Analysis A - Course Notes on the Banach Fixed Point Theorem, Distributions, Fourier and Laplace transforms.

Albert J. Milani
in cooperation with Jörg Seiler

Dipartimento di Matematica
Università di Torino

Contents

Chapter 1

The Banach Fixed Point Theorem.

In this chapter we prove one version of the Banach fixed point theorem, and we report some applications of this result to the solution of algebraic and differential equations.

1.1 Contractions.

1. In this section we give the definition of contractions in metric spaces, and prove a fixed point theorem for strict contractions.

Definition 1.1.1 *1) A map φ from a metric space (M, d) into itself is said to be Lipschitz continuous if there is $L > 0$ such that*

$$\mathrm{d}(\varphi(x), \varphi(y)) \leq L\,\mathrm{d}(x, y) \tag{1.1.1}$$

for all x, $y \in M$. We set

$$L_\varphi := \inf\{L > 0 \ \mid \ \text{(1.1.1) holds}\}\,, \tag{1.1.2}$$

and call L_φ the Lipschitz constant of the map.
2) A Lipschitz continuous map from a metric space into itself is a contraction if $L_\varphi \in\,]0, 1]$; if $L_\varphi \in\,]0, 1[$, φ is called a strict contraction.
3) A point $z \in M$ is a fixed point of $\varphi : M \to M$ if

$$\varphi(z) = z\,. \tag{1.1.3}$$

2. We wish to prove the following version of the Banach fixed point theorem.

Theorem 1.1.1 *Let $(M; \mathrm{d})$ be a complete metric space, and $\varphi : M \to M$ be a strict contraction. Then, φ admits a unique fixed point, which can be found as the limit of the recursive sequence $(x_n)_{n\geq 0}$ defined by*

$$x_{n+1} = \varphi(x_n) , \tag{1.1.4}$$

starting from an arbitrary point $x_0 \in M$. Moreover, each x_n satisfies the estimates

$$\mathrm{d}(x_n, z) \leq \tfrac{L^n}{1-L} \mathrm{d}(x_1, x_0) , \tag{1.1.5}$$

$$\mathrm{d}(x_n, z) \leq \tfrac{L}{1-L} \mathrm{d}(x_n, x_{n-1}) \qquad (n \geq 1) . \tag{1.1.6}$$

Proof. 1) Uniqueness. If φ has two fixed points a and b, then

$$\mathrm{d}(a, b) = \mathrm{d}(\varphi(a), \varphi(b)) \leq L \,\mathrm{d}(a, b) , \tag{1.1.7}$$

from which

$$0 \leq (1 - L) \,\mathrm{d}(a, b) \leq 0 . \tag{1.1.8}$$

Since $L \neq 1$, (1.1.8) implies that $\mathrm{d}(a, b) = 0$, i.e. $a = b$.

2) Existence. By iteration, we find that for all $n \geq 0$,

$$\mathrm{d}(x_{n+1}, x_n) = \mathrm{d}(\varphi(x_n), \varphi(x_{n-1}) \leq L \,\mathrm{d}(x_n, x_{n-1}) \leq \cdots \leq L^n \,\mathrm{d}(x_1, x_0) . \tag{1.1.9}$$

Consequently, given any pair of indices m and n, with $m > n$,

$$\begin{aligned} \mathrm{d}(x_m, x_n) &\leq \mathrm{d}(x_m, x_{m-1}) + \mathrm{d}(x_{m-1}, x_{m-2}) + \cdots + \mathrm{d}(x_{n+1}, x_n) \\ &\leq \sum_{k=n}^{m-1} L^k \,\mathrm{d}(x_1, x_0) \leq \mathrm{d}(x_1, x_0)\, L^n \sum_{k=0}^{m-1} L^k \leq \frac{\mathrm{d}(x_1, x_0)}{1 - L} L^n . \end{aligned} \tag{1.1.10}$$

From this it follows that $(x_n)_{n\geq 0}$ is a Cauchy sequence in M. Since M is complete, there is $z \in M$ such that $x_n \to z$. Since φ is continuous, $\varphi(x_n) \to \varphi(z)$. Thus, we conclude from (1.1.4) that $z = \varphi(z)$; that is, z is the desired fixed point of φ.

3) Estimate (1.1.5) follows from (1.1.10), keeping n fixed and letting $m \to \infty$. Likewise, (1.1.10) also implies that

$$\mathrm{d}(x_m, x_n) \leq \sum_{k=0}^{m-n-1} \mathrm{d}(x_{n+1+k}, x_{n+k}) \leq L \,\mathrm{d}(x_n, x_{n-1}) \sum_{k=0}^{m-n-k} L^k . \tag{1.1.11}$$

Letting $m \to \infty$, (1.1.6) follows. $\square$

REMARKS. 1) In general, the fixed point z is not easy to determine exactly, so one is content to consider, instead of z, any of its approximations x_n; estimates (1.1.5) and (1.1.6) give information on the accuracy of this approximation, and on the speed of the convergence $x_n \to z$.

2) Theorem 1.1.1 gives sufficient conditions for the existence and uniqueness of the fixed point of a strict contraction. Yet, a map $\varphi : M \to M$ may have a fixed point, possibly not unique, even without being a contraction (see remark (3) below). For example, consider the equations

$$x = x\,, \qquad\qquad x = 1 - x \tag{1.1.12}$$

in $[0,1]$ [1]. Equations (1.1.12) are of the form (1.1.3), with $\varphi : [0,1] \to [0,1]$ defined by, respectively, $\varphi(x) = x$ and $\varphi(x) = 1 - x$. In both cases, φ is a contraction, with $L_\varphi = 1$, but not a strict contraction. For the first equation, all numbers in $[0,1]$ are fixed points, while the second equation has the unique fixed point $x = \frac{1}{2}$. Note, however, that this fixed point cannot be determined as the limit of the recursive sequence (1.1.4) (unless, of course, one starts from the fixed point itself; that is, if $x_0 = \frac{1}{2}$, in which case $x_n = \frac{1}{2}$ for all n). Indeed, in this case the sequence (1.1.4) is the union of the two subsequences $(x_{2n})_{n\ge 0}$ and $(x_{2n+1})_{n\ge 0}$, which are constant (more precisely, $x_{2n} = x_0$ and $x_{2n+1} = x_1 = 1 - x_0$ for each $n \in I\!N$; note that $x_{2n} = x_{2n+1} = \frac{1}{2}$ if $x_0 = \frac{1}{2}$).

3) The existence (but not necessarily the uniqueness) of solutions to equation (1.1.3) in an interval $[a,b]$, with φ continuous and $\varphi([a,b]) \subseteq [a,b]$ (i.e., with φ mapping $[a,b]$ into itself) is a consequence of the mean value theorem of continuous functions. Indeed, the assumption that φ maps $[a,b]$ into itself translates into the inequalities

$$a \le \min\{\varphi(a), \varphi(b)\} \le \max\{\varphi(a), \varphi(b)\} \le b\,; \tag{1.1.13}$$

letting $f(x) := x - \varphi(x)$, it follows that

$$f(a) = a - \varphi(a) \le 0\,, \qquad f(b) = b - \varphi(b) \ge 0\,. \tag{1.1.14}$$

Consequently, there is at least one $z \in [a,b]$ such that $f(z) = 0$, which is equivalent to (1.1.3). For example, consider $\varphi(x) = x^2$ in $[0,1]$: φ is not a contraction (take

[1] We consider intervals $[a,b] \subseteq I\!R$ as metric spaces with the distance induced by the standard euclidean distance of $I\!R$.

$x = 1 - \varepsilon$ and $y = 1 - 2\varepsilon$, with $\varepsilon \le \frac{1}{3}$), but equation (1.1.3) has the two solutions $x = 0$ and $x = 1$. Likewise, consider the case $\varphi(x) = 4\,x\,(1-x)$ in $[0,1]$. Then, φ maps $[0,1]$ into $[0,1]$, but is not a strict contraction on $[0,1]$. Still, the corresponding equation (1.1.3) has the solutions $x = 0$, $x = \frac{3}{4}$, and $x = 1$. $\diamond$

1.2 Algebraic Equations

1. The Banach fixed point theorem can be used to solve equations of the form (1.1.3), i.e.

$$x = g(x)\,, \tag{1.2.1}$$

as well as equations of the more general form

$$f(x) = 0\,; \tag{1.2.2}$$

indeed, the latter can be reduced to the format (1.2.1), by setting

$$x = x - \alpha\, f(x) =: g_\alpha(x)\,, \tag{1.2.3}$$

where $0 \neq \alpha \in I\!R$ is chosen so that g_α is a strict contraction on an interval $[a,b]$ in which one knows, e.g. by graphical methods, that (1.2.2) has a unique solution z of (1.2.2). We assume g and f to be of class C^1; then, the confirmation that $[a,b]$ contains a unique solution of (1.2.2) follows in two steps, usually consisting in first verifying the condition $f(a)\,f(b) < 0$, which, by the intermediate value theorem for continuous functions, yields the existence of a solution in $[a,b]$, and then invoking the monotonicity of f in $[a,b]$ to ensure that there are no other solutions of (1.2.2) in $[a,b]$. At this point, theorem 1.1.1 yields that the solution can be approximated by means of a sequence like (1.1.4).

2. We illustrate this procedure with a few examples; in practice, the challenge is to find suitable values for a, b, and α; the constant L of (1.1.1) will in general depend on these quantities.

Example 1. We know that the equation

$$x = \cos x \tag{1.2.4}$$

admits a unique solution $z \in \left[0, \frac{\pi}{2}\right]$ (the reader is encouraged to compare the graphs of the functions $x \mapsto \cos x$ and $x \mapsto x$ for $x \in \left[0, \frac{\pi}{2}\right]$). Equation (1.2.4) is already in the form (1.2.1), with $g(x) = \cos x$. This g is a contraction on $\left[0, \frac{\pi}{2}\right]$, albeit not necessarily a strict contraction. Indeed, by the mean value theorem for differentiable functions, given $x,\, y \in \left[0, \frac{\pi}{2}\right]$, there is $\theta \in \left[0, \frac{\pi}{2}\right]$ such that

$$\cos x - \cos y = (-\sin\theta)\,(x-y)\,; \tag{1.2.5}$$

but since we do not know the value of θ, the best estimate we can derive from (1.2.5) is

$$|\cos x - \cos y| \le |\sin\theta|\,|x-y| \le 1\,|x-y|\,, \tag{1.2.6}$$

which is (1.1.1) with $L = 1$ (for completeness sake, we prove explicitly that g is not a strict contraction on $\left[0, \frac{\pi}{2}\right]$ at the end of this section). To overcome this difficulty, we realize that, since neither $x = 0$ nor $x = \frac{\pi}{2}$ are solutions of (1.2.4), we can restrict our attention to a smaller interval $[\alpha, \beta]$, with $0 < \alpha = \cos\beta < 1 < \beta < \frac{\pi}{2}$ (so that g maps $[\alpha, \beta]$ into itself). Then, if $x,\, y \in [\alpha, \beta]$, also $\theta \in [\alpha, \beta]$ (because θ lies between x and y); thus,

$$|\sin\theta| \le \sin\beta =: L < 1\,. \tag{1.2.7}$$

In other words, g is a strict contraction on $[\alpha, \beta]$; thus, z can be determined as the limit of the sequence (1.1.4), which here reads

$$x_{n+1} = \cos x_n\,, \tag{1.2.8}$$

starting from an arbitrary $x_0 \in [\alpha, \beta]$. For example, taking $\beta = \frac{\pi}{3} \in \left]1, \frac{\pi}{2}\right[$, and starting from $x_0 = \frac{\pi}{4}$, we compute that:

$$\begin{aligned} x_1 &= \cos\left(\tfrac{\pi}{4}\right) = \tfrac{1}{\sqrt{2}}\,; \qquad & L = \sin\left(\tfrac{\pi}{3}\right) = \tfrac{\sqrt{3}}{2}\,; \\ \tfrac{1}{1-L} &= 2(2+\sqrt{3})\,; & \mathrm{d}(x_0, x_1) = \left|\tfrac{1}{\sqrt{2}} - \tfrac{\pi}{4}\right|\,; \end{aligned} \tag{1.2.9}$$

thus, if we want to find an approximate solution of (1.2.4) with an error not exceeding, e.g., $\varepsilon = 10^{-5}$, by (1.1.5) it is sufficient to consider x_n with n such that

$$\left(\tfrac{\sqrt{3}}{2}\right)^n \underbrace{2(2+\sqrt{3})\,\left|\tfrac{1}{\sqrt{2}} - \tfrac{\pi}{4}\right|}_{=:\,a} \le 10^{-5}\,; \tag{1.2.10}$$

that is,

$$n \ge \frac{\ln(a\,10^5)}{\ln(2/\sqrt{3})}\,. \tag{1.2.11}$$

To prove that $g(x) = \cos x$ is not a strict contraction on $\left[0 \frac{\pi}{2}\right]$, we proceed by contradiction. Thus, we assume that there is $L \in]0, 1[$ such that

$$|\cos x - \cos y| \le L\,|x-y| \tag{1.2.12}$$

for all x, $y \in \left[0, \frac{\pi}{2}\right]$. For small $\varepsilon > 0$, we choose $x = \frac{\pi}{2} - \varepsilon$ and $y = \frac{\pi}{2} - 2\varepsilon$. Then, (1.2.12) yields

$$\begin{aligned} 2\cos\left(\tfrac{3\varepsilon}{2}\right)\sin\left(\tfrac{\varepsilon}{2}\right) &= |\sin(\varepsilon) - \sin(2\varepsilon)| \\ &= \left|\cos\left(\tfrac{\pi}{2} - \varepsilon\right) - \cos\left(\tfrac{\pi}{2} - 2\varepsilon\right)\right| \leq L\,\varepsilon\,, \end{aligned} \tag{1.2.13}$$

from which

$$\cos\left(\tfrac{3\varepsilon}{2}\right)\frac{\sin\left(\frac{\varepsilon}{2}\right)}{\frac{\varepsilon}{2}} \leq L\,. \tag{1.2.14}$$

Letting then $\varepsilon \to 0$ we obtain the contradiction

$$1 \leq L < 1\,. \tag{1.2.15}$$

This confirms that g, while being a contraction on $\left[0, \frac{\pi}{2}\right]$, is not a strict one. □

Example 2. Consider the equation

$$f(x) = x^2 - \mathrm{e}^{-x} = 0 \tag{1.2.16}$$

in the interval $[0,1]$. Since $f(0) = -1 < 0$ and $f(1) = 1 - \frac{1}{\mathrm{e}} > 0$, there is at least one $z \in \,]0,1[$ such that $f(z) = 0$; since $f\,'(x) = 2x + \mathrm{e}^{-x} > 0$, such solution of (1.2.16) is unique. We also note that $f\,''(x) = 2 - \mathrm{e}^{-x} \geq 2 - 1 > 0$ in $[0,1]$; thus, $f\,'$ is monotone increasing in $[0,1]$, with

$$f\,'(0) = \tfrac{1}{\mathrm{e}} \leq f\,'(t) \leq f\,'(1) = 2 + \tfrac{1}{\mathrm{e}} \tag{1.2.17}$$

for all $t \in [0,1]$. For $\alpha > 0$, set

$$g_\alpha(x) := x - \alpha\, f(x) = x - \alpha\,(x^2 - \mathrm{e}^{-x})\,, \tag{1.2.18}$$

in accord with (1.2.3). Then,

$$g_\alpha(0) = \alpha > 0\,, \qquad g_\alpha(1) = 1 - \alpha\left(1 - \tfrac{1}{\mathrm{e}}\right) < 1\,, \tag{1.2.19}$$

and, by (1.2.17),

$$g_\alpha'(x) = 1 - \alpha\, f\,'(x) > 0 \qquad \text{if} \quad \alpha \leq \tfrac{1}{3}\,. \tag{1.2.20}$$

This implies that

$$0 < g_\alpha(0) \leq g_\alpha(x) \leq g_\alpha(1) < 1\,, \tag{1.2.21}$$

which means that g_α maps the interval $[0,1]$ into itself. Next, let $0 \le x < y \le 1$. By the mean value theorem for differentiable functions, there is $\theta \in [x,y]$ such that

$$g_\alpha(x) - g_\alpha(y) = (x-y) - \alpha\,(f(x) - f(y)) = (x-y)(1-\alpha\, f'(\theta))\,. \tag{1.2.22}$$

By (1.2.17), it follows that, if e.g. $\alpha \le \frac{1}{3}$,

$$\begin{aligned} 1-\alpha\, f'(\theta) &\le 1 - \tfrac{\alpha}{\mathrm{e}} =: L < 1\,, \\ 1-\alpha\, f'(\theta) &\ge 1 - \tfrac{1}{3}\left(2 + \tfrac{1}{\mathrm{e}}\right) > 0\,. \end{aligned} \tag{1.2.23}$$

In conclusion, we have found that, if $\alpha \le \frac{1}{3}$ and L is as in (1.2.23), if $0 \le x \le y \le 1$ (if $y \le x$, we just invert the roles of x and y),

$$|g_\alpha(x) - g_\alpha(y)| = g_\alpha(x) - g_\alpha(y) = (x-y)(1-\alpha\, f'(\theta)) \le L(x-y)\,. \tag{1.2.24}$$

Hence, g_α is a strict contraction on $[0,1]$, and the solution z of (1.2.16) can be found as the limit of the iterative sequence

$$x_{n+1} = g_\alpha(x_n) = x_n - \alpha(x_n^2 - \mathrm{e}^{-x_n})\,. \tag{1.2.25}$$

Note that letting $n \to \infty$ in (1.2.25) yields the identity

$$z = z - \alpha(z^2 - \mathrm{e}^{-z})\,, \tag{1.2.26}$$

so that, indeed, $z^2 = \mathrm{e}^{-z}$, as desired in (1.2.16). □

Example 3. We apply the procedure outlined in section (3) above to the second equation in (1.1.12), i.e.

$$x = 1 - x\,, \tag{1.2.27}$$

written as

$$f(x) := 1 - 2x = 0\,. \tag{1.2.28}$$

Choosing (e.g.) $\alpha = -\frac{1}{4}$, (1.2.3) reads

$$x = x + \tfrac{1}{4}\,(1-2x) = \tfrac{1}{2}\,x + \tfrac{1}{4} =: g(x)\,, \tag{1.2.29}$$

in which g maps $[0,1]$ into $[0,1]$, and is a strict contraction, with $L = \frac{1}{2}$. As expected, $g\left(\frac{1}{2}\right) = \frac{1}{2}$; that is, $\frac{1}{2}$ is the fixed point of g. □

Example 4. We consider the equation

$$x^2 - 5\,x + 6 = 0\,, \tag{1.2.30}$$

and find its solutions without using the quadratic formula. Admittedly, this is overkill; but we do this in order to further illustrate the use of the fixed point method. We write (1.2.30) in the form (1.2.2), as

$$x = \tfrac{1}{5}\,(x^2 + 6) =: g(x)\,. \tag{1.2.31}$$

This g maps the interval $\left[\frac{6}{5}, \frac{12}{5}\right]$ into itself, because

$$\tfrac{6}{5} < g\left(\tfrac{6}{5}\right) = \tfrac{186}{155} < g\left(\tfrac{12}{5}\right) = \tfrac{294}{125} < \tfrac{12}{5}\,, \tag{1.2.32}$$

and is a strict contraction on this interval, because

$$0 < \tfrac{2}{5}\,\tfrac{6}{5} < g'(x) = \tfrac{2}{5}\,x \leq \tfrac{2}{5}\,\tfrac{12}{5} = \tfrac{24}{25} < 1\,. \tag{1.2.33}$$

Thus, there is a unique solution $z \in \left[\frac{6}{5}, \frac{12}{5}\right]$ of equation (1.2.31), which can be obtained, e.g., as the limit of the sequence

$$x_0 = \tfrac{6}{5}\,, \qquad x_{n+1} = g(x_n) = \tfrac{1}{5}\,(x_n^2 + 6)\,. \tag{1.2.34}$$

we now show, by induction on n, that, for all $n \geq 0$,

$$x_0 < x_n < x_{n+1} < 2\,; \tag{1.2.35}$$

that is, the sequence $(x_n)_{n\geq 0}$ is strictly increasing, and bounded above. Indeed, for $n = 0$,

$$x_0 = \tfrac{6}{5} < 2\,, \qquad x_1 = g\left(\tfrac{6}{5}\right) > \tfrac{6}{5} = x_0\,, \tag{1.2.36}$$

having recalled the first inequality of (1.2.32). Since g is strictly increasing in $\left[\frac{6}{5}, \frac{12}{5}\right]$,

$$x_n < x_{n+1} \Longleftrightarrow g(x_{n-1}) < g(x_n) \Longleftrightarrow x_{n-1} < x_n\,, \tag{1.2.37}$$

which confirms the monotonicity of the sequence. Finally, if $x_n < 2$, from (1.2.34) we deduce that also

$$x_{n+1} = \tfrac{1}{5}\,(x_n^2 + 6) < \tfrac{1}{5}\,(4 + 6) = 2\,. \tag{1.2.38}$$

Since the sequence is bounded above by 2, it converges to its upper limit z, with $z \leq 2$. In fact, we see that $z = 2$; for, if $z < 2$, we would deduce the contradiction

$$0 < g(z) < g(2) = 0 \tag{1.2.39}$$

(recall that g is increasing in $\left[\frac{6}{5}, \frac{12}{5}\right]$). Thus, $z = 2$, which is a solution of (1.2.30), and the only solution in $\left[\frac{6}{5}, \frac{12}{5}\right]$. We leave it as an exercise to show that the same conclusion holds if the first term of the sequence $(x_n)_{n\geq 0}$ is chosen in the interval $\left]\frac{6}{5}, 2\right[$ (the argument is the same), or in the interval $\left]2, \frac{12}{5}\right]$ (in this case, the sequence $(x_n)_{n\geq 0}$ is decreasing).

We point out that writing (1.2.30) in the form (1.2.31) does not allow us to capture the other solution $z = 3$, even though $g(3) = 3$. This is because there is no interval $J = [3-\alpha, 3+\beta]$, with α and $\beta > 0$, such that g maps J into itself; that is, such that

$$3-\alpha \leq g(3-\alpha) < g(3+\beta) \leq 3+\beta\,. \tag{1.2.40}$$

Indeed, while the first inequality of (1.2.40) can be satisfied if $\alpha \geq 1$, the second is equivalent to

$$(3+\beta)^2 + 6 \leq 5\,(3+\beta)\,, \tag{1.2.41}$$

which cannot be satisfied for if $\beta > 0$. To capture the solution $z = 3$ of (1.2.30), we need to write it in the form (1.2.3); we leave this process as an exercise.

1.3 Applications to IVPs.

1. The Banach fixed point theorem can be successfully applied to determine local or global solutions to the Cauchy problem for nonlinear first order ODEs, of the form

$$\begin{cases} y' &= f(t,y)\,, \\ y(t_0) &= y_0\,. \end{cases} \tag{1.3.1}$$

Here, we assume that $f : [a,b] \times I\!R^N \to I\!R^N$ is a continuous function, satisfying a Lipschitz condition in y, uniform with respect to t; namely, that there is $C > 0$ such that, for all $t \in [a,b]$, and all y, $\tilde{y} \in I\!R^N$,

$$|f(t,y) - f(t,\tilde{y})| \;\leq\; C\,|y - \tilde{y}| \tag{1.3.2}$$

(that is, the Lipschitz constant C is independent of t). In these conditions, we set $X_1 := C^1([a,b] \to I\!R^N)$, and claim:

Theorem 1.3.1 [*Picard - Lindelöf.*] *For all $(t_0, y_0) \in [a,b] \times I\!R^N$, there exists a unique function $y \in X_1$, solution of the initial-value problem* (1.3.1).

Proof. 1) By the fundamental theorem of integral calculus, finding solutions of (1.3.1) in X_1 is equivalent to finding solutions to the integral equation

$$y(t) = y_0 + \int_{t_0}^{t} f(\theta, y(\theta))\, \mathrm{d}\theta \tag{1.3.3}$$

in the space $X_0 := C([a,b] \to I\!R^N)$. We endow this space with the metric defined by

$$\mathrm{d}(g,h) := \max_{a \le t \le b} \mathrm{e}^{-(C+1)|t-t_0|}\, |g(t) - h(t)|\,, \tag{1.3.4}$$

leaving it to the reader to verify that (1.3.4) does define a distance in X_0. Let $M = (X_0, \mathrm{d})$. We note that the exponential factor $\mathrm{e}^{-(C+1)|t-t_0|}$ in (1.3.4) is bounded from above and below, because the estimate $|t - t_0| \le b - a$ yields that

$$0 < \mathrm{e}^{-(C+1)(b-a)} \le \mathrm{e}^{-(C+1)|t-t_0|} \le 1\,. \tag{1.3.5}$$

Hence, convergence of a sequence of functions in X_0 with respect to the distance (1.3.4) is equivalent to the uniform convergence of the sequence in X_0 (the reader should verify this explicitly). It follows that M is a complete metric space.

2) We define $\varphi : X_0 \to X_0$ by

$$[\varphi(g)](t) := y_0 + \int_0^t f(\theta, g(\theta))\, \mathrm{d}\theta\,, \qquad t \in [a,b]\,, \tag{1.3.6}$$

and proceed to verify that φ is a strict contraction in X_0, with respect to the metric (1.3.4). To see this, we show that

$$\mathrm{d}(\phi(g), \phi(\tilde{g})) \le \tfrac{C}{C+1}\, \mathrm{d}(g, \tilde{g}) \tag{1.3.7}$$

for all g, $\tilde{g} \in M$ (that is, (1.1.1) holds, with $L = \frac{C}{C+1} < 1$). Let first $a < t_0 < t \le b$. Then, $|t - t_0| = t - t_0$, so we compute and estimate

$$\begin{aligned}
&\mathrm{e}^{-(C+1)(t-t_0)}\, |[\phi(g)](t) - [\phi(\tilde{g})](t)| \\
&\le\ \mathrm{e}^{-(C+1)(t-t_0)} \int_{t_0}^{t} |f(\theta, g(\theta)) - f(\theta, \tilde{g}(\theta))|\, \mathrm{d}\theta \\
&\le\ \mathrm{e}^{-(C+1)(t-t_0)} \int_{t_0}^{t} \mathrm{e}^{(C+1)(s-t_0)}\, \mathrm{e}^{-(C+1)(s-t_0)}\, C\, |g(\theta) - \tilde{g}(\theta)|\, \mathrm{d}\theta
\end{aligned}$$

$$\leq \; C\,\mathrm{e}^{-(C+1)(t-t_0)} \max_{a\leq s\leq b} \Big(\mathrm{e}^{-(C+1)|\theta-t_0|}\,|g(\theta)-\tilde{g}(\theta)|\Big)$$

$$\cdot \int_{t_0}^{t} \mathrm{e}^{(C+1)(\theta-t_0)}\,\mathrm{d}\theta \tag{1.3.8}$$

$$= \; C\,\mathrm{d}(g,\tilde{g})\,\mathrm{e}^{-(C+1)(t-t_0)} \int_{t_0}^{t} \mathrm{e}^{(C+1)(\theta-t_0)}\,\mathrm{d}\theta$$

$$\leq \; \tfrac{C}{C+1}\,\mathrm{d}(g,\tilde{g})\,\mathrm{e}^{-(C+1)(t-t_0)} \Big(\mathrm{e}^{(C+1)(t-t_0)}-1\Big)$$

$$= \; \tfrac{C}{C+1}\,\mathrm{d}(g,\tilde{g}) \Big(1-\mathrm{e}^{-(C+1)(t-t_0)}\Big)$$

$$\leq \; \tfrac{C}{C+1}\,\mathrm{d}(g,\tilde{g})\,.$$

We leave it to the reader to show that the same estimate holds if $a \leq t < t_0 \leq b$. In conclusion, we have found that

$$\mathrm{e}^{-(C+1)(t-t_0)}\,\Big|[\varphi(g)](t)-[\varphi(\tilde{g})](t)\Big| \leq \tfrac{C}{C+1}\,\mathrm{d}(g,\tilde{g}) \tag{1.3.9}$$

for all $t \in [a,b]$. The right side of (1.3.9) is a number independent of t; hence, also

$$\max_{a\leq t\leq b} \mathrm{e}^{-(C+1)(t-t_0)}\,\Big|[\varphi(g)](t)-[\varphi(\tilde{g})](t)\Big| \leq \tfrac{C}{C+1}\,\mathrm{d}(g,\tilde{g})\,, \tag{1.3.10}$$

which means exactly that (1.3.7) holds. It follows that φ is a strict contraction on M, as claimed.

3) By theorem 1.1.1, we conclude that there exists a unique $y \in X_0$, which is a fixed point of φ; that is, y is a solution of the integral equation (1.3.3). It follows that, in fact, $y \in X_1$, and is the desired solution of the initial-value problem (1.3.1). □

2: An example. Let $b > 0$, and define $f : [0,b] \times I\!R \to I\!R$ by $f(t,x) = x$. Then, f is Lipschitz continuous on $[0,b]$, uniformly in t, with Lipschitz constant $C = 1$. Consider the initial value problem

$$\left\{ \begin{array}{rcl} x' & = & x\,, \\ x(0) & = & 1 \end{array} \right. \tag{1.3.11}$$

(i.e., with $t_0 = 0$ and $x_0 = 1$). Thus, in this example,

$$[\phi(g)](t) = 1 + \int_0^t g(\theta)\,\mathrm{d}\theta\,, \tag{1.3.12}$$

ϕ being a strict contraction on $X_0 = C([0,b] \to \mathbb{R})$, with $L = \frac{C}{C+1} = \frac{1}{2}$. In accord with the Picard-Lindelöf theorem, the solution of the initial value problem (1.2.4) is given by the limit of the iterative sequence

$$x_{n+1} = \phi(x_n)\,; \tag{1.3.13}$$

that is,

$$x_{n+1}(t) = 1 + \int_0^t x_n(\theta)\,\mathrm{d}\theta\,. \tag{1.3.14}$$

Starting with $x_0(t) \equiv 1$ (consistently with the initial condition $x(0) = 1$), we compute:

$$x_1(t) = 1 + \int_0^t x_0(\theta)\,\mathrm{d}\theta = 1 + \int_0^t 1\,\mathrm{d}\theta = 1 + t\,; \tag{1.3.15}$$

$$x_2(t) = 1 + \int_0^t x_1(\theta)\,\mathrm{d}\theta = 1 + \int_0^t (1+\theta)\,\mathrm{d}\theta = 1 + t + \tfrac{1}{2}\,t^2\,; \tag{1.3.16}$$

$$x_3(t) = 1 + \int_0^t x_2(\theta)\,\mathrm{d}\theta = 1 + \int_0^t \left(1 + \theta + \tfrac{1}{2}\,\theta^2\right)\mathrm{d}\theta = 1 + t + \tfrac{1}{2}\,t^2 + \tfrac{1}{6}\,t^3\,; \tag{1.3.17}$$

in general,

$$x_n(t) = \;\cdots\; = 1 + t + \tfrac{1}{2}\,t^2 + \;\cdots\; + \tfrac{1}{n!}\,t^n = \sum_{k=0}^{n} \tfrac{1}{k!}\,t^k\,. \tag{1.3.18}$$

It is well-known that the sequence $(x_n)_{n\geq 0}$ converges, uniformly on $[0,b]$, to the function

$$x(t) = \sum_{k=0}^{\infty} \tfrac{1}{k!}\,t^k = \mathrm{e}^t\,, \tag{1.3.19}$$

which is indeed the solution of the initial value problem (1.3.11). ◇

Chapter 2

Basic Distribution Theory.

In this chapter we present some basic results in the theory of distributions, and their application to PDEs. The main ideas of the theory of distributions were investigated and developed by Laurent Schwartz in the late 1940s; apparently, he introduced the term "distribution" in analogy with the distribution of an electrical point charge in a region of space. A particular class of distributions includes those distributions that can be constructed from functions; this construction is based on considering integrals of functions, rather than the functions themselves. This is one of the reasons for the rather extensive study of the Lebesgue spaces L^p in Functional Analysis.

2.1 Motivations from PDEs.

Our goal in this section is to extend the meaning of solution to a PDE, introducing a generalized notion of derivative of a function, and to explain how a function having derivatives only in such generalized sense can be a solution to a PDE. The argument we follow to introduce the definition of generalized derivatives is quite standard, and consists in finding first some necessary conditions, expressed by means of integrals, that a sufficiently smooth function will satisfy, and then investigating if such conditions are also sufficient. When not, the conditions themselves will be taken as the basis of the definition of the new type of derivative. This is the same procedure we follow when introducing the definition of real numbers as equivalence classes of Cauchy sequences of rational numbers: For a sequence of rational numbers to converge, it is necessary that it be a Cauchy sequence, but the latter is sufficient only if the limit is a rational number. Otherwise, the Cauchy sequence itself is taken as the definition of (a representative of) a type of number, called a real number.

1. We start with an example from PDEs. The simplest model of the so-called transport equation is the linear equation

$$u_t + u_x = 0\,, \tag{2.1.1}$$

for $u = u(t,x),\ (t,x) \in \mathbb{R}^2$. If $u_0 \in C^1(\mathbb{R})$, the function

$$u(t,x) := u_0(x-t) \tag{2.1.2}$$

is a solution of (2.1.1), with $u \in C^1(\mathbb{R}^2)$. However, the definition (2.1.2) of u makes sense even if u_0 is only continuous, or even if $u_0 \in L^1_{\mathrm{loc}}(\mathbb{R})$ only. In these cases, we would still want to say that u is, in some sense we want to specify, a generalized solution to (2.1.1).

To this end, we consider the following necessary conditions. Assume that (2.1.1) does have a classical solution $u \in C^1(\mathbb{R}^2)$. Then, using integration by parts, we find that for all functions $\varphi \in C^1_0(\mathbb{R}^2)$

$$0 = \int_{\mathbb{R}^2} (u_t + u_x)\,\varphi\,\mathrm{d}A = -\int_{\mathbb{R}^2} u\,(\varphi_t + \varphi_x)\,\mathrm{d}A\,. \tag{2.1.3}$$

That is, if (2.1.1) has a C^1 solution, this solution must satisfy the infinite set of identities

$$\int_{\mathbb{R}^2} u\,(\varphi_t + \varphi_x)\,\mathrm{d}A = 0\,, \qquad \varphi \in C^1_0(\mathbb{R}^2)\,. \tag{2.1.4}$$

Conversely, this condition is also sufficient only if $u \in C^1(\mathbb{R}^2)$; that is, if $u \in C^1(\mathbb{R}^2)$ is such that (2.1.4) holds, then, using the fact that φ is arbitrary, we can deduce that u is a solution of (2.1.1). Note that the identities in (2.1.4) involve integrals, while the identity in (2.1.1) involves derivatives.

The key point in our argument is now to realize that each identity (2.1.4) makes sense even if $u \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ only. This motivates the following

Definition 2.1.1 *A function $u \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ is a generalized solution of* (2.1.1) *if* (2.1.4) *holds.*

Clearly, classical solutions of (2.1.1) are also generalized ones.

Proposition 2.1.1 *Let* $u_0 \in L^1_{\mathrm{loc}}(\mathbb{R})$*, and define* u *as in* (2.1.2)*. Then,* u *is a generalized solution of* (2.1.1)*. (Note that, if we only know that* $u_0 \in L^1_{\mathrm{loc}}(\mathbb{R})$*, then* u *need not be a classical solution to* (2.1.1)*).*

Proof. We first show that $u \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$. Let $K \subset \mathbb{R}^2$ be compact, and $R > 0$ be such that $K \subseteq [-R, R] \times [-R, R] =: K_R$. Then,

$$\begin{aligned} \int_K |u(t,x)| \, \mathrm{d}A &= \int_K |u_0(x-t)| \, \mathrm{d}A \\ &\leq \int_{K_R} |u_0(x-t)| \, \mathrm{d}A \\ &= \int_{-R}^{R} \int_{-R-t}^{R-t} |u_0(r)| \, \mathrm{d}r \, \mathrm{d}t \qquad (2.1.5) \\ &\leq \int_{-R}^{R} \int_{-2R}^{2R} |u_0(r)| \, \mathrm{d}r \, \mathrm{d}t \\ &= 2R \int_{-2R}^{2R} |u_0(r)| \, \mathrm{d}r \;\leq\; 2R \, |u_0|_1 \, . \end{aligned}$$

This confirms that $u \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$. To show that u satisfies (2.1.4), fix an arbitrary $\varphi \in C^1_0(\mathbb{R}^2)$, let $K = \operatorname{supp} \varphi$, and, as above, let $R > 0$ be such that $K \subseteq [-R, R] \times [-R, R] =: K_R$. Since $C^\infty_0(]-R, R[)$ is dense in $L^1(]-R, R[)$, there exists a sequence $(u^k)_{k \in \mathbb{N}} \subset C^\infty_0(]-R, R[)$, such that

$$\int_{-R}^{R} |u^k(x) - u_0(x)| \, \mathrm{d}x \to 0 \, . \qquad (2.1.6)$$

Let then $\psi := \varphi_t + \varphi_x$, and

$$\psi_0 := \max\{ \, |\psi(t,x)| \;\mid\; (t,x) \in K \, \} \, . \qquad (2.1.7)$$

Proceeding as in (2.1.5), we estimate

$$\int_{\mathbb{R}^2} |u^k(x-t) - u_0(x-t)| \, |\psi(t,x)| \, \mathrm{d}A$$

$$= \int_K |u^k(x-t) - u_0(x-t)| \, |\psi(t,x)| \, \mathrm{d}A$$

$$\leq \psi_0 \int_{K_R} |u^k(x-t) - u_0(x-t)| \, \mathrm{d}A \tag{2.1.8}$$

$$\leq 2R\,\psi_0 \int_{-R}^{R} |u^k(r) - u_0(r)| \, \mathrm{d}r \, .$$

By (2.1.6), this last term tends to 0 as $k \to \infty$. Recalling the definition of ψ, this implies that

$$\int_{I\!R^2} u_0(x-t)\,(\varphi_t(t,x) + \varphi_x(t,x))\ \mathrm{d}A$$

$$= \lim_{k\to+\infty} \int_{I\!R^2} u^k(x-t)\,(\varphi_t(t,x) + \varphi_x(t,x))\ \mathrm{d}A \tag{2.1.9}$$

$$= -\lim_{k\to+\infty} \int_{I\!R^2} \Big(u^k_t(x-t) + u^k_x(x-t)\Big)\,\varphi(t,x)\,\mathrm{d}A \quad = 0 \, .$$

Consequently, (2.1.4) holds, and u is indeed a generalized solution of (2.1.1). □

2. The procedure of this example is rather typical; that is, in order to find a generalized solution u to a PDE, one generally looks for approximate, smooth solutions u^k, and then hopes that these converge to a possibly non smooth function, which would then be recognized as a solution in the desired generalized sense. Of course, the choice of the most convenient topology to control the convergence (as in the L^1 norm of (2.1.6)) is extremely important; in general one wants to consider the weakest possible topology. In particular, one often deals with sequences of approximating solutions which converge only weakly.

