with the generalized integral $\int_s^{+\infty} \widehat{u}(\tau)\,d\tau$, which might not converge absolutely. The statements (5.15) and (5.16) respectively follow from (5.14) and (5.13). [Ex]

Inversion of the Laplace Transform. Let us fix any $u \in D_{\mathcal{L}}$ and set

$$\varphi_x(y) := \int_{\mathbb{R}} e^{-(x+iy)t} u(t)\,dt = [\mathcal{L}(u)](x+iy) \overset{(5.8)}{=} \sqrt{2\pi}[\mathcal{F}(e^{-xt}u(t))](y) \qquad \forall x > \lambda_a(u), \forall y \in \mathbb{R}.$$

By Theorem 3.8, formally we then have

$$e^{-xt}u(t) = \frac{1}{\sqrt{2\pi}}\big[\mathcal{F}^{-1}(\{\varphi_x(y)\})\big](t) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{ity}\varphi_x(y)\,dy \qquad \forall t \in \mathbb{R}\,.$$

This yields the *Riemann-Fourier formula* for the inversion of the Laplace transform:

$$u(t) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{(x+iy)t}[\mathcal{L}(u)](x+iy)\,dy \qquad \forall t \in \mathbb{R}, \forall x > \lambda_a(u), \forall u \in D_{\mathcal{L}}, \tag{5.17}$$

which sometimes is also written as

$$
\begin{aligned}
u(t) &= \lim_{R\to+\infty} \frac{1}{2\pi i} \int_{x-iR}^{x+iR} e^{st}[\mathcal{L}(u)](s)\,ds \\
&=: \frac{1}{2\pi i} p.v. \int_{x-i\infty}^{x+i\infty} e^{st}[\mathcal{L}(u)](s)\,ds.
\end{aligned}
\tag{5.18}
$$

The above argument shows that this integral does not depend on the abscissa x.

Laplace Transform of Distributions. Next we extend this transform to a class of distributions. First we define:
the linear space of the **(Laplace) transformable distributions** [15]

$$D'_{\mathcal{L}} := \{T \in \mathcal{D}' : \operatorname{supp}(T) \subset \mathbb{R}^+, \exists x \in \mathbb{R} : \{e^{-xt}T(t)\} \in \mathcal{S}'\}, \tag{5.19}$$

the **convergence abscissa**

$$\lambda(T) := \inf\{x \in \mathbb{R} : \{e^{-xt}T(t)\} \in \mathcal{S}'\} \ (\in [-\infty, +\infty[) \quad \forall T \in D'_{\mathcal{L}}, \tag{5.20}$$

and define the **half-plane of convergence** $\mathbb{C}_{\boldsymbol{\lambda(T)}}$ as in (5.3). For any $T \in D'_{\mathcal{L}}$, we intend to give a meaning to a formula like (5.4), with the integral replaced by a duality pairing that we are tempted to write as

$$[\widetilde{\mathcal{L}}(T)](s) := \langle T(t), e^{-st}\rangle \qquad \forall s \in \mathbb{C}_{\lambda(T)}. \tag{5.21}$$

But $e^{-st} \notin \mathcal{D}$, so that this pairing is meaningless in the duality either between $\mathcal{D}'$ and $\mathcal{D}$, or between $\mathcal{S}'$ and $\mathcal{S}$. In order to provide a correct definition, we notice that the function $t \mapsto e^{-st}$ and all of its derivatives decay exponentially as $t \to +\infty$. (The behavior for $t \to -\infty$ is immaterial, as the support of T is confined to $\mathbb{R}^+$.) We then choose any

$$\zeta \in C^\infty \quad \text{such that} \quad \zeta = 1 \text{ in } \mathbb{R}^+, \quad \inf\operatorname{supp}(\zeta) = -1. \tag{5.22}$$

For any $T \in D'_{\mathcal{L}}$ and any $s \in \mathbb{C}_{\lambda(T)}$, let us select any $x \in \,]\lambda(T), \operatorname{Re}(s)[$. As $\{e^{-xt}T(t)\} \in \mathcal{S}'$ and $\{e^{-(s-x)t}\zeta(t)\} \in \mathcal{S}$, we may set

$$[\widetilde{\mathcal{L}}(T)](s) := {}_{\mathcal{S}'}\langle e^{-xt}T(t), e^{-(s-x)t}\zeta(t)\rangle_{\mathcal{S}} \qquad \forall s \in \mathbb{C}_{\lambda(T)}, \forall T \in D'_{\mathcal{L}}. \tag{5.23}$$

This value depends neither on $x \in \,]\lambda(T), \operatorname{Re}(s)[$ nor on ζ (as $\zeta = 1$ in $\operatorname{supp}(T)$).

Note that here we made no reference to absolute convergence: we replaced the Laplace integral (5.4) by a duality pairing, and in this case it has no meaning to speak of absolute convergence. The operator $\mathcal{L}$ extends the Laplace transform of functions as defined in (5.4). Henceforth we shall omit the tilde, and write $\mathcal{L}(T)$, or $\widehat{T}$, in place of $\widetilde{\mathcal{L}}(T)$.

Here are some basic examples:

$$
\begin{aligned}
&\widehat{\delta_0}(s) = 1, \quad \widehat{D^k\delta_0}(s) = s^k \qquad \forall s \in \mathbb{C}, \forall k \in \mathbb{N}, \\
&\widehat{\delta_a}(s) = e^{-as} \qquad \forall s \in \mathbb{C}_a, \forall a \in \mathbb{R}.
\end{aligned}
\tag{5.24}
$$

Any sequence $\{T_n\} \subset D'_{\mathcal{L}}$ is said to converge to $T \in D'_{\mathcal{L}}$ in the sense of $D'_{\mathcal{L}}$ iff

$$
\begin{cases}
\exists x \in \mathbb{R} : \quad x \ge \lambda(T), \quad x \ge \lambda(T_n) \ \forall n, \\
e^{-xt}T_n(t) \to e^{-xt}T(t) \quad \text{in } \mathcal{S}'.
\end{cases}
\tag{5.25}
$$

By (5.23), this entails that

$$\widehat{T_n}(s) \to \widehat{T}(s) \qquad \forall s \in \mathbb{C} : \operatorname{Re}(s) \ge \lambda(T)\,. \tag{5.26}$$

The next statement may be compared with Proposition 5.2.

[15] Of course, $D'_{\mathcal{L}} \neq (\mathcal{D}_{\mathcal{L}})'$ (the latter one has no meaning).

Proposition 5.6 *For any $T \in D'_{\mathcal{L}}$,*

(i) $\widehat{T}$ is analytic in $\mathbb{C}_{\lambda(T)}$,

(ii) $\widehat{T}$ has at most polynomial growth:

$$\exists M, m \in \mathbb{N} : \forall s \in \mathbb{C}_{\lambda(T)} \qquad |\widehat{T}| \leq M(1+|s|)^m. \tag{5.27}$$

Conversely any function that fulfills these properties is the Laplace transform of one and only one $T \in D'_{\mathcal{L}}$. []

At variance with what we saw in Proposition 5.2 for the Laplace transform of functions, for $T \in D'_{\mathcal{L}}$

$$\widehat{T}(s) \text{ need not vanish as } \mathrm{Re}(s) \to +\infty;$$

for instance we just pointed out that $\widehat{D^k\delta_0}(s) = s^k$ for any $s \in \mathbb{C}$ and $k \in \mathbb{N}$. As Proposition 5.1 carries over to distributions, this is consistent with the Fourier transform of distributions; e.g., $\widehat{\delta}_0 = 1$ identically in $\mathbb{C}$ (actually, $\lambda(\delta_0) = -\infty$), and $\mathcal{F}(\delta_0) = 1/\sqrt{2\pi}$ identically in $\mathbb{R}$. Proposition 5.3 and the first part of Proposition 5.4 also carry over to $D'_{\mathcal{L}}$, whereas the second part of Proposition 5.4 must be amended as follows.

