DISTRIBUTIONS AND FOURIER TRANSFORM

MIKKO SALO

Introduction. The theory of *distributions*, or *generalized functions*, provides a unified framework for performing standard calculus operations on nonsmooth functions, measures (such as the Dirac delta function), and even more general measure-like objects in the same way as they are done for smooth functions. In this theory, any distribution can be differentiated arbitrarily many times, a large class of distributions have well-defined Fourier transforms, and general linear operators can be expressed as integral operators with distributional kernel. The distributional point of view is very useful since it easily allows to perform such operations in a certain weak sense. However, often additional work is required if stronger statements are needed.

The theory in its modern form arose from the work of Laurent Schwartz in the late 1940s, although it certainly had important precursors such as Heaviside's operational calculus in the 1890s and Sobolev's generalized functions in the 1930s. The approach of Schwartz had the important feature of being completely mathematically rigorous while retaining the ease of calculation of the operational methods. Distributions have played a prominent role in the modern theory of partial differential equations, and they will be used heavily in the chapter on Microlocal methods in this encyclopedia.

The idea behing distribution theory is easily illustrated by the standard example, the Dirac delta function. On the real line, the Dirac delta is a "function $\delta(x)$ which is zero for $x \neq 0$ with an infinitely high peak at $x = 0$, with area equal to one". Thus, if $f(x)$ is a smooth function then integrating $\delta(x)f(x)$ is supposed to give

$$\int_{-\infty}^{\infty} \delta(x)f(x) = f(0).$$

The Dirac delta is not a well defined function (in fact it is a measure), but integration against $\delta(x)$ may be thought of as a linear operator defined on some class of test functions which for any test function f gives out the number $f(0)$. After suitable choices of test function spaces, distributions are introduced as continuous linear functionals on these spaces.

The following will be a quick introduction to distributions and the Fourier transform, mostly avoiding proofs. Further details can be found in [1], [2], [3].

Test functions and distributions. Let $\Omega \subset \mathbf{R}^n$ be an open set. We recall that if f is a continuous function on Ω, the support of f is the set

$$\mathrm{supp}(f) := \Omega \setminus V, \quad V \text{ is the largest open subset in } \Omega \text{ with } f|_V = 0.$$

Some notation: any n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbf{N}^n$ where $\mathbf{N} = \{0, 1, 2, \ldots\}$ is called a *multi-index*, and its norm is $|\alpha| = \alpha_1 + \ldots + \alpha_n$. We write

$$\partial^\alpha f(x) = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} f(x).$$

A function f on Ω is called C^∞, or infinitely differentiable, if $\partial^\alpha f$ is a continuous function on Ω for all $\alpha \in \mathbf{N}^n$. The following test function space will be used to define distributions.

Definition. The space of infinitely differentiable functions with compact support in Ω is defined as

$$C_c^\infty(\Omega) := \{f : \Omega \to \mathbf{C}\,;\, f \text{ is } C^\infty \text{ and } \mathrm{supp}(f) \text{ is compact in } \Omega\}.$$

If Ω is a domain with smooth boundary, then $\mathrm{supp}(f)$ is compact in Ω if and only if f vanishes near $\partial\Omega$. The space $C_c^\infty(\Omega)$ contains many functions, for instance it is not hard to see that

$$\eta(x) := \begin{cases} e^{-1/(1-|x|)^2}, & |x| < 1, \\ 0, & |x| \geq 1 \end{cases}$$

is in $C_c^\infty(\mathbf{R}^n)$. Generally, if $K \subset V \subset \overline{V} \subset \Omega$ where K is compact and V is open, there exists $\varphi \in C_c^\infty(\Omega)$ such that $\varphi = 1$ on K and $\mathrm{supp}(\varphi) \subset V$.

