WIKIPEDIA

Distribution (mathematics)

Distributions, also known as **Schwartz distributions** or **generalized functions**, are objects that generalize the classical notion of functions in mathematical analysis. Distributions make it possible to differentiate functions whose derivatives do not exist in the classical sense. In particular, any locally integrable function has a distributional derivative.

Distributions are widely used in the theory of partial differential equations, where it may be easier to establish the existence of distributional solutions than classical solutions, or where appropriate classical solutions may not exist. Distributions are also important in physics and engineering where many problems naturally lead to differential equations whose solutions or initial conditions are singular, such as the Dirac delta function.

A function f is normally thought of as *acting* on the *points* in its domain by sending a point x in its domain to the point $f(x)$. Instead, distribution theory reinterprets functions as being equivalent to their dual linear functionals: if every nicely behaved test function integrated with a function gives out the same result, this defines a linear functional equivalent to every conventional function.

However, this kind of definition is much laxer, and admits mathematical objects beyond functions. In particular, these kinds of generalised functions can be used to represent singular measures, such as the delta function, and all of its derivatives. Since the distributional framework is localised, linear, and shift-invariant, it can represent almost all of the compositions of the basis waveforms as well. It also admits a full Fourier theory, and since the theory is fully localised, there is no need to have functions fall off in time or shift.

In applications to physics and engineering, ***test functions*** are usually infinitely differentiable complex-valued (or real-valued) functions with compact support that are defined on some given non-empty open subset $U \subseteq \mathbb{R}^n$ (bump functions are examples of test functions). The set of all such test functions forms a vector space that is denoted by $C_c^\infty(U)$ or $\mathcal{D}(U)$.

Most commonly encountered functions, including all continuous maps $f : \mathbb{R} \to \mathbb{R}$ if using $U := \mathbb{R}$, can be canonically reinterpreted as acting via "integration against a test function." Explicitly, this means that f "acts on" a test function $\psi \in \mathcal{D}(\mathbb{R})$ by "sending" it to the number $\int_{\mathbb{R}} f\,\psi\,dx$, which is often denoted by $D_f(\psi)$. This new action $\psi \mapsto D_f(\psi)$ of f is a scalar-valued map, denoted by D_f, whose domain is the space of test functions $\mathcal{D}(\mathbb{R})$. This functional D_f turns out to have the two defining properties of what is known as a *distribution on* $\mathbb{R}$: it is linear and also continuous when $\mathcal{D}(\mathbb{R})$ is given a certain topology called *the canonical LF topology*. Distributions like D_f that arise from functions in this way are prototypical examples of distributions, but there are many which cannot be defined by integration against any function. Examples of the latter include the Dirac delta function and distributions defined to act by integration of test functions against certain measures. It is nonetheless still possible to reduce any arbitrary distribution down to a simpler *family* of related distributions that do arise via such actions of integration.

More generally, a ***distribution*** *on* U is by definition a linear functional on $C_c^\infty(U)$ that is continuous when $C_c^\infty(U)$ is given a topology called the ***canonical LF topology***. This leads to *the* space of (all) distributions on U, usually denoted by $\mathcal{D}'(U)$ (note the prime), which by definition is the space of all distributions on U (that is, it is the continuous dual space of $C_c^\infty(U)$); it is these distributions that are the main focus of this article.

Definitions of the appropriate topologies on spaces of test functions and distributions are given in the article on spaces of test functions and distributions. This article is primarily concerned with the definition of distributions, together with their properties and some important examples.

Contents

History

The practical use of distributions can be traced back to the use of Green functions in the 1830s to solve ordinary differential equations, but was not formalized until much later. According to Kolmogorov & Fomin (1957), generalized functions originated in the work of Sergei Sobolev (1936) on second-order hyperbolic partial differential equations, and the ideas were developed in somewhat extended form by Laurent Schwartz in the late 1940s. According to his autobiography, Schwartz introduced the term "distribution" by analogy with a distribution of electrical charge, possibly including not only point charges but also dipoles and so on. Gårding (1997) comments that although the ideas in the transformative book by Schwartz (1951) were not entirely new, it was Schwartz's broad attack and conviction that distributions would be useful almost everywhere in analysis that made the difference.

Notation

The following notation will be used throughout this article:

- n is a fixed positive integer and U is a fixed non-empty open subset of Euclidean space $\mathbb{R}^n$.
- $\mathbb{N} = \{0, 1, 2, \ldots\}$ denotes the natural numbers.
- k will denote a non-negative integer or ∞.
- If f is a function then $\operatorname{Dom}(f)$ will denote its domain and the ***support*** of f, denoted by $\operatorname{supp}(f)$, is defined to be the closure of the set $\{x \in \operatorname{Dom}(f) : f(x) \neq 0\}$ in $\operatorname{Dom}(f)$.
- For two functions $f, g : U \to \mathbb{C}$, the following notation defines a canonical pairing:

$$\langle f, g \rangle := \int_U f(x) g(x) \, dx.$$

- A ***multi-index*** **of size** n is an element in $\mathbb{N}^n$ (given that n is fixed, if the size of multi-indices is omitted then the size should be assumed to be n). The ***length*** of a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ is defined as $\alpha_1 + \cdots + \alpha_n$ and denoted by $|\alpha|$. Multi-indices are particularly useful when dealing with functions of several variables, in particular we introduce the following notations for a given multi-index $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$:

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$
$$\partial^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$$

We also introduce a partial order of all multi-indices by $\beta \geq \alpha$ if and only if $\beta_i \geq \alpha_i$ for all $1 \leq i \leq n$. When $\beta \geq \alpha$ we define their multi-index binomial coefficient as:

$$\binom{\beta}{\alpha} := \binom{\beta_1}{\alpha_1} \cdots \binom{\beta_n}{\alpha_n}.$$

Definitions of test functions and distributions

In this section, some basic notions and definitions needed to define real-valued distributions on U are introduced. Further discussion of the topologies on the spaces of test functions and distributions are given in the article on spaces of test functions and distributions.

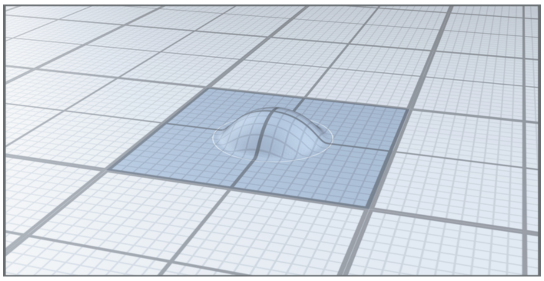
Notation:

1. Let $k \in \{0, 1, 2, \ldots, \infty\}$.
2. Let $C^k(U)$ denote the vector space of all k-times continuously differentiable real or complex-valued functions on U.
3. For any compact subset $K \subseteq U$, let $C^k(K)$ and $C^k(K; U)$ both denote the vector space of all those functions $f \in C^k(U)$ such that $\operatorname{supp}(f) \subseteq K$.
   - Note that $C^k(K)$ depends on both K and U but we will only indicate K, where in particular, if $f \in C^k(K)$ then the domain of f is U rather than K. We will use the notation $C^k(K; U)$ only when the notation $C^k(K)$ risks being ambiguous.
   - Every $C^k(K)$ contains the constant 0 map, even if $K = \varnothing$.
4. Let $C_c^k(U)$ denote the set of all $f \in C^k(U)$ such that $f \in C^k(K)$ for some compact subset K of U.
   - Equivalently, $C_c^k(U)$ is the set of all $f \in C^k(U)$ such that f has compact support.
   - $C_c^k(U)$ is equal to the union of all $C^k(K)$ as $K \subseteq U$ ranges over all compact subsets of U.
   - If f is a real-valued function on U, then f is an element of $C_c^k(U)$ if and only if f is a C^k bump function. Every real-valued test function on U is always also a complex-valued test function on U.

For all $j, k \in \{0, 1, 2, \ldots, \infty\}$ and any compact subsets K and L of U, we have:

$$\begin{aligned}
&C^k(K) \subseteq C_c^k(U) \subseteq C^k(U) && \\
&C^k(K) \subseteq C^k(L) && \text{if } K \subseteq L \\
&C^k(K) \subseteq C^j(K) && \text{if } j \leq k \\
&C_c^k(U) \subseteq C_c^j(U) && \text{if } j \leq k \\
&C^k(U) \subseteq C^j(U) && \text{if } j \leq k
\end{aligned}$$

Definition: Elements of $C_c^\infty(U)$ are called ***test functions*** on U and $C_c^\infty(U)$ is called the ***space of test functions*** on U. We will use both $\mathcal{D}(U)$ and $C_c^\infty(U)$ to denote this space.



The graph of the bump function $(x, y) \in \mathbb{R}^2 \mapsto \Psi(r)$, where $r = \left(x^2 + y^2\right)^{\frac{1}{2}}$ and $\Psi(r) = e^{-\frac{1}{1-r^2}} \cdot \mathbf{1}_{\{|r|<1\}}$. This function is a test function on $\mathbb{R}^2$ and is an element of $C_c^\infty\left(\mathbb{R}^2\right)$. The support of this function is the closed unit disk in $\mathbb{R}^2$. It is non-zero on the open unit disk and it is equal to 0 everywhere outside of it.

Distributions on U are continuous linear functionals on $C_c^\infty(U)$ when this vector space is endowed with a particular topology called the ***canonical LF-topology***. The following proposition states two necessary and sufficient conditions for the continuity of a linear functional on $C_c^\infty(U)$ that are often straightforward to verify.

Proposition: A linear functional T on $C_c^\infty(U)$ is continuous, and therefore a distribution, if and only if either of the following equivalent conditions are satisfied:

1. For every compact subset $K \subseteq U$ there exist constants $C > 0$ and $N \in \mathbb{N}$ (dependent on K) such that for all $f \in C_c^\infty(U)$ with support contained in K,[1][2]
$$|T(f)| \leq C \sup\{|\partial^\alpha f(x)| : x \in U, |\alpha| \leq N\}.$$
2. For every compact subset $K \subseteq U$ and every sequence $\{f_i\}_{i=1}^\infty$ in $C_c^\infty(U)$ whose supports are contained in K, if $\{\partial^\alpha f_i\}_{i=1}^\infty$ converges uniformly to zero on U for every multi-index α, then $T(f_i) \to 0$.

Topology on $C^k(U)$

We now introduce the seminorms that will define the topology on $C^k(U)$. Different authors sometimes use different families of seminorms so we list the most common families below. However, the resulting topology is the same no matter which family is used.

Suppose $k \in \{0, 1, 2, \ldots, \infty\}$ and K is an arbitrary compact subset of U. Suppose i an integer such that $0 \leq i \leq k$.[note 1] and p is a multi-index with length $|p| \leq k$. For $K \neq \varnothing$, define:

$$
\begin{aligned}
(1)\ \ & s_{p,K}(f) := \sup_{x_0 \in K} \left|\partial^p f(x_0)\right| \\
(2)\ \ & q_{i,K}(f) := \sup_{|p| \leq i} \left(\sup_{x_0 \in K} \left|\partial^p f(x_0)\right|\right) = \sup_{|p| \leq i} \left(s_{p,K}(f)\right) \\
(3)\ \ & r_{i,K}(f) := \sup_{\substack{|p| \leq i \\ x_0 \in K}} \left|\partial^p f(x_0)\right| \\
(4)\ \ & t_{i,K}(f) := \sup_{x_0 \in K} \left(\sum_{|p| \leq i} \left|\partial^p f(x_0)\right|\right)
\end{aligned}
$$

while for $K = \varnothing$, define all the functions above to be the constant 0 map.

Each of the functions above are non-negative $\mathbb{R}$-valued[note 2] seminorms on $C^k(U)$.

Each of the following families of seminorms generates the same locally convex vector topology on $C^k(U)$:

$$
\begin{array}{ll}
(1) & \{q_{i,K} \ :\ K \text{ compact and } \ i \in \mathbb{N} \text{ satisfies } \ 0 \leq i \leq k\} \\
(2) & \{r_{i,K} \ :\ K \text{ compact and } \ i \in \mathbb{N} \text{ satisfies } \ 0 \leq i \leq k\} \\
(3) & \{t_{i,K} \ :\ K \text{ compact and } \ i \in \mathbb{N} \text{ satisfies } \ 0 \leq i \leq k\} \\
(4) & \{s_{p,K} \ :\ K \text{ compact and } \ p \in \mathbb{N}^n \text{ satisfies } \ |p| \leq k\}
\end{array}
$$

The vector space $C^k(U)$ is endowed with the locally convex topology induced by any one of the four families of seminorms described above; equivalently, the topology is the vector topology induced by *all* of the seminorms above (that is, by the union of the four families of seminorms).