Proposition 5.7 *For any $T \in \mathcal{D}'$,*

$$T \in D'_{\mathcal{L}} \qquad \Leftrightarrow \qquad DT \in D'_{\mathcal{L}}. \tag{5.28}$$

Moreover, if this condition is fulfilled, then

$$\lambda(DT) \leq \lambda(T), \qquad \widehat{DT}(s) = s\,\widehat{T}(s) \quad \forall s \in \mathbb{C}_{\lambda(T)}.[] \tag{5.29}$$

It may occur that the (distributional) derivative Du of a transformable function u is not transformable as a function, although by the latter statement u is necessarily transformable as a distribution. An example is provided by the function

$$\begin{aligned} &u(t) = \sin(e^{t^2} - 1)H(t) \in D_{\mathcal{L}} \quad \to \\ &Du(t) = 2te^{t^2}\cos(e^{t^2} - 1)H(t) \in D'_{\mathcal{L}} \setminus D_{\mathcal{L}}. \end{aligned} \tag{5.30}$$

Actually, the function u is bounded, hence Laplace transformable. On the other hand, for any $s \in \mathbb{C}$, the function $t \mapsto e^{-st}Du(t)$ fails to be absolutely integrable. [Ex] Nevertheless, by the latter proposition, if Du is regarded as a distribution then it is transformable.

A Glance at Operational Calculus. The Laplace transform may be used for the analysis of initial-value problems, like

$$\begin{aligned} &\sum_{n=0}^{m} a_n u^{(n)} = f(t) \qquad \forall t \in [0,T]\ (T > 0) \\ &u^{(n)}(0^+)\ \big(:= \lim_{t\to 0^+} u^{(n)}(t)\big) = u_n^0 \qquad \text{for } n = 0, ..., m-1, \end{aligned} \tag{5.31}$$

for any integer $m > 0$ and prescribed $a_0, ..., a_m$ (with $a_m \neq 0$), $u_0^0, ..., u_{m-1}^0 \in \mathbb{C}$, and $f \in C^0([0,T])$. Here we search for a solution $u \in C^m([0,T])$, and by $u^{(n)}$ we denote the nth classical derivative of u. The functions u and f are both defined just in the interval $[0,T]$; we extend them into Laplace-transformable functions (thus with support confined to $\mathbb{R}^+$), that we do not relabel.

We shall use the next result.

Proposition 5.8 *Let m be an integer > 0, and $u \in D'_{\mathcal{L}}$ be such that $u|_{\mathbb{R}^+} \in C^m(\mathbb{R}^+)$. Then*

$$
\begin{aligned}
&\widehat{u^{(n)}}(s) = s^n \widehat{u}(s) - \sum_{\ell=0}^{n-1} s^{n-1-\ell} u^{(\ell)}(0^+) \\
&\forall s > M := \max\{\lambda(u), ..., \lambda(u^{(n-1)})\}, \ \textit{for } n = 1, ..., m.
\end{aligned}
\tag{5.32}
$$

E.g., $\widehat{u'}(s) = s\widehat{u}(s) - u(0^+)$, $\widehat{u''}(s) = s^2\widehat{u}(s) - su'(0^+) - u(0^+)$.

Proof. We shall still denote by D the derivative in the sense of $\mathcal{D}'(\mathbb{R})$. Notice that, for any $w \in D'_{\mathcal{L}}$ such that $w|_{\mathbb{R}^+} \in C^1(\mathbb{R}^+)$,

$$Dw = w' + w(0^+)\delta_0 \qquad \text{in } \mathcal{D}'(\mathbb{R}).$$

For any $n \in \{1, ..., m\}$, selecting $w = u^{(n-1)}$ in the latter formula we have

$$Du^{(n-1)} = u^{(n)} + u^{(n-1)}(0^+)\delta_0 \qquad \text{in } \mathcal{D}'(\mathbb{R}).$$

Applying the Laplace transform (in the sense of distributions), we get

$$
\begin{aligned}
s\,\widehat{u^{(n-1)}}(s) &\overset{(5.29)}{=} \widehat{Du^{(n-1)}}(s) \\
&= \widehat{u^{(n)}}(s) + u^{(n-1)}(0^+)\,\widehat{\delta_0}(s) \quad \text{for } \mathrm{Re}(s) > M,
\end{aligned}
$$

that is, as $\widehat{\delta_0} = 1$,

$$\widehat{u^{(n)}}(s) = s\,\widehat{u^{(n-1)}}(s) - u^{(n-1)}(0^+) \quad \text{for } \mathrm{Re}(s) > M.$$

By applying this formula iteratively, (5.32) is easily derived. □

Let us come back to the problem (5.31), set

$$p(s) := \sum_{n=0}^{m} a_n s^n, \quad \Phi(s) := \sum_{n=0}^{m} a_n \sum_{\ell=0}^{n-1} s^{n-1-\ell} u^0_\ell,$$

and denote by $s_1, ..., s_m$ the (not necessarily distinct) roots of the polynomial $p(s)$. By Proposition 5.8, the initial-value problem (5.31) is then transformed into the single algebraic equation

$$p(s)\widehat{u}(s) = \Phi(s) + \widehat{f}(s) \quad \text{for } \mathrm{Re}(s) > \max\{\lambda_a(u), \lambda_a(f)\}. \tag{5.33}$$

Its solution reads

$$
\begin{aligned}
&\widehat{u}(s) = \frac{\Phi(s)}{p(s)} + \frac{\widehat{f}(s)}{p(s)} \\
&\text{for } \mathrm{Re}(s) > M := \max\Big\{\max_{i=1,...,m}\{\mathrm{Re}(s_i)\}, \lambda_a(f)\Big\}.
\end{aligned}
\tag{5.34}
$$

By applying the inverse Laplace transform to both members and (5.12), we then get

$$u = \mathcal{L}^{-1}\left(\frac{\Phi(s)}{p(s)}\right) + \mathcal{L}^{-1}\left(\frac{1}{p(s)}\right) * f \qquad \forall t \in [0, T]. \tag{5.35}$$

The first of these addenda only depends on the initial data and on the differential operator via the characteristic polynomial $p(s)$; this represents the *free answer* of the system that is represented by the equation (5.31). The second addendum depends on the forcing function f and on $p(s)$; this represents the *forced answer* of the system.

Notice that $\mathcal{L}^{-1}(1/p(s)) \in C^{m-1}\mathbb{R}^+$ is the solution of the problem (5.31 corresponding to $u_0 = ... = u^{m-1} = 0$ and $f = \delta_0$. This is called the *fundamental solution* of that problem.

6 P.D.E.s with Constant Coefficients

In this section we briefly illustrate the use of the Fourier transform in the analysis of P.D.E.s with constant coefficients set on the whole $\mathbb{R}^N$.

