INTRODUCTION TO DISTRIBUTIONS

TSOGTGEREL GANTUMUR

Abstract. We introduce locally convex spaces by the seminorms approach, and present the fundamentals of distributions.

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 may not be 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.

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

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

Date: March 7, 2014.

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 $\mathscr{C}_c^k(\mathbb{R})$. Moreover, this specifies u uniquely, meaning that if there is some $v \in \mathscr{C}(\mathbb{R})$ such that $T_u(\varphi) = T_v(\varphi)$ for all $\varphi \in \mathscr{C}_c^k(\mathbb{R})$, then $u = v$. If we replace the space $\mathscr{C}(\mathbb{R})$ 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 $\mathscr{C}_c^k(\mathbb{R})$. 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

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

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

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

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

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

If we want to get more derivatives of T, we need k to be large, which leads us to consider the space $\mathscr{C}_c^\infty(\mathbb{R})$ 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 $\mathscr{C}_c^\infty(\mathbb{R})$, 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

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

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

Definition 3. Given a seminorm p on a vector space X, we call the set

$$B_{y,\varepsilon}(p) = \{x \in X : p(x-y) < \varepsilon\}, \tag{7}$$

the *p-semiball* of radius $\varepsilon > 0$ centred at $y \in X$. Moreover, given a finite collection of seminorms $p_1, \ldots, p_k$, let us call the set

$$B_{y,\varepsilon}(p_1, \ldots, p_k) = B_{y,\varepsilon}(p_1) \cap \ldots \cap B_{y,\varepsilon}(p_k), \tag{8}$$

the $(p_1, \ldots, p_k)$*-semiball* of radius $\varepsilon > 0$ centred at $y \in X$.

Exercise 1. Let $B = B_{0,\varepsilon}(p)$ be a semiball in a vector space X. Show that

a) B is convex.
b) B is balanced, in the sense that if $x \in B$ then $\lambda x \in B$ for all $\lambda \in [-1, 1]$.
c) B is absorbing, in the sense that for any $x \in X$, there is $\lambda \in \mathbb{R}$ such that $x \in \lambda B$.

Definition 4. 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 subsets of X which are the unions of semiballs of the form (8), that is, the latter group of semiballs form a base of the topology. 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 that the topology of X is the coarsest (i.e., smallest) topology containing all semiballs of the form (7). In other words, the semiballs of the form (7) are a subbase of the topology.

Remark 5. If p_1 and p_2 are seminorms, so is the function $x \mapsto \max\{p_1(x), p_2(x)\}$. Since

$$B_{y,\varepsilon}(p_1, \ldots, p_k) = B_{y,\varepsilon}(\max\{p_1, \ldots, p_k\}), \tag{9}$$

we see that if we define

$$\bar{\mathscr{P}} = \mathscr{P} \cup \{\max\{p_1, \ldots, p_k\} : p_1, \ldots, p_k \in \mathscr{P},\, k \in \mathbb{N}\}, \tag{10}$$

then we have $(X, \mathscr{P}) = (X, \bar{\mathscr{P}})$, and moreover, that the open subsets of X are nothing but the unions of the semiballs of the form (7) with $p \in \bar{\mathscr{P}}$, i.e., the collection of semiballs $\{B_{y,\varepsilon}(p) : y \in X,\, \varepsilon > 0,\, p \in \bar{\mathscr{P}}\}$ is a base for the topological space $(X, \mathscr{P})$. Therefore, it is easier to think about the topological properties of $(X, \mathscr{P})$ in terms of the seminorms $\bar{\mathscr{P}}$, because it frees one from having to form the finite intersections (8) "on the go".

Remark 6. Let $(X, \mathscr{P})$ be a locally convex space and let $p \in \mathscr{P}$. Suppose that p' is a seminorm on X, with $p' \notin \mathscr{P}$ and

$$p'(x) \leq cp(x), \qquad x \in X, \tag{11}$$

for some constant $c > 0$. Then $B_{y,\varepsilon}(p) \subset B_{y,c\varepsilon}(p')$, and hence $(X, \mathscr{P}) = (X, \mathscr{P}')$, where $\mathscr{P}' = \mathscr{P} \cup \{p'\}$. Thinking of $\mathscr{P}'$ as a family of seminorms that are given initially, this result can be used to remove redundant seminorms from $\mathscr{P}'$ to "clean it up".

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, i.e., if there is a base of the topology consisting of convex sets. 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 fact that any locally convex topological vector space has a family of seminorms that induces its topology (A proof of this fact can be found in Walter Rudin's *Functional analysis*).

Remark 7. 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 $p \in \mathscr{P}$.

In the following lemma, without loss of generality, we assume that $\mathscr{P} = \bar{\mathscr{P}}$.

Lemma 8. *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 seminorm $p \in \mathscr{P}$ and a number $\delta > 0$ such that*

$$z \in B_{x,\delta}(p) \quad \Rightarrow \quad \|f(x) - f(z)\|_Y \leq \varepsilon. \tag{12}$$

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

$$\|f(x)\|_Y \leq Cp(x), \qquad x \in X. \tag{13}$$

c) A seminorm $q : X \to \mathbb{R}$ (not necessarily $q \in \mathscr{P}$) is continuous if and only if there is a seminorm $p \in \mathscr{P}$ and a constant $C > 0$ such that

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

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 $B_{x,\delta}(p)$ 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 $B_{x,\delta}(p)$ such that $f(B_{x,\delta}(p)) \subset B_{y,\varepsilon}$. This means that $f^{-1}(U)$ is open.

For b), the condition associated to (13) immediately implies the condition associated to (12) by linearity. Now suppose that we have the condition associated to (12). Hence there is $\delta > 0$ and $p \in \mathscr{P}$ such that

$$z \in B_{0,\delta}(p) \quad \Rightarrow \quad \|f(z)\|_Y \leq 1. \tag{15}$$

Let $x \in X$, and define $z = \frac{\delta}{2p(x)}x$. Then we have $p(z) = \frac{\delta}{2} < \delta$, leading to

$$1 \geq \|f(z)\|_Y = \frac{\delta}{2p(x)}\|f(x)\|_Y, \tag{16}$$

which is (13) with $C = \frac{2}{\delta}$.

Part c) is completely analogous and follows from positive homogeneity of seminorms and the property in Lemma 2c). □

Remark 9. The preceding lemma can easily be extended to the case where Y is a locally convex space 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 seminorm $p \in \mathscr{P}$ and a constant $C > 0$ such that

$$q(f(x)) \leq Cp(x), \qquad x \in X. \tag{17}$$

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 seminorm $p \in \mathscr{P}$ and a constant $C > 0$ such that

$$q(x) \leq Cp(x), \qquad x \in X. \tag{18}$$

By part c) of the preceding lemma, 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 10. 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 11. 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.

Recall that a *metric* on a set X is a function $\rho : X \times X \to \mathbb{R}$ that is symmetric: $\rho(x, y) = \rho(y, x)$, nonnegative: $\rho(x, y) \geq 0$, nondegenerate: $\rho(x, y) = 0 \Leftrightarrow x = y$, and satisfies the triangle inequality: $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$. We call the pair (X, ρ) a *metric space*. In the special case where the underlying set X is a vector space, if the metric satisfies $\rho(x + z, y + z) = \rho(x, y)$, then the metric is said to be *translation invariant*. The *metric balls* $B_{y,\varepsilon} = \{x \in X : \rho(x, y) < r\}$ induce a topology on the set X, meaning that metric spaces are topological spaces. In the converse direction, if a topological space X can be given a metric so that the original topology coincides with the topology induced by the metric, then we say that X is *metrizable*.

