1

PRELIMINARIES

1.1 (Introduction) Sobolev spaces are vector spaces whose elements are functions defined on domains in n-dimensional Euclidean space $\mathbb{R}^n$ and whose partial derivatives satisfy certain integrability conditions. In order to develop and elucidate the properties of these spaces and mappings between them we require some of the machinery of general topology and real and functional analysis. We assume that readers are familiar with the concept of a vector space over the real or complex scalar field, and with the related notions of dimension, subspace, linear transformation, and convex set. We also expect the reader will have some familiarity with the concept of topology on a set, at least to the extent of understanding the concepts of an open set and continuity of a function.

In this chapter we outline, mainly without any proofs, those aspects of the theories of topological vector spaces, continuity, the Lebesgue measure and integral, and Schwartz distributions that will be needed in the rest of the book. For a reader familiar with the basics of these subjects, a superficial reading to settle notations and review the main results will likely suffice.

Notation

1.2 Throughout this monograph the term *domain* and the symbol Ω will be reserved for a nonempty open set in n-dimensional real Euclidean space $\mathbb{R}^n$. We shall be concerned with the differentiability and integrability of functions defined on Ω; these functions are allowed to be complex-valued unless the contrary is

explicitly stated. The complex field is denoted by $\mathbb{C}$. For $c \in \mathbb{C}$ and two functions u and v, the scalar multiple cu, the sum $u + v$, and the product uv are always defined pointwise:

$$(cu)(x) = cu(x),$$

$$(u + v)(x) = u(x) + v(x),$$

$$(uv)(x) = u(x)v(x)$$

at all points x where the right sides make sense.

A typical point in $\mathbb{R}^n$ is denoted by $x = (x_1, \ldots, x_n)$; its norm is given by $|x| = (\sum_{j=1}^{n} x_j^2)^{1/2}$. The inner product of two points x and y in $\mathbb{R}^n$ is $x \cdot y = \sum_{j=1}^{n} x_j y_j$.

If $\alpha = (\alpha_1, \ldots, \alpha_n)$ is an n-tuple of nonnegative integers α_j, we call α a *multi-index* and denote by x^α the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, which has degree $|\alpha| = \sum_{j=1}^{n} \alpha_j$. Similarly, if $D_j = \partial/\partial x_j$, then

$$D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n}$$

denotes a differential operator of order $|\alpha|$. Note that $D^{(0,\ldots,0)}u = u$.

If α and β are two multi-indices, we say that $\beta \leq \alpha$ provided $\beta_j \leq \alpha_j$ for $1 \leq j \leq n$. In this case $\alpha - \beta$ is also a multi-index, and $|\alpha - \beta| + |\beta| = |\alpha|$. We also denote

$$\alpha! = \alpha_1! \cdots \alpha_n!$$

and if $\beta \leq \alpha$,

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha - \beta)!} = \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n}.$$

The reader may wish to verify the Leibniz formula

$$D^\alpha(uv)(x) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta u(x) D^{\alpha - \beta} v(x)$$

valid for functions u and v that are $|\alpha|$ times continuously differentiable near x.

1.3   If $G \subset \mathbb{R}^n$ is nonempty, we denote by $\overline{G}$ the closure of G in $\mathbb{R}^n$. We shall write $G \Subset \Omega$ if $\overline{G} \subset \Omega$ and $\overline{G}$ is a compact (that is, closed and bounded) subset of $\mathbb{R}^n$. If u is a function defined on G, we define the *support* of u to be the set

$$\text{supp}\,(u) = \overline{\{x \in G : u(x) \neq 0\}}.$$

We say that u has *compact support* in Ω if $\text{supp}\,(u) \Subset \Omega$. We denote by "bdry G" the boundary of G in $\mathbb{R}^n$, that is, the set $\overline{G} \cap \overline{G^c}$, where G^c is the complement of G in $\mathbb{R}^n$; $G^c = \mathbb{R}^n - G = \{x \in \mathbb{R}^n : x \notin G\}$.