Exercise 2.1.1 *Let u_0, $u_1 \in L^1_{\mathrm{loc}}(I\!R)$. Show that d'Alembert's formula*

$$u(t,x) := \tfrac{1}{2}\left\{u_0(x+t) + u_0(x-t) + \int_{x-t}^{x+t} u_1(r)\,\mathrm{d}r\right\} \tag{2.1.10}$$

defines a generalized solution to the linear wave equation

$$u_{tt} - u_{xx} = 0 \, , \tag{2.1.11}$$

with initial values

$$u(0,x) = u_0(x)\,, \qquad u_t(0,x) = u_1(x)\,, \tag{2.1.12}$$

for almost all $x \in I\!R$.

2.2 Towards Distributions.

Our goal in this section is to introduce a generalization of the notion of the first order derivative of a function $f \in L^1_{\mathrm{loc}}(I\!R)$, which is not assumed to be differentiable in the classical sense. We will call this new derivative the distributional derivative of f. To this end, we consider the space $C^1_0(I\!R)$ consisting of those real-valued functions of a real variable, which are at least once differentiable, with continuous first order derivative, and whose support is compact in $I\!R$. We recall that the support of a continuous function $\varphi : I\!R \to I\!R$ is the closure of the set on which φ does not vanish; that is,

$$\mathrm{supp}\,(\varphi) := \overline{\{x \in I\!R \;\mid\; \varphi(x) \neq 0\}}\,. \tag{2.2.1}$$

For example, if

$$\varphi(x) = \begin{cases} \exp\left(\frac{-1}{1-|x|^2}\right) & \text{if} \quad |x| < 1\,, \\ 0 & \qquad\text{otherwise}\,, \end{cases} \tag{2.2.2}$$

the support of φ is the interval $[-1,1]$. Note that if $\varphi \in C^1_0(I\!R)$, the support of φ' is also compact; in fact,

$$\mathrm{supp}\,(\varphi') \subseteq \mathrm{supp}\,(\varphi)\,. \tag{2.2.3}$$

We show this in the simple situation when $\mathrm{supp}\,(\varphi') = [c,d]$ and $\mathrm{supp}\,(\varphi) = [a,b]$ (for the general case, see proposition 2.4.1 below). Arguing by contradiction, assume e.g. that $b < d$ (the other cases are proven similarly). Let $z \in \,]b,d[$. Then, z is in the interior of the support of φ', so that $\varphi'(z) \neq 0$. On the other hand, consider an open neighborhood I of z such that $I \subset \,]b,d[$. Then $I \cap [a,b] = \emptyset$; consequently, $\varphi(x) = 0$ for all $x \in I$, and this implies that $\varphi'(z) = 0$.

1. In this introductory section, we provisionally define a distribution on $I\!R$ to be a linear map T from $C^1_0(I\!R)$ into $I\!R$, which is sequentially continuous, in the sense that if a sequence $(\varphi_n)_{n\geq 0} \subset C^1_0(I\!R)$ and a function $\varphi \in C^1_0(I\!R)$ are such that φ and all the φ_n vanish outside a compact set $K \subset I\!R$, and $\varphi_m \to \varphi$ and ${\varphi'}_n \to \varphi'$ uniformly on K, then

$$T(\varphi_n) \;\to\; T(\varphi) \qquad (\text{in} \quad I\!R)\,. \tag{2.2.4}$$

(This is not the precise definition of sequential continuity for distributions, which will be given in part (2) of section 2.4, but it will do for the present introduction.)

Given $f \in L^1_{\mathrm{loc}}(I\!\!R)$, we can define two linear maps T_f, $\tilde{T}_f : C^1_0(I\!\!R) \to I\!\!R$, by

$$T_f(\varphi) \;:=\; \int_{-\infty}^{+\infty} f(x)\varphi(x)\,\mathrm{d}x\,, \tag{2.2.5}$$

$$\tilde{T}_f(\varphi) \;:=\; -\int_{-\infty}^{+\infty} f(x)\varphi'(x)\,\mathrm{d}x\,, \tag{2.2.6}$$

for all $\varphi \in C^1_0(I\!\!R)$, and these maps are sequentially continuous. Hence, T_f and $\tilde{T}_f$ are distributions, in the provisional sense we have adopted. Note that

$$\tilde{T}_f(\varphi) \;=\; -T_f(\varphi')\,. \tag{2.2.7}$$

We call T_f the distribution generated by f, and $\tilde{T}_f$ the distributional derivative of T_f. This terminology is motivated by the following remark. Let $f \in C^1(I\!\!R)$. Then, both f and $f' \in L^1_{\mathrm{loc}}(I\!\!R)$, so that these functions generate the distributions T_f and $T_{f'}$, via the first of (2.2.5). Integrating by parts, we see that, for all $\varphi \in C^1_0(I\!\!R)$,

$$T_{f'}(\varphi) = \int_{-\infty}^{+\infty} f'(x)\varphi(x)\,\mathrm{d}x = -\int_{-\infty}^{+\infty} f(x)\varphi'(x)\,\mathrm{d}x = \tilde{T}_f(\varphi)\,; \tag{2.2.8}$$

that is, if $f \in C^1(I\!\!R)$, then

$$T_{f'} = \tilde{T}_f\,, \tag{2.2.9}$$

which means that the distributional derivative of the distribution generated by f (i.e., the right side of (2.2.9)), coincides with the distribution generated by the classical derivative f' (i.e., the left side of (2.2.9)). In addition, by (2.2.7),

$$T_{f'}(\varphi) \;=\; -T_f(\varphi')\,, \tag{2.2.10}$$

for all $\varphi \in C^1_0(I\!\!R)$. The key point in this argument is that the distribution $\tilde{T}_f$ can be defined even if $f \in L^1_{\mathrm{loc}}(I\!\!R)$ only; that is, the right side of (2.2.10) still makes sense, while the left side in general does not. Thus, with the same logic that led to definition 2.1.1, we identify the function $f \in L^1_{\mathrm{loc}}(I\!\!R)$ with the map T_f defined in (2.2.5), and call $\tilde{T}_f$ the distributional derivative of f. It is important to realize that the distributional derivative of an integrable function is a distribution, and not a function. On the other hand, given $f \in L^1_{\mathrm{loc}}(I\!\!R)$, it may or it may not happen that there is another $g \in L^1_{\rm loc}(I\!\!R)$ such that $\tilde{T}_f = T_g$; that is, by (2.2.6) and (2.2.5), such that the identities

$$-\int_{-\infty}^{+\infty} f(x)\,\varphi'(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} g(x)\varphi(x)\,\mathrm{d}x \tag{2.2.11}$$

hold, for all $\varphi \in C^1_0(I\!\!R)$. For example, this is the case when $f \in C^1(I\!\!R)$, with $g = f'$. We call g the generalized derivative of f; thus, the generalized derivative of an integrable function is a function.

2. We summarize the notions introduced so far in

Definition 2.2.1 *Let $f \in L^1_{\rm loc}(I\!\!R)$.*
1) The linear, sequentially continuous map $T_f : C^1_0(I\!\!R) \to I\!\!R$ defined by (2.2.5) *(i.e., the distribution T_f) is called the distribution defined by f.*
2) The linear, sequentially continuous map $\tilde{T}_f : C^1_0(I\!\!R) \to I\!\!R$ defined by (2.2.6) *is another distribution, called the distributional derivative of f. To emphasize the fact that $\tilde{T}_f$ is a derivative, we change notations and denote this map by $f'_{\rm d}$; that is, we set $f'_{\rm d} := \tilde{T}_f$, and rewrite* (2.2.6) *as*

$$f'_{\rm d}(\varphi) := -\int_{-\infty}^{+\infty} f(x)\varphi'(x)\,\mathrm{d}x = -\,T_f(\varphi')\,. \tag{2.2.12}$$

3) If there is $g \in L^1_{\rm loc}(I\!\!R)$ such that $f'_{\rm d} = T_g$, i.e. (2.2.11) *holds, we call g the generalized derivative of f, and set $g =: f'_{\rm g}$; in this case,* (2.2.7) *reads*

$$f'_{\rm d}(\varphi) = T_g(\varphi) = -\,T_f(\varphi')\,. \tag{2.2.13}$$

To repeat in a slightly different way:

1) A locally integrable function f defines, via (2.2.5), a distribution T_f, which is a linear, sequentially continuous map from $C^1_0(I\!\!R)$ into $I\!\!R$;
2) The distribution T_f defines another distribution $f'_{\rm d}$, via (2.2.12);
3) We identify f with T_f, and say that $f'_{\rm d}$ is the distributional derivative of f;
4) If there is another locally integrable function g such that the two distributions $f'_{\rm d}$ and T_g coincide, we say that g is the generalized derivative of f, and write $g =: f'_{\rm g}$.
5) The distributional derivative of f is a distribution, while its generalized derivative is a function.

It may be worth to note explicitly that, while the independent variable for the functions f and f' (if the latter exists) is a real number, the independent variable for the maps T_f and f'_{d} is a function in $C_0^1(\mathbb{R})$. The motivation of definition 2.2.1 is similar to the one that led to the definition 2.1.1 of a generalized solution of the PDE (2.1.1): (2.2.10) is a necessary condition that the derivative f' of a function $f \in C^1(\mathbb{R})$ does satisfy, and that the distributional derivative of f should satisfy. Therefore, it is natural to take the right side of (2.2.10) as the *definition* of such new kind of derivative of f (it is not difficult to prove that the distributional derivative of f is uniquely determined by f). It is also worth to keep in mind that classical differentiation is a local concept, while the definition (2.2.12) of distributional derivative is a global one, as it involves integrals.

3. A function $f \in L^1_{\mathrm{loc}}(\mathbb{R})$ can therefore have two types of derivatives: the distributional and the generalized one. However, we should really be careful about the fact that the former is in fact the derivative of the *map* T_f. The advantage of considering distributional derivatives should be obvious, because while a function $f \in L^1_{\mathrm{loc}}(\mathbb{R})$ may fail to have a generalized derivative, it (or, rather, the corresponding map T_f) always has a distributional derivative, which is the distribution f'_{d}.

A natural question is whether definition 2.2.1 is consistent, in the sense that if f has a locally integrable classical derivative f', then f' is the generalized derivative of f. This means that the maps f'_{d} and $T_{f'}$ should coincide on $C_0^1(\mathbb{R})$; that is, by (2.2.9) and (2.2.10),

$$f'_{\mathrm{d}}(\varphi) = T_{f'}(\varphi) \tag{2.2.14}$$

for all $\varphi \in C_0^1(\mathbb{R})$. In turn, this means that identity (2.2.8) (which is true if $f \in C^1(\mathbb{R})$) should hold. This is not always the case: as shown in Rudin, [7], § 6.14, this is indeed the case if, and in fact only if, f is absolutely continuous (in which case f is differentiable almost everywhere, with derivative $f' \in L^1(\mathbb{R})$).

4. We now consider a few examples of distributional and generalized derivatives.

4.1. We start with the computation of the distributional derivative of the Heaviside functions.

Example 2.2.1 *Let $\alpha \in \mathbb{R}$, and define $H_\alpha : \mathbb{R} \setminus \{\alpha\} \to \mathbb{R}$ by*

$$H_\alpha(x) := \begin{cases} 0 & \text{if} \quad x < \alpha\,, \\ 1 & \text{if} \quad x > \alpha\,. \end{cases} \tag{2.2.15}$$

The function H_α is called the Heaviside function centered at α. Clearly, $H_\alpha \in L^1_{\mathrm{loc}}(\mathbb{R})$, so we can consider the linear maps T_{H_α} defined by (2.2.5). We wish to *compute the distributional derivative of H_α; that is, in the notation of* (2.2.12), *the distribution* $(H_\alpha)'_{\mathrm{d}}$.

Proposition 2.2.1 *Given $\alpha \in I\!R$, define a map $\delta_\alpha : C^1_0(I\!R) \to I\!R$ by*

$$\delta_\alpha(\varphi) := \varphi(\alpha)\,. \tag{2.2.16}$$

Then, δ_α is linear and sequentially continuous, hence a distribution on $I\!R$, and

$$(H_\alpha)'_{\mathrm{d}} = \delta_\alpha\,. \tag{2.2.17}$$

Proof. 1) The linearity of δ_α is obvious. As for its sequential continuity, we have to show that δ_α transforms converging sequences $\varphi_m \to \varphi$ in $C^1_0(I\!R)$ into converging sequences in $I\!R$. Thus, we fix a compact set $K \subset I\!R$ such that φ and all the φ_m vanish outside K, and $\varphi_m \to \varphi$ uniformly on K. Then, if $\alpha \in K$,

$$\delta_\alpha(\varphi_m) = \varphi_m(\alpha) \to \varphi(\alpha) = \delta_\alpha(\varphi)\,; \tag{2.2.18}$$

the same is obviously true if $\alpha \notin K$, since in this case

$$\delta_\alpha(\varphi_m) = 0 = \delta_\alpha(\varphi)\,. \tag{2.2.19}$$

Recalling (2.2.4), this means that δ_α is sequentially continuous.

2) To prove (2.2.17), we must show that, for all $\varphi \in C^1_0(I\!R)$,

$$(H_\alpha)'_{\mathrm{d}}(\varphi) = \delta_\alpha(\varphi)\,. \tag{2.2.20}$$

Recalling (2.2.6) of definition 2.2.1, this follows from

$$\begin{aligned}(H_\alpha)'_{\mathrm{d}}(\varphi) &= -\,T_{H_\alpha}(\varphi') = -\int_{-\infty}^{+\infty} H_\alpha(x)\varphi'(x)\,\mathrm{d}x \\ &= -\int_{-\infty}^{\alpha} 0\,\varphi'(x)\,\mathrm{d}x - \int_{\alpha}^{+\infty} 1\,\varphi'(x)\,\mathrm{d}x \\ &= -\int_{\alpha}^{+\infty} \varphi'(x)\,\mathrm{d}x = \varphi(\alpha) = \delta_\alpha(\varphi)\,.\end{aligned} \tag{2.2.21}$$

The distribution δ_α is called the DIRAC mass centered at α. □

4.2. More generally [1],

[1] Theorem 1.14 and example 1.15 of Seiler, [8], with reversed roles of v and w in the theorem.

Theorem 2.2.1 *Let $x_0 \in I\!R$, and $v \in C^1(]-\infty, x_0])$, $w \in C^1([x_0, +\infty[)$. Define u and $\tilde{u} : I\!R \setminus \{x_0\} \to I\!R$ by*

$$u(x) := \begin{cases} v(x) & \text{if } x < x_0 \,, \\ w(x) & \text{if } x > x_0 \,, \end{cases} \qquad \tilde{u}(x) := \begin{cases} v\,'(x) & \text{if } \ x < x_0 \,, \\ w\,'(x) & \text{if } \ x > x_0 \,. \end{cases} \tag{2.2.22}$$

Then, u and $\tilde{u} \in L^1_{\mathrm{loc}}(I\!R)$; $\tilde{u} = u\,'$ separately on $]-\infty, x_0[$ and $]x_0, +\infty[$, and

$$(T_u)\,' = T_{\tilde{u}} - (v(x_0) - w(x_0))\,\delta_{x_0} \,. \tag{2.2.23}$$

Proof. The claims on u, $\tilde{u}$, and $u\,'$ are clear. To show (2.2.23), let $\varphi \in C^1_0(I\!R)$. Recalling (2.2.12), we compute

$$\begin{aligned} [(T_u)\,'](\varphi) &= [-T_u](\varphi\,') = -\int_{-\infty}^{+\infty} u(x)\,\varphi\,'(x)\,\mathrm{d}x \\ &= -\int_{-\infty}^{x_0} v(x)\,\varphi\,'(x)\,\mathrm{d}x - \int_{x_0}^{+\infty} w(x)\,\varphi\,'(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{x_0} v\,'(x)\,\varphi(x)\,\mathrm{d}x - v(x_0)\,\varphi(x_0) \\ &\quad + \int_{x_0}^{+\infty} w\,'(x)\,\varphi(x)\,\mathrm{d}x + w(x_0)\,\varphi(x_0) \\ &= \int_{-\infty}^{+\infty} \tilde{u}(x)\,\varphi(x)\,\mathrm{d}x - (v(x_0) - w(x_0))\,\varphi(x_0) \\ &= T_{\tilde{u}}(\varphi) - (v(x_0) - w(x_0))\,\delta_{x_0}(\varphi) \\ &= [T_{\tilde{u}} - (v(x_0) - w(x_0))\,\delta_{x_0}]\,(\varphi)\,. \end{aligned} \tag{2.2.24}$$

This proves (2.2.23). Note that the number

$$-\,(v(x_0) - w(x_0)) = u(x_0^+) - u(x_0^-) \tag{2.2.25}$$

measures the height of the jump of u at the discontinuity x_0. □

For example, consider the Heaviside function $u = H_0$, with (e.g.) $v(x) = 0$ for $x \leq 0$ and $w(x) = 1$ for $x \geq 0$. Then, $\tilde{u}(x) = 0$, and (2.2.23) reads

$$(H_u)\,' = 0 + (w(0) - v(0))\,\delta_0 = (1 - 0)\,\delta_0 = \delta_0 \,, \tag{2.2.26}$$

as we know from (2.2.17) with $\alpha = 0$.

Example 2.2.2 *Let $u(x) := |1 - x^2|$. Since u is continuous on $I\!R$, $(T_u)' = T_{u'}$, where u' is the classical derivative of u, defined in $I\!R \setminus \{\mp 1\}$; that is,*

$$u'(x) = \begin{cases} -2x & \text{if} \quad |x| < 1\,, \\ +2x & \text{if} \quad |x| > 1\,. \end{cases} \tag{2.2.27}$$

We wish to use (2.2.22)*, with u replaced by u', to compute $(T_{u'})'$. To this end, we set $\tilde{u}_1 := \widetilde{u'}$, that is,*

$$\tilde{u}_1(x) := \begin{cases} -2 & \text{if} \quad |x| < 1\,, \\ +2 & \text{if} \quad |x| > 1\,; \end{cases} \tag{2.2.28}$$

from which it follows that:

$$\begin{cases} at \quad x_0 = -1\,, & v(x) = +2x\,, \quad w(x) = -2x\,, \\ at \quad x_0 = +1\,, & v(x) = -2x\,, \quad w(x) = +2x\,. \end{cases} \tag{2.2.29}$$

Consequently,

$$\begin{aligned} (T_{u'})' &= T_{\tilde{u}_1} - (2(-1) - (-2)(-1))\,\delta_{-1} - ((-2)1 - 2(1))\,\delta_1 \\ &= T_{\tilde{u}_1} + 4\,\delta_{-1} + 4\,\delta_1\,, \end{aligned} \tag{2.2.30}$$

which records the fact that u' has two jumps, at $x = -1$ and $x = 1$, each of height 4.

4.3. The next example illustrates the relationship between the classical derivative of a piecewise differentiable function and its distributional derivative.

Proposition 2.2.2 *Define f and g by*

$$f(x) := \begin{cases} 0 & \text{if} \quad x \le 0\,, \\ x & \text{if} \quad 0 < x \le 1\,, \\ 1 & \text{if} \quad 1 < x\,, \end{cases} \qquad g(x) := \begin{cases} 0 & \text{if} \quad x \le 0\,, \\ 1 & \text{if} \quad 0 < x \le 1\,, \\ 0 & \text{if} \quad 1 < x\,. \end{cases} \tag{2.2.31}$$

Then, f and $g \in L^1_{\text{loc}}(I\!R)$, $g = f'$ in $I\!R \setminus \{0, 1\}$, and $T_g = f'_{\text{g}}$; that is, g is the generalized derivative of f.

Proof. It is sufficient to verify (2.2.6), which in this case reads

$$-\int_{-\infty}^{+\infty} f(x)\varphi'(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} g(x)\,\varphi(x)\,\mathrm{d}x = \int_0^1 \varphi(x)\,\mathrm{d}x \tag{2.2.32}$$

for all $\varphi \in C_0^1(I\!R)$. Thus, fix $\varphi \in C_0^1(I\!R)$, and choose a, $b \in I\!R$ such that $\operatorname{supp}\varphi \subseteq [a,b]$. Since $g = f'$ on $]-\infty, 0[$, $]0,1[$, and $]1,+\infty[$, we see, by classical integration by parts, that (2.2.32) does hold, if $a < b < 0$ or $1 < a < b$ (both sides of (2.2.32) equal 0). Otherwise: if $0 < b \le 1$,

$$\begin{aligned}
-\int_{-\infty}^{+\infty} f(x)\varphi'(x)\,\mathrm{d}x &= -\int_0^b x\varphi'(x)\,\mathrm{d}x \\
&= -\left[x\varphi(x)\right]_0^b + \int_0^b \varphi(x)\,\mathrm{d}x \\
&= -b\,\varphi(b) + \int_0^b \varphi(x)\,\mathrm{d}x \\
&= \int_0^b \varphi(x)\,\mathrm{d}x \quad = \int_0^1 \varphi(x)\,\mathrm{d}x\,,
\end{aligned} \tag{2.2.33}$$

in accord with (2.2.32). If instead $b > 1$, with a similar computation we find

$$\begin{aligned}
-\int_{-\infty}^{+\infty} f(x)\varphi'(x)\,\mathrm{d}x &= -\int_0^1 x\varphi'(x)\,\mathrm{d}x - \int_1^b \varphi'(x)\,\mathrm{d}x \\
&= -\varphi(1) + \int_0^1 \varphi(x)\,\mathrm{d}x - \varphi(b) + \varphi(1) = \int_0^1 \varphi(x)\,\mathrm{d}x\,,
\end{aligned} \tag{2.2.34}$$

again in accord with (2.2.32). This confirms that $T_g = f'_{\mathrm{g}}$. □

4.4. From this example it would seem that the distributional derivative f'_{d} of a piecewise differentiable function f, with derivative g defined almost everywhere, would coincide with the distribution T_g (that is, the distribution defined by the classical derivative of f, wherever this is defined). But this is not necessarily the case, as example 2.2.1 shows. Indeed, each H_α is piecewise differentiable, with classical derivative $h_\alpha \equiv 0$ separately in $]-\infty, \alpha[$ or in $]\alpha, +\infty[$. Now, $h_\alpha \in L^1_{\mathrm{loc}}(I\!R)$, but T_{h_α} is *not* the distributional derivative of T_{H_α}. Indeed, for $\varphi \in C_0^1(I\!R)$,

$$T_{h_\alpha}(\varphi) = \int_{-\infty}^{+\infty} h_\alpha(x)\,\varphi(x)\,\mathrm{d}x = 0\,, \tag{2.2.35}$$

in contrast with (2.2.17), which yields that, if α is in the interior of the support of φ,

$$(H_\alpha)'_{\mathrm{d}}(\varphi) = \delta_\alpha(\varphi) = \varphi(\alpha) \neq 0\,. \tag{2.2.36}$$

4.5. A natural question is whether H_α, extended to all of $I\!R$ by choosing an arbitrary value for $H_\alpha(\alpha)$, possesses a generalized derivative. We know that the answer to this question is negative, because H_α is not absolutely continuous on $I\!R$ (as it is not even continuous) [2].

Proposition 2.2.3 *Let H be the Heaviside function H_0 of proposition 2.2.1. There is no $h \in L^1_{\mathrm{loc}}(I\!R)$ such that $T_h = H'_{\mathrm{d}}$.*

Proof. Proceeding by contradiction, assume such h exists. Then, as in (2.2.21), for all $\varphi \in C^1_0(I\!R)$

$$\begin{aligned} T_h(\varphi) &= \int_{-\infty}^{+\infty} h(x)\,\varphi(x)\,\mathrm{d}x = H'_{\mathrm{d}}(\varphi) \\ &= -\int_{-\infty}^{+\infty} H(x)\,\varphi'(x)\,\mathrm{d}x = \varphi(0)\,. \end{aligned} \tag{2.2.37}$$

Take now a sequence $(\varphi_k)_{k\in I\!N} \subset C^1_0(I\!R)$ such that for each $k \in I\!N$

$$|\,\varphi_k(x)| \le 1 \quad \forall\, x \in I\!R\,, \qquad \varphi_k(0) = 1\,, \qquad \mathrm{supp}(\varphi_k) = \left[-\tfrac{1}{k}, \tfrac{1}{k}\right]. \tag{2.2.38}$$

From (2.2.37) we deduce that for each k,

$$\begin{aligned} 1 = \varphi_k(0) &= \int_{-\infty}^{+\infty} h(x)\varphi_k(x)\,\mathrm{d}x \le \int_{-\infty}^{+\infty} |h(x)|\,|\varphi_k(x)|\,\mathrm{d}x \\ &\le \int_{-1/k}^{1/k} |h(x)|\,\mathrm{d}x =: \lambda_k\,. \end{aligned} \tag{2.2.39}$$

But $\lambda_k \to 0$ by the absolute continuity of the Lebesgue integral, so we reach a contradiction. □

4.6. We conclude with an example which gives another illustration of how the lack of generalized derivatives, due to jump discontinuities in a function, can be overcome with the introduction of distributional derivatives. This example also shows that linearity holds for differentiation in the distributional sense as well.

[2] The moral of the examples of propositions 2.2.1 and 2.2.2 is that jumps are "bad" for generalized derivatives, while points where a function is continuous, but not differentiable, behave somewhat better.

Proposition 2.2.4 *Let f and g be as in proposition 2.2.2, and define $\tilde{f} \in L^1_{\mathrm{loc}}(I\!R)$ by*

$$\tilde{f}(x) := \begin{cases} 0 & \text{if} \quad x \leq 0\,, \\ x & \text{if} \quad 0 < x \leq 1\,, \\ 2 & \text{if} \quad 1 < x\,, \end{cases} \tag{2.2.40}$$

Then, g (which is the generalized derivative of f) is not the generalized derivative of $\tilde{f}$. On the other hand,

$$\tilde{f}\,'_{\mathrm{d}} = f\,'_{\mathrm{d}} + \delta_1\,, \tag{2.2.41}$$

in accord with (2.2.42) *below.*

Proof. Since $\delta_1 = (H_1)'_{\mathrm{d}}$, (2.2.41) is a consequence of the identity

$$\tilde{f}(x) = f(x) + H_1(x) \qquad \text{for} \;\; x \neq 1\,, \tag{2.2.42}$$

with H_1 defined as in example 2.2.1. If g were the generalized derivative of $\tilde{f}$, then, by (2.2.41), and since, as we know, $T_g = f\,'_{\mathrm{d}}$ as well,

$$T_g = \tilde{f}\,'_{\mathrm{d}} = f\,'_{\mathrm{d}} + \delta_1 = T_g + \delta_1\,. \tag{2.2.43}$$

However, δ_1 is not the zero distribution (as we see by applying it to a $\varphi \in C^1_0(I\!R)$ with $\varphi(1) \neq 0$). In other words, g cannot be the generalized derivative of $\tilde{f}$, since it does not record the contribution of the jump of $\tilde{f}$ at $x = 1$. □

5. In conclusion, we mention that, traditionally, an abuse of notations is tolerated, whereby the distribution T_f constructed from a function $f \in L^1_{\mathrm{loc}}(I\!R)$ via (2.2.5) is still denoted by f; distributions of this kind are called regular, and, when there is no risk of confusion, their distributional derivative is denoted by $T\,'$. In particular, propositions 2.2.1 and 2.2.1 implies that the distributions $\delta_\alpha = H\,'_\alpha$, distributional derivatives of (the distributions defined by) the Heaviside functions H_α, are not regular distributions.

2.3 Further Remarks on Dirac's δ Distribution.

2.3.1 The Restrictions of δ to $I\!R_{<0}$ and $I\!R_{>0}$.

We recall that, on $I\!R$, the Dirac δ distribution is the distribution $\delta \in \mathcal{D}'(I\!R)$ defined by

$$\langle \delta, \varphi \rangle_{\mathcal{D}} = \varphi(0), \qquad \forall\, \varphi \in \mathcal{D}(I\!R), \tag{2.3.1}$$

and that δ is the distributional derivative of the regular distribution T_H defined by the locally integrable Heaviside function

$$x \mapsto H(x) = \begin{cases} 0 & \text{if} \quad x < 0\,, \\ 1 & \text{if} \quad x > 0\,. \end{cases} \tag{2.3.2}$$

We have seen that δ is not a regular distribution, i.e. there is no $h \in L^1_{\text{loc}}(\mathbb{R})$ such that $\delta = T_h$. It may be worth to recall that, in earlier days, it was thought that H did have some sort of derivative, which supposedly was a function h such that

$$h(x) = \begin{cases} 0 & \text{if} \qquad x \neq 0\,, \\ +\infty & \text{if} \qquad x = 0\,, \end{cases} \tag{2.3.3}$$

(the latter because of the vertical jump of H at $x = 0$, at which H has "infinite slope"), and, in order to somehow compensate the infinite jumps from 0 to $+\infty$ and then back to 0 of h at $x = 0$, such that also

$$\int_{-\infty}^{+\infty} h(x)\,\mathrm{d}x = 1 \tag{2.3.4}$$

(see the remark at the end of section 2.3.3). The way out of these ambiguities is precisely to keep in mind that the only way we can differentiate H is in the distributional sense, and that $\delta = H'$ is not a function, but a distribution. More precisely, $\delta = (T_H)'$, and this distribution is known to be non-regular.

The latter statement is a global one, on $\mathbb{R}$. In contrast, the restrictions of δ to either interval $\mathbb{R}_{<0} =]-\infty, 0[$ or $\mathbb{R}_{>0} =]0, +\infty[$ are regular distributions. To see this, we denote respectively by H_- and H_+ the restrictions of H to $\mathbb{R}_{<0}$ and $\mathbb{R}_{>0}$. Then, both functions H_- and H_+ are differentiable in their domains, with $(H_\pm)'(x) \equiv 0$. Consequently, the corresponding regular distributions

$$T_- := T_{(H_-)'} \in \mathcal{D}'(\mathbb{R}_{<0})\,, \qquad T_+ := T_{(H_+)'} \in \mathcal{D}'(\mathbb{R}_{>0})\,, \tag{2.3.5}$$

generated, respectively, by $(H_-)'$ and $(H_+)'$, vanish. On the other hand, we obviously have that

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^-) \cup \mathcal{D}(\mathbb{R}^+)\,, \qquad \langle \delta, \varphi \rangle_{\mathcal{D}} = 0\,. \tag{2.3.6}$$

This confirms that $\delta = T_-$ in $\mathcal{D}'(\mathbb{R}_{<0})$, and $\delta = T_+$ in $\mathcal{D}'(\mathbb{R}_{>0})$, and implies that the restrictions of δ to $\mathcal{D}(\mathbb{R}^\pm)$ are regular, as claimed.

2.3.2 Approximation by Piecewise Differentiable Functions.

Our goal is now to approximate the Heaviside function H (extended to all of $\mathbb{R}$ by setting $H(0) = \frac{1}{2}$) by a sequence of absolutely continuous functions $(F_n)_{n\geq 1}$, and to show that their derivatives $f_n = {F_n}'$, which are in $L^1(\mathbb{R})$, while not converging in $L^1(\mathbb{R})$, do converge to δ in $\mathcal{D}'(\mathbb{R})$ (more precisely, $T_{f_n} \to \delta$ in $\mathcal{D}'(\mathbb{R})$). For $n \in \mathbb{N}_{>0}$ define $F_n : \mathbb{R} \to \mathbb{R}$ by

$$F_n(x) := \begin{cases} 0 & \text{if} \quad x < -\frac{1}{n}\,, \\ \frac{1}{2}(1+nx) & \text{if} \quad |x| \leq \frac{1}{n}\,, \\ 1 & \text{if} \quad x > \frac{1}{n}\,. \end{cases} \tag{2.3.7}$$

Then, each F_n is clearly absolutely continuous on $\mathbb{R}$, and differentiable for all $x \neq \pm\frac{1}{n}$, with derivative $f_n = {F_n}' \in L^1(\mathbb{R})$ defined by

$$f_n(x) = \begin{cases} 0 & \text{if} \quad |x| > \frac{1}{n}\,, \\ \frac{n}{2} & \text{if} \quad |x| < \frac{1}{n}\,. \end{cases} \tag{2.3.8}$$

We claim:

Proposition 2.3.1 *Let F_n and f_n be as above. Then, as $n \to +\infty$:*

1. *$F_n \to H$ pointwise on $\mathbb{R}$, and also uniformly on compact sets of $\mathbb{R}_{<0}$ or $\mathbb{R}_{>0}$.*

2. *$f_n \to 0$ pointwise on $\mathbb{R}_{<0}$ or $\mathbb{R}_{>0}$, and also uniformly on compact sets of $\mathbb{R}_{<0}$ or $\mathbb{R}_{>0}$.*

3. *$F_n \to H$ in $L^1_{\mathrm{loc}}(\mathbb{R})$.*

4. *$f_n = {F_n}'$ does not converge to 0 (the only possible candidate for H' if H had a generalized derivative) in $L^1(\mathbb{R})$; in fact,*

5. *The sequence $(f_n)_{n\geq 1}$ is not a Cauchy sequence in $L^1(\mathbb{R})$.*

6. *However, $f_n = {F_n}' \to \delta = H'$ in $\mathcal{D}'(\mathbb{R})$.*

Proof. (1) and (2) are obvious, and (3) is a consequence of the Lebesgue's dominated convergence theorem, since $0 \le F_n(x) \le 1$ for all $x \in I\!R$ and all $n \in I\!N$. (4) is a consequence of (5).

To show (5), take for example $m > n > 1$. Then,

$$\begin{aligned} \|f_m - f_n\|_{L^1(I\!R)} &= \int_{-\infty}^{+\infty} |f_m(x) - f_n(x)|\,\mathrm{d}x = \int_{-1}^{1} |f_m(x) - f_n(x)|\,\mathrm{d}x \\ &= \int_{-1/n}^{-1/m} \tfrac{n}{2}\,\mathrm{d}t + \int_{-1/m}^{1/m} \left(\tfrac{m}{2} - \tfrac{n}{2}\right)\mathrm{d}t + \int_{1/m}^{1/n} \tfrac{n}{2}\,\mathrm{d}t \qquad (2.3.9) \\ &= 2\left(1 - \tfrac{n}{m}\right) . \end{aligned}$$

Thus, choosing (e.g.) $m \ge 2n$, it follows that $\|f_m - f_n\|_{L^1(I\!R)} \ge 1$.

Finally, to show (6), fix $\varphi \in \mathcal{D}(I\!R)$, and let $a,\, b \in I\!R$ be such that $\operatorname{supp}\varphi \subseteq [a,b]$. Since

$$\langle f_n, \varphi\rangle_{\mathcal{D}} = \int_{-\infty}^{+\infty} f_n(x)\varphi(x)\,\mathrm{d}x = \int_{-1/n}^{1/n} f_n(x)\varphi(x)\,\mathrm{d}x\ , \qquad (2.3.10)$$

it follows that

$$\langle f_n, \varphi\rangle_{\mathcal{D}} \;=\; 0 \;=\; \langle \delta, \varphi\rangle_{\mathcal{D}} \qquad (2.3.11)$$

if $a < b < 0$ and $n > -\frac{1}{b}$, or if $0 < a < b$ and $n > \frac{1}{a}$. If instead $a < 0 < b$, by the mean value theorem

$$\langle f_n, \varphi\rangle_{\mathcal{D}} = \frac{n}{2}\int_{-1/n}^{1/n} \varphi(x)dx = \varphi(x_n)\ , \qquad (2.3.12)$$

for some $x_n \in \left[-\frac{1}{n}, \frac{1}{n}\right]$. From (2.3.12) we deduce that

$$\langle f_n, \varphi\rangle_{\mathcal{D}} = \varphi(x_n) \to \varphi(0) = \langle \delta, \varphi\rangle_{\mathcal{D}}\ , \qquad (2.3.13)$$

and this concludes the proof of proposition 2.3.1. $\square$

As a concluding remark, note that while each F_n can be represented as an integral function, by

$$F_n(x) = \int_{-\infty}^{x} f_n(t)\,\mathrm{d}t\ , \qquad (2.3.14)$$

the limit function H cannot be represented as an integral; that is, there is no function $h \in L^1(I\!R)$ such that

$$H(x) = \int_{-\infty}^{x} h(t)\,\mathrm{d}t\,. \tag{2.3.15}$$

Indeed, H would otherwise be absolutely continuous on all of $I\!R$ (see e.g. Rudin, [6], §8.17), which it is not.

2.3.3 Approximation by C^∞ Functions.

In this section we show that, in a certain sense, the results of §2.3.2 cannot be improved. More precisely, we now approximate H by a sequence of functions $\mathrm{P}_n \in C^\infty(I\!R)$, and show that their derivatives $\rho_n = {\mathrm{P}_n}'$, which are in $\mathcal{D}(I\!R)$, again do not converge in $L^1(I\!R)$, but still converge to δ in $\mathcal{D}'(I\!R)$.

To this end, we refer to the test function ρ defined by [3]

$$\rho(x) := \begin{cases} \rho_0 \exp\left(\frac{1}{1-|x|^2}\right) & \text{if} \quad |x| < 1\,, \\ 0 & \text{if} \quad |x| \geq 1\,, \end{cases} \tag{2.3.16}$$

whose support is the interval $[-1, 1]$, and where the constant ρ_0 is chosen so that

$$\int_{-\infty}^{+\infty} \rho(x)\,\mathrm{d}x = \int_{-1}^{1} \rho(x)\,\mathrm{d}x = 1 \tag{2.3.17}$$

(compare to (2.4.3)). For $n \geq 1$ we define

$$\rho_n(x) := n\,\rho(nx)\,, \qquad \mathrm{P}_n(x) := \int_{-\infty}^{x} \rho_n(t)\,\mathrm{d}t\,. \tag{2.3.18}$$

We note that the support of ρ_n is the interval $\left[-\frac{1}{n}, \frac{1}{n}\right]$, that

$$\rho_n(0) = \frac{n\,\rho_0}{\mathrm{e}} \to +\infty \qquad \text{as } n \to +\infty\,, \tag{2.3.19}$$

and that, by (2.3.17),

$$\int_{-\infty}^{+\infty} \rho_n(t)\,\mathrm{d}t = \int_{-1/n}^{1/n} \rho_n(t)\,\mathrm{d}t = n \int_{-1/n}^{1/n} \rho(nt)\,\mathrm{d}t = \int_{-1}^{1} \rho(\tau)\,d\tau = 1\,. \tag{2.3.20}$$

[3]This test function is "prototipical", as it is used extensively; for example, in the definition of the Friedrichs mollifiers.

In addition,

$$\forall\, x \le -\tfrac{1}{n}\,, \qquad \mathrm{P}_n(x) \;=\; \int_{-\infty}^{x} 0\,\mathrm{d}t = 0\,, \tag{2.3.21}$$

$$\forall\, x \ge \tfrac{1}{n}\,, \qquad \mathrm{P}_n(x) \;=\; \int_{-\infty}^{1/n} \rho_n(t)\,\mathrm{d}t + \int_{1/n}^{x} 0\,\mathrm{d}t = 1\,, \tag{2.3.22}$$

because of (2.3.20). Finally, since each ρ_n is an even function,

$$\mathrm{P}_n(0) = \tfrac{1}{2}\int_{-1/n}^{1/n} \rho_n(t)\,\mathrm{d}t = \tfrac{1}{2}\,. \tag{2.3.23}$$

Then, the same conclusions of proposition 2.3.1 hold, with F_n and f_n replaced by P_n and ρ_n.