With this topology, $C^k(U)$ becomes a locally convex (*non*-normable) Fréchet space and all of the seminorms defined above are continuous on this space. *All* of the seminorms defined above are continuous functions on $C^k(U)$. Under this topology, a net $(f_i)_{i\in I}$ in $C^k(U)$ converges to $f \in C^k(U)$ if and only if for every multi-index p with $|p| < k+1$ and every compact K, the net of partial derivatives $(\partial^p f_i)_{i\in I}$ converges uniformly to $\partial^p f$ on K.[3] For any $k \in \{0,1,2,\ldots,\infty\}$, any (von Neumann) bounded subset of $C^{k+1}(U)$ is a relatively compact subset of $C^k(U)$.[4] In particular, a subset of $C^\infty(U)$ is bounded if and only if it is bounded in $C^i(U)$ for all $i \in \mathbb{N}$.[4] The space $C^k(U)$ is a Montel space if and only if $k = \infty$.[5]

A subset W of $C^\infty(U)$ is open in this topology if and only if there exists $i \in \mathbb{N}$ such that W is open when $C^\infty(U)$ is endowed with the subspace topology induced on it by $C^i(U)$.

Topology on $C^k(K)$

As before, fix $k \in \{0,1,2,\ldots,\infty\}$. Recall that if K is any compact subset of U then $C^k(K) \subseteq C^k(U)$.

> **Assumption**: For any compact subset $K \subseteq U$, we will henceforth assume that $C^k(K)$ is endowed with the subspace topology it inherits from the Fréchet space $C^k(U)$.

If k is finite then $C^k(K)$ is a Banach space[6] with a topology that can be defined by the norm

$$r_K(f) := \sup_{|p|<k}\left(\sup_{x_0\in K}|\partial^p f(x_0)|\right).$$

And when $k = 2$, then $C^k(K)$ is even a Hilbert space.[6]

Trivial extensions and independence of $C^k(K)$'s topology from U

The definition of $C^k(K)$ depends on U so we will let $C^k(K;U)$ denote the topological space $C^k(K)$, which by definition is a topological subspace of $C^k(U)$. Suppose V is an open subset of $\mathbb{R}^n$ containing U and for any compact subset $K \subseteq V$, let $C^k(K;V)$ is the vector subspace of $C^k(V)$ consisting of maps with support contained in K. Given $f \in C_c^k(U)$, its ***trivial extension*** *to* V is by definition, the function $I(f) := F : V \to \mathbb{C}$ defined by:

$$F(x) = \begin{cases} f(x) & x \in U, \\ 0 & \text{otherwise}, \end{cases}$$

so that $F \in C^k(V)$. Let $I : C_c^k(U) \to C^k(V)$ denote the map that sends a function in $C_c^k(U)$ to its trivial extension on V. This map is a linear injection and for every compact subset $K \subseteq U$ (where K is also a compact subset of V since $K \subseteq U \subseteq V$) we have

$$\begin{aligned} I\left(C^k(K;U)\right) &= C^k(K;V) \qquad \text{and thus} \\ I\left(C_c^k(U)\right) &\subseteq C_c^k(V) \end{aligned}$$

If I is restricted to $C^k(K;U)$ then the following induced linear map is a homeomorphism (and thus a TVS-isomorphism):

$$\begin{aligned} C^k(K;U) &\to C^k(K;V) \\ f &\mapsto I(f) \end{aligned}$$

and thus the next map is a topological embedding:

$$\begin{aligned} C^k(K;U) &\to C^k(V) \\ f &\mapsto I(f). \end{aligned}$$

Using the injection

$$I : C_c^k(U) \to C^k(V)$$

the vector space $C_c^k(U)$ is canonically identified with its image in $C_c^k(V) \subseteq C^k(V)$. Because $C^k(K;U) \subseteq C_c^k(U)$, through this identification, $C^k(K;U)$ can also be considered as a subset of $C^k(V)$. Thus the topology on $C^k(K;U)$ is independent of the open subset U of $\mathbb{R}^n$ that contains K.[7] This justifies the practice of written $C^k(K)$ instead of $C^k(K;U)$.

Canonical LF topology

Recall that $C_c^k(U)$ denote all those functions in $C^k(U)$ that have compact support in U, where note that $C_c^k(U)$ is the union of all $C^k(K)$ as K ranges over all compact subsets of U. Moreover, for every k, $C_c^k(U)$ is a dense subset of $C^k(U)$. The special case when $k = \infty$ gives us the space of test functions.

> $C_c^\infty(U)$ is called the ***space of test functions*** *on* U and it may also be denoted by $\mathcal{D}(U)$. Unless indicated otherwise, it is endowed with a topology called ***the canonical LF topology***, whose definition is given in the article: Spaces of test functions and distributions.

The canonical LF-topology is *not* metrizable and importantly, it is ***strictly*** *finer* than the subspace topology that $C^\infty(U)$ induces on $C_c^\infty(U)$. However, the canonical LF-topology does make $C_c^\infty(U)$ into a complete reflexive nuclear[8] Montel[9] bornological barrelled Mackey space; the same is true of its strong dual space (that is, the space of all distributions with its usual topology). The canonical LF-topology can be defined in various ways.

Distributions

As discussed earlier, continuous linear functionals on a $C_c^\infty(U)$ are known as distributions on U. Other equivalent definitions are described below.

> By definition, a ***distribution*** *on* U is a continuous linear functional on $C_c^\infty(U)$. Said differently, a distribution on U is an element of the continuous dual space of $C_c^\infty(U)$ when $C_c^\infty(U)$ is endowed with its canonical LF topology.

There is a canonical duality pairing between a distribution T on U and a test function $f \in C_c^\infty(U)$, which is denoted using angle brackets by

$$\begin{cases} \mathcal{D}'(U) \times C_c^\infty(U) \to \mathbb{R} \\ (T, f) \mapsto \langle T, f \rangle := T(f) \end{cases}$$

One interprets this notation as the distribution T acting on the test function f to give a scalar, or symmetrically as the test function f acting on the distribution T.

Characterizations of distributions

Proposition. If T is a linear functional on $C_c^\infty(U)$ then the following are equivalent:

1. T is a distribution;
2. T is continuous;
3. T is continuous at the origin;
4. T is uniformly continuous;
5. T is a bounded operator;
6. T is sequentially continuous;
   - explicitly, for every sequence $(f_i)_{i=1}^\infty$ in $C_c^\infty(U)$ that converges in $C_c^\infty(U)$ to some $f \in C_c^\infty(U)$, $\lim_{i\to\infty} T(f_i) = T(f)$;[note 3]
7. T is sequentially continuous at the origin; in other words, T maps null sequences[note 4] to null sequences;
   - explicitly, for every sequence $(f_i)_{i=1}^\infty$ in $C_c^\infty(U)$ that converges in $C_c^\infty(U)$ to the origin (such a sequence is called a *null sequence*), $\lim_{i\to\infty} T(f_i) = 0$;
   - a *null sequence* is by definition any sequence that converges to the origin;
8. T maps null sequences to bounded subsets;
   - explicitly, for every sequence $(f_i)_{i=1}^\infty$ in $C_c^\infty(U)$ that converges in $C_c^\infty(U)$ to the origin, the sequence $(T(f_i))_{i=1}^\infty$ is bounded;
9. T maps Mackey convergent null sequences to bounded subsets;
   - explicitly, for every Mackey convergent null sequence $(f_i)_{i=1}^\infty$ in $C_c^\infty(U)$, the sequence $(T(f_i))_{i=1}^\infty$ is bounded;
   - a sequence $f_\bullet = (f_i)_{i=1}^\infty$ is said to be *Mackey convergent to the origin* if there exists a divergent sequence $r_\bullet = (r_i)_{i=1}^\infty \to \infty$ of positive real number such that the sequence $(r_i f_i)_{i=1}^\infty$ is bounded; every sequence that is Mackey convergent to the origin necessarily converges to the origin (in the usual sense);
10. The kernel of T is a closed subspace of $C_c^\infty(U)$;
11. The graph of T is closed;

12. There exists a continuous seminorm g on $C_c^\infty(U)$ such that $|T| \leq g$;
13. There exists a constant $C > 0$ and a finite subset $\{g_1, \ldots, g_m\} \subseteq \mathcal{P}$ (where $\mathcal{P}$ is any collection of continuous seminorms that defines the canonical LF topology on $C_c^\infty(U)$) such that $|T| \leq C(g_1 + \cdots + g_m)$;[note 5]
14. For every compact subset $K \subseteq U$ there exist constants $C > 0$ and $N \in \mathbb{N}$ such that for all $f \in C^\infty(K)$,[1]
$$|T(f)| \leq C \sup\{|\partial^\alpha f(x)| : x \in U, |\alpha| \leq N\};$$
15. For every compact subset $K \subseteq U$ there exist constants $C_K > 0$ and $N_K \in \mathbb{N}$ such that for all $f \in C_c^\infty(U)$ with support contained in K,[10]
$$|T(f)| \leq C_K \sup\{|\partial^\alpha f(x)| : x \in K, |\alpha| \leq N_K\};$$
16. For any compact subset $K \subseteq U$ and any sequence $\{f_i\}_{i=1}^\infty$ in $C^\infty(K)$, if $\{\partial^p f_i\}_{i=1}^\infty$ converges uniformly to zero for all multi-indices p, then $T(f_i) \to 0$;

Topology on the space of distributions and its relation to the weak-* topology

The set of all distributions on U is the continuous dual space of $C_c^\infty(U)$, which when endowed with the strong dual topology is denoted by $\mathcal{D}'(U)$. Importantly, unless indicated otherwise, the topology on $\mathcal{D}'(U)$ is the strong dual topology; if the topology is instead the weak-* topology then this will be clearly indicated. Neither topology is metrizable although unlike the weak-* topology, the strong dual topology makes $\mathcal{D}'(U)$ into a complete nuclear space, to name just a few of its desirable properties.

Neither $C_c^\infty(U)$ nor its strong dual $\mathcal{D}'(U)$ is a sequential space and so neither of their topologies can be fully described by sequences (in other words, defining only what sequences converge in these spaces is *not* enough to fully/correctly define their topologies). However, a *sequence* in $\mathcal{D}'(U)$ converges in the strong dual topology if and only if it converges in the weak-* topology (this leads many authors to use pointwise convergence to actually *define* the convergence of a sequence of distributions; this is fine for sequences but this is *not* guaranteed to extend to the convergence of nets of distributions because a net may converge pointwise but fail to converge in the strong dual topology). More information about the topology that $\mathcal{D}'(U)$ is endowed with can be found in the article on spaces of test functions and distributions and in the articles on polar topologies and dual systems.

A *linear* map from $\mathcal{D}'(U)$ into another locally convex topological vector space (such as any normed space) is continuous if and only if it is sequentially continuous at the origin. However, this is no longer guaranteed if the map is not linear or for maps valued in more general topological spaces (for example, that are not also locally convex topological vector spaces). The same is true of maps from $C_c^\infty(U)$ (more generally, this is true of maps from any locally convex bornological space).

Localization of distributions

There is no way to define the value of a distribution in $\mathcal{D}'(U)$ at a particular point of U. However, as is the case with functions, distributions on U restrict to give distributions on open subsets of U. Furthermore, distributions are *locally determined* in the sense that a distribution on all of U can be assembled from a distribution on an open cover of U satisfying some compatibility conditions on the overlaps. Such a structure is known as a sheaf.

Extensions and restrictions to an open subset

Let U and V be open subsets of $\mathbb{R}^n$ with $V \subseteq U$. Let $E_{VU} : \mathcal{D}(V) \to \mathcal{D}(U)$ be the operator which *extends by zero* a given smooth function compactly supported in V to a smooth function compactly supported in the larger set U. The transpose of E_{VU} is called the restriction mapping and is denoted by $\rho_{VU} := {}^t E_{VU} : \mathcal{D}'(U) \to \mathcal{D}'(V)$.