If $x \in \mathbb{R}^n$ and $G \subset \mathbb{R}^n$, we denote by "dist(x, G)" the distance from x to G, that is, the number $\inf_{y \in G} |x - y|$. Similarly, if F, $G \subset \mathbb{R}^n$ are both nonempty,

$$\text{dist}(F, G) = \inf_{y \in F} \text{dist}(y, G) = \inf_{\substack{x \in G \\ y \in F}} |y - x|.$$

Topological Vector Spaces

1.4 (Topological Spaces)   If X is any set, a *topology* on X is a collection $\mathcal{O}$ of subsets of X which contains

(i) the whole set X and the empty set $\emptyset$,

(ii) the union of any collection of its elements, and

(iii) the intersection of any finite collection of its elements.

The pair $(X, \mathcal{O})$ is called a *topological space* and the elements of $\mathcal{O}$ are the *open sets* of that space. An open set containing a point x in X is called a *neighbourhood* of x. The complement $X - U = \{x \in X : x \notin U\}$ of any open set U is called a *closed* set. The closure $\overline{S}$ of any subset $S \subset X$ is the smallest closed subset of X that contains S.

Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be two topologies on the same set X. If $\mathcal{O}_1 \subset \mathcal{O}_2$, we say that $\mathcal{O}_2$ is *stronger* than $\mathcal{O}_1$, or that $\mathcal{O}_1$ is *weaker* than $\mathcal{O}_2$.

A topological space $(X, \mathcal{O})$ is called a *Hausdorff space* if every pair of distinct points x and y in X have disjoint neighbourhoods.

The *topological product* of two topological spaces $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ is the topological space $(X \times Y, \mathcal{O})$, where $X \times Y = \{(x, y) : x \in X, y \in Y\}$ is the Cartesian product of the sets X and Y, and $\mathcal{O}$ consists of arbitrary unions of sets of the form $\{O_X \times O_Y : O_X \in \mathcal{O}_X, O_Y \in \mathcal{O}_Y\}$.

Let $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ be two topological spaces. A function f from X into Y is said to be *continuous* if the preimage $f^{-1}(O) = \{x \in X : f(x) \in O\}$ belongs to $\mathcal{O}_X$ for every $O \in \mathcal{O}_Y$. Evidently the stronger the topology on X or the weaker the topology on Y, the more such continuous functions f there will be.

1.5 (Topological Vector Spaces)   We assume throughout this monograph that all vectors spaces referred to are taken over the complex field unless the contrary is explicitly stated.

A *topological vector space*, hereafter abbreviated TVS, is a Hausdorff topological space that is also a vector space for which the vector space operations of addition and scalar multiplication are continuous. That is, if X is a TVS, then the mappings

$$(x, y) \to x + y \qquad \text{and} \qquad (c, x) \to cx$$

from the topological product spaces $X \times X$ and $\mathbb{C} \times X$, respectively, into X are continuous. (Here $\mathbb{C}$ has its usual topology induced by the Euclidean metric.)

X is a *locally convex* TVS if each neighbourhood of the origin in X contains a convex neighbourhood of the origin.

We outline below those aspects of the theory of topological and normed vector spaces that play a significant role in the study of Sobolev spaces. For a more thorough discussion of these topics the reader is referred to standard textbooks on functional analysis, for example [Ru1] or [Y].

1.6 (Functionals) A scalar-valued function defined on a vector space X is called a *functional*. The functional f is linear provided

$$f(ax + by) = af(x) + bf(y), \qquad x, y \in X, \quad a, b \in \mathbb{C}.$$

If X is a TVS, a functional on X is continuous if it is continuous from X into $\mathbb{C}$ where $\mathbb{C}$ has its usual topology induced by the Euclidean metric.

The set of all continuous, linear functionals on a TVS X is called the *dual* of X and is denoted by X'. Under pointwise addition and scalar multiplication X' is itself a vector space:

$$(f + g)(x) = f(x) + g(x), \quad (cf)(x) = cf(x), \qquad f, g \in X', \ x \in X, \ c \in \mathbb{C}.$$

X' will be a TVS provided a suitable topology is specified for it. One such topology is the *weak-star topology*, the weakest topology with respect to which the functional F_x, defined on X' by $F_x(f) = f(x)$ for each $f \in X'$, is continuous for each $x \in X$. This topology is used, for instance, in the space of Schwartz distributions introduced in Paragraph 1.57. The dual of a normed vector space can be given a stronger topology with respect to which it is itself a normed space. (See Paragraph 1.11.)