Proposition 2.3.2 *Let P_n and ρ_n be as above. Then, as $n \to +\infty$,*

1. *$\mathrm{P}_n \to H$ pointwise on $I\!R$, and also uniformly on compact sets of $I\!R_{<0}$ or $I\!R_{>0}$.*
2. *$\rho_n \to 0$ pointwise on $I\!R_{<0}$ or $I\!R_{>0}$, and also uniformly on compact sets of $I\!R_{<0}$ or $I\!R_{>0}$.*
3. *$\mathrm{P}_n \to H$ in $L^1_{\mathrm{loc}}(I\!R)$.*
4. *$\rho_n = {\mathrm{P}_n}'$ does not converge to 0 in $L^1(I\!R)$; in fact,*
5. *The sequence $(\rho_n)_{n\ge 1}$ is not a Cauchy sequence in $L^1(I\!R)$.*
6. *However, $\rho_n = {\mathrm{P}_n}' \to \delta = H'$ in $\mathcal{D}'(I\!R)$.*

Proof. (1) and (2) are obvious, and (3) is a consequence of the Lebesgue's dominated convergence theorem, since $0 \le \mathrm{P}_n(x) \le 1$ for all $x \in I\!R$ and all $n \in I\!N$. (4) is a consequence of (5), which we can prove by contradiction if we assume (6). Indeed, if there existed $\rho := \lim \rho_n$ in $L^1(I\!R)$, by (2) it would be $\rho(x) \equiv 0$ for all $x \ne 0$, and it would follow that, for all $\varphi \in \mathcal{D}(I\!R)$,

$$I_n := \int_{-\infty}^{+\infty} \rho_n(x)\varphi(x)dx \to 0\,. \tag{2.3.24}$$

Choose then $\varphi \in \mathcal{D}(I\!R)$ with $\varphi(0) \neq 0$: then, (6) implies that

$$I_n \to \langle \delta, \varphi \rangle_{\mathcal{D}} = \varphi(0) \neq 0 \,, \tag{2.3.25}$$

contradicting (2.3.24).

Finally, to show (6), it is sufficient to show that

$$\langle \rho_n, \varphi \rangle_{\mathcal{D}} \to \langle \delta, \varphi \rangle_{\mathcal{D}} \tag{2.3.26}$$

for all $\varphi \in \mathcal{D}(I\!R)$ as above (if $a < b < 0$ or $0, a < b$, (2.3.26) is an obvious consequence of (2)). Setting

$$J_n := \int_{-\infty}^{+\infty} \rho_n(x)\varphi(x)dx - \varphi(0) \,, \tag{2.3.27}$$

recalling (2.3.20) we compute that

$$\begin{aligned} J_n &= \int_{-\infty}^{+\infty} \rho_n(x)\varphi(x)\,\mathrm{d}x - \varphi(0)\int_{-\infty}^{+\infty} \rho_n(t)\,\mathrm{d}t \\ &= n\int_{-1/n}^{1/n} \rho(nx)(\varphi(x) - \varphi(0))\,\mathrm{d}x \\ &= \int_{-1}^{1} \rho(t)\left(\varphi\left(\tfrac{t}{n}\right) - \varphi(0)\right)\,\mathrm{d}t \,. \end{aligned} \tag{2.3.28}$$

Consequently,

$$|J_n| \leq \sup_{|t|\leq 1} \left|\varphi\left(\tfrac{t}{n}\right) - \varphi(0)\right| \int_{-1}^{1} \rho(t)\,\mathrm{d}t = \sup_{|t|\leq 1} \left|\varphi\left(\tfrac{t}{n}\right) - \varphi(0)\right| \,, \tag{2.3.29}$$

and, therefore, $J_n \to 0$. This proves (2.3.26), and the proof of proposition 2.3.2 is complete. □

As a concluding remark, in propositions 2.3.1 and 2.3.2 we have proven that $f_n \to \delta$ and $\rho_n \to \delta$ in $\mathcal{D}'(I\!R)$. Since in both cases

$$\int_{-\infty}^{+\infty} f_n(t)\,\mathrm{d}t = 1 = \int_{-\infty}^{+\infty} \rho_n(t)\,\mathrm{d}t \,, \tag{2.3.30}$$

this was taken as a sort of justification of identity (2.3.4) (which, of course, does not make sense).

2.4 Test Functions.

1. In this section, $\Omega \subset I\!R^N$ denotes a domain, that is, an open and connected set, not necessarily bounded. We introduce the linear space $\mathcal{D}(\Omega)$ of the so-called TEST FUNCTIONS, and introduce in $\mathcal{D}(\Omega)$ a notion of sequential convergence, that allows us to consider the set $\mathcal{D}'(\Omega)$ of all linear, sequentially continuous maps from $\mathcal{D}(\Omega)$ into $I\!R$, which we call DISTRIBUTIONS on Ω.

1. We start with the set $C_0^\infty(\Omega)$ consisting of those functions $f : \Omega \to I\!R$ which are infinitely differentiable and have compact support in Ω.

Proposition 2.4.1 *$C_0^\infty(\Omega)$ is a linear space with respect to the usual definitions of the sum of two functions and of the product of a function by a scalar. For all multi-index $\alpha \in I\!N^N$ and all $\varphi \in C_0^\infty(\Omega)$,*

$$\mathrm{supp}(D^\alpha \varphi) \subseteq \mathrm{supp}(\varphi) \,. \tag{2.4.1}$$

In addition, $C_0^\infty(I\!R^N)$ is dense in $L^p(I\!R^N)$ if $1 \leq p < +\infty$.

Proof. 1) Let f, $g \in C_0^\infty(\Omega)$ and $\lambda \in I\!R$. Then, $f + \lambda\, g \in C^\infty(I\!R)$ (this is clear); to see that the support of $f + \lambda\, g$ is compact, we show the inclusion

$$A := \mathrm{supp}(f + \lambda\, g) \subseteq \mathrm{supp}(f) \cup \mathrm{supp}(g) =: B \,; \tag{2.4.2}$$

indeed, B is compact, and A, being a closed subset of a compact set, is also compact. To show the inclusion (2.4.2), let $x \in A$. Then, there is a sequence $(x_n)_{n\geq 1} \subset A^o$ (the interior of A), such that $x_n \to x$ (if $x \in A^o$, this step is not needed, as we can take $x_n = x$ for all n). Then, for each $n \geq 1$, $f(x_n) + \lambda\, g(x_n) \neq 0$, so that either $f(x_n) \neq 0$ or $g(x_n) \neq 0$. In the first case, $x_n \in \mathrm{supp}(f)$, while in the other case, $x_n \in \mathrm{supp}(g)$. In either case, $x_n \in B$, which is closed; thus, $x = \lim x_n \in B$, which proves (2.4.2).
2) We prove (2.4.1) as we did for (2.2.3). Let $K := \mathrm{supp}(\varphi)$. Then, $\varphi \equiv 0$ in $\Omega \setminus K$; since this set is open, this implies that $D^\alpha \varphi \equiv 0$ on $\Omega \setminus K$; in turn, this implies that $\mathrm{supp}(D^\alpha \varphi) \subseteq K$, as claimed.
3) For a proof of the density claim, see, e.g., [1, Cor. 2.30]. $\square$

Of course, the function $\varphi(x) \equiv 0$ is trivially a test function on any open domain. The prototypical example of a test function on an interval $]\alpha, \beta[\subset I\!R$ (not necessarily bounded), with support equal to $[a, b] \subset \,]\alpha, \beta[$, is

$$\varphi(x) := \begin{cases} \exp\left(-\frac{1}{x-a} - \frac{1}{b-x}\right) & \text{if} \quad a < x < b\,, \\ 0 & \text{if} \quad x \leq a \quad \text{or} \quad x \geq b\,. \end{cases} \tag{2.4.3}$$

Similarly, if $B(x_0, R)$ is an open ball contained in $\Omega \subseteq I\!R^N$, the function

$$\varphi(x) = \begin{cases} \exp\left(-\frac{1}{R^2 - |x - x_0|^2}\right) & \text{if} \quad x \in B(x_0, R)\,, \\ 0 & \quad \text{otherwise}\,, \end{cases} \tag{2.4.4}$$

is a test function on Ω, with $\operatorname{supp}\varphi = \overline{B(x_0, R)}$ (compare to (2.2.2)).

Exercise 2.4.1 *Prove that the functions φ defined by* (2.4.3) *and* (2.4.4) *are test functions.*

Obviously, test functions can be differentiated, and their derivatives of any order are again test functions. Since they are continuous, test functions can also be integrated, and their integral functions are again infinitely differentiable functions. However, the latter are not necessarily test functions; for example, the support of the integral function $\int_a^x \varphi(y)\,dy$ of the test function defined in (2.4.3) is the interval $[a, +\infty[$, which is unbounded. The following result characterizes test functions whose integrals are test functions.

Proposition 2.4.2 *Let $]\alpha, \beta[\; \subseteq I\!R$ be an interval, and $\varphi \in C_0^\infty(]\alpha, \beta[)$. There exists $\psi \in C_0^\infty(]\alpha, \beta[)$, with $\operatorname{supp}\psi \subseteq \operatorname{supp}\varphi$, such that $\psi' = \varphi$, if and only if*

$$\int_\alpha^\beta \varphi(x)\,\mathrm{d}x = 0\,. \tag{2.4.5}$$

This result can be generalized to N dimensions in the natural way (recall that we assume that Ω is connected).

Proof. Let $\varphi \in C_0^\infty(]\alpha, \beta[)$. Let a, $b \in I$ be such that $\operatorname{supp}\varphi \subseteq [a, b] \subset \,]\alpha, \beta[$. Then: If ψ exists as claimed,

$$\int_\alpha^\beta \varphi(x)\,\mathrm{d}x = \int_a^b \psi'(x)\,\mathrm{d}x = \psi(b) - \psi(a) = 0\,. \tag{2.4.6}$$

Conversely, let

$$\psi(t) := \int_\alpha^t \varphi(s)\, ds\,. \tag{2.4.7}$$

Then, $\psi \in C^\infty(]\alpha,\beta[)$, $\psi' = \varphi$, and for all $t \in]\alpha, a]$,

$$\psi(t) = \int_\alpha^t 0\, ds = 0\,, \tag{2.4.8}$$

while for all $t \in [b, \beta[$,

$$\psi(t) = \int_\alpha^b \varphi(s)\, ds = \int_\alpha^\beta \varphi(s)\, ds = 0\,, \tag{2.4.9}$$

by (2.4.5). This shows that $\psi \in C_0^\infty(]\alpha,\beta[)$, and $\operatorname{supp}\psi \subseteq \operatorname{supp}\varphi$. □

2. We now introduce the notion of sequential convergence of test functions. For $m \in I\!N$, we set

$$\|\varphi\|_m := \max_{\substack{|\alpha|\le m\\ x\in\Omega}} |\, D^\alpha\varphi(x)\,|\,, \qquad m \in I\!N\,. \tag{2.4.10}$$

Definition 2.4.1 *Let $(\varphi^k)_{k\ge 0} \subseteq C_0^\infty(\Omega)$, and $\varphi \in C_0^\infty(\Omega)$. We say that $\varphi^k \to \varphi$ in $C_0^\infty(\Omega)$ if:*

1) There is a compact set $K \subset \Omega$ such that both $\operatorname{supp}\varphi \subset K$ and $\operatorname{supp}\varphi^k \subset K$ for all $k \in I\!N$, and

2) For each $m \in I\!N$,

$$\|\varphi^k - \varphi\|_m \ \to\ 0\,; \tag{2.4.11}$$

that is, for all multi-indices $\alpha \in I\!N^N$, the derivatives $D^\alpha\varphi^k$ converge to $D^\alpha\varphi$ uniformly on K.

We remark that the definition of convergence $\varphi^k \to \varphi$ in $C_0^\infty(\Omega)$ requires *more* than just the uniform vanishing, as $k \to +\infty$, of all the derivatives of the functions $\varphi^k - \varphi$, as required in part (2) of definition 2.4.1. To see this, consider the the function φ defined by (2.4.3) on the interval $[-1,1]$, and for $k \ge 1$ set

$$\varphi^k(x) := \tfrac{1}{k}\,\varphi\left(\tfrac{x}{k}\right)\,. \tag{2.4.12}$$

Then, $\varphi^k \in C_0^\infty(I\!R)$ for each $k \ge 1$, and $D^m\varphi^k \to 0$ uniformly on $I\!R$ for each $m \in I\!N$ (Exercise; the claim is immediate for $m = 0$). However, φ^k does not converge to 0 in $C_0^\infty(I\!R)$, because there is no compact interval $K \subset I$ which contains the supports of all the φ^k. Indeed, $\operatorname{supp}\varphi^k = [-k,k]$. That is, part (1) of definition 2.4.1 fails.

Proposition 2.4.3 *Let $(\varphi^n)_{n\geq 1} \subset C_0^\infty(\Omega)$, and $\varphi \in C_0^\infty(\Omega)$. If $\varphi^n \to \varphi$ (in the sense of definition 2.4.1), then $D^\alpha \varphi^n \to D^\alpha \varphi$ for all multi-index $\alpha \in \mathbb{N}^N$.*

Proof. Let $K \subset \Omega$ be as in definition 2.4.1. Then, for each $n \geq 1$, $\varphi^n \equiv 0$ on $\Omega \setminus K$, which an open set. Thus, $D^\alpha \varphi^n \equiv 0$ on $\Omega \setminus K$, which means that $\operatorname{supp}(\varphi^n) \subseteq K$. Next, recalling (2.4.11), we see that for each $m \in \mathbb{N}$,

$$\|D^\alpha(\varphi^n - \varphi)\|_m \leq \|\varphi^n - \varphi\|_{m+|\alpha|} \to 0 \qquad \text{as} \quad n \to \infty\,. \tag{2.4.13}$$

Thus, $D^\alpha(\varphi^n) \to D^\alpha\varphi$ in the sense of definition 2.4.1. □

3. As we have stated, the set $C_0^\infty(\Omega)$ of all test functions on Ω is a linear space over $\mathbb{R}$, with respect to the usual operations of sum of two functions and product of a function by a scalar. Following Rudin, [7, Ch. 6, Sct. 2], we recall that, on this linear space, one can define a locally convex and metrizable topology $\mathcal{T}_1$, by means of the family of norms (2.4.10). However,

Proposition 2.4.4 *The topology $\mathcal{T}_1$ is not complete.*

Proof. Consider a function $\varphi \in C_0^\infty(\mathbb{R})$, with $\operatorname{supp}\varphi = [0,1]$ and $\varphi(x) \geq 0$ for all $x \in [0,1]$, and construct the sequence of functions $(\varphi^m)_{m\geq 1} \subset C_0^\infty(\mathbb{R})$, defined by

$$\varphi^m(x) := \sum_{j=1}^{m} \tfrac{1}{j}\, \varphi(x-j)\,. \tag{2.4.14}$$

Then, $(\varphi^m)_{m\geq 1}$ is a Cauchy sequence in $C_0^\infty(\mathbb{R})$. Indeed, for all $k \in \mathbb{N}$ and $m > n$,

$$|D^k(\varphi^m(x) - \varphi^n(x))| \leq \sum_{j=n+1}^{m} \tfrac{1}{j}\, |D^k(\varphi(x-j))| \leq \tfrac{1}{n+1} \max_{y\in\mathbb{R}} |D^k\varphi(y)|\,, \tag{2.4.15}$$

and the latter quantity of (2.4.15) vanishes, uniformly in $x \in \mathbb{R}$ (but not in k), as $n \to \infty$. We show below that, for each $m \geq 1$,

$$\operatorname{supp}(\varphi^m) = [1, m+1]\,; \tag{2.4.16}$$

thus, in this case too there is no compact interval $K \subset \mathbb{R}$ which contains the supports of all the φ^m. This means that the sequence $(\varphi^m)_{m\geq 1}$ does not converge in $C_0^\infty(I\!R)$, in the sense of definition 2.4.1. To show (2.4.16), we first note that, by (2.4.2),

$$\operatorname{supp}(\varphi^m) \subseteq \bigcup_{j=1}^{m} \operatorname{supp}(\varphi(\cdot - j)) = \bigcup_{j=1}^{m} [j, j+1] = [1, m+1] \,. \tag{2.4.17}$$

Conversely, let $x \in [1, m+1]$. There is k, $1 \le k \le m$, such that $x \in [k, k+1] = \operatorname{supp}(\varphi(\cdot - k)$; thus, there is a sequence $(x_n)_{n\ge1} \subset\,]k, k+1[$, such that $x_n \to x$. Then, $\varphi(x_n - k) \neq 0$, and since $\varphi(\cdot) \ge 0$,

$$\varphi^m(x) \ge \tfrac{1}{k}\, \varphi(x-k) > 0 \,. \tag{2.4.18}$$

It follows that $x_n \in \operatorname{supp}(\varphi^m)$; hence, $x = \lim x_n \in \operatorname{supp}(\varphi^m)$, which shows that $[1, m+1] \subseteq \operatorname{supp}(\varphi^m)$. $\square$

On the other hand, it is also possible to endow $C_0^\infty(\Omega)$ with a topology $\mathcal{T}_2$, which is locally convex and complete, but non-metrizable. Moreover, $\mathcal{T}_2$ is such that, if a map $T : C_0^\infty(\Omega) \to I\!R$ is linear, then T is continuous with respect to this topology, if and only if T is sequentially continuous. It is then customary to denote the topological space $(C_0^\infty(\Omega), \mathcal{T}_2)$ by $\mathcal{D}(\Omega)$, and to call each element of $\mathcal{D}(\Omega)$ a TEST FUNCTION.

2.5 Distributions.

1. In this section we extend and refine the provisional definition of distributions given in section 2.2.

Definition 2.5.1 *A linear, sequentially continuous map $T : \mathcal{D}(\Omega) \to I\!R$ is called a* DISTRIBUTION *on Ω. The set of all distributions on Ω is denoted by $\mathcal{D}'(\Omega)$. This set possesses the natural linear structure defined by*

$$[T_1 + \alpha T_2](\varphi) = T_1(\varphi) + \alpha T_2(\varphi) \,. \tag{2.5.1}$$

Thus, we call $\mathcal{D}'(\Omega)$ the SPACE OF DISTRIBUTIONS *on Ω. We shall often use the alternative notation $\langle T, \varphi\rangle_{\mathcal{D}}$ to denote the number $T(\varphi)$, $T \in \mathcal{D}'(\Omega)$, $\varphi \in \mathcal{D}(\Omega)$.*

The notation $\mathcal{D}'(\Omega)$ for the space of distributions is justified by the fact that $\mathcal{D}'(\Omega)$ is indeed the topological dual of the space $\mathcal{D}(\Omega)$; that is, the linear space $C_0^\infty(\Omega)$

equipped with the above mentioned topology $\mathcal{T}_2$. Then, $\mathcal{D}'(\Omega)$ is itself a topological space. For our purposes, it turns out that it is sufficient to consider in $\mathcal{D}'(\Omega)$ the weak* convergence of sequences [4]; that is:

Definition 2.5.2 *Let $(T_k)_{k\in I\!N} \subset \mathcal{D}'(\Omega)$, and $T \in \mathcal{D}'(\Omega)$. We say that $T_k \to T$ in $\mathcal{D}'(\Omega)$ if $T_k(\varphi) \to T(\varphi)$ (in $I\!R$) for all $\varphi \in \mathcal{D}(\Omega)$.*

It is essential to remark that a distribution is *not* a function, although there are distributions that are defined by means of a function. For example, we have seen that Dirac's δ distribution is not a function, and cannot be defined by means of any locally integrable function. The zero functional on $\mathcal{D}(\Omega)$ is a trivial distribution on Ω, which we also denote by 0. Other distributions can be constructed from a given distribution T by multiplication by a C^∞ function; more precisely, given $\zeta \in C^\infty(\Omega)$ and $T \in \mathcal{D}'(\Omega)$, we define a distribution ζT by

$$\langle \zeta T, \varphi\rangle_{\mathcal{D}} := \langle T, \zeta\varphi\rangle_{\mathcal{D}}\,, \qquad \forall\, \varphi \in \mathcal{D}(\Omega)\,. \tag{2.5.2}$$

This makes sense, because $\zeta\,\varphi \in \mathcal{D}(\Omega)$ if $\varphi \in \mathcal{D}(\Omega)$.

Just like for functions, there is a notion of the restriction of a distribution $T \in \mathcal{D}'(\Omega)$ to open subsets $A \subset \Omega$: noting that $\mathcal{D}(A) \subset \mathcal{D}(\Omega)$, it is natural to define the restriction T_A of T to A as the distribution $T_A \in \mathcal{D}'(A)$ defined by

$$\langle T_A, \varphi\rangle_{\mathcal{D}} = \langle T, \varphi\rangle_{\mathcal{D}}\,, \qquad \forall\, \varphi \in \mathcal{D}(A). \tag{2.5.3}$$

2. A fundamental class of distributions is the class of regular distributions, which we have already mentioned in section 2.2. The definition of this class is based on the following result, the proof of which we leave as an exercise.

Proposition 2.5.1 *Let $f \in L^1_{\rm loc}(\Omega)$. Then, the functional $T_f : \mathcal{D}(\Omega) \to I\!R$ defined by*

$$\langle T_f, \varphi\rangle_{\mathcal{D}} := \int_\Omega f(x)\varphi(x)\,\mathrm{d}x\,, \qquad \varphi \in \mathcal{D}(\Omega) \tag{2.5.4}$$

(*compare with* (2.2.5)), *is a distribution. Moreover,*

$$T_f \equiv 0 \quad \text{in } \mathcal{D}'(\Omega) \iff f \equiv 0 \quad \text{a.e. in } \Omega\,. \tag{2.5.5}$$

[4] This is the analogous of the point-wise convergence of a sequence of functions.

For example, if $T = T_f$ for some $f \in L^1_{\mathrm{loc}}(\Omega)$, (2.5.2) reads

$$\langle \zeta\, T_f, \varphi \rangle_{\mathcal{D}} \;= \int_\Omega f(x)\, \zeta(x)\, \varphi(x)\, \mathrm{d}x\,. \tag{2.5.6}$$

Proposition 2.5.1 allows us to define a map $\Psi : L^1_{\mathrm{loc}}(\Omega) \to \mathcal{D}'(\Omega)$, by

$$\Psi(f) := T_f\,, \qquad f \in L^1_{\mathrm{loc}}(\Omega)\,; \tag{2.5.7}$$

this map is obviously linear.

Definition 2.5.3 *A distribution $T \in \mathcal{D}'(\Omega)$ is called* REGULAR *if T is in the range of Ψ; that is, if there is $f \in L^1_{\mathrm{loc}}(\Omega)$ such that $T = T_f$.*

An example of a class of regular distributions is that consisting of constant distributions: since constant functions are locally integrable, given $c \in \mathbb{R}$, in accord with (2.5.4) we can define the distribution T_c, by

$$\langle T_c, \varphi \rangle_{\mathcal{D}} = \int_\Omega c\, \varphi(x)\, \mathrm{d}x = c \int_\Omega \varphi(x)\, \mathrm{d}x\,, \tag{2.5.8}$$

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

The meaning of (2.5.5) is that the map Ψ defined in (2.5.7) is injective. This allows us to identify $L^1_{\mathrm{loc}}(\Omega)$ with a linear subspace of $\mathcal{D}'(\Omega)$; this subspace consists precisely of those distributions we have called regular. As we know, Ψ is not surjective, that is, not every distribution is a regular one, as the example of Dirac's δ shows.

3. We conclude with a criterion that characterizes when a linear map from $\mathcal{D}(\Omega)$ into $\mathbb{R}$ is a distribution.

Proposition 2.5.2 *A linear map $T : \mathcal{D}(\Omega) \to \mathbb{R}$ is a distribution if and only if for every compact set $K \subset \Omega$ there are a constant $C_K > 0$ and an integer $j_K \geq 0$ such that, for all $\varphi \in \mathcal{D}(\Omega)$ with* $\mathrm{supp}(\varphi) \subseteq K$,

$$|T(\varphi)| \leq C_K\, \|\varphi\|_{j_K}\,, \tag{2.5.9}$$

(*the norm defined in* (2.4.10)).

Proof. Note that (2.5.9) is obviously true if $\varphi \equiv 0$.
1) Assume first that $T \in \mathcal{D}'(\Omega)$, but there is a compact set $\tilde{K} \subset \Omega$ with the property that for all $C > 0$ and all integers j, it is possible to find a test function $\varphi \in \mathcal{D}(\Omega)$ such that (2.5.9) does not hold; i.e.,

$$|T(\varphi)| > C\,\|\varphi\|_j\,. \tag{2.5.10}$$

Then, taking $C = j$, we can construct a sequence $(\varphi_j)_{j\geq 1} \subset \mathcal{D}(\Omega)$, with $\mathrm{supp}(\varphi_j) \subseteq K$, such that

$$|T(\varphi_j)| > j\,\|\varphi_j\|_j \geq 0\,. \tag{2.5.11}$$

In particular, (2.5.11) implies that $T(\varphi_j) \neq 0$, so that we can consider the function $\tilde{\varphi}_j := \frac{\varphi_j}{|T(\varphi_j)|}$. Clearly, $\tilde{\varphi}_j \in \mathcal{D}(\Omega)$, and $T(\tilde{\varphi}_j) = 1$. Then, for each $k \in \mathbb{N}$, we deduce from (2.5.11) that, for $j \geq k$,

$$\|\tilde{\varphi}_j\|_k \leq \|\tilde{\varphi}_j\|_j \leq \tfrac{1}{j}\,|T(\tilde{\varphi}_j)| = \tfrac{1}{j}\,; \tag{2.5.12}$$

consequently, $\varphi_j \to 0$ in $\mathcal{D}(\Omega)$. Since $T \in \mathcal{D}'(\Omega)$, this implies that $T(\tilde{\varphi}_j) \to 0$ (in $\mathbb{R}$), which contradicts the fact that $T(\tilde{\varphi}_j) = 1$.
2) Conversely, assume that $\varphi_m \to \varphi$ in $\mathcal{D}(\Omega)$, and let $K \subset \Omega$ be as in definition 2.4.1. Then, from (2.5.9),

$$|T(\varphi_m) - T(\varphi)| = |T(\varphi_m - \varphi)| \leq C_K\,\|\varphi_m - \varphi\|_{j_K} \to 0 \tag{2.5.13}$$

as $m \to \infty$. Hence, T is (sequentially) continuous, which means that $T \in \mathcal{D}'(\Omega)$. □

Definition 2.5.4 *Let $T \in \mathcal{D}'(\Omega)$, and assume that there is an integer j such that for all compact sets $K \subset \Omega$ there is $C_K > 0$ such that, for all $\varphi \in \mathcal{D}(\Omega)$ with* $\mathrm{supp}(\varphi) \subseteq K$*,*

$$|T(\varphi)| \leq C_K\,\|\varphi\|_j \tag{2.5.14}$$

(*that is,* (2.5.9) *holds for some j independent of K*)*. Let*

$$j_0 := \min\{j \in \mathbb{N} \ \mid\ \text{(2.5.14) holds}\}\,. \tag{2.5.15}$$

We call j_0 the order of the distribution T.

For example, regular distributions and the Dirac distributions are distributions of order 0. Indeed, let $f \in L^1_{\mathrm{loc}}(\Omega)$, and, given $\varphi \in \mathcal{D}(\Omega)$, let $K := \mathrm{supp}(\varphi)$. Then,

$$|T_f(\varphi)| \leq \int_\Omega |f(x)\,\varphi(x)|\,\mathrm{d}x \leq \max_{x\in K}\,|\varphi(x)| \int_K |f(x)|\,\mathrm{d}x = C_K\,\|\varphi\|_0\,. \tag{2.5.16}$$

Likewise,

$$|\delta_{x_0}(\varphi)| = |\varphi(x_0)| \leq \max_{x\in K} |\varphi(x)| \leq \|\varphi\|_0 \, . \tag{2.5.17}$$

An example [5] of distribution of order 1 on $I\!R$ is given by the map defined by

$$T(\varphi) := \lim_{\varepsilon\to 0^+} \int_{|x|\geq\varepsilon} \frac{\varphi(x)}{x} \, \mathrm{d}x \, . \tag{2.5.18}$$

Indeed, we immediately note that

$$\int_{|x|\geq\varepsilon} \frac{|\varphi(x)|}{|x|} \, \mathrm{d}x \leq \tfrac{1}{\varepsilon} \int_{\mathrm{supp}(\varphi)} |\varphi(x)| \, \mathrm{d}x \, , \tag{2.5.19}$$

so the integrals at the right side of (2.5.18) are well defined for each $\varepsilon > 0$ (note that the function $x \mapsto \frac{1}{x}$ is not in $L^1_{\mathrm{loc}}(I\!R)$). T is clearly linear; we prove that T is a distribution by means of proposition 2.5.2. To this end, we decompose

$$T(\varphi) = \int_{|x|\geq 1} \frac{\varphi(x)}{x} \, \mathrm{d}x + \int_{\varepsilon\leq|x|\leq 1} \frac{\varphi(x)}{x} \, \mathrm{d}x =: A(\varphi) + B_\varepsilon(\varphi) \, . \tag{2.5.20}$$

Since the function

$$x \mapsto f(x) := \begin{cases} 0 & \text{if} \quad |x| \leq 1 \, , \\ \frac{1}{x} & \text{if} \quad |x| \geq 1 \, , \end{cases} \tag{2.5.21}$$

is in $L^1_{\mathrm{loc}}(I\!R)$, it follows that $A = T_f$, and is a distribution of order 0. As for B_ε, we use Taylor's expansion at $x = 0$ to write

$$\varphi(x) = \varphi(0) + x \, r_\varphi(x) \, , \tag{2.5.22}$$

with $r_\varphi \in C^\infty(I\!R)$, and $\varphi \mapsto r_\varphi$ linear. We compute that

$$\int_{\varepsilon\leq|x|\leq 1} \frac{\varphi(0)}{x} \, \mathrm{d}x = \varphi(0) \left(\int_{-1}^{-\varepsilon} \frac{1}{x} \, \mathrm{d}x + \int_{\varepsilon}^{1} \frac{1}{x} \, \mathrm{d}x \right) = 0 \, ; \tag{2.5.23}$$

thus,

$$B_\varepsilon(\varphi) = \int_{\varepsilon\leq|x|\leq 1} r_\varphi(x) \, \mathrm{d}x \;\to\; \int_{-1}^{1} r_\varphi(x) \, \mathrm{d}x =: R(\varphi) \tag{2.5.24}$$

as $\varepsilon \to 0$. The map $\varphi \mapsto R(\varphi)$ is linear; in addition, by the mean value theorem for differentiable functions, for each $x \in [-1, 1]$ there is $\theta_x \in [[0, 1]$ such that

$$\varphi(x) - \varphi(0) = \varphi'(\theta_x) \, x \, ; \tag{2.5.25}$$

[5] Example 1.10 of Seiler, [8].

hence, recalling (2.5.22)

$$|r_\varphi(x)| = \left|\tfrac{1}{x}\left(\varphi(x) - \varphi(0)\right)\right| = |\varphi'(\theta_x)| \leq \max_{x \in \mathrm{supp}(\varphi)} |\varphi'(x)| \leq \|\varphi\|_1 \,, \tag{2.5.26}$$

so that

$$|R(\varphi)| \leq 2 \max_{|x|\leq 1} |r_\varphi(x)| \leq 2\,\|\varphi\|_1 \,. \tag{2.5.27}$$

By proposition 2.5.2, it follows that R is a distribution of order 1; consequently, also $T = A + R = T_f + R$ is a distribution of order 1. ($\square$)

2.5.1 Derivatives of Distributions.

1. As we have discussed in our informal introduction, the main motivation for distributions is that a notion of derivative can be introduced in $\mathcal{D}'(\Omega)$, whereby all distributions possess derivatives of any order, which are also distributions. This is justified by the following result, whose proof we leave as an exercise.

Proposition 2.5.3 *Let $T \in \mathcal{D}'(\Omega)$, and $\alpha \in \mathbb{N}^N$. The functionals $T^\alpha : \mathcal{D}(\Omega) \to \mathbb{R}$ defined by*

$$< T^\alpha, \varphi >_{\mathcal{D}} := (-1)^{|\alpha|} < T, D^\alpha \varphi >_{\mathcal{D}} \,, \qquad \forall\, \varphi \in \mathcal{D}(\Omega) \,, \tag{2.5.28}$$

(*recall the alternative notation introduced in definition 2.5.1*) *are distributions* (*i.e.,* $T^\alpha \in \mathcal{D}'(\Omega)$*; note that the right side of* (2.5.28) *makes sense, since* $D^\alpha\varphi \in \mathcal{D}(\Omega)$*.*)

Definition 2.5.5 *The distribution T^α defined by* (2.5.28) *is called the α-th distributional derivative of T, and is denoted by $D^\alpha T$.*

In particular for $N = 1$ and $\alpha = 1$, the derivative T' is defined by the identities

$$\langle T', \varphi \rangle_{\mathcal{D}} = -\,\langle T, \varphi' \rangle_{\mathcal{D}} \,, \qquad \forall\, \varphi \in \mathcal{D}(I) \,, \tag{2.5.29}$$

$I \subseteq \mathbb{R}$ an open interval. When T is regular, and is defined by a function $f \in L^1_{\mathrm{loc}}(I)$ (i.e., $T = T_f$), definition (2.5.29) coincides with (2.2.6).

Definition 2.5.5 is consistent, in the sense that the identity

$$D^\alpha(T_f) = T_{D^\alpha f} \tag{2.5.30}$$

holds in $\mathcal{D}'(\Omega)$, at least when $f \in C^{|\alpha|}(\Omega)$. (Identity (2.5.30) may be false if $D^\alpha f$ is not continuous; see Rudin, [7], §6.14). To show (2.5.30), let $f \in C^{|\alpha|}(\Omega)$. Then, both f and $D^\alpha f \in L^1_{\mathrm{loc}}(\Omega)$, so they define regular distributions T_f and $T_{D^\alpha f}$. Starting from (2.5.28), and then using classical integration by parts, we compute that, for all $\varphi \in \mathcal{D}(\Omega)$,

$$\begin{aligned} \langle D^\alpha(T_f), \varphi\rangle_{\mathcal{D}} &= (-1)^{|\alpha|}\langle T_f, D^\alpha\varphi\rangle_{\mathcal{D}} \\ &= (-1)^{|\alpha|}\int_\Omega f(x)[D^\alpha\varphi](x)\,\mathrm{d}x \\ &= (-1)^{2|\alpha|}\int_\Omega [D^\alpha f](x)\varphi(x)\,\mathrm{d}x \\ &= \langle T_{D^\alpha f}, \varphi\rangle_{\mathcal{D}}\,; \end{aligned} \tag{2.5.31}$$

this proves (2.5.31). □

2. We now briefly review some results that show how the definition of distributional derivative maintains many of the familiar features of classical derivatives of functions.

Proposition 2.5.4 *Let $T \in \mathcal{D}'(\Omega)$, $\zeta \in C^\infty(\Omega)$, and α, $\beta \in I\!N$. Then:*

1. *Distributional derivatives commute; that is, for all multi-indices α and $\beta \in I\!N^N$,*
$$D^{\alpha+\beta}T = D^\alpha[D^\beta T] = D^\beta[D^\alpha T]\,. \tag{2.5.32}$$
2. *Leibniz' formula holds:*
$$D^\alpha[\zeta T] = \sum_{\beta\le\alpha}\binom{\alpha}{\beta}(D^\beta\zeta)\,D^{\alpha-\beta}T\,. \tag{2.5.33}$$
(In (2.5.33), the distributions ζT and $(D^\beta\zeta)D^{\alpha-\beta}T$ are defined via (2.5.2)).
3. *If $T \in \mathcal{D}'(\Omega)$ is a constant distribution, and $|\alpha| > 0$, $D^\alpha T = 0$.*
4. *Conversely, if Ω is connected (which we assume), and $T \in \mathcal{D}'(\Omega)$ is such that $D^\alpha T = 0$ for all $\alpha \in I\!N$ with $|\alpha| = 1$, T is a constant distribution.*
5. *If $T_k \to T$ in $\mathcal{D}'(\Omega)$ (in the sense of definition 2.5.2), then $D^\alpha T_k \to D^\alpha T$ in $\mathcal{D}'(\Omega)$.*

Proof. 1) In accord with (2.5.28), we compute

$$\begin{aligned}
\langle D^{\alpha+\beta}T,\varphi\rangle_{\mathcal{D}} &= (-1)^{|\alpha+\beta|}\langle T,D^{\alpha+\beta}\varphi\rangle_{\mathcal{D}} \\
&= (-1)^{|\alpha|}(-1)^{|\beta|}\langle T,D^{\alpha}(D^{\beta}\varphi)\rangle_{\mathcal{D}} \\
&= (-1)^{|\alpha|}(-1)^{|\beta|}\langle T,D^{\beta}(D^{\alpha}\varphi)\rangle_{\mathcal{D}} \qquad (2.5.34)\\
&= (-1)^{|\alpha|}\langle D^{\beta}T,D^{\alpha}\varphi\rangle_{\mathcal{D}} \\
&= \langle D^{\alpha}[D^{\beta}T],\varphi\rangle_{\mathcal{D}}\,.
\end{aligned}$$

This proves that $D^{\alpha+\beta}T = D^{\alpha}[D^{\beta}T]$ in $\mathcal{D}'(\Omega)$. Reversing the role of α and β completes the proof of the first claim.

