Helge Holden

Tools from the Toolbox

Functional Analysis for Partial Differential Equations

January 9, 2017

Helge Holden
Department of Mathematical Sciences
Norwegian University of Science and Technology
NO–7491 Trondheim
Norway
E-mail address: holden@math.ntnu.no
URL: www.math.ntnu.no/~holden/

CHAPTER 3

Distributions

It is déjà vu all over again.
— Yogi Berra[1]

We are used to the traditional definition of a function f as a *rule* that to any element x in the domain X assigns a unique element, denoted $f(x)$, in the range Y, that is, $f(x) \in Y$. We write $f\colon X \to Y$ and $x \mapsto f(x)$. In mathematical analysis we frequently consider pointwise properties like continuity and differentiability (in cases where the sets X and Y have additional structure). Let us in following restrict ourselves to the case $X = Y = \mathbb{R}$. Often the requirement of a unique pointwise value is inconvenient. A familiar example in analysis is the collection of Lebesgue functions, that is, the collection of functions $f \in L^p(\mathbb{R})$ (discussed in detail in Chapter 4). To make $L^p(\mathbb{R})$ into a Banach space we have to give up the notion of pointwise values of its elements. Each element in $L^p(\mathbb{R})$ is rather an equivalence class of functions that are equal *almost everywhere* (in the measure theoretic sense). Even if we have to give up the notion of pointwise values of elements in $L^p(\mathbb{R})$, we can still add elements and multiply with scalars in a welldefined manner. We can also multiply elements; while this may violate the integrability condition and take the product outside $L^p(\mathbb{R})$, the product is associated with a unique equivalence class.

Consider next a function $f \in L^1_{\text{loc}}(\mathbb{R})$; we can associate this function with a locally finite and absolutely continuous measure $\mathrm{d}\mu_f$ by defining $\mathrm{d}\mu_f = f\,\mathrm{d}x$, where $\mathrm{d}x$ is the standard Lebesgue measure. But instead of considering measures induced in this way, we can consider the family of Radon measures $\mathscr{M}(\mathbb{R})$ on $\mathbb{R}$. Now we lose even more the control of pointwise values. Rather we can compare two measures by considering how they act when tested against selected function, called *test functions*, as follows:

$$\langle g, \mu \rangle = \int_{\mathbb{R}} g\,\mathrm{d}\mu, \quad g \in C_0(\mathbb{R}), \tag{3.1}$$

where $C_0(\mathbb{R})$ denotes the set of compactly supported continuous functions on $\mathbb{R}$. Indeed, we can still add two measures in $\mathscr{M}(\mathbb{R})$, and multiply with nonnegative constants. However, we cannot in general multiply two elements in $\mathscr{M}(\mathbb{R})$. Yet the definition (3.1) separates measures in the sense that if $\langle g, \mu \rangle = \langle g, \nu \rangle$ for all $g \in C_0(\mathbb{R})$, then $\mu = \nu$ as measures. Observe that (3.1) acts as linear functional in g for each fixed measure μ.

[1]American baseballplayer.

Now we can take this idea even one step further, by considering linear functionals $g \mapsto \langle g, \mu\rangle$ for "generalized functions" μ. To allow for the usual operations of mathematical analysis like continuity and differentiability, we will make the set of test functions smaller, and thus making the set of generalized functions, or distributions, bigger. The standard choice of test functions is the set $C_0^\infty(\mathbb{R})$. We say that a distribution is a continuous linear functional on $C_0^\infty(\mathbb{R})$. What have we gained by this? By giving up the notion of pointwise values while apparently keeping only the vector space structure of addition and multiplication by scalars, we have introduced a very rich toolbox that somewhat unexpectedly restores the familiar operations of continuity and differentiability without pointwise values. The argument goes as follows. We easily see that functions f in $L^1_{\mathrm{loc}}(\mathbb{R})$ can be considered distributions by the definition $g \mapsto \langle g, f\rangle = \int_{\mathbb{R}} f g \, \mathrm{d}\nu$. Assume for a moment that $f \in C^1(\mathbb{R}) \subset L^1_{\mathrm{loc}}(\mathbb{R})$. Then we have by integration by parts that

$$\langle g, f'\rangle = \int_{\mathbb{R}} f' g \, \mathrm{d}\nu = -\int_{\mathbb{R}} f g' \, \mathrm{d}\nu, \quad g \in C_0^\infty(\mathbb{R}). \tag{3.2}$$

We can now *define* the linear functional $g \mapsto -\langle g', f\rangle$, which does make sense even for nondifferentiable functions f. Furthermore, it turns out that this functional is continuous, and thus a distribution, and hence we take this as the *definition* of the distributional derivative of f. By studying problems, e.g., differential equations, looking for solutions in the larger space of distributions, one may hope that the existence of solutions is easier. Next, by showing more regularity of attained solution, one may hope to restore even pointwise values of the solution.

This is the idea behind distributions; next follows precise definitions.

3.1. Definitions and basic properties

We begin by making the formal definition of a distribution.

DEFINITION 3.1. A distribution T is a linear and continuous functional on $C_0^\infty(\mathbb{R}^d)$. More precisely, we have

$$T \colon C_0^\infty(\mathbb{R}^d) \to \mathbb{C}, \quad T(\alpha_1\phi_1 + \alpha_2\phi_2) = \alpha_1 T(\phi_1) + \alpha_2 T(\phi_2). \tag{3.3}$$

Functions in $C_0^\infty(\mathbb{R}^d)$ are denoted *test functions*. We define continuity as follows. Let $\phi_j \in C_0^\infty(\mathbb{R}^d)$ be such that there exists a compact set $K \subset \mathbb{R}^d$ with

$$\operatorname{supp}\phi_j \subseteq K, \quad j \in \mathbb{N} \tag{3.4}$$

and

$$\|\partial^\alpha \phi_j\|_\infty \to 0 \text{ for all multi-indices } \alpha. \tag{3.5}$$

Then

$$T(\phi_j) \to 0. \tag{3.6}$$

The set of all distribution is denoted $\mathscr{D}'$.