To define continuous linear functionals on $C_c^\infty(\Omega)$, we need a notion of convergence:

Definition. We say that a sequence $(\varphi_j)_{j=1}^\infty$ converges to φ in $C_c^\infty(\Omega)$ if there is a compact set $K \subset \Omega$ such that $\mathrm{supp}(\varphi_j) \subset K$ for all j, and if

$$\|\partial^\alpha(\varphi_j - \varphi)\|_{L^\infty(K)} \to 0 \text{ as } j \to \infty, \text{ for all } \alpha \in \mathbf{N}^n.$$

More precisely, one can define a topology on $C_c^\infty(\Omega)$ such that this space becomes a complete locally convex topological vector space, and a linear functional $u : C_c^\infty(\Omega) \to \mathbf{C}$ is continuous if and only if $u(\varphi_j) \to 0$ for any sequence (φ_j) such that $\varphi_j \to 0$ in $C_c^\infty(\Omega)$. We will not go further on this since the convergence of sequences is sufficient for most practical purposes.

We can now give a precise definition of distributions.

Definition. The set of distributions on Ω, denoted by $\mathcal{D}'(\Omega)$, is the set of all continuous linear functionals $u : C_c^\infty(\Omega) \to \mathbf{C}$.

Examples. 1. (Locally integrable functions) Let f be a locally integrable function in Ω, that is, $f : \Omega \to \mathbf{C}$ is Lebesgue measurable and $\int_K |f|\,dx < \infty$ for any compact $K \subset \Omega$. (In particular, any continuous or $L^1(\Omega)$ function is locally integrable.) We define

$$u_f : C_c^\infty(\Omega) \to \mathbf{C}, \quad u_f(\varphi) = \int_\Omega f(x)\varphi(x)\,dx.$$

By the definition of convergence of sequences, u_f is a well-defined distribution. If f_1, f_2 are two locally integrable functions and $u_{f_1} = u_{f_2}$, then $f_1 = f_2$ almost everywhere by the du Bois-Reymond lemma. Thus a locally integrable function f can be identified with the corresponding distribution u_f.

2. (Dirac mass) Fix $x_0 \in \Omega$ and define

$$\delta_{x_0} : C_c^\infty(\Omega) \to \mathbf{C}, \quad \delta_{x_0}(\varphi) = \varphi(x_0).$$

This is a well-defined distribution, called the *Dirac mass* at x_0.

3. (Measures) If μ is a positive or complex Borel measure in Ω such that the total variation $\int_K d|\mu| < \infty$ for any compact $K \subset \Omega$, then the operator

$$u_\mu : \varphi \mapsto \int_\Omega \varphi(x)\,d\mu(x)$$

is a distribution that can be identified with μ.

4. (Derivative of Dirac mass) On the real line, the operator

$$\delta_0' : \varphi \mapsto -\varphi'(0)$$

is a distribution which is not a measure.

We now wish to extend some operations, defined for smooth functions, to the case of distributions. This is possible via the duality of test functions and distributions. To emphasize this point, we employ the notation

$$\langle u, \varphi \rangle := u(\varphi), \quad u \in \mathcal{D}'(\Omega), \varphi \in C_c^\infty(\Omega).$$

Note that if u is a smooth function, then the duality is given by

$$\langle u, \varphi \rangle = \int_\Omega u(x)\varphi(x)\,dx.$$

Multiplication by functions. Let a be a C^∞ function in Ω. If $u, \varphi \in C_c^\infty(\Omega)$ we clearly have

$$\langle au, \varphi \rangle = \langle u, a\varphi \rangle.$$

Since $\varphi \mapsto a\varphi$ is continuous on $C_c^\infty(\Omega)$, we may for any $u \in \mathcal{D}'(\Omega)$ define the product au as the distribution given by

$$\langle au, \varphi \rangle := \langle u, a\varphi \rangle, \quad \varphi \in C_c^\infty(\Omega).$$

Distributional derivatives. Similarly, motivated by the corresponding property for smooth functions, if $u \in \mathcal{D}'(\Omega)$ and $\beta \in \mathbf{N}^n$ is a multi-index then the (distributional) derivative $\partial^\beta u$ is the distribution given by

$$\langle \partial^\beta u, \varphi \rangle := (-1)^{|\beta|} \langle u, \partial^\beta \varphi \rangle, \quad \varphi \in C_c^\infty(\Omega).$$

(If u is a smooth function this is true by the integration by parts formula

$$\int_\Omega u(x) \partial_{x_j} \varphi(x)\, dx = -\int_\Omega \partial_{x_j} u(x) \varphi(x)\, dx.)$$

The last fact is one of the most striking features of distributions: in this theory, any distribution (no matter how irregular) has infinitely many well defined derivatives!