2) We prove (2.5.33) by induction on $m := |\alpha|$, with a repeated application of definitions (2.5.28) and (2.5.2), and resorting to the well-known identity

$$\binom{m}{j} + \binom{m}{j-1} = \binom{m+1}{j}\,. \qquad (2.5.35)$$

For $m = 1$, let $j \in \{1,\dots,n\}$ and $\varphi \in \mathcal{D}(\Omega)$. Then,

$$\begin{aligned}
\langle \partial_j(\zeta T),\varphi\rangle_{\mathcal{D}} &= -\langle \zeta T,\partial_j\varphi\rangle_{\mathcal{D}} = -\langle T,\zeta(\partial_j\varphi)\rangle_{\mathcal{D}} \\
&= -\langle T,\partial_j(\zeta\varphi)\rangle_{\mathcal{D}} + \langle T,(\partial_j\zeta)\varphi\rangle_{\mathcal{D}} \\
&= \langle \partial_j T,\zeta\varphi\rangle_{\mathcal{D}} + \langle T,(\partial_j\zeta)\varphi\rangle_{\mathcal{D}} \qquad (2.5.36)\\
&= \langle \zeta(\partial_j T),\varphi\rangle_{\mathcal{D}} + \langle (\partial_j\zeta)T,\varphi\rangle_{\mathcal{D}} \\
&= \langle \zeta(\partial_j T) + (\partial_j\zeta)T,\varphi\rangle_{\mathcal{D}}\,.
\end{aligned}$$

This means that

$$\partial_j(\zeta T) = \zeta(\partial_j T) + (\partial_j\zeta)T \qquad (2.5.37)$$

in $\mathcal{D}'(\Omega)$; that is, the first step in the induction process holds. Assume then that (2.5.33) holds for all α, with $m = |\alpha| \geq 1$, and let β be such that $|\beta| = m+1$. There exists then $\alpha \in \mathbb{N}$, and $j \in \{1,\dots,n\}$, such that $|\alpha| = m$ and $D^{\beta} = \partial_j D^{\alpha}$ (i.e., $\beta_k = \alpha_k$ if $k \neq j$, and $\beta_j = \alpha_j + 1$). Then, as above,

$$\langle D^{\beta}(\zeta T),\varphi\rangle_{\mathcal{D}} = -\langle D^{\alpha}(\zeta T),\partial_j\varphi\rangle_{\mathcal{D}}$$

$$\begin{aligned} &= -\sum_{\gamma\le\alpha}\binom{\alpha}{\gamma}\langle(D^\gamma\zeta)D^{\alpha-\gamma}T,\partial_j\varphi\rangle_{\mathcal{D}} \qquad (2.5.38)\\ &= \sum_{\gamma\le\alpha}\binom{\alpha}{\gamma}\langle\partial_j[(D^\gamma\zeta)D^{\alpha-\gamma}T],\varphi\rangle_{\mathcal{D}}\,, \end{aligned}$$

having used (2.5.33) for $|\alpha| = m$. Since (2.5.33) also holds for $m = 1$,

$$\begin{aligned} \langle D^\beta(\zeta T),\varphi\rangle_{\mathcal{D}} &= \sum_{\gamma\le\alpha}\binom{\alpha}{\gamma}\langle\Big((D^\gamma\zeta)\partial_j D^{\alpha-\gamma}T + (\partial_j D^\gamma\zeta)D^{\alpha-\gamma}T\Big),\varphi\rangle_{\mathcal{D}}\\ &= \sum_{\substack{\gamma\le\alpha\\ \gamma_j=0}}\binom{\alpha}{\gamma}\langle(D^\gamma\zeta)\partial_j D^{\alpha-\gamma}T,\varphi\rangle_{\mathcal{D}}\\ &\quad + \sum_{\substack{\gamma\le\alpha\\ \gamma_j\ge1}}\binom{\alpha}{\gamma}\langle(D^\gamma\zeta)\partial_j D^{\alpha-\gamma}T,\varphi\rangle_{\mathcal{D}} \qquad (2.5.39)\\ &\quad + \sum_{\substack{\gamma\le\alpha\\ \gamma_j\le\alpha_j-1}}\binom{\alpha}{\gamma}\langle(\partial_j D^\gamma\zeta)D^{\alpha-\gamma}T,\varphi\rangle_{\mathcal{D}}\\ &\quad + \sum_{\substack{\gamma\le\alpha\\ \gamma_j=\alpha_j}}\binom{\alpha}{\gamma}\langle(\partial_j D^\gamma\zeta)D^{\alpha-\gamma}T,\varphi\rangle_{\mathcal{D}}\\ &=: S_1 + S_{21} + S_{22} + S_3\,. \end{aligned}$$

We now show that this sum can be written as

$$\langle D^\beta(\zeta T),\varphi\rangle_{\mathcal{D}} = \sum_{\rho\le\beta}\langle(D^\rho\zeta)D^{\beta-\rho}T,\varphi\rangle_{\mathcal{D}}\,, \qquad (2.5.40)$$

with S_1 corresponding to the terms with $\rho_j = 0$, $S_2 := S_{21} + S_{22}$ corresponding to the terms with $1 \le \rho_j \le \beta_j - 1$, and S_3 to the terms with $\rho_j = \beta_j$. Indeed, consider S_1. Since

$$\binom{\alpha_j}{\gamma_j} = \binom{\alpha_j}{0} = 1 = \binom{\beta_j}{0}\,, \qquad (2.5.41)$$

the coefficients of the terms in S_1 are

$$\binom{\alpha}{\gamma} = \binom{\alpha_1}{\gamma_1}\dots\binom{\alpha_j}{0}\dots\binom{\alpha_n}{\gamma_n} = \binom{\beta_1}{\gamma_1}\dots\binom{\beta_j}{0}\dots\binom{\beta_n}{\gamma_n} = \binom{\beta}{\gamma}\,. \qquad (2.5.42)$$

The terms of S_1 corresponding to these coefficients can be written as

$$\langle(D^\gamma\zeta)\partial_j D^{\alpha-\gamma}T,\varphi\rangle_{\mathcal{D}} = \langle(D^\gamma\zeta)D^{\beta-\gamma}T,\varphi\rangle_{\mathcal{D}}\,; \qquad (2.5.43)$$

consequently, choosing $\rho = \gamma$, we see that, indeed,

$$S_1 = \sum_{\substack{\rho \le \beta \\ \rho_j = 0}} \binom{\beta}{\rho} \langle (D^\rho \zeta) D^{\beta-\rho} T, \varphi \rangle_{\mathcal{D}} \, . \tag{2.5.44}$$

As for S_{21}, setting $g := |\gamma|$, we have

$$\langle (D^\gamma \zeta) \partial_j D^{\alpha-\gamma} T, \varphi \rangle_{\mathcal{D}} = (-1)^{m-g-1} \langle T, \partial_j D^{\alpha-\gamma} \left[(D^\gamma \zeta) \varphi \right] \rangle_{\mathcal{D}} \, ; \tag{2.5.45}$$

choosing $\rho = \gamma$ we have $0 \le \gamma_k = \rho_k \le \alpha_k = \beta_k$ if $k \ne j$, $1 \le \gamma_j = \rho_j \le \alpha_j = \beta_j - 1$ and, obviously,

$$\partial_j D^{\alpha-\gamma} \left[(D^\gamma \zeta) \varphi \right] = D^{\beta-\gamma} \left[(D^\gamma \zeta) \varphi \right] = D^{\beta-\rho} \left[(D^\rho \zeta) \varphi \right] \, . \tag{2.5.46}$$

Analogously, for S_{22} we have

$$\langle (\partial_j D^\gamma \zeta) D^{\alpha-\gamma} T, \varphi \rangle_{\mathcal{D}} = (-1)^{m-g} \langle T, D^{\alpha-\gamma} \left[(\partial_j D^\gamma \zeta) \varphi \right] \rangle_{\mathcal{D}} \, ; \tag{2.5.47}$$

thus, choosing now ρ so that $\rho_k = \gamma_k$ if $k \ne j$ and $\rho_j = \gamma_j + 1$, we have $0 \le \rho_k \le \beta_k$ if $k \ne j$, $1 \langle \rho_j \le \beta_j - 1$, so that $\alpha - \gamma = \beta - \rho$, and, therefore,

$$D^{\alpha-\gamma} \left[(\partial_j D^\gamma \zeta) \varphi \right] = D^{\beta-\rho} \left[(D^\rho \zeta) \varphi \right] \, . \tag{2.5.48}$$

Comparing (2.5.46) and (2.5.48), we see that, indeed, these terms correspond to the terms $\langle (D^\rho \zeta) D^{\beta-\rho} T, \varphi \rangle_{\mathcal{D}}$, with $\rho \le \beta$ and $\rho_j \le \beta_j - 1$. Recalling (2.5.35), we see that the coefficients of these terms are

$$\begin{aligned}
&\left[\binom{\alpha_1}{\rho_1} \dots \binom{\alpha_j}{\rho_j} \dots \binom{\alpha_n}{\rho_n} \right] + \left[\binom{\alpha_1}{\rho_1} \dots \binom{\alpha_j}{\rho_j - 1} \dots \binom{\alpha_n}{\rho_n} \right] = \\
&\qquad = \binom{\alpha_1}{\rho_1} \dots \left[\binom{\alpha_j}{\rho_j} + \binom{\alpha_j}{\rho_j - 1} \right] \dots \binom{\alpha_n}{\rho_n} \\
&\qquad = \binom{\alpha_1}{\rho_1} \dots \binom{\alpha_j + 1}{\rho_j} \dots \binom{\alpha_n}{\rho_n} \\
&\qquad = \binom{\beta_1}{\rho_1} \dots \binom{\beta_j}{\rho_j} \dots \binom{\beta_n}{\rho_n} = \binom{\beta}{\rho} \, ,
\end{aligned} \tag{2.5.49}$$

as desired. Finally, consider S_3: since $\gamma_j = \alpha_j$, choosing ρ as we did for S_{22} we have $\rho_j = \gamma_j + 1 = \alpha_j + 1 = \beta_j$; therefore, since

$$\binom{\alpha_j}{\gamma_j} = \binom{\alpha_j}{\alpha_j} = 1 = \binom{\beta_j}{\beta_j} = \binom{\beta_j}{\rho_j} \, , \tag{2.5.50}$$

the coefficients in S_3 are

$$\binom{\alpha}{\gamma} = \binom{\alpha_1}{\gamma_1} \dots \binom{\alpha_j}{\gamma_j} \dots \binom{\alpha_n}{\gamma_n} = \binom{\beta_1}{\rho_1} \dots \binom{\beta_j}{\rho_j} \dots \binom{\beta_n}{\rho_n} = \binom{\beta}{\rho} \, . \tag{2.5.51}$$

The corresponding terms can be written as $\langle (D^\rho\zeta)D^{\beta-\rho}T,\varphi\rangle_{\mathcal{D}}$; thus,

$$S_3 = \sum_{\substack{\rho\le\beta \\ \rho_j=\beta_j}} \binom{\beta}{\rho}\langle (D^\rho\zeta)D^{\beta-\rho}T,\varphi\rangle_{\mathcal{D}}\,. \tag{2.5.52}$$

This concludes the proof of (2.5.33).

3) We have to show that, if $T = T_c$ for some $c \in I\!R$, then $\partial_j T = 0$ for all $j \in \{1,\dots,N\}$. Thus, fix $\varphi \in \mathcal{D}(\Omega)$ and, if $K := \operatorname{supp}\varphi$, choose $\Omega' \subseteq \Omega$ such that $\partial\Omega'$ is of class C^1 and $K \subset \Omega' \subseteq \Omega$. Let ν denote the outward normal to $\partial\Omega'$. Then, by the divergence theorem,

$$\begin{aligned}\langle \partial_j T,\varphi\rangle_{\mathcal{D}} &= -\langle T,\partial_j\varphi\rangle_{\mathcal{D}} = -c\int_\Omega \partial_j\varphi(x)\,\mathrm{d}x \\ &= -c\int_{\Omega'}\partial_j\varphi(x)\,\mathrm{d}x = -c\int_{\partial\Omega'}\nu_j(x)\varphi(x)\,\mathrm{d}x = 0\,.\end{aligned} \tag{2.5.53}$$

4) We prove this result in the simpler case of one dimension, when $\Omega = I$ is an interval of $I\!R$; for the general case, see Renardy & Rogers, [5], §5.1. We show first that if ζ_1, $\zeta_2 \in \mathcal{D}(I)$ are such that

$$\int_I \zeta_1(x)\,\mathrm{d}x = \int_I \zeta_2(x)\,\mathrm{d}x\,, \tag{2.5.54}$$

then

$$\langle T,\zeta_1\rangle_{\mathcal{D}} = \langle T,\zeta_2\rangle_{\mathcal{D}}\,; \tag{2.5.55}$$

that is, T assigns the same value to test functions which have the same average. This is an immediate consequence of proposition 2.4.2. By (2.5.54), there exists $\psi \in \mathcal{D}(I)$ such that $\psi' = \zeta_1 - \zeta_2$. Therefore,

$$\langle T,\zeta_1-\zeta_2\rangle_{\mathcal{D}} = \langle T,\psi'\rangle_{\mathcal{D}} = -\langle T',\psi\rangle_{\mathcal{D}} = 0\,. \tag{2.5.56}$$

We fix then $\zeta \in \mathcal{D}(I)$ such that $\displaystyle\int_I \zeta(x)\,\mathrm{d}x = 1$ and, given $\varphi \in \mathcal{D}(I)$, consider the function

$$x \;\mapsto\; \alpha(x) := \varphi(x) - \left(\int_I \varphi(t)dt\right)\zeta(x)\,. \tag{2.5.57}$$

Then, $\alpha \in \mathcal{D}(I)$, and since

$$\int_I \alpha(x)\,\mathrm{d}x = \int_I \varphi(x)\,\mathrm{d}x - \int_I \varphi(t)dt\int_I \zeta(x)\,\mathrm{d}x = 0\,, \tag{2.5.58}$$

by proposition 2.4.2 there is $\beta \in \mathcal{D}(I)$ such that $\alpha = \beta'$. Consequently, as in (2.5.56), $\langle T, \alpha \rangle_{\mathcal{D}} = 0$. In turn, this implies that

$$\langle T, \varphi \rangle_{\mathcal{D}} = \left(\int_I \varphi(x)\, \mathrm{d}x \right) \langle T, \zeta \rangle_{\mathcal{D}} \,. \tag{2.5.59}$$

Letting c denote the value of $\langle T, \zeta \rangle_{\mathcal{D}}$ common to all ζ with $\int_I \zeta(x)\, \mathrm{d}x = 1$, we conclude that

$$\langle T_c, \varphi \rangle_{\mathcal{D}} = c \int_I \varphi(x)\, \mathrm{d}x = \left(\int_I \varphi(x)\, \mathrm{d}x \right) \langle T, \zeta \rangle_{\mathcal{D}} \,. \tag{2.5.60}$$

Recalling (2.5.59), this shows that $T = T_c$, as claimed.

5) The proof of the last claim is immediate. In fact, for each $\varphi \in \mathcal{D}(\Omega)$, as $k \to +\infty$,

$$\langle D^\alpha T_k, \varphi \rangle_{\mathcal{D}} = (-1)^{|\alpha|} \langle T_k, D^\alpha \varphi \rangle_{\mathcal{D}} \;\to\; (-1)^{|\alpha|} \langle T, D^\alpha \varphi \rangle_{\mathcal{D}} \;=\; \langle D^\alpha T, \varphi \rangle_{\mathcal{D}} \,. \tag{2.5.61}$$

This concludes the proof of proposition 2.5.4. $\square$

2.5.2 An Example: the Equation $-\Delta u = \delta_0$ in $\mathbb{R}^3$.

Let $\Omega := \mathbb{R}^3 \setminus \{0\}$, and consider the function $f : \Omega \to \mathbb{R}$ defined by

$$x \;\mapsto\; f(x) := \frac{1}{|x|} \,. \tag{2.5.62}$$

Then, $f \in C^\infty(\Omega)$, and f and all components of ∇f are in $L^1_{\mathrm{loc}}(\mathbb{R}^3)$. In fact, using spherical coordinates, we see that, for any $R > 0$,

$$\int_{|x| \le R} \frac{1}{|x|}\, \mathrm{d}x \le C_1\, R^2 \,, \tag{2.5.63}$$

and, recalling that, for $j = 1, 2, 3$,

$$\partial_j \frac{1}{|x|} = \frac{-x_j}{|x|^3} \,, \qquad \text{i.e.} \qquad \nabla \frac{1}{|x|} = -\frac{x}{|x|^3} \,, \tag{2.5.64}$$

that

$$\int_{|x| \le R} |[\partial_j f](x)|\, \mathrm{d}x \le C_2\, R \,, \tag{2.5.65}$$

for suitable constants C_1, C_2 independent of R. However, no components of the Hessian matrix $D^2 f$ are in $L^1_{\mathrm{loc}}(I\!R^3)$. Indeed, since

$$\partial_j \partial_k \frac{1}{|x|} = \frac{3x_j x_k}{|x|^5} - \frac{\delta_{jk}}{|x|^3}\,, \tag{2.5.66}$$

we also see that, for $\varepsilon \in\,]0, R[$,

$$\int\limits_{\varepsilon \le |x| \le R} |[\partial_j \partial_k f](x)|\ \mathrm{d}x = \begin{cases} \quad 0 & \text{if} \quad j \neq k\,, \\ \mathcal{O}\left(|\ln \varepsilon|\right) & \text{if} \quad j = k\,, \end{cases} \tag{2.5.67}$$

as $\varepsilon \to 0$. Our goal is to compute Δf in distributional sense, interpreting

$$\Delta f := \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2} \tag{2.5.68}$$

as a sum of second-order distributional derivatives of the distribution $T_f \in \mathcal{D}'(I\!R^3)$ defined by the locally integrable function f, via (2.5.4). We claim that

$$\Delta f = -4\,\pi\,\delta_0 \qquad \text{in}\ \ \mathcal{D}'(I\!R^3)\,. \tag{2.5.69}$$

Admitting this, it is then common to say that the function

$$x \ \mapsto\ u(x) := \frac{1}{4\pi|x|} \tag{2.5.70}$$

satisfies the Laplace equation

$$-\Delta u = \delta_0 \tag{2.5.71}$$

in distributional sense. (In other contexts, u is called the fundamental solution of the Laplace operator $-\Delta$ in $I\!R^3$; see definition 2.6.2 below.) Identity (2.5.69) may be somewhat surprising, given that (2.5.66) implies that, for all $x \in \Omega$,

$$\Delta f(x) = 0\,. \tag{2.5.72}$$

The situation is somewhat similar to that of the derivative of the Heaviside function; in a vague sense, the right side of (2.5.69) "keeps track" of the different behavior, with respect to integrability, of the singularities of f and its derivatives at $x = 0$.

To show (2.5.69), we fix arbitrary $\varphi \in \mathcal{D}(I\!R^3)$, and start computing

$$\begin{aligned} \langle \Delta T_f, \varphi\rangle_{\mathcal{D}} \ &=\ \langle T_f, \Delta\varphi\rangle_{\mathcal{D}} = \int_{I\!R^3} \frac{1}{|x|}\,\Delta\varphi(x)\,\mathrm{d}x \\ &=\ \lim_{\varepsilon\to 0} \int_{|x|\ge\varepsilon} \frac{1}{|x|}\,\Delta\varphi(x)\,\mathrm{d}x\,. \end{aligned} \tag{2.5.73}$$

Integrating by parts, we proceed with

$$\begin{aligned}\langle \Delta T_f, \varphi\rangle_{\mathcal{D}} &= \lim_{\varepsilon\to 0}\Big\{ \int_{|x|\geq\varepsilon} \Big(\Delta \tfrac{1}{|x|}\Big)\,\varphi(x)\,\mathrm{d}x \\ &\quad + \int_{|x|=\varepsilon} (\nu(x)\cdot\nabla\varphi(x))\,\tfrac{1}{|x|}\,\mathrm{d}x \\ &\quad - \int_{|x|=\varepsilon} \Big(\nu(x)\cdot\nabla\tfrac{1}{|x|}\Big)\,\varphi(x)\,\mathrm{d}x\Big\} \\ &=: \lim_{\varepsilon\to 0}\left(J_{1\varepsilon}+J_{2\varepsilon}-J_{3\varepsilon}\right)\ ,\end{aligned} \tag{2.5.74}$$

where ν is the outward normal to the boundary of the exterior domain $I\!R^3 - B(0,\varepsilon)$; that is,

$$\nu(x) = \frac{-x}{|x|} = -\,\frac{x}{\varepsilon}, \qquad |x| = \varepsilon\ . \tag{2.5.75}$$

By (2.5.72), $J_{1\varepsilon} = 0$. Next, we estimate

$$|J_{2\varepsilon}| \leq \int_{|x|=\varepsilon} |\nabla\varphi(x)|\,\frac{1}{\varepsilon}\,\mathrm{d}x \leq \frac{1}{\varepsilon}\,\max_{|x|=\varepsilon}|\nabla\varphi(x)|\int_{|x|=\varepsilon} 1\,\mathrm{d}x \leq C_\varphi\, 4\pi\varepsilon\ , \tag{2.5.76}$$

for suitable constant C_φ depending on φ. Hence, $J_{2\varepsilon} \to 0$ as $\varepsilon \to 0$. Similarly, recalling (2.5.75) and (2.5.64), we compute that, on the boundary $\{|x| = \varepsilon\}$,

$$\nu(x)\cdot\nabla\frac{1}{|x|} = \frac{1}{\varepsilon^2}\ . \tag{2.5.77}$$

Therefore,

$$J_{3\varepsilon} = \frac{1}{\varepsilon^2}\int_{|x|=\varepsilon}\varphi(x)\,\mathrm{d}x = 4\pi\frac{1}{4\pi\varepsilon^2}\int_{|x|=\varepsilon}\varphi(x)\,\mathrm{d}x\ . \tag{2.5.78}$$

Since $4\pi\varepsilon^2$ is the area of the spherical surface $\{|x| = \varepsilon\}$, and φ is continuous, $J_{3\varepsilon} \to 4\pi\varphi(0)$ as $\varepsilon \to 0$. It follows that

$$\langle \Delta T_f, \varphi\rangle_{\mathcal{D}} = \lim_{\varepsilon\to 0}\left(J_{2\varepsilon} - J_{3\varepsilon}\right) = -4\pi\varphi(0) = -4\pi\,\langle\delta_0,\varphi\rangle_{\mathcal{D}}\ , \tag{2.5.79}$$

from which (2.5.69) follows. $\square$

2.6 Applications to PDEs.

In this section[6] we present an application of distribution theory to linear PDEs with constant coefficients.

[6] Section 1.5 of Seiler, [8].

2.6.1 Convolution.

We start with the formal definition of the convolution of two functions, and that of a distribution with a test function [7].

Definition 2.6.1 *Let $f,\ g : \mathbb{R}^N \to \mathbb{R}$. The* CONVOLUTION *of f and g is the function $f * g : \mathbb{R}^N \to \mathbb{R}$ defined by*

$$[f * g](x) := \int_{\mathbb{R}^N} f(x-y)\, g(y)\, \mathrm{d}y = \int_{\mathbb{R}^N} f(y)\, g(x-y)\, \mathrm{d}y\,. \tag{2.6.1}$$

(The integrals in (2.6.1) are seen to be the same by means of the change of variable $z = x - y$, and then renaming $z = y$). Definition (2.6.1) is formal; one important case when the integrals at its right side are well defined is described in the following result, where we denote by $|\cdot|_p$ the usual norm in $L^p(\mathbb{R}^N)$, $1 \le p \le +\infty$.

Theorem 2.6.1 *Given p and $q \in [1, \infty]$, define $r \in [0, \infty]$ by*

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1\,, \tag{2.6.2}$$

where we interpret $\frac{1}{\infty} := 0$. If $f \in L^p(\mathbb{R}^N)$ and $g \in L^q(\mathbb{R}^N)$, then $f * g \in L^r(\mathbb{R}^N)$, and satisfies YOUNG*'s inequality*

$$|f * g|_r \le |f|_p\, |g|_q\,. \tag{2.6.3}$$

Note that (2.6.3) implies the continuity of the map $(f, g) \mapsto f * g$ from $L^p(\mathbb{R}^N) \times L^q(\mathbb{R}^N)$ into $L^r(\mathbb{R}^N)$, and that this map has unit norm. Theorem 2.6.1 implies in particular that $L^1(\mathbb{R}^N)$ is an algebra with respect to the convolution product; that is, if f and $g \in L^1(\mathbb{R}^N)$, then $f * g \in L^1(\mathbb{R}^N)$, and

$$|f * g|_1 \le |f|_1\, |g|_1\,. \tag{2.6.4}$$

[7] One can also define the convolution of two distributions, when at least one of them has compact support (see (2.5.3) for the definition of the support of a distribution). We refer to § 6.36 of Rudin, [7], for all details.

Given a distribution $T \in \mathcal{D}'(I\!R^N)$, and $\varphi \in \mathcal{D}(I\!R^N)$, we define a function $u : I\!R^N \to I\!R$ by

$$u(x) := T(\varphi(x-\cdot)) = \langle T, \varphi(x-\cdot)\rangle_{\mathcal{D}} \,; \tag{2.6.5}$$

that is, for each fixed $x \in I\!R^N$, the distribution T is applied to the test function $y \mapsto \varphi(x-y)$. We write $u =: T * \varphi$, and call $T * \varphi$ the convolution of T with φ. Again, this convolution is a function; in fact, using differentiation under the integral sign [8], we can prove

Theorem 2.6.2 *Let u be defined by* (2.6.5)*. Then, $u \in C^\infty(I\!R^N)$, with*

$$[D^\alpha u](x) = \langle D^\alpha T, (\varphi(x-\cdot))\rangle_{\mathcal{D}} = \langle T, [D^\alpha \varphi](x-\cdot)\rangle_{\mathcal{D}} \,. \tag{2.6.6}$$

Definition (2.6.5) is motivated by the observation that if T is a regular distribution generated by a function $f \in L^1_{\mathrm{loc}}(I\!R^N)$, then for $\varphi \in \mathcal{D}(I\!R^N)$

$$T_f(\varphi(x-\cdot)) = \int_{I\!R^N} f(y)\,\varphi(x-y)\,\mathrm{d}y = [f * \varphi](x) \,; \tag{2.6.7}$$

that is, the convolution of T_f with φ coincides with the convolution of the functions f and φ (recall that we often identify $f$with T_f).

As an example of convolution, we show that, for all $\varphi \in \mathcal{D}(I\!R^N)$,

$$\delta_0 * \varphi = \varphi \,. \tag{2.6.8}$$

Indeed, for all $x \in I\!R^N$, by (2.6.5),

$$[\delta_0 * \varphi](x) = \langle \delta_0, \varphi(x-\cdot)\rangle = \varphi(x-0) = \varphi(x) \,. \tag{2.6.9}$$

Note that (2.6.8) confirms that $\delta_0 * \varphi \in C^\infty(I\!R^N)$, in accord with theorem 2.6.2.

2.6.2 Distributions and Linear PDEs.

1. A differential operator

$$A := \sum_{|\alpha| \le m} a_\alpha \, D^\alpha \tag{2.6.10}$$

[8] We consider the integrals in (2.6.1) as integrals depending on the parameter x

of order m, with constant coefficients $a_\alpha \in I\!R$ induces a map $\tilde{A} : \mathcal{D}'(I\!R^N) \to \mathcal{D}'(I\!R^N)$, defined by

$$\tilde{A}\,T := \sum_{|\alpha|\le m} a_\alpha\, D^\alpha T\,, \qquad T \in \mathcal{D}'(I\!R^N)\,, \tag{2.6.11}$$

in which the derivatives $D^\alpha T$ are defined in (2.5.28). Thus, the distribution $\tilde{A}\,T$ is defined by

$$\langle \tilde{A}\,T, \varphi\rangle_{\mathcal{D}} = \langle T, A^t\,(\varphi)\rangle_{\mathcal{D}}\,, \tag{2.6.12}$$

where the opertor A^t, called the formal adjoint of A, is defined by

$$A^t := \sum_{|\alpha|\le m} (-1)^{|\alpha|}\, a_\alpha\, D^\alpha\,. \tag{2.6.13}$$

It is usual practice to identify $\tilde{A}$ with A, although it should be kept in mind that $\tilde{A}$ acts on distributions, while A (and A^t) acts on C^∞ functions.

2. Given a distribution $F \in \mathcal{D}'(I\!R^N)$, we may ask whether there exists a distribution $U \in \mathcal{D}'(I\!R^N)$, solution of the distributional PDE

$$A\,U = F\,; \tag{2.6.14}$$

that is, by (2.6.12), such that for all $\varphi \in \mathcal{D}(I\!R^N)$,

$$\langle T, A^t\,\varphi\rangle_{\mathcal{D}} = \langle F, \varphi\rangle_{\mathcal{D}}\,. \tag{2.6.15}$$

Definition 2.6.2 *A distribution $E \in \mathcal{D}'(I\!R^N)$ is called a fundamental solution of A if*

$$A\,E = \delta_0\,. \tag{2.6.16}$$

For example, we saw in section 2.5.2 that the distribution T_u defined by the function $u(x) = \frac{1}{4\pi|x|}$, which is in $L^1_{\mathrm{loc}}(I\!R^3 \setminus \{0\})$, is a fundamental solution of the Laplacian $-\Delta$ in $I\!R^3$ (note that $(-\Delta)^t = -\Delta$; that is, $-\Delta$ is a formally self-adjoint operator).[9]

The importance of the fundamental solution E of the operator A of (2.6.10) is illustrated by the following result.

[9] The existence of a fundamental solution of a differential operator A with constant coefficients, as in (2.6.10) is guaranteed by a theorem of Ehrenpreis and Malgrange (1955-56).

Theorem 2.6.3 *Given $f \in \mathcal{D}(\mathbb{R}^N)$, let $u := E * f$. Then u, which is in $C^\infty(\mathbb{R}^N)$, is a solution of the PDE*

$$A\,u = f\,. \tag{2.6.17}$$

Proof. In order to avoid confusion, we denote by D^α the distributional derivative of order α, and by ∂_x^α the classical derivative of order α with respect to the variable x. Recalling (2.6.6) and (2.6.5), we compute that

$$\begin{aligned}
[A(E*f)](x) &= \sum_{|\alpha|\le m} a_\alpha\,[\partial_x^\alpha(E*f)](x) \\
&= \sum_{|\alpha|\le m} a_\alpha\,\partial_x^\alpha\langle E, f(x-\cdot)\rangle_{\mathcal{D}} \\
&= \sum_{|\alpha|\le m} a_\alpha\,(-1)^{|\alpha|}\langle E, \partial_y^\alpha f(x-\cdot)\rangle_{\mathcal{D}} \\
&= \sum_{|\alpha|\le m} a_\alpha\,\langle D^\alpha E, f(x-\cdot)\rangle_{\mathcal{D}} \\
&= \langle \sum_{|\alpha|\le m} a_\alpha\,D^\alpha E, f(x-\cdot)\rangle_{\mathcal{D}} \\
&= \langle AE, f(x-\cdot)\rangle_{\mathcal{D}} = [(AE)*f](x) \\
&= [\delta_0 * f](x) = f(x)\,.
\end{aligned} \tag{2.6.18}$$

In the next to last line of (2.6.18), recall that AE stands for $\tilde{A}E$, with $\tilde{A}$ defined in (2.6.11). □

For example, the function

$$w(x) := \left[\tfrac{1}{4\pi\,|\cdot|} * f\right](x) = \tfrac{1}{4\pi}\int_{\mathbb{R}^3} \frac{f(y)}{|x-y|}\,\mathrm{d}y \tag{2.6.19}$$

is a solution to the Laplace equation in $\mathbb{R}^3$

$$-\,(\partial_1^2 + \partial_2^2 + \partial_3^2)\,w = f\,. \tag{2.6.20}$$

Exercise 2.6.1 *Show that the fundamental solution of the Laplacian* $\Delta := \partial_1^2 + \partial_2^2$ *in* $I\!R^2$ *is the distribution* T_u*, with* $u \in L^1_{\mathrm{loc}}(I\!R^2 \setminus \{0\})$ *defined by*

$$u(x) := \tfrac{-1}{2\pi} \ln|x| \,. \tag{2.6.21}$$

Write the corresponding solution of the Laplace equation in $I\!R^2$

$$-\,(\partial_1^2 + \partial_2^2)\, w = f \,. \tag{2.6.22}$$

2.7 Tempered Distributions.

As it turns out, in order to define generalized solutions to various types of PDEs (such as, for example, solutions in the Sobolev spaces $H^s(I\!R^N)$, which are spaces of distributions that are regular in a sense similar to that of §2.2), we need to consider spaces of distributions that are smaller than the space $\mathcal{D}'(I\!R^N)$. In a certain sense, the space $\mathcal{D}'(I\!R^N)$ is "too large"; or, correspondingly, the space $\mathcal{D}(I\!R^N)$ of test functions is "too small". For example, the requirement that test functions have compact support is too strong for a satisfactory theory of the Fourier transform (see chapter 3): the Fourier transform of a test function is a C^∞ function whose support is, in general, not compact. On the other hand, simply doing away with the requirement of compact support is not satisfactory either, since functions in $C^\infty(I\!R^N)$ may grow too fast to be integrable; thus, this space cannot be imbedded into any space $L^p(I\!R^N)$, with the consequence that the Fourier transform cannot be defined for all functions in $C^\infty(I\!R^N)$. Finally, the structure of the topological dual of this space is too complicated to be of general use in applications.

2.7.1 Rapidly Decreasing Functions.

To overcome the above mentioned obstacles, we introduce a sort of intermediate space between $C_0^\infty(I\!R^N)$ and $C^\infty(I\!R^N)$: this space is called the Schwartz space of RAPIDLY DECREASING functions, and is defined by

$$\mathcal{S}(I\!R^N) := \Big\{ f \in C^\infty(I\!R^N) \;\mid\; \forall\, \alpha,\ \beta \in I\!N^N,\ (\cdot)^\beta D^\alpha f \in L^\infty(I\!R^N) \Big\} \,, \tag{2.7.1}$$

or, equivalently,

$$\mathcal{S}(I\!R^N) = \Big\{ f \in C^\infty(I\!R^N) \mid \forall\, k \in I\!N,\ \forall\, \alpha \in I\!N^N,\ |\cdot|^k D^\alpha f \in L^\infty(I\!R^N) \Big\} \,. \tag{2.7.2}$$

Thus, functions in $\mathcal{S}(I\!R^N)$ are infinitely differentiable, and all their derivatives decay faster than any polynomial as $|x| \to +\infty$; an example of a function in $\mathcal{S}(I\!R^N)$ is $f(x) = \mathrm{e}^{-|x|^2}$.

We first record the following obvious properties of $\mathcal{S}(I\!R^N)$.

Proposition 2.7.1 *Let $f \in \mathcal{S}(I\!R^N)$. Then, for all $k \in I\!N$ and $\alpha \in I\!N$,*

$$\lim_{|x|\to+\infty} (1+|x|^2)^k |(D^\alpha f)(x)| = 0. \tag{2.7.3}$$

Moreover, if P is an arbitrary polynomial with constant coefficients, both $P(\cdot)f$ and $P(D)f \in \mathcal{S}(I\!R^N)$.

Proof. (2.7.3) follows from the estimate

$$\begin{aligned}(1+|x|^2)^k |(D^\alpha f)(x)| &\leq (1+|x|^2)^{k+1} |(D^\alpha f)(x)| (1+|x|^2)^{-1} \\ &\leq C \, \| \, |\cdot|^{2(k+1)} D^\alpha f \, \|_{L^\infty(I\!R^N)} (1+|x|^2)^{-1} \, ;\end{aligned} \tag{2.7.4}$$

next, if P has degree r,

$$(1+|x|^2)^k |(D^\alpha (P(x) f(x))| \leq C \, \| \, |\cdot|^{2k+r} D^\alpha f \, \|_{L^\infty(I\!R^N)} \, , \tag{2.7.5}$$

$$(1+|x|^2)^k |(D^\alpha P(D)(f(x))| \leq C \, \| \, |\cdot|^{2k} D^{\alpha+\beta} f \, \|_{L^\infty(I\!R^N)} \, , \tag{2.7.6}$$

with $|\beta| = r$. This ends the proof of proposition 2.7.1. □

Obviously, $C_0^\infty(I\!R^N) \subset \mathcal{S}(I\!R^N)$ (if $\varphi \in C_0^\infty(I\!R^N)$, then $(\cdot)^\beta \, D^\alpha \varphi \in C_0^\infty(I\!R^N)$ and is bounded on $I\!R^N$, because $\mathrm{supp}(\varphi)$ is compact). However, functions in $\mathcal{S}(I\!R^N)$ are not required to have compact support; moreover, $S(I\!R^N) \subset L^p(I\!R^N)$ for all $p \in [1, +\infty]$ (see proposition 2.7.3 below). It is possible to prove that $\mathcal{S}(I\!R^N)$ is a Fréchet space, endowed with a locally convex metrizable complete topology defined by the sequence of seminorms

$$p_{k,m}(u) := \max_{|\alpha|\leq m} \sup_{x\in I\!R^N} \left| (1+|x|^2)^k \, D^\alpha u(x) \right| \, , \tag{2.7.7}$$

which are all bounded if $u \in \mathcal{S}(I\!R^N)$. For our purposes, it is sufficient to consider the notion of sequential convergence, as in $\mathcal{D}(I\!R^N)$ (see definition 2.4.1).