Normed Spaces

1.7 (Norms) A *norm* on a vector space X is a real-valued function f on X satisfying the following conditions:

   (i) $f(x) \geq 0$ for all $x \in X$ and $f(x) = 0$ if and only if $x = 0$,

   (ii) $f(cx) = |c| f(x)$ for every $x \in X$ and $c \in \mathbb{C}$,

   (iii) $f(x + y) \leq f(x) + f(y)$ for every $x, y \in X$.

A *normed space* is a vector space X provided with a norm. The norm will be denoted $\|\cdot\, ; X\|$ except where other notations are introduced.

If $r > 0$, the set

$$B_r(x) = \{y \in X \ : \ \|y - x \, ; X\| < r\}$$

is called the *open ball* of radius r with center at $x \in X$. Any subset $A \subset X$ is called *open* if for every $x \in A$ there exists $r > 0$ such that $B_r(x) \subset A$. The open sets thus defined constitute a topology for X with respect to which X is a TVS. This topology is the *norm topology* on X. The closure of $B_r(x)$ in this topology is

$$\overline{B_r(x)} = \{y \in X \ : \ \|y - x \,;\, X\| \leq r\}.$$

A TVS X is *normable* if its topology coincides with the topology induced by some norm on X. Two different norms on a vector space X are equivalent if they induce the same topology on X. This is the case if and only if there exist two positive constants a and b such that,

$$a \, \|x\|_1 \leq \|x\|_2 \leq b \, \|x\|_1$$

for all $x \in X$, where $\|x\|_1$ and $\|x\|_2$ are the two norms.

Let X and Y be two normed spaces. If there exists a one-to-one linear operator L mapping X onto Y having the property $\|L(x) \,;\, Y\| = \|x \,;\, X\|$ for every $x \in X$, then we call L an *isometric isomorphism* between X and Y, and we say that X and Y are *isometrically isomorphic*. Such spaces are often identified since they have identical structures and only differ in the nature of their elements.

1.8 A sequence $\{x_n\}$ in a normed space X is *convergent* to the limit x_0 if and only if $\lim_{n \to \infty} \|x_n - x_0 \,;\, X\| = 0$ in $\mathbb{R}$. The norm topology of X is completely determined by the sequences it renders convergent.

A subset S of a normed space X is said to be *dense* in X if each $x \in X$ is the limit of a sequence of elements of S. The normed space X is called *separable* if it has a countable dense subset.

1.9 **(Banach Spaces)** A sequence $\{x_n\}$ in a normed space X is called a *Cauchy sequence* if and only if for every $\epsilon > 0$ there exists an integer N such that $\|x_m - x_n \,;\, X\| < \epsilon$ holds whenever $m, n > N$. We say that X is *complete* and a *Banach space* if every Cauchy sequence in X converges to a limit in X. Every normed space X is either a Banach space or a dense subset of a Banach space Y called the *completion* of X whose norm satisfies

$$\|x \,;\, Y\| = \|x \,;\, X\| \qquad \text{for every } x \in X.$$

1.10 **(Inner Product Spaces and Hilbert Spaces)** If X is a vector space, a functional $(\cdot, \cdot)_X$ defined on $X \times X$ is called an *inner product* on X provided that for every $x, y \in X$ and $a, b \in \mathbb{C}$

   (i) $(x, y)_X = \overline{(y, x)}_X$, (where $\overline{c}$ denotes the complex conjugate of $c \in \mathbb{C}$)

   (ii) $(ax + by, z)_X = a(x, z)_X + b(y, z)_X$,

(iii) $(x, x)_X = 0$ if and only if $x = 0$,

Equipped with such a functional, X is called an *inner product space,* and the functional

$$\|x \, ; X\| = \sqrt{(x, x)_X} \tag{1}$$

is, in fact, a norm on X If X is complete (i.e. a Banach space) under this norm, it is called a *Hilbert space.* Whenever the norm on a vector space X is obtained from an inner product via (1), it satisfies the *parallelogram law*

$$\|x + y \, ; X\|^2 + \|x - y \, ; X\|^2 = 2 \, \|x \, ; X\|^2 + 2 \, \|y \, ; X\|^2. \tag{2}$$

Conversely, if the norm on X satisfies (2) then it comes from an inner product as in (1).

