AN ELEMENTARY INTRODUCTION TO DISTRIBUTIONS

TSOGTGEREL GANTUMUR

Abstract. Textbooks on PDE usually introduce distributions as linear functionals satisfying certain properties, without saying much about where those conditions come from. The reason is that it would become a book by itself if one starts with the general setting of topological vector spaces. We take here an intermediate approach, that regards families of seminorms on vector spaces as the primary objects to generate topologies.

Contents

1. Introduction

It is well known that differentiation of functions is not a well behaved operation. For instance, continuous, nowhere differentiable functions exist. The derivative of an integrable function maybe not locally integrable. A related difficulty is that if a sequence f_k converges to some function f pointwise or uniformly, then in general it is not true that f_k' converges to f' in the same sense. In order to use differentiation freely, one has to restrict to a class of functions that are many times differentiable, and in the extreme this process leads us to smooth and analytic classes. The latter classes alleviate the aforementioned difficulties somewhat, but they are too small and cumbersome for the purposes of studying PDEs. The idea behind *distributions* is that instead of restricting ourselves to a small subclass of functions, we should *expand* the class of functions to include hypothetical objects that are derivatives of ordinary functions. This will force us to extend the notion of functions, a process that is not dissimilar to extending the reals to complex numbers. The analogy can be pushed a bit further, in that by using distributions, we end up revealing deep and hidden truths even about ordinary functions that would otherwise be difficult to discover or could not be expressed naturally in the language of functions. A precise formulation of the theory of distributions was given by Laurent Schwartz during 1940's, with some crucial precursor ideas by Sergei Lvovich Sobolev.

Date: February 9, 2019.

To explain what distributions are, we start with a continuous function $u \in C(\mathbb{R})$ defined on the real line $\mathbb{R}$. Let $C_c^k(\mathbb{R})$ denote the space of k-times continuously differentiable functions with compact support, and define

$$T_u(\varphi) = \int u\varphi, \qquad \varphi \in C_c^k(\mathbb{R}). \tag{1}$$

We required φ to be compactly supported so that the above integral is finite for any continuous function u. It is clear that T_u is a linear functional acting on the space $C_c^k(\mathbb{R})$. Moreover, this specifies u uniquely, meaning that if there is some $v \in C(\mathbb{R})$ such that $T_u(\varphi) = T_v(\varphi)$ for all $\varphi \in C_c^k(\mathbb{R})$, then $u = v$. If we replace the space $C(\mathbb{R})$ by the space $L^1_{\text{loc}}(\mathbb{R})$ of locally integrable functions, the conclusion would be that $u = v$ almost everywhere, which of course means that they are equal as the elements of $L^1_{\text{loc}}(\mathbb{R})$. So we can regard ordinary functions as linear functionals on $C_c^k(\mathbb{R})$. Then the point of departure now is to consider linear functionals that are not necessarily of the form (1) as *functions in a generalized sense*. For example, the *Dirac delta*, which is just the point evaluation

$$\delta(\varphi) = \varphi(0), \qquad \varphi \in C_c^k(\mathbb{R}), \tag{2}$$

is one such functional. In order to differentiate generalized functions, let us note that

$$T_{u'}(\varphi) = \int u'\varphi = -\int u\varphi' = -T_u(\varphi'), \qquad \varphi \in C_c^k(\mathbb{R}), \tag{3}$$

for any differentiable function u, and then make the observation that the right hand side actually makes sense even if u was just a continuous function. This motivates us to define the derivative of a generalized function T by

$$T'(\varphi) := -T(\varphi'), \qquad \varphi \in C_c^k(\mathbb{R}). \tag{4}$$

If we want to get more derivatives of T, we need k to be large, which leads us to consider the space $C_c^\infty(\mathbb{R})$ of compactly supported smooth functions as the space on which the functionals T act. This space is called the space of *test functions*. A *distribution* (on $\mathbb{R}$) is simply a continuous linear functional on $C_c^\infty(\mathbb{R})$, the latter equipped with a certain topology. In order to describe this topology, we need some preparation.

2. Locally convex spaces

In this section, we will discuss how to introduce a topology on a vector space by using a family of seminorms.

Definition 1. A function $p : X \to \mathbb{R}$ on a vector space X is called a *seminorm* if

i) $p(x + y) \leq p(x) + p(y)$ for $x, y \in X$, and
ii) $p(\lambda x) = |\lambda| p(x)$ for $\lambda \in \mathbb{R}$ and $x \in X$.

It is called a *norm* if in addition $p(x) = 0$ implies $x = 0$.

The property i) is *subadditivity* or the *triangle inequality*, and ii) is *positive homogeneity.*

Lemma 2. *Let p be a seminorm on a vector space X. Then we have*

a) $p(0) = 0$,
b) $p(x) \geq 0$,
c) $|p(x) - p(y)| \leq p(x - y)$, *and*
d) $\{x \in X : p(x) = 0\}$ *is a linear space.*

Proof. Part a) follows from positive homogeneity with $\lambda = 0$. Then we have

$$0 = p(0) = p(x - x) \leq p(x) + p(-x) = p(x) + p(x), \tag{5}$$

which gives b). While c) is obvious, d) is a consequence of

$$0 \le p(\alpha x + \beta y) \le |\alpha| p(x) + |\beta| p(y), \tag{6}$$

for $\alpha, \beta \in \mathbb{R}$ and $x, y \in X$. □

Let X be a vector space, and let $\mathscr{P}$ be a family of seminorms on X. Then given a finite collection $p_1, \ldots, p_k \in \mathscr{P}$ and given $\varepsilon > 0$, let us call the set

$$B_{y,\varepsilon}(p_1, \ldots, p_k) = \{x \in X : p_i(x - y) < \varepsilon,\, i = 1, \ldots, k\}, \tag{7}$$

the *semiball* of radius ε, centred at y, corresponding to the seminorms $p_1, \ldots, p_k$.

Definition 3. Let X be a vector space, and let $\mathscr{P}$ be a family of seminorms on X. Then we define a topology on X by calling $A \subset X$ *open* if for any $x \in A$, there exists a semiball $B_{x,\varepsilon}(p_1, \ldots, p_k) \subset A$ with $p_1, \ldots, p_k \in \mathscr{P}$ and $\varepsilon > 0$. We say that $(X, \mathscr{P})$ is a *locally convex space* (LCS).

The open sets in $(X, \mathscr{P})$ are precisely those which are the unions of semiballs. It is easy to verify that X itself is open, intersection of any two open sets is open, and that the union of any collection of open sets is open. The empty set is open, because any element of the empty set, of which there is none, satisfies any desired property. Therefore the preceding definition indeed defines a topology on X, making it a topological space. Note also that the topology on X does not change if we replace a seminorm $p \in \mathscr{P}$ by another seminorm p' satisfying $p(x) \le cp'(x)$ for $x \in X$ and for some constant $c > 0$.

Remark 4. The reason we called X a *locally convex* space is that it agrees with the same notion from the theory of topological vector spaces. A *topological vector space* is a vector space which is also a topological space, with the property that the vector addition and scalar multiplication are continuous. Then a topological vector space X is called *locally convex* if $A \subset X$ is open and if $x \in A$ then there is a convex open set $C \subset A$ containing x. We choose not to go into details here, and use families of seminorms as primary objects to specify topological properties of X. This simplifies presentation and gives a quicker way to achieve our aim, and moreover does not lose generality, because of the (nontrivial) fact that any locally convex topological vector space has a family of seminorms that induces its topology.

Recall that a sequence $\{x_k\} \subset X$ is said to converge to $x \in X$ if for any open set $\omega \subset X$ containing x, we have $x_k \in \omega$ for all large k. In terms of seminorms, this is equivalent to saying that $p(x_k - x) \to 0$ for any $p \in \mathscr{P}$.

Lemma 5. *a) Let Y be a normed space, and let X be as above. Then a function $f : X \to Y$ is continuous if and only if for any $x \in X$ and any $\varepsilon > 0$, there is a finite collection $p_1, \ldots, p_k \in \mathscr{P}$ and $\delta > 0$ such that*

$$z \in B_{x,\delta}(p_1, \ldots, p_k) \quad \Rightarrow \quad \|f(x) - f(z)\|_Y \le \varepsilon. \tag{8}$$

b) In addition to what has been assumed, suppose that f is linear. Then f is continuous if and only if there is a finite collection $p_1, \ldots, p_k \in \mathscr{P}$ and a constant $C > 0$ such that

$$\|f(x)\|_Y \le C \max_i p_i(x), \qquad x \in X. \tag{9}$$

Proof. Recall that a map is called continuous if the preimage of any open set is open. Suppose that f is continuous. Then for any $\varepsilon > 0$ and $y = f(x)$ with $x \in X$, the preimage of $B_{y,\varepsilon} \subset Y$ contains a semiball $B_{x,\delta}(p_1, \ldots, p_k)$ with $\delta = \delta(\varepsilon, x) > 0$. In the other direction, let $U \subset Y$ be open and let $x \in f^{-1}(U)$. Then with $y = f(x) \in U$, there exist a nonempty ball $B_{y,\varepsilon} \subset U$, and a nonempty semiball $B_{x,\delta}(p_1, \ldots, p_k)$ such that $f(B_{x,\delta}(p_1, \ldots, p_k)) \subset B_{y,\varepsilon}$. This means that $f^{-1}(U)$ is open.

For b), the condition associated to (9) immediately implies the condition associated to (8) by linearity. Now suppose that we have the condition associated to (8). Hence there is $\delta > 0$ and $p_1, \ldots, p_k \in \mathscr{P}$ such that

$$z \in B_{0,\delta}(p_1, \ldots, p_k) \quad \Rightarrow \quad \|f(z)\|_Y \leq 1. \tag{10}$$

Note that $z \in B_{0,\delta}(p_1, \ldots, p_k)$ is equivalent to $p(z) := \max_i p_i(z) < \delta$. Let $x \in X$, and define $z = \frac{\delta}{2p(x)} x$. Then we have $p(z) = \frac{\delta}{2} < \delta$, leading to

$$1 \geq \|f(z)\|_Y = \frac{\delta}{2p(x)} \|f(x)\|_Y, \tag{11}$$

which is (9) with $C = \frac{2}{\delta}$. □

Remark 6. The preceding lemma can easily be extended to the case where Y is a LCS endowed with a family $\mathscr{Q}$ of seminorms. For instance, part b) would read: f is continuous iff for any $q \in \mathscr{Q}$, there is a finite collection $p_1, \ldots, p_k \in \mathscr{P}$ and a constant $C > 0$ such that

$$q(f(x)) \leq C \max_i p_i(x), \qquad x \in X. \tag{12}$$

Notice how the quantifiers differ on the domain and the range of the function. If $X \subset Y$ as sets, by taking $f : X \to Y$ to be the inclusion map $f(x) = x$ we derive the following criterion: the embedding $X \subset Y$ is continuous iff for any $q \in \mathscr{Q}$, there is a finite collection $p_1, \ldots, p_k \in \mathscr{P}$ and a constant $C > 0$ such that

$$q(x) \leq C \max_i p_i(x), \qquad x \in X. \tag{13}$$

Remark 7. Part b) of Lemma 5 is valid for checking continuity of seminorms $q : X \to \mathbb{R}$, because of their positive homogeneity and the property in Lemma 2c). So a seminorm q on $(X, \mathscr{P})$ is continuous iff there is a finite collection $p_1, \ldots, p_k \in \mathscr{P}$ and a constant $C > 0$ such that

$$q(x) \leq C \max_i p_i(x), \qquad x \in X. \tag{14}$$

Comparing this with the previous remark, we conclude that the embedding $X \subset Y$ is continuous iff the restriction of every seminorm of $(Y, \mathscr{Q})$ to X is continuous on $(X, \mathscr{P})$.

Definition 8. Let $(X, \mathscr{P})$ be a locally convex space. We define the following notions.

- $\{x_k\}$ is *Cauchy* if for any $p \in \mathscr{P}$, $p(x_j - x_k) \to 0$ as $j, k \to \infty$.
- $A \subset X$ is *bounded* if for any $p \in \mathscr{P}$, $\sup_{x \in A} p(x) < \infty$.

A straightforward but useful observation is that every Cauchy sequence is bounded. Indeed, if $\{x_k\}$ is Cauchy then, with an arbitrary $p \in \mathscr{P}$, for a sufficiently large j we have $p(x_j - x_k) < 1$ hence $p(k) < p(j) + 1$ for all $k \geq j$.

Definition 9. The family $\mathscr{P}$ of seminorms on X is called *separating* if for any $x \in X \setminus \{0\}$, there exists $p \in \mathscr{P}$ such that $p(x) \neq 0$.

The significance of this is that if $(X, \mathscr{P})$ is a LCS with $\mathscr{P}$ separating, then the topology of X is *Hausdorff*, meaning that for any $x, y \in X$ distinct, there are open sets $A \subset X$ and $B \subset X$ with $x \in A$ and $y \in B$. Indeed, let $p \in \mathscr{P}$ be such that $\delta := p(x - y) > 0$. Then $A = \{z \in X : p(z - x) < \frac{\delta}{2}\}$ and $B = \{z \in X : p(z - y) < \frac{\delta}{2}\}$ satisfy the desired properties.