Lemma 12. *A locally convex space $(X, \mathscr{P})$ is metrizable with a translation invariant metric 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 the translation invariant metric

$$d(x, y) = \max_k \frac{\alpha_k p_k(x - y)}{1 + p_k(x - y)}, \tag{19}$$

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

$$a \leq b + c \qquad \Rightarrow \qquad \frac{a}{1 + a} \leq \frac{b}{1 + b} + \frac{c}{1 + c} \qquad (a, b, c \geq 0), \tag{20}$$

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

For each k, we have

$$\frac{\alpha_k p_k(x - y)}{1 + p_k(x - y)} \leq d(x, y), \tag{21}$$

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

$$d(x, y) \leq \varepsilon + \max_{1\leq k\leq n} \frac{\alpha_k p_k(x - y)}{1 + p_k(x - y)} \leq \varepsilon + \alpha \max_{1\leq k\leq n} p_k(x - y), \tag{22}$$

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 13. 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 14. 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

$$x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}, \qquad \text{and} \qquad \partial^\alpha = \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n}, \tag{23}$$

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 $\mathscr{C}(A)$ be the space of continuous functions on A. That is, $u \in \mathscr{C}(A)$ iff

$$u(x) = \lim_{A \ni y \to x} u(y), \qquad \text{for all } x \in A. \tag{24}$$

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

$$\mathscr{C}^\infty(A) = \bigcap_m \mathscr{C}^m(A). \tag{25}$$

Remark 15. An often used alternative notation is $\mathscr{E}(\Omega) = \mathscr{C}^\infty(\Omega)$ and $\mathscr{E}^m(\Omega) = \mathscr{C}^m(\Omega)$.

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

$$\|u\|_{\mathscr{C}(K)} = \sup_{x \in K} |u(x)|, \qquad u \in \mathscr{C}(K). \tag{26}$$

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

$$\|u\|_{\mathscr{C}^m(K)} = \max_{|\alpha| \leq m} \|\partial^\alpha u\|_{\mathscr{C}^0(K)}, \qquad u \in \mathscr{C}^m(K). \tag{27}$$

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

$$p_m(\varphi) = \|\varphi\|_{\mathscr{C}^m(K)}, \qquad m = 0, 1, \ldots, \qquad u \in \mathscr{C}^\infty(K). \tag{28}$$

Lemma 16. *The space $\mathscr{C}^\infty(K)$ is metrizable and complete, i.e., it is a Fréchet space.*

Proof. The space $\mathscr{C}^\infty(K)$ is metrizable by Lemma 12, since $\{p_m\}$ is countable and separating. Let $\{\varphi_k\}$ be a Cauchy sequence in $\mathscr{C}^\infty(K)$. This means that $\{\varphi_k\}$ is Cauchy in $\mathscr{C}^m(K)$ for each m. Hence by completeness of $\mathscr{C}^m(K)$, for each m there exists $\psi_m \in \mathscr{C}^m(K)$ such that $p_m(\varphi_k - \psi_m) \to 0$ as $k \to \infty$. Since $p_0(\psi) \leq p_m(\psi)$ for any $\psi \in \mathscr{C}^m(K)$, we have $\psi_m = \psi_0$ for any m. We conclude that $\psi_0 \in \mathscr{C}^\infty(K)$ and that $\varphi_k \to \psi_0$ in $\mathscr{C}^\infty(K)$ 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 $\mathscr{C}(\Omega)$ with the topology of locally uniform convergence, i.e., the topology induced by the seminorms

$$p_K(\varphi) = \|\varphi\|_{\mathscr{C}(K)}, \tag{29}$$

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

$$K_j = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) \geq \frac{1}{j}\} \cap \overline{B}_j, \tag{30}$$

where

$$B_j = \{x \in \mathbb{R}^n : |x| < j\}, \tag{31}$$

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 $\varphi \in \mathscr{C}(\Omega)$, which means, in light of Remark 6, that we can use the family $\{p_{K_j} : j = 1, 2, \ldots\}$ to generate the topology of $\mathscr{C}(\Omega)$. This family is countable and separating, and metrizability follows. Introducing the seminorms

$$p_{j,k}(\varphi) = \|\varphi\|_{\mathscr{C}^k(K_j)}, \tag{32}$$

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

Lemma 17. *Let $0 \leq m \leq \infty$. Then $\mathscr{C}^m(\Omega)$ is a Fréchet space.*

Proof. The proof is similar to the proof of Lemma 16. Let $\{\varphi_i\}$ be a Cauchy sequence in $\mathscr{C}^m(\Omega)$. Then by completeness of $\mathscr{C}^m(K_j)$, for each j there exists $\psi_j \in \mathscr{C}^m(K_j)$ such that $\varphi_i \to \psi_j$ in $\mathscr{C}^m(K_j)$ as $i \to \infty$. Since $p_{j,k}(\psi) \leq p_{j+1,k}(\psi)$ for any $\psi \in \mathscr{C}^k(K_{j+1})$ 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 $\psi \in \mathscr{C}^m(\Omega)$ 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 $\mathscr{C}^m(\Omega)$, we can introduce local versions of L^p-spaces, as

$$L^p_{\text{loc}}(\Omega) = \{u : \Omega \to \mathbb{R} \text{ measurable}, u|_K \in L^p(K) \text{ for any compact } K \subset \Omega\}. \tag{33}$$

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

$$q_j(\varphi) = \|\varphi\|_{L^p(K_j)}. \tag{34}$$

Relying on the completeness of $L^p(K)$, one can easily show that $L^p_{\text{loc}}(\Omega)$ 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

$$\operatorname{supp} \varphi = \overline{\{x \in \Omega : \varphi(x) \neq 0\}}. \tag{35}$$

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

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

and endow it with the seminorms

$$p_m(\varphi) = \|\varphi\|_{\mathscr{C}^m}, \qquad m = 0, 1, \ldots. \tag{37}$$

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

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

$$\varphi(x) = \begin{cases} e^{-1/(1-|x|^2)} & \text{for } |x| < 1, \\ 0 & \text{for } |x| \geq 1, \end{cases} \tag{38}$$

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

$$\partial^\alpha \varphi(x) = \frac{p(x)e^{-1/(1-|x|^2)}}{(1-|x|^2)^{|\alpha|}}, \qquad |x| < 1, \tag{39}$$

where p is some polynomial. From this it is also clear that $\partial^\alpha\varphi(x) \to 0$ as $|x| \nearrow 1$. So $\varphi \in \mathscr{C}^\infty(\mathbb{R}^n)$. 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 $\mathscr{C}^\infty(K)$.

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

$$\mathscr{D}_K^m = \{\varphi \in \mathscr{C}^m(\mathbb{R}^n) : \operatorname{supp}\varphi \subset K\}, \tag{40}$$

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

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

$$\mathscr{D}(\Omega) = \bigcup_{K \Subset \Omega} \mathscr{D}_K, \tag{41}$$

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

$$\mathscr{D}(\Omega) = \bigcup_m \mathscr{D}_{K_m}. \tag{42}$$

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 9 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 19. $\mathscr{D}(\Omega)$ 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

$$\varphi_k(x) = \varphi(x) + 2^{-1}\varphi(x-1) + \ldots + 2^{-k}\varphi(x-k), \qquad k = 1, 2, \ldots. \tag{43}$$

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 20. 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 21. 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 22. *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.,

$$B_{\psi,\varepsilon}(p_m; \mathscr{D}_K) = B_{\psi,\varepsilon}(p_m) \cap \mathscr{D}_K. \tag{44}$$

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 $\mathscr{D}_K$. $\square$

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

$$p(\varphi) = \sup_j c_j |\varphi(x_j)|, \qquad \varphi \in \mathscr{D}(\Omega), \tag{45}$$

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 23. *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

$$p(\varphi) \leq Cp_m(\varphi), \qquad \varphi \in \mathscr{D}_K. \tag{46}$$

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

$$p(\varphi) = \sup_m \frac{m|\varphi(x_m)|}{|\varphi_m(x_m)|}, \qquad \varphi \in \mathscr{D}(\Omega). \tag{47}$$