The map $E_{VU} : \mathcal{D}(V) \to \mathcal{D}(U)$ is a continuous injection where if $V \subseteq U$ then it is *not* a topological embedding and its range is *not* dense in $\mathcal{D}(U)$, which implies that this map's transpose is neither injective nor surjective.[11] A distribution $S \in \mathcal{D}'(V)$ is said to be ***extendible to U*** if it belongs to the range of the transpose of E_{VU} and it is called ***extendible*** if it is extendable to $\mathbb{R}^n$.[11]

For any distribution $T \in \mathcal{D}'(U)$, the restriction $\rho_{VU}(T)$ is a distribution in $\mathcal{D}'(V)$ defined by:

$$\langle \rho_{VU} T, \phi \rangle = \langle T, E_{VU} \phi \rangle \quad \text{for all } \phi \in \mathcal{D}(V).$$

Unless $U = V$, the restriction to V is neither injective nor surjective. Lack of surjectivity follows since distributions can blow up towards the boundary of V. For instance, if $U = \mathbb{R}$ and $V = (0, 2)$, then the distribution

$$T(x) = \sum_{n=1}^\infty n\, \delta\left(x - \frac{1}{n}\right)$$

is in $\mathcal{D}'(V)$ but admits no extension to $\mathcal{D}'(U)$.

Gluing and distributions that vanish in a set

Theorem[12] — Let $(U_i)_{i\in I}$ be a collection of open subsets of $\mathbb{R}^n$. For each $i \in I$, let $T_i \in \mathcal{D}'(U_i)$ and suppose that for all $i, j \in I$, the restriction of T_i to $U_i \cap U_j$ is equal to the restriction of T_j to $U_i \cap U_j$ (note that both restrictions are elements of $\mathcal{D}'(U_i \cap U_j)$). Then there exists a unique $T \in \mathcal{D}'(\bigcup_{i\in I} U_i)$ such that for all $i \in I$, the restriction of T to U_i is equal to T_i.

Let V be an open subset of U. $T \in \mathcal{D}'(U)$ is said to ***vanish in** V* if for all $f \in \mathcal{D}(U)$ such that $\operatorname{supp}(f) \subseteq V$ we have $Tf = 0$. T vanishes in V if and only if the restriction of T to V is equal to 0, or equivalently, if and only if T lies in the kernel of the restriction map ρ_{VU}.

Corollary[12] — Let $(U_i)_{i\in I}$ be a collection of open subsets of $\mathbb{R}^n$ and let $T \in \mathcal{D}'(\bigcup_{i\in I} U_i)$. $T = 0$ if and only if for each $i \in I$, the restriction of T to U_i is equal to 0.

Corollary[12] — The union of all open subsets of U in which a distribution T vanishes is an open subset of U in which T vanishes.

Support of a distribution

This last corollary implies that for every distribution T on U, there exists a unique largest subset V of U such that T vanishes in V (and does not vanish in any open subset of U that is not contained in V); the complement in U of this unique largest open subset is called *the **support** of T*.[12] Thus

$$\operatorname{supp}(T) = U \setminus \bigcup\{V \mid \rho_{VU}T = 0\}.$$

If f is a locally integrable function on U and if D_f is its associated distribution, then the support of D_f is the smallest closed subset of U in the complement of which f is almost everywhere equal to 0.[12] If f is continuous, then the support of D_f is equal to the closure of the set of points in U at which f does not vanish.[12] The support of the distribution associated with the Dirac measure at a point x_0 is the set $\{x_0\}$.[12] If the support of a test function f does not intersect the support of a distribution T then $Tf = 0$. A distribution T is 0 if and only if its support is empty. If $f \in C^\infty(U)$ is identically 1 on some open set containing the support of a distribution T then $fT = T$. If the support of a distribution T is compact then it has finite order and furthermore, there is a constant C and a non-negative integer N such that:[7]

$$|T\phi| \leq C\|\phi\|_N := C\sup\{|\partial^\alpha \phi(x)| : x \in U, |\alpha| \leq N\} \quad \text{for all } \phi \in \mathcal{D}(U).$$

If T has compact support then it has a unique extension to a continuous linear functional $\widehat{T}$ on $C^\infty(U)$; this functional can be defined by $\widehat{T}(f) := T(\psi f)$, where $\psi \in \mathcal{D}(U)$ is any function that is identically 1 on an open set containing the support of T.[7]

If $S, T \in \mathcal{D}'(U)$ and $\lambda \neq 0$ then $\operatorname{supp}(S+T) \subseteq \operatorname{supp}(S) \cup \operatorname{supp}(T)$ and $\operatorname{supp}(\lambda T) = \operatorname{supp}(T)$. Thus, distributions with support in a given subset $A \subseteq U$ form a vector subspace of $\mathcal{D}'(U)$.[13] Furthermore, if P is a differential operator in U, then for all distributions T on U and all $f \in C^\infty(U)$ we have $\operatorname{supp}(P(x,\partial)T) \subseteq \operatorname{supp}(T)$ and $\operatorname{supp}(fT) \subseteq \operatorname{supp}(f) \cap \operatorname{supp}(T)$.[13]

Distributions with compact support

Support in a point set and Dirac measures

For any $x \in U$, let $\delta_x \in \mathcal{D}'(U)$ denote the distribution induced by the Dirac measure at x. For any $x_0 \in U$ and distribution $T \in \mathcal{D}'(U)$, the support of T is contained in $\{x_0\}$ if and only if T is a finite linear combination of derivatives of the Dirac measure at x_0.[14] If in addition the order of T is $\leq k$ then there exist constants α_p such that:[15]

$$T = \sum_{|p|\leq k} \alpha_p \partial^p \delta_{x_0}.$$

Said differently, if T has support at a single point $\{P\}$, then T is in fact a finite linear combination of distributional derivatives of the δ function at P. That is, there exists an integer m and complex constants a_α such that

$$T = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha (\tau_P \delta)$$

where τ_P is the translation operator.

Distribution with compact support

Theorem[7] — Suppose T is a distribution on U with compact support K. There exists a continuous function f defined on U and a multi-index p such that

$$T = \partial^p f,$$

where the derivatives are understood in the sense of distributions. That is, for all test functions ϕ on U,

$$T\phi = (-1)^{|p|} \int_U f(x)(\partial^p \phi)(x)\, dx.$$

Distributions of finite order with support in an open subset

Theorem[7] — Suppose T is a distribution on U with compact support K and let V be an open subset of U containing K. Since every distribution with compact support has finite order, take N to be the order of T and define $P := \{0, 1, \ldots, N+2\}^n$. There exists a family of continuous functions $(f_p)_{p \in P}$ defined on U with support in V such that

$$T = \sum_{p \in P} \partial^p f_p,$$

where the derivatives are understood in the sense of distributions. That is, for all test functions ϕ on U,

$$T\phi = \sum_{p \in P} (-1)^{|p|} \int_U f_p(x)(\partial^p \phi)(x)\, dx.$$

Global structure of distributions

The formal definition of distributions exhibits them as a subspace of a very large space, namely the topological dual of $\mathcal{D}(U)$ (or the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ for tempered distributions). It is not immediately clear from the definition how exotic a distribution might be. To answer this question, it is instructive to see distributions built up from a smaller space, namely the space of continuous functions. Roughly, any distribution is locally a (multiple) derivative of a continuous function. A precise version of this result, given below, holds for distributions of compact support, tempered distributions, and general distributions. Generally speaking, no proper subset of the space of distributions contains all continuous functions and is closed under differentiation. This says that distributions are not particularly exotic objects; they are only as complicated as necessary.

Distributions as sheafs

Theorem[16] — Let T be a distribution on U. There exists a sequence $(T_i)_{i=1}^\infty$ in $\mathcal{D}'(U)$ such that each T_i has compact support and every compact subset $K \subseteq U$ intersects the support of only finitely many T_i, and the sequence of partial sums $(S_j)_{j=1}^\infty$, defined by $S_j := T_1 + \cdots + T_j$, converges in $\mathcal{D}'(U)$ to T; in other words we have:

$$T = \sum_{i=1}^\infty T_i.$$

Recall that a sequence converges in $\mathcal{D}'(U)$ (with its strong dual topology) if and only if it converges pointwise.

Decomposition of distributions as sums of derivatives of continuous functions

By combining the above results, one may express any distribution on U as the sum of a series of distributions with compact support, where each of these distributions can in turn be written as a finite sum of distributional derivatives of continuous functions on U. In other words for arbitrary $T \in \mathcal{D}'(U)$ we can write:

$$T = \sum_{i=1}^{\infty} \sum_{p \in P_i} \partial^p f_{ip},$$

where $P_1, P_2, \ldots$ are finite sets of multi-indices and the functions f_{ip} are continuous.

Theorem[17] — Let T be a distribution on U. For every multi-index p there exists a continuous function g_p on U such that

1. any compact subset K of U intersects the support of only finitely many g_p, and
2. $T = \sum_p \partial^p g_p$.

Moreover, if T has finite order, then one can choose g_p in such a way that only finitely many of them are non-zero.

Note that the infinite sum above is well-defined as a distribution. The value of T for a given $f \in \mathcal{D}(U)$ can be computed using the finitely many g_α that intersect the support of f.

Operations on distributions

Many operations which are defined on smooth functions with compact support can also be defined for distributions. In general, if $A : \mathcal{D}(U) \to \mathcal{D}(U)$ is a linear map which is continuous with respect to the weak topology, then it is possible to extend A to a map $A : \mathcal{D}'(U) \to \mathcal{D}'(U)$ by passing to the limit.[note 6]

Preliminaries: Transpose of a linear operator

Operations on distributions and spaces of distributions are often defined by means of the transpose of a linear operator. This is because the transpose allows for a unified presentation of the many definitions in the theory of distributions and also because its properties are well-known in functional analysis.[18] For instance, the well-known Hermitian adjoint of a linear operator between Hilbert spaces is just the operator's transpose (but with the Riesz representation theorem used to identify each Hilbert space with its continuous dual space). In general the transpose of a continuous linear map $A : X \to Y$ is the linear map

$${}^{t}A : Y' \to X' \qquad \text{defined by} \qquad {}^{t}A(y') := y' \circ A,$$

or equivalently, it is the unique map satisfying $\langle y', A(x)\rangle = \left\langle {}^{t}A(y'), x\right\rangle$ for all $x \in X$ and all $y' \in Y'$ (the prime symbol in y' does not denote a derivative of any kind; it merely indicates that y' is an element of the continuous dual space Y'). Since A is continuous, the transpose ${}^{t}A : Y' \to X'$ is also continuous when both duals are endowed with their respective strong dual topologies; it is also continuous when both duals are endowed with their respective weak* topologies (see the articles polar topology and dual system for more details).

In the context of distributions, the characterization of the transpose can be refined slightly. Let $A : \mathcal{D}(U) \to \mathcal{D}(U)$ be a continuous linear map. Then by definition, the transpose of A is the unique linear operator $A^t : \mathcal{D}'(U) \to \mathcal{D}'(U)$ that satisfies:

$$\langle {}^{t}A(T), \phi\rangle = \langle T, A(\phi)\rangle \quad \text{for all } \phi \in \mathcal{D}(U) \text{ and all } T \in \mathcal{D}'(U).$$

However, since the image of $\mathcal{D}(U)$ is dense in $\mathcal{D}'(U)$, it is sufficient that the above equality hold for all distributions of the form $T = D_\psi$ where $\psi \in \mathcal{D}(U)$. Explicitly, this means that the above condition holds if and only if the condition below holds:

$$\langle {}^{t}A(D_\psi), \phi\rangle = \langle D_\psi, A(\phi)\rangle = \langle \psi, A(\phi)\rangle = \int_U \psi(A\phi)\,dx \quad \text{for all } \phi, \psi \in \mathcal{D}(U).$$

Differential operators

Differentiation of distributions

Let $A : \mathcal{D}(U) \to \mathcal{D}(U)$ be the partial derivative operator $\frac{\partial}{\partial x_k}$. In order to extend A we compute its transpose:

$$\begin{aligned}\langle {}^tA(D_\psi), \phi\rangle &= \int_U \psi(A\phi)\,dx && \text{(See above.)}\\ &= \int_U \psi \frac{\partial \phi}{\partial x_k}\,dx \\ &= -\int_U \phi \frac{\partial \psi}{\partial x_k}\,dx && \text{(integration by parts)}\\ &= -\left\langle \frac{\partial \psi}{\partial x_k}, \phi \right\rangle \\ &= -\langle A\psi, \phi\rangle = \langle -A\psi, \phi\rangle\end{aligned}$$

Therefore ${}^tA = -A$. Therefore the partial derivative of T with respect to the coordinate x_k is defined by the formula

$$\left\langle \frac{\partial T}{\partial x_k}, \phi \right\rangle = -\left\langle T, \frac{\partial \phi}{\partial x_k} \right\rangle \qquad \text{for all } \phi \in \mathcal{D}(U).$$

With this definition, every distribution is infinitely differentiable, and the derivative in the direction x_k is a linear operator on $\mathcal{D}'(U)$.