Theorem 10. *A locally convex space $(X, \mathscr{P})$ is metrizable if $\mathscr{P}$ is countable and separating.*

Proof. Let $\mathscr{P} = \{p_1, p_2, \ldots\}$, and let $\{\alpha_k\}$ be a sequence of positive numbers satisfying $\alpha_k \to 0$. Then we claim that

$$d(x, y) = \max_k \frac{\alpha_k p_k(x - y)}{1 + p_k(x - y)}, \tag{15}$$

defines a metric that induces the topology of X. First observe that the maximum is well-defined, since $p_k/(1+p_k) < 1$ and $\alpha_k \to 0$. Also, because $\alpha_k > 0$ for all k, $d(x,y) = 0$ implies $p_k(x-y) = 0$ for all k, which then gives $x = y$ by the separating property. The triangle inequality for d follows from the elementary fact

$$a \leq b + c \qquad \Rightarrow \qquad \frac{a}{1+a} \leq \frac{b}{1+b} + \frac{c}{1+c} \qquad (a,b,c \geq 0), \tag{16}$$

which can easily be verified, e.g., by contradiction.

For each k, we have

$$\frac{p_k(x-y)}{1+p_k(x-y)} \leq d(x-y), \tag{17}$$

which tells us that any semiball contains a metric ball. To get the other direction, let $\varepsilon > 0$, and let n be an index such that $\alpha_k < \varepsilon$ for all $k > n$. Then we have

$$d(x-y) \leq \varepsilon + \max_{1\leq k\leq n} \frac{\alpha_k p_k(x-y)}{1+p_k(x-y)} \leq \varepsilon + \alpha \max_{1\leq k\leq n} p_k(x-y), \tag{18}$$

where $\alpha = \max \alpha_k$. This means that the semiball $B_{x,\varepsilon}(p_1,\ldots,p_n)$ is contained in the metric ball $B_{(1+\alpha)\varepsilon}(x) = \{y \in X : d(x-y) < (1+\alpha)\varepsilon\}$. □

Remark 11. In fact, the converse statement is also true: If $(X, \mathscr{P})$ is metrizable then $\mathscr{P}$ is countable and separating. For a proof, we refer to Walter Rudin's *Functional analysis*.

3. Examples of Fréchet spaces

In this section, we study some important examples of Fréchet spaces, which will serve as stepping stones to test functions and distributions.

Definition 12. A *Fréchet space* is a locally convex space that is metrizable with a complete, translation invariant metric.

An equivalent definition can be obtained from the fact that a locally convex space is metrizable if and only if its topology is induced by a countable and separating family of seminorms.

Let us recall the multi-index notation, which is a convenient shorthand notation for partial derivatives and multivariate polynomials. A *multi-index* is a vector $\alpha = (\alpha_1,\ldots,\alpha_n) \in \mathbb{N}_0^n$ whose components are nonnegative integers. Then we use

$$x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}, \qquad \text{and} \qquad \partial^\alpha = \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n}, \tag{19}$$

for multivariate monomials and partial derivatives. The *length* of a multi-index α is defined as $|\alpha| = \alpha_1 + \ldots + \alpha_n$, which corresponds to the total degree of a monomial or the order of a differential operator.

Given an arbitrary set $A \subset \mathbb{R}^n$, let $C(A)$ be the space of continuous functions on A. That is, $u \in C(A)$ iff

$$u(x) = \lim_{A\ni y\to x} u(y), \qquad \text{for all } x \in A. \tag{20}$$

If A is compact, all functions in $C(A)$ are bounded, which is not the case if A is open. Generalizing $C(A)$, we let $C^m(A)$ be the space of functions all of whose m-th order partial derivatives are continuous on A. In particular, $C^0(A) = C(A)$. The space of infinitely differentiable functions (i.e., smooth functions) on A is defined as

$$C^\infty(A) = \bigcap_m C^m(A). \tag{21}$$

Remark 13. An often used alternative notation is $\mathscr{E}(\Omega) = C^\infty(\Omega)$ and $\mathscr{E}^m(\Omega) = C^m(\Omega)$.

Let $K \subset \mathbb{R}^n$ be a compact set. Then $C(K)$ has the natural Banach space topology induced by the *uniform norm*

$$\|u\|_{C(K)} = \sup_{x\in K} |u(x)|, \qquad u \in C(K). \tag{22}$$

Moreover, the space $C^m(K)$ is a Banach space with the norm

$$\|u\|_{C^m(K)} = \max_{|\alpha|\leq m} \|\partial^\alpha u\|_{C^0(K)}, \qquad u \in C^m(K). \tag{23}$$

We equip $C^\infty(K)$ with the family $\{p_m : m = 0, 1, \ldots\}$ of seminorms

$$p_m(\varphi) = \|\varphi\|_{C^m(K)}, \qquad m = 0, 1, \ldots, \qquad u \in C^\infty(K). \tag{24}$$

Lemma 14. *The space $C^\infty(K)$ is metrizable and complete, i.e., it is a Fréchet space.*

Proof. The space $C^\infty(K)$ is metrizable by Theorem 10, since $\{p_m\}$ is countable and separating. Let $\{\varphi_k\}$ be a Cauchy sequence in $C^\infty(K)$. This means that $\{\varphi_k\}$ is Cauchy in $C^m(K)$ for each m. Hence by completeness of $C^m(K)$, for each m there exists $\psi_m \in C^m(K)$ such that $p_m(\varphi_k - \psi_m) \to 0$ as $k \to \infty$. Since $p_0(\psi) \leq p_m(\psi)$ for any $\psi \in C^m(K)$, we have $\psi_m = \psi_0$ for any m. We conclude that $\psi_0 \in C^\infty(K)$ and that $\varphi_k \to \psi_0$ in $C^\infty(K)$ as $k \to \infty$. □

Next we turn to function spaces defined on open sets. Let $\Omega \subset \mathbb{R}^n$ be an open set. We equip $C(\Omega)$ with the topology of locally uniform convergence, i.e., the topology induced by the seminorms

$$p_K(\varphi) = \|\varphi\|_{C(K)}, \tag{25}$$

where K runs over the compact subsets of Ω. That this topology is metrizable can be seen as follows. Suppose that $K_1 \subset K_2 \subset \ldots \subset \Omega$ are compact sets and $\bigcup_j K_j = \Omega$. Such a sequence $\{K_j\}$ can be constructed easily, for instance, by

$$K_j = \{x \in \Omega : \mathrm{dist}(x, \partial\Omega) \geq \frac{1}{j}\} \cap \overline{B}_j, \tag{26}$$

where

$$B_j = \{x \in \mathbb{R}^n : |x| < j\}, \tag{27}$$

is the open ball of radius j, centred at the origin. Obviously, if $K \subset \Omega$ is compact, then $K \subset K_j$ for some j. Hence $p_K(\varphi) \leq p_{K_j}(\varphi)$ for all $\varphi \in C(\Omega)$, meaning that we can use the family $\{p_{K_j} : j = 1, 2, \ldots\}$ to generate the topology of $C(\Omega)$. This family is countable and separating, and metrizability follows. Introducing the seminorms

$$p_{j,k}(\varphi) = \|\varphi\|_{C^k(K_j)}, \tag{28}$$

we topologize $C^m(\Omega)$ by $\{p_{j,m} : j \in \mathbb{N}\}$, and topologize $C^\infty(\Omega)$ by $\{p_{j,k} : j, k \in \mathbb{N}\}$.

Lemma 15. *Let $0 \leq m \leq \infty$. Then $C^m(\Omega)$ is a Fréchet space.*

Proof. The proof is similar to the proof of Lemma 14. Let $\{\varphi_i\}$ be a Cauchy sequence in $C^m(\Omega)$. Then by completeness of $C^m(K_j)$, for each j there exists $\psi_j \in C^m(K_j)$ such that $\varphi_i \to \psi_j$ in $C^m(K_j)$ as $i \to \infty$. Since $p_{j,k}(\psi) \leq p_{j+1,k}(\psi)$ for any $\psi \in C^k(K_{j+1})$ and any j and k, we have $\psi_j = \psi_{j+1}|_{K_j}$ for any j. This means that the function ψ defined on Ω by $\psi(x) = \psi_j(x)$ if $x \in K_j \setminus K_{j-1}$, with the convention $K_0 = \varnothing$, will satisfy $\psi \in C^m(\Omega)$ and $\psi|_{K_j} = \psi_j$ for all j. So by construction, $p_{j,k}(\varphi_i - \psi) \to 0$ as $i \to \infty$ for any j and any k, with the restriction $k \leq m$ if $m < \infty$. □

Similarly to the construction of $C^m(\Omega)$, we can introduce local versions of L^p-spaces, as

$$L^p_{\mathrm{loc}}(\Omega) = \{u : \Omega \to \mathbb{R} \text{ measurable}, \, u|_K \in L^p(K) \text{ for any compact } K \subset \Omega\}. \tag{29}$$

Here we assume $1 \leq p \leq \infty$, and equip it with the seminorms

$$q_j(\varphi) = \|\varphi\|_{L^p(K_j)}. \tag{30}$$

Relying on the completeness of $L^p(K)$, one can easily show that $L^p_{\mathrm{loc}}(\Omega)$ is a Fréchet space.

We end this section by considering function spaces with restrictions on where a function can be nonzero. If $\varphi : \Omega \to \mathbb{R}$ is a continuous function, we define its *support* as

$$\operatorname{supp}\varphi = \overline{\{x \in \Omega : \varphi(x) \neq 0\}}. \tag{31}$$

Note that $x \in \Omega$ is in $\operatorname{supp}\varphi$ if and only if x has no open neighbourhood on which φ vanishes. For $K \subset \mathbb{R}^n$ compact, we define the space

$$\mathscr{D}_K = \{\varphi \in C^\infty(\mathbb{R}^n) : \operatorname{supp}\varphi \subset K\}, \tag{32}$$

and endow it with the seminorms

$$p_m(\varphi) = \|\varphi\|_{C^m}, \qquad m = 0, 1, \ldots. \tag{33}$$

Note that this topology is the one induced by the embedding $\mathscr{D}_K \subset C^\infty(K)$.

The question arises if there exists any infinitely differentiable function with compact support. This is something we should check since a nonzero analytic function cannot have compact support, and being smooth is apparently only slightly weaker than being analytic. We claim that the function φ on $\mathbb{R}^n$ defined by

$$\varphi(x) = \begin{cases} e^{-1/(1-|x|^2)} & \text{for } |x| < 1, \\ 0 & \text{for } |x| \geq 1, \end{cases} \tag{34}$$

is in $C^\infty(\mathbb{R}^n)$. It is clear that $\varphi(x) \to 0$ as $|x| \nearrow 1$. As for the derivatives, we have

$$\partial^\alpha \varphi(x) = \frac{p(x)e^{-1/(1-|x|^2)}}{(1-|x|^2)^{|\alpha|}}, \qquad |x| < 1, \tag{35}$$

where p is some polynomial. From this it is also clear that $\partial^\alpha\varphi(x) \to 0$ as $|x| \nearrow 1$. So $\varphi \in C^\infty(\mathbb{R}^n)$. If K contains an open ball, we can fit infinitely many open balls inside K. Then scaling and translating φ, we can place them in K so that their supports are contained in K and do not intersect with each other. This implies that $\mathscr{D}_K$ is infinite dimensional. The space $\mathscr{D}_K$ is also Fréchet, since it is a closed subspace of $C^\infty(K)$.

For any integer $m \geq 0$ we can also introduce

$$\mathscr{D}^m_K = \{\varphi \in C^m(\mathbb{R}^n) : \operatorname{supp}\varphi \subset K\}, \tag{36}$$

and endow it with the subspace topology inherited from $C^m(K)$. Then $\mathscr{D}^m_K$ is a closed subspace of $C^m(K)$.

4. The inductive limit topology

In this section, we will establish some basic properties of the so-called inductive limit topology on the space of test functions.

Definition 16. Let $\Omega \subset \mathbb{R}^n$ be an open set. Then we define the *space of test functions* by

$$\mathscr{D}(\Omega) = \bigcup_{K \Subset \Omega} \mathscr{D}_K, \tag{37}$$

where we used the notation $K \Subset \Omega$ to mean that K is compact and is a subset of Ω.

Note that if $K_1 \subset K_2 \subset \ldots \subset \Omega$ are compact sets and $\bigcup_m K_m = \Omega$, then

$$\mathscr{D}(\Omega) = \bigcup_m \mathscr{D}_{K_m}. \tag{38}$$

We have discussed a construction of such a sequence in the preceding section.

Our next task is to introduce a topology on $\mathscr{D}(\Omega)$. In doing so, we want the inclusions $\mathscr{D}_K \subset \mathscr{D}(\Omega)$ to be continuous. This means, by Remark 6 that for every seminorm p from $(\mathscr{D}(\Omega), \mathscr{P})$, where $\mathscr{P}$ is the hypothetical family inducing a topology on $\mathscr{D}(\Omega)$, the restriction

$p|_{\mathscr{D}_K}$ must be continuous on $\mathscr{D}_K$. The family $\mathscr{P} = \{p_m\}$ has the desired property, but the following remark shows that it would not be a very convenient choice.