1.11 (The Normed Dual) A norm on the dual X' of a normed space X can be defined by setting

$$\|x' \, ; X'\| = \sup\{|x'(x)| \, : \, \|x \, ; X\| \le 1\},$$

for each $x' \in X'$. Since $\mathbb{C}$ is complete, with the topology induced by this norm X' is a Banach space (whether or not X is) and it is called the *normed dual* of X. If X is infinite dimensional, the norm topology of X' is stronger (has more open sets) than the weak-star topology defined in Paragraph 1.6.

The following theorem shows that if X is a Hilbert space, it can be identified with its normed dual.

1.12 THEOREM (The Riesz Representation Theorem) Let X be a Hilbert space. A linear functional x' on X belongs to X' if and only if there exists $x \in X$ such that for every $y \in X$ we have

$$x'(y) = (y, x)_X,$$

and in this case $\|x' \, ; X'\| = \|x \, ; X\|$. Moreover, x is uniquely determined by $x' \in X'$. ∎

A vector subspace M of a normed space X is itself a normed space under the norm of X, and so normed is called a *subspace* of X. A closed subspace of a Banach space is itself a Banach space.

1.13 THEOREM (The Hahn-Banach Extension Theorem) Let M be a subspace of the normed space X. If $m' \in M'$, then there exists $x' \in X'$ such that $\|x' \, ; X'\| = \|m' \, ; M'\|$ and $x'(m) = m'(m)$ for every $m \in M$. ∎

1.14 (Reflexive Spaces) A natural linear injection of a normed space X into its second dual space $X'' = (X')'$ is provided by the mapping J whose value Jx at $x \in X$ is given by

$$Jx(x') = x'(x), \qquad x' \in X'.$$

Since $|Jx(x')| \le \|x' \,;\, X'\| \, \|x \,;\, X\|$, we have

$$\|Jx \,;\, X''\| \le \|x \,;\, X\|.$$

However, the Hahn-Banach Extension Theorem assures us that for any $x \in X$ we can find $x' \in X'$ such that $\|x' \,;\, X'\| = 1$ and $x'(x) = \|x \,;\, X\|$. Therefore J is an isometric isomorphism of X into X''.

If the range of the isomorphism J is the entire space X'', we say that the normed space X is *reflexive*. A reflexive space must be complete, and hence a Banach space.

1.15 THEOREM Let X be a normed space. X is reflexive if and only if X' is reflexive. X is separable if X' is separable. Hence if X is separable and reflexive, so is X'. ∎

1.16 (Weak Topologies and Weak Convergence) The *weak topology* on a normed space X is the weakest topology on X that still renders continuous each x' in the normed dual X' of X. Unless X is finite dimensional, the weak topology is weaker than the norm topology on X. It is a consequence of the Hahn-Banach Theorem that a closed, convex set in a normed space is also closed in the weak topology of that space.

A sequence convergent with respect to the weak topology on X is said to *converge weakly*. Thus x_n converges weakly to x in X provided $x'(x_n) \to x'(x)$ in $\mathbb{C}$ for every $x' \in X'$. We denote norm convergence of a sequence $\{x_n\}$ to x in X by $x_n \to x$, and we denote weak convergence by $x_n \rightharpoonup x$. Since we have $|x'(x_n - x)| \le \|x' \,;\, X'\| \, \|x_n - x \,;\, X\|$, we see that $x_n \to x$ implies $x_n \rightharpoonup x$. The converse is generally not true (unless X is finite dimensional).