REMARK 3.2. The topology of $C_0^\infty(\mathbb{R}^d)$ is complicated. However, for us it suffices to discuss convergence of sequences as described by the definition.

We often write

$$T(\phi) = \langle T, \phi \rangle.$$

REMARK 3.3. One realizes that there is nothing magical about the space $C_0^\infty(\mathbb{R})$; indeed one can take other vector spaces where one has a notion of what is meant by sequential limits, and define the corresponding distributions as the set of all linear and continuous functionals on that space. Several classes of distributions are introduced in that manner, the most prominent being the set of tempered distributions, starting with the set of test functions in the Schwartz space $\mathscr{S}$, the space consisting of all infinitely differentiable functions such that the function as well as all of its derivatives are rapidly decreasing, that is, faster than any polynomial.
Furthermore, one can replace the set $\mathbb{R}^d$ by some open set $\Omega \subseteq \mathbb{R}^d$, and develop the analogous theory.

There are two prime classes of examples of distributions. We define the *Dirac δ function* at a point a by

$$\delta_a(\phi) = \phi(a), \quad \phi \in C_0^\infty(\mathbb{R}^d). \tag{3.7}$$

Clearly it is a linear functional, and it remains to prove continuity. To that end consider a sequence $\phi_j \in C_0^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\phi_j \subseteq K$ and $\partial^\alpha \phi_j \to 0$ uniformly. Then

$$|\delta_a(\phi_j)| = |\phi_j(a)| \le \|\phi_j\|_\infty \to 0. \tag{3.8}$$

The other key family of examples are the so-called *regular distributions*. Let $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, and define

$$T_f(\phi) = \int f\phi \, \mathrm{d}x, \quad \phi \in C_0^\infty(\mathbb{R}^d). \tag{3.9}$$

Again linearity is clear. Consider a sequence $\phi_j \in C_0^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\phi_j \subseteq K$ and $\partial^\alpha \phi_j \to 0$ uniformly. Then

$$|T_f(\phi_j)| \le \int_K |f||\phi_j| \, \mathrm{d}x \le \|\phi_j\|_\infty \int_K |f| \, \mathrm{d}x \to 0. \tag{3.10}$$

Fortunately, we have the following identification result (see Exercise 3)

$$T_f = 0 \text{ if and only if } f = 0 \text{ almost everywhere.} \tag{3.11}$$

We will often ignore the distinction between T_f and f when $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, and talk about the distribution f.

We are now free to define the various properties we want for the distributions. Our guiding principle will be the following. Assume first that the distribution is regular, and with sufficient smoothness of the function f. Apply the property to f, manipulate the expression formally, trying to transfer the property to the test function ϕ, and use the resulting expression as the definition of the property for a general distribution T. The following example will clarify the approach. Let us try to make a reasonable definition of a translate of a distribution (which in general has no pointwise values). Define the shift operator

$$\tau_a f(x) = f(x-a), \tag{3.12}$$

and the aim is to make the analog definition for a distribution T. Assume first that T is regular, that is, there exists an $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ with $f \in C(\mathbb{R}^d)$ such that $T = T_f$. Let $\phi \in C_0^\infty(\mathbb{R}^d)$, and consider the following formal manipulations

$$\begin{aligned}\tau_a T_f(\phi) &= T_{\tau_a f}(\phi) \text{ (here we transfer the property to } f)\\ &= \int \tau_a f(x)\phi(x)\,\mathrm{d}x\\ &= \int f(x)\phi(x+a)\,\mathrm{d}x \text{ (a change of variables)}\\ &= T_f(\tau_{-a}\phi).\end{aligned}$$

Since $\tau_{-a}\phi \in C_0^\infty$, we can *define* the translate of any distribution by

$$\tau_a T(\phi) = T(\tau_{-a}\phi). \tag{3.13}$$

Similarly one can define *odd* and *even* distributions (see Exercise 1) as well as *periodic* distributions (see Exercise 2). Clearly, a key concept for us is the definition of differentiability of distributions. Again we use the same approach. If $f \in C^{|\alpha|}(\Omega)$, then we define

$$\begin{aligned}\partial^\alpha T_f(\phi) &= T_{\partial^\alpha f}(\phi) \quad \text{(this is our definition)}\\ &= \int_\Omega \partial^\alpha f(x)\phi(x)\,\mathrm{d}x\\ &= (-1)^{|\alpha|}\int_\Omega f(x)\partial^\alpha\phi(x)\,\mathrm{d}x.\end{aligned}$$

In the general case we define the partial derivative of a distribution T as follows

$$\partial^\alpha T(\phi) = (-1)^{|\alpha|}T(\partial^\alpha\phi) \tag{3.14}$$

for any multi-index α. Observe that with this definition any distribution is infinitely differentiable. We also use the standard notation for derivatives, that is, T', $\nabla T = (T_{x_1}, \dots, T_{x_d})$, and $T^{(k)}$, etc. If f is smooth, the distributional derivative coincides with the regular derivative.

We say that a function $g \in L^1_{\text{loc}}(\mathbb{R}^d)$ is the distributional derivative of $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ if the distributional derivative of T_f, that is, $T'_f = (T_f)'$, equals the regular distribution induces by g, that is, T_g. Thus, if $T'_f = T_g$. Without reference to distributions, we can write this condition as

$$\int_{\mathbb{R}^d} g\phi\,\mathrm{d}x = -\int_{\mathbb{R}^d} f\phi'\,\mathrm{d}x, \quad \phi \in C_0^\infty(\mathbb{R}^d). \tag{3.15}$$

More precise statements will follow later.