Definition 2.7.1 *We say that $f_r \to f$ in $\mathcal{S}(I\!R^N)$ if $D^\alpha f_r \to D^\alpha f$ uniformly on $I\!R^N$ for all $\alpha \in I\!N^N$, and if the sequences $p_{k,m}(f_r)$ are all bounded independently of r; that is, if for all k, $m \in I\!N$, there is $C_{k,m}$ such that $p_{k,m}(f_r) \le C_{k,m}$ for all $r \in I\!N$.*

This allows us to prove

Proposition 2.7.2 *The space $C_0^\infty(I\!R^N)$ is dense in $\mathcal{S}(I\!R^N)$.*

Proof. Let $\psi \in C_0^\infty(I\!R^N)$ be such that $0 \le \psi(x) \le 1$ for all $x \in I\!R^N$, $\psi(x) \equiv 1$ iff $|x| \le 1$, $\psi(x) \equiv 0$ iff $|x| \ge 2$. Given $f \in \mathcal{S}(I\!R^N)$ and $\varepsilon > 0$, set

$$f_\varepsilon(x) := \psi(\varepsilon x) f(x)\,. \tag{2.7.8}$$

Then $f_\varepsilon \in C_0^\infty(I\!R^N)$, with $\operatorname{supp} f_\varepsilon = B\left(0, \frac{2}{\varepsilon}\right)$. We now show that $f_\varepsilon \to f$ in $\mathcal{S}(I\!R^N)$. Indeed, for each $k \in I\!N$ and $\alpha \in I\!N$, by Leibniz' formula it follows that

$$(1+|x|^2)^k\, D^\alpha f_\varepsilon(x) = \sum_{\beta\le\alpha} \binom{\alpha}{\beta} (1+|x|^2)^k D^\beta f(x)\, D^{\alpha-\beta}\psi(\varepsilon x)\,, \tag{2.7.9}$$

so that, recalling the definition of $p_{k,m}$ in (2.7.7),

$$\begin{aligned} p_{k,m}(f_\varepsilon) &\le C\, p_{k,m}(f) \max_{|\alpha-\beta|\le m} \sup_{x\in I\!R^N} \left| D^{\alpha-\beta}\psi(\varepsilon x)\right| \\ &\le C\, p_{k,m}(f) \max_{|\alpha-\beta|\le m} \varepsilon^{|\alpha-\beta|} \sup_{1\le\varepsilon|x|\le 2} \left| (D^{\alpha-\beta}\psi)(\varepsilon x)\right| \\ &\le C_1\, p_{k,m}(f) \max_{|\alpha-\beta|\le m} \varepsilon^{|\alpha-\beta|}\,, \end{aligned} \tag{2.7.10}$$

with C_1 depending on ψ. This shows that each $p_{k,m}(f_\varepsilon)$ is uniformly bounded with respect to ε if (e.g.) $\varepsilon \le 1$. In the same way, from

$$\begin{aligned} &D^\alpha(f(x) - f_\varepsilon(x)) \\ &\quad= \sum_{\beta\le\alpha} \binom{\alpha}{\beta} D^\beta f(x)\, D^{\alpha-\beta}(1-\psi(\varepsilon x)) \\ &\quad= (1+|x|^2)^{-k} \sum_{\beta\le\alpha} \binom{\alpha}{\beta} (1+|x|^2)^k D^\beta f(x)\, D^{\alpha-\beta}(1-\psi(\varepsilon x)) \end{aligned} \tag{2.7.11}$$

we deduce that

$$|D^\alpha(f(x)-f_\varepsilon(x))| \le C \sum_{\beta\le\alpha} \left|(1+|x|^2)^k D^\beta f(x)\right| \left|D^{\alpha-\beta}(1-\psi(\varepsilon x))\right|. \tag{2.7.12}$$

By (2.7.3), each function $(1+|x|^2)^k D^\beta f(x)$ vanishes as $|x|\to+\infty$: thus, given $\eta>0$ there is $M>0$ such that $|(1+|x|^2)^k D^\beta f(x)| \le \eta$ if $|x|\ge M$. As above then, for all $\varepsilon \le \min\left\{1, \frac{1}{M}\right\}$, if $|x|\ge \frac{1}{\varepsilon}$ it follows that

$$\begin{aligned} |D^\alpha(f(x)-f_\varepsilon(x))| &\le C\,\eta \sup_{\varepsilon|x|\ge 1} \sum_{\beta\le\alpha} |D^{\alpha-\beta}(1-\psi(\varepsilon x))| \\ &\le C_1\,\eta\,; \end{aligned} \tag{2.7.13}$$

since the right side of (2.7.12) equals 0 if $|x|\le\frac{1}{\varepsilon}$, this shows that

$$(1+|x|^2)^k\, D^\alpha(f(x)-f_\varepsilon(x)) \to 0\,, \tag{2.7.14}$$

uniformly in x as $\varepsilon\to 0$. Hence, $f_\varepsilon \to f$ in $\mathcal{S}(I\!R^N)$ as claimed. □

Proposition 2.7.3 *$\mathcal{S}(I\!R^N)$ is dense in $L^p(I\!R^N)$ for all $p\in[1,+\infty[$.*

Proof. Since $C_0^\infty(I\!R^N)$ is dense in $L^p(I\!R^N)$, it is sufficient to prove that $\mathcal{S}(I\!R^N)\subset L^p(I\!R^N)$ (this is obvious if $p=\infty$, with

$$|\varphi|_\infty \le p_{0,0}(\varphi)\,; \tag{2.7.15}$$

however, $\mathcal{S}(I\!R^N)$ is not dense in $L^\infty(I\!R^N)$). If $1\le p<\infty$, we choose an integer $m\ge\frac{N}{2p}$ and, for $\varphi\in\mathcal{S}(I\!R^N)$, and, recalling (2.7.7), estimate

$$\begin{aligned} |\varphi|_p^p &= \int_{I\!R^N} (1+|x|^2)^{-mp}\,(1+|x|^2)^{mp}\,|\varphi(x)|^p\,\mathrm{d}x \\ &\le \sup_{x\in I\!R^N} \Big((1+|x|^2)^m\,|\varphi(x)|\Big)^p \int_{I\!R^N} (1+|x|^2)^{-mp}\,\mathrm{d}x \\ &\le (p_{m,0}(\varphi))^p \int_0^{+\infty} \frac{r^{N-1}}{(1+r^2)^{mp}}\,\mathrm{d}r \;\le (p_{m,0}(\varphi))^p\,M^p\,, \end{aligned} \tag{2.7.16}$$

where the constant M depends on φ, N, and p; the last integral of (2.7.16) converges, because of the choice $m>\frac{N}{2p}$.For future reference, we note explicitly that from (2.7.16) it follows that

$$|\varphi|_p \le M\,p_{m,0}(\varphi)\,, \tag{2.7.17}$$

which also shows that convergence in $\mathcal{S}$ implies convergence in $L^p(I\!R^N)$. □

2.7.2 Tempered Distributions.

1. Just as we have defined the space $\mathcal{D}'(I\!R^N)$ of distributions as the space of all linear, sequentially continuous functionals on $\mathcal{D}(I\!R^N)$, we define the space $\mathcal{S}'(I\!R^N)$ of tempered distributions as the space of all linear, sequentially continuous functionals on $\mathcal{S}(I\!R^N)$, and we say that $F_m \to F$ in $S'(I\!R^N)$ if for all $\varphi \in \mathcal{S}(I\!R^N)$, $F_m(\varphi) \to F(\varphi)$ in $I\!R$, in the sense of definition 2.7.1. As for general distributions, we often denote the duality between $\mathcal{S}'(I\!R^N)$ and $\mathcal{S}(I\!R^N)$ by $\langle\cdot,\cdot\rangle_{\mathcal{S}}$.

Example 2.7.1 *Let $f \in L^p(I\!R^N)$, $1 \le p \le +\infty$. The distribution T_f defined by f via* (2.5.4)*, which now reads*

$$\langle T_f, \varphi\rangle_{\mathcal{S}} = \int_{I\!R^N} f(x)\,\varphi(x)\,\mathrm{d}x, \qquad \varphi \in \mathcal{S}(I\!R^N)\,, \tag{2.7.18}$$

is a tempered distribution. More generally, if $f : I\!R^N \to I\!R$ is a measurable function such that there are $C > 0$ and $r \in I\!N$ such that, for all $x \in I\!R^N$,

$$|f(x)| \le C\,(1+|x|^2)^r, \tag{2.7.19}$$

then f also defines a tempered distribution T_f by means of the same formula (2.7.18).

Proof. Each T_f is obviously linear; if $f \in L^p(I\!R^N)$, the sequential continuity of T_f follows from (2.7.17) and (2.7.15). If f only satisfies (2.7.19), we estimate

$$\begin{aligned} |\langle T_f, \varphi\rangle_{\mathcal{S}}| &\le \int_{I\!R^N} \frac{|f(x)|(1+|x|^2)^{r+N}}{(1+|x|^2)^{r+N}} |\varphi(x)|\,\mathrm{d}x \\ &\le C \sup_{x\in I\!R^N} \Big((1+|x|^2)^{r+N}|\varphi(x)|\Big) \int_{I\!R^N} \frac{1}{(1+|x|^2)^N}\,\mathrm{d}x \\ &\le C\,M\,p_{r+N,0}(\varphi). \end{aligned} \tag{2.7.20}$$

This proves that $T_f \in \mathcal{S}'(I\!R^N)$ in either case. □

In particular, taking $f = 1$ (which is in $L^\infty(I\!R^N)$, and satisfies (2.7.19)) yields that the map $T_1 : \mathcal{S}(I\!R^N) \to I\!R$, defined by

$$\langle T_1, \varphi\rangle_{\mathcal{S}} = \int_{I\!R^N} \varphi(x)\,\mathrm{d}x, \tag{2.7.21}$$

is in $\mathcal{S}'(\mathbb{R}^N)$.

2. Perhaps the most important example of tempered distributions is given by the family of the Dirac distributions δ_α. More precisely:

Theorem 2.7.1 *For $\alpha \in \mathbb{R}^N$, define $\tilde{\delta}_\alpha : \mathcal{S}(\mathbb{R}^N) \to \mathbb{R}$ by*

$$\langle \tilde{\delta}_\alpha, \varphi \rangle_{\mathcal{S}} = \varphi(\alpha), \qquad \forall\, \varphi \in \mathcal{S}(\mathbb{R}^N). \tag{2.7.22}$$

Then each $\tilde{\delta}_\alpha$ is in $\mathcal{S}'(\mathbb{R}^N)$, and is an extension to $\mathcal{S}(\mathbb{R}^N)$ of the Dirac δ-distributions $\delta_\alpha \in \mathcal{D}'(\mathbb{R}^N)$.

Proof. It is sufficient to show that the maps $\varphi \mapsto \varphi(\alpha)$ are sequentially continuous on $S(\mathbb{R}^N)$. But, according to definition 2.7.1, the convergence of a sequence $(\varphi_k)_{k\in\mathbb{N}}$ to φ in $\mathcal{S}(\mathbb{R}^N)$ implies, in particular, that $\varphi_k \to \varphi$ uniformly on $\mathbb{R}^N$; hence, $\varphi_k(\alpha) \to \varphi(\alpha)$. This shows that each $\tilde{\delta}_\alpha \in \mathcal{S}'(\mathbb{R}^N)$. On $\mathcal{D}(\mathbb{R}^N)$, $\tilde{\delta}_\alpha$ obviously coincides with the corresponding Dirac δ-distribution; since $\mathcal{D}(\mathbb{R}^N)$ is dense in $\mathcal{S}(\mathbb{R}^N)$, this shows that $\tilde{\delta}_\alpha$ is an extension of δ_α to $\mathcal{S}(\mathbb{R}^N)$. $\square$

3. Finally, we remark that, as the example of the Dirac δ-distributions shows, tempered distributions are indeed distributions; that is, $\mathcal{S}'(\mathbb{R}^N) \subset \mathcal{D}'(\mathbb{R}^N)$. This inclusion is dense, since $\mathcal{D}(\mathbb{R}^N)$ is dense in $\mathcal{S}(\mathbb{R}^N)$; indeed, sequential convergence in $\mathcal{D}(\mathbb{R}^N)$ implies sequential convergence in $\mathcal{S}(\mathbb{R}^N)$, since the factors $x \mapsto (1+|x|^2)^k$ are bounded on any compact set of $\mathbb{R}^N$. In other words, if $\varphi_m \to \varphi$ in $\mathcal{D}(\mathbb{R}^N)$ and $T \in \mathcal{S}'(\mathbb{R}^N) \subset D'(\mathbb{R}^N)$, then $\langle T, \varphi_m - \varphi \rangle_{\mathcal{D}} \to 0$, because $\varphi_m \to \varphi$ also in $\mathcal{S}(\mathbb{R}^N)$. From this, it also follows that distributional derivatives of tempered distributions are again tempered distributions. On the other hand, there are distributions that are not tempered, as we see from

Proposition 2.7.4 *The distribution $T_f \in \mathcal{D}'(\mathbb{R})$ defined by the locally integrable function $f(x) = \mathrm{e}^x$ is not in $\mathcal{S}'(\mathbb{R})$.*

Proof. Assuming the contrary, consider the non-negative function $\varphi \in C^\infty(\mathbb{R})$ defined by

$$\varphi(x) := \begin{cases} 0 & \text{if} \quad x \le -1\,, \\ \psi(x) & \text{if} \quad -1 < x < 0\,, \\ \mathrm{e}^{-x} & \text{if} \quad x \ge 0\,, \end{cases} \tag{2.7.23}$$

where $\psi \in C^\infty([-2,1])$ with $\psi(-1) = 0 = \varphi(-1)$ and $\psi(0) = 1 = \varphi(0)$. Then, $\varphi \in \mathcal{S}(I\!R)$. Since $\varphi \geq 0$,

$$T_f(\varphi) = \int_{-\infty}^{+\infty} f(x)\,\varphi(x)\,\mathrm{d}x \geq \int_0^{+\infty} \mathrm{e}^x\,\mathrm{e}^{-x}\,\mathrm{d}x = +\infty\,; \tag{2.7.24}$$

thus, $T_f(\varphi) = +\infty$, contradicting the fact that $T_f(\varphi)$ is a finite real number if $T_f \in \mathcal{S}'(I\!R)$. $\square$

Chapter 3

The Fourier Transform.

In this chapter we briefly recall the basic results on the Fourier transform in $L^1(\mathbb{R}^N)$, $L^2(\mathbb{R}^N)$, and $\mathcal{S}(\mathbb{R}^N)$, and its inverse in $L^2(\mathbb{R}^N)$ and $\mathcal{S}(\mathbb{R}^N)$. Most of the material of this chapter is taken from chapter 9 of Rudin, [6].

3.1 Fourier Transform in $L^1(\mathbb{R}^N)$.

We start with the definition of the Fourier transform of functions in $L^1(\mathbb{R}^N)$. If $\varphi \in L^1(\mathbb{R}^N)$, its Fourier transform is the function $\hat{\varphi} : \mathbb{R}^N \to \mathbb{C}$ defined by

$$\hat{\varphi}(y) := \tfrac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \mathrm{e}^{-ix\cdot y} \varphi(x)\, \mathrm{d}x\,, \tag{3.1.1}$$

where $i^2 := -1$; we set $\tfrac{1}{(2\pi)^{N/2}} =: c_N$. A first example is

Example 3.1.1 *For $m \in \mathbb{N}$, $m \geq 1$, let χ_m denote the characteristic function of the interval $[-m, m]$. Then $\chi_m \in L^1(\mathbb{R})$, and*

$$\hat{\chi}_m(y) = \begin{cases} \sqrt{\dfrac{2}{\pi}}\, \dfrac{\sin(my)}{y} & \text{if} \quad y \neq 0\,, \\ \sqrt{\dfrac{2}{\pi}}\, m & \text{if} \quad y = 0\,. \end{cases} \tag{3.1.2}$$

Proof. If $y \neq 0$,

$$\begin{aligned}\hat{\chi}_m(y) &= c_1 \int_{-m}^{m} \mathrm{e}^{-ixy}\,\mathrm{d}x \;=\; c_1 \left[\frac{\mathrm{e}^{-ixy}}{-iy}\right]_{x=-m}^{x=m} \\ &= c_1\,\frac{\mathrm{e}^{imy} - \mathrm{e}^{-imy}}{iy} \;=\; \sqrt{\frac{2}{\pi}}\,\frac{\sin my}{y}\,,\end{aligned} \tag{3.1.3}$$

while if $y = 0$,

$$\hat{\chi}_m(0) = c_1 \int_{-m}^{m} 1\,\mathrm{d}x = \sqrt{\frac{2}{\pi}}\,m\,. \tag{3.1.4}$$

Thus, (3.1.2) follows. Note that $\hat{\chi}_m$ is continuous at $y = 0$. □

The following results are of immediate proof.

Proposition 3.1.1 *Let $f \in L^1(\mathbb{R}^N)$, $a \in \mathbb{R}^N$ and $\alpha \neq 0$. Then:*

1) If $f_a(x) := f(x-a)$, then $\hat{f}_a(y) = \mathrm{e}^{-ia\cdot y}\hat{f}(y)$;

2) If $f_\alpha(x) := f\left(\frac{x}{\alpha}\right)$, then $\hat{f}_\alpha(y) = \alpha^N \hat{f}(\alpha\, y)$.

We usually write $\mathcal{F}$ to denote the map $f \mapsto \hat{f}$, which is linear; that is, we set $\hat{f} = \mathcal{F}f$. $\mathcal{F}$ is defined on all of $L^1(\mathbb{R}^N)$; to characterize its range, we show that if $f \in L^1(\mathbb{R}^N)$, then $\hat{f}$ is continuous, but not necessarily integrable.

Theorem 3.1.1 *Let $f \in L^1(\mathbb{R}^N)$. Then $\hat{f} \in \mathcal{UCB}(\mathbb{R}^N)$ (that is, $\hat{f}$ is uniformly continuous and bounded on $\mathbb{R}^N$); however, $\hat{f}$ is not necessarily in $L^1(\mathbb{R}^N)$. The map $\mathcal{F}$ is continuous from $L^1(\mathbb{R}^N)$ into $\mathcal{UCB}(\mathbb{R})$. The range space $\mathcal{F}\left(L^1(\mathbb{R}^N)\right)$ is a proper subspaces of $\mathcal{UCB}(\mathbb{R}^N)$.*

Proof. 1) Let $f \in L^1(\mathbb{R}^N)$. Then, for all $y \in \mathbb{R}^N$,

$$|\hat{f}(y)| \leq c_N \int_{\mathbb{R}^N} |\mathrm{e}^{-iy\cdot x}|\,|f(x)|\,\mathrm{d}x \leq |f|_1\,, \tag{3.1.5}$$

which means that $\hat{f}$ is bounded.

2) Fix $\varepsilon > 0$. Since $f \in L^1(\mathbb{R}^N)$, by the absolute continuity of the Lebesgue integral we can find $R > 0$ such that

$$\int_{|y|\geq R} |f(y)|dy \leq \frac{1}{4c_N}\,\varepsilon\,. \tag{3.1.6}$$

With R so chosen, we choose δ such that

$$\delta \leq \frac{1}{4c_N R\,|f|_1}\,\varepsilon\,, \tag{3.1.7}$$

and assume that $|x-\bar{x}| \leq \delta$. Recalling that $|\mathrm{e}^{-iz\cdot y}-1| \leq 2|z\cdot y|$ for all z, $y \in \mathbb{R}^N$, we compute that

$$\begin{aligned}
|\hat{f}(x)-\hat{f}(\bar{x})| &\leq c_N \int_{\mathbb{R}^N} |\mathrm{e}^{-ix\cdot y}-\mathrm{e}^{-i\bar{x}\cdot y}||f(y)|\,\mathrm{d}y \\
&= c_N \int_{\mathbb{R}^N} |\mathrm{e}^{-i\bar{x}\cdot y}|\,|\mathrm{e}^{-i(x-\bar{x})\cdot y}-1|\,|f(y)|\,\mathrm{d}y \\
&\leq 2\,c_N \int_{|y|\geq R} |f(y)|\,\mathrm{d}y \\
&\qquad + 2\,c_N \int_{|y|\leq R} |x-\bar{x}||y||f(y)|\,\mathrm{d}y \\
&\leq 2\,c_N \left(\int_{|y|\geq R} |f(y)|\,\mathrm{d}y + \delta R \int_{\mathbb{R}^N} |f(y)|\,\mathrm{d}y\right) \\
&\leq 2\,c_N \left(\tfrac{1}{4c_N}\,\varepsilon + \delta R|f|_1\right)\,,
\end{aligned} \tag{3.1.8}$$

having used (3.1.6). By (3.1.7) we conclude that

$$|\hat{f}(x)-\hat{f}(\bar{x})| \leq \varepsilon \qquad \text{whenever} \quad |x-\bar{x}| \leq \delta\,. \tag{3.1.9}$$

This proves the uniform continuity of $\hat{f}$. In addition, (3.1.5) shows that $\mathcal{F}$ is continuous from $L^1(\mathbb{R}^N)$ into $C(\overline{\mathbb{R}})$. In fact, $\mathcal{F}$ is a contraction on $L^1(\mathbb{R}^N)$. Indeed, from (3.1.5), and recalling that $c_N < 1$, we see that if f, $g \in L^1(\mathbb{R}^N)$,

$$\begin{aligned}
\sup_{y\in\mathbb{R}^N} |\hat{f}(y)-\hat{g}(y)| &\leq \sup_{y\in\mathbb{R}^N} c_N \int_{\mathbb{R}^N} |f(x)-g(x)|\,\mathrm{d}x \\
&= c_N\,|f-g|_1\,.
\end{aligned} \tag{3.1.10}$$

3) We show that the function $\hat{\chi}_1$, transform of the characteristic function χ_1 of the interval $[-1,1]$, is not in $L^1(\mathbb{R})$ (while, in accord with part (1), $\hat{\chi}_1$ is uniformly

continuous on $I\!R$). To show this, recall from (3.1.2) that, if $y \neq 0$,

$$\hat{\chi}_1(y) = \sqrt{\frac{2}{\pi}} \frac{\sin y}{y} \,. \tag{3.1.11}$$

Then, for $n \in I\!N$,

$$\begin{aligned} I_n &:= \int_{-(n+1)\pi}^{(n+1)\pi} \left| \frac{\sin y}{y} \right| \mathrm{d}y \geq 2 \sum_{k=0}^{n} \int_{(k+1/6)\pi}^{(k+5/6)\pi} \left| \frac{\sin y}{y} \right| \mathrm{d}y \\ &\geq 2 \sum_{k=0}^{n} \int_{(k+1/6)\pi}^{(k+5/6)\pi} \frac{1}{2y} \, \mathrm{d}y = \sum_{k=0}^{n} \ln \left(\frac{6k+5}{6k+1} \right) \\ &\geq \sum_{k=0}^{n} \ln \left(1 + \frac{4}{6k+1} \right) =: S_n \,. \end{aligned} \tag{3.1.12}$$

Since

$$\lim_{n\to\infty} S_n = \sum_{k=0}^{\infty} \ln \left(1 + \frac{4}{6k+1} \right) = +\infty \,, \tag{3.1.13}$$

also $I_n \to +\infty$, so that $\hat{\chi}_1 \notin L^1(I\!R)$.

4) Finally, we prove that there is *no* function $f \in L^1(I\!R)$ such that $\hat{f}(y) \equiv 1$. Assuming otherwise, by the density of $C_0^\infty(I\!R)$ in $L^1(I\!R)$, we can find a function $\varphi \in C_0^\infty(I\!R)$ such that

$$c_1 \, |f - \varphi|_1 \leq \tfrac{1}{3} \,. \tag{3.1.14}$$

Then for each $y \in I\!R$,

$$\begin{aligned} 1 = \hat{f}(y) &= c_1 \int_{-\infty}^{+\infty} \mathrm{e}^{-ixy} f(x) \, \mathrm{d}x \\ &= c_1 \int_{-\infty}^{+\infty} \mathrm{e}^{-ixy} (f(x) - \varphi(x)) \, \mathrm{d}x \\ &\qquad + c_1 \int_{-\infty}^{+\infty} \mathrm{e}^{-ixy} \varphi(x) \, \mathrm{d}x \\ &=: I(y) + J(y) \,. \end{aligned} \tag{3.1.15}$$

At first, by (3.1.14), for all $y \in I\!R$,

$$|I(y)| \leq c_1 \, |f - \varphi|_1 \leq \tfrac{1}{3} \,; \tag{3.1.16}$$

to estimate $J(y)$, we integrate by parts: since φ has compact support, we find that, if $y \neq 0$,

$$\begin{aligned} J(y) &= \frac{c_1}{-iy} \int_{-\infty}^{+\infty} \left(\frac{d}{\mathrm{d}x} \mathrm{e}^{-ixy} \right) \varphi(x) \, \mathrm{d}x \\ &= \frac{c_1}{iy} \int_{-\infty}^{+\infty} \mathrm{e}^{-ixy} \varphi'(x) \, \mathrm{d}x \, . \end{aligned} \tag{3.1.17}$$

Thus, if $y > 0$,

$$|J(y)| \leq \frac{c_1}{y} \int_{-\infty}^{+\infty} \varphi'(x) \, \mathrm{d}x \leq \frac{C_\varphi}{y} \; ; \tag{3.1.18}$$

choosing then y so large that

$$\frac{C_\varphi}{y} \leq \frac{1}{3} \, , \tag{3.1.19}$$

(3.1.15), (3.1.16) and (3.1.18) yield a contradiction. This concludes the proof of theorem 3.1.1. □

3.2 The Heat Kernel.

The following result is fundamental for the sequel, and is at the basis of the solution theory of the heat equation.

Proposition 3.2.1 *Define* $E : \mathbb{R}^N \to \mathbb{R}_{>0}$ *by*

$$E(x) := \mathrm{e}^{-|x|^2/2} \, . \tag{3.2.1}$$

Then $E \in L^1(\mathbb{R}^N)$*, and* $\hat{E} = E$*; that is,* E *is an invariant for* $\mathcal{F}$*. Moreover,*

$$E(0) = \hat{E}(0) = c_N \int_{\mathbb{R}^N} \mathrm{e}^{-|x|^2/2} \, \mathrm{d}x = 1 \, . \tag{3.2.2}$$

Finally, for all $y \in \mathbb{R}^N$ *and* $t > 0$*,*

$$E\left(\tfrac{y}{\sqrt{2t}} \right) = c_N \left(\sqrt{2t} \right)^N \int_{\mathbb{R}^N} \mathrm{e}^{iy \cdot z - t|z|^2} \, \mathrm{d}z \, . \tag{3.2.3}$$

Proof. Let first $N = 1$: then $E(x) = \mathrm{e}^{-x^2/2}$, and solves the Cauchy problem

$$\begin{cases} h' + x\,h = 0\,, \\ h(0) = 1\,. \end{cases} \tag{3.2.4}$$

By definition,

$$\hat{E}(y) = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-ixy} \mathrm{e}^{-x^2/2}\,\mathrm{d}x\,, \tag{3.2.5}$$

and we note that we can differentiate with respect to y within this integral, since

$$\frac{\partial}{\partial y}\left(\mathrm{e}^{-ixy}\mathrm{e}^{-x^2/2}\right) = (-ix)\mathrm{e}^{-ixy}\mathrm{e}^{-x^2/2}\,, \tag{3.2.6}$$

the right side of (3.2.6) satisfies the estimate

$$|(-ix)\mathrm{e}^{-ixy}\mathrm{e}^{-x^2/2}| \le |x| E(x)\,, \tag{3.2.7}$$

and the function $x \mapsto |x|\,E(x)$ is in $L^1(I\!R)$. Thus, by (3.2.4), and integrating by parts,

$$\begin{aligned} \hat{E}'(y) &= \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-ixy}\,(-ix)E(x)\,\mathrm{d}x \\ &= \tfrac{i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-ixy} E'(x)\,\mathrm{d}x \\ &= \tfrac{-i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (-iy)\,\mathrm{e}^{-ixy} E(x)\,\mathrm{d}x \\ &= \tfrac{-y}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-ixy} E(x)\,\mathrm{d}x \\ &= -y\,\hat{E}(y)\,. \end{aligned} \tag{3.2.8}$$

Since also

$$\hat{E}(0) = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2}\,\mathrm{d}x = 1\,, \tag{3.2.9}$$

we conclude that $\hat{E}$ is also a solution of (3.2.4); hence, $\hat{E} = E$. This proves that $\hat{E} = E$ when $N = 1$. When $N > 1$, letting $h(r) := \mathrm{e}^{-r^2/2}$, $r \in I\!R$, we decompose

$$E(x) = \exp\left(-\frac{1}{2}\sum_{j=1}^{N} x_j^2\right) = \prod_{j=1}^{N} \mathrm{e}^{-x_j^2/2} = \prod_{j=1}^{N} h(x_j)\,, \tag{3.2.10}$$

so that

$$\begin{aligned}\hat{E}(y) &= \frac{1}{(2\pi)^{n/2}} \prod_{j=1}^{N} \int_{-\infty}^{+\infty} \mathrm{e}^{-iy_j x_j} \mathrm{e}^{-x_j^2/2} \,\mathrm{d}x_j \\ &= \prod_{j=1}^{N} \hat{h}(y_j) = \prod_{j=1}^{N} h(y_j) = E(y) \,.\end{aligned} \tag{3.2.11}$$

This proves that $\hat{E} = E$ for all N. Next, we compute directly that, by (3.2.9),

$$\begin{aligned}E(0) &= \hat{E}(0) = \frac{1}{(2\pi)^{N/2}} \int_{I\!R^N} \mathrm{e}^{-|x|^2/2} \,\mathrm{d}x \\ &= \prod_{j=1}^{N} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x_j^2/2} \,\mathrm{d}x_j \\ &= 1 \,,\end{aligned} \tag{3.2.12}$$

which proves (3.2.2); note that this implies the well-known identity

$$\int_{I\!R^N} \mathrm{e}^{-|x|^2/2} \,\mathrm{d}x = (2\pi)^{N/2} \,. \tag{3.2.13}$$

As for (3.2.3), we compute

$$\begin{aligned}c_N \int_{I\!R^N} \mathrm{e}^{iy\cdot z - t|z|^2} \,\mathrm{d}z &= c_N \int_{I\!R^N} \mathrm{e}^{-iy\cdot(-z)} \mathrm{e}^{-t|-z|^2} \,\mathrm{d}z \\ &= c_N \int_{I\!R^N} \mathrm{e}^{-iy\cdot z} \mathrm{e}^{-t|z|^2} \,\mathrm{d}z \\ &= \frac{c_N}{(\sqrt{2t})^N} \int_{I\!R^N} \mathrm{e}^{-iy\cdot(u/\sqrt{2t})} \mathrm{e}^{-|u|^2/2} \,\mathrm{d}u \,,\end{aligned} \tag{3.2.14}$$

so that the right side of (3.2.3) equals

$$c_N \int_{I\!R^N} \mathrm{e}^{-i(y/\sqrt{2t})\cdot u} \, E(u) \,\mathrm{d}u = \hat{E}\left(\tfrac{y}{\sqrt{2t}}\right) = E\left(\tfrac{y}{\sqrt{2t}}\right) . \tag{3.2.15}$$

Thus, (3.2.3) holds, and the proof of proposition 3.2.1 is complete. □

In particular, for $t = \frac{1}{2}$ (3.2.3) yields the identity

$$\begin{aligned} E(y) &= c_N \int_{\mathbb{R}^N} \mathrm{e}^{iy\cdot x - |x|^2/2}\,\mathrm{d}x \\ &= c_N \int_{\mathbb{R}^N} \mathrm{e}^{iy\cdot x} E(x)\,\mathrm{d}x \\ &= c_N \int_{\mathbb{R}^N} \mathrm{e}^{iy\cdot x} \hat{E}(x)\,\mathrm{d}x\,; \end{aligned} \tag{3.2.16}$$

as we shall see in §3.3 below, this means that E also coincides with its own inverse Fourier transform. For $t > 0$, we now set [1]

$$E_{(t)}(x) := \frac{c_N}{(\sqrt{2t})^N}\,\mathrm{e}^{-|x|^2/4t} = \frac{1}{(\sqrt{4\pi t})^N}\,E\left(\frac{x}{\sqrt{2t}}\right)\,. \tag{3.2.17}$$

For each $t > 0$, the function $x \mapsto E_{(t)}(x)$ is obviously in $L^q(\mathbb{R}^N)$ for all $q \in [1, +\infty[$. The function $(t, x) \mapsto E_{(t)}(x) =: H(t, x)$ is called the HEAT KERNEL, because, as it is immediate to verify, H satisfies, on the set $\mathbb{R}_{>0} \times \mathbb{R}^N$, the heat equation

$$H_t - \Delta H = 0\,. \tag{3.2.18}$$

We list the most important properties of the heat kernel.

Proposition 3.2.2 *Let $E_{(t)}$ be the heat kernel defined in* (3.2.17)*. Then:*

1) For all $t > 0$,

$$\int_{\mathbb{R}^N} E_{(t)}(x)\,\mathrm{d}x = 1\,; \tag{3.2.19}$$

$$\hat{E}_{(t)}(y) = c_N\,\mathrm{e}^{-t\,|y|^2}\,; \tag{3.2.20}$$

$$E_{(t)}(x) = c_N \int_{\mathbb{R}^N} \mathrm{e}^{ix\cdot y} E\left(\sqrt{2t}\,y\right)\,\mathrm{d}y\,. \tag{3.2.21}$$

[1] We adopt the somewhat cumbersome notation $E_{(t)}$, instead of the usual one E_t, in order not to confuse the latter with the partial derivative of E, considered as a function of the two variables t and x, with respect to t (compare to (3.2.18)).

2) If $f \in L^1(I\!R^N)$*,*

$$[f * E_{(t)}](x) = \int_{I\!R^N} \mathrm{e}^{-t|y|^2 + ix\cdot y} \hat{f}(y)\,\mathrm{d}y\,. \tag{3.2.22}$$

3) If $f \in L^p(I\!R^N)$*,* $1 \le p < +\infty$*, then* $f * E_{(t)} \in L^p(I\!R^N)$*, and*

$$f * E_{(t)} \to f \qquad \text{in } \ L^p(I\!R^N) \qquad \text{as } \ t \to 0\,. \tag{3.2.23}$$

4) If $f \in L^\infty(I\!R^N)$ *and is continuous at a point* $x \in I\!R^N$*, then*

$$\lim_{t\to 0}[f * E_{(t)}](x) = f(x)\,. \tag{3.2.24}$$

Proof. We follow chapter 9 of Rudin, [6]. (3.2.19) follows immediately from (3.2.2), via the change of variable $x = \sqrt{2t}\,z$; likewise, (3.2.20) follows from (3.2.17) and part (2) of proposition 3.1.1, with $\alpha = \sqrt{2t}$, recalling that $\hat{E} = E$. From (3.2.17) and (3.2.3) it also follows that

$$\begin{aligned} E_{(t)}(x) &= \frac{c_N}{(\sqrt{2t})^N} E(x/\sqrt{2t}) \\ &= c_N \int_{I\!R^N} \mathrm{e}^{ix\cdot y - t|y|^2}\,\mathrm{d}y \\ &= c_N \int_{I\!R^N} \mathrm{e}^{ix\cdot y} E\left(\sqrt{2t}\,y\right)\,\mathrm{d}y\,, \end{aligned} \tag{3.2.25}$$

which proves (3.2.21). Using this, we compute that

$$\begin{aligned} [f * E_{(t)}](x) &= \int_{I\!R^N} f(x-y) E_{(t)}(y)\,\mathrm{d}y \\ &= c_N \int_{I\!R^N} f(x-y) \int_{I\!R^N} \mathrm{e}^{iy\cdot z} E\left(\sqrt{2t}\,z\right)\,\mathrm{d}z\,\mathrm{d}y \\ &= c_N \int_{I\!R^N} E\left(\sqrt{2t}\,z\right) \int_{I\!R^N} f(x-y)\mathrm{e}^{iy\cdot z}\,\mathrm{d}y\,\mathrm{d}z \\ &= c_N \int_{I\!R^N} E\left(\sqrt{2t}\,z\right) \mathrm{e}^{ix\cdot z} \int_{I\!R^N} f(u)\mathrm{e}^{-iu\cdot z}\,\mathrm{d}u\,\mathrm{d}z \\ &= \int_{I\!R^N} E\left(\sqrt{2t}\,z\right) \mathrm{e}^{ix\cdot z} \hat{f}(z)\,\mathrm{d}z \\ &= \int_{I\!R^N} \mathrm{e}^{-t|z|^2 + ix\cdot z} \hat{f}(z)\,\mathrm{d}z\,, \end{aligned} \tag{3.2.26}$$

which is (3.2.22). Next, let q be the conjugate index of p (that is, $\frac{1}{p}+\frac{1}{q}=1$). By (3.2.19) and Hölder's inequality,

$$\begin{aligned} |[f*E_{(t)}](x)| &\leq \int_{\mathbb{R}^N} |f(x-y)| \left(E_{(t)}(y)\right)^{1/p} \left(E_{(t)}(y)\right)^{1/q} \,\mathrm{d}y \\ &\leq \left(\int_{\mathbb{R}^N} |f(x-y)|^p E_{(t)}(y)\,\mathrm{d}y\right)^{1/p} , \end{aligned} \tag{3.2.27}$$

so that, again by (3.2.19),

$$\begin{aligned} \int_{\mathbb{R}^N} |[f*E_{(t)}](x)|^p \,\mathrm{d}x &\leq \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |f(x-y)|^p E_{(t)}(y)\,\mathrm{d}y\,\mathrm{d}x \\ &\leq \int_{\mathbb{R}^N} E_{(t)}(y) \int_{\mathbb{R}^N} |f(x-y)|^p \,\mathrm{d}x\,\mathrm{d}y \\ &\leq |f|_p^p ; \end{aligned} \tag{3.2.28}$$

thus, $f*E_{(t)} \in L^p(\mathbb{R}^N)$. To prove (3.2.23), we first note that (3.2.19) allows us to write

$$[f*E_{(t)}](x) - f(x) = \int_{\mathbb{R}^N} f(x-y)E_{(t)}(y)\,\mathrm{d}y - \int_{\mathbb{R}^N} f(x)E_{(t)}(y)\,\mathrm{d}y ; \tag{3.2.29}$$

thus, proceeding as above,

$$\begin{aligned} \Big|[f*E_{(t)}](x) - f(x)\Big| &\leq \int_{\mathbb{R}^N} |f(x-y)-f(x)|\, E_{(t)}(y)\,\mathrm{d}y \\ &= \left(\int_{\mathbb{R}^N} |f(x-y)-f(x)|^p \left(E_{(t)}(y)\right) \,\mathrm{d}y\right)^{1/p} , \end{aligned} \tag{3.2.30}$$

from which we deduce that

$$\begin{aligned} &\int_{\mathbb{R}^N} \Big|[f*E_{(t)}](x) - f(x)\Big|^p \,\mathrm{d}x \\ &\leq \int_{\mathbb{R}^N} E_{(t)}(y) \int_{\mathbb{R}^N} |f(x-y)-f(x)|^p \,\mathrm{d}x\,\mathrm{d}y \\ &\leq \int_{\mathbb{R}^N} E_{(t)}(y)\, |f(\cdot - y) - f|_p^p \,\mathrm{d}y . \end{aligned} \tag{3.2.31}$$

Since traslations are uniformly continuous in $L^p(I\!\!R^N)$, given $\varepsilon > 0$ there is $\delta > 0$ such that if $|y| \le \delta$,

$$g(y) := |f(\cdot - y) - f|_p^p \le \varepsilon\,; \tag{3.2.32}$$

with this δ now fixed, we split the last integral in (3.2.31) and write

$$\begin{aligned}
&\int_{I\!\!R^N} \left|\left(f * E_{(t)}\right)(x) - f(x)\right|^p \,\mathrm{d}x \\
&\le \int_{|y|\le\delta} E_{(t)}(y) g(y)\,\mathrm{d}y + \int_{|y|\ge\delta} E_{(t)}(y) g(y)\,\mathrm{d}y \\
&=: I_t + J_t\,.
\end{aligned} \tag{3.2.33}$$

Because of (3.2.19) and (3.2.32),

$$0 \le I_t \le \varepsilon \int_{|y|\le\delta} E_{(t)}(y)\,\mathrm{d}y = \varepsilon\,; \tag{3.2.34}$$

since g is also bounded on $I\!\!R^N$, we can estimate

$$\begin{aligned}
0 \le J_t \quad &\le \quad \sup_{y\in I\!\!R^N} \int_{|y|\ge\delta} E_{(t)}(y)\,\mathrm{d}y \\
&= \quad \sup_{y\in I\!\!R^N} \frac{c_N}{(\sqrt{2t})^N} \int_{|y|\ge\delta} \mathrm{e}^{-|y|^2/4t}\,\mathrm{d}y \\
&= \quad c_N \sup_{y\in I\!\!R^N} \int_{|z|\ge\delta/\sqrt{2t}} \mathrm{e}^{-|z|^2/2}\,\mathrm{d}z\,.
\end{aligned} \tag{3.2.35}$$

The last integral vanishes as $t \to 0$; together with (3.2.34), this implies (3.2.23). Finally, to prove (3.2.24) we start as in (3.2.30) with

$$\begin{aligned}
\left|\left(f * E_{(t)}\right)(x) - f(x)\right| \quad &\le \quad \int_{I\!\!R^N} |f(x-y) - f(x)|\, E_{(t)}(y)\,\mathrm{d}y \\
&= \quad \tfrac{1}{(\sqrt{2t})^N} \int_{I\!\!R^N} |f(x-y) - f(x)| E\left(\tfrac{y}{\sqrt{2t}}\right)\,\mathrm{d}y \\
&= \quad \int_{I\!\!R^N} |f(x - \sqrt{2t}\,z) - f(x)| E(z)\,\mathrm{d}z\,,
\end{aligned} \tag{3.2.36}$$

having used (3.2.17). Since f is continuous at x, the last integrand in (3.2.36) converges pointwise to 0 as $t \to 0$; since $f \in L^\infty(I\!\!R^N)$, it is also bounded by the function $2\,|f|_\infty\, E(\cdot)$, which is in $L^1(I\!\!R^N)$. Hence, (3.2.24) follows by the dominated convergence theorem. This completes the proof of proposition 3.2.2. $\square$

3.3 Towards $\mathcal{F}^{-1}$.

1. Our goal is now to define an inverse of the map $\mathcal{F}$, that is, to define the inverse Fourier transform. The fact that the range of $\mathcal{F}$ is not even a subspace of $L^1(\mathbb{R}^N)$ creates unsurmountable difficulties, and will lead us to eventually redefine $\mathcal{F}$ in a suitable way on the space $L^2(\mathbb{R}^N)$ (note that since $\mathbb{R}^N$ has infinite measure, $L^2(\mathbb{R}^N)$ is also not a subspace of $L^1(\mathbb{R}^N)$). Thus, for the moment we restrict our attention to the space of those integrable functions whose Fourier transform is also integrable, that is, the space

$$\Lambda^1(\mathbb{R}^N) := \{f \in L^1(\mathbb{R}^N) \mid \hat{f} \in L^1(\mathbb{R}^N)\} . \tag{3.3.1}$$

This space is not empty: from proposition 3.2.1 we know that $E \in \Lambda^1(\mathbb{R}^N)$.