More generally, if α is an arbitrary multi-index, then the partial derivative $\partial^\alpha T$ of the distribution $T \in \mathcal{D}'(U)$ is defined by

$$\langle \partial^\alpha T, \phi\rangle = (-1)^{|\alpha|} \langle T, \partial^\alpha \phi\rangle \qquad \text{for all } \phi \in \mathcal{D}(U).$$

Differentiation of distributions is a continuous operator on $\mathcal{D}'(U)$; this is an important and desirable property that is not shared by most other notions of differentiation.

If T is a distribution in $\mathbb{R}$ then

$$\lim_{x\to 0} \frac{T - \tau_x T}{x} = T' \in \mathcal{D}'(\mathbb{R}),$$

where T' is the derivative of T and τ_x is translation by x; thus the derivative of T may be viewed as a limit of quotients.[19]

Differential operators acting on smooth functions

A linear differential operator in U with smooth coefficients acts on the space of smooth functions on U. Given such an operator $P := \sum_\alpha c_\alpha \partial^\alpha$, we would like to define a continuous linear map, D_P that extends the action of P on $C^\infty(U)$ to distributions on U. In other words, we would like to define D_P such that the following diagram commutes:

$$\begin{matrix} \mathcal{D}'(U) & \stackrel{D_P}{\longrightarrow} & \mathcal{D}'(U) \\ \uparrow & & \uparrow \\ C^\infty(U) & \stackrel{P}{\longrightarrow} & C^\infty(U) \end{matrix}$$

where the vertical maps are given by assigning $f \in C^\infty(U)$ its canonical distribution $D_f \in \mathcal{D}'(U)$, which is defined by:

$$D_f(\phi) = \langle f, \phi\rangle := \int_U f(x)\phi(x)\,dx \quad \text{for all } \phi \in \mathcal{D}(U).$$

With this notation the diagram commuting is equivalent to:

$$D_{P(f)} = D_P D_f \qquad \text{for all } f \in C^\infty(U).$$

In order to find D_P, the transpose ${}^tP : \mathcal{D}'(U) \to \mathcal{D}'(U)$ of the continuous induced map $P : \mathcal{D}(U) \to \mathcal{D}(U)$ defined by $\phi \mapsto P(\phi)$ is considered in the lemma below. This leads to the following definition of the differential operator on U called *the* ***formal transpose*** *of* P, which will be denoted by P_* to avoid confusion with the transpose map, that is defined by

$$P_* := \sum_\alpha b_\alpha \partial^\alpha \quad \text{where} \quad b_\alpha := \sum_{\beta \geq \alpha} (-1)^{|\beta|} \binom{\beta}{\alpha} \partial^{\beta-\alpha} c_\beta.$$

Lemma — Let P be a linear differential operator with smooth coefficients in U. Then for all $\phi \in \mathcal{D}(U)$ we have

$$\left\langle {}^tP(D_f), \phi \right\rangle = \left\langle D_{P_*(f)}, \phi \right\rangle,$$

which is equivalent to:

$${}^tP(D_f) = D_{P_*(f)}.$$

Proof

As discussed above, for any $\phi \in \mathcal{D}(U)$, the transpose may be calculated by:

$$\begin{aligned}
\left\langle {}^tP(D_f), \phi \right\rangle &= \int_U f(x) P(\phi)(x)\, dx \\
&= \int_U f(x) \left[\sum\nolimits_\alpha c_\alpha(x) (\partial^\alpha \phi)(x)\right] dx \\
&= \sum\nolimits_\alpha \int_U f(x) c_\alpha(x) (\partial^\alpha \phi)(x)\, dx \\
&= \sum\nolimits_\alpha (-1)^{|\alpha|} \int_U \phi(x) (\partial^\alpha (c_\alpha f))(x)\, dx
\end{aligned}$$

For the last line we used integration by parts combined with the fact that ϕ and therefore all the functions $f(x) c_\alpha(x) \partial^\alpha \phi(x)$ have compact support.[note 7] Continuing the calculation above, for all $\phi \in \mathcal{D}(U)$:

$$\begin{aligned}
\left\langle {}^tP(D_f), \phi \right\rangle &= \sum\nolimits_\alpha (-1)^{|\alpha|} \int_U \phi(x) (\partial^\alpha (c_\alpha f))(x)\, dx && \text{As shown above} \\
&= \int_U \phi(x) \sum\nolimits_\alpha (-1)^{|\alpha|} (\partial^\alpha (c_\alpha f))(x)\, dx \\
&= \int_U \phi(x) \sum_\alpha \left[\sum_{\gamma \leq \alpha} \binom{\alpha}{\gamma} (\partial^\gamma c_\alpha)(x) (\partial^{\alpha-\gamma} f)(x)\right] dx && \text{Leibniz rule} \\
&= \int_U \phi(x) \left[\sum_\alpha \sum_{\gamma \leq \alpha} (-1)^{|\alpha|} \binom{\alpha}{\gamma} (\partial^\gamma c_\alpha)(x) (\partial^{\alpha-\gamma} f)(x)\right] dx \\
&= \int_U \phi(x) \left[\sum_\alpha \left[\sum_{\beta \geq \alpha} (-1)^{|\beta|} \binom{\beta}{\alpha} \left(\partial^{\beta-\alpha} c_\beta\right)(x)\right] (\partial^\alpha f)(x)\right] dx && \text{Grouping terms by derivatives of } f \\
&= \int_U \phi(x) \left[\sum\nolimits_\alpha b_\alpha(x) (\partial^\alpha f)(x)\right] dx && b_\alpha := \sum_{\beta \geq \alpha} (-1)^{|\beta|} \binom{\beta}{\alpha} \partial^{\beta-\alpha} c_\beta \\
&= \left\langle \left(\sum\nolimits_\alpha b_\alpha \partial^\alpha\right)(f), \phi \right\rangle
\end{aligned}$$

The Lemma combined with the fact that the formal transpose of the formal transpose is the original differential operator, that is, $P_{**} = P$,[20] enables us to arrive at the correct definition: the formal transpose induces the (continuous) canonical linear operator $P_* : C_c^\infty(U) \to C_c^\infty(U)$ defined by $\phi \mapsto P_*(\phi)$. We claim that the transpose of this map, ${}^tP_* : \mathcal{D}'(U) \to \mathcal{D}'(U)$, can be taken as D_P. To see this, for every $\phi \in \mathcal{D}(U)$, compute its action on a distribution of the form D_f with $f \in C^\infty(U)$:

$$\begin{aligned}\left\langle {}^{t}P_{*}\left(D_{f}\right),\phi\right\rangle &= \left\langle D_{P_{**}(f)},\phi\right\rangle && \text{Using Lemma above with } P_{*} \text{ in place of } P \\ &= \left\langle D_{P(f)},\phi\right\rangle && P_{**}=P\end{aligned}$$

We call the continuous linear operator $D_P := {}^tP_* : \mathcal{D}'(U) \to \mathcal{D}'(U)$ the ***differential operator on distributions extending P***.[20] Its action on an arbitrary distribution S is defined via:

$$D_P(S)(\phi) = S\left(P_*(\phi)\right) \quad \text{for all } \phi \in \mathcal{D}(U).$$

If $(T_i)_{i=1}^{\infty}$ converges to $T \in \mathcal{D}'(U)$ then for every multi-index α, $(\partial^\alpha T_i)_{i=1}^{\infty}$ converges to $\partial^\alpha T \in \mathcal{D}'(U)$.

Multiplication of distributions by smooth functions

A differential operator of order 0 is just multiplication by a smooth function. And conversely, if f is a smooth function then $P := f(x)$ is a differential operator of order 0, whose formal transpose is itself (that is, $P_* = P$). The induced differential operator $D_P : \mathcal{D}'(U) \to \mathcal{D}'(U)$ maps a distribution T to a distribution denoted by $fT := D_P(T)$. We have thus defined the multiplication of a distribution by a smooth function.

We now give an alternative presentation of multiplication by a smooth function. If $m : U \to \mathbb{R}$ is a smooth function and T is a distribution on U, then the product mT is defined by

$$\langle mT, \phi\rangle = \langle T, m\phi\rangle \qquad \text{for all } \phi \in \mathcal{D}(U).$$

This definition coincides with the transpose definition since if $M : \mathcal{D}(U) \to \mathcal{D}(U)$ is the operator of multiplication by the function m (that is, $(M\phi)(x) = m(x)\phi(x)$), then

$$\int_U (M\phi)(x)\psi(x)\,dx = \int_U m(x)\phi(x)\psi(x)\,dx = \int_U \phi(x)m(x)\psi(x)\,dx = \int_U \phi(x)(M\psi)(x)\,dx,$$

so that ${}^tM = M$.

Under multiplication by smooth functions, $\mathcal{D}'(U)$ is a module over the ring $C^\infty(U)$. With this definition of multiplication by a smooth function, the ordinary product rule of calculus remains valid. However, a number of unusual identities also arise. For example, if δ' is the Dirac delta distribution on $\mathbb{R}$, then $m\delta = m(0)\delta$, and if δ' is the derivative of the delta distribution, then

$$m\delta' = m(0)\delta' - m'\delta = m(0)\delta' - m'(0)\delta.$$

The bilinear multiplication map $C^\infty(\mathbb{R}^n) \times \mathcal{D}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$ given by $(f,T) \mapsto fT$ is *not* continuous; it is however, hypocontinuous.[21]

Example. For any distribution T, the product of T with the function that is identically 1 on U is equal to T.

Example. Suppose $(f_i)_{i=1}^{\infty}$ is a sequence of test functions on U that converges to the constant function $1 \in C^\infty(U)$. For any distribution T on U, the sequence $(f_iT)_{i=1}^{\infty}$ converges to $T \in \mathcal{D}'(U)$.[22]

If $(T_i)_{i=1}^{\infty}$ converges to $T \in \mathcal{D}'(U)$ and $(f_i)_{i=1}^{\infty}$ converges to $f \in C^\infty(U)$ then $(f_iT_i)_{i=1}^{\infty}$ converges to $fT \in \mathcal{D}'(U)$.

Problem of multiplying distributions

It is easy to define the product of a distribution with a smooth function, or more generally the product of two distributions whose singular supports are disjoint. With more effort it is possible to define a well-behaved product of several distributions provided their wave front sets at each point are compatible. A limitation of the theory of distributions (and hyperfunctions) is that there is no associative product of two distributions extending the product of a distribution by a smooth function, as has been proved by Laurent Schwartz in the 1950s. For example, if $\operatorname{p.v.} \frac{1}{x}$ is the distribution obtained by the Cauchy principal value

$$\left(\operatorname{p.v.} \frac{1}{x}\right)(\phi) = \lim_{\varepsilon\to 0^+} \int_{|x|\geq\varepsilon} \frac{\phi(x)}{x}\,dx \quad \text{for all } \phi \in \mathcal{S}(\mathbb{R}).$$

If δ is the Dirac delta distribution then

$$(\delta \times x) \times \mathrm{p.v.}\frac{1}{x} = 0$$

but,

$$\delta \times \left(x \times \mathrm{p.v.}\frac{1}{x}\right) = \delta$$

so the product of a distribution by a smooth function (which is always well defined) cannot be extended to an associative product on the space of distributions.

Thus, nonlinear problems cannot be posed in general and thus not solved within distribution theory alone. In the context of quantum field theory, however, solutions can be found. In more than two spacetime dimensions the problem is related to the regularization of divergences. Here Henri Epstein and Vladimir Glaser developed the mathematically rigorous (but extremely technical) *causal perturbation theory*. This does not solve the problem in other situations. Many other interesting theories are non linear, like for example the Navier–Stokes equations of fluid dynamics.