Examples. 1. In Example 4 above, the distribution δ_0' is in fact the distributional derivative of the Dirac mass δ_0.

2. Let $u(x) := |x|,\ x \in \mathbf{R}$. Since u is continuous, we have $u \in \mathcal{D}'(\mathbf{R})$. We claim the one has the distributional derivatives

$$u' = \mathrm{sign}(x),$$
$$u'' = 2\delta_0.$$

In fact, if $\varphi \in C_c^\infty(\mathbf{R})$, one has

$$\langle u', \varphi \rangle = -\langle u, \varphi' \rangle = \int_{-\infty}^0 x\varphi'(x)\, dx - \int_0^\infty x\varphi'(x)\, dx$$
$$= \int_{\mathbf{R}} \mathrm{sign}(x)\varphi(x)\, dx = \langle \mathrm{sign}(x), \varphi \rangle,$$

using integration by parts and the compact support of φ. Similarly,

$$\langle u'', \varphi \rangle = -\langle u', \varphi' \rangle = \int_{-\infty}^0 \varphi'(x)\, dx - \int_0^\infty \varphi'(x)\, dx$$
$$= 2\varphi(0) = \langle 2\delta_0, \varphi \rangle.$$

Homogeneous distributions. We wish to discuss homogeneous distributions, which are useful in representing fundamental solutions of differential operators for instance. We concentrate on a particular example following [2, Section 3.2]. If $a > -1$, define

$$f_a(x) := \begin{cases} x^a, & x > 0, \\ 0, & x < 0. \end{cases}$$

This is a locally integrable function, and positively homogeneous of degree a in the sense that $f_a(tx) = t^a f(x)$ for $t > 0$. For $a > -1$ we can define the distribution $x_+^a := f_a$. If $a > 0$ it has the properties

$$x x_+^{a-1} = x_+^a,$$
$$(x_+^a)' = a x_+^{a-1}.$$

We would like to define x_+^a for any real number a as an element of $\mathcal{D}'(\mathbf{R})$ so that some of these properties remain valid.

First note that if $a > -1$, then for $k \in \mathbf{N}$ by repeated differentiation

$$\begin{aligned}\langle x_+^a, \varphi\rangle &= -\frac{1}{a+1}\langle x_+^{a+1}, \varphi'\rangle = \dots \\ &= (-1)^k \frac{1}{(a+1)(a+2)\cdots(a+k)}\langle x_+^{a+k}, \varphi^{(k)}\rangle.\end{aligned}$$

If $a \notin \{-1, -2, \dots\}$ we can define $x_+^a \in \mathcal{D}'(\mathbf{R})$ by the last formula.

If a is a negative integer, we need to regularize the expression x_+^a to obtain a valid distribution. For fixed $\varphi \in C_c^\infty(\mathbf{R})$, the quantity $F(a) = \langle x_+^a, \varphi\rangle = \int f_a(x)\varphi(x)\,dx$ can be extended as an analytic function for complex a with $\mathrm{Re}(a) > -1$. The formula above for x_+^a with negative a then shows that F is analytic is $\mathbf{C} \setminus \{-1, -2, \dots\}$, and it has simple poles at the negative integers with residues

$$\lim_{a \to -k}(a+k)F(a) = \frac{(-1)^k\langle x_+^0, \varphi^{(k)}\rangle}{(-k+1)(-k+2)\cdots(-1)} = \frac{\varphi^{(k-1)}(0)}{(k-1)!}.$$