Given $\varphi \in L^1(\mathbb{R}^N)$, in analogy to (3.1.1) we define the function $\check{\varphi} : \mathbb{R}^N \to \mathbb{C}$ by

$$\check{\varphi}(y) := c_N \int_{\mathbb{R}^N} \mathrm{e}^{ix\cdot y} \varphi(x)\, \mathrm{d}x . \tag{3.3.2}$$

This defines a linear map $\varphi \mapsto \check{\varphi} =: \Phi(\varphi)$ on $L^1(\mathbb{R}^N)$, which enjoys all the properties of $\mathcal{F}$ described in theorem 3.1.1. We propose to show that the map Φ defined by (3.3.2) does yield the desired inverse Fourier transform, at least on $\Lambda^1(\mathbb{R}^N)$. More precisely, we wish to show that the composition $\mathcal{F} \circ \Phi$ is the identity in $\Lambda_1(\mathbb{R}^N)$; that is, if $\varphi \in \Lambda^1(\mathbb{R}^N)$, then $\check{\varphi} \in L^1(\mathbb{R}^N)$, and

$$\mathcal{F}(\check{\varphi}) = \varphi . \tag{3.3.3}$$

For example, $E \in \Lambda^1(\mathbb{R}^N)$; since, by (3.2.16),

$$E = \Phi(\hat{E}) = \check{E} , \tag{3.3.4}$$

it follows that

$$\mathcal{F}(\check{E}) = \mathcal{F}(E) = E . \tag{3.3.5}$$

In particular, (3.3.3) would justify the notation $\Phi =: \mathcal{F}^{-1}$; that is,

$$\check{\varphi} =: \mathcal{F}^{-1}(\varphi) , \qquad \varphi \in \Lambda^1(\mathbb{R}^N) . \tag{3.3.6}$$

The justification that the map defined by (3.3.2) should in fact be the inverse map of $\mathcal{F}$ is a consequence of the following inversion theorem, a proof of which can be found in Rudin, [6, sct. 9.11].

Theorem 3.3.1 *Let $f \in \Lambda^1(\mathbb{R}^N)$, and define g on $\mathbb{R}^N$ by*

$$g(x) := c_N \int_{\mathbb{R}^N} \mathrm{e}^{ix\cdot y} \hat{f}(y)\,\mathrm{d}y\,. \tag{3.3.7}$$

Then, $g \in C(\mathbb{R}^N)$, and $g = f$ almost everywhere in $\mathbb{R}^N$.

As a consequence, $\mathcal{F}$ is injective on $L^1(\mathbb{R}^N)$. Indeed, assume that $f,\, g \in L^1(\mathbb{R}^N)$ are such that $\hat{f} = \hat{g}$, and let $h := f - g$. Then, $h \in L^1(\mathbb{R}^N)$ and, since $\hat{h} \equiv 0$, also $\hat{h} \in L^1(\mathbb{R}^N)$. Hence, $h \in \Lambda^1(\mathbb{R}^N)$; thus, by (3.3.7),

$$0 = c_N \int_{\mathbb{R}^N} \mathrm{e}^{ix\cdot y}\hat{h}(y)\,\mathrm{d}y = h(x) = f(x) - g(x)\,. \tag{3.3.8}$$

This shows that $f = g$ almost everywhere. □

Corollary 3.3.1 *Let $f \in \Lambda^1(\mathbb{R}^N)$. The inversion formula*

$$f(x) = c_N \int_{\mathbb{R}^N} \mathrm{e}^{ix\cdot y} \hat{f}(y)\,\mathrm{d}y \tag{3.3.9}$$

holds almost everywhere in $\mathbb{R}^N$. Consequently, f and $\hat{f} \in C(\mathbb{R}^N)$, and vanish at infinity. In particular f and $\hat{f}$ are bounded; consequently, f and $\hat{f} \in L^p(\mathbb{R}^N)$, for all $p \in [1, +\infty]$.

Proof. Since $f \in L^1(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N)$,we only need to show that $f \in L^p(\mathbb{R}^N)$ for $1 < p < \infty$. This follows by interpolation; that is, from the estimate

$$\begin{aligned} |f|_p^p &= \int_{\mathbb{R}^N} |f(x)|^p\,\mathrm{d}x \\ &\le \sup_{x\in\mathbb{R}^N} |f(x)|^{p-1} \int_{\mathbb{R}^N} |f(x)|\,\mathrm{d}x \le |f|_\infty^{p-1}\,|f|_1\,. \end{aligned} \qquad \square \tag{3.3.10}$$

2. The following results summarize the other major properties of the space $\Lambda^1(\mathbb{R}^N)$ that we need in the sequel; in general, if h is a complex-valued function, we denote by $\overline{h}$ its complex conjugate.

Proposition 3.3.1 *Let $f \in \Lambda^1(I\!R^N)$. Then, $f \in \mathcal{UCB}(I\!R^N)$ (that is, f is uniformly continuous and bounded on $I\!R^N$), and for all $x \in I\!R^N$,*

$$[\mathcal{F}(\hat{f})](x) = f(-x)\,. \tag{3.3.11}$$

If also $g \in \Lambda^1(I\!R^N)$,

$$\int_{I\!R^N} f(x)\overline{g(x)}\,\mathrm{d}x = \int_{I\!R^N} \hat{f}(y)\overline{\hat{g}(y)}\,\mathrm{d}y\,, \tag{3.3.12}$$

$$\int_{I\!R^N} \hat{f}(y)g(y)\,\mathrm{d}y = \int_{I\!R^N} f(x)\hat{g}(x)\,\mathrm{d}x\,. \tag{3.3.13}$$

Identity (3.3.12) *is known as* PARSEVAL'S FORMULA.

Proof. 1) The uniform continuity and boundedness of f follows from theorem 3.1.1. Next, note that since $\hat{f} \in L^1(I\!R^N)$, it makes sense to consider its Fourier transform $\mathcal{F}(\hat{f})$. Then, (3.3.11) follows from the inversion formula (3.3.9). Indeed,

$$\begin{aligned} [\mathcal{F}(\hat{f})](x) &= c_N \int_{I\!R^N} \mathrm{e}^{-ix\cdot y}\hat{f}(y)\,\mathrm{d}y \\ &= c_N \int_{I\!R^N} \mathrm{e}^{i(-x)\cdot y}\hat{f}(y)\,\mathrm{d}y = f(-x)\,. \end{aligned} \tag{3.3.14}$$

2) We first note that each term of (3.3.12), as well as of (3.3.13), makes sense, since $\hat{f}$ and $\hat{g} \in L^2(I\!R^N)$, by corollary 3.3.1. Again by the inversion formula (3.3.9), (3.3.12) follows from

$$\begin{aligned} \int_{I\!R^N} f(x)\overline{g(x)}\,\mathrm{d}x &= c_N \int_{I\!R^N}\int_{I\!R^N} \mathrm{e}^{ix\cdot y}\hat{f}(y)\,\mathrm{d}y\,\overline{g(x)}\,\mathrm{d}x \\ &= c_N \int_{I\!R^N} \hat{f}(y)\int_{I\!R^N} \mathrm{e}^{ix\cdot y}\overline{g(x)}\,\mathrm{d}x\,\mathrm{d}y \\ &= c_N \int_{I\!R^N} \hat{f}(y)\overline{\int_{I\!R^N} \mathrm{e}^{-ix\cdot y}g(x)\,\mathrm{d}x}\,\mathrm{d}y \\ &= \int_{I\!R^N} \hat{f}(y)\overline{\hat{g}(y)}\,\mathrm{d}y\,. \end{aligned} \tag{3.3.15}$$

3) Similarly, (3.3.13) follows from

$$\int_{I\!R^N} f(x)\hat{g}(x)\,\mathrm{d}x = c_N \int_{I\!R^N} f(x)\int_{I\!R^N} \mathrm{e}^{-ix\cdot y}g(y)\,\mathrm{d}y\,\mathrm{d}x$$

$$= c_N \int_{\mathbb{R}^N} g(y) \int_{\mathbb{R}^N} \mathrm{e}^{-iy\cdot x} f(x)\,\mathrm{d}x\,\mathrm{d}y \qquad (3.3.16)$$

$$= \int_{\mathbb{R}^N} g(y)\hat{f}(y)\,\mathrm{d}y\,.$$

This completes the proof of proposition 3.3.1. □

In particular, taking $g = f$ in (3.3.12) yields PLANCHEREL FORMULA for $f \in \Lambda^1(\mathbb{R}^N)$; that is,

$$|f|_2 = |\hat{f}|_2\,. \qquad (3.3.17)$$

Example 3.3.1 *For each $m \in \mathbb{N}$, $m \geq 1$,*

$$I_m := \int_{-\infty}^{+\infty} \frac{(\sin y)\,(\sin(my))}{y^2}\,\mathrm{d}y = \pi\,. \qquad (3.3.18)$$

Proof. First note that the generalized integral in (3.3.18) is well-defined. Then, recalling (3.1.2), we compute that

$$\frac{\sin(my)}{y} = \sqrt{\frac{\pi}{2}}\,\hat{\chi}_m(y)\,. \qquad (3.3.19)$$

Consequently, $\chi_m \in \Lambda^1(\mathbb{R}^N)$; thus, by (3.3.12),

$$\begin{aligned} I_m &= \int_{-\infty}^{+\infty} \frac{\sin y}{y}\frac{\sin(my)}{y}\,\mathrm{d}y \\ &= \frac{\pi}{2}\int_{-\infty}^{+\infty} \hat{\chi}_1(y)\,\hat{\chi}_m(y)\,\mathrm{d}y \\ &= \frac{\pi}{2}\int_{-\infty}^{+\infty} \chi_1(y)\,\chi_m(y)\,\mathrm{d}y \\ &= \frac{\pi}{2}\int_{-1}^{1} 1\,\mathrm{d}y\,, \end{aligned} \qquad (3.3.20)$$

from which (3.3.18) follows. □

3. A completely analogous set of results holds for functions in the space

$$\tilde{\Lambda}^1(\mathbb{R}^N) := \{f \in L^1(\mathbb{R}^N) \mid \check{f} \in L^1(\mathbb{R}^N)\}\,, \qquad (3.3.21)$$

which is also not empty because, by proposition 3.2.1, $E \in \tilde{\Lambda}^1(\mathbb{R}^N)$. More precisely, the same conclusions of theorem 3.3.1, corollary 3.3.1 and proposition 3.3.1 hold, and can be proven in the same way, with the assumptions $f,\, g \in \Lambda^1(I\!\!R^N)$ replaced by $f,\, g \in \tilde{\Lambda}^1(I\!\!R^N)$, and $\hat{f}$, $\hat{g}$ replaced by $\check{f}$, $\check{g}$ throughout. In particular, the analogous of theorem 3.3.1 guarantees that if $f \in \tilde{\Lambda}^1(I\!\!R^N)$ and

$$\tilde{g}(y) := c_N \int_{I\!\!R^N} \mathrm{e}^{-ix\cdot y} f(x)\,\mathrm{d}x\,, \tag{3.3.22}$$

then $\tilde{g} = \hat{f}$ almost everywhere. Consequently, the map $\mathcal{F}^{-1}$ is also injective; that is, if $f,\, g \in L^1(I\!\!R^N)$ are such that $\check{f} = \check{g}$, then $f = g$ almost everywhere.

3.4 Fourier Transform and Convolution.

1. In this section we consider the Fourier transform of the convolution of two functions. We recall that $L^1(I\!\!R^N)$ is an algebra with respect to the convolution product; that is, if f and $g \in L^1(I\!\!R^N)$, then $f * g \in L^1(I\!\!R^N)$, and

$$|f * g|_1 \le |f|_1\,|g|_1\,. \tag{3.4.1}$$

Thus, it makes sense to consider the Fourier transform of the convolution of two integrable functions. In the sequel, we adopt the notation $\mathcal{F}^{-1}$ for the map Φ defined in (3.3.6); that is, again, for $f \in \Lambda^1(I\!\!R^N)$, we set

$$[\mathcal{F}^{-1}(f)](x) = \check{f}(x) = c_N \int_{I\!\!R^N} \mathrm{e}^{i\,x\cdot y}\, f(y)\,\mathrm{d}y\,. \tag{3.4.2}$$

Proposition 3.4.1 *Let $f,\, g \in L^1(I\!\!R^N)$. Then,*

$$\mathcal{F}(f * g) = \tfrac{1}{c_N}\,\hat{f}\,\hat{g}, \qquad\qquad \mathcal{F}^{-1}(f * g) = \tfrac{1}{c_N}\,\check{f}\,\check{g}\,. \tag{3.4.3}$$

Conversely, assume that $f,\, g \in \Lambda^1(I\!\!R^N)$ [respectively, $\tilde{\Lambda}^1(I\!\!R^N)$]. Then,

$$\mathcal{F}(f\,g) = \tfrac{1}{c_N}\,\hat{f} * \hat{g}, \qquad [\text{resp.,}\quad \mathcal{F}^{-1}(f\,g) = \tfrac{1}{c_N}\,\check{f} * \check{g}\,]\,. \tag{3.4.4}$$

Proof. First, we remark that (3.4.3) means that the maps $\mathcal{F}$ and $\mathcal{F}^{-1}$ transform a convolution product into a pointwise product; the latter makes sense, because $\hat{f}$, $\hat{g}$, $\check{f}$, and $\check{g}$, are continuous functions. Next:

1) Identity (3.4.3) is a consequence of Fubini's theorem. Following Rudin, [6, sct. 9.2], we set $h := f * g$, and compute

$$\begin{aligned}
\hat{h}(y) &= c_N \int_{I\!R^N} \mathrm{e}^{-ix\cdot y}\, h(x)\,\mathrm{d}x \\
&= c_N \int_{I\!R^N} \mathrm{e}^{-ix\cdot y} \int_{I\!R^N} f(z)\, g(x-z)\,\mathrm{d}z\,\mathrm{d}x \\
&= c_N \int_{I\!R^N} f(z) \int_{I\!R^N} \mathrm{e}^{-ix\cdot y}\, g(x-z)\,\mathrm{d}x\,\mathrm{d}z \\
&= c_N \int_{I\!R^N} f(z) \int_{I\!R^N} \mathrm{e}^{-iy\cdot(s+z)}\, g(s)\,\mathrm{d}s\,\mathrm{d}z \qquad (3.4.5) \\
&= c_N \int_{I\!R^N} f(z)\mathrm{e}^{-iy\cdot z} \int_{I\!R^N} \mathrm{e}^{-iy\cdot s}\, g(s)\,\mathrm{d}s\,\mathrm{d}z \\
&= \left(c_N \int_{I\!R^N} \mathrm{e}^{-iy\cdot z}\, f(z)\,\mathrm{d}z\right) \tfrac{1}{c_N} \left(c_N \int_{I\!R^N} \mathrm{e}^{-iy\cdot s}\, g(s)\,\mathrm{d}s\right) \\
&= \tfrac{1}{c_N}\, \hat{f}(y)\, \hat{g}(y)\,.
\end{aligned}$$

This proves the first identity of (3.4.3); the second is proven in the same way.

2) Note first that (3.4.4) makes sense, since by corollary 3.3.1 f and $g \in L^2(I\!R^N)$, so that the product fg is in $L^1(I\!R^N)$, and its Fourier transform is defined. Moreover, also $\hat{f} * \hat{g} \in L^1(I\!R^N)$, and, by (3.4.3) and (3.3.11),

$$\begin{aligned}
\left[\mathcal{F}(\hat{f} * \hat{g})\right](x) &= \tfrac{1}{c_N}\,[\mathcal{F}\hat{f}](x)[\mathcal{F}\hat{g}](x) \\
&= \tfrac{1}{c_N}\, f(-x)g(-x) = \tfrac{1}{c_N}\,[fg](-x) \qquad (3.4.6) \\
&= \tfrac{1}{c_N}\,\left[\mathcal{F}\left(\mathcal{F}(fg)\right)\right](x)\,.
\end{aligned}$$

This implies that

$$\mathcal{F}\left[\left(\hat{f} * \hat{g}\right) - \tfrac{1}{c_N}\,\mathcal{F}(fg)\right] = 0\,, \qquad (3.4.7)$$

so by theorem 3.3.1 we conclude that

$$\hat{f} * \hat{g} = \tfrac{1}{c_N}\,\mathcal{F}(fg)\,, \qquad (3.4.8)$$

from which the first of (3.4.4) follows. The second of (3.4.4) is proven in the same way. □

2. As a first application of proposition 3.4.1, we show the so-called semigroup property of the heat kernel.

Proposition 3.4.2 *Let $E_{(t)}$ be the heat kernel defined in (3.2.17). Then, for all t and $s > 0$,*

$$E_{(t+s)} = E_{(t)} * E_{(s)} \, . \tag{3.4.9}$$

Proof. This result is an immediate consequence of (3.2.20) and (3.4.3), which imply that

$$\begin{aligned} \hat{E}_{(t+s)}(y) &= c_N \, \mathrm{e}^{-(t+s)|y|^2} = c_N^{-1} \, c_N \, \mathrm{e}^{-t|y|^2} \, c_N \, \mathrm{e}^{-s|y|^2} \\ &= \tfrac{1}{c_N} \, \hat{E}_{(t)}(y) \, \hat{E}_{(s)}(y) = \left[\mathcal{F}(E_{(t)} * E_{(s)})\right](y) \, . \qquad \square \end{aligned} \tag{3.4.10}$$

As a second application, we prove an important formula characterizing the Fourier transform of an integrable function.

Proposition 3.4.3 *Let $f \in L^1(I\!R^N)$. Then, for all $y \in I\!R^N$,*

$$\hat{f}(y) = c_N \, \left[f * \mathrm{e}^{i\, y \cdot (\cdot)}\right](0) \, . \tag{3.4.11}$$

Proof. We compute that

$$\begin{aligned} \hat{f}(y) &= c_N \int_{I\!R^N} \mathrm{e}^{-\, is \cdot y} \, f(s) \, \mathrm{d}s = c_N \int_{I\!R^N} \mathrm{e}^{i\, z \cdot y} \, f(-z) \, \mathrm{d}z \\ &= c_N \int_{I\!R^N} \mathrm{e}^{i\, z \cdot y} \, f(0 - z) \, \mathrm{d}z = \left[f * \mathrm{e}^{i\, y \cdot (\cdot)}\right](0) \, , \end{aligned} \tag{3.4.12}$$

having used the change of coordinates $z = -\, s$ and the definition of convolution of two functions, evaluated at $x = 0$. □

3. A more important consequence of proposition 3.4.1 is that it allows us to show that while the map $\mathcal{F}$ is, by theorem 3.1.1, continuous from $L^1(I\!R^N)$ into $\mathcal{UCB}(I\!R^N)$, the inverse map $\mathcal{F}^{-1}$ needs not be continuous from $\mathcal{UCB}(I\!R^N) \cap L^1(I\!R^N)$, with the topology induced by $\mathcal{UCB}(I\!R^N)$, into $L^1(I\!R^N)$. Note that we do need to consider the restriction of $\mathcal{F}^{-1}$ to the intersection $\mathcal{UCB}(I\!R^N) \cap L^1(I\!R^N)$, in order to guarantee that $\mathcal{F}^{-1}$ be defined. This highlights once more the unsuitability of the space $L^1(I\!R^N)$ as domain of the Fourier transform.

Example 3.4.1 *For $m \in I\!N$, let χ_m denote the characteristic function of the interval $[-m, m]$, and set $g_m := \chi_m * \chi_1$. Define*

$$f_m(x) := \begin{cases} \frac{2}{\pi}\frac{(\sin x)\,(\sin(mx))}{x^2} & \text{if} \quad x \neq 0\,, \\ \frac{2}{\pi}\,m & \text{if} \quad x = 0\,. \end{cases} \tag{3.4.13}$$

The functions f_m and g_m are in $\mathcal{UCB}(I\!R) \cap L^1(I\!R)$ for all $m \in I\!N$, and $f_m = \frac{1}{c_1}\,\check{g}_m$. However, $|f_m|_1 \to +\infty$ as $m \to \infty$, while $|g_m|_\infty = 1$.

Proof. It is easy to compute that

$$g_m(y) = \begin{cases} 1 & \text{if} \quad |y| \leq m\,, \\ m + 1 - |y| & \text{if} \quad m \leq |y| \leq m+1\,, \\ 0 & \text{if} \quad |y| \geq m+1\,; \end{cases} \tag{3.4.14}$$

thus, obviously $g_m \in \mathcal{UCB}(I\!R) \cap L^1(I\!R)$. We immediately verify that the same is true for f_m. Exactly as in (3.1.2), we also compute that, if $y \neq 0$,

$$\check{\chi}_m(y) = \sqrt{\frac{2}{\pi}}\frac{\sin(my)}{y}\,. \tag{3.4.15}$$

Hence, from the second of (3.4.3),

$$\check{g}_m = \mathcal{F}^{-1}(\chi_m * \chi_1) = \sqrt{2\pi}\,(\check{\chi}_m\check{\chi}_1) = \sqrt{2\pi}\,f_m\,, \tag{3.4.16}$$

as claimed. Clearly, $|g_m|_\infty = 1$; to estimate $|f_m|_1$, we first choose $\delta \in \,]0,1[$ such that $\frac{\sin x}{x} \geq \frac{1}{2}$ if $0 < x < \delta$ and then, for $m > \frac{5\pi}{6\delta}$, we set

$$K_m := \max\left\{k \in I\!N \;\middle|\; \tfrac{5\pi}{6m} + \tfrac{k\pi}{m} \leq \delta\right\} = \left\lfloor \tfrac{m\delta}{\pi} - \tfrac{5}{6}\right\rfloor\,. \tag{3.4.17}$$

Then, since in general $\lfloor x \rfloor > x - 1$,

$$K_m \geq \tfrac{m\delta}{\pi} - \tfrac{11}{6}\,. \tag{3.4.18}$$

Noting that f_m is even, for $m > \frac{5\pi}{6\delta}$ we compute that

$$\begin{aligned} |f_m|_1 &= 2\int_0^\infty |f_m(x)|\,\mathrm{d}x \\ &\geq 2\sum_{k=0}^{K_m}\int_{\frac{(1+6k)\pi}{6m}}^{\frac{(5+6k)\pi}{6m}} \frac{|\sin x|}{x}\frac{|\sin(mx)|}{x}\,\mathrm{d}x\,. \end{aligned} \tag{3.4.19}$$

Our choices of m and δ guarantee that

$$0 < \frac{\pi}{6m} + \frac{k\pi}{m} \leq x \leq \frac{5\pi}{6m} + \frac{k\pi}{m} < \delta \,, \tag{3.4.20}$$

so that $\sin x > 0$ and $\frac{\sin x}{x} \geq \frac{1}{2}$. Likewise,

$$0 < \frac{\pi}{6} + k\pi \leq mx \leq \frac{5\pi}{6} + k\pi \,, \tag{3.4.21}$$

so $|\sin(mx)| \geq \frac{1}{2}$. Consequently,

$$\begin{aligned} |f_m|_1 &\geq \frac{1}{2} \sum_{k=0}^{K_m} \int_{\frac{(1+6k)\pi}{6m}}^{\frac{(5+6k)\pi}{6m}} \frac{1}{x} \, \mathrm{d}x \\ &\geq \frac{1}{2} \sum_{k=0}^{K_m} \frac{6m}{(5+6k)\pi} \frac{4\pi}{6m} \\ &= 2 \sum_{k=0}^{K_m} \frac{1}{5+6k} \,. \end{aligned} \tag{3.4.22}$$

Since, by (3.4.18), $K_m \geq C\,m$ for a suitable constant C independent of m, it follows that $|f_m|_1 \to +\infty$, as claimed. □

3.5 Fourier Transform and Differentiation.

One of the main reasons why the Fourier transform plays such an important role in the study of PDEs (see section 3.9) is that it transforms differentiation into multiplication, and viceversa. Before we proceed, we need a result that extends the density claim of proposition 2.4.1.

Definition 3.5.1 *Let $p \in [1, \infty[$, $m \in \mathbb{N}$, and $\alpha \in \mathbb{N}^N$, with $|\alpha| \leq m$. Given $f \in L^p(\mathbb{R}^N)$, we say that $D^\alpha f \in L^p(\mathbb{R}^N)$ if there is $g_\alpha \in L^p(\mathbb{R}^N)$ such that*

$$D^\alpha T_f = T_{g_\alpha} \,; \tag{3.5.1}$$

then, we introduce the (Sobolev) space

$$W^{m,p}(\mathbb{R}^N) := \Big\{ f \in L^p(\mathbb{R}^N) \;\mid\; D^\alpha f \in L^p(\mathbb{R}^N) \quad \forall\, \alpha\,,\; |\alpha| \le m \Big\}\,. \tag{3.5.2}$$

In other words, $W^{m,p}(\mathbb{R}^N)$ consists of all p-integrable functions having regular distributional derivatives of order up to m generated by functions in $L^p(\mathbb{R}^N)$.

Proposition 3.5.1 *$W^{m,p}(\mathbb{R}^N)$ is a Banach space with respect to the norm*

$$\|f\|_{m,p} := \sum_{|\alpha|\le m} \|D^\alpha f\|_{L^p(\mathbb{R}^N)}\,. \tag{3.5.3}$$

The space $C_0^\infty(\mathbb{R}^N)$ is dense in $W^{m,p}(\mathbb{R}^N)$.

Proof. See, e.g., Adams and Fournier, [1, thm. 3.3 & cor. 3.23]. □

We are now ready to prove

Theorem 3.5.1 *Let $f \in W^{1,1}(\mathbb{R}^N)$, and $1 \le j \le N$. Then,*

$$[\mathcal{F}(\partial_j f)](y) = i\, y_j\, \hat{f}(y)\,. \tag{3.5.4}$$

Conversely, define $g_j(y) := iy_j \hat{f}(y)$. Then, if $\hat{f}$ and $g \in L^1(\mathbb{R}^N)$, f has a classical derivative $\partial_j f$ given by

$$[\partial_j f](x) = \check{g}_j(x) = [\mathcal{F}^{-1}(i(\cdot)_j \hat{f})](x)\,. \tag{3.5.5}$$

Proof. We first note that since $\partial_j f \in L^1(\mathbb{R}^N)$, it does have a Fourier transform. Because of proposition 3.5.1, it is sufficient [2] to show (3.5.4) for $f \in C_0^\infty(\mathbb{R}^N)$, in which case (3.5.4) is obtained by simple integration by parts. Indeed,

$$\begin{aligned} [\mathcal{F}(\partial_j f)](y) &= c_N \int_{\mathbb{R}^N} \mathrm{e}^{-i\,x\cdot y}\, [\partial_j f](x)\, \mathrm{d}x \\ &= -\, c_N \int_{\mathbb{R}^N} (-iy_j)\, \mathrm{e}^{-ix\cdot y}\, f(x)\, \mathrm{d}x = i\, y_j\, \hat{f}(y)\,. \end{aligned} \tag{3.5.6}$$

Conversely, note first that f is (uniformly) continuous because $\hat{f} \in L^1(\mathbb{R}^N)$, which also implies that the inversion formula (3.3.9) holds. It is then sufficient to prove the

[2] The reader is encouraged to work out the details of this (standard) density argument.

one-dimensional version of (3.5.5), that is, that $f'(a) = \check{g}(a)$ for all $a \in I\!\!R$, assuming that $\hat{f}$ and $g \in L^1(I\!\!R)$. Thus, recalling that

$$\left| \frac{e^{i\,z\cdot y} - 1}{z} \right| \le |y| \tag{3.5.7}$$

for all $y,\, z \in \mathbb{C}$ [3], we compute that

$$\begin{aligned} \frac{f(x) - f(a)}{x - a} &= \frac{c_1}{x-a} \int_{-\infty}^{+\infty} \left(e^{ixy} - e^{iay} \right) \hat{f}(y)\, dy \\ &= \frac{c_1}{x-a} \int_{-\infty}^{+\infty} e^{iay} \left(e^{iy(x-a)} - 1 \right) \hat{f}(y)\, dy \\ &= i\, c_1 \int_{-\infty}^{+\infty} e^{iay}\, \frac{e^{iy(x-a)} - 1}{i(x-a)}\, \hat{f}(y)\, dy \\ &=: i\, c_1 \int_{-\infty}^{+\infty} e^{iay}\, E(x,a,y)\, \hat{f}(y)\, dy\,. \end{aligned} \tag{3.5.8}$$

Since, by (3.5.7),

$$|e^{iay}\, E(x,a,y) f(y)| \le |y||\hat{f}(y)| = |g(y)|\,, \tag{3.5.9}$$

independently of $|x - a|$, and $g \in L^1(I\!\!R)$, and

$$E(x,a,y) \to y \qquad \text{as} \quad x \to a\,, \tag{3.5.10}$$

the Lebesgue dominated convergence theorem allows us to deduce from (3.5.8) that

$$f'(a) = i\, c_1 \int_{-\infty}^{+\infty} e^{iay} y \hat{f}(y)\, dy = c_1 \int_{-\infty}^{+\infty} e^{iay} g(y)\, dy = \check{g}(a)\,, \tag{3.5.11}$$

as claimed. □

Corollary 3.5.1 *Let $f \in W^{m,1}(I\!\!R^N)$, and $\alpha \in I\!\!N$, with $|\alpha| \le m$. Then,*

$$[\mathcal{F}(D^\alpha f)](y) = (iy)^\alpha \hat{f}(y) \tag{3.5.12}$$

(recall that $y^\alpha := y_1^{\alpha_1} \dots y_N^{\alpha_N}$). In particular, if $f \in W^{2,1}(I\!\!R^N)$,

$$\widehat{\Delta\, f}(y) = -\,|y|^2\, \hat{f}(y)\,. \tag{3.5.13}$$

[3] To prove (3.5.7), recall that $0 \le 1 - \cos\theta \le \frac{1}{2}\,\theta^2$ for all $\theta \in I\!\!R$.

Conversely, define functions $y \mapsto g_\alpha(y)$ *by the right side of* (3.5.12), *and assume that, for some* $m \in \mathbb{N}$, f *and* $g_\alpha \in L^1(\mathbb{R}^N)$ *for all* $\alpha \in \mathbb{N}$ *such that* $|\alpha| \leq m$. *Then,* $D^\alpha f \in \mathcal{UCB}(\mathbb{R}^N)$, *and for all* $x \in \mathbb{R}^N$,

$$[D^\alpha f](x) = \check{g}_\alpha(x)\,. \tag{3.5.14}$$

Proof. It is sufficient to show (3.5.13). By (3.5.12) with $D^\alpha = \partial_k^2$, $1 \leq k \leq N$ (that is, with $\alpha = 2\,\mathrm{e}_k$ (the k-th unit vector of the standard basis of $\mathbb{R}^N$), it follows that

$$[\mathcal{F}(\partial_k^2)](y) = (i\,y)^{2\,\mathrm{e}_k}\,\hat{f}(y) = -\,y_k^2\,\hat{f}(y)\,. \tag{3.5.15}$$

Consequently,

$$\widehat{\Delta\,f}(y) = -\sum_{k=1}^{N} y_k^2\,\hat{f}(y) = -\,|y|^2\,\hat{f}(y)\,, \tag{3.5.16}$$

which is (3.5.13). □

Formulas (3.5.12) and (3.5.14) express the conversion of differentiation into multiplication we mentioned above; indeed, either formula allows us to recover the derivative $D^\alpha f$ from the knowledge (admittedly a difficult question in itself!) of the inverse Fourier transform of g_α.

3.6 Fourier Transform in $L^2(\mathbb{R}^N)$.

1. In order to eliminate the unpleasant "asymmetry" between the domain (i.e. $L^1(\mathbb{R}^N)$) and the range (i.e. $\mathcal{UCB}(\mathbb{R}^N)$) of the maps $\mathcal{F}$ and $\mathcal{F}^{-1}$, which is apparent in the results of the previous section, we extend these maps to linear isometries from $L^2(\mathbb{R}^N)$ into itself, in such a way that $\mathcal{F}$ and $\mathcal{F}^{-1}$ are indeed the inverse one of the other, as maps in $L^2(\mathbb{R}^N)$. The basis of the construction of the Fourier transform in $L^2(\mathbb{R}^N)$ is the density of the space

$$L^{1\cap 2}(\mathbb{R}^N) \;:=\; L^{1\cap 2}(\mathbb{R}^N) \tag{3.6.1}$$

into $L^2(\mathbb{R}^N)$: we show that the map $\mathcal{F}$, which is defined in $L^{1\cap 2}(\mathbb{R}^N)$ because $L^{1\cap 2}(\mathbb{R}^N) \subset L^1(\mathbb{R}^N)$, when restricted to this space has range in $L^2(\mathbb{R}^N)$, and this restriction preserves the L^2-norm. Thus, we can extend this restriction to a unitary operator in $L^2(I\!R^N)$, which we finally show to be invertible.

We start with

Theorem 3.6.1 *Let $f \in L^{1\cap 2}(I\!R^N)$. Then, $\hat{f}$, which is in $\mathcal{UCB}(I\!R^N)$, is also in $L^2(I\!R^N)$, and Plancherel' formula* (3.3.17) *holds, i.e.*

$$|\hat{f}|_2 = |f|_2 \,. \tag{3.6.2}$$

Proof. Given $f \in L^{1\cap 2}(I\!R^N)$, let $f_0(x) := f(-x)$ and $g := f * f_0$; explicitly,

$$\begin{aligned} g(x) &= \int_{I\!R^N} f(x-y)f(-y)\,\mathrm{d}y \\ &= \int_{I\!R^N} f(x+z)f(z)\,\mathrm{d}z = \langle f(x+\cdot), f\rangle \,, \end{aligned} \tag{3.6.3}$$

where the last pairing denotes the scalar product in $L^2(I\!R^N)$. In particular,

$$g(0) = |f|_2^2 \,. \tag{3.6.4}$$

By the continuity of translations in $L^2(I\!R^N)$, g is continuous. By (3.6.3), g is also bounded on $I\!R^N$, with

$$|g(x)| \le |f|_2^2 \tag{3.6.5}$$

for all $x \in I\!R^N$. In addition, $g \in L^1(I\!R^N)$ as well (because $L^1(I\!R^N)$ is an algebra with respect to the convolution product), and, by (3.4.3),

$$\hat{g} = \tfrac{1}{c_N}\, \hat{f}\, \hat{f}_0 \tag{3.6.6}$$

(pointwise product). Since

$$\begin{aligned} \hat{f}_0(y) &= c_N \int_{I\!R^N} \mathrm{e}^{-iy\cdot x} f(-x)\,\mathrm{d}x = c_N \int_{I\!R^N} \mathrm{e}^{iy\cdot z} f(z)\,\mathrm{d}z \\ &= c_N \overline{\int_{I\!R^N} \mathrm{e}^{-iy\cdot z} f(z)\,\mathrm{d}z} = \overline{\hat{f}(y)} \,, \end{aligned} \tag{3.6.7}$$

(3.6.6) implies that

$$\hat{g} = \tfrac{1}{c_N}\, \hat{f}\,\overline{\hat{f}} \,. \tag{3.6.8}$$

Let now $E_{(t)}$ be the heat kernel defined in (3.2.17). By (3.2.24) applied to g, which is bounded and continuous at $x = 0$, and (3.6.4),

$$\lim_{t\to 0}(g * E_{(t)})(0) = g(0) = |f|_2^2 \,. \tag{3.6.9}$$

On the other hand, since $g \in L^1(\mathbb{R}^N)$, by (3.2.22) we also have that

$$(g * E_{(t)})(0) = c_N \int_{\mathbb{R}^N} \mathrm{e}^{-t|z|^2}\,\hat{g}(z)\,\mathrm{d}z \,; \tag{3.6.10}$$

since for each $z \in \mathbb{R}^N$, $\mathrm{e}^{-t|z|^2} \to 1^-$ as $t \to 0$, by the dominated convergence theorem we deduce from (3.6.10) that $\hat{g} \in L^1(\mathbb{R}^N)$ (so that $g \in \Lambda^1(\mathbb{R}^N)$), and, by (3.6.9),

$$\int_{\mathbb{R}^N} \hat{g}(z)\,\mathrm{d}z = \tfrac{1}{c_N}\,\lim_{t\to 0}(g * E_{(t)})(0) = \tfrac{1}{c_N}\,|f|_2^2 \,. \tag{3.6.11}$$

Recalling (3.6.8), (3.6.2) follows from (3.6.11). □

2. Let now $f \in L^2(\mathbb{R}^N)$. Since $L^{1\cap 2}(\mathbb{R}^N)$ is dense in $L^2(\mathbb{R}^N)$, there is a sequence $(f_m)_{m\geq 0} \subset L^{1\cap 2}(\mathbb{R}^N)$ such that $f_m \to f$ in $L^2(\mathbb{R}^N)$. By theorem 3.6.1, each $\hat{f}_m$ is in $L^2(\mathbb{R}^N)$, and Plancherel's identity (3.6.2) implies that

$$|\hat{f}_m - \hat{f}_r|_2 = |f_m - f_r|_2 \,. \tag{3.6.12}$$

This means that $(\hat{f}_m)_{m\geq 0}$ is a Cauchy sequence in $L^2(\mathbb{R}^N)$; thus, it converges to a limit $\tilde{f} \in L^2(\mathbb{R}^N)$. This limit depends only on f; that is, it does not depend on the particular sequence $(f_m)_{m\geq 0}$ approximating f. Indeed, if $(g_m)_{m\geq 0} \subset L^{1\cap 2}(\mathbb{R}^N)$ also converges to f, again by (3.6.2) we see that

$$\begin{aligned} |\hat{g}_m - \tilde{f}|_2 &\leq |\hat{g}_m - \hat{f}_m|_2 + |\hat{f}_m - \tilde{f}|_2 \\ &= |g_m - f_m|_2 + |\hat{f}_m - \tilde{f}|_2 \,, \end{aligned} \tag{3.6.13}$$

which implies that $\hat{g}_m \to \tilde{f}$ as well. (For example, following Rudin, [6, sct. 9.13], we can take $f_m = f\,\chi_m$, where χ_m is the characteristic function of the ball $B(0,m)$). Since $\tilde{f}$ coincides with $\hat{f}$ if $f \in L^{1\cap 2}(\mathbb{R}^N)$ (as in this case we can take $f_m = f$ for all m), we *define* $\tilde{f}$ to be the Fourier transform of $f \in L^2(\mathbb{R}^N)$, and we set again $\tilde{f} =: \hat{f}$. We have thus defined a map

$$\Phi : L^2(\mathbb{R}^N) \to L^2(\mathbb{R}^N)\,, \qquad \Phi f = \hat{f} \,, \tag{3.6.14}$$

which coincides with $\mathcal{F}$ on $L^{1\cap 2}(I\!R^N)$; moreover, Plancherel's formula (3.6.2) still holds in $L^2(I\!R^N)$, because if f and f_m are as above,

$$|\hat{f}|_2 = \lim |\hat{f}_m|_2 = \lim |f_m|_2 = |f|_2 \,. \tag{3.6.15}$$

In a totally analogous way, we can extend the inverse Fourier transform $\mathcal{F}^{-1}$ from $L^{1\cap 2}(I\!R^N)$ into a map

$$\Psi : L^2(I\!R^N) \to L^2(I\!R^N)\,, \tag{3.6.16}$$

in such a way that $\Psi f = \check{f}$ if $f \in L^{1\cap 2}(I\!R^N)$, and

$$|\Psi f|_2 = |f|_2 \tag{3.6.17}$$

for all $f \in L^2(I\!R^N)$. Naturally, we set $\check{f} := \Psi f$ for $f \in L^2(I\!R^N)$ as well.