Differential Operators with Constant Coefficients. Any polynomial $P(\eta)$ of N complex variables of degree m is canonically associated to a linear differential operator $P(D)$ ($D := (\partial/\partial_1, ..., \partial/\partial_N)$) with constant coefficients of order m, and conversely:

$$P(\eta) := \sum_{|\alpha|\leq m} c_\alpha \eta^\alpha \quad \longleftrightarrow \quad P(D) := \sum_{|\alpha|\leq m} c_\alpha D^\alpha, \tag{6.1}$$

with $c_\alpha \in \mathbb{C}$ for any $\alpha \in \mathbb{N}^N$ and $\eta \in \mathbb{C}^N$. This establishes an isomorphism between the linear space of polynomials over $\mathbb{C}^N$ with complex coefficients and that of linear differential operators with constant complex coefficients. $P(\eta)$ ($\eta \in \mathbb{C}^N$) is called the **characteristic polynomial** of the differential operator $P(D)$. On the other hand, the polynomial $P(i\xi)$ ($\xi \in \mathbb{R}^N$) is called the *symbol* of $P(D)$, and is an element of $\mathcal{S}'$. Moreover, if $\varphi \in C^1$ is such that $P(iD\varphi(x)) = 0$ for any $x \in \mathbb{R}^N$, then the manifold $\{x \in \mathbb{R}^N : \varphi(x) = 0\}$ is called the *characteristic manifold* of the operator $P(D)$.

In this section we show how the Fourier transform exploits this isomorphism.

Proposition 6.1 *For any polynomial $P(\eta)$ with complex coefficients,*

$$P(D) = \mathcal{F}^{-1}[P(i\xi)\cdot\mathcal{F}] : \mathcal{S} \to \mathcal{S}, \tag{6.2}$$

$$P(D) = \mathcal{F}^{-1}[P(i\xi)\cdot\mathcal{F}T] : \mathcal{S}' \to \mathcal{S}'. \tag{6.3}$$

In particular

$$P(D)\delta_0 = (2\pi)^{-N/2}\mathcal{F}^{-1}[P(i\xi)] \ (\in \mathcal{S}'). \tag{6.4}$$

Proof. By Proposition 3.4 and by its extension to $\mathcal{S}'$, $\mathcal{F}[P(D)u] = P(i\xi)\mathcal{F}u$ for any $u \in \mathcal{S}$, and similarly for any $T \in \mathcal{S}$'. (6.2) and (6.3) are thus established. (6.4) follows by selecting $u = \delta_0$ in (6.3). □

Remarks. (i) By Proposition 4.2

$$P(D)u = \mathcal{F}^{-1}[P(i\xi)\cdot(\mathcal{F}u)(\xi)] = (2\pi)^{-N/2}\big(\mathcal{F}^{-1}[P(i\xi)]\big)*u \qquad \forall u \in \mathcal{S}. \tag{6.5}$$

(ii) May this equality be extended to $T \in \mathcal{S}'$? Actually, $[P(i\xi)] \in \mathcal{S}'$ so that $\mathcal{F}^{-1}[P(i\xi)] \in \mathcal{S}'$, but the convolution between arbitrary elements of $\mathcal{S}'$ need not exist. However $\mathcal{S}' * \mathcal{E}' \subset \mathcal{S}'$, [] so that

$$P(D)T = \mathcal{F}^{-1}[P(i\xi)\cdot(\mathcal{F}T)(\xi)] = (2\pi)^{-N/2}\big(\mathcal{F}^{-1}[P(i\xi)]\big)*T \qquad \forall T \in \mathcal{E}'. \tag{6.6}$$

Thus, although $P(D)$ operates in both spaces $\mathcal{S}$ and $\mathcal{S}'$, we are able to represent it in terms of convolution just in $\mathcal{S}$ and $\mathcal{E}'$.

(iii) Because of the symmetry of the representation of $\mathcal{F}$ and $\mathcal{F}^{-1}$, their role may be exchanged. In (6.4) and in the final equality of (6.5) and (6.6), $\mathcal{F}$ may thus be replaced by $\mathcal{F}^{-1}$; the same applies to other formulas of this section. □

Lemma 6.2 *If $P(\eta)$ is a polynomial with complex coefficients such that*

$$P(i\xi) \neq 0 \quad \forall \xi \in \mathbb{R}^N, \tag{6.7}$$

then [16]

$$v \mapsto P(i\xi)^{-1}v : \mathcal{S} \to \mathcal{S}, \mathcal{S}' \to \mathcal{S}'. \tag{6.8}$$

[16] Notice that $P(i\xi)^{-1} = 1/P(i\xi)$.

Proof. Notice that $|P(i\xi)| \to +\infty$ as $|\xi| \to +\infty$, and the same holds for all derivatives of $P(i\xi)$. Therefore

$$P(i\xi)^{-1} \in C^\infty \cap L^\infty, \tag{6.9}$$

and (6.8) follows. [Ex] □

The previous properties may be applied to the study of linear P.D.E.s with constant coefficients that are set on the whole $\mathbb{R}^N$. Let us fix a differential operator $P(D) \not\equiv 0$ (i.e., not vanishing identically), a function $f \in \mathcal{S}$, and consider the equation

$$u \in \mathcal{S}, \qquad P(D)u = f \qquad \text{in } \mathcal{S}. \tag{6.10}$$

By applying the Fourier transform to both sides of this differential equation, we see that it is equivalent to the algebraic equation

$$\widehat{u} \in \mathcal{S}, \qquad P(i\xi)\widehat{u}(\xi) = \widehat{f}(\xi) \qquad \text{in } \mathcal{S}. \tag{6.11}$$

If (6.7) is fulfilled, then by Lemma 6.2 the latter equation is equivalent to

$$\widehat{u}(\xi) = P(i\xi)^{-1}\widehat{f}(\xi) \qquad \text{in } \mathcal{S}. \tag{6.12}$$

Therefore, for any $f \in \mathcal{S}$, (6.10) is equivalent to

$$u = \mathcal{F}^{-1}[P(i\xi)^{-1}\widehat{f}(\xi)] \qquad \text{in } \mathcal{S}. \tag{6.13}$$

Let us now assume that $f \in \mathcal{S}'$ and consider the equation

$$T \in \mathcal{S}', \qquad P(D)T = f \qquad \text{in } \mathcal{S}'. \tag{6.14}$$

If (6.7) is fulfilled, this yields the analogous of (6.11) and (6.12) in $\mathcal{S}'$. By (6.9), $P(i\xi)^{-1}\widehat{f}(\xi) \in \mathcal{S}'$, so that

$$T = \mathcal{F}^{-1}[P(i\xi)^{-1}\widehat{f}(\xi)] \qquad \text{in } \mathcal{S}'. \tag{6.15}$$

We have thus proved the next statement.

Proposition 6.3 *Let P be a nonconstant polynomial of N complex variables. If $P(i\xi) \neq 0$ for any $\xi \in \mathbb{R}^N$, then*

$$P(D)^{-1} = (\mathcal{F}^{-1}[P(i\xi)^{-1}]\mathcal{F}) : \mathcal{S} \to \mathcal{S} \;\; \textit{and} \;\; \mathcal{S}' \to \mathcal{S}', \tag{6.16}$$

and both operators are continuous.

Remarks. (i) Still assuming that $P(i\xi) \neq 0$ for any $\xi \in \mathbb{R}^N$, by (6.9), $P(i\xi)^{-1} \in \mathcal{S}'$. By (6.5) then

$$\begin{aligned} P(D)^{-1}f &= (\mathcal{F}^{-1}[P(i\xi)^{-1}]\mathcal{F}(f)) \\ &= (2\pi)^{-N/2}(\mathcal{F}^{-1}[P(i\xi)^{-1}]) * f \qquad \forall f \in \mathcal{S}. \end{aligned} \tag{6.17}$$

(ii) As $P(i\xi)^{-1} \in \mathcal{S}'$ and $\mathcal{S}' * \mathcal{E}' \subset \mathcal{S}'$, the same formula holds for any $T \in \mathcal{E}'$. (However we are not able to extend it to $T \in \mathcal{S}'$.) □

For instance, [17]

$$P(\eta) = 1 - \sum_{j=1}^N \eta_j^2 \quad \longleftrightarrow \quad P(D) = I - \sum_{j=1}^N D_j^2 = I - \Delta \quad \text{in } \mathcal{S} \text{ and } \mathcal{S}'. \tag{6.18}$$

[17] In passing notice that $I - \Delta : v \mapsto v - \Delta v$ is a linear operator, but $1 - \Delta : v \mapsto 1 - \Delta v$ is not.

