# An introduction to some aspects of functional analysis, 5: Smooth functions and distributions

Stephen Semmes
Rice University

**Abstract**

Some basic aspects of smooth functions and distributions on open subsets of $\mathbf{R}^n$ are briefly discussed.

## Contents

# 1 Smooth functions

Let $U$ be a nonempty open set in $\mathbf{R}^n$ for some positive integer $n$, and let $f(x)$ be a continuous real or complex-valued function on $U$. Remember that $f$ is said to be *continuously differentiable* on $U$ if the first partial derivatives

$$\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \tag{1.1}$$

exist at every point in $U$ and are continuous on $U$. Similarly, if the first partial derivatives (1.1) of $f$ are also continuously-differentiable on $U$, then $f$ is said to be *twice continuously-differentiable* on $U$. If $k$ is a positive integer, and all derivatives of $f$ of order up to and including $k$ exist on $U$ and are continuous on $U$, then $f$ is said to be $k$ *times continuously-differentiable* on $U$. If derivatives of $f$ of all orders exist and are continuous on $U$, then $f$ is said to be *infinitely-differentiable* on $U$, or simply *smooth*. Let $C^k(U)$ be the space of $k$ times continuously-differentiable functions on $U$ for each positive integer $k$. This can be extended to $k = 0$ by letting $C^0(U)$ be the space $C(U)$ of continuous functions on $U$. Similarly, $C^\infty(U)$ denotes the space of smooth functions on $U$. These are all vector spaces with respect to pointwise addition and scalar multiplication, and commutative algebras with respect to pointwise multiplication of functions.

A *multi-index* is an $n$-tuple $\alpha = (\alpha_1, \dots \alpha_n)$ of nonnegative integers. The sum of two multi-indices is defined coordinatewise, and we put

$$|\alpha| = \sum_{j=1}^n \alpha_j \tag{1.2}$$

for each multi-index $\alpha$. If $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ and $\alpha$ is a multi-index, then the corresponding monomial $x^\alpha$ is defined by

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}. \tag{1.3}$$

More precisely, $x_j^{\alpha_j}$ is interpreted as being equal to 1 when $\alpha_j = 0$, even when $x_j = 0$, so that $x^\alpha = 1$ for every $x \in \mathbf{R}^n$ when $\alpha = 0$. Note that a polynomial on $\mathbf{R}^n$ is a finite linear combination of monomials.

If $f$ is a continuously-differentiable function on $U$ and $j = 1, \ldots, n$, then let $\partial_j f$ or $D_j f$ denote the partial derivative $\partial f / \partial x_j$ of $f$ in the direction $x_j$. Thus $D_j$ defines a linear mapping from $C^1(U)$ into $C(U)$ for each $j$, which maps $C^k(U)$ into $C^{k-1}(U)$ for each positive integer $k$. In particular, $D_j$ maps $C^\infty(U)$ into itself, which is one of the advantages of working with smooth functions. If $f \in C^k(U)$ for some $k \in \mathbf{Z}_+$ and $\alpha$ is a multi-index with $|\alpha| \le k$, then let

$$\partial^\alpha f = D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x^\alpha} \tag{1.4}$$

be the corresponding derivative of $f$ of order $|\alpha|$, which is equal to $f$ when $\alpha = 0$. Thus $D^\alpha$ is a linear mapping from $C^k(U)$ into $C^{k-|\alpha|}(U)$, and $D^\alpha$ maps $C^\infty(U)$ into itself for each $\alpha$.

# 2 Supremum seminorms

Let $U$ be a nonempty open set in $\mathbf{R}^n$, let $K$ be a nonempty compact subset of $U$, and let $f$ be a continuous real or complex-valued function on $U$. The corresponding supremum seminorm of $f$ is defined by

$$\|f\|_K = \sup_{x \in K} |f(x)|. \tag{2.1}$$

The collection of these seminorms defines a topology on $C(U)$ in a standard way, so that $C(U)$ becomes a locally convex topological vector space. A sequence $\{f_j\}_{j=1}^\infty$ of continuous functions on $U$ converges to a continuous function $f$ on $U$ with respect to this topology if and only if

$$\lim_{j \to \infty} \|f_j - f\|_K = 0 \tag{2.2}$$

for each nonempty compact set $K \subseteq U$, which is the same as saying that $\{f_j\}_{j=1}^\infty$ converges to $f$ uniformly on compact subsets of $U$.

Now let $K$ be a nonempty compact subset of $U$, let $k$ be a positive integer, and let $\alpha$ be a multi-index with $|\alpha| \le k$. It is easy to see that

$$\|f\|_{K,\alpha} = \|D^\alpha f\|_K = \sup_{x \in K} |D^\alpha f(x)| \tag{2.3}$$

defines a seminorm on $C^k(U)$. As before, $C^k(U)$ is a locally convex topological vector space with respect to the topology defined by the collection of these seminorms, using all nonempty compact subsets $K$ of $U$ and all multi-indices $\alpha$ with $|\alpha| \le k$. A sequence $\{f_j\}_{j=1}^\infty$ of $C^k$ functions on $U$ converges to $f \in C^k(U)$ with respect to this topology if and only if

$$\lim_{j \to \infty} \|f_j - f\|_{K,\alpha} \tag{2.4}$$

for each such $K$ and $\alpha$, which is the same as saying that $\{D^\alpha f_j\}_{j=1}^\infty$ converges to $D^\alpha f$ uniformly on compact subsets of $U$ when $|\alpha| \le k$.

Similarly, $C^\infty(U)$ is a locally convex topological vector space with respect to the topology defined by the collection of all seminorms $\|f\|_{K,\alpha}$, where $K$ is any nonempty compact subset of $U$ and $\alpha$ is any multi-index. A sequence $\{f_j\}_{j=1}^\infty$ of $C^\infty$ functions on $U$ converges to $f \in C^\infty(U)$ if and only if (2.4) for each $K$ and $\alpha$, which is the same as saying that $\{D^\alpha f_j\}_{j=1}^\infty$ converges to $D^\alpha f$ uniformly on compact subsets of $U$ for each multi-index $\alpha$. Of course, $C^\infty(U) \subseteq C^k(U)$ for each $k \ge 0$, and it is easy to see that the natural inclusion map that sends each $f \in C^\infty(U)$ to itself as an element of $C^k(U)$ is continuous with respect to these topologies. If $k$, $l$ are nonnegative integers with $k \le l$, then $C^l(U) \subseteq C^k(U)$, and again the natural inclusion of $C^l(U)$ in $C^k(U)$ is continuous with respect to these topologies.

If $\alpha$ is a multi-index, $k$ is a positive integer, and $|\alpha| \le k$, then $D^\alpha$ is a linear mapping from $C^k(U)$ into $C^{k-|\alpha|}(U)$, as in the previous section. It is easy to see that $D^\alpha$ is also continuous as a mapping from $C^k(U)$ into $C^{k-|\alpha|}(U)$ with respect to the corresponding topologies just defined. This basically comes down to the fact that

$$\|D^\alpha f\|_{K,\beta} = \|f\|_{K,\alpha+\beta} \tag{2.5}$$

for every $f \in C^k(U)$, every nonempty compact set $K \subseteq U$, and every multi-index $\beta$ such that

$$|\alpha + \beta| = |\alpha| + |\beta| \le k, \tag{2.6}$$

since

$$D^\beta(D^\alpha f) = D^{\alpha+\beta} f. \tag{2.7}$$

Similarly, $D^\alpha$ is a continuous linear mapping from $C^\infty(U)$ into itself for every multi-index $\alpha$.

# 3 Countably many seminorms

Suppose for the moment that $U = \mathbf{R}^n$, let

$$|x| = \Big(\sum_{j=1}^n x_j^2\Big)^{1/2} \tag{3.1}$$

be the standard Euclidean norm on $\mathbf{R}^n$, and let

$$\overline{B}(x,r) = \{y \in \mathbf{R}^n : |x-y| \le r\} \tag{3.2}$$

be the closed ball in $\mathbf{R}^n$ with center $x \in \mathbf{R}^n$ and radius $r \ge 0$. Thus $\overline{B}(x,r)$ is closed and bounded and hence compact in $\mathbf{R}^n$ for every $x \in \mathbf{R}^n$ and $r \ge 0$, and every compact set in $\mathbf{R}^n$ is contained in $\overline{B}(0,r)$ for some $r \ge 0$, because compact subsets of $\mathbf{R}^n$ are bounded. This implies that one can get the same topology on $C(\mathbf{R}^n)$ using the sequence of supremum seminorms associated to the closed balls $\overline{B}(0,r)$ with $r \in \mathbf{Z}_+$. Similarly, one can get the same topology on $C^k(U)$

for some positive integer $k$ by using the seminorms associated to $\overline{B}(0,r)$ with $r \in \mathbf{Z}_+$ and multi-indices $\alpha$ with $|\alpha| \le k$, and there are only countably many of these seminorms. One can also get the same topology on $C^\infty(U)$ using the seminorms associated to $\overline{B}(0,r)$ with $r \in \mathbf{Z}_+$ and arbitrary multi-indices $\alpha$, and there are only countably many of these seminorms, because there are only countably many multi-indices.

Now let $U$ be a nonempty proper subset of $\mathbf{R}^n$, and put

$$\mathrm{dist}(x, \mathbf{R}^n \backslash U) = \inf\{|x-y| : y \in \mathbf{R}^n \backslash U\} \tag{3.3}$$

for each $x \in \mathbf{R}^n$. Consider

$$A_r = \{x \in U : |x| \le r,\ \mathrm{dist}(x, \mathbf{R}^n \backslash U) \ge 1/r\} \tag{3.4}$$

for each positive integer $r$, and observe that $A_r$ is closed and bounded and hence compact in $\mathbf{R}^n$ for each $r \in \mathbf{Z}_+$. By construction, $\bigcup_{r=1}^\infty A_r = U$, and it is easy to see that every compact set $K \subseteq U$ is contained in $A_r$ when $r$ is sufficiently large. Although $A_r$ may be empty for finitely many $r$, it is nonempty for all but finitely many $r$, in which case it can be used to define seminorms as in the previous section. One can get the same topologies on $C(U)$, $C^l(U)$ for each $k \in \mathbf{Z}_+$, and $C^\infty(U)$ using the seminorms associated to these compact sets $A_r$ when $A_r \ne \emptyset$, and there are only countably many of these seminorms, as before.

Thus these topologies on $C(U)$, $C^k(U)$, and $C^\infty(U)$ can be defined by only countably many seminorms on these spaces for every nonempty open set $U$ in $\mathbf{R}^n$, which implies that there are countable local bases for the topologies of each of these spaces at 0. This leads to a lot of helpful simplifications, and in fact this implies that there are translation-invariant metrics on these spaces that determine the same topologies. In particular, it suffices to use sequences for many topological properties related to these spaces.

# 4 Cauchy sequences

Let $U$ be a nonempty open set in $\mathbf{R}^n$. A sequence $\{f_j\}_{j=1}^\infty$ of continuous real or complex-valued functions on $U$ is said to be a *Cauchy sequence* if

$$\lim_{j,l \to \infty} \|f_j - f_l\|_K = 0 \tag{4.1}$$

for every nonempty compact set $K \subseteq U$. Equivalently, this means that $f_j - f_l$ is contained in any neighborhood of 0 in $C(U)$ for all sufficiently large $j, l \in \mathbf{Z}_+$. This is also the same as saying that $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence with respect to any translation-invariant metric on $C(U)$ that defines the same topology on $C(U)$ as before.

At any rate, if $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $C(U)$, then $\{f_j(x)\}_{j=1}^\infty$ is a Cauchy sequence in the real or complex numbers, as appropriate, for each $x \in U$, since one can apply (4.1) to $K = \{x\}$. Because the real and complex numbers are complete, it follows that $\{f_j(x)\}_{j=1}^\infty$ converges for each $x \in U$.

Thus $\{f_j\}_{j=1}^{\infty}$ converges pointwise on to a real or complex-valued function $f$ on $U$, as apprpriate, and one can use (4.1) to show that $\{f_j\}_{j=1}^{\infty}$ actually converges to $f$ uniformly on compact subsets of $U$. This implies that $f$ is also continuous on $U$, by standard arguments, and hence that $\{f_j\}_{j=1}^{\infty}$ converges to $f$ in $C(U)$. It follows that $C(U)$ is complete, in the sense that every Cauchy sequence in $C(U)$ converges, so that $C(U)$ is a *Fréchet space*, since the topology on $C(U)$ can be defined by a sequence of seminorms.

Similarly, a sequence $\{f_j\}_{j=1}^{\infty}$ of continuously-differentiable functions on $U$ if (4.1) holds and

$$\lim_{j,l\to\infty} \|D_r f_j - D_r f_l\|_K = 0 \tag{4.2}$$

for every nonempty compact set $K \subseteq U$ and $r = 1, \ldots, n$, which is to say that $\{f_j\}_{j=1}^{\infty}$ is a Cauchy sequence in $C(U)$, and that $\{D_r f_j\}_{j=1}^{\infty}$ is a Cauchy sequence in $C(U)$ for each $r = 1, \ldots, n$. By the previous argument, this implies that $\{f_j\}_{j=1}^{\infty}$ converges to a continuous function $f$ on $U$ in $C(U)$, and that $\{D_r f_j\}_{j=1}^{\infty}$ converges to a continuous function $g_r$ on $U$ in $C(U)$ for $r = 1, \ldots, n$. Using well-known results in analysis, it follows that $f$ is also continuously-differentiable on $U$, and that $D_r f = g_r$ for $r = 1, \ldots, n$. One way to see this is to reduce to the $n = 1$ case by considering each variable separately, and then use the fundamental theorem of calculus to derive this from analogous results about integrals of uniformly convergent sequences of continuous functons on closed intervals in the real line. Thus $\{f_j\}_{j=1}^{\infty}$ converges to $f$ in $C^1(U)$.

If $k$ is any positive integer, then a sequence $\{f_j\}_{j=1}^{\infty}$ of $C^k$ functions on $U$ is said to be a Cauchy sequence in $C^k(U)$ if

$$\lim_{j,l\to\infty} \|f_j - f_l\|_{K,\alpha} = 0 \tag{4.3}$$

for every nonempty compact set $K \subseteq U$ and every multi-index $\alpha$ with $|\alpha| \le k$. One can show that every Cauchy sequence in $C^k(U)$ converges in $C^k(U)$, by repeating the same type of arguments as in the $k = 1$ case. Thus $C^k(U)$ is also complete for every $k \in \mathbf{Z}_+$, and hence is a Fréchet space. A sequence $\{f_j\}_{j=1}^{\infty}$ of $C^\infty$ functions on $U$ is said to be a Cauchy sequence in $C^\infty(U)$ if (4.3) holds for every nonempty compact set $K \subseteq U$ and every multi-index $\alpha$, in which case $\{f_j\}_{j=1}^{\infty}$ converges in $C^\infty(U)$ by the same type of argument as before. This implies that $C^\infty(U)$ is complete and hence a Fréchet space as well.