1.17 (Compact Sets) A subset A of a normed space X is called *compact* if every sequence of points in A has a subsequence converging in X to an element of A. (This definition is equivalent in normed spaces to the definition of compactness in a general topological space; A is compact if whenever A is a subset of the union of a collection of open sets, it is a subset of the union of a finite subcollection of those sets.) Compact sets are closed and bounded, but closed and bounded sets need not be compact unless X is finite dimensional. A is called *precompact* in X if its closure $\overline{A}$ in the norm topology of X is compact. A is called *weakly sequentially compact* if every sequence in A has a subsequence converging weakly in X to a point in A. The reflexivity of a Banach space can be characterized in terms of this property.

1.18 THEOREM A Banach space is reflexive if and only if its closed unit ball $\overline{B_1(0)} = \{x \in X \,:\, \|x \,;\, X\| \le 1\}$ is weakly sequentially compact. ∎

1.19   THEOREM   A set A is precompact in a Banach space X if and only if for every positive number ϵ there is a finite subset N_ϵ of points of X such that

$$A \subset \bigcup_{y \in N_\epsilon} B_\epsilon(y).$$

A set N_ϵ with this property is called a *finite ϵ-net* for A. ∎

1.20   (Uniform Convexity)   Any normed space is locally convex with respect to its norm topology. The norm on X is called *uniformly convex* if for every number ϵ satisfying $0 < \epsilon \leq 2$, there exists a number $\delta(\epsilon) > 0$ such that if $x, y \in X$ satisfy $\|x\,;X\| = \|y\,;X\| = 1$ and $\|x - y\,;X\| \geq \epsilon$, then $\|(x + y)/2\,;X\| \leq 1 - \delta(\epsilon)$. The normed space X itself is called "uniformly convex" in this case. It should be noted, however, that uniform convexity is a property of the norm—X may have another equivalent norm that is not uniformly convex. Any normable space is called *uniformly convex* if it possesses a uniformly convex norm. The parallelogram law (2) shows that a Hilbert space is uniformly convex.

1.21   THEOREM   A uniformly convex Banach space is reflexive. ∎

The following two theorems will be used to establish the separability, reflexivity, and uniform convexity of the Sobolev spaces introduced in Chapter 3.

1.22   THEOREM   Let X be a Banach space and M a subspace of X closed with respect to the norm topology of X. Then M is also a Banach space under the norm inherited from X. Furthermore

   (i) M is separable if X is separable,

   (ii) M is reflexive if X is reflexive,

   (iii) M is uniformly convex if X is uniformly convex. ∎

The completeness, separability, and uniform convexity of M follow easily from the corresponding properties of X. The reflexivity of M is a consequence of Theorem 1.18 and the fact that M, being closed and convex, is closed in the weak topology of X.

1.23   THEOREM   For $j = 1, 2, \ldots, n$ let X_j be a Banach space with norm $\|\cdot\|_j$. The Cartesian product $X = \prod_{j=1}^{n} X_j$, consisting of points $(x_1, \ldots, x_n)$ with $x_j \in X_j$, is a vector space under the definitions

$$x + y = (x_1 + y_1, \ldots, x_n + y_n), \qquad cx = (cx_1, \ldots, cx_n),$$

and is a Banach space with respect to any of the equivalent norms

$$\|x\|_{(p)} = \left(\sum_{j=1}^{n} \|x_j\|_j^p \right)^{1/p}, \qquad 1 \leq p < \infty,$$

$$\|x\|_{(\infty)} = \max_{1 \leq j \leq n} \|x_j\|_j.$$

Furthermore,

    (i) if X_j is separable for $1 \leq j \leq n$, then X is separable,

    (ii) if X_j is reflexive for $1 \leq j \leq n$, then X is reflexive,

    (iii) if X_j is uniformly convex for $1 \leq j \leq n$, then X is uniformly convex. More precisely, $\|\cdot\|_{(p)}$ is a uniformly convex norm on X provided $1 < p < \infty$. ∎

The functionals $\|\cdot\|_{(p)}$, $1 \leq p \leq \infty$, are norms on X, and X is complete with respect to each of them. Equivalence of these norms follows from the inequalities

$$\|x\|_{(\infty)} \leq \|x\|_{(p)} \leq \|x\|_{(1)} \leq n \, \|x\|_{(\infty)} \, .$$

The separability and uniform convexity of X are readily deduced from the corresponding properties of the spaces X_j. The reflexivity of X follows from that of X_j, $1 \leq j \leq n$, via Theorem 1.18 or via the natural isomorphism between X' and $\prod_{j=1}^{n} X_j'$.