We define $x_+^{-k} \in \mathcal{D}'(\mathbf{R})$, after a computation, by

$$\begin{aligned}\langle x_+^{-k}, \varphi\rangle &:= \lim_{a \to -k}\Big(F(a) - \frac{\varphi^{(k-1)}(0)}{(k-1)!(a+k)}\Big) \\ &= \frac{1}{(k-1)!}\left(-\int_0^\infty (\log x)\varphi^{(k)}(x)\,dx + \left(\sum_{j=1}^{k-1}\frac{1}{j}\right)\varphi^{(k-1)}(0)\right).\end{aligned}$$

Then $xx_+^{a-1} = x_+^a$ is valid for all $a \in \mathbf{R}$, and $(x_+^a)' = ax_+^{a-1}$ holds true except for nonpositive integers a.

Schwartz kernel theorem. One of the important features of distribution theory is that it allows to write almost any linear operator as an integral operator, at least in a weak sense. If $\Omega, \Omega' \subset \mathbf{R}^n$ are open sets and if $K \in L^2(\Omega \times \Omega')$, by Cauchy-Schwarz one has a bounded linear operator

$$T : L^2(\Omega') \to L^2(\Omega), \quad T\varphi(x) := \int_{\Omega'} K(x,y)\varphi(y)\,dy.$$

The function K is called the integral kernel of the operator T. There is a general one-to-one correspondence between continuous linear operators and integral kernels.

Theorem. *If $T : C_c^\infty(\Omega') \to \mathcal{D}'(\Omega)$ is a continuous linear map, then there is $K \in \mathcal{D}'(\Omega \times \Omega')$ such that*

$$\langle T(\varphi), \psi\rangle = \langle K, \psi \otimes \varphi\rangle, \quad \varphi \in C_c^\infty(\Omega'), \psi \in C_c^\infty(\Omega).$$

Here $(\psi \otimes \varphi)(x, y) = \psi(x)\varphi(y)$. *Conversely, any* $K \in \mathcal{D}'(\Omega \times \Omega')$ *gives rise to a continuous linear map* T *by the above formula.*

Tempered distributions. In the following, we will give a brief review of the Fourier transform in the general setting of tempered distributions. We introduce a test function space designed for the purposes of Fourier analysis.

Definition. The *Schwartz space* $\mathscr{S}(\mathbf{R}^n)$ is the set of all infinitely differentiable functions $f : \mathbf{R}^n \to \mathbf{C}$ such that the seminorms

$$\|f\|_{\alpha,\beta} := \|x^\alpha \partial^\beta f(x)\|_{L^\infty(\mathbf{R}^n)}$$

are finite for all multi-indices $\alpha, \beta \in \mathbf{N}^n$. If $(f_j)_{j=1}^\infty$ is a sequence in $\mathscr{S}$, we say that $f_j \to f$ in $\mathscr{S}$ if $\|f_j - f\|_{\alpha,\beta} \to 0$ for all α, β.

It follows from the definition that a smooth function f is in $\mathscr{S}(\mathbf{R}^n)$ iff for all β and N there exists $C > 0$ such that

$$|\partial^\beta f(x)| \leq C\langle x\rangle^{-N}, \quad x \in \mathbf{R}^n.$$

Here and below, $\langle x\rangle := (1 + |x|^2)^{1/2}$. Based on this, Schwartz space is sometimes called the space of *rapidly decreasing test functions.*

Example. Any function in $C_c^\infty(\mathbf{R}^n)$ is in Schwartz space, and functions like $e^{-\gamma|x|^2}$, $\gamma > 0$, also belong to $\mathscr{S}$. The function $e^{-\gamma|x|}$ is not in Schwartz space because it is not smooth at the origin, and also $\langle x\rangle^{-N}$ is not in $\mathscr{S}$ because it does not decrease sufficiently rapidly at infinity.