One subtlety should be addressed explicitly. We have just seen that $(T_f)' = T_{f'}$ if $f \in C^1(\mathbb{R})$. But we may still define the right-hand side assuming only $f' \in L^1_{\text{loc}}(\mathbb{R})$. What is then the relation between $(T_f)' = T'_f$ and $T_{f'}$? To study this, we assume slightly more than $f, f' \in L^1_{\text{loc}}(\mathbb{R})$, namely that f is continuously differentiable except at points $a_1 < \cdots < a_m$ where it enjoys jump discontinuities. Then we find for $\phi \in C_0^\infty(\mathbb{R})$ that

$$\langle T'_f, \phi\rangle = -\int_{\mathbb{R}} f\phi'\,\mathrm{d}x$$

$$
\begin{aligned}
&= -\int_{-\infty}^{a_1} f\phi' \, \mathrm{d}x - \sum_{j=1}^{m-1} \int_{a_j}^{a_{j+1}} f\phi' \, \mathrm{d}x - \int_{a_m}^{\infty} f\phi' \, \mathrm{d}x \\
&= -f(a_1-)\phi(a_1) + \int_{-\infty}^{a_1} f'\phi \, \mathrm{d}x \\
&\quad + \sum_{j=1}^{m-1} \Big(\int_{a_j}^{a_{j+1}} f'\phi \, \mathrm{d}x - f(a_{j+1}-)\phi(a_{j+1}) + f(a_j+)\phi(a_j) \Big) \\
&\quad + f(a_m+)\phi(a_m) + \int_{a_m}^{\infty} f'\phi \, \mathrm{d}x \\
&= \int f'\phi \, \mathrm{d}x + \sum_{j=1}^{m} (f(a_j+) - f(a_j-))\phi(a_j) \\
&= \langle T_{f'}, \phi \rangle + \sum_{j=1}^{m} \big(f(a_j+) - f(a_j-) \big) \delta_{a_j}(\phi)
\end{aligned}
$$

or in other words

$$
T_f' = T_{f'} + \sum_{j=1}^{m} \big(f(a_j+) - f(a_j-) \big) \delta_{a_j}. \tag{3.16}
$$

As an example of this result we note the following. Let H be the *Heaviside function*

$$
H(x) = \chi_{(0,\infty)}(x).
$$

Clearly $H \in L^1_{\text{loc}}(\mathbb{R})$ with $H' = 0$ almost everywhere, which clearly is in $L^1_{\text{loc}}(\mathbb{R})$. Applying (3.16) we find that

$$
T_H' = (H(0+) - H(0-))\delta = \delta,
$$

or

$$
H' = \delta. \tag{3.17}
$$

Another useful result is

$$
|x|' = \operatorname{sgn} x \tag{3.18}
$$

where the left-hand side is the distributional derivative of the absolute value function, while the right-hand side equals the pointwise almost everywhere derivative of the same function. If we continue to take another derivative, we find

$$
\frac{d}{dx} \operatorname{sgn} x = 2\delta. \tag{3.19}
$$

The left-hand side denotes the distributional derivative of $\operatorname{sgn} x$, which equals the Dirac delta function at the origin. However, considered as function, its derivative vanishes almost everywhere.

From the definition of translates we observe that the function

$$
\psi(x) = \tau_x T(\phi) \tag{3.20}
$$

is infinitely differentiable, that is, $\psi \in C^\infty(\mathbb{R}^d)$, indeed

$$
\partial^\alpha \tau_x T(\phi) = (-1)^{|\alpha|} \tau_x \partial^\alpha T(\phi). \tag{3.21}
$$

It suffices to study the first derivative in the case $n = 1$. To that end we note that

$$\frac{1}{h}\big(\psi(x+h)-\psi(x)\big) = T\big(\frac{1}{h}(\phi(\cdot+x+h)-\phi(\cdot+x))\big),$$

and $\frac{1}{h}(\phi(\cdot+x+h)-\phi(\cdot+x)) \to \phi'(\cdot+x)$ in $C_0^\infty(\mathbb{R})$ as $h \to 0$. Thus we conclude that

$$\psi'(x) = \partial_x \tau_x T(\phi) = T(\partial_x \tau_{-x}\phi) = T(\tau_{-x}\partial\phi).$$

This generalizes to

$$\partial^\alpha \tau_x T(\phi) = T(\tau_{-x}(\partial^\alpha \phi)) = T(\partial^\alpha \tau_{-x}\phi) = (-1)^{|\alpha|}(\partial^\alpha T)(\tau_{-x}\phi) = (-1)^{|\alpha|}\tau_x \partial^\alpha T(\phi) \tag{3.22}$$

for any multi-index α. Furthermore, we have that

$$\int \psi(x) T(\tau_x \phi)\,\mathrm{d}x = T(\psi * \phi), \quad \phi,\psi \in C_0^\infty(\mathbb{R}^d), \tag{3.23}$$

due to the following argument: Approximate $\psi*\phi(y)$ by Riemann sums, that is, define $\omega_{\Delta x}(y) = \Delta x \sum_j \psi(x_j)\phi(y-x_j)$. Then we have that $\omega_{\Delta x} \to \psi*\phi$ in C_0^∞. Furthermore, we see that

$$T(\omega_{\Delta x}) = \Delta x \sum_j T(\psi(x_j)\phi(\cdot - x_j)) = \Delta x \sum_j \psi(x_j) T(\phi(\cdot - x_j)). \tag{3.24}$$

Continuity of T shows that the left-hand side of (3.24) converges to $T(\psi*\phi)$, while the right-hand side approaches $\int \psi(x) T(\tau_x\phi)\,\mathrm{d}x$. Thus (3.23) holds. By using the commutativity of the convolution product, we may write this as

$$\langle \psi, T(\tau_{(\cdot)}\phi)\rangle = \langle T(\tau_{(\cdot)}\psi), \phi\rangle. \tag{3.25}$$

We have the following fundamental theorem for distributions. It extends the well-known result

$$\psi(x) - \psi(0) = \int_0^1 x\cdot\nabla\psi(tx)\,\mathrm{d}t, \quad \psi \in C^1(\mathbb{R}^d). \tag{3.26}$$

PROPOSITION 3.4. *Let $T \in \mathscr{D}'$, $\phi \in C_0^\infty(\mathbb{R}^d)$, and $x \in \mathbb{R}^d$. Then*

$$T(\tau_x\phi) - T(\phi) = \int_0^1 x\cdot\nabla T(\tau_{tx}\phi)\,\mathrm{d}t. \tag{3.27}$$