# 5 Compact support

Let $U$ be a nonempty open set in $\mathbf{R}^n$, and let $f$ be a continuous real or complex-valued function on $U$. The *support* of $f$ is defined by

$$\operatorname{supp} f = \overline{\{x \in U : f(x) \neq 0\}}, \tag{5.1}$$

where more precisely one takes the closure of the set where $f(x) \neq 0$ relative to $U$. We say that $f$ has compact support in $U$ if $\operatorname{supp} f$ is a compact set in $U$, which implies in particular that $\operatorname{supp} f$ is a closed set in $\mathbf{R}^n$. The space of continuous functions on $U$ with compact support is denoted $C_{com}(U)$, which is a linear subspace of $C(U)$. Similarly, the space $C^k_{com}(U)$ of $C^k$ functions on $U$ with compact support is a linear subspace of $C^k(U)$, and the space $C^\infty_{com}(U)$ of $C^\infty$ functions on $U$ with compact support is a linear subspace of $C^\infty(U)$.

Let $\eta(x)$ be the real-valued function on the real line defined by $\eta(x) = 0$ when $x \le 0$, and

$$\eta(x) = \exp(-1/x) \tag{5.2}$$

when $x > 0$. It is well known and not difficult to check that $\eta$ is a $C^\infty$ function on $\mathbf{R}$. Of course, $\eta$ is obviously $C^\infty$ on $\mathbf{R}\backslash\{0\}$, and so the main point is that the derivatives of $\eta$ at 0 are all equal to 0, and that the derivatives of $\eta$ at $x > 0$ all tend to 0 as $x \to 0$. If $a, b \in \mathbf{R}$ and $a < b$, then

$$\eta_{a,b}(x) = \eta(x - a)\,\eta(b - x) \tag{5.3}$$

is a $C^\infty$ function on $\mathbf{R}$ such that $\eta_{a,b}(x) > 0$ when $a < x < b$ and $\eta_{a,b}(x) = 0$ otherwise. If $a_1, b_1, \ldots, a_n, b_n \in \mathbf{R}$ and $a_j < b_j$ for $j = 1, \ldots, n$, then

$$\prod_{j=1}^{n} \eta_{a_j,b_j}(x_j) \tag{5.4}$$

is a $C^\infty$ function on $\mathbf{R}^n$ which is strictly positive when $a_j < x_j < b_j$ for $j = 1, \ldots, n$ and is equal to 0 otherwise. Similarly, if $p \in \mathbf{R}^n$ and $r > 0$, then

$$\eta(r^2 - |x - p|^2) \tag{5.5}$$

is a $C^\infty$ function of $x$ on $\mathbf{R}^n$ which is positive when $|x - p| < r$ and equal to 0 otherwise. This shows that there are a lot of $C^\infty$ functions with compact support on $\mathbf{R}^n$.

Note that $\eta(x) < 1$ for every $x \in \mathbf{R}$, $\eta(x)$ is strictly increasing for $x \ge 0$, and that $\eta(x) \to 1$ as $x \to \infty$. Put

$$\xi_c(x) = 1 - \eta(c - x) \tag{5.6}$$

for each $x, c \in \mathbf{R}$, so that $\xi_c(x)$ is a $C^\infty$ function of $x$ on $\mathbf{R}$, $\xi_c(x) = 1$ when $x \ge c$, $0 < \xi_c(x) < 1$ when $x < c$, and $\xi_c(x)$ is strictly increasing in $x$ when $x \le c$. Suppose now that $0 < c < 1$, and consider

$$\zeta_c(x) = (1 - \xi_c(0))^{-1}\,(\xi_c(\eta(x)) - \xi_c(0)). \tag{5.7}$$

This is a $C^\infty$ function on the real line that satisfies $\zeta_c(x) = 0$ when $x \le 0$, since $\eta(x) = 0$. If $x > 0$, then $\eta(x) > 0$, which implies that $\xi_c(\eta(x)) > \xi_c(0)$, and hence $\zeta_c(x) > 0$. If $x > 0$ and $\eta(x) \ge c$, then $\xi_c(\eta(x)) = 1$, so that $\zeta_c(x) = 1$. Equivalently, $\zeta_c(x) = 1$ when $x \ge \eta^{-1}(c) > 0$. If $x > 0$ and $\eta(x) < c$, then $\zeta_c(x) < 1$, and in fact $\zeta_c$ is strictly increasing on $[0, \eta^{-1}(c)]$.

Let $K$ be a nonempty compact set in $\mathbf{R}^n$, and let $V$ be an open set in $\mathbf{R}^n$ such that $K \subseteq V$. By compactness, there are finitely many open balls $B_1, \ldots, B_l$ in $\mathbf{R}^n$ such that $K \subseteq \bigcup_{j=1}^{l} B_l$ and the closure $\overline{B_j}$ of $B_j$ is contained in $V$ for each

$j$. Let $\phi_j$ be a nonnegative real-valued $C^\infty$ function on $\mathbf{R}^n$ such that $\phi_j(x) > 0$ for each $x \in B_j$ and $\phi_j(x) = 0$ when $x \in \mathbf{R}^n \backslash B_j$, as in (5.5). If $\phi = \sum_{j=1}^n \phi_j$, then $\phi$ is a nonnegative real-valued $C^\infty$ function on $\mathbf{R}^n$, $\phi(x) > 0$ for every $x \in K$, and the support of $\phi$ is equal to $\bigcup_{j=1}^l \overline{B_j}$, which is a compact subset of $V$. Let $c$ be a positive real number, and consider

$$\psi_c(x) = (1 - \xi_c(0))^{-1} \, (\xi_c(\phi(x)) - \xi_c(0)). \tag{5.8}$$

This is a real-valued $C^\infty$ function on $\mathbf{R}^n$ such that $\psi_c(x) = 0$ when $\phi(x) = 0$, which happens when $x \in \mathbf{R}^n \backslash \Big( \bigcup_{j=1}^l B_j \Big)$. If $x \in \bigcup_{j=1}^l B_j$, then $\phi(x) > 0$, and hence $\psi_c(x) > 0$ too. More precisely, $0 < \psi_c(x) < 1$ when $0 < \phi(x) < c$, and $\psi_c(x) = 1$ when $\phi(x) \geq c$, since $\xi_c(\phi(x)) = 1$. In particular, $\psi_c(x) = 1$ for every $x \in K$ when $c > 0$ is sufficiently small, because $K$ is compact and $\phi > 0$ on $K$. If $c$ is a bit smaller, then $\psi_c(x) = 1$ for every $x$ in a neighborhood of $K$.

# 6 Inductive limits

Let $U$ be a nonempty open set in $\mathbf{R}^n$, and let $K$ be a nonempty compact subset of $U$. The space $C_K(U)$ of continuous real or complex-valued functions on $U$ with support contained in $K$ is a linear subspace of $C_{com}(U)$, and a closed linear subspace of $C(U)$. Similarly, if $k$ is any positive integer, then the space $C^k_K(U)$ of $C^k$ functions on $U$ with support contained in $K$ is a linear subspace of $C^k_{com}(U)$, and a closed linear subspace of $C^k(U)$. The space $C^\infty_K(U)$ of $C^\infty$ functions on $U$ with support contained in $K$ is also a linear subspace of $C^\infty_{com}(U)$, and a closed linear subspace of $C^\infty(U)$. Of course, these subspaces are trivial unless $K$ has nonempty interior in $\mathbf{R}^n$.

The topologies on $C_K(U)$, $C^k_K(U)$, and $C^\infty_K(U)$ induced by those on $C(U)$, $C^k(U)$, and $C^\infty(U)$, respectively, can be described more simply than before. The restriction of the supremum seminorm $\|f\|_K$ on $C(U)$ to $C_K(U)$ is a norm, and it is easy to see that the topology on $C_K(U)$ determined by this norm is the same as the one induced by the topology on $C(U)$. The topology on $C^k_K(U)$ induced by the one on $C^k(U)$ is also determined by the finitely many seminorms $\|f\|_{K,\alpha}$ with $|\alpha| \leq k$, and the sum or maximum of these seminorms is a norm on $C^k_K(U)$ that defines the same topology. It suffices to use the collection of seminorms $\|f\|_{K,\alpha}$ on $C^\infty_K(U)$, where $\alpha$ is any multi-index, which is still a countable collection of seminorms.

Note that $C_{com}(U)$ is the same as the union of $C_K(U)$ over all compact subsets $K$ of $U$, and similarly for $C^k_{com}(U)$ and $C^\infty_{com}(U)$. In fact, it is enough to use a suitable sequence of compact subsets of $U$, as in Section 3. Let us focus on $C^\infty_{com}(U)$, since that is the case of primary interest here. There is a well-known construction of an "inductive limit" topology on $C^\infty_{com}(U)$, which makes $C^\infty_{com}(U)$ into a locally convex topological vector space with some additional properties. One of these properties is that the topology on $C^\infty_K(U)$ induced by the one on $C^\infty_{com}(U)$ is the same as before for each nonempty compact set $K \subseteq U$. Another important property is that a linear mapping from $C^\infty_{com}(U)$

into a locally convex topological vector space is continuous if and only if its restriction to $C^\infty_K(U)$ is continuous for each nonempty compact set $K \subseteq U$. It is easy to see that this topology on $C^\infty_{com}(U)$ is uniquely determined by these properties, using the continuity of the identity mapping on $C^\infty_{com}(U)$ with respect to any two such topologies to compare them.

In particular, the obvious inclusion mapping of $C^\infty_{com}(U)$ in $C^\infty(U)$ should be continuous, since its restriction to $C^\infty_K(U)$ is automatically continuous for each nonempty compact set $K \subseteq U$. This implies that $C^\infty_K(U)$ should be a closed linear subspace of $C^\infty_{com}(U)$, since $C^\infty_K(U)$ is a closed linear subspace of $C^\infty(U)$. However, this topology on $C^\infty_{com}(U)$ is much stronger than the one induced on it by $C^\infty(U)$, and in fact there is no countable local base for the topology of $C^\infty_{com}(U)$ at 0, for instance.

# 7 Distributions

By definition, a *distribution* on a nonempty open set $U$ in $\mathbf{R}^n$ is a continuous linear functional $\lambda$ on $C^\infty_{com}(U)$, with respect to the topology discussed in the previous section. Because of the properties of this topology, it suffices to check that the restriction of $\lambda$ to $C^\infty_K(U)$ is continuous for every nonempty set $K \subseteq U$. Here we shall simply use this as the working definition of continuity of a linear functional on $C^\infty_{com}(U)$, and hence of a distribution on $U$. Equivalently, a linear functional $\lambda$ on $C^\infty_{com}(U)$ is continuous if for each compact set $K \subseteq U$ and every sequence $\{\phi_j\}_{j=1}^\infty$ of smooth functions on $U$ supported in $K$ that converges to a smooth function $\phi$ on $U$ in the $C^\infty$ topology, we have that

$$\lim_{j \to \infty} \lambda(\phi_j) = \lambda(\phi). \tag{7.1}$$

Of course, the support of $\phi$ is also contained in $K$ under these conditions, and we may as well restrict our attention to the case where $\phi = 0$, because $\lambda$ is supposed to be linear on $C^\infty_{com}(U)$.

As a basic class of examples, let $f$ be a continuous real or complex-valued function on $U$, and consider

$$\lambda_f(\phi) = \int_U \phi(x)\, f(x)\, dx \tag{7.2}$$

for each test function $\phi \in C^\infty_{com}(U)$. If the support of $\phi$ is contained in a nonempty compact set $K \subseteq U$, then we get that

$$|\lambda_f(\phi)| \le \int_K |\phi(x)|\, |f(x)|\, dx \le \|\phi\|_K \int_K |f(x)|\, dx \tag{7.3}$$

where $\|\phi\|_K$ is the supremum of $|\phi(x)|$ over $x \in K$, as in (2.1). This implies that $\lambda_f$ is continuous on $C^\infty_K(U)$, and hence that $\lambda_f$ is a distribution on $U$. If $f$ is a locally integrable function on $U$, then $\lambda_f$ defines a distribution on $U$ in the same way. Now let $p$ be an element of $U$, and put

$$\delta_p(\phi) = \phi(p) \tag{7.4}$$

for each $\phi \in C^\infty_{com}(U)$, which is known as the *Dirac delta distribution* at $p$. It is easy to see that this is a distribution on $U$, since $|\delta_p(\phi)| \le \|\phi\|_K$ when the support of $\phi$ is contained in a compact set $K \subseteq U$ that contains $p$, and $\delta_p(\phi) = 0$ otherwise. One can also think of the Dirac mass at $p$ as a Borel measure on $U$, and every locally finite Borel measure on $U$ defines a distribution on $U$.

To be a bit more precise, if one takes $C^\infty_{com}(U)$ to be the space of real-valued smooth functions on $U$ with compact support, then a linear functional on $C^\infty_{com}(U)$ means a real-linear mapping from $C^\infty_{com}(U)$ into $\mathbf{R}$, and the continuous linear functionals on $C^\infty_{com}(U)$ may be described as real distributions on $U$. If one take $C_{com}(U)$ to be the space of complex-valued smooth functions on $U$ with compact support, then the linear functionals on $C^\infty_{com}(U)$ are the complex-linear mappings from $C^\infty_{com}(U)$ into $\mathbf{C}$, and the continuous linear functionals on $C^\infty_{com}(U)$ may be described as complex distributions on $U$.

Let $W$ be a nonempty open set in $\mathbf{R}^n$ which is a subset of $U$. If $\phi \in C^\infty_{com}(W)$, then we can extend $\phi$ to a function on $U$ by putting $\phi(x) = 0$ for every $x$ in $U \backslash W$, and this extension of $\phi$ is a smooth function on $U$ with compact support contained in $W$. This defines a natural linear mapping from $C^\infty_{com}(W)$ into $C^\infty_{com}(U)$, and it is easy to see that this linear mapping is continuous with respect to the corresponding inductive limit topologies, because $C^\infty_K(W)$ is essentially the same as $C^\infty_K(U)$ for each compact set $K \subseteq W$, with the same topology. If $\lambda$ is a distribution on $U$, then we can compose $\lambda$ as a continuous linear functional on $C^\infty_{com}(U)$ with this linear mapping to get a continuous linear functional on $C^\infty_{com}(W)$, and thus a distribution on $W$, known as the restriction of $\lambda$ to $W$. If $\lambda = \lambda_f$ is associated to a locally integrable function $f$ on $U$ as in (7.2), for example, then the restriction of $\lambda$ to $W$ corresponds exactly to the restriction of $f$ to $W$ as a locally integrable function.