As $P(i\xi) = 1 + |\xi|^2 \neq 0$ for any $\xi \in \mathbb{R}^N$, by Proposition 6.3 we infer that

$$u = \mathcal{F}^{-1}[(1 + |\xi|^2)^{-1}\widehat{f}] = (2\pi)^{-N/2}\mathcal{F}^{-1}[(1 + |\xi|^2)^{-1}]*f \tag{6.19}$$

is the unique solution in $\mathcal{S}$ (in $\mathcal{S}'$, resp.) of the equation $u - \Delta u = f$, for any $f \in \mathcal{S}$ ($f \in \mathcal{E}'$, resp.).

A similar conclusion does not hold for $P(D) = -\Delta$; indeed in this case $P(i\xi) = |\xi|^2$, so that $P(i\xi) = 0$ for $\xi = 0$. Similarly, it fails for $P(D) = D_t - \Delta_x$ as, denoting by $\xi \in \mathbb{R}^{N+1}$ the conjugate variable of (x, t), $P(i\xi) = 0$ for $\xi = 0$.

Remark. So far we used the Fourier transform, and thus restricted ourselves to the spaces $\mathcal{S}$ and $\mathcal{S}'$. More caution is needed when dealing with these equations in $\mathcal{D}'$. For instance, Tychonov pointed out that the Cauchy problem

$$\begin{aligned} &D_t u - D_x^2 u = 0 \qquad &&\text{in } \mathbb{R}, \text{ for } t > 0\\ &u(\cdot, 0) = 0 &&\text{in } \mathbb{R} \end{aligned} \tag{6.20}$$

has a nontrivial solution $u \in \mathcal{D}' \setminus \mathcal{S}'$:

$$\begin{aligned} &u(x, t) = \sum_{n=1}^{\infty} g^{(n)}(t)\; x^{2n}/(2n)!\\ &\text{for} \quad g(t) = e^{-1/t^2} \quad \forall t > 0, \quad g(t) = 0 \quad \forall t \leq 0. \end{aligned} \tag{6.21}$$

□

As $\mathcal{E}' * \mathcal{D}' \subset \mathcal{D}'$,

$$P(D)u = P(D)[\delta_0 * u] = [P(D)\delta_0] * u \in \mathcal{D}' \qquad \forall u \in \mathcal{D}', \tag{6.22}$$

that is,

$$P(D) = [P(D)\delta_0] * \quad \text{as a continuous operator } \; \mathcal{D}' \to \mathcal{D}'. \tag{6.23}$$

Thus any differential operator (with constant coefficients) may be represented as the convolution with a distribution supported in $\{0\}$.

Pseudo-Differential Operators. The formulae (6.2) and (6.3) also read as identities between operators:

$$\begin{aligned} &P(D) = \mathcal{F}^{-1}[P(i\xi)\mathcal{F}] = (2\pi)^{-N/2}(\mathcal{F}^{-1}[P(i\xi)])*\\ &\mathcal{S} \to \mathcal{S} \;\text{ and }\; \mathcal{E}' \to \mathcal{S}'. \end{aligned} \tag{6.24}$$

More generally, in this way one may define the **pseudo-differential operator** $P(D)$ in $\mathcal{S}$ also for nonpolynomial functions P, provided that (at least) $P(i\xi)\mathcal{S}' \subset \mathcal{S}'$. For instance, one may define any positive power and certain negative powers of the Laplace operator $P(D) = -\sum_{j=1}^{N} D_j^2 = -\Delta$, which is associated to the polynomial $P(i\xi) = |\xi|^2$. For any $s > 0$, let us set

$$(-\Delta)^{s/2}u := \mathcal{F}^{-1}\big[|\xi|^s\mathcal{F}(u)\big] \in \mathcal{S}' \qquad \forall u \in \mathcal{S}, \forall s > 0. \tag{6.25}$$

As we already pointed out $\mathcal{S}' * \mathcal{E}' \subset \mathcal{S}'$; moreover $\mathcal{F}^{-1}(|\xi|^s) \in \mathcal{S}'$, since $|\xi|^s \in \mathcal{S}'$. Therefore

$$(-\Delta)^{s/2}u = (2\pi)^{N/2}\mathcal{F}^{-1}(|\xi|^s)*u \in \mathcal{S}' \qquad \forall u \in \mathcal{S}. \tag{6.26}$$

Still for any $s > 0$, one may also set

$$(-\Delta)^{s/2}T = (2\pi)^{N/2}\mathcal{F}^{-1}(|\xi|^s)*T \qquad \forall T \in \mathcal{E}'. \tag{6.27}$$

These definitions may be extended to any $s \in \,]-N,0[$, since in this case the function $\xi \mapsto |\xi|^s$ is an element of $L^1_{\rm loc} \cap \mathcal{S}'$. For any $s \in \,]-N,0[$, one may show that

$$\exists c_s > 0 : \mathcal{F}^{-1}(|\xi|^s) = c_s|x|^{-(s+N)} \ (\in \mathcal{S}' \cap L^1_{\rm loc}); \qquad [] \tag{6.28}$$

one may then also reformulate (6.27) as [18]

$$(-\Delta)^{s/2}u := \mathcal{F}^{-1}(|\xi|^s)*u \ \big(= c_s|x|^{-(s+N)}*u\big) \in \mathcal{S}' \qquad \forall u \in \mathcal{S}. \tag{6.29}$$

The latter formula also applies if u is replaced by any $T \in \mathcal{E}'$, as $\mathcal{S}' * \mathcal{E}' \subset \mathcal{S}'$.

In conclusion,

$$(-\Delta)^{s/2} : \mathcal{E}' \to \mathcal{S}' \qquad \forall s > -N; \tag{6.30}$$

one may also check that this operator is linear and continuous.

A similar statement applies to any power of the operator $I - \Delta$, as $(1 + |\xi|)^{s/2} \in \mathcal{S}'$ whence $\mathcal{F}^{-1}[(1 + |\xi|)^{s/2}] \in \mathcal{S}'$ for any $s \in \mathbb{R}$. [Ex] The (in general nonintegral) order of a pseudo-differential operator $P(D)$ is defined in terms of the asymptotic behavior of the function $P(i\xi)$ as $|\xi| \to +\infty$. For instance $(I - \Delta)^{s/2}$ has order s, for any $s \in \mathbb{R}$.

The Fundamental Theorem. An operator $P(D)$ is said to be **elliptic** iff its principal part (i.e., the sum of the terms of leading order) vanishes only at the origin of $\mathbb{C}^N$. This class includes e.g. the operator $(-\Delta)^m + \widetilde{P}(D)$ for any $m \in \mathbb{N}$ and any polynomial $\widetilde{P}$ of degree lower than m. Any elliptic operator is of even order. [Ex]

The next theorem is one of the great achievements of the theory of linear differential operators with constant coefficients.

Theorem 6.4 (Ehrenpreis-Malgrange-Hörmander)
Let P be a nonconstant polynomial of N complex variables, and Ω be a convex domain in $\mathbb{R}^N$. Then: [19]

$$P(D)\mathcal{E}(\Omega) = \mathcal{E}(\Omega), \tag{6.31}$$
$$P(D)\mathcal{D}'_F(\Omega) = \mathcal{D}'_F(\Omega), \tag{6.32}$$
$$P(D)\mathcal{D}'(\Omega) = \mathcal{D}'(\Omega) \quad \textit{(Ehrenpreis-Malgrange)}, \tag{6.33}$$
$$P(D)\mathcal{S}' = \mathcal{S}' \quad \textit{(Hörmander)}. \tag{6.34}$$

If $P(D)$ is elliptic, then (6.31) and (6.32) also hold for nonconvex domains. []

For any of these four equalities the inclusion "$\subset$" is trivial and holds for any domain Ω; the element of interest lies in the inclusion "$\supset$", which corresponds to the surjectivity of the operator, namely the existence of a solution of the P.D.E. (6.10) (or (6.14)).