PROOF. Let $R(x)$ denote the right-hand side of (3.27). Then

$$\begin{aligned}
\partial_{x_j} R(x) &= \int_0^1 \partial_j T(\tau_{tx}\phi)\,\mathrm{d}t + \int_0^1 x\cdot\partial_{x_j}\nabla T(\tau_{tx}\phi)\,\mathrm{d}t \\
&= \int_0^1 \partial_j T(\tau_{tx}\phi)\,\mathrm{d}t - \int_0^1 tx\cdot\nabla T(\tau_{tx}\partial_j\phi)\,\mathrm{d}t \\
&= \int_0^1 \partial_j T(\tau_{tx}\phi)\,\mathrm{d}t - \int_0^1 t\frac{d}{dt} T(\tau_{tx}\partial_j\phi)\,\mathrm{d}t \\
&= \int_0^1 \partial_j T(\tau_{tx}\phi)\,\mathrm{d}t - T(\tau_x\partial_j\phi) + \int_0^1 T(\tau_{tx}\partial_j\phi)\,\mathrm{d}t \\
&= \partial_{x_j} T(\tau_x\phi).
\end{aligned}$$

Thus the left-hand and the right-hand side of (3.27) share the same partial derivatives, and since both sides vanish at $x = 0$, we have proved (3.27). This proof is based on [96, Thm. 6.9], where one can find a proof in the general case where $\mathbb{R}^d$ is replaced by an open set $\Omega \subseteq \mathbb{R}^d$. □

A more refined result is the following.

PROPOSITION 3.5. *Let $T \in \mathscr{D}'$. The following two statements are equivalent:*
(1) T is regular, that is, $T = T_f$, and $f \in C^1(\mathbb{R}^d)$.
(2) $\partial_{x_j} T$ is regular, that is, $\partial_{x_j} T = T_{g_j}$, and $g_j \in C(\mathbb{R}^d)$ for $j = 1, \dots, d$.
In each case $g_j = \partial_{x_j} f$ for $j = 1, \dots, d$.

PROOF. That (1) implies (2) follows directly from the definition of derivatives, and integration by parts, given the regularity of f. Let us now assume (2). Using Proposition 3.4 we can write, introducing $g = (g_1, \dots, g_d)$,

$$T(\tau_x\phi) - T(\phi) = \int_0^1 \int x \cdot g(y)\tau_{tx}\phi(y)\,\mathrm{d}y\,\mathrm{d}t = \int \int_0^1 x \cdot g(y + tx)\phi(y)\,\mathrm{d}t\,\mathrm{d}y. \tag{3.28}$$

Let $\psi \in C_0^\infty$ be nonnegative with $\int \psi\,\mathrm{d}x = 1$. Then

$$\int \psi(x)T(\tau_x\phi)\,\mathrm{d}x - T(\phi) = \int \psi(x)\Big(\int \int_0^1 x \cdot g(y + tx)\phi(y)\,\mathrm{d}t\,\mathrm{d}y\Big)\,\mathrm{d}x. \tag{3.29}$$

Observe that from (3.25) we infer

$$\int \psi(x)T(\tau_x\phi)\,\mathrm{d}x = \int \phi(y)T(\tau_y\psi)\,\mathrm{d}y.$$

If we advocate this in (3.29), we find that

$$T(\phi) = \int \Big(T(\tau_y\psi) - \int_0^1 \int \psi(x)x \cdot g(y + tx)\,\mathrm{d}x\,\mathrm{d}t\Big)\phi(y)\,\mathrm{d}y = \int f(y)\phi(y)\,\mathrm{d}y = T_f(\phi),$$

where we have defined $f(y) = T(\tau_y\psi) - \int_0^1 \int \psi(x)x \cdot g(y + tx)\,\mathrm{d}x\,\mathrm{d}t$. Thus we can write (3.28) as

$$\begin{aligned}
T(\tau_x\phi) - T(\phi) &= \int f(y)\tau_x\phi(y)\,\mathrm{d}x - \int f(y)\phi(y)\,\mathrm{d}y \\
&= \int \phi(y)\big(f(y + x) - f(y)\big)\,\mathrm{d}y \\
&= \int \int_0^1 x \cdot g(y + tx)\phi(y)\,\mathrm{d}t\,\mathrm{d}y,
\end{aligned}$$

which implies that

$$\int \phi(y)\Big(f(y + x) - f(y) - \int_0^1 x \cdot g(y + tx)\,\mathrm{d}t\Big)\,\mathrm{d}y = 0,$$

from which we conclude that

$$f(y + x) - f(y) = \int_0^1 x \cdot g(y + tx)\,\mathrm{d}t = x \cdot g(y) + o(|x|),$$

and hence

$$\nabla f(y) = g(y).$$

Our proof is based on [96, Thm. 6.10], where the extension to general domains can be found. □

We may use the result (3.25) to define the convolution of a distribution with a smooth function. To find the proper definition we first assume that T is regular, i.e., $T = T_f$ with sufficiently smooth f. Then

$$\begin{aligned}(\psi * T_f)(\phi) &= T_{\psi * f}(\phi) \text{ (this is our definition)}\\ &= \int\int \psi(y) f(x-y)\,\mathrm{d}y\,\phi(x)\,\mathrm{d}x\\ &= \int\int \psi(x-y) f(y)\,\mathrm{d}y\,\phi(x)\,\mathrm{d}x\\ &= \int \langle \tau_{-x} T_f, \psi_\sigma\rangle \phi(x)\,\mathrm{d}x\end{aligned}$$

where

$$\psi_\sigma(x) = \psi(-x), \quad x \in \mathbb{R}^d.$$

Thus we define the convolution of a distribution $T \in \mathscr{D}'$ with a smooth function $\psi \in C_0^\infty$, that is, $\psi * T$, as the *function*

$$(\psi * T)(x) = \langle \tau_{-x} T, \psi_\sigma\rangle = T(\tau_x \psi_\sigma), \quad \psi \in C_0^\infty(\mathbb{R}^d). \tag{3.30}$$

Observe that $(\psi * T) \in C^\infty(\mathbb{R})$.