# 8 Differentiation of distributions

Let $U$ be a nonempty open set in $\mathbf{R}^n$, let $k$ be a positive integer, and let $\alpha$ be a multi-index with $|\alpha| \le k$. If $f$ and $\phi$ are $C^k$ functions on $U$, and if $\phi$ has compact support in $U$, then

$$\int_U (D^\alpha f)(x)\, \phi(x)\, dx = (-1)^{|\alpha|} \int_U f(x)\, (D^\alpha \phi)(x)\, dx. \tag{8.1}$$

This follows from integration by parts, and is especially clear when $U = \mathbf{R}^n$. Otherwise, one can use suitable cut-off functions to reduce to that case, for instance.

If $\lambda$ is a distribution on $U$, then we can define $D^\alpha \lambda$ initially as a linear functional on $C^\infty_{com}(U)$ by

$$(D^\alpha \lambda)(\phi) = (-1)^{|\alpha|}\, \lambda(D^\alpha\, \phi). \tag{8.2}$$

More precisely, $D^\alpha$ is a linear mapping from $C^\infty_{com}(U)$ into itself, and so (8.2) makes sense as a linear functional on $C^\infty_{com}(U)$, since $\lambda$ is a linear functional on

$C^\infty_{com}(U)$. If $\lambda$ is a distribution on $U$, and thus a continuous linear functional on $C^\infty_{com}(U)$, then it is easy to see that $D^\alpha \lambda$ is also continuous on $C^\infty_{com}(U)$, and hence a distribution on $U$, because $D^\alpha$ is a continuous linear mapping from $C^\infty_{com}(U)$ into itself. To see that $D^\alpha$ is continuous on $C^\infty_{com}(U)$, it suffices to check that the restriction of $D^\alpha$ is continuous on $C^\infty_K(U)$ for each nonempty compact set $K \subseteq U$. In fact, $D^\alpha$ is a continuous linear mapping from $C^\infty(U)$ into itself, as in Section 2, and $D^\alpha$ clearly maps $C^\infty_K(U)$ into itself for each nonempty compact set $K \subseteq U$,

If $f \in C^k(U)$ and $\lambda = \lambda_f$ is the distribution on $U$ corresponding to $f$ as in (7.2), then $D^\alpha \lambda$ is the distribution corresponding to $D^\alpha f$, as in (8.1). If $p \in U$ and $\delta_p$ is the Dirac delta distribution on $U$ at $p$ as in (7.4), then

$$(D^\alpha \delta_p)(\phi) = (-1)^{|\alpha|}\, (D^\alpha \phi)(p) \tag{8.3}$$

for every test function $\phi \in C^\infty_{com}(U)$. Suppose now that $n = 1$ and $U = \mathbf{R}$, and let $f(x)$ be the Heaviside function on $\mathbf{R}$ equal to 1 when $x > 0$ and to 0 when $x \le 0$. If $\lambda = \lambda_f$ is the distribution on $\mathbf{R}$ associated to $f$ as in (7.2), then the derivative of $\lambda$ is the Dirac delta distribution $\delta_0$ at $p = 0$. Indeed,

$$(D\lambda)(\phi) = -\lambda(D\phi) = -\int_0^\infty (D\phi)(x)\, dx = \phi(0) = \delta_0(\phi) \tag{8.4}$$

for every $\phi \in C^\infty_{com}(U)$, by the fundamental theorem of calculus.

Note that the space of distributions on $U$ is a vector space over the real or complex numbers, as appropriate, because the sum of two continuous linear functionals on $C^\infty_{com}(U)$ is also a continuous linear functional on $C^\infty_{com}(U)$, and similarly one can multiply continuous linear functionals on $C^\infty_{com}(U)$ by real or complex numbers. It is easy to see that $D^\alpha$ defines a linear mapping on the space of distributions on $U$ for each multi-index $\alpha$, because of the way that $D^\alpha$ is defined on distributions, and because $D^\alpha$ is linear on $C^\infty_{com}(U)$.

# 9 Multiplication by smooth functions

Let $U$ be a nonempty open set in $\mathbf{R}^n$, and let $g$ be a smooth real or complex-valued function on $U$. If $\phi$ is a smooth function on $U$ with compact support, then the product $g\,\phi$ of $g$ and $\phi$ is also a smooth function on $U$ with compact support contained in the support of $\phi$. More precisely, the mapping that sends $\phi$ to $g\,\phi$ is a linear mapping from $C^\infty_{com}(U)$ into itself, and one can check that this mapping is continuous. To see this, it suffices to show that this mapping is continuous as a mapping from $C^\infty_K(U)$ into itself for each nonempty compact set $K \subseteq U$. Basically, one can use Leibniz' product rule for derivatives to show that for each multi-index $\alpha$, $\|g\,\phi\|_{K,\alpha} = \|D^\alpha(g\,\phi)\|_K$ is less than or equal to a constant depending on $K$, $\alpha$, and $g$ times the sum of finitely many terms of the form $\|g\|_{K,\beta}$, where $\beta$ runs through the multi-indices such that $\beta_j \le \alpha_j$ for each $j = 1, \ldots, n$.

If $\lambda$ is a distribution on $U$, then it follows that

$$\widetilde{\lambda}(\phi) = \lambda(g\,\phi) \tag{9.1}$$

also defines a distribution on $U$, which is the product of $\lambda$ by $g$, and which may be denoted by $g\,\lambda$. If $\lambda$ is associated to a continuous or locally-integrable function $f$ on $U$ as in (7.2), then it is easy to see that $g\,\lambda$ is the distribution on $U$ corresponding to the product $f\,g$ of $f$ and $g$ on $U$ in the same way. Note that the mapping that sends $\lambda$ to $g\,\lambda$ is linear as a mapping from the space of distributions on $U$ into itself for each $g \in C^\infty(U)$. Similarly, the mapping that sends $g$ to $g\,\lambda$ is linear as a mapping from $C^\infty(U)$ into the space of distributions on $U$ for each distribution $\lambda$ on $U$.

There is a version of the product rule in this context, which states that

$$D_j(g\,\lambda) = (D_j g)\,\lambda + g\,(D_j \lambda) \tag{9.2}$$

for each $j = 1, \ldots, n$. Here $D_j(g\,\lambda)$ is the derivative of $g\,\lambda$ as a distribution on $U$, while $D_j g$ is the ordinary derivative of $g$ as a smooth function on $U$, which is also a smooth function on $U$. Thus $(D_j g)\,\lambda$ is defined as a distribution on $U$, as the product of the smooth function $D_j g$ on $U$ and the distribution $\lambda$. The derivative $D_j \lambda$ of $\lambda$ is defined as a distribution on $U$ as in the previous section, and then the product $g\,(D_j \lambda)$ is defined as a distribution on $U$ as before.

In order to verify (9.2), it suffices to check that

$$(D_j(g\,\lambda))(\phi) = ((D_j g)\,\lambda)(\phi) + (g\,(D_j \lambda))(\phi) \tag{9.3}$$

for each test function $\phi \in C^\infty_{com}(U)$. By definitions,

$$(D_j(g\,\lambda))(\phi) = -(g\,\lambda)(D_j\,\phi) = -\lambda(g\,(D_j \phi)), \tag{9.4}$$

where $D_j \phi$ is the ordinary derivative of $\phi$. Similarly,

$$((D_j g)\lambda)(\phi) = \lambda((D_j g)\,\phi), \tag{9.5}$$

and

$$(g\,(D_j \lambda))(\phi) = (D_j \lambda)(g\,\phi) = -\lambda(D_j(g\,\phi)), \tag{9.6}$$

where $D_j(g\,\phi)$ is the ordinary derivative of $g\,\phi$. The ordinary product rule implies that $D_j(g\,\phi) = (D_j g)\,\phi + g\,(D_j \phi)$, so that (9.6) becomes

$$(g\,(D_j \lambda))(\phi) = -\lambda((D_j g)\,\phi) - \lambda(g\,(D_j \phi)). \tag{9.7}$$

Thus (9.3) follows by combining (9.4), (9.5), and (9.7).

# 10 Partitions of unity

Let $K$ be a nonempty compact set in $\mathbf{R}^n$, and suppose that for each $p \in K$ we have an open ball $B(p)$ in $\mathbf{R}^n$ that contains $p$ as an element. Because $K$

is compact, there are finitely many elements $p_1, \ldots, p_l$ of $K$ such that $K$ is contained in $V = \bigcup_{j=1}^{l} B(p_j)$. As in Section 5, there is a real-valued smooth function $\theta$ on $\mathbf{R}^n$ with compact support contained in $V$ such that $0 \le \theta(x) \le 1$ for every $x \in \mathbf{R}^n$ and $\theta(x) = 1$ for every $x \in K$. In particular, $1 - \theta(x)$ is a nonnegative smooth function on $\mathbf{R}^n$ that vanishes on $K$ and is positive on the complement of a compact subset of $V$, which is all that we shall need here.

Let $\phi_j$ be a nonnegative real-valued smooth function on $\mathbf{R}^n$ which is positive on $B(p_j)$ and vanishes on $\mathbf{R}^n \backslash B_j$ for each $j = 1, \ldots, l$, as in (5.5). Thus

$$(10.1) \qquad (1 - \theta(x)) + \sum_{k=1}^{l} \phi_k(x)$$

is a real valued smooth function on $\mathbf{R}^n$ that is positive everywhere, since it is positive on each $B_j$ and on $\mathbf{R}^n \backslash V$. It follows that

$$(10.2) \qquad \psi_j(x) = \phi_j(x) \Big( (1 - \theta(x)) + \sum_{k=1}^{l} \phi_k(x) \Big)^{-1}$$

is also a nonnegative smooth function on $\mathbf{R}^n$ that is positive on $B_j$ and vanishes on $\mathbf{R}^n \backslash B_j$ for each $j = 1, \ldots, l$. By construction, $\psi_j(x) \le 1$ for every $x \in \mathbf{R}^n$, and

$$(10.3) \qquad \sum_{j=1}^{l} \psi_j(x) = 1$$

for every $x \in K$, since $1 - \theta(x) = 0$ when $x \in K$.

Now let $\lambda$ be a distribution on a nonempty open set $U$ in $\mathbf{R}^n$, and let $W$ be an open set contained in $U$. We say that $\lambda$ vanishes on $W$, or $\lambda = 0$ on $W$, if $\lambda(\phi) = 0$ for every $\phi \in C^\infty_{com}(U)$ with $\operatorname{supp} \phi \subseteq W$. If $\lambda$ is the distribution associated to a continuous function $f$ on $U$ as in (7.2), then this is equivalent to the condition that $f(x) = 0$ for every $x \in W$. If $f$ is a locally integrable function on $U$, then this is equivalent to asking that $f(x) = 0$ for almost every $x \in W$. If $\lambda$ is the Dirac delta distribution $\delta_p$ at some point $p \in U$, or any derivative of $\delta_p$, then $\lambda$ vanishes on $U \backslash \{p\}$.

Let $\{W_\alpha\}_{\alpha \in A}$ be a collection of open subsets of $U$, and suppose that a distribution $\lambda$ vanishes on $W_\alpha$ for each $\alpha \in A$. We would like to show that $\lambda$ also vanishes on $W = \bigcup_{\alpha \in A} W_\alpha$ under these conditions. Suppose that $\phi \in C^\infty_{com}(U)$ satisfies $\operatorname{supp} \phi \subseteq W$, and put $K = \operatorname{supp} \phi$. If $p \in K$, then $p \in W_{\alpha(p)}$ for some $\alpha(p) \in A$, and hence there is an open ball $B(p)$ in $\mathbf{R}^n$ such that $p \in B(p)$ and $\overline{B(p)} \subseteq W_{\alpha(p)}$. Because $K$ is compact, there are finitely many elements $p_1, \ldots, p_l$ of $K$ such that $K \subseteq \bigcup_{j=1}^{l} B(p_j)$, as before. Thus we get smooth functions $\psi_1, \ldots, \psi_l$ on $\mathbf{R}^n$ with $\operatorname{supp} \psi_j = \overline{B(p_j)} \subseteq W_{\alpha(p_j)}$ for each $j$ and which satisfy (10.3) for each $x \in K = \operatorname{supp} \phi$, as in the previous discussion. In particular, $\operatorname{supp}(\psi_j \, \phi) \subseteq W_{\alpha(p_j)}$ for each $j$, so that $\lambda(\psi_j \, \phi) = 0$ for each $j$, since $\lambda = 0$ on $W_\alpha$ for each $\alpha \in A$ by hypothesis. This implies that $\lambda(\phi) = 0$, as desired, because $\phi = \sum_{j=1}^{l} \psi_j \, \phi$, by (10.3).

# 11 Compactly-supported distributions

Let $U$ be a nonempty open set in $\mathbf{R}^n$, and let $\lambda$ be a distribution on $U$. As in the previous section, $\lambda$ vanishes on an open set $W \subseteq U$ if $\lambda(\phi) = 0$ for every $\phi \in C^\infty_{com}(U)$ with $\operatorname{supp} \phi \subseteq W$. If $V$ is the union of all of the open sets $W \subseteq U$ such that $\lambda = 0$ on $W$, then $\lambda = 0$ on $V$, as before. The support $\operatorname{supp} \lambda$ of $\lambda$ is defined to be $U \backslash V$, which is relatively closed in $U$, since $V$ is an open set.

Suppose that the support of $\lambda$ is contained in a compact set $K \subseteq U$, so that $\lambda = 0$ on $U \backslash K$. This means that $\lambda(\phi) = 0$ when $\phi \in C^\infty_{com}(U)$ satisfies $\operatorname{supp} \phi \subseteq U \backslash K$, which is the same as saying that $\phi$ vanishes on a neighborhood of $K$. As in Section 5, there is a smooth function $\psi_K$ on $\mathbf{R}^n$ with compact support contained in $U$ such that $\psi_K(x) = 1$ for every $x$ in a neighborhood of $K$. If $\phi$ is any smooth function on $U$ with compact support, then $(1 - \psi_K)\, \phi$ is a smooth function on $U$ with compact support contained in $U \backslash K$, and hence

$$\lambda((1 - \psi_K)\, \phi) = 0. \tag{11.1}$$

Equivalently,

$$\lambda(\phi) = \lambda(\psi_K\, \phi) \tag{11.2}$$

for every $\phi \in C^\infty_{com}(U)$.