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

$$p(\varphi) \leq C\|\varphi\|_{\mathscr{C}^0}, \qquad \varphi \in \mathscr{D}_{K'}. \tag{48}$$

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

Corollary 24. *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 $\|\varphi_j - \varphi_k\|_{\mathscr{C}^m} \to 0$ 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 $\|\varphi_j\|_{\mathscr{C}^m} \to 0$ 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 25. *Let $(Y, \mathscr{Q})$ be a locally convex space, and let $f : \mathscr{D}(\Omega) \to Y$ be a linear map. Then the followings 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

$$q(f(\varphi)) \leq p(\varphi), \qquad \varphi \in \mathscr{D}(\Omega). \tag{49}$$

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 (49) holds. Given q, let us define the function

$$p(\varphi) = q(f(\varphi)), \qquad \varphi \in \mathscr{D}(\Omega). \tag{50}$$

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

$$q(f(\varphi)) \leq C p_m(\varphi), \qquad \varphi \in \mathscr{D}_K, \tag{51}$$

for some C and m possibly depending on K. Therefore $p \in \mathscr{P}$, which clearly implies (49). □

Example 26. 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

$$p_m(\partial_j \varphi) \leq p_{m+1}(\varphi), \qquad \varphi \in \mathscr{D}_K. \tag{52}$$

Remark 27. $\mathscr{D}(\Omega)$ 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

$$\varphi_{km}(x) = \frac{1}{m} \varphi\left(\frac{x}{k}\right), \qquad k, m = 1, 2, \ldots. \tag{53}$$

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

$$\mathscr{D}^m(\Omega) = \bigcup_{K \Subset \Omega} \mathscr{D}^m_K, \tag{54}$$

and the space of compactly supported L^p-functions

$$L^p_{\text{comp}}(\Omega) = \bigcup_{K \Subset \Omega} L^p(K), \tag{55}$$

where in the right hand side, the elements of $L^p(K)$ are extended by zero outside K. Then 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 29. 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 25 tailored to distributions is the following.

Lemma 30. *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

$$|u(\varphi)| \leq C\|\varphi\|_{\mathscr{C}^m} \qquad \textit{for} \quad \varphi \in \mathscr{D}_K. \tag{56}$$

Definition 31. Let $u \in \mathscr{D}'(\Omega)$. If we have

$$|u(\varphi)| \leq C\|\varphi\|_{\mathscr{C}^m} \qquad \text{for} \quad \varphi \in \mathscr{D}_K, \tag{57}$$

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

$$|T_u(\varphi)| \leq \text{vol}(K)\|u\|_{\mathscr{C}^0(K)}\|\varphi\|_{\mathscr{C}^0}, \qquad \text{for} \quad \varphi \in \mathscr{D}_K. \tag{58}$$

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

Definition 33. 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* 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

$$u_j(\varphi) \to 0 \qquad \text{for each} \quad \varphi \in \mathscr{D}(\Omega). \tag{59}$$

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

$$|T_u(\varphi)| \leq \|u\|_{L^1(K)} \|\varphi\|_{\mathscr{C}^0}. \qquad \text{for} \quad \varphi \in \mathscr{D}_K, \tag{60}$$

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 $\mathscr{C}(\Omega) \subset \mathscr{D}'(\Omega)$ either directly or through the continuous embedding $\mathscr{C}(\Omega) \subset L^1_{\mathrm{loc}}(\Omega)$.

Example 36. 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

$$\int \sin(jx)\varphi(x)\mathrm{d}x = \frac{1}{j}\int \cos(jx)\varphi'(x)\mathrm{d}x \to 0 \qquad \text{as} \quad j \to \infty, \tag{61}$$

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 $\mathscr{E}(\Omega) = \mathscr{C}^\infty(\Omega)$ and $\mathscr{E}^m(\Omega) = \mathscr{C}^m(\Omega)$. 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 37. We have $u \in \mathscr{D}'^m(\Omega)$ iff for any compact $K \subset \Omega$, there exists $C > 0$ such that

$$|u(\varphi)| \leq C\|\varphi\|_{\mathscr{C}^m} \qquad \text{for} \quad \varphi \in \mathscr{D}^m_K. \tag{62}$$

In light of Definition 31, 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 (62) 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

$$\langle Ax, y'\rangle = \langle x, A'y'\rangle, \qquad x \in X,\, y' \in Y'. \tag{63}$$

This situation can be described by the diagram

$$X \xrightarrow{A} Y \xrightarrow{y'} \mathbb{R}\,, \qquad X \overset{A'y'}{\dashrightarrow} \mathbb{R} \tag{64}$$

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

$$p_x(A'y') \equiv |\langle A'y', x\rangle| = |\langle y', Ax\rangle| \equiv p_y(y'), \tag{65}$$

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

Theorem 38. *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 $y' = 0$.

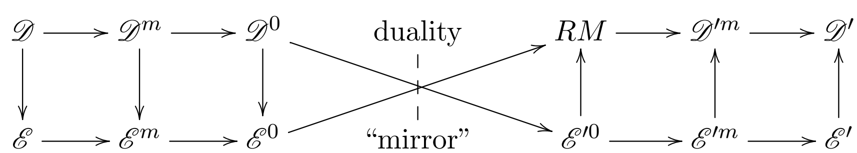
In the other direction, assume that $A(X)$ is *not* dense in Y. We want to produce a nonzero element $y' \in Y'$ such that $A'y' = 0$. By assumption, there is an element $y \in Y$ such that $y \notin \overline{A(X)}$. Consider the quotient $Z = Y/\overline{A(X)}$ with $\pi : Y \to Z$ the canonical projection. 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.,

$$\langle A'y', x\rangle = \langle y', Ax\rangle = 0, \qquad x \in X, \tag{66}$$

implying that $A'y' = 0$. □

Corollary 39. *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.



(67)

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 $\mathscr{E}(\Omega) = \mathscr{C}^\infty(\Omega)$ and $\mathscr{E}^m(\Omega) = \mathscr{C}^m(\Omega)$. 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

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. Let X and Y be locally convex spaces, and $Z \subset X'$ be a linear subspace. Suppose that $T : Z \to Y'$ is a linear map, and that $T' : Y \to X$ is a linear continuous map, satisfying

$$\langle Tz, y\rangle = \langle z, T'y\rangle, \qquad z \in Z,\ y \in Y. \tag{68}$$

The basic example we have in mind is $X = Y = Z = \mathscr{D}(\Omega)$, where we want to extend a linear operator $T : \mathscr{D}(\Omega) \to \mathscr{D}'(\Omega)$ to $\tilde{T} : \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega)$. In this case, we are required to find a continuous linear map $T' : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ satisfying (68), that is,

$$\int (T\psi)\varphi = \int \psi(T'\varphi), \qquad \psi, \varphi \in \mathscr{D}(\Omega). \tag{69}$$

For instance, if $T = \partial_j$, we have

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

hence we may take $T' = -\partial_j$, upon noting that the continuity of $\partial_j : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$, and therefore of T', has been confirmed in Example 26.

Returning to the general setting, we define $\tilde{T} : X' \to Y'$ by

$$\langle \tilde{T}x', y\rangle = \langle x', T'y\rangle, \qquad x' \in X',\, y \in Y, \tag{71}$$

which turns out to be an extension of T that has favourable properties.