Furthermore, one can define the support of a distribution. First one says that T vanishes on an open set Ω, writing $T|_\Omega = 0$ if

$$T(\phi) = 0, \quad \phi \in C_0^\infty, \operatorname{supp}\phi \subseteq \Omega.$$

We then define

$$\operatorname{supp} T = \mathbb{R}^d \setminus \cup\{\Omega \mid T|_\Omega = 0, \Omega \text{ open}\}. \tag{3.31}$$

A consequence of this is that for regular distributions we have (see, e.g., [59, Prop. 28.2.3])

$$\operatorname{supp} T_f = \operatorname{ess\,supp} f, \quad f \in L^1_{\text{loc}}. \tag{3.32}$$

Given a sequence $T_n \in \mathscr{D}'$, we say that $T_n \to T$ in $\mathscr{D}'$ if

$$T_n(\phi) \to T(\phi), \quad \phi \in C_0^\infty(\mathbb{R}^d).$$

Immediately from the definition we infer that if $T_n \to T$, then all derivatives converge, that is, $\partial^\alpha T_n \to \partial^\alpha T$. However, there are pitfalls also in this context. Consider the situation where $f_n \in L^1_{\text{loc}}$, and $T_{f_n} \to T_f$. We do *not* have in general that $f_n \to f$, e.g., pointwise.

Consider namely the family $f_n(x) = \sin(2\pi n x)$. Certainly f_n has no pointwise limit. However, the Riemann–Lebesgue lemma implies that

$$\langle T_{f_n}, \phi\rangle = \int f_n \phi\,\mathrm{d}x \to 0.$$

Thus $T_{f_n} \to T_0$ while $f_n \not\to 0$. However, we do have the following simple result.

PROPOSITION 3.6. *(i) Let $f_n \in L^1(\mathbb{R})$ be such that*

$$\|f_n - f\|_1 \to 0$$

for some $f \in L^1(\mathbb{R})$. *Then* $T_{f_n} \to T_f$.
(ii) Let $f_n, f \in L^1_{\mathrm{loc}}(\mathbb{R})$ *be such that*

$$\|f_n - f\|_{L^1(B)} \to 0$$

for all bounded sets B. *Then* $T_{f_n} \to T_f$.

PROOF. (i) The proof is simple:

$$|T_{f_n}(\phi) - T_f(\phi)| \le \int |f_n(x) - f(x)||\phi(x)|\,\mathrm{d}x \le \|f_n - f\|_1 \|\phi\|_\infty.$$

(ii) This proof is also simple:

$$|T_{f_n}(\phi) - T_f(\phi)| \le \int_B |f_n(x) - f(x)||\phi(x)|\,\mathrm{d}x \le \|(f_n - f)\chi_B\|_{L^1(B)} \|\phi\|_\infty$$

when we assume that $\operatorname{supp}\phi \subset B$. □

One can perform the usual algebraic manipulations with distributions — addition of distributions and multiplication with constants. However multiplication of distributions *per se* is not admissible. The reason is the following. Let g be a function that we want to multiply the distribution T. Following our general recipe, we first assume that T is regular, thus $T = T_f$ for some f. Then

$$\begin{aligned}\langle gT_f, \phi\rangle = \langle T_{gf}, \phi\rangle &= \int g(x) f(x)\phi(x)\,\mathrm{d}x \\ &= \int f(x) g(x)\phi(x)\,\mathrm{d}x \\ &= \langle T_f, g\phi\rangle\end{aligned}$$

which requires that $g\phi$ is a valid test function. This is certainly not true in general and requires that $g \in C^\infty(\mathbb{R})$. We say that distributions can be multiplied by smooth functions, and advocate the definition

$$\langle gT, \phi\rangle = \langle T, g\phi\rangle, \quad g \in C^\infty(\mathbb{R}).$$

With this definition we get the following simple, but useful result

$$D^\alpha(fT) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} f D^\beta T, \quad f \in C^\infty(\mathbb{R}),$$

for any multi-index α.

As for functions, integration theory is more complicated than derivation. A standard result is the following.

THEOREM 3.7. *Let* $T \in \mathscr{D}'(\mathbb{R})$. *Then there exists a* $U \in \mathscr{D}'(\mathbb{R})$ *such that* $T = U'$. *If* U *and* V *share this property, then there exists a constant* c *such that* $U = V + c$.

PROOF. Let us first see what properties U has to possess, granted that it exists. Clearly

$$\langle U, \phi'\rangle = -\langle U', \phi\rangle = -\langle T, \phi\rangle \tag{3.33}$$

for all $\phi \in C_0^\infty(\mathbb{R})$. Hence we know how U acts on $\mathscr{D}_0$ where

$$\mathscr{D}_0 = \{\phi' \mid \phi \in C_0^\infty(\mathbb{R})\} = \{\psi \in C_0^\infty(\mathbb{R}) \mid \int_{\mathbb{R}} \psi \,\mathrm{d}x = 0\}.$$

Let now $\phi \in C_0^\infty(\mathbb{R})$, and introduce $\psi \in \mathscr{D}_0$ by

$$\psi = \phi - \omega \int_{\mathbb{R}} \phi \,\mathrm{d}x,$$

where $\omega \in C_0^\infty(\mathbb{R})$ satisfies $\int_{\mathbb{R}} \omega \,\mathrm{d}x = 1$. Define

$$\Phi_\phi(x) = \int_{-\infty}^x \psi(y)\,\mathrm{d}y = \int_{-\infty}^x \big(\phi(y) - \omega(y)\int_{\mathbb{R}} \phi \,\mathrm{d}x\big)\,\mathrm{d}y.$$