If $\phi$ is any smooth function on $U$, then $\psi_K\, \phi$ is a smooth function on $U$ with compact support contained in the support of $\psi_K$, and hence $\lambda(\psi_K\, \phi)$ is defined. Thus (11.2) can be used to extend $\lambda$ to a linear functional on $C^\infty(U)$, and one can check that this linear functional is continuous with respect to the topology on $C^\infty(U)$ described in Section 2. More precisely, if $K_1 = \operatorname{supp} \psi_K$, then the mapping that sends $\phi \in C^\infty(U)$ to $\psi_K\, \phi$ is continuous as a linear mapping from $C^\infty(U)$ into $C^\infty_{K_1}(U)$. In particular, the mapping that sends $\phi \in C^\infty(U)$ to $\psi_K\, \phi$ is continuous as a linear mapping from $C^\infty(U)$ into $C^\infty_{com}(U)$. This implies that $\lambda(\psi_K\, \phi)$ is continuous as a linear functional of $\phi \in C^\infty(U)$, since $\lambda$ is supposed to be continuous on $C^\infty_{com}(U)$.

Conversely, suppose now that $\lambda$ is a continuous linear functional on $C^\infty(U)$. Because of the way that the topology on $C^\infty(U)$ is defined, this means that there are finitely many nonempty compact subsets $K_1, \ldots, K_l$ of $U$, finitely many multi-indices $\alpha_1, \ldots, \alpha_l$, and a nonnegative real number $A$ such that

$$|\lambda(\phi)| \le A \sum_{j=1}^{l} \|\phi\|_{K_j, \alpha_j} \tag{11.3}$$

for every $\phi \in C^\infty(U)$, where $\|\phi\|_{K,\alpha}$ is as in (2.3). If $K = \bigcup_{j=1}^{l} K_j$, then $K$ is a compact subset of $U$, and (11.3) implies that $\lambda(\phi) = 0$ for every $\phi \in C^\infty(U)$ that vanishes on a neighborhood of $K$. In particular, the restriction of $\lambda$ to $C^\infty_{com}(U)$ is a distribution on $U$ with support contained in $K$. Note that $\lambda$ is uniquely determined on $C^\infty(U)$ by its restriction to $C^\infty_{com}(U)$, as in (11.2).

## 12 Nonnegative distributions

Let $U$ be a nonempty open set in $\mathbf{R}^n$. A linear functional $\lambda$ on $C^\infty_{com}(U)$ is said to be *nonnegative* if

$$\lambda(\phi) \ge 0 \tag{12.1}$$

for every nonnegative real-valued smooth function $\phi$ on $U$ with compact support. More precisely, if $\lambda$ is complex, then $\lambda(\phi)$ is supposed to be a real number under these conditions. A distribution on $U$ is said to be nonnegative if it is nonnegative as a linear functional on $C^\infty_{com}(U)$. If $f$ is a nonnegative real-valued locally-integrable function on $U$, then the corresponding distribution $\lambda_f$ on $U$ as in (7.2) is nonnegative, for instance.

Let $\lambda$ be a nonnegative linear functional on $C^\infty_{com}(U)$, and let us show that $\lambda$ is automatically continuous on $C^\infty_{com}(U)$. Let us consider the real case first, so that $\lambda$ is a real-linear functional on the space of real-valued smooth functions on $U$ with compact support. Let $K$ be a nonempty compact subset of $U$, and let $\psi_K$ be a nonnegative real-valued smooth function on $U$ with compact support such that $\psi_K(x) \ge 1$ for every $x \in K$, as in Section 5. If $\phi$ is a real-valued smooth function on $U$ with compact support contained in $K$ such that $\|\phi\|_K \le 1$, where $\phi_K$ is as in (2.1), then $\psi_K + \phi$ and $\psi_K - \phi$ are both nonnegative smooth functions on $U$ with compact support. By hypothesis,

$$\lambda(\psi_K + \phi),\ \lambda(\psi_K - \phi) \ge 0, \tag{12.2}$$

which implies that $\pm\lambda(\phi) \le \lambda(\psi_K)$, and hence

$$|\lambda(\phi)| \le \lambda(\psi_K). \tag{12.3}$$

Because of linearity, we get that

$$|\lambda(\phi)| \le \lambda(\psi_K)\, \|\phi\|_K \tag{12.4}$$

for every real-valued smooth function $\phi$ on $U$ with compact support contained in $K$. This implies that $\lambda$ is continuous on $C^\infty_{com}(U)$, since this works for every compact set $K \subseteq U$.

Suppose now that $\lambda$ is a complex-linear functional on the space of complex-valued smooth functions on $U$ with compact support, and that $\lambda$ is nonnegative as before. Let $K$ be a nonempty compact subset of $U$ again, and let $\psi_K$ be as in the preceding paragraph. If $\phi$ is a real-valued smooth function on $U$ with compact support contains in $K$ such that $\|\phi\|_K \le 1$, so that $\psi_K \pm \phi \ge 0$ on $U$, then the nonnegativity condition includes the requirement that $\lambda(\psi_K \pm \phi)$ be a real number, which implies that $\lambda(\phi) \in \mathbf{R}$. In particular, it follows that $\lambda(\phi) \in \mathbf{R}$ for every real-valued smooth function $\phi$ on $U$ with compact support, since this holds for each compact set $K \subseteq U$. If $\phi$ is a complex-valued smooth function on $U$ with compact support contained in $K$, then we can apply the previous argument to the real part of $a\,\phi$ for every $a \in \mathbf{C}$ with $|a| \le 1$. This implies that the real part of $\lambda(a\,\phi) = a\,\lambda(\phi)$ is less than or equal to $\lambda(\psi_K)$, as before. It follows that (12.3) still holds in this case, because this works for every $a \in \mathbf{C}$ with $|a| \le 1$. Thus (12.4) also holds for every complex-valued smooth function $\phi$ on $U$ with compact support contained in $K$, by linearity, which implies that $\lambda$ is continuous on $C^\infty_{com}(U)$, as before.

Because these continuity conditions only involve supremum seminorms of $\phi$, and not derivatives of $\phi$, one can show that there is a unique extension of $\lambda$ to a nonnegative linear functional on $C_{com}(U)$. The Riesz representation theorem then implies that $\lambda(\phi)$ can be expressed as the integral of $\phi$ with respect to a unique locally-finite nonnegative Borel measure on $U$. Conversely, any locally finite nonnegative Borel measure on $U$ determines a nonnegative distribution on $U$ in this way.

# 13 Continuity conditions

Let $U$ be a nonempty open set in $\mathbf{R}^n$, and let $\lambda$ be a distribution on $U$. This means that for each nonempty compact set $K \subseteq U$, the restriction of $\lambda$ to $C^\infty_K(U)$ is continuous, with respect to the topology on $C^\infty_K(U)$ induced by the one on $C^\infty(U)$ discussed in Section 2. Because of the way that this topology is defined, this implies that there are a nonnegative real number $A(K)$ and a nonnegative integer $N(K)$ such that

$$|\lambda(\phi)| \le A(K) \sum_{|\alpha| \le N(K)} \|\phi\|_{K,\alpha} \tag{13.1}$$

for every $\phi \in C^\infty_K(U)$, where $\|\phi\|_{K,\alpha}$ is as in (2.3). More precisely, the sum in (13.1) is taken over the finite set of all multi-indices $\alpha$ such that $|\alpha| \le N(K)$. Conversely, if $\lambda$ is a linear functional on $C^\infty_{com}(U)$, and if for each nonempty compact set $K \subseteq U$ there are $A(K), N(K) \ge 0$ such that (13.1) holds for every $\phi \in C^\infty_K(U)$, then the restriction of $\phi$ to $C^\infty_K(U)$ is continuous for each nonempty compact set $K \subseteq U$, and hence $\phi$ is a distribution on $U$.

If one can take $N(K) = 0$ for every nonempty compact set $K \subseteq U$, then $\lambda$ is said to be a distribution of order 0 on $U$. Thus distributions associated to locally-integrable functions on $U$ as in (7.2) have order 0, as well as nonnegative distributions on $U$. In particular, Dirac delta distributions have order 0, but their derivatives do not.

If $\lambda$ is a distribution on $U$ and $g$ is a smooth function on $U$, then their product $g\,\lambda$ is also a distribution on $U$, as in Section 9. Note that the support of $g\,\lambda$ is contained in the support of $g$, because $(g\,\lambda)(\phi) = \lambda(g\,\phi) = 0$ whenever $\phi \in C^\infty_{com}(U)$ satisfies $g(x)\,\phi(x) = 0$ for every $x \in U$. In particular, $g\,\lambda$ has compact support in $U$ when $g$ has compact support in $U$. Observe also that $g\,\lambda$ has order 0 on $U$ for every $g \in C^\infty(U)$ when $\lambda$ has order 0 on $U$. Thus $g\,\lambda$ is a distribution of order 0 on $U$ with compact support when $\lambda$ has order 0 on $U$ and $g \in C^\infty_{com}(U)$.

If $\lambda$ is a distribution of order 0 on $U$ with compact support, then one can show that there is a unique extension of $\lambda$ to a continuous linear functional on $C(U)$, and which still has compact support in a suitable sense. Using another version of the Riesz representation theorem, one can express $\lambda$ in terms of integration with respect to a unique real or complex Borel measure with compact support on $U$, as appropriate. Conversely, any real or complex Borel measure on $U$ defines a distribution of order 0 on $U$ by integration, and this distribution on $U$ has compact support when the inital measure does.

# 14 Sequences of distributions

Let $U$ be a nonempty open set in $\mathbf{R}^n$. A sequence $\{\lambda_j\}_{j=1}^{\infty}$ of distributions on $U$ is said to converge to a distribution $\lambda$ on $U$ if

$$\lim_{j \to \infty} \lambda_j(\phi) = \lambda(\phi) \tag{14.1}$$

for every $\phi \in C^{\infty}_{com}(U)$. More precisely, this is the same as convergence with respect to the weak$^*$ topology on the space of distributions on $U$, as the dual of $C^{\infty}_{com}(U)$. If $\{\lambda_j\}_{j=1}^{\infty}$ converges to $\lambda$ in this sense and $\alpha$ is any multi-index, then it is easy to see that $\{D^{\alpha}\lambda_j\}_{j=1}^{\infty}$ converges to $D^{\alpha}\lambda$ in the same sense, by applying (14.1) to $D^{\alpha}\phi$ for each $\phi \in C^{\infty}_{com}(U)$.

Let $\{\lambda_j\}_{j=1}^{\infty}$ be a sequence of distributions on $U$ again, and suppose that for each nonempty compact set $K \subseteq U$ there are a nonnegative real number $A(K)$ and a nonnegative integer $N(K)$ such that

$$|\lambda_j(\phi)| \le A(K) \sum_{|\alpha| \le N(K)} \|\phi\|_{K,\alpha} \tag{14.2}$$

for each $\phi \in C^{\infty}_K(U)$ and $j \in \mathbf{Z}_+$. As usual, $\|\phi\|_{K,\alpha}$ is as in (2.3), and the sum in (14.2) is taken over all multi-indices $\alpha$ with $|\alpha| \le N(K)$. As in the previous section, $\lambda_j$ should satisfy conditions like this for each $j$, and the point here is to ask for conditions that are uniform in $j$, which is to say that $A(K)$ and $N(K)$ depend on $K$ and not on $\phi$ or $j$. Suppose also that $\{\lambda_j(\phi)\}_{j=1}^{\infty}$ converges as a sequence of real or complex numbers, as appropriate, for each $\phi \in C^{\infty}_{com}(U)$. Let $\lambda(\phi)$ be the limit of $\{\lambda_j(\phi)\}_{j=1}^{\infty}$ for each $\phi \in C^{\infty}_{com}(U)$. It is easy to see that $\lambda$ is a linear functional on $C^{\infty}_{com}(U)$, since $\lambda_j$ is linear for each $j$. Under these conditions, $\lambda$ also satisfies the analogue of (14.2) with the same choices of $A(K)$ and $N(K)$, which is (13.1). This implies that $\lambda$ is a distribution on $U$, so that $\{\lambda_j\}_{j=1}^{\infty}$ converges to $\lambda$ in the sense described in the previous paragraph.

Suppose now that $\{\lambda_j\}_{j=1}^{\infty}$ is a sequence of distributions on $U$ such that $\{\lambda_j(\phi)\}_{j=1}^{\infty}$ is a bounded sequence in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, for each $\phi$ in $C^{\infty}_{com}(U)$. In particular, this happens when $\{\lambda_j(\phi)\}_{j=1}^{\infty}$ converges in $\mathbf{R}$ or $\mathbf{C}$ for each $\phi \in C^{\infty}_{com}(U)$, since convergent sequences of real or complex numbers are automatically bounded. Under these conditions, a version of the Banach–Steinhaus theorem implies that for each nonempty compact set $K \subseteq U$ there are $A(K), N(K) \ge 0$ such that (14.2) holds for all $\phi \in C^{\infty}_K(U)$ and $j \in \mathbf{Z}_+$. More precisely, one applies the Banach–Steinhaus theorem to the restrictions of the $\lambda_j$'s to $C^{\infty}_K(U)$ for each nonempty compact set $K \subseteq U$, which is possible because $C^{\infty}_K(U)$ is an $F$-space, since it is a Fréchet space. This means that $C^{\infty}_K(U)$ may be considered as a complete metric space for each $K$, which permits one to use the Baire category theorem. Because the $\lambda_j$'s are pointwise bounded on $C^\infty_K(U)$, the Baire category theorem implies that they are uniformly bounded on a nonempty open subset of $C^\infty_K(U)$ for each $K$. Using linearity, one can check that the $\lambda_j$'s are uniformly bounded on an open subset of $C^\infty_K(U)$ that contains 0 for each $K$. This implies that there are $A(K), N(K) \ge 0$ such that (14.2) holds for every $\phi \in C^\infty_K(U)$ and $j \in \mathbf{Z}_+$, because of linearity and the way that the topology on $C^\infty_K(U)$ is defined.

Suppose that $\{\lambda_j\}_{j=1}^\infty$ is a sequence of distributions on $U$ that converges to a distribution $\lambda$ on $U$ as before. Let $K$ be a nonempty compact subset of $U$ again, and suppose also that $\{\phi_j\}_{j=1}^\infty$ is a sequence of smooth functions on $U$ with supports contained in $K$ for each $j$ that converges in the $C^\infty$ topology to a smooth function $\phi$ on $U$, whose support is also contained in $K$. Under these conditions, one can check that

$$\lim_{j\to\infty} \lambda_j(\phi_j) = \phi. \tag{14.3}$$

To see this, it suffices to show that

$$\lim_{j\to\infty} \lambda_j(\phi_j - \phi) = 0, \tag{14.4}$$

since $\{\lambda_j(\phi)\}_{j=1}^\infty$ converges to $\lambda(\phi)$ by hypothesis. In fact, $\lambda_l(\phi_j - \phi) \to 0$ as $j \to \infty$ uniformly in $l$ when $\{\phi_j\}_{j=1}^\infty$ converges to $\phi$ in $C^\infty_K(U)$, because of the uniform boundedness condition (14.2).