Several not entirely satisfactory theories of algebras of generalized functions have been developed, among which Colombeau's (simplified) algebra is maybe the most popular in use today.

Inspired by Lyons' rough path theory,[23] Martin Hairer proposed a consistent way of multiplying distributions with certain structure (regularity structures[24]), available in many examples from stochastic analysis, notably stochastic partial differential equations. See also Gubinelli–Imkeller–Perkowski (2015) for a related development based on Bony's paraproduct from Fourier analysis.

Composition with a smooth function

Let T be a distribution on $U.$ Let V be an open set in $\mathbb{R}^n,$ and $F : V \to U.$ If F is a submersion, it is possible to define

$$T \circ F \in \mathcal{D}'(V).$$

This is *the* ***composition*** *of the distribution* T *with* F, and is also called *the* ***pullback*** *of* T *along* F, sometimes written

$$F^\sharp : T \mapsto F^\sharp T = T \circ F.$$

The pullback is often denoted $F^*,$ although this notation should not be confused with the use of '*' to denote the adjoint of a linear mapping.

The condition that F be a submersion is equivalent to the requirement that the Jacobian derivative $dF(x)$ of F is a surjective linear map for every $x \in V.$ A necessary (but not sufficient) condition for extending $F^\#$ to distributions is that F be an open mapping.[25] The Inverse function theorem ensures that a submersion satisfies this condition.

If F is a submersion, then $F^\#$ is defined on distributions by finding the transpose map. Uniqueness of this extension is guaranteed since $F^\#$ is a continuous linear operator on $\mathcal{D}(U).$ Existence, however, requires using the change of variables formula, the inverse function theorem (locally) and a partition of unity argument.[26]

In the special case when F is a diffeomorphism from an open subset V of $\mathbb{R}^n$ onto an open subset U of $\mathbb{R}^n$ change of variables under the integral gives:

$$\int_V \phi \circ F(x)\psi(x)\,dx = \int_U \phi(x)\psi\left(F^{-1}(x)\right)\left|\det dF^{-1}(x)\right|\,dx.$$

In this particular case, then, $F^\#$ is defined by the transpose formula:

$$\left\langle F^\sharp T, \phi\right\rangle = \left\langle T, \left|\det d(F^{-1})\right|\phi \circ F^{-1}\right\rangle.$$

Convolution

Under some circumstances, it is possible to define the convolution of a function with a distribution, or even the convolution of two distributions. Recall that if f and g are functions on $\mathbb{R}^n$ then we denote by $f * g$ *the* ***convolution*** *of* f *and* g, defined at $x \in \mathbb{R}^n$ to be the integral

$$(f*g)(x) := \int_{\mathbb{R}^n} f(x-y)g(y)\,dy = \int_{\mathbb{R}^n} f(y)g(x-y)\,dy$$

provided that the integral exists. If $1 \leq p, q, r \leq \infty$ are such that $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$ then for any functions $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$ we have $f * g \in L^r(\mathbb{R}^n)$ and $\|f*g\|_{L^r} \leq \|f\|_{L^p}\|g\|_{L^q}$.[27] If f and g are continuous functions on $\mathbb{R}^n$, at least one of which has compact support, then $\operatorname{supp}(f*g) \subseteq \operatorname{supp}(f) + \operatorname{supp}(g)$ and if $A \subseteq \mathbb{R}^n$ then the value of $f*g$ on A do *not* depend on the values of f outside of the Minkowski sum $A - \operatorname{supp}(g) = \{a - s : a \in A, s \in \operatorname{supp}(g)\}$.[27]

Importantly, if $g \in L^1(\mathbb{R}^n)$ has compact support then for any $0 \leq k \leq \infty$, the convolution map $f \mapsto f * g$ is continuous when considered as the map $C^k(\mathbb{R}^n) \to C^k(\mathbb{R}^n)$ or as the map $C_c^k(\mathbb{R}^n) \to C_c^k(\mathbb{R}^n)$.[27]

Translation and symmetry

Given $a \in \mathbb{R}^n$, the translation operator τ_a sends $f : \mathbb{R}^n \to \mathbb{C}$ to $\tau_a f : \mathbb{R}^n \to \mathbb{C}$, defined by $\tau_a f(y) = f(y-a)$. This can be extended by the transpose to distributions in the following way: given a distribution T, *the* ***translation*** *of* T *by* a is the distribution $\tau_a T : \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}$ defined by $\tau_a T(\phi) := \langle T, \tau_{-a}\phi\rangle$.[28][29]

Given $f : \mathbb{R}^n \to \mathbb{C}$, define the function $\tilde{f} : \mathbb{R}^n \to \mathbb{C}$ by $\tilde{f}(x) := f(-x)$. Given a distribution T, let $\tilde{T} : \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}$ be the distribution defined by $\tilde{T}(\phi) := T\left(\tilde{\phi}\right)$. The operator $T \mapsto \tilde{T}$ is called ***the symmetry with respect to the origin***.[28]

Convolution of a test function with a distribution

Convolution with $f \in \mathcal{D}(\mathbb{R}^n)$ defines a linear map:

$$\begin{aligned} C_f : \,& \mathcal{D}(\mathbb{R}^n) && \to \mathcal{D}(\mathbb{R}^n) \\ & g && \mapsto f * g \end{aligned}$$

which is continuous with respect to the canonical LF space topology on $\mathcal{D}(\mathbb{R}^n)$.

Convolution of f with a distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ can be defined by taking the transpose of C_f relative to the duality pairing of $\mathcal{D}(\mathbb{R}^n)$ with the space $\mathcal{D}'(\mathbb{R}^n)$ of distributions.[30] If $f, g, \phi \in \mathcal{D}(\mathbb{R}^n)$, then by Fubini's theorem

$$\langle C_f g, \phi\rangle = \int_{\mathbb{R}^n} \phi(x) \int_{\mathbb{R}^n} f(x-y)g(y)\,dy\,dx = \left\langle g, C_{\tilde{f}}\,\phi\right\rangle.$$

Extending by continuity, the convolution of f with a distribution T is defined by

$$\langle f * T, \phi\rangle = \left\langle T, \tilde{f} * \phi\right\rangle, \quad \text{for all } \phi \in \mathcal{D}(\mathbb{R}^n).$$

An alternative way to define the convolution of a test function f and a distribution T is to use the translation operator τ_a. The convolution of the compactly supported function f and the distribution T is then the function defined for each $x \in \mathbb{R}^n$ by

$$(f * T)(x) = \left\langle T, \tau_x \tilde{f}\right\rangle.$$

It can be shown that the convolution of a smooth, compactly supported function and a distribution is a smooth function. If the distribution T has compact support then if f is a polynomial (resp. an exponential function, an analytic function, the restriction of an entire analytic function on $\mathbb{C}^n$ to $\mathbb{R}^n$, the restriction of an entire function of exponential type in $\mathbb{C}^n$ to $\mathbb{R}^n$) then the same is true of $T * f$.[28] If the distribution T has compact support as well, then $f * T$ is a compactly supported function, and the Titchmarsh convolution theorem Hörmander (1983, Theorem 4.3.3) implies that:

$$\operatorname{ch}(\operatorname{supp}(f * T)) = \operatorname{ch}(\operatorname{supp}(f)) + \operatorname{ch}(\operatorname{supp}(T))$$

where ch denotes the convex hull and supp denotes the support.

Convolution of a smooth function with a distribution

Let $f \in C^\infty(\mathbb{R}^n)$ and $T \in \mathcal{D}'(\mathbb{R}^n)$ and assume that at least one of f and T has compact support. The ***convolution*** of f and T, denoted by $f * T$ or by $T * f$, is the smooth function:[28]

$$\begin{alignedat}{4} f * T : \,& \mathbb{R}^n && \to \,&& \mathbb{C} \\ & x && \mapsto\,&& \left\langle T, \tau_x \tilde{f} \right\rangle \end{alignedat}$$

satisfying for all $p \in \mathbb{N}^n$:

$$\operatorname{supp}(f * T) \subseteq \operatorname{supp}(f) + \operatorname{supp}(T)$$

$$\text{for all } p \in \mathbb{N}^n: \quad \begin{cases} \partial^p \left\langle T, \tau_x \tilde{f} \right\rangle = \left\langle T, \partial^p \tau_x \tilde{f} \right\rangle \\ \partial^p (T * f) = (\partial^p T) * f = T * (\partial^p f). \end{cases}$$

If T is a distribution then the map $f \mapsto T * f$ is continuous as a map $\mathcal{D}(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ where if in addition T has compact support then it is also continuous as the map $C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ and continuous as the map $\mathcal{D}(\mathbb{R}^n) \to \mathcal{D}(\mathbb{R}^n)$.[28]

If $L : \mathcal{D}(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ is a continuous linear map such that $L \partial^\alpha \phi = \partial^\alpha L \phi$ for all α and all $\phi \in \mathcal{D}(\mathbb{R}^n)$ then there exists a distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ such that $L \phi = T \circ \phi$ for all $\phi \in \mathcal{D}(\mathbb{R}^n)$.[7]

Example.[7] Let H be the Heaviside function on $\mathbb{R}$. For any $\phi \in \mathcal{D}(\mathbb{R})$,

$$(H * \phi)(x) = \int_{-\infty}^x \phi(t) \, dt.$$

Let δ be the Dirac measure at 0 and δ' its derivative as a distribution. Then $\delta' * H = \delta$ and $1 * \delta' = 0$. Importantly, the associative law fails to hold:

$$1 = 1 * \delta = 1 * (\delta' * H) \neq (1 * \delta') * H = 0 * H = 0.$$

Convolution of distributions

It is also possible to define the convolution of two distributions S and T on $\mathbb{R}^n$, provided one of them has compact support. Informally, in order to define $S * T$ where T has compact support, the idea is to extend the definition of the convolution $*$ to a linear operation on distributions so that the associativity formula

$$S * (T * \phi) = (S * T) * \phi$$

continues to hold for all test functions ϕ.[31]

It is also possible to provide a more explicit characterization of the convolution of distributions.[30] Suppose that S and T are distributions and that S has compact support. Then the linear maps

$$\begin{alignedat}{9} \bullet * \tilde{S} : \,& \mathcal{D}(\mathbb{R}^n) && \to \,&& \mathcal{D}(\mathbb{R}^n) && \quad \text{ and } \quad && \bullet * \tilde{T} : \,&& \mathcal{D}(\mathbb{R}^n) && \to \,&& \mathcal{D}(\mathbb{R}^n) \\ & f && \mapsto\,&& f * \tilde{S} && && && f && \mapsto\,&& f * \tilde{T} \end{alignedat}$$

are continuous. The transposes of these maps:

$${}^t\left(\bullet * \tilde{S}\right) : \mathcal{D}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n) \qquad {}^t\left(\bullet * \tilde{T}\right) : \mathcal{E}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$$

are consequently continuous and it can also be shown that[28]

$${}^t\left(\bullet * \tilde{S}\right)(T) = {}^t\left(\bullet * \tilde{T}\right)(S).$$

This common value is called *the **convolution** of S and T* and it is a distribution that is denoted by $S * T$ or $T * S$. It satisfies $\operatorname{supp}(S * T) \subseteq \operatorname{supp}(S) + \operatorname{supp}(T)$.[28] If S and T are two distributions, at least one of which has compact support, then for any $a \in \mathbb{R}^n$, $\tau_a(S * T) = (\tau_a S) * T = S * (\tau_a T)$.[28] If T is a distribution in $\mathbb{R}^n$ and if δ is a Dirac measure then $T * \delta = T = \delta * T$;[28] thus δ is the identity element of the convolution operation. Moreover, if f is a function then $f * \delta' = f' = \delta' * f$ where now the associativity of convolution implies that $f' * g = g' * f$ for all functions f and g.