Theorem 40. *Let X and Y be locally convex spaces, and $Z \subset X'$ be a linear subspace. Suppose that $T : Z \to Y'$ is a linear map, and that $T' : Y \to X$ is a linear continuous map, satisfying*

$$\langle Tz, y\rangle = \langle z, T'y\rangle, \qquad z \in Z,\, y \in Y. \tag{72}$$

Then the linear map $\tilde{T} : X' \to Y'$ defined by (71) is continuous, and $\tilde{T}z = Tz$ for $z \in Z$. Moreover, if Z is dense in X' then $\tilde{T}$ is the unique continuous extension of T.

Proof. It is easily checked that $\tilde{T}x' \in Y'$ for $x' \in X'$, since by continuity of $x' : X \to \mathbb{R}$ and of $T' : Y \to X$, we have

$$|\langle \tilde{T}x', y\rangle| = |\langle x', T'y\rangle| \leq cp(T'y) \leq c'q(y), \tag{73}$$

where p is some seminorm of X, and q is some seminorm of Y. Moreover, $\tilde{T} : X' \to Y'$ is continuous, because

$$p_y(\tilde{T}x') = |\langle \tilde{T}x', y\rangle| = |\langle x', T'y\rangle| = p_x(x'), \qquad y \in Y, \tag{74}$$

where $x = T'y \in X$. If $z \in Z$, then

$$\langle \tilde{T}z, y\rangle = \langle z, T'y\rangle = \langle Tz, y\rangle, \qquad \text{for} \quad y \in Y, \tag{75}$$

hence $\tilde{T}$ is indeed an extension of T.

To prove uniqueness, suppose that T_1 and T_2 are two continuous extensions of T. Then the operator $S = T_1 - T_2 : X' \to Y'$ is continuous and $Sz = 0$ for $z \in Z$. Since $S^{-1}(\{0\})$ is closed and $Z \subset S^{-1}(\{0\})$, we have $S^{-1}(\{0\}) \supset \overline{Z}$, the closure or Z in X'. But Z is dense in X', hence $S^{-1}(\{0\}) = X'$, or $S = 0$. □

Remark 41. Note that in the preceding theorem, there is no continuity requirement on T. The assumed continuity of $T' : Y \to X$ guarantees that $\tilde{T} : X' \to Y'$ is continuous, which, by restriction, implies that $T : Z \to Y'$ is continuous, where Z is taken with the subspace topology induced by the embedding $Z \subset X'$.

Let us consider some applications of this device.

Theorem 42 (Differential operators with smooth coefficients)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open set, let $a \in \mathscr{E}(\Omega)$, and let α be a multi-index. Then the operator $a\partial^\alpha : \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega)$ defined by*

$$\langle a\partial^\alpha u, \varphi\rangle = \langle u, (-1)^{|\alpha|}\partial^\alpha(a\varphi)\rangle, \qquad \varphi \in \mathscr{D}(\Omega), \tag{76}$$

is the unique continuous extension of the operator $a\partial^\alpha$ acting on $\mathscr{D}(\Omega)$.

Proof. We would like to apply Theorem 40 with $X = Y = Z = \mathscr{D}(\Omega)$ and $T = a\partial^\alpha$. For $u, \varphi \in \mathscr{D}(\Omega)$, integration by parts yields

$$\langle a\partial^\alpha u, \varphi\rangle = \int_\Omega a\varphi\partial^\alpha u = (-1)^{|\alpha|}\int_\Omega u\partial^\alpha(a\varphi), \tag{77}$$

which confirms that $T' : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ given by $T'\varphi = (-1)^{|\alpha|}\partial^\alpha(a\varphi)$ would work, if it is continuous. Continuity of $T' : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ means that for any compact $K \subset \Omega$, the map $T' : \mathscr{D}_K \to \mathscr{D}(\Omega)$ is continuous. Since $\operatorname{supp}\varphi \subset K$ implies $\operatorname{supp}(T'\varphi) \subset K$, and the topology of $\mathscr{D}_K$ coincides with the subspace topology $\mathscr{D}_K \subset \mathscr{D}(\Omega)$, this in turn is equivalent to continuity of $T' : \mathscr{D}_K \to \mathscr{D}_K$. The latter follows from the bound

$$\|\partial^\alpha(a\varphi)\|_{\mathscr{C}^m(K)} \leq C\|a\|_{\mathscr{C}^{m+|\alpha|}(K)}\|\varphi\|_{\mathscr{C}^{m+|\alpha|}(K)}, \qquad \varphi \in \mathscr{D}_K, \tag{78}$$

where m is arbitrary and C may depend on m.

For uniqueness, we simply state the fact that $\mathscr{D}(\Omega)$ is (sequentially) dense in $\mathscr{D}'(\Omega)$, which will be given as an exercise once we define convolutions. □

Example 43 (Differentiation)**.** Taking $a \equiv 1$ in the previous theorem, we get

$$\langle \partial^\alpha u, \varphi\rangle = (-1)^{|\alpha|}\langle u, \partial^\alpha\varphi\rangle, \qquad \varphi \in \mathscr{D}(\Omega). \tag{79}$$

Example 44. Let $\theta \in L^1_{\text{loc}}(\mathbb{R})$ 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

$$\langle \theta', \varphi\rangle = -\langle \theta, \varphi'\rangle = -\int_0^\infty \varphi'(x)\mathrm{d}x = \varphi(0), \tag{80}$$

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

Example 45 (Multiplication by a smooth function)**.** Taking $\alpha = 0$ in Theorem 42, we get

$$\langle au, \varphi\rangle = \langle u, a\varphi\rangle, \qquad \varphi \in \mathscr{D}(\Omega). \tag{81}$$

Theorem 46 (Change of variables)**.** *Let $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \mathbb{R}^n$ be open sets, and let $\Phi : \Omega \to \Omega'$ be a diffeomorphism. Then the map $\Phi^* : \mathscr{D}'(\Omega') \to \mathscr{D}'(\Omega)$ defined by*

$$\langle \Phi^* u, \varphi\rangle = \langle u, |\det D\Phi^{-1}|(\Phi^{-1})^*\varphi\rangle, \qquad \varphi \in \mathscr{D}(\Omega), \tag{82}$$

is the unique continuous extension of the pull-back operator $\Phi^* : \mathscr{D}(\Omega') \to \mathscr{D}(\Omega)$.

Proof. We would like to apply Theorem 40 with $X = Z = \mathscr{D}(\Omega')$, $Y = \mathscr{D}(\Omega)$, and $T = \Phi^*$. Recalling that the pull-back is defined for ordinary functions by $\Phi^* u = u \circ \Phi$, we have

$$\langle \Phi^* u, \varphi\rangle = \int_\Omega u(\Phi(x))\varphi(x)\mathrm{d}x = \int_{\Omega'} u(y)\varphi(\Phi^{-1}(y))|\det D\Phi^{-1}(y)|\mathrm{d}y, \tag{83}$$

for $u \in \mathscr{D}(\Omega')$ and $\varphi \in \mathscr{D}(\Omega)$, which suggests that we take $T' = |\det D\Phi^{-1}|(\Phi^{-1})^*$. What remains is to establish the continuity of $T' : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega')$, that is, to show that for any compact $K \subset \Omega$, the map $T' : \mathscr{D}_K \to \mathscr{D}(\Omega')$ is continuous. Note that $K' := \Phi(K)$ is compact, and that $\operatorname{supp}\varphi \subset K$ implies $\operatorname{supp}(T'\varphi) \subset K'$. Therefore taking into account the fact that the topology of $\mathscr{D}_{K'}$ coincides with the subspace topology $\mathscr{D}_{K'} \subset \mathscr{D}(\Omega')$, it suffices to check the continuity of $T' : \mathscr{D}_K \to \mathscr{D}_{K'}$. Since $|\det D\Phi^{-1}| > 0$ on K', we have

$$\||\det D\Phi^{-1}|\varphi \circ \Phi^{-1}\|_{\mathscr{C}^m(K')} \leq C\|\det D\Phi^{-1}\|_{\mathscr{C}^m(K')}\|\varphi \circ \Phi^{-1}\|_{\mathscr{C}^m(K')}, \qquad \varphi \in \mathscr{D}_K, \tag{84}$$

where m is arbitrary and C may depend on m. The last term can be estimated as

$$\|\varphi \circ \Phi^{-1}\|_{\mathscr{C}^m(K')} \leq C\big(1 + \|D\Phi^{-1}\|_{\mathscr{C}^m(K')} + \ldots + \|D\Phi^{-1}\|^m_{\mathscr{C}^m(K')}\big)\|\varphi\|_{\mathscr{C}^m(K)}. \tag{85}$$

confirming the continuity of $T' : \mathscr{D}_K \to \mathscr{D}_{K'}$.