# 15 The Schwartz class

Let us now restrict our attention to $U = \mathbf{R}^n$. A smooth function $f$ on $\mathbf{R}^n$ belongs in the Schwartz class $\mathcal{S}(\mathbf{R}^n)$ if for each nonnegative integer $k$ and multi-index $\alpha$, $(1+|x|)^k\,|D^\alpha f|$ is a bounded function on $\mathbf{R}^n$. In this case, we put

$$N_k(f) = \sup_{x\in\mathbf{R}^n} (1+|x|)^k\,|f(x)| \tag{15.1}$$

for each nonnegative integer $k$, and

$$N_{k,\alpha}(f) = N_k(D^\alpha f) = \sup_{x\in\mathbf{R}^n} (1+|x|)^k\,|D^\alpha f(x)| \tag{15.2}$$

for each $k \ge 0$ and multi-index $\alpha$. The collection of all of the $N_{k,\alpha}$'s defines a topology on $\mathcal{S}(\mathbf{R}^n)$ in the usual way, so that $\mathcal{S}(\mathbf{R}^n)$ becomes a locally convex topological vector space. Because the collection of the $N_{k,\alpha}$'s is countable, there is a countable local base for this topology on $\mathcal{S}(\mathbf{R}^n)$ at 0, and in particular this topology can be described by a translation-invariant metric on $\mathcal{S}(\mathbf{R}^n)$.

Of course, every element of $\mathcal{S}(\mathbf{R}^n)$ is a smooth function on $\mathbf{R}^n$, and every smooth function on $\mathbf{R}^n$ with compact support satisfies the requirements in the previous paragraph, so that

$$C^\infty_{com}(\mathbf{R}^n) \subseteq \mathcal{S}(\mathbf{R}^n) \subseteq C^\infty(\mathbf{R}^n). \tag{15.3}$$

Note that $f(x) = \exp(-|x|^2)$ is a smooth function on $\mathbf{R}^n$ in $\mathcal{S}(\mathbf{R}^n)$ that does not have compact support. If $K$ is any nonempty compact subset of $\mathbf{R}^n$, then

$$\|f\|_{K,\alpha} \le N_{0,\alpha}(f) \tag{15.4}$$

for every $f \in \mathcal{S}(\mathbf{R}^n)$ and multi-index $\alpha$, where $\|f\|_{K,\alpha}$ is as in (2.3). This implies that the natural inclusion mapping of $\mathcal{S}(\mathbf{R}^n)$ in $C^\infty(\mathbf{R}^n)$ is continuous with respect to the topolgies that have been defined on these spaces.

Similarly, if $f$ is a smooth function on $\mathbf{R}^n$ with support contained in a nonempty compact set $K \subseteq \mathbf{R}^n$, then

$$N_{k,\alpha}(f) \le \Big(\sup_{x \in K} (1 + |x|)^k\Big) \, \|f\|_{K,\alpha} \tag{15.5}$$

for every multi-index $\alpha$. This implies that the natural inclusion of $C^\infty_K(\mathbf{R}^n)$ in $\mathcal{S}(\mathbf{R}^n)$ is continuous, using the topology on $C^\infty_K(\mathbf{R}^n)$ induced by the usual topology on $C^\infty(\mathbf{R}^n)$. It follows that the natural inclusion of $C^\infty_{com}(\mathbf{R}^n)$ in $\mathcal{S}(\mathbf{R}^n)$ is also continuous, since its restriction to $C^\infty_K(\mathbf{R}^n)$ is continuous for each nonempty compact set $K \subseteq \mathbf{R}^n$.

As usual, a sequence $\{f_j\}_{j=1}^\infty$ of functions in $\mathcal{S}(\mathbf{R}^n)$ converges to $f \in \mathcal{S}(\mathbf{R}^n)$ with respect to the topology on $\mathcal{S}(\mathbf{R}^n)$ defined by the $N_{k,\alpha}$'s if and only if

$$\lim_{j \to \infty} N_{k,\alpha}(f_j - f) = 0 \tag{15.6}$$

for each $k \ge 0$ and multi-index $\alpha$. Similarly, a sequence $\{f_j\}_{j=1}^\infty$ of functions in $\mathcal{S}(\mathbf{R}^n)$ is a Cauchy sequence if

$$\lim_{j,l \to \infty} N_{k,\alpha}(f_j - f_l) = 0 \tag{15.7}$$

for each $k \ge 0$ and multi-index $\alpha$. If $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $\mathcal{S}(\mathbf{R}^n)$, then $\{f_j\}_{j=1}^\infty$ is a Cauchy sequence in $C^\infty(\mathbf{R}^n)$ in particular, which implies that $\{f_j\}_{j=1}^\infty$ converges to a smooth function $f$ on $\mathbf{R}^n$ in $C^\infty(\mathbf{R}^n)$, as in Section 4. Under these conditions, one can check that $f \in \mathcal{S}(\mathbf{R}^n)$, and that $\{f_j\}_{j=1}^\infty$ also converges to $f$ in $\mathcal{S}(\mathbf{R}^n)$ in the sense of (15.6), using (15.7). Thus every Cauchy sequence in $\mathcal{S}(\mathbf{R}^n)$ converges, so that $\mathcal{S}(\mathbf{R}^n)$ is complete, and hence a Fréchet space, since the topology on $\mathcal{S}(\mathbf{R}^n)$ is defined by a countable collection of seminorms.

Let $\theta$ be a smooth function on $\mathbf{R}^n$ with compact support such that $\theta(x) = 1$ when $|x| \le 1$. Put

$$\theta_R(x) = \theta(x/R) \tag{15.8}$$

for each $R > 0$, so that $\theta_R$ is a smooth function on $\mathbf{R}^n$ with compact support such that $\theta_R(x) = 1$ when $|x| \le R$. If $f \in \mathcal{S}(\mathbf{R}^n)$, then one can check that $\theta_R\, f \to f$ in $\mathcal{S}(\mathbf{R}^n)$ as $R \to \infty$, which means that

$$\lim_{R \to \infty} N_{k,\alpha}(f - \theta_R\, f) = 0 \tag{15.9}$$

for each $k \ge 0$ and multi-index $\alpha$. This implies that $C^\infty_{com}(\mathbf{R}^n)$ is dense in $\mathcal{S}(\mathbf{R}^n)$, since $\theta_R\, f$ has compact support in $\mathbf{R}^n$ for each $R > 0$.

If $f \in \mathcal{S}(\mathbf{R}^n)$ and $\alpha$ is a multi-index, then it is easy to see that $D^\alpha f \in \mathcal{S}(\mathbf{R}^n)$ too. Moreover, $D^\alpha$ defines a continuous linear mapping from $\mathcal{S}(\mathbf{R}^n)$ into itself. This follows from the fact that

$$N_{k,\beta}(D^\alpha f) = N_{k,\alpha+\beta}(f) \tag{15.10}$$

for each $k \geq 0$, multi-index $\beta$, and $f \in \mathcal{S}(\mathbf{R}^n)$, since $D^\beta(D^\alpha f) = D^{\alpha+\beta} f$.

If $f \in \mathcal{S}(\mathbf{R}^n)$ and $p(x)$ is a polynomial on $\mathbf{R}^n$, then one can check that the product $p\, f$ is in $\mathcal{S}(\mathbf{R}^n)$ as well. As an extension of this, let $g$ be a smooth function on $\mathbf{R}^n$, and suppose that for each multi-index $\alpha$ there are a nonnegative integer $k(\alpha)$ and a nonnegative real number $C(\alpha)$ such that

$$|D^\alpha g(x)| \leq C(\alpha)\, (1 + |x|)^{k(\alpha)} \tag{15.11}$$

for every $x \in \mathbf{R}^n$. Under these conditions, one can check that the product $f\, g$ is also in $\mathcal{S}(\mathbf{R}^n)$. More precisely, the mapping that sends $f \in \mathcal{S}(\mathbf{R}^n)$ to $f\, g$ is a continuous linear mapping from $\mathcal{S}(\mathbf{R}^n)$ into itself.

# 16 Tempered distributions

A *tempered distribution* on $\mathbf{R}^n$ is a continuous linear functional on $\mathcal{S}(\mathbf{R}^n)$, which is to say a linear functional on $\mathcal{S}(\mathbf{R}^n)$ that is continuous with respect to the topology defined by the $N_{k,\alpha}$'s, as in the previous section. If $\lambda$ is a tempered distribution on $\mathbf{R}^n$, then $\lambda$ determines an ordinary distribution on $\mathbf{R}^n$, since the restriction of $\lambda$ to $C^\infty_{com}(\mathbf{R}^n) \subseteq \mathcal{S}(\mathbf{R}^n)$ is a continuous linear functional on $C^\infty_{com}(\mathbf{R}^n)$. More precisely, this uses the fact that the natural inclusion of $C^\infty_{com}(\mathbf{R}^n)$ in $\mathcal{S}(\mathbf{R}^n)$ is continuous. Note that a continuous linear functional $\lambda$ on $\mathcal{S}(\mathbf{R}^n)$ is uniquely determined by its restriction to $C^\infty_{com}(\mathbf{R}^n)$, because $C^\infty_{com}(\mathbf{R}^n)$ is dense in $\mathcal{S}(\mathbf{R}^n)$, as in the previous section.

Suppose that $f$ is a locally integrable function on $\mathbf{R}^n$ such that

$$\int_{\mathbf{R}^n} (1 + |x|)^{-k}\, |f(x)|\, dx < +\infty \tag{16.1}$$

for some nonnegative integer $k$. In this case, the product $\phi\, f$ is an integrable function on $\mathbf{R}^n$ for every $\phi \in \mathcal{S}(\mathbf{R}^n)$, and we put

$$\lambda_f(\phi) = \int_{\mathbf{R}^n} \phi(x)\, f(x)\, dx. \tag{16.2}$$

This defines a continuous linear functional on $\mathcal{S}(\mathbf{R}^n)$, and hence a tempered distribution on $\mathbf{R}^n$, because

$$|\lambda_f(\phi)| \leq \int_{\mathbf{R}^n} |\phi(x)|\, |f(x)|\, dx \leq N_{k,0}(\phi) \int_{\mathbf{R}^n} (1 + |x|)^{-k} |f(x)|\, dx \tag{16.3}$$

for every $\phi \in \mathcal{S}(\mathbf{R}^n)$. Similarly, if $\mu$ is a positive Borel measure on $\mathbf{R}^n$ such that

$$\int_{\mathbf{R}^n} (1 + |x|)^{-k}\, d\mu(x) < +\infty \tag{16.4}$$

for some nonnegative integer $k$, then

$$\lambda_\mu(\phi) = \int_{\mathbf{R}^n} \phi(x)\, d\mu(x) \tag{16.5}$$

defines a tempered distribution on $\mathbf{R}^n$. If $\lambda$ is an ordinary distribution on $\mathbf{R}^n$ with compact support, then we have seen in Section 11 that $\lambda$ has a natural continuous extension to $C^\infty(\mathbf{R}^n)$, whose restriction to $\mathcal{S}(\mathbf{R}^n)$ defines a tempered distribution on $\mathbf{R}^n$.

If $\lambda$ is a tempered distribution on $\mathbf{R}^n$ and $\alpha$ is a multi-index, then we can define $D^\alpha \lambda$ in the same way as for ordinary distributions, as in (8.2). More precisely, this is a continuous linear functional on $\mathcal{S}(\mathbf{R}^n)$ when $\lambda$ is, because $D^\alpha$ is a continuous linear mapping from $\mathcal{S}(\mathbf{R}^n)$ into itself, as in the previous section. Now let $g$ be a smooth function on $\mathbf{R}^n$ with the property that for each multi-index $\alpha$ there are $k(\alpha), C(\alpha) \geq 0$ such that (15.11) holds for every $x \in \mathbf{R}^n$. If $\lambda$ is a tempered distribution on $\mathbf{R}^n$, then the product of $\lambda$ and $g$ can be defined as a tempered distribution on $\mathbf{R}^n$, as in Section 9. Note that the derivatives of $g$ satisfy analogous conditions, and that product rule for the derivative of a product also works in this situation, as before.

If $\lambda$ is a tempered distribution on $\mathbf{R}^n$, then there are nonnegative integers $k$, $L$ and a nonnegative real number $A$ such that

$$|\lambda(\phi)| \leq A \sum_{|\alpha| \leq L} N_{k,\alpha}(\phi) \tag{16.6}$$

for every $\phi \in \mathcal{S}(\mathbf{R}^n)$, where the sum is taken over all multi-indices $\alpha$ such that $|\alpha| \leq L$. This follows from the way that the topology on $\mathcal{S}(\mathbf{R}^n)$ is defined, and conversely a linear functional $\lambda$ on $\mathbf{S}(\mathbf{R}^n)$ is continuous if it satisfies a condition of this type.

A sequence $\{\lambda_j\}_{j=1}^\infty$ of tempered distributions on $\mathbf{R}^n$ is said to converge to a tempered distribution $\lambda$ on $\mathbf{R}^n$ if

$$\lim_{j \to \infty} \lambda_j(\phi) = \lambda(\phi) \tag{16.7}$$

for every $\phi \in \mathcal{S}(\mathbf{R}^n)$. As before, this is the same as convergence with respect to the weak$^*$ topology on the space of tempered distributions on $\mathbf{R}^n$, as the dual of $\mathcal{S}(\mathbf{R}^n)$. If $\{\lambda_j\}_{j=1}^\infty$ converges to $\lambda$ in this sense, and if $\alpha$ is any multi-index, then it is easy to see that $\{D^\alpha \lambda_j\}_{j=1}^\infty$ converges to $D^\alpha \lambda$ in the same sense, as in the case of ordinary distributions on $\mathbf{R}^n$.

Suppose that $\{\lambda_j\}_{j=1}^\infty$ is a sequence of tempered distributions on $\mathbf{R}^n$ such that $\{\lambda_j(\phi)\}_{j=1}^\infty$ is a bounded sequence in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, for each $\phi \in \mathcal{S}(\mathbf{R}^n)$. As in Section 14, a version of the Banach–Steinhaus theorem implies that there are nonnegative integers $k$, $L$ and a nonnegative real number $A$ such that

$$|\lambda_j(\phi)| \leq A \sum_{|\alpha| \leq L} N_{k,\alpha}(\phi) \tag{16.8}$$

for every $\phi \in \mathcal{S}(\mathbf{R}^n)$ and $j \in \mathbf{Z}_+$, where the sum is again taken over all multi-indices $\alpha$ with $|\alpha| \le L$. This uses the fact that $\mathcal{S}(\mathbf{R}^n)$ is a Fréchet space, as in the previous section.