Remark 17. $\mathscr{D}(\Omega)$ is *not* complete with respect to the topology induced by $\{p_m\}$. We illustrate it in the case $\Omega = \mathbb{R}$. Take a nonzero function $\varphi \in \mathscr{D}(\mathbb{R})$ whose support is small and concentrated near 0, and consider the sequence

$$\varphi_k(x) = \varphi(x) + 2^{-1}\varphi(x-1) + \ldots + 2^{-k}\varphi(x-k), \qquad k = 1, 2, \ldots. \tag{39}$$

Obviously, this sequence is Cauchy with respect to the family $\{p_m\}$, but the support of the limit function is not compact.

This failure indicates that the family $\{p_m\}$ has not enough seminorms to prevent Cauchy sequences from "leaking" towards the boundary of Ω. So we can add more seminorms to the family, and hope that things get better. Having a large family of seminorms will have the added benefit that it becomes easier for a function $f : \mathscr{D}(\Omega) \to Y$ to be continuous, meaning that we will have a large supply of continuous functions on $\mathscr{D}(\Omega)$. Of course there is a limit in expanding the family $\mathscr{P}$ because of the aforementioned requirement that $p|_{\mathscr{D}_K}$ be continuous. These two competing requirements give rise to a unique family $\mathscr{P}$ as follows.

Definition 18. We define the collection $\mathscr{P}$ of seminorms on $\mathscr{D}(\Omega)$ by the condition that a seminorm p on $\mathscr{D}(\Omega)$ is in $\mathscr{P}$ iff $p|_{\mathscr{D}_K}$ is continuous for each compact $K \subset \Omega$.

The topology generated by $\mathscr{P}$ on $\mathscr{D}(\Omega)$ is called the *inductive limit topology*. Looking back, this topology is completely natural, given that $\mathscr{D}(\Omega)$ is the union of $\{\mathscr{D}_K : K \Subset \Omega\}$, and that each $\mathscr{D}_K$ has its own topology.

Remark 19. In general, if $X_1 \subset X_2 \subset \ldots$ are locally convex spaces, then the inductive limit topology on the union $X = \bigcup_j X_j$ is the finest topology that leaves the embeddings $X_j \to X$ continuous. If each of the spaces X_j is Fréchet, we call the resulting space X an *LF space*. If each X_j is Banach, we call X an *LB space*.

Lemma 20. *The topology of $\mathscr{D}_K$ is exactly the one induced by the embedding $\mathscr{D}_K \subset \mathscr{D}(\Omega)$.*

Proof. Let $A \subset \mathscr{D}(\Omega)$ be open and let $K \subset \Omega$ be compact. We will show that $A \cap \mathscr{D}_K$ is open in $\mathscr{D}_K$. Let $\psi \in A \cap \mathscr{D}_K$. Let us denote the semiballs in $\mathscr{D}_K$ by $B_{\psi,\varepsilon}(p_m; \mathscr{D}_K)$ etc., and the semiballs in $\mathscr{D}(\Omega)$ by $B_{\psi,\varepsilon}(p)$ etc. Then there exists $p \in \mathscr{P}$ such that $B_{\psi,\varepsilon}(p) \subset A$ with $\varepsilon > 0$. By construction, there exists p_m such that $p \leq cp_m$ on $\mathscr{D}_K$, with some constant $c > 0$. Hence $B_{\psi,\varepsilon/c}(p_m, \mathscr{D}_K) \subset B_{\psi,\varepsilon}(p) \cap \mathscr{D}_K \subset A \cap \mathscr{D}_K$, showing that $A \cap \mathscr{D}_K$ is open in $\mathscr{D}_K$.

On the other hand, since $\{p_m\} \subset \mathscr{P}$, any semiball $B_{\psi,\varepsilon}(p_m; \mathscr{D}_K)$ in $\mathscr{D}_K$ is equal to the intersection of the semiball $B_{\psi,\varepsilon}(p_m)$ in $\mathscr{D}(\Omega)$ with $\mathscr{D}_K$, i.e.,

$$B_{\psi,\varepsilon}(p_m; \mathscr{D}_K) = B_{\psi,\varepsilon}(p_m) \cap \mathscr{D}_K. \tag{40}$$

This immediately implies that any open set in $\mathscr{D}_K$ can be written as the intersection of an open set of $\mathscr{D}(\Omega)$ with $\mathscr{D}_K$. □

Let us ask the question: Does $\mathscr{P}$ have any seminorm that is not one of $\{p_m\}$? An example of such a seminorm is given by

$$p(\varphi) = \sup_j c_j |\varphi(x_j)|, \qquad \varphi \in \mathscr{D}(\Omega), \tag{41}$$

where $\{x_j\} \subset \Omega$ is a sequence having no accumulation points in Ω, and $\{c_j\}$ is a sequence of positive numbers. We can easily check that p is a seminorm, and that $p|_{\mathscr{D}_K}$ is continuous on $\mathscr{D}_K$ for any compact $K \subset \Omega$, so that $p \in \mathscr{P}$. Seminorms such as this give a very strong control near the boundary of Ω, because $\{x_j\}$ concentrate towards the boundary and c_j can grow arbitrarily fast. The following result illustrates this phenomenon.

Theorem 21. *The set $A \subset \mathscr{D}(\Omega)$ is bounded if and only if there is a compact set $K \subset \Omega$ such that $A \subset \mathscr{D}_K$ and that A is bounded in $\mathscr{D}_K$. Recall that the latter means that each p_m is bounded on A.*

Proof. Suppose that A is bounded in $\mathscr{D}_K$ for some compact set $K \subset \Omega$. We claim that continuity of the embedding $\mathscr{D}_K \subset \mathscr{D}(\Omega)$ implies that A is also bounded in $\mathscr{D}(\Omega)$. To prove it, let $p \in \mathscr{P}$. Then there is p_m such that

$$p(\varphi) \leq Cp_m(\varphi), \qquad \varphi \in \mathscr{D}_K. \tag{42}$$

By assumption, $p_m(\varphi) \leq M$ for $\varphi \in A$ and for some constant M, which implies that p is bounded on A.

To prove the other direction, suppose that $A \not\subset \mathscr{D}_K$ for any compact $K \subset \Omega$. Then there exist sequences $\{\varphi_m\} \subset A$ and $\{x_m\} \subset \Omega$ such that $\varphi(x_m) \neq 0$, and that $\{x_m\}$ has no accumulation points in Ω. Let

$$p(\varphi) = \sup_m \frac{m|\varphi(x_m)|}{|\varphi_m(x_m)|}, \qquad \varphi \in \mathscr{D}(\Omega). \tag{43}$$

Obviously it is a seminorm, and $p \in \mathscr{P}$ because for any compact $K' \subset \Omega$ there is a constant C such that

$$p(\varphi) \leq C\|\varphi\|_{C^0}, \qquad \varphi \in \mathscr{D}_{K'}. \tag{44}$$

However, we have $p(\varphi_m) \geq m$, so p is not bounded on A, leading to a contradiction. □

Corollary 22. *a) The sequence $\{\varphi_j\}$ is Cauchy in $\mathscr{D}(\Omega)$ iff $\{\varphi_j\} \subset \mathscr{D}_K$ for some compact $K \subset \Omega$, and $\|\varphi_j - \varphi_k\|_{C^m} \to 0$ as $j, k \to \infty$, for each m.*

b) We have $\varphi_j \to 0$ in $\mathscr{D}(\Omega)$ if and only if $\{\varphi_j\} \subset \mathscr{D}_K$ for some compact $K \subset \Omega$, and $\|\varphi_j\|_{C^m} \to 0$ as $j \to \infty$, for each m.

c) $\mathscr{D}(\Omega)$ is sequentially complete.

Proof. a) If $\{\varphi_j\} \subset \mathscr{D}_K$ is Cauchy in $\mathscr{D}_K$ for some compact $K \subset \Omega$, then it is Cauchy in $\mathscr{D}(\Omega)$ by continuity of the embedding $\mathscr{D}_K \subset \mathscr{D}(\Omega)$. Now let $\{\varphi_j\} \subset \mathscr{D}(\Omega)$ be Cauchy in $\mathscr{D}(\Omega)$. Since Cauchy sequences are bounded, by the preceding theorem we have $\{\varphi_j\} \subset \mathscr{D}_K$ for some compact $K \subset \Omega$. But then $\{p_m\} \subset \mathscr{P}$, which means that $p_m(\varphi_j - \varphi_k) \to 0$ as $j, k \to \infty$, for each p_m.

b) Left as an exercise.

c) Let $\{\varphi_j\} \subset \mathscr{D}(\Omega)$ be Cauchy in $\mathscr{D}(\Omega)$. Then by a) it is Cauchy in some $\mathscr{D}_K$. But $\mathscr{D}_K$ is Fréchet, so the limit exists in $\mathscr{D}_K$. This limit is valid also in $\mathscr{D}(\Omega)$, since a convergent sequence in $\mathscr{D}_K$ is convergent in $\mathscr{D}(\Omega)$. □

Theorem 23. *Let $(Y, \mathscr{Q})$ be a locally convex space, and let $f : \mathscr{D}(\Omega) \to Y$ be a linear map. Then the following are equivalent.*

(a) f is continuous.
(b) $\varphi_j \to 0$ in $\mathscr{D}(\Omega)$ implies $f(\varphi_j) \to 0$ in Y.
(c) For any compact $K \subset \Omega$, $f : \mathscr{D}_K \to Y$ is continuous.

Proof. a) ⇒ b). The continuity of f means that for any $q \in \mathscr{Q}$, there is $p \in \mathscr{P}$ such that

$$q(f(\varphi)) \leq p(\varphi), \qquad \varphi \in \mathscr{D}(\Omega). \tag{45}$$

Since $\varphi_j \to 0$ in $\mathscr{D}(\Omega)$, we have $p(\varphi_j) \to 0$, hence $q(f(\varphi_j)) \to 0$. As $q \in \mathscr{Q}$ is arbitrary, we conclude that $f(\varphi_j) \to 0$ in Y.

b) ⇒ c). Let $K \subset \Omega$ be compact. If b) holds then for any sequence $\varphi_j \to 0$ in $\mathscr{D}_K$ we have $f(\varphi_j) \to 0$ in Y. Then continuity of $f : \mathscr{D}_K \to Y$ follows from the general fact that for a metric space X and a topological space Y, a map $f : X \to Y$ is continuous if whenever $x_j \to x$ in X we have $f(x_j) \to f(x)$ in Y. To prove this fact, supposing that f is *not* continuous at

$x \in X$, we want to show that there is a sequence $x_n \to x$ with $f(x_n) \not\to f(x)$. Let $U \subset Y$ be an open set such that $f(x) \in U$ and that $f^{-1}(U)$ is not open. Hence $f^{-1}(U)$ does not contain any metric ball $B_\varepsilon(x) = \{z \in X : d(z, x) < \varepsilon\}$ with $\varepsilon > 0$, where d is the metric of X. This means that for any $\varepsilon > 0$, there is $z \in B_\varepsilon(x)$ with $f(z) \notin U$, i.e., there exists a sequence $x_n \to x$ with $f(x_n) \notin U$ for all n.

c) ⇒ a). We want to show that for any $q \in \mathscr{Q}$, there is $p \in \mathscr{P}$ such that (45) holds. Given q, let us define the function

$$p(\varphi) = q(f(\varphi)), \qquad \varphi \in \mathscr{D}(\Omega). \tag{46}$$

It is a seminorm on $\mathscr{D}(\Omega)$, and moreover for each compact $K \subset \Omega$, the restriction $p|_{\mathscr{D}_K}$ is continuous since

$$q(f(\varphi)) \leq C p_m(\varphi), \qquad \varphi \in \mathscr{D}_K, \tag{47}$$

for some C and m possibly depending on K. Therefore $p \in \mathscr{P}$, which clearly implies (45). □

Example 24. The partial differentiation operator $\partial_j : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ is continuous, since for any compact $K \subset \Omega$ and any m, we have

$$p_m(\partial_j \varphi) \leq p_{m+1}(\varphi), \qquad \varphi \in \mathscr{D}_K. \tag{48}$$

Remark 25. $\mathscr{D}(\Omega)$ is *not* metrizable. We illustrate this in the case $\Omega = \mathbb{R}$. Pick a function $\varphi \in \mathscr{D}(\mathbb{R})$ with $\operatorname{supp} \varphi = [-1, 1]$, and define the double-indexed sequence

$$\varphi_{km}(x) = \frac{1}{m} \varphi\big(\frac{x}{k}\big), \qquad k, m = 1, 2, \ldots. \tag{49}$$

It is clear that for each fixed k, the sequence $\varphi_{k,1}, \varphi_{k,2}, \ldots$ converges to 0 in $\mathscr{D}(\mathbb{R})$. Then if $\mathscr{D}(\mathbb{R})$ was metrizable, say with metric d, we can extract a sequence $m_1, m_2, \ldots$, such that $\varphi_{k,m_k} \to 0$ in $\mathscr{D}(\mathbb{R})$. This can be done, for instance, by choosing m_k sufficiently large so that $d(\varphi_{k,m_k}, 0) < \frac{1}{k}$, for each k. But it is not possible for such a sequence to converge in $\mathscr{D}(\mathbb{R})$, because the support of φ_{k,m_k} is $[-k, k]$, which eventually becomes larger than any compact set in $\mathbb{R}$.

Remark 26. We define the space of compactly supported C^m-functions

$$\mathscr{D}^m(\Omega) = \bigcup_{K \Subset \Omega} \mathscr{D}^m_K, \tag{50}$$

and the space of compactly supported L^p-functions

$$L^p_{\text{comp}}(\Omega) = \bigcup_{K \Subset \Omega} L^p(K), \tag{51}$$

where in the right hand side, the elements of $L^p(K)$ are extended by zero outside K. They have natural inductive limit topologies, and all the results of the current section apply to these spaces, with obvious modifications.