Remarks. (i) The equalities (6.31) – (6.34) are of the form $P(D)X = X$, for various selections of the space X. This entails $X \subset P(D)^{-1}X$, but not $X = P(D)^{-1}X$, since $P(D)$ need not be injective. For instance:

— in general $P(D)^{-1}\mathcal{E} \not\subset \mathcal{E}$. E.g., if $P(D) = D_t^2 - D_x^2$, the wave equation $P(D)u = 0$ has solutions $u \notin \mathcal{E}$ (u may even be discontinuous!).

— in general $P(D)^{-1}\mathcal{S}' \not\subset \mathcal{S}'$. Actually, for certain operators the equation $P(D)u = 0$ has a nontrivial solution $u \in \mathcal{D}' \setminus \mathcal{S}'$. The Tychonov function (6.21) provides an example for the heat operator in a single space dimension.

[18] This holds also for $s \leq -N$, provided that the function $g(x) = |x|^{-(s+N)}$ ($\notin L^1_{\rm loc}$) is replaced by a *regularization*, namely by a distribution T whose restriction to $\mathbb{R}^N \setminus \{0\}$ coincides with g.

[19] By $\mathcal{D}'_F(\Omega)$ we denote the subspace of distributions of finite order.

(ii) Obviously $P(D)\mathcal{E}'(\Omega) \subset \mathcal{E}'(\Omega)$, but this inclusion is strict. For instance, for $N = 1$, $\Omega = \mathbb{R}$, $P(D) = D$ and $f = \chi_{]0,1[}$ (the characteristic function of the interval $]0, 1[$), the solutions of the equation $Du = \chi_{]0,1[}$ simply read $u(x) = \int_0^x \chi_{]0,1[}(s)\, ds$ + constant for any $x \in \mathbb{R}$, and have linear growth at infinity. Consistently with (6.32) and (6.34), they are elements of $\mathcal{D}'_F \cap \mathcal{S}'$, but not of $\mathcal{E}'$. Similarly, one may see that

$$\mathcal{D}(\Omega) \not\subset P(D)\mathcal{D}(\Omega), \quad \mathcal{S}(\Omega) \not\subset P(D)\mathcal{S}(\Omega). \quad [Ex] \tag{6.35}$$

E.g., no harmonic function has compact support. □

Fundamental Solutions. Let $P(\eta)$ be a polynomial of N complex variables of degree m, as in (6.1). A distribution $E \in \mathcal{D}'$ is called a **fundamental solution** of $P(D)$ iff

$$P(D)E = \delta_0 \qquad \text{in } \mathcal{D}'. \tag{6.36}$$

By (6.34), $P(D)$ has at least one fundamental solution in $E \in \mathcal{S}'$. For such a E we may apply the Fourier transform, getting $P(i\xi)\widehat{E} = 1$. If $P(i\xi)^{-1} \in L^1_{\rm loc}$ then clearly $E = P(i\xi)^{-1}$; otherwise the distribution $\widehat{E}$ is a *regularization* of the function $P(i\xi)^{-1}$. [20]

Proposition 6.5 *If P is a nonconstant polynomial of N complex variables, then no fundamental solution E of $P(D)$ has compact support.*

Proof. Let us assume that E has compact support. By the Paley-Wiener-Schwartz Theorem 4.6, $\widehat{E}$ is then an entire function. By applying the Fourier transform to the equation $P(D)E = \delta_0$, we get $P(i\xi)\widehat{E} = 1$ for any $\xi \in \mathbb{R}^N$, and this equality may be extended to any $\xi \in \mathbb{C}^N$. But $E = P(i\xi)^{-1}$ is an entire function only if P is a constant, because any other polynomial has roots in $\mathbb{C}^N$. □

Next we see that any nontrivial operator $P(D)$ has a fundamental solution in any bounded domain Ω, and that this allows one to construct a solution of the equation (6.14) for any $f \in \mathcal{E}'$.

Proposition 6.6 *Let P be a nonconstant polynomial of N complex variables. Then:*
(i) there exists a fundamental solution $E \in \mathcal{D}'_F \cap \mathcal{S}'$;
(ii) $E* : \mathcal{E}' \to \mathcal{S}'$, and

$$P(D)(E * T) = T \qquad \forall T \in \mathcal{E}'; \tag{6.37}$$

(iii) let Ω be a bounded domain of $\mathbb{R}^N$. For any $f \in L^1(\Omega)$, setting $f_0 := f$ in Ω and $f_0 := 0$ outside Ω, we have

$$P(D)[(E * f_0)|_\Omega] = f \qquad \text{in } \mathcal{D}'(\Omega), \forall f \in L^1(\Omega). \tag{6.38}$$

Proof. (i) As $\delta_0 \in \mathcal{D}'_F \cap \mathcal{S}'$, by (6.32) and (6.34) the equation $P(D)E = \delta_0$ has a solution $E \in \mathcal{D}'_F \cap \mathcal{S}'$.

(ii) As $\mathcal{S}' * \mathcal{E}' \subset \mathcal{S}'$ [] and $E \in \mathcal{S}'$, it follows that $E * f \in \mathcal{S}'$ for any $f \in \mathcal{E}'$. Moreover, by the differential properties of the convolution,

$$P(D)(E * T) = [P(D)E] * T = \delta_0 * T = T. \tag{6.39}$$

(iii) As $f_0 \in \mathcal{E}'$, by (6.37) we have $P(D)(E * f_0) = f_0$ in $\mathcal{E}'(\mathbb{R}^N)$. [21] (6.38) then follows. □

The next two results respectively deal with the existence and the uniqueness of the solution of the equation (6.14). Let us first set

$$\begin{aligned} &\varepsilon_\eta(x) := e^{i\eta\cdot x} \qquad \forall \eta \in \mathbb{C}^N, \forall x \in \mathbb{R}^N, \\ &\mathcal{Z}_P := \{\xi \in \mathbb{R}^N : P(i\xi) = 0\}, \\ &\Sigma_P := \text{ span of } \{\varepsilon_\xi : \xi \in \mathcal{Z}_P\} \ (\subset L^\infty). \end{aligned} \tag{6.40}$$

[20] As we saw in Section **??**, this is a distribution on $\mathbb{R}^N$ whose restriction to $\Omega := \{\xi \in \mathbb{R}^N : P(i\xi) \neq 0\}$ coincides with the function $P(i\xi)^{-1}$ (which is an element of $L^1_{\rm loc}(\Omega)$). Although the regularization is never unique, we shall see that the fundamental solution is unique iff $P(i\xi) \neq 0$ for any $\xi \in \mathbb{R}^N$.

[21] In other terms, $E*$ is a *right inverse* of the operator $P(D) : \mathcal{E}'(\mathbb{R}^N) \to \mathcal{E}'(\mathbb{R}^N)$. (It is also a *left inverse.)*

(Notice that $0 \in \mathcal{Z}_P$, by definition.) Let P be a nonconstant polynomial of N complex variables. By the fundamental theorem of algebra, the equation $P(i\eta) = 0$ has at least one root $\eta \in \mathbb{C}^N$. As $|\varepsilon_\eta(x)| = |e^{i\eta\cdot x}| = e^{-\mathrm{Im}(\eta)\cdot x}$, the distribution ε_η has exponential growth iff $\eta \notin \mathbb{R}^N$. Thus $\varepsilon_\eta \in \mathcal{S}'$ iff $\eta \in \mathbb{R}^N$; this corresponds to $\varepsilon_\eta \in L^\infty$. The next statement is then easily established.