As in the proof of Theorem 42, uniqueness follows from the density of $\mathscr{D}(\Omega')$ in $\mathscr{D}'(\Omega')$. □

Example 47 (Translation)**.** Taking $\Omega = \Omega' = \mathbb{R}^n$ and $\Phi(x) = x - a$ for some fixed $a \in \mathbb{R}^n$, and denoting $\tau_a\phi(x) = \Phi^*\phi(x) = \phi(x-a)$, we get

$$\langle \tau_a u, \varphi\rangle = \langle u, \tau_{-a}\varphi\rangle, \qquad \varphi \in \mathscr{D}(\mathbb{R}^n). \tag{86}$$

Example 48 (Linear transformation)**.** Taking $\Omega = \Omega' = \mathbb{R}^n$ and $\Phi(x) = A^{-1}x$ for some invertible matrix $A \in \mathbb{R}^{n\times n}$, and denoting $A\phi(x) = \Phi^*\phi(x) = \phi(A^{-1}x)$, we get

$$\langle Au, \varphi\rangle = \langle u, |\det A| A^{-1}\varphi\rangle, \qquad \varphi \in \mathscr{D}(\mathbb{R}^n). \tag{87}$$

Example 49 (Convolution with a test function)**.** We would like to extend the map $\psi \mapsto a * \psi$, where $a \in \mathscr{D}(\mathbb{R}^n)$. For $\psi, \varphi \in \mathscr{D}(\mathbb{R}^n)$, we have

$$\int (a*\psi)\varphi = \int\int a(x-z)\psi(z)\varphi(x)\mathrm{d}z\mathrm{d}x = \int \psi(\tilde{a}*\varphi), \tag{88}$$

where $\tilde{a}(x) = a(-x)$ denotes the reflection through the origin. Postponing the continuity question to a later section, at least formally for now, we get

$$\langle a*u, \varphi\rangle = \langle u, \tilde{a}*\varphi\rangle, \qquad u \in \mathscr{D}'(\mathbb{R}^n),\, \varphi \in \mathscr{D}(\mathbb{R}^n). \tag{89}$$

7. Sheaf structure of distributions

Definition 50. 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

$$\langle u|_\omega, \varphi\rangle = \langle u, \varphi\rangle, \qquad \varphi \in \mathscr{D}(\omega). \tag{90}$$

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 $\mathscr{D}'(\Omega)$. Of course, this situation is entirely analogous to what happens between $\mathscr{C}(\omega)$ and $\mathscr{C}(\Omega)$, or between $L^1_{\mathrm{loc}}(\omega)$ and $L^1_{\mathrm{loc}}(\Omega)$.

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 51. *In a), b), and c), 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*

$$\forall\alpha,\, u|_{\omega_\alpha} = 0 \qquad \Rightarrow \qquad u = 0. \tag{91}$$

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

$$u_\alpha|_{\omega_\alpha\cap\omega_\beta} = u_\beta|_{\omega_\alpha\cap\omega_\beta} \qquad \forall\alpha,\beta. \tag{92}$$

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

$$\langle u, \varphi\rangle = \langle u, \sum_\alpha \chi_\alpha\varphi\rangle = \sum_\alpha \langle u, \chi_\alpha\varphi\rangle = \sum_\alpha \langle u|_{\omega_\alpha}, \chi_\alpha\varphi\rangle = 0, \tag{93}$$

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

$$u(\varphi) := \sum_\alpha \langle u_\alpha, \chi_\alpha\varphi\rangle. \tag{94}$$

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

$$\sum_\alpha \langle u_\alpha, \chi_\alpha\varphi\rangle = \sum_{\alpha,\beta} \langle u_\alpha, \xi_\beta\chi_\alpha\varphi\rangle = \sum_{\alpha,\beta} \langle u_\beta, \xi_\beta\chi_\alpha\varphi\rangle = \sum_\beta \langle u_\beta, \xi_\beta\varphi\rangle, \tag{95}$$

where in the second step we used the property (92). 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

$$|u(\varphi)| \le \sum_\alpha |\langle u_\alpha, \chi_\alpha\varphi\rangle| \le \sum_\alpha C_\alpha \|\chi_\alpha\varphi\|_{C^{m_\alpha}} \le C\|\varphi\|_{C^m}, \tag{96}$$

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

8. Compactly supported distributions

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

$$\operatorname{supp} u = \Omega \setminus \bigcup\{\omega \subset \Omega \text{ open} : u|_\omega = 0\}. \tag{97}$$

Lemma 53. *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 54. $\operatorname{supp}\delta = \{0\}$.

The following theorem characterizes compactly supported distributions.

Theorem 55. *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. Similarly, we also have $\mathscr{E}'^m(\Omega) = \{u \in \mathscr{D}'^m(\Omega) : \operatorname{supp} u \Subset \Omega\}$.*

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

$$p_{m,K}(\varphi) = \|\varphi\|_{\mathscr{C}^m(K)}, \qquad K \Subset \Omega,\, m \in \mathbb{N}_0. \tag{98}$$

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

$$|u(\varphi)| \leq C\|\varphi\|_{\mathscr{C}^m(K)}, \qquad \varphi \in \mathscr{E}(\Omega). \tag{99}$$

Hence $u(\varphi) = 0$ if $\operatorname{supp} \varphi \subset \Omega \backslash K$, 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

$$u(\varphi) = u(\rho\varphi) + u(\varphi - \rho\varphi) = u(\rho\varphi), \qquad \varphi \in \mathscr{D}(\Omega), \tag{100}$$

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

$$|u(\varphi)| = |u(\rho\varphi)| \leq C\|\rho\varphi\|_{\mathscr{C}^m(K')} \leq C'\|\varphi\|_{\mathscr{C}^m(K')}, \qquad \varphi \in \mathscr{D}(\Omega), \tag{101}$$

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 $u \in \mathscr{E}'(\Omega)$. We leave the third statement of the theorem as an exercise. □

Exercise 2. 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 (101) 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 56. 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

$$u(\varphi) = \sum_j [\varphi(x_j) - \varphi(0)], \qquad \varphi \in \mathscr{E}(\mathbb{R}). \tag{102}$$

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

$$|u(\varphi)| = \sum_j |x_j| \|\varphi'\|_{\mathscr{C}^0([0,1])} \leq \|\varphi'\|_{\mathscr{C}^0([0,1])}, \qquad \varphi \in \mathscr{E}(\mathbb{R}). \tag{103}$$

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

$$|u(\varphi)| \leq C \sum_{k=0}^{m} \|\varphi^{(k)}\|_{\mathscr{C}^0(K)}, \qquad \varphi \in \mathscr{E}(\mathbb{R}), \tag{104}$$

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

$$u(\varphi_j) = j. \tag{105}$$

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

$$\sum_{k=0}^{m} \|\varphi_j^{(k)}\|_{\mathscr{C}^0(K)} = 1. \tag{106}$$

Therefore the bound (104) cannot hold.