5. Subspaces of distributions

From now on the space $\mathscr{D}(\Omega)$ is equipped with its inductive limit topology.

Definition 27. A *distribution* on Ω is a continuous linear functional on $\mathscr{D}(\Omega)$. The space of all distributions on Ω is denoted by $\mathscr{D}'(\Omega)$.

We denote the action $u(\varphi)$ of $u \in \mathscr{D}'(\Omega)$ also by $\langle u, \varphi \rangle$. Theorem 23 tailored to distributions is the following.

Lemma 28. *A linear functional $u : \mathscr{D}(\Omega) \to \mathbb{R}$ is in $\mathscr{D}'(\Omega)$ iff any of the following holds.*
(a) $\varphi_j \to 0$ in $\mathscr{D}(\Omega)$ implies $u(\varphi_j) \to 0$.

(b) For any compact $K \subset \Omega$, there exist m and C such that

$$|u(\varphi)| \leq C\|\varphi\|_{C^m} \qquad \textit{for} \quad \varphi \in \mathscr{D}_K. \tag{52}$$

Definition 29. Let $u \in \mathscr{D}'(\Omega)$. If we have

$$|u(\varphi)| \leq C\|\varphi\|_{C^m} \qquad \text{for} \quad \varphi \in \mathscr{D}_K, \tag{53}$$

with the same m for all compact $K \subset \Omega$, with C possibly depending on K, then u is said to be a *distribution of order* $\leq m$. The smallest such m is called the *order of* u.

The rationale for this definition is the idea that lower order distributions are more regular, because in general the parameter m in (52) will depend on the compact set K, and will grow unboundedly as K approaches the boundary of Ω. We shall make this intuition more precise later.

Example 30. For $u \in C(\Omega)$, the functional $T_u : \varphi \mapsto \int u\varphi$ is a distribution of order 0 since

$$|T_u(\varphi)| \leq \mathrm{vol}(K)\|u\|_{C^0(K)}\|\varphi\|_{C^0}, \qquad \text{for} \quad \varphi \in \mathscr{D}_K. \tag{54}$$

Similarly, δ is a distribution of order 0, and the derivative evaluation $\varphi \mapsto \varphi'(0)$ is a distribution of order 1.

Definition 31. Let X be a locally convex space equipped with the family $\mathscr{P}$ of seminorms, and let X' be its topological dual. Then the *weak dual topology* (or *weak-$*$ topology*) on X' is the one induced by the family of seminorms $\mathscr{P}' = \{p_x : x \in X\}$, where $p_x(u) = |u(x)|$.

Thus $u_j \to 0$ in the weak dual topology of $\mathscr{D}'(\Omega)$ iff

$$u_j(\varphi) \to 0 \qquad \text{for each} \quad \varphi \in \mathscr{D}(\Omega). \tag{55}$$

We see that this is simply the pointwise convergence. The family $\mathscr{P}'$ is separating, since if $u \in X'$ is nonzero, there is $x \in X$ such that $u(x) \neq 0$. Hence the weak dual topology is Hausdorff.

Remark 32. Another natural topology on $\mathscr{D}'(\Omega)$ is the *strong dual topology* that is described by the seminorms $p_{\mathscr{B}}(u) = \sup_{\varphi\in\mathscr{B}} |u(\varphi)|$, where $\mathscr{B}$ varies over bounded subsets of $\mathscr{D}(\Omega)$. This topology is the topology of uniform convergence on bounded sets (which is a generalization of locally uniform convergence). It turns out that the weak and strong dual topologies produce the same bounded subsets for $\mathscr{D}'(\Omega)$, and these two topologies themselves coincide on bounded subsets of $\mathscr{D}'(\Omega)$. So in particular, a sequence converges in the weak topology if and only if it converges in the strong topology. In these notes we will be concerned only with the weak dual topology.

Example 33. For $u \in L^1_{\mathrm{loc}}(\Omega)$, the functional $T_u : \varphi \mapsto \int u\varphi$ is a distribution of order 0 since

$$|T_u(\varphi)| \leq \|u\|_{L^1(K)}\|\varphi\|_{C^0}. \qquad \text{for} \quad \varphi \in \mathscr{D}_K, \tag{56}$$

We have seen that the map $u \mapsto T_u : L^1_{\mathrm{loc}}(\Omega) \to \mathscr{D}'(\Omega)$ is an injection, so that $L^1_{\mathrm{loc}}(\Omega)$ can be regarded as a subspace of $\mathscr{D}'(\Omega)$. Thus we will identify T_u with u. Then with the (Fréchet) topology on $L^1_{\mathrm{loc}}(\Omega)$ defined by the seminorms $\{\|\cdot\|_{L^1(K)} : K \Subset \Omega\}$, from the above inequality we infer that $u_j \to 0$ in $L^1_{\mathrm{loc}}(\Omega)$ implies $\langle u_j, \varphi\rangle \to 0$ for any fixed $\varphi \in \mathscr{D}(\Omega)$. Hence the embedding $L^1_{\mathrm{loc}}(\Omega) \subset \mathscr{D}'(\Omega)$ is continuous. We can also infer the continuity of the embedding $C(\Omega) \subset \mathscr{D}'(\Omega)$ either directly or through the continuous embedding $C(\Omega) \subset L^1_{\mathrm{loc}}(\Omega)$.

Example 34. Consider $u_j(x) = \sin(jx)$. Then $u_j \to 0$ in $\mathscr{D}'(\mathbb{R})$, since for any $\varphi \in \mathscr{D}(\mathbb{R})$ we have

$$\int \sin(jx)\varphi(x)\mathrm{d}x = \frac{1}{j}\int \cos(jx)\varphi'(x)\mathrm{d}x \to 0 \qquad \text{as} \quad j \to \infty, \tag{57}$$

Analogously to $\mathscr{D}'(\Omega)$, we can define the dual spaces $\mathscr{E}'(\Omega)$, $\mathscr{E}'^m(\Omega)$, and $\mathscr{D}'^m(\Omega)$, that are the topological duals of $\mathscr{E}(\Omega)$, $\mathscr{E}^m(\Omega)$, and $\mathscr{D}^m(\Omega)$, respectively. Recall that $\mathscr{E}(\Omega) = C^\infty(\Omega)$ and $\mathscr{E}^m(\Omega) = C^m(\Omega)$. Here $\mathscr{E}(\Omega)$ and $\mathscr{E}^m(\Omega)$ carry their natural Fréchet space topologies, and $\mathscr{D}^m(\Omega)$ is equipped with its inductive limit topology.

Remark 35. We have $u \in \mathscr{D}'^m(\Omega)$ iff for any compact $K \subset \Omega$, there exists $C > 0$ such that

$$|u(\varphi)| \leq C\|\varphi\|_{C^m} \qquad \text{for} \quad \varphi \in \mathscr{D}^m_K. \tag{58}$$

In light of Definition 29, this reveals that the elements of $\mathscr{D}'^m(\Omega)$ are distributions of order at most m. Conversely, if $u \in \mathscr{D}'(\Omega)$ is of order at most m, then for each compact $K \subset \Omega$ we have the bound (58) for $\varphi \in \mathscr{D}_K$. This means that the map $u : \mathscr{D}(\Omega) \to \mathbb{R}$ is continuous when we take $\mathscr{D}(\Omega)$ with the subspace topology induced by the embedding $\mathscr{D}(\Omega) \subset \mathscr{D}^m(\Omega)$. Hence u can be extended to continuous $u : \mathscr{D}^m(\Omega) \to \mathbb{R}$ in a unique way, because the embedding $\mathscr{D}(\Omega) \subset \mathscr{D}^m(\Omega)$ is dense. To conclude, the space $\mathscr{D}'^m(\Omega)$ is precisely the subspace of $\mathscr{D}(\Omega)$ consisting of distributions of order at most m.

We shall see an analogous characterization of the spaces $\mathscr{E}'(\Omega)$ and $\mathscr{E}'^m(\Omega)$ in a later section. For now, let us ascertain that they are indeed subspaces of $\mathscr{D}'(\Omega)$.

Let X and Y be locally convex spaces, and let $A : X \to Y$ be a linear continuous operator. Then the *transpose* $A' : Y' \to X'$ is defined by

$$\langle Ax, y'\rangle = \langle x, A'y'\rangle, \qquad x \in X,\, y' \in Y'. \tag{59}$$

This situation can be described by the diagram

$$X \xrightarrow{\;A\;} Y \xrightarrow{\;y'\;} \mathbb{R}\,, \qquad X \overset{A'y'}{\dashrightarrow} \mathbb{R} \tag{60}$$

In other words, given $y' \in Y'$, $A'y' \in X'$ is simply $y' \circ A$, i.e., the pullback of y' under A.

Equip X' and Y' with their weak dual topologies. Then for any $x \in X$, we have

$$p_x(A'y') \equiv |\langle A'y', x\rangle| = |\langle y', Ax\rangle| \equiv p_y(y'), \tag{61}$$

where $y = Ax$, showing that $A' : Y' \to X'$ is continuous.

Theorem 36. *In this setting, $A' : Y' \to X'$ is injective if and only if $A(X)$ is dense in Y.*

Proof. Suppose that $A(X)$ is dense in Y. We want to show that $A'y' = 0$ implies $y' = 0$. By definition, $A'y' = 0$ means that $y'(Ax) = 0$ for all $x \in X$, i.e., that $A(X) \subset (y')^{-1}(\{0\})$. But the set $(y')^{-1}(\{0\})$ is closed, so it must contain the closure of $A(X)$, which is Y by the density assumption. Hence $y' = 0$.

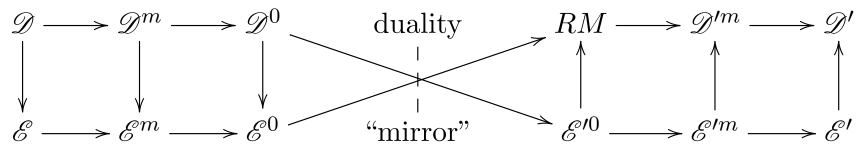
In the other direction, assume that $A(X)$ is *not* dense in Y. We want to produce a nonzero element $y' \in Y'$ such that $A'y' = 0$. By assumption, there is an element $y \in Y$ such that $y \notin \overline{A(X)}$. Consider the quotient $Z = Y/\overline{A(X)}$ with $\pi : Y \to Z$ the canonical projection. Then since $\pi(y) \neq 0$, there is $z' \in Z'$ such that $z'(\pi(y)) = 1$. So if we define $y' = z' \circ \pi$, we get $y \in Y'$ and $y'(y) = 1$. On the other hand, we have $y'(\eta) = 0$ for $\eta \in A(X)$, i.e.,

$$\langle A'y', x\rangle = \langle y', Ax\rangle = 0, \qquad x \in X, \tag{62}$$

implying that $A'y' = 0$. ∎

Corollary 37. *If X is a dense subspace of Y and if the embedding $X \subset Y$ is continuous, then Y' is canonically identified with a subspace of X' such that the embedding $Y' \subset X'$ is continuous.*

As an application of this corollary, in the following diagram, we arrange several important function spaces, and derive embedding relationships between their duals.



(63)

The spaces on the right hand side are the duals of the spaces on the left hand side. Each arrow represents a continuous and dense embedding. The domain Ω in each space is understood, e.g., $\mathscr{D} = \mathscr{D}(\Omega)$. Recall that $\mathscr{E}(\Omega) = C^\infty(\Omega)$ and $\mathscr{E}^m(\Omega) = C^m(\Omega)$. It can be taken as a definition that $RM(\Omega)$, the space of *Radon measures* on Ω, is equal to the topological dual of $\mathscr{D}^0(\Omega)$, the space of continuous functions with compact support in Ω, equipped with its inductive limit (LB) topology.