Hence

$$\Phi_{\phi'} = \phi, \quad \Phi_\phi' = \psi$$

and thus $\Phi_\phi \in C_0^\infty(\mathbb{R})$. Now we define $U \colon C_0^\infty(\mathbb{R}) \to \mathbb{C}$ by

$$\langle U, \phi\rangle = -\langle T, \Phi_\phi\rangle. \tag{3.34}$$

We see that

$$\langle U', \phi\rangle = -\langle U, \phi'\rangle = \langle T, \Phi_{\phi'}\rangle = \langle T, \phi\rangle, \tag{3.35}$$

and thus it remains to show that indeed this defines a distribution. Linearity is clear, and as usual it is the continuity that requires attention. To that end let ϕ_n be a sequence that goes to zero in $C_0^\infty(\mathbb{R})$, and we have to show that the same holds for Φ_{ϕ_n}. Since $\int_{\mathbb{R}} \phi_n \,\mathrm{d}x \to 0$, we note that $\psi_n \to 0$ in $C_0^\infty(\mathbb{R})$. Assume that $\operatorname{supp}\psi_n \subseteq [a,b]$. By possibly increasing the interval, we may assume that $\operatorname{supp}\omega \subseteq [a,b]$. For $x > b$ we see that

$$\Phi_{\phi_n}(x) = \int_{-\infty}^x \big(\phi_n(y) - \omega(y)\int_{\mathbb{R}} \phi_n \,\mathrm{d}x\big)\,\mathrm{d}y = \int_{\mathbb{R}} \phi_n \,\mathrm{d}x - \int_{\mathbb{R}} \phi_n \,\mathrm{d}x \int_{\mathbb{R}} \omega \,\mathrm{d}x = 0$$

and thus $\operatorname{supp}\Phi_{\phi_n} \subseteq [a,b]$. Furthermore,

$$\|\Phi_{\phi_n}\|_\infty \le \|\psi_n\|_1 \le (b-a)\|\psi_n\|_\infty$$

which goes to zero. Similarly, we see that derivatives of Φ_{ϕ_n} vanish uniformly.

If U and V both satisfy $U' = V' = T$, we see that $W = U - V$ satisfies $W' = 0$. If we construct $\psi \in \mathscr{D}_0$ as above, we infer that

$$0 = \langle W, \psi\rangle = \langle W, \phi\rangle - \int \phi(x)\,\mathrm{d}x \langle W, \omega\rangle,$$

or

$$\langle W, \phi\rangle = \langle W, \omega\rangle \int \phi(x)\,\mathrm{d}x$$

which indeed shows that W is a constant, namely $\langle W, \omega\rangle$. □

Let us discuss some simple examples of applications of the theory of distributions to (ordinary) differential equations.

EXAMPLE 3.8. (i) The simplest equation of all (see also Theorem 3.7),

$$T' = 0,$$

is to be solved in the sense of distributions. The expected solution is that T has to be equal to a constant. But this follows directly from Theorem 3.7. Thus

$$T(\phi) = K \int \phi \, dx, \quad \phi \in C_0^\infty(\mathbb{R}),$$

for some $K \in \mathbb{C}$. If we consider the multi-dimensional case where

$$\nabla T = 0,$$

we still get, applying Proposition 3.5, that T equals a constant.
(ii) Consider now

$$xT = 0.$$

Not a differential equation, this will still give some insight. Observe first that for any $\phi \in C_0^\infty(\mathbb{R})$ we may write (cf. (3.26))

$$\phi(x) = \phi(0) + \int_0^x \phi'(t)\, dt = \phi(0) + x \int_0^1 \phi'(xt)\, dt = \phi(0) + x\psi(x).$$

Thus $\phi(0) = 0$ if and only if we can write $\phi(x) = x\psi(x)$ for some $\psi \in C_0^\infty(\mathbb{R})$. Let now $\omega \in C_0^\infty(\mathbb{R})$ with $\omega(0) = 1$. Then any $\phi \in C_0^\infty(\mathbb{R})$ can be written

$$\phi(x) = \phi(0)\omega(x) + x\psi(x).$$

Then

$$0 = xT(\psi) = T(\phi - \phi(0)\omega) = T(\phi) - \phi(0)T(\omega),$$

or

$$T(\phi) = \phi(0)T(\omega).$$

But this is nothing but

$$T = K\delta_0$$

where $K = T(\omega)$ is a constant.
(iii) Consider now

$$T' + cT = 0$$

where c is a nonzero constant. Introduce $U = e^{cx}T$. Then U satisfies $U' = 0$ with unique solution $U = K$, where K is a constant. Thus $T = e^{-cx}K$.
(iv) Consider

$$T' + cT = \delta.$$

Again we introduce $U = e^{cx}T$ which satisfies $U' = e^{cx}\delta$. Observe that $e^{cx}\delta(\phi) = e^{c0}\phi(0) = \phi(0)$, and hence $e^{cx}\delta = \delta$. Since the Heaviside function H satisfies $H' = \delta$ (cf. (3.17)), we find that $(U - H)' = 0$ and thus $U = H + K$ for some constant K. Finally, we find

$$T = e^{-cx}H + e^{-cx}K.$$

(v) Consider

$$T'' + T = \delta.$$

Note first that the distribution $V = H(x)\sin(x)$ satisfies

$$V' = \delta(x)\sin(x) + H(x)\cos(x), \quad V'' = \delta'(x)\sin(x) + 2\delta(x)\cos(x) - H(x)\sin(x).$$