Even though in general we cannot get a bound of the form (104) 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 57. *Let $u \in \mathscr{E}'^m(\Omega)$ and let $\varphi \in \mathscr{C}^m(\Omega)$ 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 $\|\rho_\varepsilon\|_{\mathscr{C}^k} \leq c\varepsilon^{-k}$ 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

$$\begin{aligned} |u(\varphi)| &\leq c \sum_{|\alpha| \leq m} \|\partial^\alpha(\rho_\varepsilon \varphi)\|_{\mathscr{C}^0} \leq c \sum_{|\alpha+\beta| \leq m} \|\partial^\beta \rho_\varepsilon\|_{\mathscr{C}^0} \|\partial^\alpha \varphi\|_{\mathscr{C}^0(K+B_\varepsilon)} \\ &\leq c \sum_{|\alpha| \leq m} \varepsilon^{|\alpha| - m} \|\partial^\alpha \varphi\|_{\mathscr{C}^0(K+B_\varepsilon)}, \end{aligned} \tag{107}$$

where the constant $c > 0$ may have different values at its different occurrences. We have $\|\partial^\alpha \varphi\|_{\mathscr{C}^0(K+B_\varepsilon)} \to 0$ as $\varepsilon \to 0$ for $|\alpha| \leq m$, because $\varphi \in \mathscr{C}^m(\Omega)$. This means that in the sum on the right hand side of (107), 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

$$|\partial^\alpha \varphi(y)| \leq c \sup_{0<t<1} |f^{(m-|\alpha|)}(t)|. \tag{108}$$

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

$$\sup_{0<t<1} |f^{(m-|\alpha|)}(t)| \leq c\varepsilon^{m-|\alpha|} \|\varphi\|_{\mathscr{C}^m(K+B_\varepsilon)}, \tag{109}$$

confirming that the sum on the right hand side of (107) tends to 0 as $\varepsilon \to 0$. $\square$

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

Corollary 58. *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*

$$u(\varphi) = \sum_{|\alpha| \leq m} a_\alpha \partial_\alpha \varphi(0), \qquad \varphi \in \mathscr{C}^m(\Omega), \tag{110}$$

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

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

$$\psi(x) = \varphi(x) - \sum_{|\alpha| \leq m} \frac{\partial^\alpha \varphi(0)}{\alpha!} x^\alpha, \qquad x \in \Omega. \tag{111}$$

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

$$u(\varphi) = \sum_{|\alpha| \leq m} \frac{\partial^\alpha \varphi(0)}{\alpha!} u(p_\alpha) + u(\psi) = \sum_{|\alpha| \leq m} \frac{u(p_\alpha)}{\alpha!} \partial^\alpha \varphi(0), \tag{112}$$

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 59. *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

$$|u(\varphi)| \leq C\|\varphi\|_{\mathscr{C}^m(K)} = C \max_{|\alpha|\leq m} \|\partial^\alpha \varphi\|_{\mathscr{C}^0(K)}, \qquad \varphi \in \mathscr{E}^m(\Omega). \tag{113}$$

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

$$\|\psi\|_{\mathscr{C}^0} \leq C\|\partial_j \psi\|_{\mathscr{C}^0}, \tag{114}$$

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 (113) 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.,

$$|u(\varphi)| \leq C\|\partial^\beta \varphi\|_{\mathscr{C}^0}, \qquad \varphi \in \mathscr{D}(\Omega). \tag{115}$$

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

$$\psi(x) = \int_{y<x} \partial_1 \ldots \partial_n \psi(y)\, \mathrm{d}y, \tag{116}$$

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

$$|u(\varphi)| \leq C \int |\partial^\beta \varphi|, \qquad \varphi \in \mathscr{D}(\Omega), \tag{117}$$

now with $\beta = (m+1, m+1, \ldots, m+1)$. 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

$$T(\partial^\beta \varphi) := u(\varphi), \tag{118}$$

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.

$$\begin{array}{ccc} \mathscr{D}(\Omega) & \xrightarrow{\ u\ } & \mathbb{R} \\ {\scriptstyle \partial^\beta}\big\downarrow & \nearrow_{T} & \\ X & & \end{array} \tag{119}$$

Now the estimate (117) simply says that

$$|T(\xi)| \leq C\|\xi\|_{L^1(\omega)}, \qquad \xi \in X, \tag{120}$$

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

$$T(\xi) = \int g\xi, \qquad \xi \in L^1(\Omega). \tag{121}$$

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

$$u(\varphi) = T(\partial^\beta \varphi) = \int g \partial^\beta \varphi = (-1)^{|\beta|} \langle \partial^\beta g, \varphi \rangle, \tag{122}$$

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 3. Let $\Omega \subset \mathbb{R}^n$ be a bounded open set, and let $u \in \mathscr{E}'^m(\Omega)$. Show that there exists $g \in \mathscr{C}(\mathbb{R}^n)$ 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 4. Let $u \in \mathscr{D}'(\Omega)$. Show that for any open $\omega \subset \Omega$ with $\overline{\omega} \Subset \Omega$, there exists $g \in \mathscr{C}(\mathbb{R}^n)$ 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 60. *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 $|\alpha_j|_\infty \leq m$.*

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

$$\Omega = \bigcup_j \omega_j, \qquad \Omega = \bigcup_j \sigma_j, \qquad \text{and} \qquad \bar{\sigma}_j \subset \omega_j \quad \text{for all } j. \tag{123}$$

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

$$\langle u, \varphi \rangle = \sum_j \langle \rho_j u, \varphi \rangle = \sum_j \langle \rho_j u, \psi_j \varphi \rangle = \sum_j \int g_j \partial^{\alpha_j} (\psi_j \varphi_j), \tag{124}$$

where in the last step we employed the representation from Exercise 3 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 $\operatorname{supp} \phi_j$, and it would not affect the integral in (124). The proof is established, since any compact set $K \subset \Omega$ intersects only finitely many of ω_j. □

10. Vector-valued calculus

Let $\Omega \subset \mathbb{R}^n$, and let $(X, \mathscr{P})$ be a sequentially complete locally convex space. We want to develop a basic calculus for functions $f : \Omega \to X$. We denote by $\mathscr{C}(\Omega, X)$ the space of *continuous functions* on Ω having values in X, and define the space of *bounded functions*

$$\mathscr{B}(\Omega, X) = \{f : \Omega \to X : f(\Omega) \text{ is bounded}\}. \tag{125}$$

Note that when we assess the continuity of $f : \Omega \to X$, we think of Ω as a topological space in itself, whose topology is induced by the embedding $\Omega \subset \mathbb{R}^n$. For instance, even when Ω is a closed subset of $\mathbb{R}^n$, the sets of the form $B_\varepsilon(x) \cap \Omega$ with $x \in \Omega$ are *open* subsets of the topological space Ω.

Exercise 5. Show that if X is a Banach space, and $\Omega \subset \mathbb{R}^n$ is open, then $\mathscr{B}(\Omega, X) \cap \mathscr{C}(\Omega, X)$ is a Banach space under the norm

$$\|f\|_{\Omega} = \sup_{x\in\Omega} \|f(x)\|_X. \tag{126}$$

In what follows, it will be notationally convenient to assume that the family $\mathscr{P}$ of seminorms is given by the collection $\{|\cdot|_\lambda : \lambda \in \Lambda\}$, where Λ is some index set. Thus for instance, $f \in \mathscr{B}(\Omega, X)$ if and only if

$$|f|_{\Omega,\lambda} := \sup_{x\in\Omega} |f(x)|_\lambda < \infty, \qquad \text{for each} \quad \lambda \in \Lambda. \tag{127}$$

This suggests that $\{|\cdot|_{\Omega,\lambda} : \lambda \in \Lambda\}$ may serve as a family of seminorms on $\mathscr{B}(\Omega, X)$ inducing a natural topology on it. Similarly, if Ω is open, the family

$$\{|\cdot|_{K,\lambda} : K \Subset \Omega,\ \lambda \in \Lambda\}, \qquad \text{where} \quad |f|_{K,\lambda} = \sup_{x\in K} |f(x)|_\lambda, \tag{128}$$

is a natural candidate for the space $\mathscr{C}(\Omega, X)$.