6. Basic operations on distributions

Now we want to extend some basic operations on functions to distributions. This is usually achieved by means of a simple duality device that can be described as follows. Suppose that $T, T' : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ are continuous linear maps, satisfying

$$\int (T\psi)\varphi = \int \psi(T'\varphi), \qquad \psi, \varphi \in \mathscr{D}(\Omega). \tag{64}$$

Then we define $\tilde{T} : \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega)$, which is the intended extension of T, by

$$\langle \tilde{T}u, \varphi\rangle = \langle u, T'\varphi\rangle, \qquad u \in \mathscr{D}'(\Omega),\ \varphi \in \mathscr{D}(\Omega). \tag{65}$$

It is easily checked that $\tilde{T}u \in \mathscr{D}'(\Omega)$, since by continuity of T', $\varphi_j \to 0$ in $\mathscr{D}(\Omega)$ implies $T'\varphi_j \to 0$ in $\mathscr{D}(\Omega)$, which then implies that $\langle u, T'\varphi_j\rangle \to 0$. Moreover, $\tilde{T} : \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega)$ is continuous, because

$$p_\varphi(\tilde{T}u) = |\langle \tilde{T}u, \varphi\rangle| = |\langle u, T'\varphi\rangle| = p_\psi(u), \tag{66}$$

where $\psi = T'\varphi \in \mathscr{D}(\Omega)$. If $u \in \mathscr{D}(\Omega)$, then

$$\langle \tilde{T}u, \varphi\rangle = \langle u, T'\varphi\rangle = \int u(T'\varphi) = \int (Tu)\varphi = \langle Tu, \varphi\rangle, \tag{67}$$

hence $\tilde{T}$ is indeed an extension of T. In fact, $\tilde{T}$ is the *unique* continuous extension of T. To see this, we will use the (nontrivial) fact that $\mathscr{D}(\Omega)$ is sequentially dense in $\mathscr{D}'(\Omega)$, i.e., for any $u \in \mathscr{D}'(\Omega)$, there exists a sequence $\{u_j\} \subset \mathscr{D}(\Omega)$ such that $u_j \to u$ in $\mathscr{D}'(\Omega)$. Let T_1 and T_2 be two continuous extensions of T. Then with u and $\{u_j\}$ as above, since $T_1u_j = T_2u_j$, we have

$$T_1u - T_2u = T_1(u - u_j) + T_2(u_j - u), \tag{68}$$

which implies for any $\varphi \in \mathscr{D}(\Omega)$ that

$$\begin{aligned} |\langle T_1u - T_2u, \varphi\rangle| &\leq p_\varphi(T_1(u-u_j)) + p_\varphi(T_2(u-u_j)) \\ &\leq C_1 p_{\psi_1}(u-u_j) + C_2 p_{\psi_2}(u-u_j), \end{aligned} \tag{69}$$

with some $\psi_1, \psi_2 \in \mathscr{D}(\Omega)$, and some constants $C_1, C_2 > 0$. Now sending $j \to \infty$ we get $\langle T_1u - T_2u, \varphi\rangle = 0$ for any $\varphi \in \mathscr{D}(\Omega)$, hence $T_1u = T_2u$.

Let us consider now some applications of this device.

Differentiation: $T = \partial_j$. As we have already discussed, the operator $\partial_j : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ is continuous, and integration by parts gives

$$\int \varphi\partial_j\psi = -\int \psi\partial_j\varphi, \qquad \psi, \varphi \in \mathscr{D}(\Omega). \tag{70}$$

Hence $T' = -\partial_j$, and the derivative of $u \in \mathscr{D}'(\Omega)$ is given by

$$\langle \partial_j u, \varphi \rangle = -\langle u, \partial_j \varphi \rangle, \qquad \varphi \in \mathscr{D}(\Omega). \tag{71}$$

For any multi-index α, this generalizes to

$$\langle \partial^\alpha u, \varphi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle, \qquad \varphi \in \mathscr{D}(\Omega). \tag{72}$$

Multiplication by a smooth function: $T\psi = a\psi$, where $a \in C^\infty(\Omega)$. One can easily see that $T' = T$, so up to showing continuity of T on $\mathscr{D}(\Omega)$, we infer

$$\langle au, \varphi \rangle = \langle u, a\varphi \rangle, \qquad u \in \mathscr{D}'(\Omega),\, \varphi \in \mathscr{D}(\Omega). \tag{73}$$

The continuity of $T : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ is left as an exercise.

Translation: $(T\psi)(x) = \psi(x + a)$, where $a \in \mathbb{R}^n$. We take $\Omega = \mathbb{R}^n$. By change of variables, we have

$$\int \psi(x + a)\varphi(x)\mathrm{d}x = \int \psi(x)\varphi(x - a)\mathrm{d}x, \qquad \psi, \varphi \in \mathscr{D}(\mathbb{R}^n), \tag{74}$$

so with $(\tau_a \psi)(x) = \psi(x + a)$, we infer

$$\langle \tau_a u, \varphi \rangle = \langle u, \tau_{-a} \varphi \rangle, \qquad u \in \mathscr{D}'(\mathbb{R}^n),\, \varphi \in \mathscr{D}(\mathbb{R}^n). \tag{75}$$

The continuity of τ_a on test functions is left as an exercise.

Convolution with a test function: $T\psi = a * \psi$, where $a \in \mathscr{D}(\mathbb{R}^n)$. We have

$$\begin{aligned} \int (a * \psi)\varphi &= \int \int a(x - z)\psi(z)\varphi(x)\mathrm{d}z\mathrm{d}x \\ &= \int \psi(\tilde{a} * \varphi), \qquad \psi, \varphi \in \mathscr{D}(\mathbb{R}^n), \end{aligned} \tag{76}$$

where $\tilde{a}(x) = a(-x)$ denotes the reflection through the origin. Again leaving the continuity question as an exercise, we get

$$\langle a * u, \varphi \rangle = \langle u, \tilde{a} * \varphi \rangle, \qquad u \in \mathscr{D}'(\mathbb{R}^n),\, \varphi \in \mathscr{D}(\mathbb{R}^n). \tag{77}$$

Example 38. Let $\theta \in L^1_{\mathrm{loc}}(\mathbb{R})$ be the Heaviside step function, defined by $\theta(x) = 1$ for $x > 0$ and $\theta(x) = 0$ for $x < 0$. Then its distributional derivative is given by

$$\langle \theta', \varphi \rangle = -\langle \theta, \varphi' \rangle = -\int_0^\infty \varphi'(x)\mathrm{d}x = \varphi(0), \tag{78}$$

for any $\varphi \in \mathscr{D}(\mathbb{R})$. Hence $\theta' = \delta$.

7. Support of distributions

Definition 39. Let $u \in \mathscr{D}'(\Omega)$ and let $\omega \subset \Omega$ be open. The *restriction* $u|_\omega \in \mathscr{D}'(\omega)$ of u to ω is defined by

$$\langle u|_\omega, \varphi \rangle = \langle u, \varphi \rangle, \qquad \varphi \in \mathscr{D}(\omega). \tag{79}$$

We say that $u = 0$ on ω if $u|_\omega = 0$.

This gives us a possibility to talk about distributions locally, meaning that we can focus on small open sets, one at a time. In order for this to be meaningful, we expect some natural properties to be satisfied by the restriction process. First, let us check if the above definition indeed makes sense, i.e., if $u|_\omega \in \mathscr{D}'(\omega)$. So let $\varphi_j \to 0$ in $\mathscr{D}(\omega)$. Then $\varphi_j \to 0$ in $\mathscr{D}(\Omega)$, because there is a compact $K \subset \omega$ such that $\varphi_j \to 0$ in $\mathscr{D}_K$. Since $u \in \mathscr{D}'(\Omega)$, we have $\langle u|_\omega, \varphi_j \rangle = \langle u, \varphi_j \rangle \to 0$, showing that $u|_\omega \in \mathscr{D}'(\omega)$. Note that the same argument also demonstrates that the embedding $\mathscr{D}(\omega) \subset \mathscr{D}(\Omega)$ is continuous. However, unless $\omega = \Omega$, the topology of $\mathscr{D}(\omega)$ is *strictly finer* than that induced by the embedding $\mathscr{D}(\omega) \subset \mathscr{D}(\Omega)$, i.e., there are more open sets in $\mathscr{D}(\omega)$ than those inherited from $\mathscr{D}(\Omega)$. The reason is that for instance, the seminorm $p(\varphi) = \sup j|\varphi(x_j)|$ with $\{x_j\}$ having no accumulation points in

ω, is compatible with the topology of $\mathscr{D}(\omega)$, while it is in general not with the topology of $\mathscr{D}(\Omega)$. This results in the fact that not every distribution in $\mathscr{D}'(\omega)$ is the restriction of some distribution in $\mathscr{D}'(\Omega)$.

The following theorem shows that as far as restrictions are concerned, we can work with distributions as if they were functions. The properties a)-d) in the theorem are called the *sheaf properties*.

Theorem 40. *Let $u \in \mathscr{D}'(\Omega)$.*

a) $u|_\Omega = u$.
b) $(u|_\omega)|_\sigma = u|_\sigma$ *for open sets* $\sigma \subset \omega \subset \Omega$.
c) *If* $\{\omega_\alpha\}$ *is an open cover of* Ω, *then*

$$\forall\alpha,\, u|_{\omega_\alpha} = 0 \qquad \Rightarrow \qquad u = 0. \tag{80}$$

d) *With* $\{\omega_\alpha\}$ *as in c), let* $u_\alpha \in \mathscr{D}'(\omega_\alpha)$ *is given for each* α, *satisfying*

$$u_\alpha|_{\omega_\alpha\cap\omega_\beta} = u_\beta|_{\omega_\alpha\cap\omega_\beta} \qquad \forall\alpha,\beta. \tag{81}$$

Then there exists a unique $u \in \mathscr{D}'(\Omega)$ *such that* $u|_\alpha = u_\alpha$ *for each* α.

Proof. a) and b) are trivial.

For c), let $\varphi \in \mathscr{D}(\Omega)$, and let $K = \operatorname{supp}\varphi$. Let $\{\chi_\alpha\}$ be a $\mathscr{D}(\Omega)$-partition of unity over K subordinate to $\{\omega_\alpha\}$. This means that

- $\chi_\alpha \in \mathscr{D}(\Omega)$ is nonnegative for each α,
- χ_α is nonzero for only finitely many α,
- there is an open set $V \supset K$ such that $\sum_\alpha \chi_\alpha = 1$ on V, and
- $\operatorname{supp}\chi_\alpha \subset \omega_\alpha$ for each α.

Note that we use the same index set for $\{\chi_\alpha\}$ as that of $\{\omega_\alpha\}$ at the expense of keeping some unnecessary zero functions in $\{\chi_\alpha\}$. We employ the existence of such a partition of unity without proof. We compute

$$\langle u,\varphi\rangle = \langle u, \sum_\alpha \chi_\alpha\varphi\rangle = \sum_\alpha \langle u,\chi_\alpha\varphi\rangle = \sum_\alpha \langle u|_{\omega_\alpha},\chi_\alpha\varphi\rangle = 0, \tag{82}$$

showing that $u = 0$, since $\varphi \in \mathscr{D}(\Omega)$ was arbitrary.

The uniqueness part of d) follows immediately from c). For existence, let $\varphi \in \mathscr{D}(\Omega)$, and keep the setting of the previous paragraph. We define

$$u(\varphi) := \sum_\alpha \langle u_\alpha,\chi_\alpha\varphi\rangle. \tag{83}$$

Before anything, we need to show that this definition does not depend on the partition of unity $\{\chi_\alpha\}$. Let $\{\xi_\alpha\}$ be another such partition of unity. Then we have

$$\sum_\alpha \langle u_\alpha,\chi_\alpha\varphi\rangle = \sum_{\alpha,\beta} \langle u_\alpha,\xi_\beta\chi_\alpha\varphi\rangle = \sum_{\alpha,\beta} \langle u_\beta,\xi_\beta\chi_\alpha\varphi\rangle = \sum_\beta \langle u_\beta,\xi_\beta\varphi\rangle, \tag{84}$$

where in the second step we used the property (81). Linearity can be verified for $\varphi_1,\varphi_2 \in \mathscr{D}(\Omega)$ by taking a partition of unity on $\operatorname{supp}\varphi_1 \cup \operatorname{supp}\varphi_2$. For continuity, let $K \subset \Omega$ be a compact set, and let $\varphi \in \mathscr{D}_K$. Then by using the fact that $u_\alpha \in \mathscr{D}'(\omega_\alpha)$ and $\chi_\alpha\varphi \in \mathscr{D}(\omega_\alpha)$, we have

$$|u(\varphi)| \leq \sum_\alpha |\langle u_\alpha,\chi_\alpha\varphi\rangle| \leq \sum_\alpha C_\alpha\|\chi_\alpha\varphi\|_{C^{m_\alpha}} \leq C\|\varphi\|_{C^m}, \tag{85}$$

showing that $u \in \mathscr{D}'(\Omega)$. □

Recall that the support of a continuous function is the closure of the set on which the function is nonzero. In other words, the support is the complement of the largest open set on which the function vanishes. This latter formulation makes sense even for distributions.

Definition 41. The *support* of $u \in \mathscr{D}'(\Omega)$ is given by

$$\operatorname{supp} u = \Omega \setminus \bigcup \{\omega \subset \Omega \text{ open} : u|_\omega = 0\}. \tag{86}$$

Lemma 42. *It is easy to check that the following properties hold.*

(a) $u|_{\Omega \setminus \operatorname{supp} u} = 0$.
(b) $x \in \operatorname{supp} u$ *iff* $x \in \Omega$ *and* x *does not have any open neighbourhood on which* u *vanishes.*
(c) $\operatorname{supp} u$ *agrees with the usual notion when* u *is a continuous function.*
(d) $\operatorname{supp} u$ *is relatively closed in* Ω.
(e) $\operatorname{supp} u = \varnothing$ *implies* $u = 0$.
(f) If $\rho \in C^\infty(\Omega)$ *is* $\rho \equiv 1$ *in a neighbourhood of* $\operatorname{supp} u$, *then* $\rho u = u$.
(g) $\operatorname{supp}(u + v) \subset \operatorname{supp} u \cup \operatorname{supp} v$.
(h) $\operatorname{supp}(au) \subset \operatorname{supp} a \cap \operatorname{supp} u$ *for* $a \in \mathscr{E}(\Omega)$ *and* $u \in \mathscr{D}'(\Omega)$.
(i) $\operatorname{supp} \partial^\alpha u \subset \operatorname{supp} u$.

Example 43. $\operatorname{supp} \delta = \{0\}$.