When we observe that $\delta'(x)\sin(x) = -\delta(x)$ and $\delta(x)\cos(x) = \delta(x)$ (just apply a test function and see what you get), we conclude that

$$V'' + V = \delta.$$

Hence $U = T - V$ satisfies

$$U'' + U = 0.$$

Write

$$(\frac{d}{dx} + iI)(\frac{d}{dx} - iI) = \frac{d^2}{dx^2} + I,$$

and let $W = (\frac{d}{dx} - iI)U = U' - iU$, which implies $(\frac{d}{dx} + iI)W = W' + iW = 0$. Part (iii) of this example implies that $W = e^{-ix}K$ for some constant K. Now we have to solve

$$U' - iU = e^{-ix}K.$$

A particular solution (classical!) is $iKe^{-ix}/2$, and hence

$$\big(U - \frac{iK}{2}e^{-ix}\big)' - i\big(U - \frac{iK}{2}e^{-ix}\big) = 0.$$

Advocating part (iii) in this example, we see that

$$U = \frac{iK}{2}e^{-ix} + K_1 e^{ix}$$

for some constant K_1. Introducing more conveniently chosen constants we may finally write the solution as

$$T = a\sin(x) + b\cos(x) + H(x)\sin(x).$$

Next we turn to the application of distributions to linear partial differential equations. We restrict our attention to linear partial differential operators with constant coefficients. More precisely, define the linear operator

$$L = \sum_{|\alpha|\le m} c_\alpha D^\alpha \tag{3.36}$$

where we have used the multi-index notation and $c_\alpha \in \mathbb{R}$. If at least one of c_α with $|\alpha| = m$ is nonzero, we say that L has order m. We say that $E \in \mathscr{D}'$ is a fundamental solution if

$$LE = \delta_0. \tag{3.37}$$

We have the following result.

THEOREM 3.9. *Let $f \in C_0^\infty(\mathbb{R}^d)$ and let $E \in \mathscr{D}'$ be a fundamental solution of* (3.37). *Then $u = f * E \in C^\infty(\mathbb{R}^d)$ is a solution of*

$$Lu = f. \tag{3.38}$$

PROOF. Observe first that

$$(f * D^\alpha E)(x) = D^\alpha E(\tau_x f_\sigma) = D^\alpha(f * E)(x),$$

which by linearity implies that

$$f * LE = L(f * E). \tag{3.39}$$

By assumption we have that $LE = \delta_0$ which implies

$$f * LE = f * \delta_0 = f, \tag{3.40}$$

using the definition of the convolution for distributions, viz. (3.30). This proves the statement. □

The fundamental theorem in this context the Malgrange–Ehrenpreis theorem.

THEOREM 3.10 (Malgrange–Ehrenpreis theorem). *Every constant coefficient linear partial differential operator on $\mathbb{R}^d$ has a fundamental solution.*

PROOF. For a proof, see [121, Thm. 8.5]. □

EXAMPLE 3.11. Let L be the Laplacian, viz.,

$$L = \Delta = \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2}. \tag{3.41}$$

Our goal is to solve the Poisson equation

$$Lu = \Delta u = f$$

for $f \in C_0^\infty(\mathbb{R}^d)$.

The fundamental solution equals[2]

$$E(x) = \begin{cases} \frac{1}{2}|x|, & \text{if } d = 1, \\ \frac{1}{2\pi}\ln|x|, & \text{if } d = 2, \\ \frac{1}{d\omega_d(2-d)}|x|^{2-d}, & \text{if } d \geq 3, \end{cases} \tag{3.42}$$

where

$$\omega_d = \frac{2\pi^{d/2}}{d\Gamma(d/2)}.$$

We will now show this result. For $d = 1$ it is fairly straightforward (using (3.18), (3.19))

$$\partial_x^2 \frac{1}{2}|x| = \frac{1}{2}\partial_x \operatorname{sgn}(x) = \delta_0. \tag{3.43}$$

For $d \geq 2$ we will show directly that Theorem 3.9 holds, that is, show that

$$\langle \Delta E, \phi\rangle = \langle E, \Delta\phi\rangle = \phi(0). \tag{3.44}$$

We first prove that E indeed defines a regular distribution, that is, $E \in L^1_{\text{loc}}(\mathbb{R}^d)$. For $d = 2$ we find

$$\begin{aligned} \int_{|x|\leq 1} \left|\frac{1}{2\pi}\ln|x|\right| \mathrm{d}x &= -\frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \ln r \, r \, dr d\theta \\ &= \frac{1}{4} < \infty, \end{aligned}$$

[2] We here identify E and T_E.

where we have introduced polar coordinates $x = r\cos\theta$ and $y = r\sin\theta$. For $d \geq 3$ we find[3]

$$\int_{|x|\leq 1} \frac{1}{d\omega_d(2-d)}|x|^{2-d}\,\mathrm{d}x = -\frac{1}{d\omega_d(2-d)}\int_{|x|=1}\int_0^1 r^{2-d}r^{d-1}dr\,dS_x$$
$$= \frac{1}{2(d-2)} < \infty,$$

by introducing spherical coordinates. Thus the function $E \in L^1_{\text{loc}}(\mathbb{R}^d)$ for all $d \in \mathbb{N}$ and defines a regular distribution.

We will next show

$$\langle E, \Delta\phi\rangle = \phi(0). \tag{3.45}$$

Let $\varepsilon > 0$ and choose R such that $\operatorname{supp}\phi \subset B_R(0)$. By applying the divergence theorem we find that for smooth functions ϕ, g that

$$\int_{\varepsilon<|x|<R} g(x)\Delta\phi(x)\,\mathrm{d}x = -\int_{\varepsilon<|x|<R}\nabla g(x)\cdot\nabla\phi(x)\,\mathrm{d}x + \int_{|x|=\varepsilon} g(x)\nabla\phi(x)\cdot\nu\,dS_x \tag{3.46}$$

where ν is the unit normal, pointing towards the origin, at the surface of the sphere $|x| = \varepsilon$. By first interchanging g and ϕ in the previous result, and then eliminate the term with first derivatives on both functions, we find

$$\begin{aligned}\int_{\varepsilon<|x|<R} g(x)\Delta\phi(x)\,\mathrm{d}x = \int_{\varepsilon<|x|<R}\Delta g(x)\phi(x)\,\mathrm{d}x - \int_{|x|=\varepsilon}\phi(x)\nabla g(x)\cdot\nu\,dS_x\\ + \int_{|x|=\varepsilon} g(x)\nabla\phi(x)\cdot\nu\,dS_x.\end{aligned} \tag{3.47}$$