There is a topology on $\mathscr{S}$ such that $\mathscr{S}$ becomes a complete metric space. The operations $f \mapsto Pf$ and $f \mapsto \partial^\beta f$ are continuous maps $\mathscr{S} \to \mathscr{S}$, if P is any polynomial and β any multi-index. More generally, let

$$\mathscr{O}_M(\mathbf{R}^n) := \{f \in C^\infty(\mathbf{R}^n) \text{ ; for all } \beta \text{ there exist } C, N > 0 \text{ such that } |\partial^\beta f(x)| \leq C\langle x\rangle^N\}.$$

It is easy to see that the map $f \mapsto af$ is continuous $\mathscr{S} \to \mathscr{S}$ if $a \in \mathscr{O}_M$.

Definition. If $f \in \mathscr{S}(\mathbf{R}^n)$, then the *Fourier transform* of f is the function $\mathscr{F}f = \hat{f} : \mathbf{R}^n \to \mathbf{C}$ defined by

$$\hat{f}(\xi) := \int_{\mathbf{R}^n} e^{-ix\cdot\xi} f(x)\,dx, \quad \xi \in \mathbf{R}^n.$$

The importance of Schwartz space is based on the fact that it is invariant under the Fourier transform.

Theorem. *(Fourier inversion formula) The Fourier transform is an isomorphism from* $\mathscr{S}(\mathbf{R}^n)$ *onto* $\mathscr{S}(\mathbf{R}^n)$. *A Schwartz function* f *may be recovered from its Fourier transform by the inversion formula*

$$f(x) = \mathscr{F}^{-1}\hat{f}(x) = (2\pi)^{-n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \hat{f}(\xi)\,d\xi, \quad x \in \mathbf{R}^n.$$

After introducing the Fourier transform on nicely behaving functions, it is possible to define it in a very general setting by duality.

Definition. Let $\mathscr{S}'(\mathbf{R}^n)$ be the set of continuous linear functionals $\mathscr{S}(\mathbf{R}^n) \to \mathbf{C}$. The elements of $\mathscr{S}'$ are called *tempered distributions*, and their action on test functions is written as

$$\langle u, \varphi \rangle := u(\varphi), \qquad u \in \mathscr{S}', \ \varphi \in \mathscr{S}.$$

Since the embedding $C_c^\infty(\mathbf{R}^n) \subset \mathscr{S}(\mathbf{R}^n)$ is continuous, it follows that $\mathscr{S}'(\mathbf{R}^n) \subset \mathcal{D}'(\mathbf{R}^n)$, that is, tempered distributions are distributions. The elements in $\mathscr{S}'$ are somewhat loosely also called *distributions of polynomial growth.* The following examples are similar to the case of $\mathcal{D}'(\mathbf{R}^n)$ above.

Examples. 1. Let $f : \mathbf{R}^n \to \mathbf{C}$ be a measurable function, such that for some $C, N > 0$ one has

$$|f(x)| \leq C\langle x \rangle^N, \quad \text{for a.e. } x \in \mathbf{R}^n.$$

Then the corresponding distribution u_f is in $\mathscr{S}'(\mathbf{R}^n)$. The function f is usually identified with the tempered distribution u_f.

2. In a similar way, any function $f \in L^p(\mathbf{R}^n)$ with $1 \leq p \leq \infty$ is a tempered distribution (with the identification $f = u_f$).
3. Let μ be a positive Borel measure in $\mathbf{R}^n$ such that

$$\int_{\mathbf{R}^n} \langle x \rangle^{-N} \, d\mu(x) < \infty$$

for some $N > 0$. Then the corresponding distribution u_μ is tempered. In particular, the Dirac mass δ_{x_0} at $x_0 \in \mathbf{R}^n$ is in $\mathscr{S}'$.
4. The function $e^{\gamma x}$ is in $\mathcal{D}'(\mathbf{R})$ but not in $\mathscr{S}'(\mathbf{R})$ if $\gamma \neq 0$, since it is not possible to define $\int_{\mathbf{R}} e^{\gamma x} \varphi(x) \, dx$ for all Schwartz functions φ. However, $e^{\gamma x} \cos(e^{\gamma x})$ belongs to $\mathscr{S}'$ since it is the distributional derivative (see below) of the bounded function $\sin(e^{\gamma x}) \in \mathscr{S}'$.