8. Compactly supported distributions

The following theorem characterizes compactly supported distributions.

Theorem 44. *We have* $\mathscr{E}'(\Omega) = \{u \in \mathscr{D}'(\Omega) : \operatorname{supp} u \Subset \Omega\}$, *i.e., the space* $\mathscr{E}'(\Omega)$ *is precisely the space of compactly supported distributions in* Ω. *Moreover,* $\mathscr{E}'(\Omega) = \bigcup_m \mathscr{E}'^m(\Omega)$, *meaning that any compactly supported distribution is of finite order.*

Proof. Recall that $\mathscr{E}(\Omega)$ is a Fréchet space with the seminorms

$$p_{m,K}(\varphi) = \|\varphi\|_{C^m(K)}, \qquad K \Subset \Omega,\ m \in \mathbb{N}_0. \tag{87}$$

Suppose that $u \in \mathscr{E}'(\Omega)$. This means that there exists $K \Subset \Omega$, $m \in \mathbb{N}_0$, and $C > 0$ such that

$$|u(\varphi)| \leq C\|\varphi\|_{C^m(K)}, \qquad \varphi \in \mathscr{E}(\Omega). \tag{88}$$

Hence $u(\varphi) = 0$ if $\operatorname{supp} \varphi \subset \Omega \setminus K$, i.e., $\operatorname{supp} u \subset K$. This bound also implies that $u : \mathscr{E}(\Omega) \to \mathbb{R}$ is continuous in the subspace topology induced by the (dense) embedding $\mathscr{E}(\Omega) \subset \mathscr{E}^m(\Omega)$. Therefore u has a unique continuous extension $u : \mathscr{E}^m(\Omega) \to \mathbb{R}$, establishing the second statement of the theorem.

Now suppose that $u \in \mathscr{D}'(\Omega)$ and that $K \equiv \operatorname{supp} u$ is compact. Let $\rho \in \mathscr{D}(\Omega)$ be such that $\rho \equiv 1$ in a neighbourhood of K. Then we have

$$u(\varphi) = u(\rho\varphi) + u(\varphi - \rho\varphi) = u(\rho\varphi), \qquad \varphi \in \mathscr{D}(\Omega), \tag{89}$$

because $\operatorname{supp}(\varphi - \rho\varphi) \subset \Omega \setminus K$. Since u is a distribution there exist $m \in \mathbb{N}_0$ and $C > 0$ such that

$$|u(\varphi)| = |u(\rho\varphi)| \leq C\|\rho\varphi\|_{C^m(K')} \leq C'\|\varphi\|_{C^m(K')}, \qquad \varphi \in \mathscr{D}(\Omega), \tag{90}$$

where $K' = \operatorname{supp}\rho$, i.e., the map $u : \mathscr{D}(\Omega) \to \mathbb{R}$ is continuous in the subspace topology induced by the (dense) embedding $\mathscr{D}(\Omega) \subset \mathscr{E}(\Omega)$. So u has a unique extension $u \in \mathscr{E}'(\Omega)$. $\square$

Exercise 45. Prove that $\mathscr{E}'^m(\Omega) = \{u \in \mathscr{D}'^m(\Omega) : \operatorname{supp} u \Subset \Omega\}$.

The preceding proof shows that for a compactly supported distribution u, the bound (90) holds with any compact K' that contains the support of u in its interior. The following example illustrates the interesting phenomenon that in general one cannot get the same bound with $K' = \operatorname{supp} u$. Informally speaking, this means that distributions can "feel" regions slightly outside of their support.

Example 46. Let $x_j = 2^{-j}$ for $j = 1, 2, \ldots$, and let $K = \{0, x_1, x_2, \ldots\}$. Obviously $K \subset \mathbb{R}$ is compact. Consider

$$u(\varphi) = \sum_j [\varphi(x_j) - \varphi(0)], \qquad \varphi \in \mathscr{E}(\mathbb{R}). \tag{91}$$

We have $u \in \mathscr{E}'(\mathbb{R})$, since

$$|u(\varphi)| = \sum_j |x_j| \|\varphi'\|_{C^0([0,1])} \leq \|\varphi'\|_{C^0([0,1])}, \qquad \varphi \in \mathscr{E}(\mathbb{R}). \tag{92}$$

We also have $\operatorname{supp} u = K$. Now, suppose that

$$|u(\varphi)| \leq C \sum_{k=0}^{m} \|\varphi^{(k)}\|_{C^0(K)}, \qquad \varphi \in \mathscr{E}(\mathbb{R}), \tag{93}$$

holds, with some constants $C > 0$ and m. Let us compute the both sides of the above inequality for $\varphi_j \in \mathscr{E}(\mathbb{R})$ satisfying $\varphi_j \equiv 1$ in a neighbourhood of $[x_j, x_1]$, and $\varphi_j \equiv 0$ in a neighbourhood of $[0, x_{j+1}]$. Since $\varphi_j(0) = 0$ we have

$$u(\varphi_j) = j. \tag{94}$$

On the other hand, all $\varphi_j^{(k)}$ except the case $k = 0$ vanish on K, and we get

$$\sum_{k=0}^{m} \|\varphi_j^{(k)}\|_{C^0(K)} = 1. \tag{95}$$

Therefore the bound (93) cannot hold.

Even though in general we cannot get a bound of the form (93) with $K = \operatorname{supp} u$, it is still true that $u(\varphi) = 0$ if sufficiently many derivatives of φ vanish on $\operatorname{supp} u$.

Theorem 47. *Let $u \in \mathscr{E}'^m(\Omega)$ and let $\varphi \in C^m(\Omega)$ with $\partial^\alpha \varphi = 0$ on $K = \operatorname{supp} u$ for $|\alpha| \leq m$. Then $u(\varphi) = 0$.*

Proof. Let $\varepsilon > 0$ be arbitrary, and let $\rho_\varepsilon \in \mathscr{D}(K + B_\varepsilon)$ be a cut-off function satisfying $\rho_\varepsilon \equiv 1$ in a neighbourhood of K and $\|\rho_\varepsilon\|_{C^k} \leq c\varepsilon^{-k}$ for all $k \leq m$, where $K + B_\varepsilon = \{x \in \mathbb{R}^n : \operatorname{dist}(x, K) < \varepsilon\}$ and $c > 0$ is a constant independent of ε. We have $u(\varphi) = u(\rho_\varepsilon \varphi)$, and hence

$$\begin{aligned} |u(\varphi)| &\leq c \sum_{|\alpha| \leq m} \|\partial^\alpha(\rho_\varepsilon \varphi)\|_{C^0} \leq c \sum_{|\alpha+\beta| \leq m} \|\partial^\beta \rho_\varepsilon\|_{C^0} \|\partial^\alpha \varphi\|_{C^0(K+B_\varepsilon)} \\ &\leq c \sum_{|\alpha| \leq m} \varepsilon^{|\alpha|-m} \|\partial^\alpha \varphi\|_{C^0(K+B_\varepsilon)}, \end{aligned} \tag{96}$$

where the constant $c > 0$ may have different values at its different occurrences. We have $\|\partial^\alpha \varphi\|_{C^0(K+B_\varepsilon)} \to 0$ as $\varepsilon \to 0$ for $|\alpha| \leq m$, because $\varphi \in C^m(\Omega)$. This means that in the sum on the right hand side of (96), the terms with $|\alpha| = m$ tend to 0 as $\varepsilon \to 0$. The remaining terms can be treated as follows. Let $y \in (K + B_\varepsilon) \setminus K$, and let $x \in K$ be such that $|x - y| < 2\varepsilon$. Fix some α with $|\alpha| < m$, and set $f(t) = (\partial^\alpha \varphi)(x + t(y - x))$. Then taking into account that $f^{(k)}(0) = 0$ for $k \leq m - |\alpha|$, from Cauchy's repeated integration formula we get

$$|\partial^\alpha \varphi(y)| \leq c \sup_{0<t<1} |f^{(m-|\alpha|)}(t)|. \tag{97}$$

Since $|x - y| < 2\varepsilon$, the chain rule gives

$$\sup_{0<t<1} |f^{(m-|\alpha|)}(t)| \leq c\varepsilon^{m-|\alpha|} \|\varphi\|_{C^m(K+B_\varepsilon)}, \tag{98}$$

confirming that the sum on the right hand side of (96) tends to 0 as $\varepsilon \to 0$. □

An easy application of this theorem shows that point-supported distributions are nothing but finite linear combinations of derivatives of the Dirac distribution.

Corollary 48. *Let $u \in \mathscr{E}'^m(\Omega)$ and let* $\operatorname{supp} u = \{0\}$*, where we assume $0 \in \Omega$. Then there exist coefficients $a_\alpha \in \mathbb{R}$, $|\alpha| \leq m$, such that*

$$u(\varphi) = \sum_{|\alpha|\leq m} a_\alpha \partial_\alpha \varphi(0), \qquad \varphi \in C^m(\Omega), \tag{99}$$

that is, $u = \sum (-1)^{|\alpha|} a_\alpha \partial^\alpha \delta$.

Proof. Let $\varphi \in C^m(\Omega)$, and let

$$\psi(x) = \varphi(x) - \sum_{|\alpha|\leq m} \frac{\partial^\alpha \varphi(0)}{\alpha!} x^\alpha, \qquad x \in \Omega. \tag{100}$$

We have $\partial^\alpha \psi(0) = 0$ for $|\alpha| \leq m$, and so the preceding theorem implies $u(\psi) = 0$. Consequently, we conclude

$$u(\varphi) = \sum_{|\alpha|\leq m} \frac{\partial^\alpha \varphi(0)}{\alpha!} u(p_\alpha) + u(\psi) = \sum_{|\alpha|\leq m} \frac{u(p_\alpha)}{\alpha!} \partial^\alpha \varphi(0), \tag{101}$$

where the functions $p_\alpha \in \mathscr{E}(\Omega)$ are defined by $p_\alpha(x) = x^\alpha$. □

9. Distributions are derivatives of functions

In this section, we will prove that every distribution is locally a (possibly high order) derivative of a function. This means that distributions are not much more than derivatives of functions. The heart of the matter is the following representation of compactly supported distributions as derivatives of functions.

Theorem 49. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set, and let $u \in \mathscr{E}'^m(\Omega)$. Then there exists $f \in L^\infty(\Omega)$ such that $u = \partial_1^{m+1} \ldots \partial_n^{m+1} f$.*

Proof. By definition, there exist $K \Subset \Omega$ and a constant $C > 0$ such that

$$|u(\varphi)| \leq C\|\varphi\|_{C^m(K)} = C \max_{|\alpha|\leq m} \|\partial^\alpha \varphi\|_{C^0(K)}, \qquad \varphi \in \mathscr{E}^m(\Omega). \tag{102}$$

For any $\psi \in \mathscr{D}(\Omega)$, we have

$$\|\psi\|_{C^0} \leq C\|\partial_j \psi\|_{C^0}, \tag{103}$$

with some constant $C > 0$, because Ω is bounded. Hence assuming that $\varphi \in \mathscr{D}(\Omega)$, we can replace the derivatives in the right hand side of (102) by higher order derivatives so as to have only one term in the maximum. This term would of course be the norm of $\partial^\beta \varphi$ with $\beta = (m, m, \ldots, m)$, i.e.,

$$|u(\varphi)| \leq C\|\partial^\beta \varphi\|_{C^0}, \qquad \varphi \in \mathscr{D}(\Omega). \tag{104}$$

We want to replace the uniform norm in the right hand side by the L^1-norm of a derivative of φ. For any $\psi \in \mathscr{D}(\Omega)$ and for $x \in \mathbb{R}^n$, we have

$$\psi(x) = \int_{y<x} \partial_1 \ldots \partial_n \psi(y)\,\mathrm{d}y, \tag{105}$$

where $y < x$ should be read componentwise. Using this, we finally get

$$|u(\varphi)| \leq C \int |\partial^\beta \varphi|, \qquad \varphi \in \mathscr{D}(\Omega), \tag{106}$$

now with $\beta = (m+1, m+1, \dots, m+1)$. This inequality in particular implies that the distribution u cannot distinguish two functions $\varphi, \psi \in \mathscr{D}(\Omega)$ if they satisfy $\partial^\beta \varphi = \partial^\beta \psi$. Therefore the map

$$T(\partial^\beta \varphi) := u(\varphi), \tag{107}$$

as a linear functional on the space $X = \{\partial^\beta \psi : \psi \in \mathscr{D}(\Omega)\}$, is well-defined. The following commutative diagram illustrates the setting.

$$\begin{array}{ccc} \mathscr{D}(\Omega) & \xrightarrow{u} & \mathbb{R} \\ {\scriptstyle \partial^\beta}\downarrow & \nearrow_{T} & \\ X & & \end{array} \tag{108}$$

Now the estimate (106) simply says that

$$|T(\xi)| \leq C\|\xi\|_{L^1(\omega)}, \qquad \xi \in X, \tag{109}$$

and so we can employ the Hahn-Banach theorem to extend T as a bounded linear functional on all of $L^1(\Omega)$. Hence by the duality between L^1 and L^∞, there is $g \in L^\infty(\Omega)$ such that

$$T(\xi) = \int g\xi, \qquad \xi \in L^1(\Omega). \tag{110}$$

Finally, putting $\xi = \partial^\beta \varphi$ with $\varphi \in \mathscr{D}(\Omega)$ and unraveling the definitions, we get

$$u(\varphi) = T(\partial^\beta \varphi) = \int g \partial^\beta \varphi = (-1)^{|\beta|}\langle \partial^\beta g, \varphi\rangle, \tag{111}$$

which means that $u = (-1)^{|\beta|}\partial^\beta g$ on Ω. □

It is not difficult to improve the preceding result so that one can represent any compactly supported distribution as a derivative of a *continuous* function.