Let now $g = E$ in (3.47), and conclude that

$$\begin{aligned}\int_{\varepsilon<|x|<R} E(x)\Delta\phi(x)\,\mathrm{d}x = \int_{\varepsilon<|x|<R}\Delta E(x)\phi(x)\,\mathrm{d}x - \int_{|x|=\varepsilon}\phi(x)\nabla E(x)\cdot\nu\,dS_x\\ + \int_{|x|=\varepsilon} E(x)\nabla\phi(x)\cdot\nu\,dS_x\end{aligned} \tag{3.48}$$

holds. But a straightforward computation shows that

$$\Delta E(x) = 0, \quad x \neq 0. \tag{3.49}$$

Furthermore, we find for $d > 2$ that

$$\int_{|x|=\varepsilon} E(x)\nabla\phi(x)\cdot\nu\,dS_x = \frac{1}{d\omega_d(2-d)}\int_{|z|=1}\varepsilon^{2-d}\nabla\phi(z\varepsilon)\cdot\nu\varepsilon^{d-1}dS_z \to 0 \tag{3.50}$$

as $\varepsilon \to 0$. A similar argument proves the case $d = 2$.

Next we consider the term

$$\begin{aligned}-\int_{|x|=\varepsilon}\phi(x)\nabla E(x)\cdot\nu\,dS_x &= \int_{|z|=1}\frac{\partial E}{\partial r}(z\varepsilon)\phi(z\varepsilon)\varepsilon^{d-1}dS_z\\ &= \frac{1}{d\omega_d}\int_{|z|=1}\varepsilon^{1-d}\phi(z\varepsilon)\varepsilon^{d-1}dS_z \to \phi(0)\end{aligned} \tag{3.51}$$

[3] We have that $\int_{|x|=1} dS_x = d\omega_d$.

as $\varepsilon \to 0$. Thus we conclude that

$$\int E(x)\Delta\phi(x)\,\mathrm{d}x = \lim_{\varepsilon\to 0}\int_{\varepsilon<|x|<R} E(x)\Delta\phi(x)\,\mathrm{d}x = \phi(0). \tag{3.52}$$

We can hence write the solution of the Poisson equation

$$\Delta u = f$$

as

$$u(x) = \int_{\mathbb{R}^d} E(x-y)f(y)\,\mathrm{d}y, \quad x \in \mathbb{R}^d$$

with $E(x)$ given by (3.42).

How close are distributions to being functions? The following structure theorem gives the best answer in the general situation.

THEOREM 3.12. *Let $T \in \mathscr{D}'$ and $K \subseteq \mathbb{R}^d$ compact. Then there exists a continuous function f and a multi-index α such that*

$$\langle T, \phi\rangle = (-1)^{|\alpha|}\int f(x)\partial^\alpha\phi(x)\,\mathrm{d}x \tag{3.53}$$

for each $\phi \in C_0^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\phi \subseteq K$. In other words, each distribution is locally

$$T = \partial^\alpha f$$

in the sense of distributions for a continuous function f.

PROOF. For a proof, see [121, Thm. 6.26]. □

Thus we see that distributions are just locally the weak derivative of a continuous function.

A different way to characterize distributions is to ask how they can be approximated by smoother objects. The next theorem shows that indeed smooth functions are dense in the set of distributions.

THEOREM 3.13. *The set $C_0^\infty(\mathbb{R}^d)$ is dense in $\mathscr{D}'$.*

PROOF. Let $T \in \mathscr{D}'$. Consider any $\psi \in C_0^\infty(\mathbb{R}^d)$ with $\int \psi(x)\,\mathrm{d}x = 1$ and define

$$\psi_j(x) = j^d\psi(jx), \quad j \in \mathbb{N}. \tag{3.54}$$

Then $\psi_j \to \delta$ in $\mathscr{D}'$ as $j \to \infty$ and the support will shrink: If $\operatorname{supp}\psi \subset B_R(0)$, the ball of radius R centered at the origin, then $\operatorname{supp}\psi_j \subset B_{j^{-1}R}(0)$. Furthermore

$$\psi_j * T \to T, \quad j \to \infty \tag{3.55}$$

in $\mathscr{D}'$. Namely,

$$\int \psi_j * T(x)\phi(x)\,\mathrm{d}x = \int T(\tau_x\psi_{j,\sigma})\phi(x)\,\mathrm{d}x = \int \psi_{j,\sigma}(x)T(\tau_x\phi)\,\mathrm{d}x \to T(\phi), \tag{3.56}$$

for $\phi \in C_0^\infty(\mathbb{R}^d)$.

To obtain a function with compact support we proceed as follows. Take $\eta \in C_0^\infty(\mathbb{R}^d)$ such that $\eta(x) = 1$ for $|x| \le 1$ and define

$$\tilde{\psi}_j(x) = \eta(\frac{x}{j})\big(\psi_j * T\big)(x), \quad j \in \mathbb{N}. \tag{3.57}$$

Then

$$\langle T_{\tilde{\psi}_j}, \phi\rangle = \int \tilde{\psi}_j(x)\phi(x)\,\mathrm{d}x \to \langle T, \phi\rangle, \quad j \to \infty. \tag{3.58}$$

□

Often one will be able to prove that a certain differential equation has a distributional solution. The name of the game is then to show as much regularity of the distribution to make it into a function with pointwise values (almost everywhere) or even continuity or some differentiability. To that end one introduces various subsets of distributions, for instance, the set of tempered distributions.

Notes

Our presentation owes a lot to Gasquet and Witomski [59].

Exercises

(1) Show how to define an odd or even distribution.
(2) Show how to define a periodic distribution.
(3) Show (3.11).
(4) Show that $T(\phi) = \sum_{j\in\mathbb{Z}} \phi(n)$ for $\phi \in C_0^\infty$ defines a distribution.
(5) The functions $1/x$ and $1/x^2$ are not in L^1_{loc}, and hence do not define regular distributions. However, consider the function $f(x) = \ln|x| \in L^1_{\text{loc}}$. Use the definition of derivative of a distribution to derive explicit expressions in terms of the test function (and not its derivatives) that can be used to define distributions formally given by f' and $-f''$, respectively.