In particular, if $\{\lambda_j(\phi)\}_{j=1}^\infty$ converges in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, for each $\phi$ in $\mathcal{S}(\mathbf{R}^n)$, then $\{\lambda_j(\phi)\}_{j=1}^\infty$ is bounded for each $\phi \in \mathcal{S}(\mathbf{R}^n)$. In this case, one can define a linear functional $\lambda$ on $\mathcal{S}(\mathbf{R}^n)$ by (16.7), and the uniform boundedness condition (16.8) implies that $\lambda$ satisfies (16.6) for each $\phi \in \mathcal{S}(\mathbf{R}^n)$, with the same choices of $A$, $k$, and $L$. This implies that $\lambda$ is also continuous on $\mathcal{S}(\mathbf{R}^n)$.

If $\{\lambda_j\}_{j=1}^\infty$ is a sequence of tempered distributions on $\mathbf{R}^n$ that converges to a tempered distribution $\lambda$ on $\mathbf{R}^n$ as in (16.7), and if $\{\phi_j\}_{j=1}^\infty$ is a sequence of elements of $\mathcal{S}(\mathbf{R}^n)$ that converges to $\phi \in \mathcal{S}(\mathbf{R}^n)$, then

$$\lim_{j\to\infty} \lambda_j(\phi_j) = \lambda(\phi). \tag{16.9}$$

As in Section 14, one can use the uniform boundedness condition (16.8), to get that $\lambda_l(\phi_j - \phi) \to 0$ as $j \to \infty$ uniformly over $l$. This permits (16.9) to be reduced to (16.7), as before.

# 17 Bounded sets of functions

Let $U$ be a nonempty open set in $\mathbf{R}^n$, and let $E$ be a subset of $C^k(U)$ for some $k$, $0 \le k \le +\infty$. We say that $E$ is *bounded* in $C^k(U)$ if for each nonempty compact set $K \subseteq U$ and multi-index $\alpha$ with $|\alpha| \le k$ there is a nonnegative real number $C(K, \alpha)$ such that

$$\|f\|_{K,\alpha} \le C(K, \alpha) \tag{17.1}$$

for every $f \in E$, where $\|f\|_{K,\alpha}$ is as in (2.3). There is a more abstract definition of the boundedness of a subset $E$ of a topological vector space $V$ over the real or complex numbers, which reduces to this condition in this situation. More precisely, if the topology on $V$ is determined by a collection $\mathcal{N}$ of seminorms on $V$, then $E \subseteq V$ is bounded if each seminorm $N \in \mathcal{N}$ is bounded on $E$.

Suppose now that $E \subseteq C^k_{com}(U)$ for some $k$, $0 \le k \le \infty$. In this case, we say that $E$ is bounded in $C^k_{com}(U)$ if there is a compact set $K \subseteq U$ such that $E$ is a bounded subset of $C^k_K(U)$. This means that $E \subseteq C^k_K(U)$, and that for each multi-index $\alpha$ with $|\alpha| \le k$ there is a nonnegative real number $C(\alpha)$ such that

$$\|f\|_{K,\alpha} \le C(\alpha) \tag{17.2}$$

for every $f \in E$, which is the same as saying that $E \subseteq C^k_K(U)$ is bounded as a subset of $C^k(U)$. It is well known that this is equivalent to the boundedness of $E$ as a subset of $C^k_{com}(U)$ as a topological vector space, with respect to the appropriate inductive limit topology.

Now let $E$ be a subset of the Schwartz class $\mathcal{S}(\mathbf{R}^n)$. We say that $E$ is bounded in $\mathcal{S}(\mathbf{R}^n)$ if for each nonnegative integer $k$ and multi-index $\alpha$ there is a nonnegative real number $C(k, \alpha)$ such that

$$N_{k,\alpha}(f) \le C(k, \alpha) \tag{17.3}$$

for every $f \in E$, where $N_{k,\alpha}(f)$ is as in (15.2). As before, this is equivalent to the boundedness of $E$ as a subset of $\mathcal{S}(\mathbf{R}^n)$ as a topological vector space. Note that bounded subsets of $C^\infty_{com}(\mathbf{R}^n)$ are bounded as subsets of $\mathcal{S}(\mathbf{R}^n)$, and that bounded subsets of $\mathcal{S}(\mathbf{R}^n)$ are bounded as subsets of $C^\infty(\mathbf{R}^n)$.

Let $\{f_j\}_{j=1}^\infty$ be a sequence of functions in $\mathcal{S}(\mathbf{R}^n)$ which is bounded as in the previous paragraph, so that for each nonnegative integer $k$ and multi-index $\alpha$ there is a nonnegative real number $C(k,\alpha)$ such that

$$N_{k,\alpha}(f_j) \le C(k,\alpha) \tag{17.4}$$

for every $j \ge 1$. Suppose that $\{f_j\}_{j=1}^\infty$ converges to a smooth function $f$ on $\mathbf{R}^n$ in $C^\infty(\mathbf{R}^n)$, which means that $\{D^\alpha f_j\}_{j=1}^\infty$ converges to $D^\alpha f$ uniformly on compact subsets of $\mathbf{R}^n$ for every multi-index $\alpha$. Under these conditions, it is easy to see that $f \in \mathcal{S}(\mathbf{R}^n)$ too, with $N_{k,\alpha}(f) \le C(k,\alpha)$ for each $k \ge 0$ and multi-index $\alpha$. One can also check that $\{f_j\}_{j=1}^\infty$ converges to $f$ in $\mathcal{S}(\mathbf{R}^n)$, so that $N_{k,\alpha}(f_j - f) \to 0$ as $j \to \infty$ for every $k \ge 0$ and multi-index $\alpha$. More precisely, for a fixed $k$ and $\alpha$, one can use the boundedness of $N_{k+1,\alpha}(f_j)$ to get the convergence of $\{f_j\}_{j=1}^\infty$ to $f$ with respect to $N_{k,\alpha}$ from the uniform convergence of $\{D^\alpha f_j\}_{j=1}^\infty$ to $D^\alpha f$ on compact subsets of $\mathbf{R}^n$. Similarly, if $E$ is a bounded subset of $\mathcal{S}(\mathbf{R}^n)$, then the topology on $E$ induced by the one on $\mathcal{S}(\mathbf{R}^n)$ is the same as the topology on $E$ induced by the one on $C^\infty(\mathbf{R}^n)$. This can be derived from the previous remarks about sequences, or verified more directly from the definitions using analogous arguments.

# 18 Compactness and equicontinuity

Let $(M, d(x,y))$ be a metric space. Remember that a set $E \subseteq M$ is said to be *sequentially compact* if every sequence of elements of $E$ has a subsequence that converges to an element of $E$. It is well known that sequential compactness is equivalent to the usual definition of compactness in terms of open coverings in the context os subsets of metric spaces. A set $E \subseteq M$ is said to be *totally bounded* if for every $\epsilon > 0$, $E$ is contained in the union of finitely many open balls in $M$ with radius $\epsilon$. If $M$ is a complete metric space, in the sense that every Cauchy sequence in $M$ converges to an element of $M$, then it is also well known that a set $E \subseteq M$ is compact if and only if $E$ is closed in $M$ and totally bounded. Another result of this type states that $E \subseteq M$ is totally bounded if and only if every sequence of elements of $E$ has a subsequence which is a Cauchy sequence. Indeed, if $E$ is not totally bounded, then there is an $\epsilon > 0$ and a sequence $\{x_j\}_{j=1}^\infty$ of elements of $E$ such that $d(x_j, x_l) \ge \epsilon$ when $j \ne l$, and it is easy to see that such a sequence cannot have a Cauchy sequence. Conversely, if $E$ is totally bounded and $\{x_j\}_{j=1}^\infty$ is a sequence of elements of $E$, then for each $\epsilon > 0$ there is a subsequence of $\{x_j\}_{j=1}^\infty$ whose terms are contained in a ball of radius $\epsilon$. To get a subsequence of $\{x_j\}_{j=1}^\infty$ which is a Cauchy sequence, one can use Cantor's diagonalization argument, applied to a sequence of successive subsequences of $\{x_j\}_{j=1}^\infty$ whose terms are contained in open balls in $M$ with radii converging to 0.

Now let $X$ be a compact topological space, and let $C(X)$ be the space of real or complex-valued continuous functions on $X$, equipped with the supremum norm. In particular, $C(X)$ is a metric space with respect to the supremum metric, and it is well known that $C(X)$ is complete. Let $E$ be a subset of $C(X)$, and suppose that $E$ is *pointwise bounded* on $X$, in the sense that for each $x \in X$,

$$E(x) = \{f(x) : f \in E\} \tag{18.1}$$

is a bounded subset of the real or complex numbers, as appropriate. Suppose also that $E$ is equicontinuous at every point in $X$, which means that for each $\epsilon > 0$ and $x \in X$ there is an open set $U$ in $X$ such that $x \in X$ and

$$|f(x) - f(y)| < \epsilon \tag{18.2}$$

for every $f \in E$ and $y \in U$. Under these conditions, it is well known that $E$ is totally bounded in $C(X)$ with respect to the supremum metric.

Alternatively, let $X$ be a topological space which is not necessarily compact, and let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued continuous functions on $X$. Suppose that $\{f_j\}_{j=1}^\infty$ is bounded pointwise on $X$, so that $\{f_j(x)\}_{j=1}^\infty$ is a bounded sequence of real or complex numbers, as appropriate, for each $x \in X$. This implies that for each $x \in X$ there is a subsequence $\{f_{j_l}\}_{l=1}^\infty$ of $\{f_j\}_{j=1}^\infty$ such that $\{f_{j_l}(x)\}_{j=1}^\infty$ converges in $\mathbf{R}$ or $\mathbf{C}$, since closed and bounded subsets of the real and complex numbers are compact. If $A$ is a countable subset of $X$, then one can use a Cantor diagonalization argument to get a subsequence $\{f_{j_l}\}_{l=1}^\infty$ of $\{f_j\}_{j=1}^\infty$ such that $\{f_{j_l}(x)\}_{l=1}^\infty$ converges in $\mathbf{R}$ or $\mathbf{C}$ for every $x \in A$.

Suppose that $A$ is a countable dense subset of $X$, and that the set of $f_j$'s is equicontinuous at every point in $X$. Under these conditions, one can check that $\{f_{j_l}(x)\}_{j=1}^\infty$ is a Cauchy sequence in $\mathbf{R}$ or $\mathbf{C}$ for every $x \in X$, and hence that $\{f_{j_l}(x)\}_{j=1}^\infty$ converges in $\mathbf{R}$ or $\mathbf{C}$ for every $x \in X$. One can also show that the limit $f(x)$ of $\{f_{j_l}(x)\}_{l=1}^\infty$ is a continuous function on $X$, using the equicontinuity of the $f_j$'s again. In addition, $\{f_{j_l}\}_{l=1}^\infty$ converges to $f$ uniformly on compact subsets of $X$ in this situation. In particular, $\{f_{j_l}\}_{l=1}^\infty$ converges to $f$ uniformly on $X$ when $X$ is compact.

# 19 Compact sets of functions

Let $U$ be a nonempty open set in $\mathbf{R}^n$, and let $\{f_j\}_{j=1}^\infty$ be a sequence of functions in $C^1(U)$ such that the set of $f_j$'s is bounded as a subset of $C^1(U)$, in the sense of Section 17. This means that for each compact set $K \subseteq U$, the $f_j$'s and their first derivatives are uniformly bounded on $K$. In particular, $\{f_j(x)\}_{j=1}^\infty$ is a bounded sequence of real or complex numbers, as appropriate, for each $x \in U$. Also, for each $x \in U$, there is a positive real number $r(x)$ such that the closed ball $\overline{B}(x, r(x))$ with center $x$ and radius $r(x)$ in $\mathbf{R}^n$ is contained in $U$, and hence is a compact subset of $U$. The boundedness of the set of $f_j$'s in $C^1(U)$ implies that the first derivatives of the $f_j$'s are uniformly bounded on $\overline{B}(x, r(x))$, so that the set of $f_j$'s is equicontinuous at $x$. Of course, there is a countable dense set in $U$, such as the set of points in $U$ with rational coordinates. Under these conditions, the Arzela–Ascoli type of arguments discussed in the previous section imply that there is a subsequence $\{f_{j_l}\}_{l=1}^\infty$ of $\{f_j\}_{j=1}^\infty$ that converges to a continuous function $f$ on $U$ uniformly on compact subsets of $U$.

Now let $k$ be a positive integer, and suppose that $\{f_j\}_{j=1}^\infty$ is a sequence of functions in $C^k(U)$ such that the set of $f_j$'s is bounded in $C^k(U)$. As before, this means that for each compact set $K \subseteq U$ and multi-index $\alpha$ with $|\alpha| \le k$, the functions $D^\alpha f_j$ are uniformly bounded on $K$. Under these conditions, for each multi-index $\alpha$ with $|\alpha| \le k-1$, there is a subsequence of $\{D^\alpha f_j\}_{j=1}^\infty$ that converges uniformly on compact subsets of $U$, as in the previous paragraph. By applying this to each such multi-index $\alpha$ one at a time, and passing to suitable subsequences at each step, one can get a single subsequence $\{f_{j_l}\}_{l=1}^\infty$ of $\{f_j\}_{j=1}^\infty$ such that $\{D^\alpha f_{j_l}\}_{l=1}^\infty$ converges uniformly on compact subsets of $U$ for each multi-index $\alpha$ with $|\alpha| \le k-1$. As in Section 4, well-known results in analysis imply that the limit $f$ of $\{f_{j_l}\}_{l=1}^\infty$ is in $C^{k-1}(U)$, and that the limit of $\{D^\alpha f_{j_l}\}_{l=1}^\infty$ is equal to $D^\alpha f$ for each multi-index $\alpha$ with $|\alpha| \le k-1$.

Similarly, if $\{f_j\}_{j=1}^\infty$ is a sequence of smooth functions on $U$ such that the set of $f_j$'s is bounded in $C^\infty(U)$, then there is a subsequence $\{f_{j_l}\}_{l=1}^\infty$ of $\{f_j\}_{j=1}^\infty$ that converges to a smooth function $f$ on $U$ in $C^\infty(U)$. This uses an additional Cantor diagonalization argument, to get a subsequence of $\{f_j\}_{j=1}^\infty$ for which all derivatives converge uniformly on compact subsets of $U$. It follows that a closed and bounded set $E$ in $C^\infty(U)$ is compact with respect to the usual topology on $C^\infty(U)$. There are analogous statements for bounded sequences in the Schwartz class $\mathcal{S}(\mathbf{R}^n)$, and for closed and bounded subsets of $\mathcal{S}(\mathbf{R}^n)$. These statements can be derived from the corresponding ones for $C^\infty(\mathbf{R}^n)$, because of the compatibility between the topologies on $\mathcal{S}(\mathbf{R}^n)$ and $C^\infty(\mathbf{R}^n)$ on bounded subsets of $\mathcal{S}(\mathbf{R}^n)$, as in Section 17.