Exercise 50. Let $\Omega \subset \mathbb{R}^n$ be a bounded open set, and let $u \in \mathscr{E}'^m(\Omega)$. Show that there exists $g \in C(\mathbb{R}^n)$ such that $u = \partial_1^{m+2} \dots \partial_n^{m+2} g$ in Ω.

Provided that we work locally, the same result can be established for arbitrary distributions, since one can turn an arbitrary distribution into a compactly supported one with the help of a cut-off function.

Exercise 51. Let $u \in \mathscr{D}'(\Omega)$. Show that for any open $\omega \subset \Omega$ with $\overline{\omega} \Subset \Omega$, there exists $g \in C(\mathbb{R}^n)$ and a multi-index α such that $u = \partial^\alpha g$ in ω.

By patching together series of local representations, we get a global representation, giving a precise meaning to the assertion that distributions are derivatives of functions.

Corollary 52. *Let $u \in \mathscr{D}'(\Omega)$. Then there exist a sequence $\{g_j\} \subset \mathscr{D}^0(\Omega)$ of functions, and a sequence $\{\alpha_j\}$ of multi-indices, such that $u = \sum_j \partial^{\alpha_j} g_j$, and that $\{\operatorname{supp} g_j\}$ is locally finite, i.e., any compact set $K \subset \Omega$ intersects with only a finitely many of the support sets $\operatorname{supp} g_j$. In addition, if $u \in \mathscr{D}'^m(\Omega)$ then we can take all α_j satisfying $|\alpha_j| \leq m$.*

Proof. Let $\{\omega_j\}$ and $\{\sigma_j\}$ be locally finite coverings of Ω by bounded open sets, such that

$$\Omega = \bigcup_j \omega_j, \qquad \Omega = \bigcup_j \sigma_j, \qquad \text{and} \qquad \bar{\sigma}_j \subset \omega_j \quad \text{for all } j. \tag{112}$$

Let $\{\rho_j\}$ be a partition of unity subordinate to the covering $\{\sigma_j\}$, and for each j, let $\psi_j \in \mathscr{D}(\omega_j)$ be a function satisfying $\psi_j \equiv 1$ on $\bar{\sigma}_1$. Then for any $\varphi \in \mathscr{D}(\Omega)$, we have

$$\langle u, \varphi\rangle = \sum_j \langle \rho_j u, \varphi\rangle = \sum_j \langle \rho_j u, \psi_j \varphi\rangle = \sum_j \int g_j \partial^{\alpha_j}(\psi_j \varphi_j), \tag{113}$$

where in the last step we employed the representation from Exercise 50 to $\rho_j u$. The function g_j can be chosen so that $\operatorname{supp} g_j \subset \omega_j$, because we can multiply it by a cut-off function $\chi \in \mathscr{D}(\omega_j)$ with $\chi_j \equiv 1$ in a neighbourhood of ω_j, and it would not affect the integral in (113). The proof is established, since any compact set $K \subset \Omega$ intersects only finitely many of ω_j. □

10. Complex valued distributions

Let X and Z be complex linear spaces. Recall that a map $A : X \to Z$ is called *linear* (or *complex linear*) if

$$A(\alpha x + y) = \alpha Ax + Ay, \qquad x, y \in X,\, \alpha \in \mathbb{C}, \tag{114}$$

and *anti-linear* (or *conjugate linear*) if

$$A(\alpha x + y) = \bar{\alpha} Ax + Ay, \qquad x, y \in X,\, \alpha \in \mathbb{C}. \tag{115}$$

If X is a complex linear space with a locally convex structure (or more generally, with a topology), then the space of all linear and continuous maps $u : X \to \mathbb{C}$ is called the *dual* of X, and denoted by X'. Moreover, the space of all anti-linear and continuous maps $u : X \to \mathbb{C}$ is called the *anti-dual* of X, and denoted by X'^*. Note that the dual and anti-dual are themselves complex linear spaces. There is a simple relation between these two spaces. For a map $f : X \to \mathbb{C}$, let us call $\bar{f} : X \to \mathbb{C}$ given by $\bar{f}(x) = \overline{f(x)}$ the *conjugate* of f. Then we have $\bar{u} \in X'^*$ for $u \in X'$ and vice versa. Moreover, since $\overline{\alpha u} = \bar{\alpha}\bar{u}$, the map $u \mapsto \bar{u} : X' \to X'^*$ is anti-linear, which makes it an anti-linear topological isomorphism between X' and X'^*.

Example 53. Let X be a complex linear space, and recall that an inner product on X is a function $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$ satisfying the following properties.

- $\langle x, x \rangle \geq 0$, and $\langle x, x \rangle = 0 \;\Rightarrow\; x = 0$. (Positive definiteness)
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$. (Hermitian property)
- The map $\langle \cdot, y \rangle : X \to \mathbb{C}$ is complex linear for any fixed $y \in X$.

The second and third properties imply that the inner product is anti-linear in its second argument. We equip X with the topology induced by the norm $\|\cdot\| = \langle \cdot, \cdot \rangle^{\frac{1}{2}}$, making X an inner product space. The following two maps arise naturally.

- $R : X \to X'$, defined by $(Ry)(x) = \langle x, y \rangle$.
- $S : X \to X'^*$, defined by $(Sx)(y) = \langle x, y \rangle$.

We see that R sends X into its dual, but R itself is anti-linear. We also see that S sends X into its anti-dual, but S itself is linear. Moreover, we have $Sx = \overline{Rx}$, because

$$(Sx)(y) = \langle x, y \rangle = \overline{\langle y, x \rangle} = \overline{(Rx)(y)}. \tag{116}$$

If X is complete, i.e., if X is a Hilbert space, by the Riesz representation theorem, both R and S are invertible. Hence in this case, R is an anti-linear topological isomorphism between X and X', and S is a topological isomorphism between X and X'^*.

A possible definition of complex valued distributions is to regard any continuous linear functional $u : \mathscr{D}(\Omega) \to \mathbb{C}$ as a complex valued distribution. Since $\mathscr{D}(\Omega)$ is only a real vector space, linearity here is understood as $\mathbb{R}$-linearity, and hence with this definition, a complex valued distribution is nothing but a pair of real valued distributions. Another possibility is to consider continuous (real) linear functionals $u : \mathscr{D}(\Omega, \mathbb{C}) \to \mathbb{R}$, where $\mathscr{D}(\Omega, \mathbb{C})$ is the set of complex valued test functions. This will not lead to anything new, since if we consider $\mathscr{D}(\Omega, \mathbb{C})$ as a real vector space, then $\mathscr{D}(\Omega, \mathbb{C}) = \mathscr{D}(\Omega) \oplus \mathscr{D}(\Omega)$.

More "complex" candidates for complex valued distributions are the complex dual $\mathscr{D}'(\Omega, \mathbb{C})$ of $\mathscr{D}(\Omega, \mathbb{C})$, and the anti-dual $\mathscr{D}'^*(\Omega, \mathbb{C})$ of $\mathscr{D}(\Omega, \mathbb{C})$. There are four natural ways to embed

$L^1_{\mathrm{loc}}(\Omega,\mathbb{C})$ into either $\mathscr{D}'(\Omega,\mathbb{C})$ or $\mathscr{D}'^*(\Omega,\mathbb{C})$, as follows.

$$\begin{aligned}
\langle j_0 u,\varphi\rangle &= \int u\varphi = \langle u,\bar\varphi\rangle_{L^2}, &\quad j_0 : L^1_{\mathrm{loc}}(\Omega,\mathbb{C}) \to \mathscr{D}'(\Omega,\mathbb{C}) \text{ linear},\\
\langle j_1 u,\varphi\rangle &= \int u\bar\varphi = \langle u,\varphi\rangle_{L^2}, &\quad j_1 : L^1_{\mathrm{loc}}(\Omega,\mathbb{C}) \to \mathscr{D}'^*(\Omega,\mathbb{C}) \text{ linear},\\
\langle j_2 u,\varphi\rangle &= \int \bar u\varphi = \langle \varphi,u\rangle_{L^2}, &\quad j_2 : L^1_{\mathrm{loc}}(\Omega,\mathbb{C}) \to \mathscr{D}'(\Omega,\mathbb{C}) \text{ anti-linear},\\
\langle j_3 u,\varphi\rangle &= \int \bar u\bar\varphi = \langle \bar u,\varphi\rangle_{L^2}, &\quad j_3 : L^1_{\mathrm{loc}}(\Omega,\mathbb{C}) \to \mathscr{D}'^*(\Omega,\mathbb{C}) \text{ anti-linear},
\end{aligned} \tag{117}$$

where $\langle\cdot,\cdot\rangle_{L^2}$ is the (Hermitian) inner product of $L^2(\Omega,\mathbb{C})$. Noting that $j_3u=\overline{j_0u}$, $j_2u=j_0\bar u$, and $j_1u=\overline{j_0\bar u}$, it takes a little bookkeeping to switch between the different conventions. For concreteness, in these notes, *we are going to consider $\mathscr{D}'(\Omega,\mathbb{C})$ as the space of complex valued distributions*, and stick to j_0 as the way to embed $L^1_{\mathrm{loc}}(\Omega,\mathbb{C})$ into $\mathscr{D}'(\Omega,\mathbb{C})$. One must be a bit careful *not* to write $u(\varphi)=\langle u,\varphi\rangle_{L^2}$, when regarding a function $u\in L^2(\Omega,\mathbb{C})$ as a distribution, since $(j_0u)(\varphi)=\langle u,\bar\varphi\rangle_{L^2}\neq\langle u,\varphi\rangle_{L^2}$ in general. If we want $u(\varphi)=\langle u,\varphi\rangle_{L^2}$ to be true, then we must either replace $\mathscr{D}'(\Omega,\mathbb{C})$ by $\mathscr{D}'^*(\Omega,\mathbb{C})$, as it would be the case for j_1, or give up on linearity of the embedding $L^1_{\mathrm{loc}}(\Omega,\mathbb{C})\to\mathscr{D}'(\Omega,\mathbb{C})$, as it would be the case for j_2.

11. The Fourier transform

For $u\in L^1$, we define its *Fourier transform* by

$$\hat u(\xi)=(\mathscr{F}u)(\xi)=\int e^{-i\xi\cdot x}u(x)\mathrm{d}x,\qquad \xi\in\mathbb{R}^n. \tag{118}$$

It is immediate that $\|\hat u\|_{L^\infty}\leq\|u\|_{L^1}$, and that

$$\hat u(\xi)=\frac{1}{i\xi_k}\int e^{-i\xi\cdot x}\partial_k u(x)\mathrm{d}x\to 0,\qquad \text{as } \xi_k\to\infty, \tag{119}$$

for $u\in\mathscr{D}$. Since $\mathscr{D}$ is dense in L^1, and C_0 is a closed subspace of L^∞, where C_0 is the space of continuous functions decaying at infinity, it follows that $\hat u\in C_0$ whenever $u\in L^1$. This is a variant of the *Riemann-Lebesgue lemma.*

If in addition $\hat u\in L^1$, then (as we will prove below) u can be recovered by

$$u(x)=\frac{1}{(2\pi)^n}\int e^{i\xi\cdot x}\hat u(\xi)\mathrm{d}\xi,\qquad x\in\mathbb{R}^n, \tag{120}$$

i.e., $\hat{\hat u}=(2\pi)^n\tilde u$, where $\tilde u$ is the reflection $\tilde u(x)=u(-x)$.

Definition 54. We define the *Schwartz class* to be

$$\mathscr{S}=\{\psi\in C^\infty(\mathbb{R}^n):p_{\alpha,\beta}(\psi)<\infty\,\forall\alpha,\beta\}, \tag{121}$$

where

$$p_{\alpha,\beta}(\psi)=\sup_{x\in\mathbb{R}^n}|x^\alpha\partial^\beta\psi(x)|, \tag{122}$$

are the seminorms with which we equip $\mathscr{S}$.

It is immediate that differentiation and multiplication by polynomials are continuous operations in $\mathscr{S}$.

Exercise 55. Show that $\mathscr{S}$ is a Fréchet space, and that $\mathscr{D}\subset\mathscr{S}\subset L^1$ where all inclusions are continuous, proper, and dense.

Theorem 56. *The Fourier transform $\mathscr{F}:\mathscr{S}\to\mathscr{S}$ is a topological isomorphism with its inverse given by $\mathscr{F}^{-1}\psi=(2\pi)^{-n}\widetilde{\mathscr{F}\psi}$. Moreover, for $\phi,\psi\in\mathscr{S}$ we have*

(a) $\int \hat{\phi}\psi = \int \hat{\psi}\phi$, $\int \phi\psi = (2\pi)^{-n} \int \hat{\phi}\tilde{\hat{\psi}}$ *and* $\int \phi\bar{\psi} = (2\pi)^{-n} \int \hat{\phi}\bar{\hat{\psi}}$ *(Parseval's identity).*
(b) $\widehat{\phi * \psi} = \hat{\phi}\hat{\psi}$ *and* $\widehat{\phi\psi} = (2\pi)^{-n}\hat{\phi} * \hat{\psi}$.
(c) $\widehat{\partial^\alpha \phi} = (i\xi)^\alpha \hat{\phi}$ *and* $\widehat{x^\alpha \psi} = (i\partial)^\alpha \hat{\psi}$.
(d) $\widehat{\tau_a \phi} = e^{-ia\xi}\hat{\phi}$ *and* $\widehat{e^{iax}\psi} = \tau_a \hat{\psi}$.
(e) $\hat{\hat{\phi}} = \tilde{\hat{\phi}}$ *and* $\hat{\bar{\psi}} = \bar{\hat{\psi}}$.