Suppose that it is T that has compact support. For $\phi \in \mathcal{D}(\mathbb{R}^n)$ consider the function

$$\psi(x) = \langle T, \tau_{-x}\phi \rangle.$$

It can be readily shown that this defines a smooth function of x, which moreover has compact support. The convolution of S and T is defined by

$$\langle S * T, \phi \rangle = \langle S, \psi \rangle.$$

This generalizes the classical notion of convolution of functions and is compatible with differentiation in the following sense: for every multi-index α.

$$\partial^\alpha (S * T) = (\partial^\alpha S) * T = S * (\partial^\alpha T).$$

The convolution of a finite number of distributions, all of which (except possibly one) have compact support, is associative.[28]

This definition of convolution remains valid under less restrictive assumptions about S and T.[32]

The convolution of distributions with compact support induces a continuous bilinear map $\mathcal{E}' \times \mathcal{E}' \to \mathcal{E}'$ defined by $(S, T) \mapsto S * T$, where $\mathcal{E}'$ denotes the space of distributions with compact support.[21] However, the convolution map as a function $\mathcal{E}' \times \mathcal{D}' \to \mathcal{D}'$ is *not* continuous[21] although it is separately continuous.[33] The convolution maps $\mathcal{D}(\mathbb{R}^n) \times \mathcal{D}' \to \mathcal{D}'$ and $\mathcal{D}(\mathbb{R}^n) \times \mathcal{D}' \to \mathcal{D}(\mathbb{R}^n)$ given by $(f, T) \mapsto f * T$ both *fail* to be continuous.[21] Each of these non-continuous maps is, however, separately continuous and hypocontinuous.[21]

Convolution versus multiplication

In general, regularity is required for multiplication products and locality is required for convolution products. It is expressed in the following extension of the Convolution Theorem which guarantees the existence of both convolution and multiplication products. Let $F(\alpha) = f \in \mathcal{O}'_C$ be a rapidly decreasing tempered distribution or, equivalently, $F(f) = \alpha \in \mathcal{O}_M$ be an ordinary (slowly growing, smooth) function within the space of tempered distributions and let F be the normalized (unitary, ordinary frequency) Fourier transform[34] then, according to Schwartz (1951),

$$F(f * g) = F(f) \cdot F(g) \qquad \text{and} \qquad F(\alpha \cdot g) = F(\alpha) * F(g)$$

hold within the space of tempered distributions.[35][36][37] In particular, these equations become the Poisson Summation Formula if $g \equiv ⧢$ is the Dirac Comb.[38] The space of all rapidly decreasing tempered distributions is also called the space of *convolution operators* $\mathcal{O}'_C$ and the space of all ordinary functions within the space of tempered distributions is also called the space of *multiplication operators* $\mathcal{O}_M$. More generally, $F(\mathcal{O}'_C) = \mathcal{O}_M$ and $F(\mathcal{O}_M) = \mathcal{O}'_C$. [39][40] A particular case is the Paley-Wiener-Schwartz Theorem which states that $F(\mathcal{E}') = \mathrm{PW}$ and $F(\mathrm{PW}) = \mathcal{E}'$. This is because $\mathcal{E}' \subseteq \mathcal{O}'_C$ and $\mathrm{PW} \subseteq \mathcal{O}_M$. In other words, compactly supported tempered distributions $\mathcal{E}'$ belong to the space of *convolution operators* $\mathcal{O}'_C$ and Paley-Wiener functions PW, better known as bandlimited functions, belong to the space of *multiplication operators* $\mathcal{O}_M$.[41]

For example, let $g \equiv ⧢ \in \mathcal{S}'$ be the Dirac comb and $f \equiv \delta \in \mathcal{E}'$ be the Dirac delta then $\alpha \equiv 1 \in \mathrm{PW}$ is the function that is constantly one and both equations yield the Dirac comb identity. Another example is to let g be the Dirac comb and $f \equiv \mathbf{rect} \in \mathcal{E}'$ be the rectangular function then $\alpha \equiv \mathbf{sinc} \in \mathrm{PW}$ is the sinc function and both equations yield the Classical Sampling Theorem for suitable $\mathbf{rect}$ functions. More generally, if g is the Dirac comb and $f \in \mathcal{S} \subseteq \mathcal{O}'_C \cap \mathcal{O}_M$ is a smooth window function (Schwartz function), for example, the Gaussian, then $\alpha \in \mathcal{S}$ is another smooth window function (Schwartz function). They are known as mollifiers, especially in partial differential equations theory, or as regularizers in physics because they allow turning generalized functions into regular functions.

Tensor products of distributions

Let $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^n$ be open sets. Assume all vector spaces to be over the field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For $f \in \mathcal{D}(U \times V)$ define for every $u \in U$ and every $v \in V$ the following functions:

$$\begin{aligned} f_u :\, & V \to \mathbb{F} \\ & y \mapsto f(u, y) \end{aligned} \qquad \text{and} \qquad \begin{aligned} f^v :\, & U \to \mathbb{F} \\ & x \mapsto f(x, v) \end{aligned}$$

Given $S \in \mathcal{D}'(U)$ and $T \in \mathcal{D}'(V)$, define the following functions:

$$\begin{aligned} \langle S, f^{\bullet}\rangle :\ & V \to \mathbb{F} \\ & v \mapsto \langle S, f^{v}\rangle \end{aligned} \qquad \text{and} \qquad \begin{aligned} \langle T, f_{\bullet}\rangle :\ & U \to \mathbb{F} \\ & u \mapsto \langle T, f_{u}\rangle \end{aligned}$$

where $\langle T, f_{\bullet}\rangle \in \mathcal{D}(U)$ and $\langle S, f^{\bullet}\rangle \in \mathcal{D}(V)$. These definitions associate every $S \in \mathcal{D}'(U)$ and $T \in \mathcal{D}'(V)$ with the (respective) continuous linear map:

$$\begin{aligned} \mathcal{D}(U \times V) & \to \mathcal{D}(V) \\ f & \mapsto \langle S, f^{\bullet}\rangle \end{aligned} \qquad \text{and} \qquad \begin{aligned} \mathcal{D}(U \times V) & \to \mathcal{D}(U) \\ f & \mapsto \langle T, f_{\bullet}\rangle \end{aligned}$$

Moreover if either S (resp. T) has compact support then it also induces a continuous linear map of $C^{\infty}(U \times V) \to C^{\infty}(V)$ (resp. $C^{\infty}(U \times V) \to C^{\infty}(U)$).[42]

> **Fubini's theorem for distributions**[42] — Let $S \in \mathcal{D}'(U)$ and $T \in \mathcal{D}'(V)$. If $f \in \mathcal{D}(U \times V)$ then
>
> $$\langle S, \langle T, f_{\bullet}\rangle\rangle = \langle T, \langle S, f^{\bullet}\rangle\rangle.$$

*The **tensor product** of $S \in \mathcal{D}'(U)$ and $T \in \mathcal{D}'(V)$*, denoted by $S \otimes T$ or $T \otimes S$, is the distribution in $U \times V$ defined by:[42]

$$(S \otimes T)(f) := \langle S, \langle T, f_{\bullet}\rangle\rangle = \langle T, \langle S, f^{\bullet}\rangle\rangle.$$

Spaces of distributions

For all $0 < k < \infty$ and all $1 < p < \infty$, every one of the following canonical injections is continuous and has an image (also called the range) that is a dense subset of its codomain:

$$\begin{matrix} C_c^{\infty}(U) & \to & C_c^{k}(U) & \to & C_c^{0}(U) & \to & L_c^{\infty}(U) & \to & L_c^{p}(U) & \to & L_c^{1}(U) \\ \downarrow & & \downarrow & & \downarrow & & & & & & \\ C^{\infty}(U) & \to & C^{k}(U) & \to & C^{0}(U) & & & & & & \end{matrix}$$

where the topologies on $L_c^q(U)$ ($1 \leq q \leq \infty$) are defined as direct limits of the spaces $L_c^q(K)$ in a manner analogous to how the topologies on $C_c^k(U)$ were defined (so in particular, they are not the usual norm topologies). The range of each of the maps above (and of any composition of the maps above) is dense in its codomain.[43]

Suppose that X is one of the spaces $C_c^k(U)$ (for $k \in \{0, 1, \ldots, \infty\}$) or $L_c^p(U)$ (for $1 \leq p \leq \infty$) or $L^p(U)$ (for $1 \leq p < \infty$). Because the canonical injection $\operatorname{In}_X : C_c^{\infty}(U) \to X$ is a continuous injection whose image is dense in the codomain, this map's transpose ${}^{t}\operatorname{In}_X : X'_b \to \mathcal{D}'(U) = (C_c^{\infty}(U))'_b$ is a continuous injection. This injective transpose map thus allows the continuous dual space X' of X to be identified with a certain vector subspace of the space $\mathcal{D}'(U)$ of all distributions (specifically, it is identified with the image of this transpose map). This transpose map is continuous but it is *not* necessarily a topological embedding. A linear subspace of $\mathcal{D}'(U)$ carrying a locally convex topology that is finer than the subspace topology induced by $\mathcal{D}'(U) = (C_c^{\infty}(U))'_b$ is called ***a space of distributions***.[44] Almost all of the spaces of distributions mentioned in this article arise in this way (for example, tempered distribution, restrictions, distributions of order $\leq$ some integer, distributions induced by a positive Radon measure, distributions induced by an L^p-function, etc.) and any representation theorem about the continuous dual space of X may, through the transpose ${}^{t}\operatorname{In}_X : X'_b \to \mathcal{D}'(U)$, be transferred directly to elements of the space $\operatorname{Im}({}^{t}\operatorname{In}_X)$.

Radon measures

The inclusion map $\operatorname{In} : C_c^{\infty}(U) \to C_c^{0}(U)$ is a continuous injection whose image is dense in its codomain, so the transpose ${}^{t}\operatorname{In} : (C_c^{0}(U))'_b \to \mathcal{D}'(U) = (C_c^{\infty}(U))'_b$ is also a continuous injection.

Note that the continuous dual space $(C_c^{0}(U))'_b$ can be identified as the space of Radon measures, where there is a one-to-one correspondence between the continuous linear functionals $T \in (C_c^{0}(U))'_b$ and integral with respect to a Radon measure; that is,

- if $T \in (C_c^{0}(U))'_b$ then there exists a Radon measure μ on U such that for all $f \in C_c^{0}(U), T(f) = \int_U f\, d\mu$, and
- if μ is a Radon measure on U then the linear functional on $C_c^{0}(U)$ defined by sending $f \in C_c^{0}(U)$ to $\int_U f\, d\mu$ is continuous.

Through the injection ${}^t\,\mathbf{In} : (C_c^0(U))'_b \to \mathcal{D}'(U)$, every Radon measure becomes a distribution on U. If f is a locally integrable function on U then the distribution $\phi \mapsto \int_U f(x)\phi(x)\,dx$ is a Radon measure; so Radon measures form a large and important space of distributions.

The following is the theorem of the structure of distributions of Radon measures, which shows that every Radon measure can be written as a sum of derivatives of locally L^∞ functions on U:

Theorem.[45] — Suppose $T \in \mathcal{D}'(U)$ is a Radon measure, where $U \subseteq \mathbb{R}^n$, let $V \subseteq U$ be a neighborhood of the support of T, and let $I = \{p \in \mathbb{N}^n : |p| \leq n\}$. There exists a family $f = (f_p)_{p \in I}$ of locally L^∞ functions on U such that $\operatorname{supp} f_p \subseteq V$ for every $p \in I$, and

$$T = \sum_{p \in I} \partial^p f_p.$$

Furthermore, T is also equal to a finite sum of derivatives of continuous functions on U, where each derivative has order $\leq 2n$.

Positive Radon measures

A linear function T on a space of functions is called ***positive*** if whenever a function f that belongs to the domain of T is non-negative (that is, f is real-valued and $f \geq 0$) then $T(f) \geq 0$. One may show that every positive linear functional on $C_c^0(U)$ is necessarily continuous (that is, necessarily a Radon measure).[46] Lebesgue measure is an example of a positive Radon measure.

Locally integrable functions as distributions

One particularly important class of Radon measures are those that are induced locally integrable functions. The function $f : U \to \mathbb{R}$ is called ***locally integrable*** if it is Lebesgue integrable over every compact subset K of U.[note 8] This is a large class of functions which includes all continuous functions and all Lp space L^p functions. The topology on $\mathcal{D}(U)$ is defined in such a fashion that any locally integrable function f yields a continuous linear functional on $\mathcal{D}(U)$ – that is, an element of $\mathcal{D}'(U)$ – denoted here by T_f, whose value on the test function ϕ is given by the Lebesgue integral:

$$\langle T_f, \phi \rangle = \int_U f\phi\,dx.$$

Conventionally, one abuses notation by identifying T_f with f, provided no confusion can arise, and thus the pairing between T_f and ϕ is often written

$$\langle f, \phi \rangle = \langle T_f, \phi \rangle.$$

If f and g are two locally integrable functions, then the associated distributions T_f and T_g are equal to the same element of $\mathcal{D}'(U)$ if and only if f and g are equal almost everywhere (see, for instance, Hörmander (1983, Theorem 1.2.5)). In a similar manner, every Radon measure μ on U defines an element of $\mathcal{D}'(U)$ whose value on the test function ϕ is $\int \phi\,d\mu$. As above, it is conventional to abuse notation and write the pairing between a Radon measure μ and a test function ϕ as $\langle \mu, \phi \rangle$. Conversely, as shown in a theorem by Schwartz (similar to the Riesz representation theorem), every distribution which is non-negative on non-negative functions is of this form for some (positive) Radon measure.