Operations on tempered distributions. The operations defined above for $\mathcal{D}'(\mathbf{R}^n)$ have natural analogues for tempered distributions. For instance, if $a \in \mathscr{O}_M(\mathbf{R}^n)$ and $u \in \mathscr{S}'(\mathbf{R}^n)$ then au is a tempered distribution where

$$\langle au, \varphi \rangle = \langle u, a\varphi \rangle.$$

Similarly, if $u \in \mathscr{S}'(\mathbf{R}^n)$ then the distributional derivative $\partial^\beta u$ is also a tempered distribution.

Finally, we can define the Fourier transform of any $u \in \mathscr{S}'$ as the tempered distribution $\mathscr{F}u = \hat{u}$ with

$$\langle \hat{u}, \varphi \rangle := \langle u, \hat{\varphi} \rangle, \quad \varphi \in \mathscr{S}.$$

In fact, this identity is true if $u, \varphi \in \mathscr{S}$ and it then extends the Fourier transform on Schwartz space to the case of tempered distributions.

Example. The Fourier transform of δ_0 is the constant 1, since

$$\begin{aligned}\langle \widehat{\delta}_0, \varphi \rangle = \langle \delta_0, \hat{\varphi} \rangle = \hat{\varphi}(0) &= \int_{-\infty}^{\infty} \varphi(x)\,dx \\ &= \langle 1, \varphi \rangle.\end{aligned}$$

If $u \in L^2(\mathbf{R}^n)$ then u is a tempered distribution, and the Fourier transform $\hat{u}$ is another element of $\mathscr{S}'$. The Plancherel theorem (which is the exact analog of Parseval's theorem for Fourier series) states that in fact $\hat{u} \in L^2$, and that the Fourier transform is an isometry on L^2 up to a constant. We state the basic properties of the Fourier transform as a theorem.

Theorem. *The Fourier transform $u \mapsto \hat{u}$ is a bijective map from $\mathscr{S}'$ onto $\mathscr{S}'$, and one has the inversion formula*

$$\langle u, \varphi \rangle = (2\pi)^{-n} \langle \hat{u}, \hat{\varphi}(-\,\cdot\,) \rangle, \quad \varphi \in \mathscr{S}.$$

The Fourier transform is also an isomorphism from $L^2(\mathbf{R}^n)$ onto $L^2(\mathbf{R}^n)$, and

$$\|\hat{u}\|_{L^2} = (2\pi)^{n/2} \|u\|_{L^2}.$$

It is remarkable that any tempered distribution has a Fourier transform (thus, also any L^p function or measurable polynomially bounded function), and there is a Fourier inversion formula for recovering the original distribution from its Fourier transform.

We end this section by noting the identities

$$\begin{aligned}(\partial^\alpha u)\hat{} &= (i\xi)^\alpha \hat{u}, \\ (x^\alpha u)\hat{} &= (i\partial_\xi)^\alpha \hat{u},\end{aligned}$$

where $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. These hold for Schwartz functions u by a direct computation, and remain true for tempered distributions u by duality. Thus the Fourier transform converts derivatives into multiplication by polynomials, and vice versa. This explains why the Fourier transform is useful in the study of partial differential equations, since it can be used to convert constant coefficient differential equations into algebraic equations.

References

[1] F.G. Friedlander and M. Joshi, *Introduction to the theory of distributions*, 2nd edition, Cambridge University Press, 1998.

[2] L. Hörmander, *The analysis of linear partial differential operators, volume I*, 2nd edition, Springer-Verlag, 1990.

[3] L. Schwartz, *Théorie des distributions*, Hermann, 1950.