Proposition 6.7 *Let P be a nonconstant polynomial of N complex variables. Then:*
(i) Σ_P coincides with the kernel of $P(D)$ in $\mathcal{S}'$. [Ex]
(ii) The set of all fundamental solutions of $P(D)$ in $\mathcal{S}'$ coincides with the affine space $E + \Sigma_P$, where E is any fundamental solution in $\mathcal{S}'$. [Ex]

Proof. For any $T \in \mathcal{S}'$, $P(D)T = 0$ iff $P(i\xi)\hat{T} = 0$, that is, the support of $\hat{T}$ is included in $\mathcal{Z}_P$. This means that $\hat{T}$ is in the span of $\{\delta_\xi : \xi \in \mathcal{Z}_P\}$, that is, $T \in \Sigma_P$. Part (i) is thus established. Part (ii) easily follows. □

Remarks. (i) As we pointed out, there exists $\eta \in \mathbb{C}^N$ such that $P(i\eta) = 0$. The kernel of $P(D)$ in $\mathcal{D}'$ thus includes ε_η, so that the fundamental solution is not unique in $\mathcal{D}'$.
(ii) However this kerned may not be reduced to the span of $\{\varepsilon_\eta : \eta \in \mathbb{C}, P(i\eta) = 0\}$. For the heat operator in a single space dimension, an example is provided by the Tychonov function (6.21). □

Although the fundamental solution is not unique in $\mathcal{D}'$, next we see that it may be unique in $\mathcal{S}'$.

Proposition 6.8 *Let P be a nonconstant polynomial of N complex variables. Then:*
(i) The restriction of the operator $P(D)$ to $\mathcal{S}'$ is injective (in particular, there is at most one fundamental solution in $\mathcal{S}'$) iff $P(i\xi) \neq 0$ for any $\xi \in \mathbb{R}^N$.
(ii) If $P(i\xi) \neq 0$ for any $\xi \in \mathbb{R}^N$, then $\mathcal{F}^{-1}(P(i\xi)^{-1}) \in \mathcal{S}'$, and this is the unique fundamental solution of $P(D)$ in $\mathcal{S}'$.

Proof. (i) Note that for any $v \in \mathcal{S}'$

$$P(D)v = 0 \quad \Leftrightarrow \quad P(i\xi)\widehat{v}(\xi) = 0 \quad \forall \xi \in \mathbb{R}^N \quad \Leftrightarrow \quad \widehat{v} \text{ is supported in } \mathcal{Z}_P.$$

Therefore $v = 0$ is the only solution $v \in \mathcal{S}'$ of the equation $P(D)v = 0$ (namely, $P(D)$ is injective in $\mathcal{S}'$) iff $\mathcal{Z}_P = \emptyset$, i.e., iff $P(i\xi) \neq 0$ for any $\xi \in \mathbb{R}^N$.
(ii) By Lemma 6.2 $P(i\xi)^{-1} \in \mathcal{S}' \cap L^\infty$, whence $\mathcal{F}^{-1}(P(i\xi)^{-1}) \in \mathcal{S}'$. This is a fundamental solution of $P(D)$ in $\mathcal{S}'$, and the uniqueness was established in part (i). □

Remark. If $P(i\xi) = 0$ for some $\xi \in \mathbb{R}^N$, then $P(D)$ has either no locally integrable fundamental solution or more than one. This depends upon the multiplicity of those roots and on N. For instance, for any $m \in \mathbb{N}$, the polynomial $P(i\xi) = |\xi|^{2m}$ is the symbol of the elliptic operator $P(D) = (-\sum_{j=1}^m D_j^2)^m = (-\Delta)^m$. As

$$P(i\xi) = |\xi|^{2m} = 0 \quad \Leftrightarrow \quad \xi = 0 \quad \text{(with algebraic multiplicity } 2m\text{)},$$

it is promptly checked that $P(i\xi)^{-1} \in \mathcal{S}' \cap L^1_{\mathrm{loc}}$ iff $2m < N$ (if $2m \geq N$ then $P(i\xi)^{-1}$ must be replaced by a regularization). Therefore:
— If $2m < N$ then $(-\Delta)^m$ has a fundamental solution $E \in \mathcal{S}' \cap L^1_{\mathrm{loc}}$. For any polynomial $Q(x)$ of degree $< 2m$, $\widetilde{E} = E + Q(x)$ ($\in \mathcal{S}' \cap L^1_{\mathrm{loc}}$) is also a fundamental solution (By the next proposition, there is no other fundamental solution in $\mathcal{S}'$).
— If $2m \geq N$ then $(-\Delta)^m$ has no fundamental solution in L^1_{loc}.
Thus, for instance, the operator $(-\Delta)^m$ has a locally integrable fundamental solution in $\mathbb{R}^3$ iff $m = 1$.

Proposition 6.9 *Let P be a nonconstant polynomial of N complex variables of degree m. If $P(i\xi) \neq 0$ for any $\xi \in \mathbb{R}^N \setminus \{0\}$, then any solution $u \in \mathcal{S}'$ of the equation $P(D)u = 0$ is a polynomial of degree at most m.*

Proof. The Fourier transform of the equation $P(D)u = 0$ reads $P(i\xi)\widehat{u} = 0$ in $\mathcal{S}'$. By the assumption on P, this entails that the support of the distribution $\widehat{u}$ is confined to $\{0\}$. By a classical result of the theory of distributions, u is then of the form $u(x) = \sum_{|\alpha|\le\ell} c_\alpha x^\alpha$, for some $\ell \in \mathbb{N}$ and suitable constants c_α. It is easily seen that $\ell \le m$. $\square$

This statement generalizes the classical Liouville theorem of complex analysis which states that any bounded entire holomorphic function is constant. The latter is indeed easily retrieved by recalling that the real and imaginary parts of any holomorphic function are harmonic, and then applying Proposition 6.9 to $P(D) = \Delta$. [Ex]

This theorem also applies e.g. to the heat operator $P(D) = D_t - \Delta_x$.

Fundamental Solutions of Cauchy Problems. Next we fix a linear differential operator $P(D)$ with constant coefficients, fix any $u^0 \in \mathcal{E}'(\mathbb{R}^N)$, and consider a Cauchy problem of the form

$$\begin{aligned} &D_t u + P(D_x)u = 0 \qquad \forall t > 0 \\ &u(\cdot,0) = u^0 \end{aligned} \tag{6.41}$$

(we still set $x = (x_1, ..., x_N)$, so that $(x,t) \in \mathbb{R}^{N+1}$). For the moment we do not specify either the regularity that we require for u, or the meaning of the initial condition.