Test functions as distributions

The test functions are themselves locally integrable, and so define distributions. The space of test functions $C_c^\infty(U)$ is sequentially dense in $\mathcal{D}'(U)$ with respect to the strong topology on $\mathcal{D}'(U)$.[47] This means that for any $T \in \mathcal{D}'(U)$, there is a sequence of test functions, $(\phi_i)_{i=1}^\infty$, that converges to $T \in \mathcal{D}'(U)$ (in its strong dual topology) when considered as a sequence of distributions. Or equivalently,

$$\langle \phi_i, \psi \rangle \to \langle T, \psi \rangle \qquad \text{for all } \psi \in \mathcal{D}(U).$$

Distributions with compact support

The inclusion map $\operatorname{In} : C_c^\infty(U) \to C^\infty(U)$ is a continuous injection whose image is dense in its codomain, so the transpose map ${}^t\operatorname{In} : (C^\infty(U))'_b \to \mathcal{D}'(U) = (C_c^\infty(U))'_b$ is also a continuous injection. Thus the image of the transpose, denoted by $\mathcal{E}'(U)$, forms a space of distributions.[13]

The elements of $\mathcal{E}'(U) = (C^\infty(U))'_b$ can be identified as the space of distributions with compact support.[13] Explicitly, if T is a distribution on U then the following are equivalent,

- $T \in \mathcal{E}'(U)$.
- The support of T is compact.
- The restriction of T to $C_c^\infty(U)$, when that space is equipped with the subspace topology inherited from $C^\infty(U)$ (a coarser topology than the canonical LF topology), is continuous.[13]
- There is a compact subset K of U such that for every test function ϕ whose support is completely outside of K, we have $T(\phi) = 0$.

Compactly supported distributions define continuous linear functionals on the space $C^\infty(U)$; recall that the topology on $C^\infty(U)$ is defined such that a sequence of test functions ϕ_k converges to 0 if and only if all derivatives of ϕ_k converge uniformly to 0 on every compact subset of U. Conversely, it can be shown that every continuous linear functional on this space defines a distribution of compact support. Thus compactly supported distributions can be identified with those distributions that can be extended from $C_c^\infty(U)$ to $C^\infty(U)$.

Distributions of finite order

Let $k \in \mathbb{N}$. The inclusion map $\operatorname{In} : C_c^\infty(U) \to C_c^k(U)$ is a continuous injection whose image is dense in its codomain, so the transpose ${}^t\operatorname{In} : (C_c^k(U))'_b \to \mathcal{D}'(U) = (C_c^\infty(U))'_b$ is also a continuous injection. Consequently, the image of ${}^t\operatorname{In}$, denoted by $\mathcal{D}'^k(U)$, forms a space of distributions. The elements of $\mathcal{D}'^k(U)$ are ***the distributions of order*** $\leq k$.[16] The distributions of order ≤ 0, which are also called ***distributions of order 0***, are exactly the distributions that are Radon measures (described above).

For $0 \neq k \in \mathbb{N}$, a ***distribution of order k*** is a distribution of order $\leq k$ that is not a distribution of order $\leq k - 1$[16]

A distribution is said to be of ***finite order*** if there is some integer k such that it is a distribution of order $\leq k$, and the set of distributions of finite order is denoted by $\mathcal{D}'^F(U)$. Note that if $k \leq l$ then $\mathcal{D}'^k(U) \subseteq \mathcal{D}'^l(U)$ so that $\mathcal{D}'^F(U) := \bigcup_{n=0}^\infty \mathcal{D}'^n(U)$ is a vector subspace of $\mathcal{D}'(U)$ and furthermore, if and only if $\mathcal{D}'^F(U) = \mathcal{D}'(U)$.[16]

Structure of distributions of finite order

Every distribution with compact support in U is a distribution of finite order.[16] Indeed, every distribution in U is *locally* a distribution of finite order, in the following sense:[16] If V is an open and relatively compact subset of U and if ρ_{VU} is the restriction mapping from U to V, then the image of $\mathcal{D}'(U)$ under ρ_{VU} is contained in $\mathcal{D}'^F(V)$.

The following is the theorem of the structure of distributions of finite order, which shows that every distribution of finite order can be written as a sum of derivatives of Radon measures:

Theorem[16] — Suppose $T \in \mathcal{D}'(U)$ has finite order and $I = \{p \in \mathbb{N}^n : |p| \leq k\}$. Given any open subset V of U containing the support of T, there is a family of Radon measures in U, $(\mu_p)_{p \in I}$, such that for very $p \in I, \operatorname{supp}(\mu_p) \subseteq V$ and

$$T = \sum_{|p| \leq k} \partial^p \mu_p.$$

Example. (Distributions of infinite order) Let $U := (0, \infty)$ and for every test function f, let

$$Sf := \sum_{m=1}^\infty (\partial^m f)\left(\frac{1}{m}\right).$$

Then S is a distribution of infinite order on U. Moreover, S can not be extended to a distribution on $\mathbb{R}$; that is, there exists no distribution T on $\mathbb{R}$ such that the restriction of T to U is equal to S.[48]

Tempered distributions and Fourier transform

Defined below are the ***tempered distributions***, which form a subspace of $\mathcal{D}'(\mathbb{R}^n)$, the space of distributions on $\mathbb{R}^n$. This is a proper subspace: while every tempered distribution is a distribution and an element of $\mathcal{D}'(\mathbb{R}^n)$, the converse is not true. Tempered distributions are useful if one studies the Fourier transform since all tempered distributions have a Fourier transform, which is not true for an arbitrary distribution in $\mathcal{D}'(\mathbb{R}^n)$.

Schwartz space

The Schwartz space, $\mathcal{S}(\mathbb{R}^n)$, is the space of all smooth functions that are rapidly decreasing at infinity along with all partial derivatives. Thus $\phi : \mathbb{R}^n \to \mathbb{R}$ is in the Schwartz space provided that any derivative of ϕ, multiplied with any power of $|x|$, converges to 0 as $|x| \to \infty$. These functions form a complete TVS with a suitably defined family of seminorms. More precisely, for any multi-indices α and β define:

$$p_{\alpha,\beta}(\phi) \;=\; \sup_{x\in\mathbb{R}^n} \left|x^\alpha \partial^\beta \phi(x)\right|.$$

Then ϕ is in the Schwartz space if all the values satisfy:

$$p_{\alpha,\beta}(\phi) < \infty.$$

The family of seminorms $p_{\alpha,\beta}$ defines a locally convex topology on the Schwartz space. For $n = 1$, the seminorms are, in fact, norms on the Schwartz space. One can also use the following family of seminorms to define the topology:[49]

$$|f|_{m,k} = \sup_{|p|\le m} \left(\sup_{x\in\mathbb{R}^n} \left\{ (1+|x|)^k \left|(\partial^\alpha f)(x)\right| \right\} \right), \qquad k, m \in \mathbb{N}.$$

Otherwise, one can define a norm on $\mathcal{S}(\mathbb{R}^n)$ via

$$\|\phi\|_k \;=\; \max_{|\alpha|+|\beta|\le k} \sup_{x\in\mathbb{R}^n} \left|x^\alpha \partial^\beta \phi(x)\right|, \qquad k \ge 1.$$

The Schwartz space is a Fréchet space (that is, a complete metrizable locally convex space). Because the Fourier transform changes ∂^α into multiplication by x^α and vice versa, this symmetry implies that the Fourier transform of a Schwartz function is also a Schwartz function.

A sequence $\{f_i\}$ in $\mathcal{S}(\mathbb{R}^n)$ converges to 0 in $\mathcal{S}(\mathbb{R}^n)$ if and only if the functions $(1+|x|)^k (\partial^p f_i)(x)$ converge to 0 uniformly in the whole of $\mathbb{R}^n$, which implies that such a sequence must converge to zero in $C^\infty(\mathbb{R}^n)$.[49]

$\mathcal{D}(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$. The subset of all analytic Schwartz functions is dense in $\mathcal{S}(\mathbb{R}^n)$ as well.[50]

The Schwartz space is nuclear and the tensor product of two maps induces a canonical surjective TVS-isomorphisms

$$\mathcal{S}(\mathbb{R}^m)\ \widehat{\otimes}\ \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^{m+n}),$$

where $\widehat{\otimes}$ represents the completion of the injective tensor product (which in this case is the identical to the completion of the projective tensor product).[51]

Tempered distributions

The inclusion map $\operatorname{In} : \mathcal{D}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a continuous injection whose image is dense in its codomain, so the transpose ${}^t\operatorname{In} : (\mathcal{S}(\mathbb{R}^n))'_b \to \mathcal{D}'(\mathbb{R}^n)$ is also a continuous injection. Thus, the image of the transpose map, denoted by $\mathcal{S}'(\mathbb{R}^n)$, forms a space of distributions.

The space $\mathcal{S}'(\mathbb{R}^n)$ is called the space of *tempered distributions*. It is the continuous dual space of the Schwartz space. Equivalently, a distribution T is a tempered distribution if and only if

$$\Big(\text{for all } \alpha,\beta\in\mathbb{N}^n:\ \lim_{m\to\infty} p_{\alpha,\beta}(\phi_m)=0\Big) \Longrightarrow \lim_{m\to\infty} T(\phi_m)=0.$$

The derivative of a tempered distribution is again a tempered distribution. Tempered distributions generalize the bounded (or slow-growing) locally integrable functions; all distributions with compact support and all square-integrable functions are tempered distributions. More generally, all functions that are products of polynomials with elements of Lp space $L^p(\mathbb{R}^n)$ for $p \geq 1$ are tempered distributions.

The *tempered distributions* can also be characterized as *slowly growing*, meaning that each derivative of T grows at most as fast as some polynomial. This characterization is dual to the *rapidly falling* behaviour of the derivatives of a function in the Schwartz space, where each derivative of ϕ decays faster than every inverse power of $|x|$. An example of a rapidly falling function is $|x|^n \exp(-\lambda|x|^\beta)$ for any positive n, λ, β.

Fourier transform

To study the Fourier transform, it is best to consider complex-valued test functions and complex-linear distributions. The ordinary continuous Fourier transform $F : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a TVS-automorphism of the Schwartz space, and the ***Fourier transform*** is defined to be its transpose ${}^tF : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$, which (abusing notation) will again be denoted by F. So the Fourier transform of the tempered distribution T is defined by $(FT)(\psi) = T(F\psi)$ for every Schwartz function ψ. FT is thus again a tempered distribution. The Fourier transform is a TVS isomorphism from the space of tempered distributions onto itself. This operation is compatible with differentiation in the sense that

$$F\frac{dT}{dx} = ixFT$$

and also with convolution: if T is a tempered distribution and ψ is a *slowly increasing* smooth function on $\mathbb{R}^n$, ψT is again a tempered distribution and

$$F(\psi T) = F\psi * FT$$

is the convolution of FT and $F\psi$. In particular, the Fourier transform of the constant function equal to 1 is the δ distribution.

Expressing tempered distributions as sums of derivatives

If $T \in \mathcal{S}'(\mathbb{R}^n)$ is a tempered distribution, then there exists a constant $C > 0$, and positive integers M and N such that for all Schwartz functions $\phi \in \mathcal{S}(\mathbb{R}^n)$

$$\langle T, \phi\rangle \leq C \sum\nolimits_{|\alpha|\leq N, |\beta|\leq M} \sup_{x\in\mathbb{R}^n} \left|x^\alpha \partial^\beta \phi(x)\right| = C \sum\nolimits_{|\alpha|\leq N, |\beta|\leq M} p_{\alpha,\beta}(\phi).$$

This estimate along with some techniques from functional analysis can be used to show that there is a continuous slowly increasing function F and a multi-index α such that

$$T = \partial^\alpha F.$$

Restriction of distributions to compact sets

If $T \in \mathcal{D}'(\mathbb{R}^n)$, then for any compact set $K \subseteq \mathbb{R}^n$, there exists a continuous function Fcompactly supported in $\mathbb{R}^n$ (possibly on a larger set than K itself) and a multi-index α such that $T = \partial^\alpha F$ on $C_c^\infty(K)$.