# 20 Riemann–Stieltjes integrals

Let $\alpha(x)$ be a monotone increasing real-valued function on the real line. It is well known that the one-sided limits of $\alpha$ exist at each point $x \in \mathbf{R}$, and are given by

$$\alpha(x+) = \lim_{t \to x+} \alpha(t) = \inf\{\alpha(t) : t \in \mathbf{R},\, t > x\} \tag{20.1}$$

and

$$\alpha(x-) = \lim_{t \to x-} \alpha(t) = \sup\{\alpha(t) : t \in \mathbf{R},\, t < x\}. \tag{20.2}$$

In particular,

$$\alpha(x-) \le \alpha(x) \le \alpha(x+) \tag{20.3}$$

for every $x \in \mathbf{R}$, and $\alpha(x-) = \alpha(x) = \alpha(x+)$ if and only if $\alpha$ is continuous at $x$. Similarly,

$$\alpha(x+) \le \alpha(y-) \tag{20.4}$$

when $x < y$, because $\alpha(x+) \le \alpha(t) \le \alpha(y-)$ when $x < t < y$.

Let $A$ be the set of $x \in \mathbf{R}$ such that $\alpha$ is not continuous at $x$, which happens if and only if $\alpha(x-) < \alpha(x+)$. Put

$$I(x) = (\alpha(x-), \alpha(x+)) \tag{20.5}$$

for each $x \in A$, and observe that

$$I(x) \cap I(y) = \emptyset \tag{20.6}$$

when $x, y \in A$ and $x < y$, because of (20.4). Let $r(x)$ be a rational number contained in $I(x)$ for each $x \in A$, so that

$$r(x) < r(y) \tag{20.7}$$

when $x, y \in A$ and $x < y$, by (20.4) again. This implies that $A$ has only finitely or countably many elements, because the set $\mathbf{Q}$ of rational numbers has only finitely or countably many elements.

Let $a$, $b$ be real numbers with $a < b$, and let $\phi$ be a continuous real-valued function on $[a, b]$. The Riemann–Stieltjes integral

$$\int_a^b \phi \, d\alpha \tag{20.8}$$

can be defined in essentially the same way as for the ordinary Riemann integral, which corresponds to the case where $\alpha(x) = x$ for every $x \in \mathbf{R}$. The main difference is to measure the length of an interval using $\alpha$, by taking the difference of the values of $\alpha$ at the endpoints. As in the classical case, the existence of the integral for a continuous function $\phi$ on $[a, b]$ uses the fact that continuous functions on closed intervals in the real line are uniformly continuous. Note that (20.8) is a nonnegative real number when $\phi(x) \geq 0$ for every $x \in [a, b]$, and that

$$\left| \int_a^b \phi \, d\alpha \right| \leq \int_a^b |\phi| \, d\alpha \leq \Big( \sup_{a \leq x \leq b} |\phi(x)| \Big) \, (\alpha(b) - \alpha(a)) \tag{20.9}$$

for every continuous function $\phi$ on $[a, b]$.

Let $\lambda_\alpha$ be the distribution on the real line corresponding to $\alpha$ as a locally integrable function on $\mathbf{R}$, so that

$$\lambda_\alpha(\phi) = \int_{\mathbf{R}} \phi(x) \, \alpha(x) \, dx \tag{20.10}$$

for every $\phi \in C^\infty_{com}(\mathbf{R})$, as in (7.2). Using an appropriate version of integration by parts, one can show that

$$(D\lambda_\alpha)(\phi) = -\lambda_\alpha(D\phi) = -\int_{\mathbf{R}} (D\phi)(x) \, \alpha(x) \, dx = \int_{\mathbf{R}} \phi \, d\alpha \tag{20.11}$$

for every $\phi \in C^\infty_{com}(\mathbf{R})$, where the last integral is the Riemann–Stieltjes integral (20.8) over any interval $[a, b]$ that contains the support of $\phi$. In particular, the derivative of $\lambda_\alpha$ is nonnegative as a distribution on $\mathbf{R}$, and jump discontinuities in $\alpha$ correspond to Dirac masses in the derivative of $\lambda_\alpha$ in the usual way.

# 21 Translations

Let $U$ be a nonempty open set in $\mathbf{R}^n$, let $a$ be an element of $\mathbf{R}^n$, and put

$$U + a = \{x + a : x \in U\}, \tag{21.1}$$

which is also an open set in $\mathbf{R}^n$. If $f$ is a function on $U$, then let $\tau_a(f)$ be the function on $U + a$ defined by

$$(\tau_a(f))(x) = f(x - a) \tag{21.2}$$

for each $x \in U + a$. Of course, this preserves integrability and smoothness properties of $f$, and it is easy to see that the support of $\tau_a(f)$ is the same as the support of $f$ translated by $a$. In particular, if $f$ has compact support in $U$, then $\tau_a(f)$ has compact support in $U + a$.

Let $\lambda$ be a distribution on $U$, and define a distribution $\tau_a(\lambda)$ on $U + a$ by

$$(\tau_a(\lambda))(\phi) = \lambda(\tau_{-a}(\phi)) \tag{21.3}$$

for each $\phi \in C^\infty_{com}(U + a)$. More precisely, if $\phi \in C^\infty_{com}(U + a)$, then $\tau_{-a}(\phi)$ is an element of $C^\infty_{com}(U)$, as in the previous paragraph, so that $\lambda(\tau_{-a}(\phi))$ makes sense. It is easy to see that $\tau_{-a}$ defines a continuous linear mapping from $C^\infty_{com}(U+a)$ onto $C^\infty_{com}(U)$, which implies that $\tau_a(\lambda)$ is a continuous linear functional on $C^\infty_{com}(U+a)$ when $\lambda$ is a continuous linear functional on $C^\infty_{com}(U)$. If $\lambda_f$ is the distribution on $U$ corresponding to a locally integrable function $f$ on $U$, as in (7.2), then

$$\begin{aligned} (\tau_a(\lambda_f))(\phi) = \lambda_f(\tau_{-a}(\phi)) &= \int_U \phi(x + a)\, f(x)\, dx \\ &= \int_{U+a} \phi(x)\, f(x - a)\, dx \end{aligned} \tag{21.4}$$

for every $\phi \in C^\infty_{com}(U + a)$, and hence $\tau_a(\lambda_f) = \lambda_{\tau_a(f)}$. Similarly, if $\delta_p$ is the Dirac delta distribution corresponding to a point $p \in U$ as in (7.4), then

$$(\tau_a(\delta_p))(\phi) = \delta_p(\tau_{-a}(\phi)) = \phi(p + a) \tag{21.5}$$

for each $\phi \in C^\infty_{com}(U + a)$, so that $\tau_a(\delta_p) = \delta_{p+a}$.

Let $r$ be a positive real number, and let $U_r$ be the set of $x \in U$ such that

$$B(x, t) \subseteq U \tag{21.6}$$

for some $t > r$. Equivalently, $x \in U_r$ if and only if $\overline{B}(x, r) \subseteq U$. Note that $U_r$ is an open set in $\mathbf{R}^n$ for each $r > 0$, $U_r \subseteq U_t$ when $r < t$, and every $x \in U$ is in $U_r$ for sufficiently small $r$. If $K$ is a nonempty compact subset of $U_r$, then put

$$K(r) = \bigcup_{p \in K} \overline{B}(p, r), \tag{21.7}$$

which is the same as

$$K(r) = \{x \in \mathbf{R}^n : \mathrm{dist}(x, K) \le r\}, \tag{21.8}$$

by well known properties of compact sets. The first description of $K(r)$ implies that $K(r) \subseteq U$ when $K \subseteq U_r$, and the second description implies that $K(r)$ is closed and bounded and thus compact.

If $a \in \mathbf{R}^n$ and $|a| \le r$, then $U_r - a \subseteq U$, so that $U_r \subseteq U + a$, and the restriction of $\tau_a(\lambda)$ to $U_r$ is defined as a distribution on $U_r$ as in Section 7. Let $K$ be a nonempty compact subset of $U_r$, and let $\phi$ be a smooth function on $U_r$ with compact support contained in $K$. As usual, we can extend $\phi$ to a smooth function on $\mathbf{R}^n$ with compact support contained in $K$, by setting $\phi(x) = 0$ for every $x \in \mathbf{R}^n \backslash K$. Note that

$$\operatorname{supp} \tau_a(\phi) = (\operatorname{supp} \phi) + a \subseteq K(r) \tag{21.9}$$

under these conditions. Also, if $K$ is any compact subset of $U$, then $K \subseteq U_r$ when $r > 0$ is sufficiently small.

Let $j$ be a positive integer less than or equal to $n$, and let $e_j \in \mathbf{R}^n$ be the corresponding standard basis vector in $\mathbf{R}^n$, with $j$th coordinate equal to 1 and all other coordinates equal to 0. If $h$ is a nonzero real number with $|h| \le r$, then

$$\mu_{j,h} = h^{-1}\,(\tau_{-h\,e_j}(\lambda) - \lambda) \tag{21.10}$$

is defined as a distribution on $U_r$, by restricting $\lambda$ and $\tau_{-h\,e_j}(\lambda)$ to $U_r$ as in the previous paragraph. If $K$ is a compact subset of $U_r$ and $\phi$ is a smooth function on $\mathbf{R}^n$ with support contained in $K$, then

$$\mu_{j,h}(\phi) = h^{-1}\,((\tau_{-h\,e_j}(\lambda))(\phi) - \lambda(\phi)) = \lambda(h^{-1}(\tau_{h\,e_j}(\phi) - \phi)), \tag{21.11}$$

and

$$\operatorname{supp}\, h^{-1}\,(\tau_{h\,e_j}(\phi) - \phi) \subseteq \operatorname{supp} K(r). \tag{21.12}$$

Of course,

$$h^{-1}\,((\tau_{h\,e_j}(\phi))(x) - \phi(x)) = h^{-1}\,(\phi(x - h\,e_j) - \phi(x)) \tag{21.13}$$

tends to $-(D_j\phi)(x)$ as $h \to 0$ for each $x \in U_r$. More precisely, (21.13) converges to $-(D_j\phi)(x)$ as $h \to 0$ uniformly on $U_r$, because $\phi$ is continuously differentiable and has compact support. Similarly, the derivatives of (21.13) converge to the corresponding derivatives of $-(D_j\phi)(x)$ as $h \to 0$ uniformly on $U_r$, because $\phi$ is smooth. This and (21.12) imply that (21.13) converges to $-(D_j\phi)(x)$ as $h \to 0$ in the topology of $C^\infty_{com}(U_r)$.

It follows that

$$\lim_{h \to 0} \mu_{j,h}(\phi) = -\lambda(D_j\phi) = (D_j\lambda)(\phi) \tag{21.14}$$

when $\phi$ is a smooth function on $\mathbf{R}^n$ with support contained in a compact subset $K$ of $U_r$ for some $r > 0$. If $\phi$ is any smooth function on $U$ with support contained in a compact subset $K$ of $U$, then $\phi$ can be extended to a smooth function on $\mathbf{R}^n$ with compact support contained in $K$ as before, and $K \subseteq U_r$ when $r > 0$ is sufficiently small. Thus (21.14) holds for every $\phi \in C^\infty_{com}(U)$, with the understanding that $\mu_{j,h}(\phi)$ is defined when $|h|$ is sufficiently small, depending on $\phi$. This is all a bit simpler when $U = \mathbf{R}^n$, so that $U + a = \mathbf{R}^n$ for each $a \in \mathbf{R}^n$, and $U_r = \mathbf{R}^n$ for every $r > 0$. In this case, $\mu_{j,h}(\phi)$ is defined for every $\phi \in C^\infty_{com}(\mathbf{R}^n)$ and $h \neq 0$, and (21.14) implies that $\mu_{j,h} \to D_j\lambda$ as $h \to 0$ with respect to the weak* topology on the dual of $C^\infty_{com}(\mathbf{R}^n)$. Note that there are analogous statements for tempered distributions on $\mathbf{R}^n$. This uses the fact that (21.13) converges to $-(D_j\phi)(x)$ as $h \to 0$ with respect to the usual topology on the Schwartz class when $\phi \in \mathcal{S}(\mathbf{R}^n)$.

Suppose for instance that $n = 1$, $U = \mathbf{R}$, and $\lambda = \lambda_f$ as in (7.2) for some locally integrable function $f$ on $\mathbf{R}$. Put

$$\mu_h = h^{-1}\,(\tau_{-h}(\lambda) - \lambda) \tag{21.15}$$

for each nonzero real number $h$, as in (21.10). Thus

$$\mu_h(\phi) = \int_{\mathbf{R}} \phi(x)\,\Big(\frac{f(x+h) - f(x)}{h}\Big)\,dx \tag{21.16}$$

for every $\phi \in C^\infty_{com}(\mathbf{R})$, by (21.4). If $f$ is monotone increasing on $\mathbf{R}$, then it follows that $\mu_h$ is a nonnegative distribution on $\mathbf{R}$ for each $h \neq 0$. This implies that $D\lambda_f$ is a nonnegative distribution on $\mathbf{R}$ as well, since $\mu_h \to D\lambda_f$ as $h \to 0$ with respect to the weak* topology on the dual of $C^\infty_{com}(\mathbf{R}^n)$, as before.

# 22 Lebesgue–Stieltjes measures

Let $\alpha(x)$ be a monotone increasing real-valued function on the real line. It is well known that there is a unique nonnegative Borel measure $\mu_\alpha$ on the real line such that

$$\mu_\alpha((a, b)) = \alpha(b-) - \alpha(a+) \tag{22.1}$$

for every $a, b \in \mathbf{R}$ with $a < b$. Equivalently,

$$\mu_\alpha([a, b]) = \alpha(b+) - \alpha(a-) \tag{22.2}$$

for every $a, b \in \mathbf{R}$ with $a \leq b$, because every closed interval in the real line can be expressed as the intersection of a decreasing sequence of open intervals, and every open interval can be expressed as the union of an increasing sequence of closed intervals. Similarly, $\mu_\alpha$ can be characterized by the property that

$$\mu_\alpha((a, b]) = \alpha(b+) - \alpha(a+) \tag{22.3}$$

for every $a, b \in \mathbf{R}$ with $a, b$, or the property that

$$\mu_\alpha([a, b)) = \mu(b-) - \mu(a-) \tag{22.4}$$

for every $a, b \in \mathbf{R}$ with $a < b$. In particular, if $\alpha(x) = x$ for every $x \in \mathbf{R}$, then $\mu_\alpha$ is the same as ordinary Lebesgue measure on $\mathbf{R}$. Note that

$$\mu_\alpha(\{a\}) = \alpha(a+) - \alpha(a-) \tag{22.5}$$

for each $a \in \mathbf{R}$, by applying (22.2) with $a = b$. Thus discontinuities of $\alpha$ correspond exactly to points $a \in \mathbf{R}$ such that $\mu_\alpha > 0$.