We call a mapping $E : \mathbb{R}^+ \to \mathcal{D}'(\mathbb{R}^N)$ a **fundamental solution of the Cauchy problem** associated to the operator $D_t + P(D_x)$ iff

(i) $E : \mathbb{R}^+ \to \mathcal{D}'(\mathbb{R}^N)$ is differentiable, namely,

$$t \mapsto \langle E(t), \varphi\rangle \text{ is differentiable in } \mathbb{R}^+, \forall\varphi \in \mathcal{D}(\mathbb{R}^N), \tag{6.42}$$

(ii) denoting by $\widetilde{E}$ the distribution that is obtained by extending E with vanishing value for $t < 0$, $\widetilde{E}$ is a fundamental solution of the operator $D_t + P(D_x)$ in $\mathbb{R}^{N+1}$, that is

$$D_t\widetilde{E} + P(D_x)\widetilde{E} = \delta_{(x,t)=(0,0)} \; \big(= \delta_{x=0} \otimes \delta_{t=0}\big) \qquad \text{in } \mathcal{D}'(\mathbb{R}^{N+1}). \tag{6.43}$$

Analogously, the Cauchy problem (6.41) may be reformulated as the search for a mapping $u : \mathbb{R}^+ \to \mathcal{D}'(\mathbb{R}^N)$ such that, denoting by $\widetilde{u}$ the distribution that is obtained by extending E with vanishing value for $t < 0$,

$$D_t\widetilde{u} + P(D_x)\widetilde{u} = u^0 \otimes \delta_{t=0} \qquad \text{in } \mathcal{D}'(\mathbb{R}^{N+1}). \tag{6.44}$$

This equation is solved by the distribution

$$\widetilde{u}(x,t) = \widetilde{E} * (u^0 \otimes \delta_{t=0}) = \langle E(x-y,t), u^0(y)\rangle, \tag{6.45}$$

which indeed has support confined to $\mathbb{R}^N \times \mathbb{R}^+$. [Ex]

This definition of the fundamental solution of the Cauchy problem may easily be generalized to equations of higher order in time. [Ex]

Examples of Fundamental Solutions. The construction of a fundamental solution may not be trivial. Here we just indicate some simple examples.

(i) The Heaviside function (denoted by H) is a fundamental solution of the derivative D in $\mathbb{R}$. By Proposition 6.7, all fundamental solutions of D are of the form $H + c$, with $c \in \mathbb{R}$.

(ii) For any $C \in \mathbb{C}$, the function $E(x) = x(H(x) + C)$ is a fundamental solution of D^2 in $\mathbb{R}$. By Proposition 6.8 all fundamental solutions of D^2 are of the form $E(x) + ax + b$, for any $a, b \in \mathbb{C}$. [Ex]

(iii) The rotational invariance of the Laplace operator Δ suggests that the equation $\Delta E = \delta_0$ might have a radial solution $E(x) := \varphi(r)$ $(r := |x|)$. By representing the Laplace operator in radial coordinates, we have

$$\varphi''(r) + \frac{N-1}{r}\varphi'(r) = 0 \qquad \forall r > 0 \text{ (in } \mathbb{R}^N).$$

Prescribing the appropriate singular behaviour at the origin, a direct calculation shows that the following functions are fundamental solutions of the Laplace operator:

$$E(x) = \begin{cases} |x|/2 \qquad \forall x \in \mathbb{R}^N \setminus \{0\}, \ \text{if } N = 1, \\ (\log|x|)/(2\pi) \qquad \forall x \in \mathbb{R}^N \setminus \{0\}, \ \text{if } N = 2\,, \\ -|x|^{2-N}/\,[(N-2)\omega_N] \qquad \forall x \in \mathbb{R}^N \setminus \{0\}, \ \text{if } N > 2\,; \end{cases} \tag{6.46}$$

here ω_N is the $(N-1)$-dimensional measure of the unit sphere of $\mathbb{R}^N$. By Proposition 6.8 the fundamental solution of Δ is unique up to the sum of polynomials of degree not larger than one; that is, the kernel of this operator in $\mathcal{S}'$ consists of those polynomials.

(iv) The function

$$E(x,t) = (4\pi t)^{-N/2} \exp\left(-|x|^2/4t\right) \qquad \forall t > 0 \tag{6.47}$$

is a fundamental solution of the heat operator $D_t - \Delta_x$ in $\mathbb{R}^N \times \mathbb{R}^+$ in the above sense. By Proposition 6.8 the fundamental solution of this operator is unique up to an additive constant; that is, the kernel of $D_t - \Delta_x$ in $\mathcal{S}'$ consists of the constant functions. [22]

(v) Let us define the three functions

$$\begin{aligned} E_1(x,t) &= \frac{H(ct-|x|)}{2c} = \begin{cases} 0 & \text{if } |x| \geq ct \\ (2c)^{-1} & \text{if } |x| < ct \end{cases} \\ E_2(x,t) &= \frac{H(ct-|x|)}{2\pi\sqrt{(ct)^2-|x|^2}} = \begin{cases} 0 & \text{if } |x| \geq ct \\ \left(2\pi\sqrt{(ct)^2-|x|^2}\right)^{-1} & \text{if } |x| < ct \end{cases} \\ E_3(x,t) &= \frac{\delta_0(|x|-ct)}{4\pi ct} \end{aligned} \tag{6.48}$$

for any $t > 0$ and respectively for any $x \in \mathbb{R}^1$, $x \in \mathbb{R}^2$, $x \in \mathbb{R}^3$.

For $\ell = 1,2,3$, E_j is a fundamental solution of the Cauchy problem for the wave operator $D_t^2 - c^2\Delta$ in $\mathbb{R}^\ell \times \mathbb{R}^+$; By this we mean that, in analogy to what we saw above for first order equation, denoting by $\widetilde{E}$ the distribution that is obtained by extending E with vanishing value for $t < 0$,

$$L(D)\widetilde{E}_\ell = \delta_{(x,t)=(0,0)} \qquad \text{in } \mathcal{D}'(\mathbb{R}^{\ell+1}). \tag{6.49}$$

7 A Glance at Systems Theory (For Signal Processing)

In *signal analysis* [23] and other branches of engineering, some of the above results are used in the framework of the *theory of systems* with a language different from that of functional analysis — thus without referring to function spaces, and with *attenuated* mathematical rigor.

Filters. A system, that may represent a technical device, acting on time-dependent *signals* (i.e., functions of time) is represented by a *filter* (i.e., a linear operator) $L : u \mapsto f$. Here we just deal with linear single input – single output (SISO) systems. This filter may be represented, e.g., by a linear ODE: $P(t,D)u = f$, or inversely (and this is a typical situation) by the solution of a Cauchy problem for an ODE; other filters are represented by multiplication by a constant, by time-differentiation, by time-integration, by time-translation, and so on. An important class consists of the filters L that are translation-invariant (or *time-invariant)*; by this we mean that, setting $\rho_\tau u(t) := u(t-\tau)$ for any $t, \tau \in \mathbb{R}$,

$$L(\rho_\tau u) = \rho_\tau L u \qquad \text{for any admissible input } u \text{ and any } \tau \in \mathbb{R}. \tag{7.1}$$

[22] In passing note that the kernel of this operator in $\mathcal{D}'$ is larger; e.g., it contains the function e^{ax-a^2t} for any $a \in \mathbb{R}$.

[23] Here is a classical reference: A. Papoulis: Signal analysis. McGraw-Hill, New York 1977

For instance, if $L = P(t, D)$ this property is fulfilled iff the coefficients of P are constant (i.e., $L = P(D)$).

Convolution Filters. Henceforth we shall assume that δ_0 is an admissible input for the system, i.e., δ_0 is in the domain of L. Let us set

$$\bar{h}(t,\tau) := [L(\rho_\tau \delta_0)](t) = [L(\delta_\tau)](t) \qquad \forall t, \tau \in \mathbb{R}. \tag{7.2}$$

As $u = \langle \delta_\tau(\cdot), u(\tau) \rangle$ (that is, $u = \delta * u$), for any admissible input u we have

$$(Lu)(t) = [L(\delta_0 * u)](t) = [L\langle \delta_\tau(\cdot), u(\tau)\rangle](t) = \langle \bar{h}(t,\tau), u(\tau)\rangle. \tag{7.3}$$

In general the latter need not be a convolution. However, if he filter L is time-invariant, then setting $h = L\delta_0$ (7.2) reads

$$\bar{h}(t,\tau) = h(t-\tau) \qquad \forall t, \tau \in \mathbb{R}.$$

In this case the outcome of (7.3) is a convolution:

$$(Lu)(t) = [L(\delta_0 * u)](t) = \langle h(t-\cdot), u\rangle = [(L\delta_0) * u](t) \tag{7.4}$$

for any admissible input u (consistently with common practice in this theory, we do not specify the regularity properties). Thus $L = (L\delta_0)*$; accordingly, L is called a **convolution filter**.