Exercise 6. Show that if Ω is open, the aforementioned family (128) defines a sequentially complete, locally convex structure on $\mathscr{C}(\Omega, X)$.

Let us define the λ-*oscillation of f over $\tau \subset \Omega$* as

$$\operatorname{osc}_\lambda(f, \tau) = \sup_{\xi,\eta\in\tau} |f(\xi) - f(\eta)|_\lambda. \tag{129}$$

Then a function f is continuous at $x \in \Omega$ if and only if $\operatorname{osc}_\lambda(f, B_\varepsilon(x) \cap \Omega) \to 0$ as $\varepsilon \to 0$ for each $\lambda \in \Lambda$. Moreover, if f is continuous in Ω, and $\lambda \in \Lambda$, then $|f(\xi) - f(\eta)|_\lambda$ is a continuous function of $(\xi, \eta) \in \Omega \times \Omega$, hence the oscillation $\operatorname{osc}_\lambda(f, B_\varepsilon(x) \cap \Omega)$ is a continuous function of $x \in \Omega$ and $\varepsilon > 0$.

For $f \in \mathscr{C}(\Omega, X)$ and $\lambda \in \Lambda$, the seminorm $|f(x)|_\lambda$ is a continuous function of $x \in \Omega$. Therefore if K is compact, we have $\mathscr{C}(K, X) \subset \mathscr{B}(K, X)$. Furthermore, assuming that K is compact and $f \in \mathscr{C}(K, X)$, for $\lambda \in \Lambda$, the λ-*modulus of continuity*

$$\omega_\lambda(\varepsilon) = \max_{x\in K} \operatorname{osc}_\lambda(f, B_\varepsilon(x) \cap K), \tag{130}$$

is a continuous function of $\varepsilon \geq 0$ with $\omega_\lambda(0) = 0$, which shows that continuous functions on compact sets are *uniformly continuous*, in the sense that $\omega_\lambda(\varepsilon) \to 0$ as $\varepsilon \searrow 0$ for each $\lambda \in \Lambda$.

Definition 61. If for a function $f : \Omega \to X$, a point $x \in \Omega$, and a vector $\xi \in \mathbb{R}^n$, there is some $u \in X$ such that

$$\frac{f(x + \xi t) - f(x)}{t} \to u \qquad \text{as} \quad t \to 0, \tag{131}$$

then we say that f is *differentiable at $x \in \Omega$ along the direction ξ*, and write

$$\partial_\xi f(x) = \frac{\partial f}{\partial \xi}(x) = u, \tag{132}$$

which is called the *directional derivative of f at x along ξ*. In the special case where ξ is the unit vector along the coordinate direction x_k, the corresponding directional derivative is called the *partial derivative* with respect to x_k, and written

$$\partial_k f = \frac{\partial f}{\partial x_k} = \frac{\partial f}{\partial \xi}. \tag{133}$$

The following exercise explains why we *can* simply focus on the partial derivatives and their continuity, rather than the totality $\partial_\xi f(x)$ for $\xi \in \mathbb{R}^n$.

Exercise 7. Let Ω be an open subset of $\mathbb{R}^n$, and let $\{\xi_1, \ldots, \xi_n\}$ be a basis of $\mathbb{R}^n$. Assume that the directional derivatives $\partial_{\xi_1} f(x), \ldots, \partial_{\xi_n} f(x)$ exist at each $x \in \Omega$ and are continuous in Ω as functions of x. Show that for any $\xi \in \mathbb{R}^n$, the directional derivative $\partial_\xi f$ exists and

$$\partial_\xi f(x) = a_1 \partial_{\xi_1} f(x) + \ldots + a_1 \partial_{\xi_n} f(x), \qquad x \in \Omega, \tag{134}$$

where the numbers $a_1, \ldots, a_n \in \mathbb{R}$ are such that $\xi = a_1\xi_1 + \ldots a_n\xi_n$.

Suppose that $\Omega \subset \mathbb{R}^n$ is open. Since $\partial_\xi f$ is again a function defined on Ω with values in X, we can talk about its directional differentiability, hence for instance, $\partial_\xi \partial_\xi f$, $\partial_1 \partial_2 f$, or even $\partial^\alpha f$ can be defined (if they exist). The following exercise justifies the formula $\partial_j \partial_k f = \partial_k \partial_j f$ under the condition that the participated derivatives are continuous.

Exercise 8. Let $\xi, \eta \in \mathbb{R}^n$, and assume that $\partial_\xi \partial_\eta f$ and $\partial_\eta \partial_\xi f$ exist and are continuous in Ω, where Ω is an open subset of $\mathbb{R}^n$. Show that $\partial_\xi \partial_\eta f = \partial_\eta \partial_\xi f$ in Ω.

For $\Omega \subset \mathbb{R}^n$ open and $0 \leq m \leq \infty$, we define

$$\mathscr{C}^m(\Omega, X) = \{f \in \mathscr{C}(\Omega, X) : \partial^\alpha f \in \mathscr{C}(\Omega, X),\, |\alpha| \leq m\}. \tag{135}$$

If $\Omega \subset \mathbb{R}^n$ is not open, we can still make sense of the preceding definition, by reading the statement "$\partial^\alpha f \in \mathscr{C}(\Omega, X)$" as "$\partial^\alpha f$ exists in the interior of Ω, and can be extended to a continuous function on Ω."

Exercise 9. Define a canonical locally convex structure on $\mathscr{C}^m(\Omega, X)$, where $\Omega \subset \mathbb{R}^n$ is open and $0 \leq m \leq \infty$.

Now we want to introduce the Riemann integral. We will only be concerned with integration over the unit cube $Q = (0,1)^n$, and give pointers to how to deal with more general domains.

Definition 62. A *partition* of Q is a finite collection $P = \{\tau\}$ of parallelepipeds that is

- essentially disjoint, in the sense that $\mathrm{vol}(\tau \cap \sigma) = 0$ if τ and σ are two distinct parallelepipeds from P,
- a cover of $\bar{Q}$, that is, $\bar{Q} = \bigcup_{\tau \in P} \tau$.

The *width* of a partition P is

$$|P| = \max_{\tau \in P} \mathrm{diam}(\tau). \tag{136}$$

Definition 63. A partition P is called *tagged* if to each $\tau \in P$ there assigned a point $\xi \in \tau$. If P is a tagged partition, we overload the letter P to denote also the set of pairs (τ, ξ), where $\xi \in \tau$ is the tag of $\tau \in P$.

Definition 64. Let $\mathscr{T}$ be the set of all tagged partitions, and let $F : \mathscr{T} \to X$. We say that F has a *limit as* $|P| \to 0$, if there exists $u \in X$ such that $u = \lim_{k\to\infty} F(P_k)$ whenever $P_1, P_2, \ldots$ is a sequence of tagged partitions satisfying $|P_k| \to 0$ as $k \to \infty$. In this situation, we call u the *limit of* F *as* $|P| \to 0$, and write

$$u = \lim_{|P|\to 0} F(P). \tag{137}$$

Definition 65. Given a tagged partition P, the *Riemann sum* of $f : \bar{Q} \to X$ with respect to P is defined by

$$S_P(f) = \sum_{(\tau,\xi)\in P} f(\xi)\mathrm{vol}(\tau), \tag{138}$$

whenever it makes sense. We say that f is *Riemann integrable* if $S_P(f)$ has a limit in X as $|P| \to 0$. If f is integrable, we take its *Riemann integral* to be

$$\int f(x)\mathrm{d}x = \lim_{|P|\to 0} S_P(f). \tag{139}$$

Remark 66. The preceding definition can be generalized in several directions.