Using holomorphic functions as test functions

The success of the theory led to investigation of the idea of hyperfunction, in which spaces of holomorphic functions are used as test functions. A refined theory has been developed, in particular Mikio Sato's algebraic analysis, using sheaf theory and several complex variables. This extends the range of symbolic methods that can be made into rigorous mathematics, for example Feynman integrals.

See also

- Colombeau algebra
- Current (mathematics) — Distributions on spaces of differential forms
- Distribution (number theory)

- Distribution on a linear algebraic group — Linear function satisfying a support condition
- Gelfand triple
- Gelfand–Shilov space
- Generalized function
- Homogeneous distribution
- Hyperfunction — Type of generalized function
- Laplacian of the indicator — Limit of sequence of smooth functions
- Limit of a distribution
- Linear form — Linear map from a vector space to its field of scalars
- Malgrange–Ehrenpreis theorem
- Pseudodifferential operator
- Riesz representation theorem
- Vague topology
- Weak solution

Notes

1. Note that i being an integer implies $i \neq \infty$. This is sometimes expressed as $0 \leq i < k+1$. Since $\infty + 1 = \infty$, the inequality "$0 \leq i < k+1$" means: $0 \leq i < \infty$ if $k = \infty$, while if $k \neq \infty$ then it means $0 \leq i \leq k$.
2. The image of the compact set K under a continuous $\mathbb{R}$-valued map (for example, $x \mapsto |\partial^p f(x)|$ for $x \in U$) is itself a compact, and thus bounded, subset of $\mathbb{R}$. If $K \neq \varnothing$ then this implies that each of the functions defined above is $\mathbb{R}$-valued (that is, none of the supremums above are ever equal to ∞).
3. Even though the topology of $C_c^\infty(U)$ is not metrizable, a linear functional on $C_c^\infty(U)$ is continuous if and only if it is sequentially continuous.
4. A ***null sequence*** is a sequence that converges to the origin.
5. If $\mathcal{P}$ is also directed under the usual function comparison then we can take the finite collection to consist of a single element.
6. This approach works for non-linear mappings as well, provided they are assumed to be uniformly continuous.
7. For example, let $U = \mathbb{R}$ and take P to be the ordinary derivative for functions of one real variable and assume the support of ϕ to be contained in the finite interval (a, b), then since $\operatorname{supp}(\phi) \subseteq (a, b)$

$$\begin{aligned}\int_{\mathbb{R}} \phi'(x) f(x)\,dx &= \int_a^b \phi'(x) f(x)\,dx \\ &= \phi(x) f(x)\big|_a^b - \int_a^b f'(x)\phi(x)\,dx \\ &= \phi(b) f(b) - \phi(a) f(a) - \int_a^b f'(x)\phi(x)\,dx \\ &= \int_a^b f'(x)\phi(x)\,dx\end{aligned}$$

   where the last equality is because $\phi(a) = \phi(b) = 0$.
8. For more information on such class of functions, see the entry on locally integrable functions.

References

1. Trèves 2006, pp. 222–223.
2. Grubb 2009, p. 14
3. Trèves 2006, pp. 85–89.
4. Trèves 2006, pp. 142–149.
5. Trèves 2006, pp. 356–358.
6. Trèves 2006, pp. 131–134.
7. Rudin 1991, pp. 149–181.
8. Trèves 2006, pp. 526–534.
9. Trèves 2006, p. 357.
10. See for example Grubb 2009, p. 14.
11. Trèves 2006, pp. 245–247.
12. Trèves 2006, pp. 253–255.
13. Trèves 2006, pp. 255–257.
14. Trèves 2006, pp. 264–266.

15. Rudin 1991, p. 165.
16. Trèves 2006, pp. 258–264.
17. Rudin 1991, pp. 169–170.
18. Strichartz 1994, §2.3; Trèves 2006.
19. Rudin 1991, p. 180.
20. Trèves 2006, pp. 247–252.
21. Trèves 2006, p. 423.
22. Trèves 2006, p. 261.
23. Lyons, T. (1998). "Differential equations driven by rough signals" (https://doi.org/10.4171%2FRMI%2F240). *Revista Matemática Iberoamericana*: 215–310. doi:10.4171/RMI/240 (https://doi.org/10.4171%2FRMI%2F240).
24. Hairer, Martin (2014). "A theory of regularity structures". *Inventiones Mathematicae*. **198** (2): 269–504. arXiv:1303.5113 (https://arxiv.org/abs/1303.5113). Bibcode:2014InMat.198..269H (https://ui.adsabs.harvard.edu/abs/2014InMat.198..269H). doi:10.1007/s00222-014-0505-4 (https://doi.org/10.1007%2Fs00222-014-0505-4).
25. See for example Hörmander 1983, Theorem 6.1.1.
26. See Hörmander 1983, Theorem 6.1.2.
27. Trèves 2006, pp. 278–283.
28. Trèves 2006, pp. 284–297.
29. See for example Rudin 1991, §6.29.
30. Trèves 2006, Chapter 27.
31. Hörmander 1983, §IV.2 proves the uniqueness of such an extension.
32. See for instance Gel'fand & Shilov 1966–1968, v. 1, pp. 103–104 and Benedetto 1997, Definition 2.5.8.
33. Trèves 2006, p. 294.
34. Folland, G.B. (1989). *Harmonic Analysis in Phase Space*. Princeton, NJ: Princeton University Press.
35. Horváth, John (1966). *Topological Vector Spaces and Distributions*. Reading, MA: Addison-Wesley Publishing Company.
36. Barros-Neto, José (1973). *An Introduction to the Theory of Distributions*. New York, NY: Dekker.
37. Petersen, Bent E. (1983). *Introduction to the Fourier Transform and Pseudo-Differential Operators*. Boston, MA: Pitman Publishing.
38. Woodward, P.M. (1953). *Probability and Information Theory with Applications to Radar*. Oxford, UK: Pergamon Press.
39. Trèves 2006, pp. 318–319.
40. Friedlander, F.G.; Joshi, M.S. (1998). *Introduction to the Theory of Distributions*. Cambridge, UK: Cambridge University Press.
41. Schwartz 1951.
42. Trèves 2006, pp. 416–419.
43. Trèves 2006, pp. 150–160.
44. Trèves 2006, pp. 240–252.
45. Trèves 2006, pp. 262–264.
46. Trèves 2006, p. 218.
47. Trèves 2006, pp. 300–304.
48. Rudin 1991, pp. 177–181.
49. Trèves 2006, pp. 92–94.
50. Trèves 2006, pp. 160.
51. Trèves 2006, p. 531.

Bibliography

- Barros-Neto, José (1973). *An Introduction to the Theory of Distributions*. New York, NY: Dekker.
- Benedetto, J.J. (1997), *Harmonic Analysis and Applications*, CRC Press.
- Folland, G.B. (1989). *Harmonic Analysis in Phase Space*. Princeton, NJ: Princeton University Press.
- Friedlander, F.G.; Joshi, M.S. (1998). *Introduction to the Theory of Distributions*. Cambridge, UK: Cambridge University Press..
- Gårding, L. (1997), *Some Points of Analysis and their History*, American Mathematical Society.
- Gel'fand, I.M.; Shilov, G.E. (1966–1968), *Generalized functions*, vol. 1–5, Academic Press.
- Grubb, G. (2009), *Distributions and Operators*, Springer.
- Hörmander, L. (1983), *The analysis of linear partial differential operators I*, Grundl. Math. Wissenschaft., vol. 256, Springer, doi:10.1007/978-3-642-96750-4 (https://doi.org/10.1007%2F978-3-642-96750-4), ISBN 3-540-12104-8, MR 0717035 (https://www.ams.org/mathscinet-getitem?mr=0717035).
- Horváth, John (1966). *Topological Vector Spaces and Distributions*. Addison-Wesley series in mathematics. Vol. 1. Reading, MA: Addison-Wesley Publishing Company. ISBN 978-0201029857.
- Kolmogorov, Andrey; Fomin, Sergei V. (1957). *Elements of the Theory of Functions and Functional Analysis*. Dover Books on Mathematics. New York: Dover Books. ISBN 978-1-61427-304-2. OCLC 912495626 (https://www.worldcat.o

rg/oclc/912495626).

- Narici, Lawrence; Beckenstein, Edward (2011). *Topological Vector Spaces*. Pure and applied mathematics (Second ed.). Boca Raton, FL: CRC Press. ISBN 978-1584888666. OCLC 144216834 (https://www.worldcat.org/oclc/144216834).
- Petersen, Bent E. (1983). *Introduction to the Fourier Transform and Pseudo-Differential Operators*. Boston, MA: Pitman Publishing..
- Rudin, Walter (1991). *Functional Analysis* (https://archive.org/details/functionalanalys00rudi). International Series in Pure and Applied Mathematics. Vol. 8 (Second ed.). New York, NY: McGraw-Hill Science/Engineering/Math. ISBN 978-0-07-054236-5. OCLC 21163277 (https://www.worldcat.org/oclc/21163277).
- Schaefer, Helmut H.; Wolff, Manfred P. (1999). *Topological Vector Spaces*. GTM. Vol. 8 (Second ed.). New York, NY: Springer New York Imprint Springer. ISBN 978-1-4612-7155-0. OCLC 840278135 (https://www.worldcat.org/oclc/840278135).
- Schwartz, Laurent (1954), "Sur l'impossibilité de la multiplications des distributions", *C. R. Acad. Sci. Paris*, **239**: 847—848.
- Schwartz, Laurent (1951), *Théorie des distributions*, vol. 1—2, Hermann.
- Sobolev, S.L. (1936), "Méthode nouvelle à résoudre le problème de Cauchy pour les équations linéaires hyperboliques normales" (http://mi.mathnet.ru/msb5358), *Mat. Sbornik*, **1**: 39—72.
- Stein, Elias; Weiss, Guido (1971), *Introduction to Fourier Analysis on Euclidean Spaces* (https://archive.org/details/introductiontofo0000stei), Princeton University Press, ISBN 0-691-08078-X.
- Strichartz, R. (1994), *A Guide to Distribution Theory and Fourier Transforms*, CRC Press, ISBN 0-8493-8273-4.
- Trèves, François (2006) [1967]. *Topological Vector Spaces, Distributions and Kernels*. Mineola, N.Y.: Dover Publications. ISBN 978-0-486-45352-1. OCLC 853623322 (https://www.worldcat.org/oclc/853623322).
- Woodward, P.M. (1953). *Probability and Information Theory with Applications to Radar*. Oxford, UK: Pergamon Press.

Further reading

- M. J. Lighthill (1959). *Introduction to Fourier Analysis and Generalised Functions*. Cambridge University Press. ISBN 0-521-09128-4 (requires very little knowledge of analysis; defines distributions as limits of sequences of functions under integrals)
- V.S. Vladimirov (2002). *Methods of the theory of generalized functions*. Taylor & Francis. ISBN 0-415-27356-0
- Vladimirov, V.S. (2001) [1994], "Generalized function" (https://www.encyclopediaofmath.org/index.php?title=Generalized_function), *Encyclopedia of Mathematics*, EMS Press.
- Vladimirov, V.S. (2001) [1994], "Generalized functions, space of" (https://www.encyclopediaofmath.org/index.php?title=Generalized_functions,_space_of), *Encyclopedia of Mathematics*, EMS Press.
- Vladimirov, V.S. (2001) [1994], "Generalized function, derivative of a" (https://www.encyclopediaofmath.org/index.php?title=Generalized_function,_derivative_of_a), *Encyclopedia of Mathematics*, EMS Press.
- Vladimirov, V.S. (2001) [1994], "Generalized functions, product of" (https://www.encyclopediaofmath.org/index.php?title=Generalized_functions,_product_of), *Encyclopedia of Mathematics*, EMS Press.
- Oberguggenberger, Michael (2001) [1994], "Generalized function algebras" (https://www.encyclopediaofmath.org/index.php?title=Generalized_function_algebras), *Encyclopedia of Mathematics*, EMS Press.