Transfer Functions. By (7.3), the response of the system is determined by the response $L\delta_0$ to the unit impulse δ_0. In this section we shall denote by $\widehat{\ }$ the Laplace transform. $L\delta_0$ and $\widehat{L\delta_0}$ are respectively called the **transfer function in time** and the **transfer function in frequency** (or the **spectrum**) of the system L.

Henceforth we shall restrict ourselves to time-invariant filters, and shall restrict ourselves to Laplace-transformable signals (defined for a.e. t). In these systems signals may conveniently be represented either as functions of time or (via Laplace transform) as functions of frequency. For any admissible input u, we have

$$Lu = L(\delta_0 * u) = (L\delta_0) * u \quad \Leftrightarrow \quad \widehat{Lu} = \widehat{L\delta_0}\, \widehat{u}\,; \tag{7.5}$$

the first formula holds for a.e. time t, the second one for a.e. frequency ω (assuming that these are regular distributions). Hence

$$|\widehat{Lu}|^2 = |\widehat{L\delta_0}|^2\, |\widehat{u}|^2 \quad \text{ for a.e. frequency } \omega. \tag{7.6}$$

If $u \in L^2$, $\int |\widehat{u}(\omega)|^2\, d\omega = \int |u(t)|^2\, dt$ is interpreted as the power of the signal u; the mapping $|\widehat{u}|^2$ is accordingly called the **power spectrum** of the signal, and $|\widehat{L\delta_0}|^2$ is called the **transfer function of the energy** (both are functions of the frequency). The formula (7.6) may then be restated as follows:

> "the power spectrum of the response equals the product of the transfer function of the energy by the power spectrum of the input."

Filter Compositions. A filter L may be constructed e.g. by combining two filters L_1, L_2 in parallel (in series, resp.). This means that $L = L_1 + L_2$ ($L = L_1 \circ L_2$, resp.), and entails

$$L = [L_1\delta_0 + L_2\delta_0] * \quad \big(L = (L_1\delta_0) * (L_2\delta_0)*, \text{ resp.}\big) \quad \text{ in time}, \tag{7.7}$$

and, whenever L_1, L_2 are convolution filters,

$$\widehat{L\delta_0} = \widehat{L_1\delta_0} + \widehat{L_2\delta_0} \quad \big(\widehat{L\delta_0} = \widehat{L_1\delta_0} \cdot \widehat{L_2\delta_0}, \text{ resp.}\big) \quad \text{ in frequency}. \tag{7.8}$$

There are further ways to construct filters. For instance, a *feedback system* is represented by a mapping $f \mapsto g$ that is implictly defined by the equation

$$g = L_1(f + L_2(g)) \qquad \text{in time,} \tag{7.9}$$

L_1 and L_2 being linear filters. Let us assume that L_1 and L_2 are convolution filters, set $H_i := \widehat{L_i(\delta_0)}$ $(i = 1, 2)$, and also assume that $H_1 H_2 \neq 1$ for any frequency ω. By applying the Laplace transform to this equation we get

$$\widehat{g} = H_1\big(\widehat{f} + H_2\widehat{g}\big), \quad \text{i.e.} \quad \widehat{g}(\omega) = \frac{H_1(\omega)}{1 - H_1(\omega)H_2(\omega)}\widehat{f}(\omega) \qquad \forall \omega \in \mathbb{C}. \tag{7.10}$$

We conclude that the feedback system defines a convolution filter, and that its transfer function in frequency is $H = \frac{H_1}{1 - H_1 H_2}$.

Fundamental Relation. The next statement establishes a fundamental relation between a filter, its transfer function, and sinusoidal signals (i.e., exponential functions with imaginary exponent). Let us first set

$$\varepsilon_\omega(t) := e^{\omega t} \qquad \forall \omega \in \mathbb{C}, \forall t \in \mathbb{R}.$$

By the next statement, any time-invariant linear system acts on each frequency of the input independently of the other ones. In other terms, the system may modify the amplitude and shift the phase of any input ε_ω but it cannot modify its frequency.

Theorem 7.1 *Let Φ be a subspace of $\mathcal{S}'$ such that $\delta_0, \varepsilon_\omega \in \Phi$ for any $\omega \in \mathbb{C}$, and $L : \Phi \to \Phi$ be a time-invariant linear system. Any ε_ω is then an eigenfunction of L, and the spectrum $\widehat{L\delta_0}(\omega)$ is the corresponding eigenvalue. That is,*

$$L(\varepsilon_\omega) = \widehat{L\delta_0}(\omega)\, \varepsilon_\omega \quad \textit{in time, } \forall \omega \in \mathbb{C}. \tag{7.11}$$

Proof. For any $\omega \in \mathbb{C}$,

$$\begin{aligned} [L(\varepsilon_\omega)](t) = [L(\delta_0 * \varepsilon_\omega)](t) = [(L\delta_0) * \varepsilon_\omega](t) &= \int_{\mathbb{R}} [L\delta_0](\tau) e^{\omega(t-\tau)}\, d\tau \\ &= \int_{\mathbb{R}} [L\delta_0](\tau) e^{-\omega\tau}\, d\tau\; e^{\omega t} = \widehat{L\delta_0}(\omega)\, \varepsilon_\omega(t) \qquad \forall t \in \mathbb{R}. \end{aligned} \tag{7.12}$$

□

In several cases one may take $\Phi = \mathcal{S}'$.

Note that, in a linear time-invariant system, the input frequency has not changed, only the amplitude and the phase angle of the sinusoid has been changed by the system.

Differential Filters (ODEs). For these filters we retrieve some known results. After (6.11), if $L = P(D)$ then

$$\widehat{L\delta_0}(\xi) = P(i\xi)\, \widehat{\delta_0}(\xi) = P(i\xi). \tag{7.13}$$

Hence $L\delta_0 = \mathcal{F}^{-1}(P(i\xi))$, so that we retrieve (6.37):

$$Lu = (L\delta_0) * u = \mathcal{F}^{-1}(P(i\xi)) * u \quad \Leftrightarrow \quad \widehat{Lu} = \widehat{L\delta_0}\, \widehat{u} = P(i\xi)\widehat{u}. \tag{7.14}$$

Moreover, the transfer function of the energy of $L = P(D)$ is $|\widehat{L\delta_0}|^2 = |P(i\xi)|^2$:

$$|\widehat{Lu}|^2 = |\widehat{L\delta_0}|^2 |\widehat{u}|^2 = |P(i\xi)|^2\, |\widehat{u}|^2 \qquad \text{in frequency.} \tag{7.15}$$

For any $\xi \in \mathbb{R}$, ε_ω is an eigenfunction of any differential filter $L = P(D) = \sum_{n=0}^m c_n D^n$, and corresponds to the eigenvalue $\widehat{L\delta_0} \overset{(7.13)}{=} P(i\xi) = \sum_{n=0}^m c_n (i\xi)^n$. In this case (7.11) is reduced to the obvious formula

$$\sum_{n=0}^{m} c_n D^n e^{i\xi t} = \sum_{n=0}^{m} c_n (i\xi)^n\, e^{i\xi t}\,. \tag{7.16}$$