To summarize: Given $f \in L^2(I\!R^N)$, we define its Fourier transform $\hat{f}$ as follows. First, we take an arbitrary sequence $(f_m)_{m\geq 0} \subset L^2(I\!R^N) \cap L^1(I\!R^N)$ such that

$$|f_m - f|_2 \;\to\; 0 \qquad \text{as} \quad m \to \infty\,; \tag{3.6.18}$$

that is, $f_m \to f$ in $L^2(I\!R^N)$. We have shown that the sequence $(\hat{f}_m)_{m\geq 0}$ has a limit in $L^2(I\!R^N)$; thus, we define

$$\hat{f} := \lim \hat{f}_m \qquad \text{in} \quad L^2(I\!R^N)\,. \tag{3.6.19}$$

3. We now show that the map Φ defined in (3.6.14) is onto $L^2(I\!R^N)$ and invertible, and that, as in $\Lambda^1(I\!R^N)$, $\Phi^{-1} = \Psi$, the map defined in (3.6.16). To this end, we prepare

Proposition 3.6.1 *Let $f \in L^{1\cap 2}(I\!R^N)$, and, for $t > 0$, set $g_{(t)} := f * E_{(t)}$, where $E_{(t)}$ is the heat kernel, and $\hat{g}_{(t)} := \mathcal{F}(g_{(t)})$. Then, both $g_{(t)}$ and $\hat{g}_{(t)} \in L^{1\cap 2}(I\!R^N)$, for all $t > 0$.*

Proof. The statement for $g_{(t)}$ follows from part (3) of Propositon 3.2.2, for $p = 1$ and $p = 2$. As for $\hat{g}_{(t)}$, by (3.4.3) it follows that

$$\hat{g}_{(t)} = \mathcal{F}[f * E_{(t)}] = \tfrac{1}{c_N}\, \hat{f}\, \hat{E}_{(t)}\,, \tag{3.6.20}$$

so that the conclusion follows because $\hat{f}$ is uniformly bounded on $I\!R^N$ and $\hat{E}_{(t)} \in L^{1\cap 2}(I\!R^N)$ (by (3.2.20)). $\square$

As a consequence, we deduce that for all $f \in L^{1\cap 2}(\mathbb{R}^N)$ and $t > 0$,

$$\Psi\Phi(f * E_{(t)}) = f * E_{(t)} \; ; \tag{3.6.21}$$

thus, by (3.2.23),

$$\lim_{t\to 0} \Psi\Phi(f * E_{(t)}) = \lim_{t\to 0} f * E_{(t)} = f \; , \tag{3.6.22}$$

from which we conclude that the identity $\Psi\Phi f = f$ holds also in $L^{2\cap 1}(\mathbb{R}^N)$. By the density of this space in $L^2(\mathbb{R}^N)$, it follows that this identy also holds in $L^2(\mathbb{R}^N)$.

A completely analogous argument allows us to interchange Ψ and Φ, yielding the identity $\Phi\Psi g_{(t)} = g_{(t)}$ for all $g_{(t)} \in L^2(\mathbb{R}^N)$. Thus, we obtain that if $g_{(t)} \in L^2(\mathbb{R}^N)$ and $f_{(t)} := \Psi g_{(t)}$, then $f_{(t)} \in L^2(\mathbb{R}^N)$ and $\Phi f_{(t)} = g_{(t)}$: this shows that the map Φ is onto. Since so is Ψ, this shows that Ψ and Φ are indeed inverse one of the other, as isometries in $L^2(\mathbb{R}^N)$.

From now on we denote $\Phi = \mathcal{F}$ and $\Psi = \mathcal{F}^{-1}$ as usual.

Finally, the validitiy of Parseval's formula (3.3.12) in $L^2(\mathbb{R}^N)$ follows from the parallelogram identity

$$f\,\overline{g} = \tfrac{1}{4}\left(|f+g|^2 - |f-g|^2 + i\,|f+ig|^2 - i\,|f-ig|^2\right) \; , \tag{3.6.23}$$

and Plancherel's formula (3.6.2).

4. As a final remark, and as an application of these results, we prove the well-known identity

$$\int_{\mathbb{R}^N} \mathrm{e}^{(2\pi i)\, z\cdot y - \pi |y|^2}\, \mathrm{d}y = \mathrm{e}^{-\pi |z|^2}. \tag{3.6.24}$$

Indeed, recalling (3.2.20) we compute that

$$E_{(t)}(x) = \left[\mathcal{F}^{-1}\left(c_N\, \mathrm{e}^{-t|\cdot|^2}\right)\right](x) = c_N^2 \int_{\mathbb{R}^N} \mathrm{e}^{ix\cdot y - t\,|y|^2}\, \mathrm{d}y \; ; \tag{3.6.25}$$

comparing this to the definition (3.2.17) of $E_{(t)}$, we obtain the identity

$$c_N \int_{\mathbb{R}^N} \mathrm{e}^{ix\cdot y - t\,|y|^2}\, \mathrm{d}y = \frac{1}{(\sqrt{2t})^N}\mathrm{e}^{-|x|^2/4t}. \tag{3.6.26}$$

Since $c_N = \left(\sqrt{2\pi}\right)^{-N}$, (3.6.24) follows then from (3.6.26) for the choice $t = \pi$ and the change of variable $x = 2\pi z$.

3.7 Fourier Transform in $\mathcal{S}(I\!\!R^N)$.

In this section we show that the Fourier transform is "well behaved" in the space of the rapidly decreasing functions, in the sense that $\mathcal{F}$ maps $\mathcal{S}(I\!\!R^N)$ into itself, and is in fact a continuous bijection between $\mathcal{S}(I\!\!R^N)$ and itself. This would give us an alternative way to define the Fourier transform in $L^2(I\!\!R^N)$, defining it first on $\mathcal{S}(I\!\!R^N)$, and then extending the definition to $L^2(I\!\!R^N)$ by the density of $\mathcal{S}(I\!\!R^N)$ into $L^2(I\!\!R^N)$, as per propositiion 2.7.3. In the approach we have followed instead, the Fourier transform is obviously defined on $\mathcal{S}(I\!\!R^N)$, since $\mathcal{S}(I\!\!R^N) \subset L^1(I\!\!R^N)$.

1. Our first goal is to show that $\mathcal{F}$ maps $\mathcal{S}(I\!\!R^N)$ into itself. To this end, given $f \in \mathcal{S}(I\!\!R^N)$ and $k = 1, \ldots, N$ we set $f_k(x) := x_k\, f(x)$, and prepare

Proposition 3.7.1 *For each $f \in \mathcal{S}(I\!\!R^N)$, $k = 1, \ldots, N$, and $y \in I\!\!R^N$,*

$$[\mathcal{F}(\partial_k f)](y) \;=\; i\, y_k\, \hat{f}(y)\,, \tag{3.7.1}$$

$$[\mathcal{F}(f_k)](y) \;=\; i\, \frac{\partial \hat{f}}{\partial y_k}(y)\,. \tag{3.7.2}$$

Proof. If $f \in \mathcal{S}(I\!\!R^N)$, also $\partial_k f \in \mathcal{S}(I\!\!R^N) \subset L^1(I\!\!R^N)$; thus, by an admissible integration by parts,

$$\begin{aligned} (\partial_k f)\hat{}(y) &= c_N \int_{I\!\!R^N} \mathrm{e}^{-i\,x\cdot y}\, \partial_k f(x)\, \mathrm{d}x \\ &= -c_N \int_{I\!\!R^N} (-i\, y_k) \mathrm{e}^{-i\,x\cdot y} f(x)\, \mathrm{d}x \;=\; i\, y_k\, \hat{f}(y)\,. \end{aligned} \tag{3.7.3}$$

This proves (3.7.1) (note that this identity corresponds to (3.5.4)). To prove (3.7.2), noting that the function $x \mapsto x_k \hat{f}(x)$ is still in $\mathcal{S}(I\!\!R^N)$, we compute that

$$\begin{aligned} [\mathcal{F}(f_k)](y) &= c_N \int_{I\!\!R^N} \mathrm{e}^{-i\,x\cdot y}\, x_k f(x)\, \mathrm{d}x \\ &= -\, c_N\, \frac{1}{i} \int_{I\!\!R^N} \frac{\partial}{\partial y_k} \mathrm{e}^{-i\,x\cdot y} f(x)\, \mathrm{d}x \\ &= c_N\, i \int_{I\!\!R^N} \frac{\partial}{\partial y_k} \mathrm{e}^{-i\,x\cdot y} f(x)\, \mathrm{d}x = i\, \frac{\partial}{\partial y_k} \hat{f}(y)\,, \end{aligned} \tag{3.7.4}$$

where for the last step we note that differentiation under the integral sign is permissible, because

$$|e^{-i\,x\cdot y}f(x)| \le |f(x)|\,, \qquad |x_k e^{-i\,x\cdot y}f(x)| \le |x_k f(x)|\,, \tag{3.7.5}$$

and both $|f|$ and $|f_k| \in L^1(I\!R^N)$. This proves (3.7.2). □

By repeated application of proposition 3.7.1, we obtain that if P is a polynomial with constant coefficients, then for all $f \in \mathcal{S}(I\!R^N)$,

$$[\mathcal{F}(P(D)f)](y) \;=\; P(i\,y)\hat{f}(y)\,, \tag{3.7.6}$$

$$[\mathcal{F}(P(\cdot)f)](y) \;=\; P(i\,D)\hat{f}(y)\,; \tag{3.7.7}$$

in particular,

$$\left[\mathcal{F}(D_x^\beta((\cdot)^\alpha f))\right](y) = (i\,y)^\beta[\mathcal{F}((\cdot)^\alpha f)](y) = i^{\alpha+\beta}y^\beta D_y^\alpha\,\hat{f}(y)\,. \tag{3.7.8}$$

We are now in a position to prove

Theorem 3.7.1 *The Fourier transform $\mathcal{F}$ maps $\mathcal{S}(I\!R^N)$ continuously into itself.*

Proof. We first show that $\mathcal{F}(\mathcal{S}) \subseteq \mathcal{S}$ (that is, that $\mathcal{F}$ maps $\mathcal{S}(I\!R^N)$ into itself) and then, that $\mathcal{F}$ is bounded; since $\mathcal{F}$ is linear, this implies its continuity. To show that $\hat{f} \in \mathcal{S}$ if $f \in \mathcal{S}(I\!R^N)$, recalling (2.7.1), we need to show that, for each α, $\beta \in I\!N$, the map

$$y \mapsto (y)^\beta\, D_y^\alpha\,\hat{f}(y) \tag{3.7.9}$$

is bounded. This is a consequence of (3.7.8): indeed, the function

$$x \mapsto D_x^\beta(x^\alpha f(x)) \tag{3.7.10}$$

is still in $\mathcal{S}(I\!R^N) \subset L^1(I\!R^N)$, so (3.7.8) implies that the function in (3.7.9) is bounded. More precisely, by (3.7.8) we can estimate

$$\sup_{y\in I\!R^N}\left|y^\beta D_y^\alpha \hat{f}(y)\right| = \sup_{y\in I\!R^N}\left|D_y^\beta\left[\mathcal{F}((\cdot)^\alpha f)\right](y)\right|$$

$$\le\ \sup_{y\in I\!R^N}\int_{I\!R^N}\left|e^{-i\,x\cdot y}D_x^\beta(x^\alpha f(x))\right|\,dx$$

$$\leq \int_{\mathbb{R}^N} (1+|x|^2)^{-N} \left|(1+|x|^2)^N D_x^\beta (x^\alpha f(x))\right| \mathrm{d}x \qquad (3.7.11)$$

$$\leq C\, p_{|\alpha|+2N,|\beta|}(f) \int_{\mathbb{R}^N} (1+|x|^2)^{-N} \mathrm{d}x$$

$$\leq C_N\, p_{|\alpha|+2N,|\beta|}(f)\,,$$

where the constant C_N depends only on N. □

2. Our second goal is to show that the Fourier transform maps $\mathcal{S}(\mathbb{R}^N)$ *onto* itself and, in fact, that $\mathcal{F}$ is a bijection of $\mathcal{S}(\mathbb{R}^N)$ into itself, whose inverse is exactly the map $\mathcal{F}^{-1}(\mathbb{R}^N)$, defined by (3.3.2). To this end, note first that the map $f \mapsto \check{f}$ does map $\mathcal{S}(\mathbb{R}^N)$ into itself, and is continuous (the proof of this result is analogous to that of theorem 3.7.1). To show that $\mathcal{F}$ is onto, we prove

Theorem 3.7.2 *Given $g \in \mathcal{S}(\mathbb{R}^N)$, let $f = \check{g}$ (thus, $f \in \mathcal{S}(\mathbb{R}^N)$). Then, $\hat{f} = g$ (that is, f is a counterimage of g by $\mathcal{F}$).*

Proof. Let E be the function defined in proposition 3.2.1. Recalling that $\check{E} = E$, for $\varepsilon > 0$ we compute

$$\begin{aligned}
&\int_{\mathbb{R}^N} \check{g}(x) E(\varepsilon x)\, \mathrm{e}^{-iy\cdot x}\, \mathrm{d}x \\
&\quad = c_N \int_{\mathbb{R}^N} E(\varepsilon x)\, \mathrm{e}^{-iy\cdot x} \int_{\mathbb{R}^N} \mathrm{e}^{ix\cdot z} g(z)\, \mathrm{d}z\, \mathrm{d}x \\
&\quad = c_N \int_{\mathbb{R}^N} g(z) \int_{\mathbb{R}^N} \mathrm{e}^{ix\cdot(z-y)}\, E(\varepsilon x)\, \mathrm{d}x\, \mathrm{d}z \\
&\quad = c_N \int_{\mathbb{R}^N} g(z) \int_{\mathbb{R}^N} \mathrm{e}^{i(\varepsilon x)\cdot((z-y)/\varepsilon)}\, E(\varepsilon x)\, \mathrm{d}x\, \mathrm{d}z \\
&\quad = \frac{c_N}{\varepsilon^N} \int_{\mathbb{R}^N} g(z) \int_{\mathbb{R}^N} \mathrm{e}^{iu\cdot((z-y)/\varepsilon)}\, E(u)\, \mathrm{d}u\, \mathrm{d}z \qquad (3.7.12) \\
&\quad = \tfrac{1}{\varepsilon^N} \int_{\mathbb{R}^N} g(z)\, \check{E}\left(\tfrac{z-y}{\varepsilon}\right) \mathrm{d}z \\
&\quad = \tfrac{1}{\varepsilon^N} \int_{\mathbb{R}^N} g(z)\, E\left(\tfrac{z-y}{\varepsilon}\right) \mathrm{d}z \\
&\quad = \int_{\mathbb{R}^N} g(y+\varepsilon z)\, E(z)\, \mathrm{d}z.
\end{aligned}$$

We now let $\varepsilon \to 0$ on both sides of this estimate, which we can do by the dominated convergence theorem, since

$$|\check{g}(x)\, E(\varepsilon x)\, \mathrm{e}^{-ix\cdot y}| \le |\check{g}(x)|\,, \tag{3.7.13}$$

with $\check{g} \in \mathcal{S}(\mathbb{R}^N) \hookrightarrow L^1(\mathbb{R}^N)$, and

$$|g(y+\varepsilon z)E(z)| \le \sup_{\mathbb{R}^N} |g(\cdot)|\, E(z)\,, \tag{3.7.14}$$

with $E \in L^1(\mathbb{R}^N)$. Thus, we obtain

$$\int_{\mathbb{R}^N} \check{g}(x)E(0)\, \mathrm{e}^{-ix\cdot y}\, \mathrm{d}y \le \int_{\mathbb{R}^N} g(y)E(z)\, \mathrm{d}z\,. \tag{3.7.15}$$

Recalling (3.2.2), we multiply both sides of (3.7.15) by c_N and obtain

$$c_N \int_{\mathbb{R}^N} \mathrm{e}^{-iy\cdot x}\, \check{g}(x)\, \mathrm{d}x = c_N\, g(y)\,; \tag{3.7.16}$$

that is, $(\check{g})\hat{}(y) = g(y)$, as claimed. $\square$

3. In conclusion, we have proven that $\mathcal{F}$ is a continuous bijection of $\mathcal{S}(\mathbb{R}^N)$ into itself, with a continuous inverse given by the inversion formula (3.3.2). With analogous techniques, we can prove

Proposition 3.7.2 *The map $\mathcal{F}$ has period 4 on $\mathcal{S}(\mathbb{R}^N)$, i.e. $\mathcal{F}^4 = I_{\mathcal{S}}$, the identity map on $\mathcal{S}(\mathbb{R}^N)$; moreover, identity (3.3.11) holds for all $f \in \mathcal{S}(\mathbb{R}^N)$, that is*

$$\mathcal{F}^2 f = \mathcal{F}(\hat{f}\,) = \tilde{f}\,, \tag{3.7.17}$$

where $\tilde{f}(x) := f(-x)$.

Proof. If (3.7.17) holds, then for each $f \in \mathcal{S}(\mathbb{R}^N)$, $\mathcal{F}^4 f = \mathcal{F}^2 \tilde{f} = \tilde{\tilde{f}} = f$: this proves that $\mathcal{F}^4 = I_{\mathcal{S}}$. To show (3.7.17), using the inversion formula (3.3.2) we compute that

$$\begin{aligned} [\mathcal{F}\hat{f}](x) &= c_N \int_{\mathbb{R}^N} \mathrm{e}^{-ix\cdot y}\, \hat{f}(y)\, \mathrm{d}y \\ &= c_N \int_{\mathbb{R}^N} \mathrm{e}^{i(-x)\cdot y}\, \hat{f}(y)\, \mathrm{d}y \\ &= [\mathcal{F}^{-1}\hat{f}](-x) = f(-x) = \tilde{f}(x)\,, \end{aligned} \tag{3.7.18}$$

as claimed. □

Finally, we remark that all the properties of the Fourier transform we have proven so far obviously hold in $\mathcal{S}(I\!R^N)$; in particular, this holds for the identities of propositions 3.3.1 and 3.4.1. For example, (3.7.17) is a rewriting of (3.3.11) in $\mathcal{S}(I\!R^N)$.

3.8 Fourier Transform in $\mathcal{S}'(I\!R^N)$.

1. We now define the Fourier transform on the space $\mathcal{S}'(I\!R^N)$ of tempered distributions introduced in §2.7.2. We provisionally denote by Φ the map $\Phi : \mathcal{S}'(I\!R^N) \to \mathcal{S}'(I\!R^N)$ defined by

$$< \Phi(T), \varphi >_{\mathcal{S}'\times\mathcal{S}} \; := \; < T, \hat{\varphi} >_{\mathcal{S}'\times\mathcal{S}}, \qquad T \in \mathcal{S}'(I\!R^N)\,, \; \varphi \in \mathcal{S}(I\!R^N)\,, \tag{3.8.1}$$

where the index $\mathcal{S}' \times \mathcal{S}$, which we omit in the sequel for brevity, denotes the duality pairing between $\mathcal{S}'$ and $\mathcal{S}$. The map Φ is well-defined, because theorem 3.7.1 guarantees that $\Phi(T)$, which is obviously linear, is also continuous on $\mathcal{S}(I\!R^N)$ (note that $\hat{\varphi} \in \mathcal{S}(I\!R^N)$, again by theorem 3.7.1). Definition (3.8.1) is motivated by the observation that if $T = T_f$, i.e. if T is the regular distribution generated by a function $f \in L^1(I\!R^N)$, then $\Phi(T) = T_{\hat{f}}$; i.e., $\Phi(T)$ is the regular distribution generated by the function $\hat{f}$, which is in $L^1(I\!R^N)$. Indeed, recalling (3.3.13) we compute that for all $\varphi \in \mathcal{S}(I\!R^N)$,

$$\begin{aligned} < \Phi(T_f), \varphi > \; &= \; < T_f, \hat{\varphi} > = \int_{I\!R^N} f(x)\hat{\varphi}(x)\,\mathrm{d}x \\ &= \int_{I\!R^N} \hat{f}(y)\varphi(y)\,\mathrm{d}y \; = < T_{\hat{f}}, \varphi > \; . \end{aligned} \tag{3.8.2}$$

This justifies calling $\Phi(T)$ the Fourier transform of T in $\mathcal{S}'(I\!R^N)$; that is, given the tempered distribution T, we define its Fourier transform $\hat{T} := \mathcal{F}(T)$ as the tempered distribution defined by (3.8.1), i.e.

$$< \hat{T}, \varphi > = < T, \hat{\varphi} >, \qquad \varphi \in \mathcal{S}(I\!R^N)\,. \tag{3.8.3}$$

In an analogous way, we can define a map $\mathcal{F}^{-1}$ on $\mathcal{S}'(I\!R^N)$ by $\mathcal{F}^{-1}(T) = \check{T}$, the latter being the tempered distribution defined by

$$< \check{T}, \varphi > = < T, \check{\varphi} >, \qquad \varphi \in \mathcal{S}(I\!R^N)\,; \tag{3.8.4}$$

the fact that this map is indeed the inverse of the Fourier transform map defined in $\mathcal{S}'(\mathbb{R}^N)$ by (3.8.3) follows from the invertibility of the Fourier transform in $\mathcal{S}(\mathbb{R}^N)$. We confirm this in

Theorem 3.8.1 *Let $\mathcal{F}$ be the Fourier transform in $\mathcal{S}'(\mathbb{R}^N)$ defined in* (3.8.3)*. Then $\mathcal{F}$ is a linear, continuous bijection of $\mathcal{S}'(\mathbb{R}^N)$, endowed with the topology of weak* convergence, into itself. The inverse map $\mathcal{F}^{-1}$ is defined by* (3.8.4)*, and is continuous. Moreover, the analogous of proposition 3.7.2 holds, i.e. $\mathcal{F}^4 = I_{\mathcal{S}'}$ (the identity in $\mathcal{S}'(\mathbb{R}^N)$), and $\mathcal{F}^2 T = \tilde{T}$, the tempered distribution defined by*

$$< \tilde{T}, \varphi > = < T, \tilde{f} > \qquad \forall\, \varphi \in \mathcal{S}(\mathbb{R}^N). \tag{3.8.5}$$

In addition, if P is a polynomial with constant coefficients, then for all $T \in \mathcal{S}'(\mathbb{R}^N)$

$$\mathcal{F}(P(D)T) = P(i\cdot)\hat{T}, \qquad \mathcal{F}(P(\cdot)T) = P(iD)\hat{T} \tag{3.8.6}$$

(recall that the distribution $P(\cdot)T$ is defined as in (2.5.2)*).*

Proof. All results are an immediate consequence of the fact that the analogous statements are true in $\mathcal{S}(\mathbb{R}^N)$. For example, to show that $\mathcal{F}$ is onto, given $L \in \mathcal{S}'(\mathbb{R}^N)$ let $T = \check{L}$: then $T \in \mathcal{S}'(\mathbb{R}^N)$, and $\hat{T} = L$, because for all $\varphi \in \mathcal{S}(\mathbb{R}^N)$,

$$< \hat{T}, \varphi > = < T, \hat{\varphi} > = < \check{L}, \hat{\varphi} > = < L, \mathcal{F}^{-1}\hat{\varphi} > = < L, \varphi > \ . \tag{3.8.7}$$

In particular, the continuity of $\mathcal{F}^{-1}$ follows from that of $\mathcal{F}$, via the identity $\mathcal{F}^{-1} = \mathcal{F}^3$. As for (3.8.6), these identities follow from the analogous identities (3.7.6) and (3.7.7) in $\mathcal{S}(\mathbb{R}^N)$. For example, from (3.7.2) we obtain that

$$\begin{aligned} < \mathcal{F}[\partial_j T], \varphi > &= < \partial_j T, \hat{\varphi} > = - < T, \partial_j \hat{\varphi} > \\ &= -\tfrac{1}{i} < T, \mathcal{F}[(\cdot)_j \varphi] > = i < \hat{T}, (\cdot)_j \varphi > \\ &= i < (\cdot)_j \hat{T}, \varphi >\,, \end{aligned} \tag{3.8.8}$$

which shows that $\mathcal{F}[\partial_j T] = i(\cdot)_j \hat{T}$, in accord with the first of (3.8.6). □

2. As an important application, we show

Proposition 3.8.1 *Let P be a polynomial with constant coefficients, and define $T_P \in \mathcal{S}'(I\!R^N)$ by*

$$< T_P, \varphi > = c_N \int_{I\!R^N} P(x)\,\varphi(x)\,\mathrm{d}x, \qquad \forall\,\varphi \in \mathcal{S}(I\!R^N) \tag{3.8.9}$$

(*a slight modification of the usual definition of T_P*). *Then,*

$$\mathcal{F}(T_P) = P(iD)\delta_0, \qquad \mathcal{F}(P(-iD)\delta_0) = T_P\,, \tag{3.8.10}$$

where δ_0 is the Dirac δ-distribution on $I\!R^N$ (*which is a tempered distribution*).

Proof. (a) If $P(x) \equiv 1$, then for all $\varphi \in \mathcal{S}(I\!R^N)$

$$\begin{aligned} < \mathcal{F}(T_1), \varphi > &= < T_1, \hat{\varphi} > = c_N \int_{I\!R^N} 1\,\hat{\varphi}(x)\,\mathrm{d}x \\ &= c_N \int_{I\!R^N} \mathrm{e}^{i\,0\cdot x}\hat{\varphi}(x)\,\mathrm{d}x = \varphi(0) = < \delta_0, \varphi >\,, \end{aligned} \tag{3.8.11}$$

in accord with the first of (3.8.10). Analogously,

$$\begin{aligned} < \mathcal{F}(\delta_0), \varphi > &= < \delta_0, \hat{\varphi} > = \hat{\varphi}(0) = c_N \int_{I\!R^N} \mathrm{e}^{-i\,0\cdot y}\varphi(y)\,\mathrm{d}y \\ &= c_N \int_{I\!R^N} 1\,\varphi(y)\,\mathrm{d}y = < T_1, \varphi >\,, \end{aligned} \tag{3.8.12}$$

in accord with the second of (3.8.10).

(b) More generally, if P is an arbitrary polynomial with constant coefficients, by (3.8.6) we compute that

$$\mathcal{F}(T_P) = \mathcal{F}(P(\cdot)T_1) = P(iD)(\mathcal{F}(T_1)) = P(iD)\delta_0\,, \tag{3.8.13}$$

$$\mathcal{F}(P(-iD)\delta_0) = P((-i)(i\cdot))(\mathcal{F}(\delta_0)) = P(\cdot)T_1 = T_P\,, \tag{3.8.14}$$

from which (3.8.10) follows. □

We remark explicitly that part (a) of the proof of proposition 3.8.1 shows that

$$\delta_0 = \mathcal{F}(\mathcal{I}), \qquad \mathcal{F}(\delta_0) = \mathcal{I}\,, \tag{3.8.15}$$

where $\mathcal{I}$ denotes the distribution T_1 generated by the constant polynomial $P(x) \equiv 1$.

3. We conclude with the following result.

Proposition 3.8.2 *Let $T \in \mathcal{S}'(I\!R^N)$ be such that $\hat{T} \in L^2(I\!R^N)$. Then $T \in L^2(I\!R^N)$.*

Remark. Per se, the assumption that $T \in \mathcal{S}'(I\!R^N)$ only implies that $\hat{T} \in \mathcal{S}'(I\!R^N)$. But as $L^2(I\!R^N) \subset \mathcal{S}'(I\!R^N)$, it may or may not happen that $\hat{T} \in L^2(I\!R^N)$. Proposition 3.8.2 states that if this is the case, then in fact also the original tempered distribution T is regular; that is, $T \in L^2(I\!R^N)$. $\diamond$

Proof. To say that $\hat{T} \in L^2(I\!R^N)$ means that there is $g \in L^2(I\!R^N)$ such that $\hat{T}$ is the distribution generated by g; that is, with a small abuse of notation, that $\hat{T} = \hat{T}_g$. More precisely, for all $\varphi \in \mathcal{S}(I\!R^N)$,

$$< \hat{T}, \varphi > = \int_{I\!R^N} g(x)\varphi(x)\,\mathrm{d}x\,. \tag{3.8.16}$$

We now claim that T is the distribution generated by $\check{g}$; that is, that $T = T_{\check{g}}$. Since $\check{g} \in L^2(I\!R^N)$, this proves the proposition. Given then $\varphi \in \mathcal{S}(I\!R^N)$, we compute that

$$< \hat{g}, \varphi > = < g, \check{\varphi} > = < \hat{T}, \check{\varphi} > = < T, \mathcal{F}[\check{\varphi}] > = < T, \varphi >\,, \tag{3.8.17}$$

which proves our claim. $\square$

3.9 Applications to PDEs of Evolution.

In this section we apply the results we have seen on the Fourier transform to solve the initial-value problems for the linear heat and the wave equations in $I\!R^N$.

3.9.1 The Heat Equation.

Given a function $u_0 \in L^1(I\!R^N)$, we seek to find a function $u : I\!R_{\geq 0} \times I\!R^N \to I\!R$ which solves the Cauchy problem consisting of the heat equation

$$u_t - \Delta\, u = 0 \qquad \text{in} \quad I\!R_{>0} \times I\!R^N\,, \tag{3.9.1}$$

together with the initial condition

$$u(0, \cdot) = u_0 \qquad \text{in} \quad \{t = 0\} \times I\!R^N\,. \tag{3.9.2}$$

Theorem 3.9.1 *Let*

$$u(t,x) := \begin{cases} \frac{1}{(4\pi t)^{N/2}} \int_{\mathbb{R}^N} u_0(y)\, \mathrm{e}^{-|x-y|^2/4t}\, \mathrm{d}y & \text{if} \quad t > 0\,, \\ u_0(x) & \text{if} \quad t = 0\,. \end{cases} \tag{3.9.3}$$

Then, $u \in C^\infty(\,]0,+\infty[\,\times \mathbb{R}^N)$*;* u *solves the heat equation* (3.9.1) *for all* $(t,x) \in\,]0,+\infty[\,\times\mathbb{R}^N$*, and takes on the initial value* (3.9.2) *in the generalized sense that*

$$\lim_{t\to 0} u(t,\cdot) = u_0 \qquad \text{in} \quad L^1(\mathbb{R}^N)\,. \tag{3.9.4}$$

If $u_0 \in C_{\mathrm{b}}(\mathbb{R}^N)$*, then* $u \in C([0,+\infty[\,\times\mathbb{R}^N)$ *as well, and*

$$\lim_{(t,x)\to(0,x_0)} u(t,x) = u_0(x_0) \tag{3.9.5}$$

for all $x_0 \in \mathbb{R}^N$*.*

Proof. We first proceed formally, keeping $t > 0$ fixed and taking the Fourier transform of the two terms of equation (3.9.1) with respect to the space variable. Recalling (3.5.13), we obtain that the function

$$[\mathcal{F}u(t,\cdot)](y) = \hat{u}(t,y) \tag{3.9.6}$$

should solve the equation

$$\hat{u}_t + |y|^2\, \hat{u} = 0\,. \tag{3.9.7}$$

We look at this equation as a family of ODEs in the unknowns $t \mapsto \hat{u}(t,y)$, parametrized by $y \in \mathbb{R}^N$, and attach to these ODEs the natural initial conditions

$$\hat{u}(0,\cdot) = \hat{u}_0\,. \tag{3.9.8}$$

Thus, we are led to consider, for each $y \in \mathbb{R}^N$, the Cauchy problem

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\, v + |y|^2\, v = 0\,, \\ v(0) = \hat{u}_0(y)\,, \end{cases} \tag{3.9.9}$$

which has the solution

$$v(t,y) = \hat{u}_0(y)\, \mathrm{e}^{-|y|^2\, t}\,. \tag{3.9.10}$$

Recalling (3.2.20), we deduce that

$$v(t,y) = \tfrac{1}{c_N}\,\widehat{E_{(t)}}(y) = [\mathcal{F}(u_0 * E_{(t)})](y)\,, \tag{3.9.11}$$

where $E_{(t)}$ is the heat kernel defined in (3.2.17). For $t > 0$, we now *define* u by

$$u(t,x) := \left[\mathcal{F}^{-1}(v(t,\cdot)\right](x) = [u_0 * E_{(t)}](x)\,, \tag{3.9.12}$$

and, recalling (3.2.18), we see that u is the desired solution of the heat equation (3.9.1). The generalized taking of the initial condition (3.9.4) is a consequence of (3.2.23) of proposition 3.2.2. For (3.9.5), see, e.g., Evans, [3, sct. 2.3.b]. Note that (3.9.3) defines u uniquely, because if $z := u - \tilde{u}$ is the difference of two solutions corresponding to the same initial value u_0, then z satisfies the heat equation (3.9.1), with initial condition $z(0,\cdot) = 0$; replacing $z_0 = 0$ in (3.9.3), written with u replaced by z, we deduce that $z = 0$. □

3.9.2 The Wave Equation.

Given two functions $u_0,\, u_1 \in L^1(\mathbb{R}^N)$, we seek to find a function $u : \mathbb{R}_{\geq 0} \times \mathbb{R}^N \to \mathbb{R}$ which solves the Cauchy problem consisting of the wave equation

$$u_{tt} - \Delta\, u = 0 \qquad \text{in} \quad \mathbb{R}_{>0} \times \mathbb{R}^N\,, \tag{3.9.13}$$

together with the initial conditions

$$u(0,\cdot) = u_0\,, \quad u_t(0,\cdot) = u_1 \qquad \text{in} \quad \{t = 0\} \times \mathbb{R}^N\,. \tag{3.9.14}$$

Reasoning as in the previous section, we proceed formally, keeping $t > 0$ fixed and taking the Fourier transform of the two terms of equation (3.9.13) with respect to the space variable. Recalling (3.5.13), we obtain that the function $\hat{u}$ defined in (3.9.6) should now solve the equation

$$\hat{u}_{tt} + |y|^2\,\hat{u} = 0\,. \tag{3.9.15}$$

Again, we look at this equation as a family of ODEs in the unknowns $t \mapsto \hat{u}(t,y)$, parametrized by $y \in \mathbb{R}^N$, and attach to these ODEs the natural initial conditions

$$\hat{u}(0,\cdot) = \hat{u}_0\,, \qquad \hat{u}_t(0) = \hat{u}_1\,. \tag{3.9.16}$$

Thus, we are led to consider, for each $y \in I\!R^N$, the Cauchy problem

$$\begin{cases} \frac{\mathrm{d}^2}{\mathrm{d}t^2}\, v + |y|^2\, v = 0\,, \\ v(0) = \hat{u}_0(y)\,, \quad v\,'(0) = \hat{u}_1(y)\,, \end{cases} \tag{3.9.17}$$

which has the solution

$$v(t,y) = \begin{cases} \hat{u}_1(y)\, \frac{\sin(|y|\, t)}{|y|} + \hat{u}_0(y)\, \cos(|y|\, t) & \text{if} \quad y \neq 0\,, \\ \hat{u}_1(0)\, t + \hat{u}_0(0) & \text{if} \quad y = 0\,. \end{cases} \tag{3.9.18}$$

As in (3.9.12), we now define

$$u(t,x) := \left[\mathcal{F}^{-1}(v(t,\cdot)\right](x)\,, \tag{3.9.19}$$

and obtain [4]

Theorem 3.9.2 *Let u_0, $u_1 \in L^1(I\!R^N)$, and define u by* (3.9.19)*. Then, $v \in C^1([0, +\infty[\,, L^1(I\!R^N))$, and solves the initial-value problem* (3.9.13)+(3.9.14).

To find an explicit expression for u, involving the inverse Fourier transforms of the functions appearing at the right side of (3.9.19), we refer to Torchinsky's paper [9], where, with some candid understating, of this problem it is said that "it is not easy". Here, we limit ourselves to consider the one-dimensional case, and show that the function u defined in (3.9.19) coincides with the solution of the wave equation given by d'Alembert's formula (2.1.10). Noting that the functions

$$x \mapsto \frac{\sin(tx)}{x} \qquad \text{and} \qquad x \mapsto \cos(tx) \tag{3.9.20}$$

are even in x, from (3.9.18) and (3.1.2) we obtain that, if $y \neq 0$,

$$\begin{aligned} v(t,y) &= \hat{u}_1(y)\, \frac{\sin(y\,t)}{y} + \hat{u}_0(y)\, \cos(y\,t) \\ &= \sqrt{\tfrac{\pi}{2}}\, \hat{u}_1(y)\, \hat{\chi}_{[-t,t]}(y) + \hat{u}_0(y)\, \cos(y\,t)\,. \end{aligned} \tag{3.9.21}$$

[4]The reader is encouraged to work out the proof of theorem 3.9.2.