One way to get $\mu_\alpha$ for an arbitrary monotone increasing function $\alpha$ on $\mathbf{R}$ is to start with the corresponding Riemann–Stieltjes integral

$$\int_{\mathbf{R}} \phi \, d\alpha \tag{22.6}$$

of a continuous real-valued function $\phi$ with compact support on $\mathbf{R}$, as in Section 20. This defines a nonnegative linear functional on $C_{com}(\mathbf{R})$, and the Riesz representation theorem implies that there is a unique nonnegative Borel measure $\mu_\alpha$ on $\mathbf{R}$ such that (22.6) is equal to the Lebesgue integral

$$\int_{\mathbf{R}} \phi \, d\mu_\alpha \tag{22.7}$$

of $\phi$ with respect to $\mu_\alpha$ for each $\phi \in C_{com}(\mathbf{R})$. One check that $\mu_\alpha$ satisfies the conditions in the previous paragraph, by approximating indicator functions of intervals by nonnegative continuous functions on $\mathbf{R}$. Conversely, if $\mu_\alpha$ is a nonnegative Borel measure on $\mathbf{R}$ that satisfies the conditions described in the preceding paragraph, then it is easy to see that the Lebesgue integral (22.7) of $\phi \in C_{com}(\mathbf{R})$ is the same as the Riemann–Stieltjes integral (22.6). Alternatively, one can start by defining $\mu_\alpha$ on finite unions of intervals in $\mathbf{R}$ as in the previous paragraph, and then show that $\mu_\alpha$ can be extended to a countably-additive Borel measure on $\mathbf{R}$. As another approach, if $\alpha$ is continuous and strictly increasing, then $\alpha$ is a homeomorphism onto $\alpha(\mathbf{R})$, and $\alpha(\mathbf{R})$ is either an open interval, an open half-line, or the whole real line. In this case, one can simply take $\mu_\alpha(E)$ to be the Lebesgue measure of $\alpha(E)$, and otherwise there are some related constructions of $\mu_\alpha$.

Now let $\mu$ be a nonnegative Borel measure on $\mathbf{R}$ such that $\mu(E) < \infty$ for every bounded Borel set $E \subseteq \mathbf{R}$. It is well known that $\mu = \mu_\alpha$ for some monotone increasing real-valued function $\alpha$ on $\mathbf{R}$. This is simpler when $\mu(\mathbf{R}) < +\infty$, or at least $\mu((-\infty, 0)) < +\infty$, in which case one can take

$$\alpha(x) = \mu((-\infty, x)) \tag{22.8}$$

for each $x \in \mathbf{R}$. One could also take

$$\alpha(x) = \mu((-\infty, x]) \tag{22.9}$$

for each $x \in \mathbf{R}$, which is the same as (22.8) when $\mu(\{x\}) = 0$, and which has the same one-sided limits at every point in $\mathbf{R}$. Note that (22.8) is left continuous on $\mathbf{R}$, which means that $\alpha(x-) = \alpha(x)$ for every $x \in \mathbf{R}$, while (22.9) is right continuous on $\mathbf{R}$, in the sense that $\alpha(x+) = \alpha(x)$ for every $x \in \mathbf{R}$. Both (22.8) and (22.9) have the property that $\alpha(x) \to 0$ as $x \to -\infty$. There are analogous but more complicated choices of $\alpha$ based at any point in $\mathbf{R}$ instead of $-\infty$, which do not require any additional finiteness assumptions on $\mu$.

If $\alpha_1(x)$ and $\alpha_2(x)$ are monotone increasing real-valued functions on $\mathbf{R}$ with the same one-sided limits at each point, so that $\alpha_1(x+) = \alpha_2(x+)$ and $\alpha_1(x-) = \alpha_2(x-)$ for every $x \in \mathbf{R}$, then the corresponding Borel measures $\mu_{\alpha_1}$ and $\mu_{\alpha_2}$ are the same. In this case, $\alpha_1$ and $\alpha_2$ have the same set of discontinuities, and $\alpha_1(x) = \alpha_2(x)$ when $\alpha_1$ or equivalently $\alpha_2$ is continuous at $x$. In particular, $\alpha_1(x) = \alpha_2(x)$ for all but finitely or countably many $x \in \mathbf{R}$. If $\alpha_1$ and $\alpha_2$ are both right continuous at every $x \in \mathbf{R}$, or if they are both left continuous at every $x \in \mathbf{R}$, then it follows that $\alpha_1(x) = \alpha_2(x)$ for every $x \in \mathbf{R}$. Similarly, if $\alpha_3(x)$ is a monotone increasing real-valued function on $\mathbf{R}$ such that $\alpha_3(x) - \alpha_1(x)$ is constant, then $\mu_{\alpha_1} = \mu_{\alpha_3}$.

# 23 Differentiation of monotone functions

Let $\alpha(x)$ be a monotone increasing real-valued function on the real line. It is well known that the classical derivative

$$\alpha'(x) = \lim_{h \to 0} \frac{\alpha(x+h) - \alpha(x)}{h} \tag{23.1}$$

exists for almost every $x \in \mathbf{R}$ with respect to Lebesgue measure, in which case we obviously have that $\alpha'(x) \ge 0$. Note that

$$\alpha'(x) = \lim_{j \to \infty} j\,(\alpha(x + 1/j) - \alpha(x)) \tag{23.2}$$

for every $x \in \mathbf{R}$ for which (23.1) exists. In particular, this implies that $\alpha'(x)$ is Lebesgue measurable on $\mathbf{R}$. If $a, b \in \mathbf{R}$ and $a < b$, then we can apply Fatou's lemma to get that

$$\int_a^b \alpha'(x)\,dx \le \liminf_{j \to \infty} \int_a^b j\,(\alpha(x + 1/j) - \alpha(x))\,dx. \tag{23.3}$$

Of course,

$$\begin{aligned} \int_a^b j\,(\alpha(x+1/j) - \alpha(x))\,dx &= j \int_a^b \alpha(x + 1/j)\,dx - j \int_a^b \alpha(x)\,dx \\ &= j \int_{a+1/j}^{b+1/j} \alpha(x)\,dx - j \int_a^b \alpha(x)\,dx \\ &= j \int_b^{b+1/j} \alpha(x)\,dx - j \int_a^{a+1/j} \alpha(x)\,dx \end{aligned} \tag{23.4}$$

for each $j$. It is easy to see that

$$\lim_{j \to \infty} \int_y^{y+1/j} \alpha(x)\,dx = \alpha(y+) \tag{23.5}$$

for each $y \in \mathbf{R}$, so that

$$\int_a^b \alpha'(x)\,dx \le \alpha(b+) - \alpha(a+) \tag{23.6}$$

for every $a, b \in \mathbf{R}$ with $a < b$. In fact, we have that

$$\int_a^b \alpha'(x)\,dx \le \alpha(b-) - \alpha(a+) \tag{23.7}$$

for every $a, b \in \mathbf{R}$ with $a < b$, since one can apply (23.6) with $b$ replaced by any element of $(a, b)$, and then pass to the limit to get (23.7).

Now let $f$ be a locally integrable function on $\mathbf{R}$, and put

$$F(y) = \int_0^y f(x)\,dx \tag{23.8}$$

for each $y \in \mathbf{R}$. More precisely, this is intended as an oriented integral, so that

$$F(y) = \int_{[0,y]} f(x)\,dx \tag{23.9}$$

when $x \ge 0$, and

$$F(y) = -\int_{[y,0]} f(x)\,dx \tag{23.10}$$

when $y \le 0$, and at any rate the main point is that

$$F(b) - F(a) = \int_a^b f(x)\,dx \tag{23.11}$$

for every $a, b \in \mathbf{R}$ with $a < b$. Lebesgue's differentiation theorem implies that $F'(y)$ exists and is equal to $f(y)$ for almost every $y \in \mathbf{R}$. Note that $F(y)$ is monotone increasing on $\mathbf{R}$ when $f$ is a nonnegative real-valued function on $\mathbf{R}$.

In particular, if $\alpha(x)$ is a monotone increasing function on $\mathbf{R}$, then we can apply this to $f(x) = \alpha'(x)$, to get a monotone increasing function $F$ on $\mathbf{R}$ such that $F'(y) = \alpha'(y)$ for almost every $y \in \mathbf{R}$. Let us check that $\beta(y) = \alpha(y) - F(y)$ is also monotone increasing on $\mathbf{R}$. If $a, b \in \mathbf{R}$ and $a < b$, then

$$\begin{aligned} \beta(b) - \beta(a) &= (\alpha(b) - \alpha(a)) - (F(b) - F(a)) \\ &= (\alpha(b) - \alpha(a)) - \int_a^b f(x)\,dx \ge 0, \end{aligned} \tag{23.12}$$

by (23.7) and (20.3). Thus $\beta(y)$ is monotone increasing and $\beta'(y) = 0$ for almost every $y \in \mathbf{R}$, although $\beta$ is not constant on $\mathbf{R}$ when $\alpha$ has jump discontinuities. There are also well-known examples of continuous monotone increasing functions on $\mathbf{R}$ that are not constant but have derivative equal to 0 almost everywhere, such as the Cantor–Lebesgue function.

## References


[1] L. Baggett, *Functional Analysis: A Primer*, Dekker, 1992.

[2] R. Bass, *Real Analysis for Graduate Students: Measure and Integration Theory*, 2011. http://homepages.uconn.edu/˜rib2005/real.html

[3] R. Beals, *Advanced Mathematical Analysis*, Springer-Verlag, 1973.

[4] R. Beals, *Analysis: An Introduction*, Cambridge University Press, 2004.

[5] J. Benedetto and W. Czaja, *Integration and Modern Analysis*, Birkhäuser, 2009.

[6] M. Carter and B. van Brunt, *The Lebesgue–Stieltjes Integral*, Springer-Verlag, 2000.

[7] J. Cerdà, *Linear Functional Analysis*, American Mathematical Society, Real Sociedad Matemática Española, 2010.

[8] J. Conway, *A Course in Functional Analysis*, 2nd edition, Springer-Verlag, 1990.

[9] J. Duistermaat and J. Kolk, *Distributions: Theory and Applications*, translated from the Dutch by J. van Braam Houckgeest, Birkhäuser, 2010.

[10] R. Edwards, *Functional Analysis*, Dover, 1995.

[11] J. Ferreira, *Introduction to the Theory of Distributions*, translated from the 1993 Portuguese original by J. Sousa Pinto and R. Hoskins, Longman, 1997.

[12] G. Folland, *Real Analysis: Modern Techniques and their Applications*, 2nd edition, Wiley, 1999.

[13] G. Folland, *A Guide to Advanced Real Analysis*, Mathematical Association of America, 2009.

[14] F. Friedlander, *Introduction to the Theory of Distributions*, 2nd edition, with additional material by M. Joshi, Cambridge Universoty Press, 1998.

[15] R. Goldberg, *Methods of Real Analysis*, 2nd edition, Wiley, 1976.

[16] G. Grubb, *Distributions and Operators*, Springer-Verlag, 2009.

[17] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, 1975.

[18] F. Hirsch and G. Lacombe, *Elements of Functional Analysis*, translated from the 1997 French original by S. Levy, Springer-Verlag, 1999.

[19] J. Horváth, *Topological Vector Spaces and Distributions*, Addison-Wesley, 1966.

[20] J. Horváth, *An introduction to distributions*, American Mathematical Monthly **77** (1970), 227–240.

[21] F. Jones, *Lebesgue Integration on Euclidean Space*, Jones and Bartlett, 1993.

[22] S. Kantorovitz, *Introduction to Modern Analysis*, Oxford University Press, 2003.

[23] A. Knapp, *Basic Real Analysis*, Birkhäuser, 2005.

[24] A. Knapp, *Advanced Real Analysis*, Birkhäuser, 2005.

[25] S. Krantz, *Real Analysis and Foundations*, 2nd edition, Chapman & Hall / CRC, 2005.

[26] S. Krantz, *A Guide to Real Variables*, Mathematical Association of America, 2009.

[27] S. Lang, *Real and Functional Analysis*, 3rd edition, Springer-Verlag, 1993.

[28] S. Lang, *Undergraduate Analysis*, 2nd edition, Springer-Verlag, 1997.

[29] P. Lax, *Functional Analysis*, Wiley, 2002.

[30] I. Maddox, *Elements of Functional Analysis*, 2nd edition, Cambridge University Press, 1988.

[31] L. Narici and E. Beckenstein, *Topological Vector Spaces*, 2nd edition, CRC Press, 2011.

[32] I. Richards and H. Youn, *Theory of Distributions: A Nontechnical Introduction*, Cambridge University Press, 1990.

[33] A. Robertson and W. Robertson, *Topological Vector Spaces*, 2nd edition, Cambridge University Press, 1980.

[34] H. Royden, *Real Analysis*, 3rd edition, Macmillan, 1988.

[35] W. Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, 1976.

[36] W. Rudin, *Real and Complex Analysis*, 3rd edition, McGraw-Hill, 1987.

[37] W. Rudin, *Functional Analysis*, 2nd edition, McGraw-Hill, 1991.

[38] L. Schwartz, *Théorie des Distributions*, Hermann, 1966.

[39] E. Stein and R. Shakarchi, *Real Analysis*, Princeton University Press, 2005.

[40] E. Stein and R. Shakarchi, *Functional Analysis*, Princeton University Press, 2011.

[41] R. Strichartz, *The Way of Analysis*, Jones and Bartlett, 1995.

[42] R. Strichartz, *A Guide to Distribution Theory and Fourier Transforms*, World Scientific, 2003.

[43] K. Stromberg, *Introduction to Classical Real Analysis*, Wadsworth, 1981.

[44] D. Stroock, *Essentials of Integration Theory for Analysis*, Springer-Verlag, 2011.

[45] F. Trèves, *Topological Vector Spaces, Distributions, and Kernels*, Dover, 2006.

[46] A. Wilansky, *Modern Methods in Topological Vector Spaces*, McGraw-Hill, 1978.

[47] A. Zemanian, *Distribution Theory and Transform Analysis*, 2nd edition, Dover, 1987.

[48] R. Zimmer, *Essential Results of Functional Analysis*, University of Chicago Press, 1990.