- Obviously, the size of Q is not a restriction; The same construction can be made for an arbitrary parallelepiped Q.
- If we use simplices instead of parallelepipeds in the partitions, we can define the Riemann integral over an arbitrary bounded polyhedron.
- Some *ad hoc* constructions can be made for simple domains such as a ball.
- If $\Omega \subset Q$ is a bounded domain, and χ_Ω is its characteristic function, then we define
$$\int_\Omega f(x)\mathrm{d}x = \int \chi_\Omega(x) f(x)\mathrm{d}x, \tag{140}$$
which works at least for nice enough domains.
- For a possibly unbounded $\Omega \subset \mathbb{R}^n$ and a possibly unbounded function $f : \Omega \to X$, we define the *improper Riemann integral* to be the limit
$$\int_\Omega f(x)\mathrm{d}x = \lim_{k\to\infty} \int_{\Omega_k} f(x)\mathrm{d}x, \tag{141}$$
where Ω_k is a sequence of (simple) bounded domains satisfying $\bar{\Omega} = \overline{\cup_k \Omega_k}$.
- Note that "improper integral" typically means that the limit (141) does not depend on the sequence $\{\Omega_k\}$, as long as the sequence $\{\Omega_k\}$ stays "reasonable". If there is such a dependence, we choose one specific sequence $\{\Omega_k\}$, and call the resulting limit the *Cauchy principal value* (corresponding to that sequence). Usually, we pick a sequence with some special symmetry, such as balls centred at the origin.

Since there is no risk of confusion we will drop the adjective "Riemann" from integrability and integral. We have a simple criterion on integrability, in terms of the quantity
$$\mathrm{osc}_\lambda(f, P) = \sum_{\tau\in P} \mathrm{osc}_\lambda(f, \tau)\mathrm{vol}(\tau), \tag{142}$$
which could be called the λ-*oscillation of* f *over the partition* P.

Lemma 67. *A bounded function* $f \in \mathscr{B}(\bar{Q}, X)$ *is integrable if there is a sequence* $P_1, P_2, \ldots$ *of partitions such that* $\mathrm{osc}_\lambda(f, P_k) \to 0$ *as* $k \to \infty$ *for each* $\lambda \in \Lambda$.

Proof. Note that if P' is a refinement of P, i.e., if P' can be obtained by subdividing the elements of P into smaller parallelepipeds, then $\mathrm{osc}_\lambda(f, P') \leq \mathrm{osc}_\lambda(f, P)$ and
$$|S_P(f) - S_{P'}(f)|_\lambda \leq \sum_{\tau\in P} \mathrm{osc}_\lambda(f, \tau)\mathrm{vol}(\tau) = \mathrm{osc}_\lambda(f, P), \tag{143}$$
regardless of how P and P' are tagged. Let $P_1, P_2, \ldots$ be a sequence of partitions whose λ-oscillations tending to 0 for each $\lambda \in \Lambda$. Replacing P_i by the common refinement of P_i and P_{i-1}, we can assume that P_{i+1} is a refinement of P_i for all i. Then from (143) and the sequential completeness of X we see that the Riemann sums corresponding to the sequence $\{P_i\}$ converge to some element $u \in X$, independent of how the partitions are tagged. The convergence can be established first for some particular tagging of the sequence $\{P_i\}$, and then be extended to arbitrary tagging by (143). Now we need to show that as long as the width is small, any tagged partition gives rise to a Riemann sum that is close to u. Let Q be a tagged partition with $|Q|$ small. Let Q' be the common refinement of Q and P_k, with k large. We tag Q' so that the tags of Q coincide with those of Q' on the parallelepipeds common to both Q and Q', meaning that
$$|S_Q(f) - S_{Q'}(f)|_\lambda \leq \sum_{\tau\in Q\setminus Q'} |Q|^n \mathrm{osc}_\lambda(f, \tau) \leq 2\#P_k|Q|^n \sup_{x\in\mathbb{R}^n} |f(x)|_\lambda, \tag{144}$$

where $\#P_k$ denotes the number of elements in P_k. Given $\lambda \in \Lambda$ and $\varepsilon > 0$, we choose k so large that $|S_{P_k}(f) - u|_\lambda < \varepsilon$. Then if $|Q|$ is so small that the right hand side of (144) is less than ε, we get

$$|S_Q(f) - u|_\lambda \leq |S_Q(f) - S_{Q'}(f)|_\lambda + |S_{Q'}(f) - u|_\lambda < 2\varepsilon, \tag{145}$$

establishing the claim. □

The following lemma produces a large class of integrable functions.

Corollary 68. *Functions in $\mathscr{C}(\bar{Q}, X)$ are integrable.*

Proof. Since $\mathscr{C}(\bar{Q}, X) \subset \mathscr{B}(\bar{Q}, X)$, it suffices to produce a sequence of partitions with oscillations vanishing in the limit. As we have previously observed, the functions in $\mathscr{C}(\bar{Q}, X)$ are uniformly continuous, that is, for each $\lambda \in \Lambda$, the modulus of continuity $\omega_\lambda(\varepsilon)$ is a continuous function of $\varepsilon \geq 0$ with $\omega_\lambda(0) = 0$. This means that for instance, if we take P_k to be a partition of Q into cubes of side length 2^{-k}, then $\operatorname{osc}_\lambda(f, P_k) \to 0$ as $k \to 0$ for each $\lambda \in \Lambda$ and $f \in \mathscr{C}(\bar{Q}, X)$. □

Obviously, integration is a linear operation. A few more simple properties are as follows.

Lemma 69. *If $f \in \mathscr{B}(\bar{Q}, X)$ is integrable, then we have the bounds*

$$\Big|\int f(x)\mathrm{d}x\Big|_\lambda \leq \operatorname{vol}(\operatorname{supp} f) \sup_{x \in \mathbb{R}^n} |f(x)|_\lambda, \qquad \textit{and} \qquad \Big|\int f(x)\mathrm{d}x\Big|_\lambda \leq \int |f(x)|_\lambda \mathrm{d}x, \tag{146}$$

for each $\lambda \in \Lambda$. Moreover, if $\{f_n\}$ is a sequence of integrable functions converging to f in $\mathscr{B}(\bar{Q}, X)$, then f is integrable and

$$\lim_{n\to\infty} \int f_n(x)\mathrm{d}x = \int f(x)\mathrm{d}x. \tag{147}$$

Proof. The inequalities (146) are true because they are true when the integrals are replaced by the corresponding Riemann sums. For any partition P (with any tag), we have

$$\begin{aligned} \Big|\int f_n(x)\mathrm{d}x - S_P(f)\Big|_\lambda &\leq \Big|\int f_n(x)\mathrm{d}x - S_P(f_n)\Big|_\lambda + |S_P(f_n) - S_P(f)|_\lambda \\ &\leq \Big|\int f_n(x)\mathrm{d}x - S_P(f_n)\Big|_\lambda + \operatorname{vol}(\operatorname{supp} f) \sup_{x\in\mathbb{R}^n} |f_n(x) - f(x)|_\lambda, \end{aligned} \tag{148}$$

which implies that f is integrable and that the limit (147) holds. □

Exercise 10 (Fundamental theorem of calculus)*.* a) If $u \in \mathscr{C}^1([0,1], X)$ then

$$u(1) - u(0) = \int u'(t)\mathrm{d}t. \tag{149}$$

b) If $f \in \mathscr{C}([0,1], X)$ then the function

$$F(x) = \int_0^x f(t)\mathrm{d}t, \qquad x \in [0,1], \tag{150}$$

satisfies $F \in \mathscr{C}^1([0,1], X)$ and $F' = f$ on $[0,1]$.