Assume for the moment that $u_0 = 0$, and call $\tilde{v}$ and $\tilde{u}$ the corresponding functions, defined by (3.9.18) (3.9.19). Recalling the first of (3.4.3), we find that

$$v(t,y) = c_1 \sqrt{\tfrac{\pi}{2}} \left[\mathcal{F}(u_1 * \chi_{[-t,t]})\right](y) = \tfrac{1}{2} \left[\mathcal{F}(u_1 * \chi_{[-t,t]})\right](y) \,, \tag{3.9.22}$$

from which

$$\begin{aligned} \tilde{u}(t,x) &= \left[\mathcal{F}^{-1}\,\tilde{v}(t,\cdot)\right](x) = \tfrac{1}{2}\left[u_1 * \chi_{[-t,t]}\right](x) \\ &= \tfrac{1}{2}\int_{-\infty}^{+\infty} u_1(y)\,\chi_{[-t,t]}(x-y)\,\mathrm{d}y \\ &= \tfrac{1}{2}\int_{x-t}^{x+t} u_1(y)\,\mathrm{d}y \,, \end{aligned} \tag{3.9.23}$$

having noted that

$$\chi_{[-t,t]}(x-y) = \begin{cases} 1 & \text{if} \quad |x-y| \le t \,, \\ 0 & \text{if} \quad |x-y| > t \,, \end{cases} \tag{3.9.24}$$

that is,

$$\chi_{[-t,t]}(x-y) = \begin{cases} 1 & \text{if} \quad x-t \le y \le x+t \,, \\ 0 & \quad\text{otherwise} \,. \end{cases} \tag{3.9.25}$$

If $u_0 \neq 0$, we note that we can decompose [5] the solution of the wave equation as

$$u = u^1 + u^2_t \,, \tag{3.9.26}$$

where u^1 and u^2 are the solutions of the Cauchy problems

$$u^i_{tt} - u^i_{xx} = 0 \,, \qquad i = 1,\,,2 \,, \tag{3.9.27}$$

with initial values

$$u^1(0) = 0 \,, \quad u^1_t(0) = u_1 \,, \qquad u^2(0) = 0 \,, \quad u^2_t(0) = u_0 \,. \tag{3.9.28}$$

In fact, clearly

$$u_{tt} - u_{xx} = (u^1_{tt} - u^1_{xx}) + (u^2_{tt} - u^2_{xx})_t = 0 \,; \tag{3.9.29}$$

[5] Decomposition (3.9.26) holds in any space dimension.

moreover, by (3.9.28),

$$u(0) = u^1(0) + u_t^2(0) = 0 + u_0\,, \tag{3.9.30}$$

$$u_t(0) = u_t^1(0) + u_{tt}^2(0) = u_t^1(0) + u_{xx}^2(0) = u_1 + 0\,. \tag{3.9.31}$$

Consequently, from (3.9.23),

$$\begin{aligned} u(t,x) &= \tfrac{1}{2}\left(\int_{x-t}^{x+t} u_1(y)\,\mathrm{d}y + \frac{\partial}{\partial t}\int_{x-t}^{x+t} u_0(y)\,\mathrm{d}y\right) \\ &= \tfrac{1}{2}\left(\int_{x-t}^{x+t} u_1(y)\,\mathrm{d}y + u_0(x+t) + u_0(x-t)\right)\,, \end{aligned} \tag{3.9.32}$$

which is d'Alembert's formula (2.1.10). □

Chapter 4

Notes on the Laplace Transform.

In this chapter we briefly recall some basic results on the Laplace transform in $\mathbb{R}$. Most of the material of this chapter is taken from the notes of J. Seiler, [8].

4.1 Definition and Basic Properties.

1. We start with the formal definition of the Laplace transform. Given a complex number z, we denote by $\Re(z)$, $\Im(z)$, and $\bar{z}$, respectively, the real part, the imaginary part, and the conjugate of z.

Definition 4.1.1 *1) A function $f : \mathbb{R}_{>0} \to \mathbb{C}$ is said to be $\mathcal{L}$-transformable if there is $\sigma \in \mathbb{R}_{>0}$ such that the function*

$$x \;\mapsto\; \mathrm{e}^{-\sigma x} f(x) \quad \text{is in} \quad L^1(0,+\infty)\,; \tag{4.1.1}$$

we set

$$\sigma_f := \inf\{\sigma \in \mathbb{R}_{>0} \;\mid\; \text{(4.1.1) holds}\}\,. \tag{4.1.2}$$

2) The Laplace transform of a $\mathcal{L}$-transformable function f is the function $\mathcal{L}f : \mathbb{C} \to \mathbb{C}$ defined, on the half-plane

$$\mathbb{C}_f := \{z \in \mathbb{C} \;\mid\; \Re(z) > \sigma_f\}\,, \tag{4.1.3}$$

by

$$[\mathcal{L}f](z) := \int_0^{+\infty} \mathrm{e}^{-zx}\, f(x)\,\mathrm{d}x\,. \tag{4.1.4}$$

3) A continuous function $f : I\!R_{\geq 0} \to \mathbb{C}$ is of exponential order $\sigma > 0$, if there are M and $T > 0$ such that

$$|f(t)| \leq M\,\mathrm{e}^{\sigma\,t} \tag{4.1.5}$$

for all $t \geq T$.

Remarks. 1) The function $\mathcal{L}f$ is indeed defined for $\Re(z) > \sigma_f$, since in this case

$$|\mathrm{e}^{-zx}| = \mathrm{e}^{-x\Re(z)} \leq \mathrm{e}^{-x\,\sigma_f}\,, \tag{4.1.6}$$

so that the integral in (4.1.4) converges.
2) The operator $\mathcal{L}$ is obviously linear over $\mathbb{C}$; more precisely, for all $\mathcal{L}$-transformable functions f and g, and all $a \in \mathbb{C}$,

$$\mathcal{L}(f + a\,g) = \mathcal{L}f + a\,\mathcal{L}g \tag{4.1.7}$$

on the half-plane $\{z \in \mathbb{C} \;\mid\; \Re(z) > \max\{\sigma_f, \sigma_g\}$.
3) If f is of exponential order $\sigma \geq \sigma_f$, f is $\mathcal{L}$-transformable, because for all $\tilde{\sigma} > \sigma$, the function $\mathrm{e}^{-\tilde{\sigma}(\cdot)}\,f$ is in $L^1(0,+\infty)$. Indeed,

$$\begin{aligned}
\int_0^{+\infty} \mathrm{e}^{-\tilde{\sigma}t}\,|f(t)|\,\mathrm{d}t &= \int_0^{T} \mathrm{e}^{-\tilde{\sigma}t}\,|f(t)|\,\mathrm{d}t + \int_T^{+\infty} \mathrm{e}^{-(\tilde{\sigma}-\sigma)t}\,\mathrm{e}^{-\sigma t}\,|f(t)|\,\mathrm{d}t \\
&\leq \int_0^{T} |f(t)|\,\mathrm{d}t + M\int_T^{+\infty} \mathrm{e}^{-(\tilde{\sigma}-\sigma)t}\,\mathrm{d}t \qquad (4.1.8) \\
&= \int_0^{T} |f(t)|\,\mathrm{d}t + \tfrac{M}{\tilde{\sigma}-\sigma}\,\mathrm{e}^{-(\sigma-\sigma)\,T}\,;
\end{aligned}$$

thus, the integral at the left side of (4.1.8) is finite. $\diamond$

2. The following examples are fundamental for the sequel.

Example 4.1.1 *Let $a \in \mathbb{C}$. Then, for $\Re(z) > \Re(a)$,*

$$[\mathcal{L}(\mathrm{e}^{a(\cdot)})](z) = \frac{1}{z-a}\,. \tag{4.1.9}$$

In particular, for $a = 0$,

$$[\mathcal{L}\,1](z) = \frac{1}{z}\,, \qquad \Re(z) > 0\,. \tag{4.1.10}$$

Proof. By the definition (4.1.4),

$$\begin{aligned}[\mathcal{L}(\mathrm{e}^{a(\cdot)}](z) &= \int_0^{+\infty} \mathrm{e}^{(a-z)\,x}\,\mathrm{d}x \lim_{\substack{m\to+\infty\\ \varepsilon\to 0}} \int_\varepsilon^m \mathrm{e}^{(a-z)\,x}\,\mathrm{d}x \\ &= \lim_{\substack{m\to+\infty\\ \varepsilon\to 0}} \left(\frac{\mathrm{e}^{(a-z)\,m}}{a-z} - \frac{\mathrm{e}^{(a-z)\,\varepsilon}}{a-z}\right) .\end{aligned} \tag{4.1.11}$$

Since

$$|\mathrm{e}^{(a-z)\,m}| = \mathrm{e}^{m\,\Re(a-z)} \;\to\; 0 \qquad \text{as} \quad m\to+\infty \tag{4.1.12}$$

because $\Re(z) > \Re(a)$, from (4.1.11) it follows that

$$[\mathcal{L}(\mathrm{e}^{a(\cdot)})](z) = \frac{-\,1}{a-z}\,, \tag{4.1.13}$$

which is (4.1.9), and yields (4.1.10) for $a=0$. □

Example 4.1.2 *Let $\omega \in I\!R$. Then, for $\Re(z) > 0$,*

$$[\mathcal{L}(\cos(\omega(\cdot)))](z) = \frac{z}{z^2+\omega^2}\,, \qquad [\mathcal{L}(\sin(\omega(\cdot)))](z) = \frac{\omega}{z^2+\omega^2}\,. \tag{4.1.14}$$

Proof. We recall the Euler formulas

$$\cos(\omega\,x) = \tfrac{1}{2}\,(\mathrm{e}^{i\,\omega x} + \mathrm{e}^{-\,i\,\omega x})\,, \qquad \sin(\omega\,x) = \tfrac{1}{2i}\,(\mathrm{e}^{i\,\omega x} - \mathrm{e}^{-\,i\,\omega x})\,. \tag{4.1.15}$$

Then, by (4.1.9), with $a = \pm\,i\,\omega$, we find that

$$[\mathcal{L}(\cos(\omega(\cdot)))](z) = \frac{1}{2}\left(\frac{1}{z-i\,\omega} + \frac{1}{z+i\,\omega}\right)\,, \tag{4.1.16}$$

$$[\mathcal{L}(\sin(\omega(\cdot)))](z) = \frac{1}{2i}\left(\frac{1}{z-i\,\omega} - \frac{1}{z+i\,\omega}\right)\,, \tag{4.1.17}$$

from which (4.1.14) follows. □

Example 4.1.3 *Let $k \in I\!N$. Then, for $\Re(z) > 0$,*

$$[\mathcal{L}((\cdot)^k)](z) = \frac{k\,!}{z^{k+1}}\,. \tag{4.1.18}$$

Proof. We proceed by induction on k. The case $k = 0$ is (4.1.10). Assuming (4.1.18) true for some $k \geq 0$, we compute

$$
\begin{aligned}
[\mathcal{L}((\cdot)^{k+1})](z) &= \int_0^{+\infty} \mathrm{e}^{-zx}\, x^{k+1}\, \mathrm{d}x \\
&= \Big[-\tfrac{1}{z}\, x^{k+1}\, \mathrm{e}^{-zx}\Big]_{x=0}^{x=+\infty} + \tfrac{1}{z}\int_0^{+\infty} (k+1)\, \mathrm{e}^{-zx}\, x^k\, \mathrm{d}x \qquad (4.1.19) \\
&= 0 + \tfrac{1}{z}\,(k+1)\,[\mathcal{L}((\cdot)^k)](z) = \tfrac{k+1}{z}\, \tfrac{k\,!}{z^k}\,,
\end{aligned}
$$

form which (4.1.18) follows for k replaced by $k+1$. □

The following results are of immediate proof.

Proposition 4.1.1 *Let f be $\mathcal{L}$-transformable, $a \in \mathbb{C}$, and $n \in \mathbb{N}$. Then,*

$$
[\mathcal{L}(\mathrm{e}^{a(\cdot)}\, f)](z) = [\mathcal{L}f](z-a)\,, \qquad \Re(z-a) > \sigma_f\,, \tag{4.1.20}
$$

$$
\frac{\mathrm{d}^n}{\mathrm{d}z^n}\,[\mathcal{L}f](z) = (-1)^n\,[\mathcal{L}((\cdot)^n\, f)](z)\,, \quad \Re(z) > \sigma_f\,. \tag{4.1.21}
$$

Conversely, if $f \in C^m([0,+\infty[\to \mathbb{C})$, and f and all its derivatives $f^{(k)}$, $1 \leq k \leq m$, are of exponential order $\sigma > \sigma_f$, then, for $1 \leq k \leq m$,

$$
[\mathcal{L}(f^{(k)})](z) = z^k\,[\mathcal{L}f](z) - \sum_{i=1}^{k} f^{(i-1)}(0)\, z^{k-i}\,, \qquad \Re(z) > \sigma\,. \tag{4.1.22}
$$

Proof. 1) We compute

$$
\begin{aligned}
[\mathcal{L}(\mathrm{e}^{a(\cdot)}\, f)](z) &= \int_0^{+\infty} \mathrm{e}^{-zx}\,(\mathrm{e}^{ax}\, f(x))\, \mathrm{d}x \\
&= \int_0^{+\infty} \mathrm{e}^{-(z-a)}\, f(x)\, \mathrm{d}x\,.
\end{aligned} \tag{4.1.23}
$$

2) Similarly,

$$
\begin{aligned}
\frac{\mathrm{d}^n}{\mathrm{d}z^n}\,[\mathcal{L}f](z) &= \frac{\mathrm{d}^n}{\mathrm{d}z^n}\int_0^{+\infty} \mathrm{e}^{-zx}\, f(x)\, \mathrm{d}x \\
&= \int_0^{+\infty} (-x)^n\, \mathrm{e}^{-zx}\, f(x)\, \mathrm{d}x\,,
\end{aligned} \tag{4.1.24}
$$

from which the first two claims follow.

3) We first note that each function $f^{(k)}$, $0 \le k \le m$, is $\mathcal{L}$-transformable, as seen in the third remark after definition 4.1.1. We prove (4.1.22) by induction on k.
i) For $k = 1$, we compute that

$$\begin{aligned} [\mathcal{L}(f')](z) &= \int_0^{+\infty} \mathrm{e}^{-zx} f'(x)\,\mathrm{d}x \\ &= \left[\mathrm{e}^{-zx} f(x)\right]_{x=0}^{x=+\infty} + \int_0^{+\infty} z\,\mathrm{e}^{-zx} f(x)\,\mathrm{d}x\,. \end{aligned} \tag{4.1.25}$$

Now, we realize that, since f is of exponential order σ, for large enough x (determined by (4.1.5)) we can estimate

$$\begin{aligned} |\mathrm{e}^{-zx} f(x)| &= \mathrm{e}^{-x\,\Re(z)}\,|f(x)| \\ &= \mathrm{e}^{-x(\Re(z)-\sigma)}\,\mathrm{e}^{-x\,\sigma}\,|f(x)| \le M\,\mathrm{e}^{-x(\Re(z)-\sigma)}\,; \end{aligned} \tag{4.1.26}$$

thus, since $\Re(z) > \sigma$,

$$\mathrm{e}^{-zx} f(x) \to 0 \qquad \text{as} \quad x \to +\infty\,. \tag{4.1.27}$$

Hence, we obtain from (4.1.25) that

$$[\mathcal{L}(f')](z) = \int_0^{+\infty} z\,\mathrm{e}^{-zx} f(x)\,\mathrm{d}x - f(0)\,, \tag{4.1.28}$$

which is (4.1.22) for $k = 1$.
ii) Similarly, if $1 \le k \le m-1$,

$$\begin{aligned} \left[\mathcal{L}(f^{(k+1)})\right](z) &= \int_0^{+\infty} \mathrm{e}^{-zx} f^{(k+1)}(x)\,\mathrm{d}x \\ &= \left[\mathrm{e}^{-zx} f^{(k)}(x)\right]_{x=0}^{x=+\infty} + \int_0^{+\infty} z\,\mathrm{e}^{-zx} f^{(k)}(x)\,\mathrm{d}x \\ &= -f^{(k)}(0) + z\left[\mathcal{L}(f^{(k)})\right](z) \\ &= z\,z^k\,[\mathcal{L}f](z) - \sum_{i=1}^{k} f^{(i-1)}(0)\,z^{k-i} - f^{(k)}(0) \\ &= z^{k+1}\,[\mathcal{L}](z) - \sum_{i=1}^{k+1} f^{(i-1)}(0)\,z^{k+1-i}\,, \end{aligned} \tag{4.1.29}$$

which is (4.1.22) for k replaced by $k+1$. □

4.2 The Inverse Laplace Transform.

We proceed to give a formal definition of the inverse Laplace transform of a holomorphic function.

1. We start by observing that, given a $\mathcal{L}$-transformable function f, $\sigma > \sigma_f$, and $\tau \in I\!R$,

$$[\mathcal{L}f](\sigma + i\,\tau) = \int_0^{+\infty} \mathrm{e}^{-x\,(\sigma+i\,\tau)}\, f(x)\, \mathrm{d}x = \int_0^{+\infty} \mathrm{e}^{-\,i\,x\tau}\, (\mathrm{e}^{-\,x\sigma}\, f(x))\, \mathrm{d}x\;; \tag{4.2.1}$$

thus, defining $\tilde{f} : I\!R \to I\!R$ by

$$\tilde{f}(x) := \begin{cases} 0 & \text{if} \quad x < 0\,, \\ f(x) & \text{if} \quad x > 0\,, \end{cases} \tag{4.2.2}$$

and noting that the function $x \mapsto \mathrm{e}^{-\,\sigma x}\tilde{f}(x)$ is in $L^1(I\!R)$ (because $\sigma > \sigma_f$), we can write

$$[\mathcal{L}f](\sigma + i\,\tau) = \int_{-\infty}^{+\infty} \mathrm{e}^{-\,i\,x\tau}\, (\mathrm{e}^{-\,x\sigma}\, \tilde{f}(x))\, \mathrm{d}x = \sqrt{2\pi}\, [\mathcal{F}(\mathrm{e}^{-\,\sigma(\cdot)}\, \tilde{f})](\tau)\,. \tag{4.2.3}$$

For fixed $\sigma > \sigma_f$, define

$$g_\sigma(\tau) := [\mathcal{L}f](\sigma + i\,\tau)\,. \tag{4.2.4}$$

Then, if $g_\sigma \in L^1(I\!R)$, (4.2.3) yields the formal identity

$$\begin{aligned} \mathrm{e}^{-\,\sigma x}\, \tilde{f}(x) &= \left[\mathcal{F}^{-1}\left(\tfrac{1}{\sqrt{2\pi}}\, g_\sigma\right)\right](x) \\ &= \tfrac{1}{\sqrt{2\pi}}\, \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{i\,\tau x}\, g_\sigma(\tau)\, \mathrm{d}\tau \\ &= \tfrac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{e}^{i\,\tau x}\, [\mathcal{L}f](\sigma + i\,\tau)\, \mathrm{d}\tau\,, \end{aligned} \tag{4.2.5}$$

from which, restricting to $x > 0$,

$$f(x) = \tfrac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{e}^{(\sigma+i\,\tau)\,x}\, [\mathcal{L}f](\sigma + i\,\tau)\, \mathrm{d}\tau\,, \tag{4.2.6}$$

for $\sigma > \sigma_f$. We interpret the integral at the right side of (4.2.6) as a line integral for the function

$$\mathbb{C} \ni z \;\mapsto\; \mathrm{e}^{zx}\, [\mathcal{L}f](z) \in \mathbb{C}\,, \tag{4.2.7}$$

over the vertical line

$$\gamma_\sigma := \{z \in \mathbb{C} \ \mid \ \Re(z) = \sigma\} , \qquad \sigma > \sigma_f . \tag{4.2.8}$$

Since this line can be parametrized by

$$z = z(\tau) = \sigma + i\,\tau , \qquad -\infty < \tau < +\infty \tag{4.2.9}$$

(that is, σ is fixed, and the parameter τ varies in all of $\mathbb{R}$), and $\frac{\mathrm{d}}{\mathrm{d}\tau} z(\tau) = i$, it follows from (4.2.6) that, if $x > 0$,

$$f(x) = \frac{1}{2\pi i} \int_{\gamma_\sigma} \mathrm{e}^{zx} \, [\mathcal{L}f](z) \, \mathrm{d}z . \tag{4.2.10}$$

We point out explicitly that identity (4.2.10) holds, irrespective of the value of σ, as long as $\sigma > \sigma_f$ (see, e.g., Brown and Churchill, [2, ch. 7, sct. 66]).

2. A first consequence of (4.2.10) is that the Laplace transform is injective. More precisely, assume that f and g are $\mathcal{L}$-transformable, and $\mathcal{L}f = \mathcal{L}g$ on some vertical line (4.2.8) with $\sigma > \max\{\sigma_f, \sigma_g\}$. Then, by (4.2.10) and the linearity of $\mathcal{L}$,

$$f(x) - g(x) = \frac{1}{2\pi i} \int_{\gamma_\sigma} \mathrm{e}^{zx} \, \left([\mathcal{L}f](z) - [\mathcal{L}g](z)\right) \, \mathrm{d}z = 0 ; \tag{4.2.11}$$

that is, f coincides with g almost everywhere in $[0, +\infty[$.

3. Identity (4.2.10) suggests the following

Definition 4.2.1 *Let $\sigma \in \mathbb{R}$, and $g : \mathbb{C} \to \mathbb{C}$ be analytic on the half-plane $\Re(z) > \sigma$. The inverse Laplace transform of g is the function $\tilde{g} : \mathbb{R}_{\geq 0} \to \mathbb{C}$ defined by*

$$\tilde{g}(x) := \frac{1}{2\pi i} \int_{\gamma_\sigma} \mathrm{e}^{zx} \, g(z) \, \mathrm{d}z , \tag{4.2.12}$$

where γ_σ is the vertical line defined in (4.2.8). We set $\tilde{g} =: \mathcal{L}^{-1} g$. [1]

Example 4.2.1 *Let $a \in \mathbb{C}$, $n \in \mathbb{N}$, and*

$$g(z) = \frac{1}{(z-a)^{n+1}} . \tag{4.2.13}$$

Then,

$$[\mathcal{L}^{-1} g](x) = \frac{x^n}{n!} \, \mathrm{e}^{ax} , \qquad x \geq 0 . \tag{4.2.14}$$

[1] Note the analogy between definitions (4.1.4) and (4.2.12) with definitions (3.1.1) and (3.3.2) of the Fourier transform and its inverse.

Proof. 1) By (4.1.20) of proposition 4.1.1, and (4.1.18) of example 4.1.3,

$$\left[\mathcal{L}\left(\mathrm{e}^{a(\cdot)}\,(\cdot)^n\right)\right](z) = [\mathcal{L}(\cdot)^n](z-a) = \frac{n\,!}{(z-a)^{n+1}} = (n\,!)\,g(z)\,. \tag{4.2.15}$$

Thus, we formally have that

$$[\mathcal{L}^{-1}g](x) = \frac{1}{n\,!}\left[\mathcal{L}^{-1}\left(\mathcal{L}(\mathrm{e}^{a(\cdot)}\,(\cdot)^n)\right)\right](x) = \frac{1}{n\,!}\,\mathrm{e}^{ax}\,x^n\,, \tag{4.2.16}$$

which is (4.2.14).

2) A rigorous justification of (4.2.14) can be given with the help of Cauchy's integral formula

$$\int_\gamma \frac{f(z)}{(z-z_0)^{n+1}}\,\mathrm{d}z = \frac{2\pi i}{n\,!}\,f^{(n)}(z_0)\,, \tag{4.2.17}$$

in which f is analytic in a domain $\Omega \subseteq \mathbb{C}$, $\gamma \subset \Omega$ is a closed, piecewise smooth contour, described counterclockwise, and z_0 is interior to the region bounded by γ. To see this, for σ and $\tau > 0$ we consider the perimeter $\gamma_{\sigma,\tau}$ of the square of center $a = \alpha + i\,\beta$ and vertices $(\alpha-\sigma, \beta-\tau)$, $(\alpha+\sigma, \beta-\tau)$, $(\alpha+\sigma, \beta+\tau)$, and $(\alpha-\sigma, \beta+\tau)$. Then, by (4.2.17) with $f(z) = \mathrm{e}^{(z-a)x}$ and $z_0 = a$,

$$\int_{\gamma_{\sigma,\tau}} \frac{\mathrm{e}^{(z-a)x}}{(z-a)^{n+1}}\,\mathrm{d}z = \frac{2\pi i}{n\,!}\left[\frac{\mathrm{d}^n}{\mathrm{d}z^n}\,\mathrm{e}^{(z-a)x}\right]_{z=a} = \frac{2\pi i}{n\,!}\,x^n\,. \tag{4.2.18}$$

Letting $\sigma \to 0$ and $\tau \to +\infty$, we deduce from (4.2.18) that

$$\frac{1}{2\pi i}\int_{\gamma_\alpha} \frac{\mathrm{e}^{(z-a)x}}{(z-a)^{n+1}}\,\mathrm{d}z = \frac{x^n}{n\,!}\,, \tag{4.2.19}$$

from which, by (4.2.12) with $\sigma = \alpha$,

$$[\mathcal{L}^{-1}g](x) = \frac{1}{2\pi i}\int_{\gamma_\alpha} \frac{\mathrm{e}^{zx}}{(z-a)^{n+1}}\,\mathrm{d}z = \mathrm{e}^{ax}\,\frac{x^n}{n\,!}\,, \tag{4.2.20}$$

which is (4.2.14). □

4.3 Applications to ODEs.

1. One of the main reasons why the Laplace transform plays such an important role in the study of ODEs is that, as seen in part (3) of proposition 4.1.1, it transforms differentiation into multiplication, up to a polynomial. More precisely, consider a polynomial

$$P(\lambda) = \lambda^n + a_1\,\lambda^{n-1} + \cdots + a_{n-1}\,\lambda + a_n \tag{4.3.1}$$

with coefficients $a_i \in \mathbb{C}$, of real variable λ, and the associated ordinary differential operator of order n

$$P(D) := D^n + a_1\,D^{n-1} + \cdots + a_{n-1}\,D + a_n\,, \tag{4.3.2}$$

where $D := \frac{\mathrm{d}}{\mathrm{d}t}$. Let $y \in C^n([0,+\infty[\to \mathbb{C})$. Then, by the linearity of the Laplace transform, and a repeated application of (4.1.22),

$$\begin{aligned}
&\mathcal{L}(P(D)\,y)(z)\\
&\quad= [\mathcal{L}(y^{(n)})](z) + a_1\,[\mathcal{L}(y^{(n-1)})](z) + \cdots + a_{n-1}\,[\mathcal{L}(y')](z) + a_n\,[\mathcal{L}y](z)\\
&\quad= z^n\,[\mathcal{L}y](z) - y(0)\,z^{n-1} - y'(0)\,z^{n-2} - \cdots - y^{(n-1)}(0)\\
&\qquad + a_1\left(z^{n-1}\,[\mathcal{L}y](z) - y(0)\,z^{n-2} - y'(0)\,z^{n-3} - \cdots - y^{(n-2)}(0)\right)\\
&\qquad + \cdots\cdots\\
&\qquad + a_{n-1}\,(z\,[\mathcal{L}y](z) - y(0)) + a_n\,[\mathcal{L}y](z)\\
&= P(z)\,[\mathcal{L}y](z) - \sum_{k=1}^{n} p_k\,z^{n-k}\,,
\end{aligned} \tag{4.3.3}$$

where the coefficients p_k are defined by

$$p_k := \sum_{i=1}^{k} y^{(k-i)}(0)\,a_{i-1}\,, \qquad \text{with} \quad a_0 := 1\,. \tag{4.3.4}$$

For example, when $n = 2$, let

$$P(\lambda) = \lambda^2 + a_1\,\lambda + a_2\,. \tag{4.3.5}$$

Then,

$$P(D)\,y = y'' + a_1\,y' + a_2\,y\,, \tag{4.3.6}$$

so that

$$\begin{aligned}
[\mathcal{L}(y'' + a_1\,y' + a_2\,y)](z) &= \left(z^2\,[\mathcal{L}y](z) - y(0)\,z - y'(0)\right)\\
&\quad + a_1\,(z\,[\mathcal{L}y](z) - y(0)) + a_2\,[\mathcal{L}y](z)\\
&= (z^2 + a_1\,z + a_2)\,[\mathcal{L}y](z) - (y(0)\,z) - (y'(0) + a_1\,y(0))\,.
\end{aligned} \tag{4.3.7}$$

This is (4.3.3) for $n = 2$, with

$$p_1 = y(0)\,a_0 = y(0)\,, \qquad p_2 = y'(0)\,a_0 + y(0)\,a_1 = y'(0) + y(0)\,a_1\,. \tag{4.3.8}$$

2. We apply the previous observation to formally solve the initial value problem

$$\begin{cases} y^{(n)} + a_1\,y^{(n-1)} + \cdots + a_{n-1}\,y' + a_n\,y = f\,, \\ y(0) = y_0\,,\ y'(0) = y_1\,,\ \cdots,\ y^{(n-1)}(0) = y_{n-1}\,, \end{cases} \tag{4.3.9}$$

where f is a given $\mathcal{L}$-transformable function, and the n numbers $y_0\,, \ldots,\, y_{n-1}$ are the given initial values for the unknown y. Proceeding formally, we take the Laplace transform of all terms of (4.3.9); by (4.3.3), the initial-value problem is then equivalent to the *algebraic* equation

$$P(z)\,u = [\mathcal{L}f](z) - \sum_{k=1}^{n} q_k\,z^{n-k} \tag{4.3.10}$$

in the unknown $u := [\mathcal{L}y](z)$, in which the complex variable z plays the role of a parameter, and, in analogy with (4.3.4),

$$q_k := \sum_{i=1}^{k} y_{k-i}\,a_{i-1}\,, \qquad a_0 = 1\,. \tag{4.3.11}$$

Equation (4.3.10) has the solution

$$u = \frac{1}{P(z)}\left([\mathcal{L}f](z) + \sum_{k=1}^{n} q_k\,z^{n-k}\right) =: u(z)\,, \tag{4.3.12}$$

from which one then hopes to obtain the solution of (4.3.9), by determining the inverse Laplace transform

$$y(t) = [\mathcal{L}^{-1}u](t) = \left[\mathcal{L}^{-1}\left(\tfrac{Q}{P}\right)\right](t) + \left[\mathcal{L}^{-1}\left(\tfrac{\mathcal{L}f}{P}\right)\right](t)\,, \tag{4.3.13}$$

where

$$Q(z) := \sum_{k=1}^{n} q_k\,z^{n-k}\,. \tag{4.3.14}$$

To proceed from (4.3.13), it is useful to recall the following result on the Laplace convolution of two functions f and g, defined for $x > 0$ by

$$[f * g](x) := \int_0^x f(x-y)\,g(y)\,\mathrm{d}y\,. \tag{4.3.15}$$

Proposition 4.3.1 *1) Let f and g be $\mathcal{L}$-transformable. Then, on the half-plane $\Re(z) > \max\{\sigma_f, \sigma_g\}$,*

$$\mathcal{L}(f * g) = [\mathcal{L}f]\,[\mathcal{L}g]\,. \tag{4.3.16}$$

2) Conversely, if F and G are Laplace transforms,

$$\mathcal{L}^{-1}(F\,G) = [\mathcal{L}^{-1}F] * [\mathcal{L}^{-1}G]\,. \tag{4.3.17}$$

The proof of proposition 4.3.1 is similar to that of the analogous proposition 3.4.1 on the Fourier transform. In particular, the last term of the formal solution (4.3.13) can be written as

$$\mathcal{L}^{-1}\left(\tfrac{\mathcal{L}f}{P}\right) = f * \mathcal{L}^{-1}\left(\tfrac{1}{P}\right). \tag{4.3.18}$$

3. The process described above is particularly easy to implement when $n = 2$ (even though, admittedly, the standard methods of solutions yield the solutions more readily). We illustrate this by means of a few examples, all taken from Seiler, [8].

Example 4.3.1 *Consider the homogeneous initial-value problem*

$$\begin{cases} y'' - 2\,y' - 8\,y = 0\,, \\ y(0) = 1\,, \quad y'(0) = 2\,. \end{cases} \tag{4.3.19}$$

From (4.3.14) and (4.3.11), we first determine

$$Q(z) = q_1\,z + q_2 = y_0\,1\,z + (y_1\,1 + y_0\,a_1) = z + (2 + (-2)) = z\,; \tag{4.3.20}$$

thus, by partial fractions decomposition,

$$\frac{Q(z)}{P(z)} = \frac{z}{z^2 - 2z - 8} = \frac{1}{3(z+2)} + \frac{2}{3(z-4)}\,. \tag{4.3.21}$$

Then, we refer to the solution formula (4.3.13): recalling (4.2.13) of example 4.2.1, with $n = 1$ and $a = -2$, $a = 4$, we obtain

$$\begin{aligned} y(t) &= \left[\mathcal{L}^{-1}\left(\tfrac{Q}{P}\right)\right](t) \\ &= \tfrac{1}{3}\left[\mathcal{L}^{-1}\left(\tfrac{1}{(\cdot)+2}\right)\right](t) + \tfrac{2}{3}\left[\mathcal{L}^{-1}\left(\tfrac{1}{(\cdot)-4}\right)\right](t) \\ &= \tfrac{1}{3}\,\mathrm{e}^{-2t} + \tfrac{2}{3}\,\mathrm{e}^{4t}\,. \end{aligned} \tag{4.3.22}$$

The reader is encouraged to verify that this solution is correct, and coincides with the one obtained with the characteristic equation method. □

Example 4.3.2 *Consider the non-homogeneous initial-value problem*

$$\begin{cases} y'' - 2\,y' + 2\,y = f(t)\,, \\ y(0) = 2\,, \quad y'(0) = 3\,. \end{cases} \tag{4.3.23}$$

In this case,

$$\begin{aligned} P(\lambda) &= \lambda^2 - 2\,\lambda + 2 = (\lambda - (1+i))\,(\lambda - (1-i))\,, \\ Q(z) &= q_1\,z + q_2 = 2\,z + (3 + 2(-2)) = 2\,z - 1\,. \end{aligned} \tag{4.3.24}$$

Again by partial fractions decomposition,

$$\frac{Q(Z)}{P(z)} = \left(1 - i\,\tfrac{1}{2}\right)\frac{1}{z-(1+i)} + \left(1 + i\,\tfrac{1}{2}\right)\frac{1}{z-(1-i)}\,, \tag{4.3.25}$$

$$\frac{1}{P(z)} = \frac{1}{2\,i}\left(\frac{1}{z-(1+i)} + \frac{1}{z-(1-i)}\right); \tag{4.3.26}$$

consequently,

$$\begin{aligned} \left[\mathcal{L}^{-1}\left(\tfrac{Q}{P}\right)\right](t) &= \left(1 - i\,\tfrac{1}{2}\right)\left[\mathcal{L}^{-1}\left(\tfrac{1}{z-(1+i)}\right)\right](t) \\ &\quad + \left(1 + i\,\tfrac{1}{2}\right)\left[\mathcal{L}^{-1}\left(\tfrac{1}{z-(1-i)}\right)\right](t) \\ &= \left(1 - i\,\tfrac{1}{2}\right)\mathrm{e}^{(1+i)\,t} + \left(1 + i\,\tfrac{1}{2}\right)\mathrm{e}^{(1-i)\,t} \\ &= \mathrm{e}^{t}\left((\mathrm{e}^{it} + \mathrm{e}^{-\,it}) + \tfrac{i}{2}\,(\mathrm{e}^{-\,it} - \mathrm{e}^{it})\right) \\ &= \mathrm{e}^{t}\,(2\,\cos t + \sin t)\,. \end{aligned} \tag{4.3.27}$$

This is the solution to the homogeneous problem, corresponding to (4.3.23) when $f = 0$. Next, from (4.3.26),

$$\left[\mathcal{L}^{-1}\left(\tfrac{1}{P}\right)\right](t) = \frac{1}{2\,i}\left(\mathrm{e}^{(1+i)\,t} - \mathrm{e}^{(1-i)\,t}\right) = \mathrm{e}^{t}\,\sin t\,; \tag{4.3.28}$$

thus, by (4.3.18),

$$\begin{aligned} \left[f * \mathcal{L}^{-1}\left(\tfrac{1}{P}\right)\right](t) &= \int_0^t f(s)\left[\mathcal{L}^{-1}\left(\tfrac{1}{P}\right)\right](t-s)\,\mathrm{d}s \\ &= \int_0^t f(s)\,\mathrm{e}^{t-s}\,\sin(t-s)\,\mathrm{d}s\,. \end{aligned} \tag{4.3.29}$$

In conclusion, the solution to the initial-value problem (4.3.23) is

$$y(t) = \mathrm{e}^t \left(\sin t + 2\,\cos t + \int_0^t b(s)\,\mathrm{e}^{-s}\,\sin(t-s)\,\mathrm{d}s \right) . \tag{4.3.30}$$

Again, the reader is encouraged to verify that this solution is correct, and coincides with the one obtained with the characteristic equation method. □

Example 4.3.3 *Consider the non-homogeneous initial-value problem*

$$\begin{cases} y'' + 2\,y' + 5\,y = \mathrm{e}^{-t}\,\sin t\,, \\ y(0) = 0\,, \quad y'(0) = 1\,. \end{cases} \tag{4.3.31}$$

In this case, $Q(z) = 1$ and, recalling (4.1.20) and the second of (4.1.14),

$$[\mathcal{L}(\mathrm{e}^{-(\cdot)}\,\sin)](z) = [\mathcal{L}(\sin)](z+1) = \tfrac{1}{(z+1)^2+1}\,; \tag{4.3.32}$$

thus, from

$$(z^2 + 2\,z + 5)\,[\mathcal{L}y](z) - 1 = \tfrac{1}{(z+1)^2+1}\,, \tag{4.3.33}$$

obtained by taking the Laplace transform of all terms of (4.3.31), we find that

$$\begin{aligned} [\mathcal{L}y](z) &= \frac{z^2 + 2\,z + 3}{(z^2 + 2\,z + 2)(z^2 + 2\,z + 5)} \\ &= \frac{1}{(s+1)^2 + 1} + \frac{2}{3((s+1)^2 + 4}\,. \end{aligned} \tag{4.3.34}$$

Consequently,

$$\begin{aligned} y(t) &= \frac{1}{3}\left[\mathcal{L}^{-1}\left(\frac{1}{(1+(\cdot))^2+1}\right)\right](t) + \frac{1}{3}\left[\mathcal{L}^{-1}\left(\frac{2}{(1+(\cdot))^2+4}\right)\right](t) \\ &= \frac{1}{3}\,\mathrm{e}^{-t}\,(\sin t + \sin(2\,t))\,. \end{aligned} \tag{4.3.35}$$

Once more, the reader is encouraged to verify that this solution is correct, and coincides with the one obtained with the characteristic equation method. □

This concludes our brief presentation of the Laplace transform techniques for solving initial-boundary value problems for ODEs. We mention that the same techniques can be used, in a straightforward way, to solve initial-value problems for systems of ODEs (see, e.g., [8, ex. 3.12]).

Bibliography

[1] R. A. Adams and J. J.F. Fournier, *Sobolev Spaces.* Academic Press, New York, 2003.

[2] J. W. Brown and R. V. Churchill, *Complex Variables and Applications, 6th ed..* McGraw - Hill, NY, 1996.

[3] L. C. Evans, *Partial Differential Equations.* American Mathematical Society (G.S.M. vol. 19), Providence, 1998.

[4] J. L . Lions and E. Magenes, *Non-Homogeneous Boundary value Problems and Applications, Vol. I.* Springer Verlag, Berlin, 1972.

[5] M. Renardy and R. C. Rogers, *An Introduction to Partial Differential Equations.* Springer-Verlag, New York, 1993.

[6] W. Rudin, *Real and Complex Analysis.* McGraw-Hill, New York, 1974.

[7] W. Rudin, *Functional Analysis.* McGraw-Hill, New York, 1973.

[8] J. Seiler, *Distribution Theory, Fourier and Laplace Transform.* Notes for the course "Analysis A", MS in Stochastics and Data Science, Univ. Torino, 2019.

[9] A. Torchinsky, *The Fourier transform and the wave equation.* Am. Math. Monthly, 118/7 (2011), 599-609.