Proof. Let $\varphi \in \mathscr{S}$. Then by differentiation of (118) we obtain

$$\partial^\alpha \hat{\varphi}(\xi) = \int e^{-i\xi\cdot x}(-ix)^\alpha \varphi(x)\mathrm{d}x, \qquad \xi \in \mathbb{R}^n. \tag{123}$$

This shows that $\hat{\varphi} \in C^\infty(\mathbb{R}^n)$ and that $\partial^\alpha \hat{\varphi} = (-i)^{|\alpha|}\widehat{x^\alpha \varphi}$. Putting $\psi = (-ix)^\alpha \varphi \in \mathscr{S}$, integration by parts gives

$$\xi^\beta \partial^\alpha \hat{\varphi}(\xi) = (-i)^{|\beta|} \int e^{-i\xi\cdot x}\partial^\beta \psi(x)\mathrm{d}x, \qquad \xi \in \mathbb{R}^n. \tag{124}$$

Since $\partial^\beta \psi \in \mathscr{S}$, we have $\hat{\varphi} \in \mathscr{S}$. Also, the case $\alpha = 0$ gives $\xi^\beta \hat{\varphi}(\xi) = (-i)^{|\beta|}\widehat{\partial^\beta \varphi}$.

Continuity of $\mathscr{F} : \mathscr{S} \to \mathscr{S}$ follows from

$$\begin{aligned} |\xi^\beta \partial^\alpha \hat{\varphi}(\xi)| &\leq \int (1 + |x|)^{-n-1}(1 + |x|)^{n+1}|\partial^\beta \psi(x)|\mathrm{d}x \\ &\leq C \sup_{x\in\mathbb{R}^n} (1 + |x|)^{n+1}|\partial^\beta (x^\alpha \varphi(x))|. \end{aligned} \tag{125}$$

For the inversion formula, we need to justify the formal manipulation

$$\int e^{i\xi\cdot x}\hat{\varphi}(\xi)\mathrm{d}\xi = \int e^{i\xi\cdot x}\Big(\int e^{-i\xi\cdot y}\varphi(y)\mathrm{d}y\Big)\mathrm{d}\xi = \int \varphi(y)\Big(\int e^{-i\xi\cdot(y-x)}\mathrm{d}\xi\Big)\mathrm{d}y = c\varphi(x), \tag{126}$$

where c is some constant. Switching the integrals would cause no trouble if we introduce the factor $\psi \in \mathscr{S}$ that forces decay in the ξ-direction:

$$\begin{aligned} \int e^{i\xi\cdot x}\hat{\varphi}(\xi)\psi(\xi)\mathrm{d}\xi &= \int e^{i\xi\cdot x}\psi(\xi)\Big(\int e^{-i\xi\cdot y}\varphi(y)\mathrm{d}y\Big)\mathrm{d}\xi \\ &= \int \varphi(y)\Big(\int e^{-i\xi\cdot(y-x)}\psi(\xi)\mathrm{d}\xi\Big)\mathrm{d}y \\ &= \int \varphi(y)\hat{\psi}(y - x)\mathrm{d}y \\ &= \int \hat{\psi}(y)\varphi(x + y)\mathrm{d}y. \end{aligned} \tag{127}$$

The particular case $x = 0$ gives the identity $\int \hat{\varphi}\psi = \int \hat{\psi}\varphi$. In order to eliminate ψ from the above formula, we would like to take a sequence of ψ converging to a constant function, which we implement by replacing $\psi(\xi)$ by $\psi(\varepsilon\xi)$ with $\varepsilon > 0$ small. Noting the simple scaling law $\hat{\psi}_\varepsilon(\xi) = \varepsilon^{-n}\hat{\psi}(\xi/\varepsilon)$ where $\psi_\varepsilon(x) = \psi(\varepsilon x)$, we infer

$$\int \hat{\varphi}(\xi)\psi(\varepsilon\xi)e^{i\xi\cdot x}\mathrm{d}\xi = \int \varepsilon^{-n}\hat{\psi}(y/\varepsilon)\varphi(x + y)\mathrm{d}y = \int \hat{\psi}(y)\varphi(x + \varepsilon y)\mathrm{d}y. \tag{128}$$

Now we take the limit $\varepsilon \to 0$, and get

$$\psi(0)\int \hat{\varphi}(\xi)e^{i\xi\cdot x}\mathrm{d}\xi = \varphi(x)\int \hat{\psi}(y)\mathrm{d}y, \tag{129}$$

by dominated convergence. The constant $(2\pi)^{-n}$ can be verified by taking ψ to be a function whose Fourier transform is easily computable, e.g., the Gaussian $\psi(x) = e^{-|x|^2}$ will do.

The identity (127) can be written as $\widehat{\varphi\psi} = \hat{\psi} * \tilde{\varphi}$. Then the substitution $\varphi = \hat{\hat{\phi}}$ leads to $(2\pi)^n\widehat{\phi\psi} = \hat{\psi} * \hat{\phi}$. We leave the rest of the theorem as an exercise. □

Corollary 57. *Let $u \in \mathscr{S}$. Then the maps $\phi \mapsto u\phi : \mathscr{S} \to \mathscr{S}$ and $\phi \mapsto u * \phi : \mathscr{S} \to \mathscr{S}$ are continuous.*

Proof. It is easy to see continuity of multiplication. Then continuity of convolution follows from the formula $\widehat{u * \phi} = \hat{u}\hat{\phi}$. □

Remark 58. The Parseval's identity implies the *Plancherel formula*: $\|\hat{u}\|_{L^2} = (2\pi)^{n/2}\|u\|_{L^2}$. Interpolating this with the bound $\|\hat{u}\|_{L^\infty} \leq \|u\|_{L^1}$ gives the *Hausdorff-Young inequality*

$$\|\hat{u}\|_{L^q} \leq (2\pi)^{n/q}\|u\|_{L^p}, \tag{130}$$

which is valid for $1 \leq p \leq 2$ and $\frac{1}{p} + \frac{1}{q} = 1$. The Fourier transform of a function $u \in L^p$ with $p > 2$ is in general not a locally integrable function.

12. Tempered distributions

Definition 59. The topological dual $\mathscr{S}'$ of $\mathscr{S}$ is called the *space of tempered distributions*. We equip $\mathscr{S}'$ with its weak dual topology.

Remark 60. Since $\mathscr{D} \hookrightarrow \mathscr{S} \hookrightarrow \mathscr{E}$ with dense embeddings, we have $\mathscr{E}' \hookrightarrow \mathscr{S}' \hookrightarrow \mathscr{D}'$.

Definition 61. The Fourier transform of $u \in \mathscr{S}'$ is defined by

$$\langle \hat{u}, \varphi\rangle = \langle u, \hat{\varphi}\rangle, \qquad \varphi \in \mathscr{S}. \tag{131}$$

This definition clearly gives $\hat{u} \in \mathscr{S}'$, as the map $\varphi \mapsto \hat{\varphi}$ is continuous in $\mathscr{S}$.

Example 62. For $\varphi \in \mathscr{S}$, we have

$$\langle \hat{\delta}, \varphi\rangle = \langle \delta, \hat{\varphi}\rangle = \hat{\varphi}(0) = \int \varphi = \langle 1, \varphi\rangle, \tag{132}$$

hence $\hat{\delta} = 1$. On the other hand, from

$$\langle \hat{1}, \varphi\rangle = \langle 1, \hat{\varphi}\rangle = \int \hat{\varphi} = \hat{\hat{\varphi}}(0) = (2\pi)^n\varphi(0), \tag{133}$$

we infer $\hat{1} = (2\pi)^n\delta$ (This assertion is equivalent to the inversion formula).

Theorem 63. *The Fourier transform is a topological automorphism in $\mathscr{S}'$, and satisfies*
(a) $\hat{\hat{u}} = (2\pi)^n\tilde{u}$,
(b) $\widehat{\partial^\alpha u} = (i\xi)^\alpha\hat{u}$,
(c) $\widehat{x^\alpha u} = (i\partial)^\alpha\hat{u}$.

Proof. Continuity of the Fourier transform is obvious from

$$p_\varphi(u) = |\langle \hat{u}, \varphi\rangle| = |\langle u, \hat{\varphi}\rangle| = p_{\hat{\varphi}}(u). \tag{134}$$

The other two assertions are also straightforward. For instance, we have

$$\langle \widehat{\partial^\alpha u}, \varphi\rangle = \langle \partial^\alpha u, \hat{\varphi}\rangle = (-1)^{|\alpha|}\langle u, \partial^\alpha\hat{\varphi}\rangle = (-1)^{|\alpha|}\langle u, \partial^\alpha\widehat{(-i\xi)^\alpha\varphi}\rangle = \langle \hat{u}, \partial^\alpha(i\xi)^\alpha\varphi\rangle, \tag{135}$$

implying (b). □

We end these notes by proving a structure theorem for tempered distributions.

Theorem 64. *The following are equivalent.*
(a) $u \in \mathscr{S}'$.

(b) *There exists a finite sequence* $\{g_\alpha\}$ *of continuous functions satisfying* $|g_\alpha(x)| \leq C(1+|x|^m)$, *such that* $u = \sum \partial^\alpha g_\alpha$.

Proof. The implication (b) $\implies$ (a) is obvious. To prove (a) $\implies$ (b), let $u \in \mathscr{S}'$. Then there exist k and m such that

$$|\langle u, \varphi\rangle| \leq C \max_{|\alpha|\leq m} \sup_{x\in\mathbb{R}^n} (1+|x|^2)^k |\partial^\alpha \varphi(x)|, \qquad \varphi \in \mathscr{S}. \tag{136}$$

Introducing $\varphi_k(x) = (1+|x|^2)^k \varphi(x)$, we can derive

$$|\partial^\alpha \varphi| \leq C_\alpha (1+|x|^2)^{-k} \sum_{\beta\leq\alpha} |\partial^\beta \varphi|, \tag{137}$$

and hence

$$|\langle u, \varphi\rangle| \leq C \max_{|\alpha|\leq m} \sup_{x\in\mathbb{R}^n} |\partial^\alpha \varphi_k(x)|, \qquad \varphi \in \mathscr{S}. \tag{138}$$

Now, as in the proof of Theorem 49 the supremum can be replaced by an L^1-norm:

$$|\langle u, \varphi\rangle| \leq C \max_{|\alpha|\leq m+n} \|\partial^\alpha \varphi_k(x)\|_{L^1(\mathbb{R}^n)}, \qquad \varphi \in \mathscr{S}. \tag{139}$$

Therefore the map

$$T(\{\partial^\alpha \varphi_k : |\alpha| \leq m+n\}) := u(\varphi), \tag{140}$$

as a linear functional on the space $X = \{(\partial^\alpha \psi)_{|\alpha|\leq m+n} : \psi \in \mathscr{S}\}$, is well-defined. Then the estimate (139) says that

$$|T(\xi)| \leq C \max_{|\alpha|\leq m+n} \|\xi_\alpha\|_{L^1(\mathbb{R}^n)}, \qquad \xi \in X, \tag{141}$$

and so the Hahn-Banach theorem guarantees an extension of T as a bounded linear functional on $[L^1(\mathbb{R}^n)]^N$, where $N = \#\{\alpha : |\alpha| \leq m+n\}$. Hence by duality, there is $g_\alpha \in L^\infty(\mathbb{R}^n)$ for each α with $|\alpha| \leq m+n$, such that

$$T(\xi) = \sum_{|\alpha|\leq m+n} \int g_\alpha \xi_\alpha, \qquad \xi \in [L^1(\mathbb{R}^n)]^N. \tag{142}$$

Finally, putting $\xi_\alpha = \partial^\alpha \varphi_k = \partial^\alpha\big((1+|x|^2)^k \varphi\big)$ with $\varphi \in \mathscr{S}$, we get

$$\langle u, \varphi\rangle = T(\{\partial^\alpha \varphi_k\}) = \sum_\alpha \int g_\alpha \partial^\alpha \varphi_k = \sum_\alpha \int g_\alpha \partial^\alpha((1+|x|^2)^k \varphi) \tag{143}$$

which can be further processed to yield

$$\langle u, \varphi\rangle = \sum_{\alpha,\beta} \int c_{\alpha\beta} g_\alpha x^\beta \partial^\beta \varphi = \sum_{\alpha,\beta} (-1)^{|\beta|} c_{\alpha\beta} \big\langle \partial^\beta (x^\beta g_\alpha), \varphi\big\rangle, \tag{144}$$

with some coefficients $c_{\alpha\beta}$, $|\alpha| \leq m+n$, $|\beta| \leq 2k$. It is straightforward to turn $x^\beta g_\alpha$ into a continuous function of polynomial growth by integration. □