1.24   **(Operators)**   Since the topology of a normed space X is determined by the sequences it renders convergent, an operator f defined on X into a topological space Y is continuous if and only if $f(x_n) \to f(x)$ in Y whenever $x_n \to x$ in X. Such is also the case for any topological space X whose topology is determined by the sequences it renders convergent. (These are called *first countable spaces*.)

Let X, Y be normed spaces and f an operator from X into Y. We say that f is *compact* if $f(A)$ is precompact in Y whenever A is bounded in X. (A bounded set in a normed space is one which is contained in the ball $B_R(0)$ for some R.) If f is continuous and compact, we say that f is *completely continuous*. We say that f is *bounded* if $f(A)$ is bounded in Y whenever A is bounded in X.

Every compact operator is bounded. Every bounded linear operator is continuous. Therefore, every compact linear operator is completely continuous. The norm of a linear operator f is $\sup\{\|f(x)\,;\,Y\| : \|x\,;\,X\| \leq 1\}$.

1.25   **(Imbeddings)**   We say the normed space X is *imbedded* in the normed space Y, and we write $X \to Y$ to designate this imbedding, provided that

    (i) X is a vector subspace of Y, and

    (ii) the identity operator I defined on X into Y by $Ix = x$ for all $x \in X$ is continuous.

Since I is linear, (ii) is equivalent to the existence of a constant M such that

$$\|Ix\,;\,Y\| \leq M \, \|x\,;\,X\|\,, \qquad x \in X.$$

Sometimes the requirement that X be a subspace of Y and I be the identity map is weakened to allow as imbeddings certain canonical transformations of X into Y. Examples are trace imbeddings of Sobolev spaces as well as imbeddings of Sobolev spaces into spaces of continuous functions. See Chapter 5.

We say that X is *compactly imbedded* in Y if the imbedding operator I is compact.

Spaces of Continuous Functions

1.26   Let Ω be a domain in $\mathbb{R}^n$. For any nonnegative integer m let $C^m(\Omega)$ denote the vector space consisting of all functions ϕ which, together with all their partial derivatives $D^\alpha \phi$ of orders $|\alpha| \leq m$, are continuous on Ω. We abbreviate $C^0(\Omega) \equiv C(\Omega)$. Let $C^\infty(\Omega) = \bigcap_{m=0}^{\infty} C^m(\Omega)$.

The subspaces $C_0(\Omega)$ and $C_0^\infty(\Omega)$ consist of all those functions in $C(\Omega)$ and $C^\infty(\Omega)$, respectively, that have compact support in Ω.

1.27   (Spaces of Bounded, Continuous Functions)   Since Ω is open, functions in $C^m(\Omega)$ need not be bounded on Ω. We define $C_B^m(\Omega)$ to consist of those functions $\phi \in C^m(\Omega)$ for which $D^\alpha u$ is bounded on Ω for $0 \leq |\alpha| \leq m$. $C_B^m(\Omega)$ is a Banach space with norm given by

$$\left\| \phi \, ; C_B^m(\Omega) \right\| = \max_{0 \leq \alpha \leq m} \sup_{x \in \Omega} |D^\alpha \phi(x)|.$$

1.28   (Spaces of Bounded, Uniformly Continuous Functions)   If $\phi \in C(\Omega)$ is bounded and uniformly continuous on Ω, then it possesses a unique, bounded, continuous extension to the closure $\overline{\Omega}$ of Ω. We define the vector space $C^m(\overline{\Omega})$ to consist of all those functions $\phi \in C^m(\Omega)$ for which $D^\alpha \phi$ is bounded and uniformly continuous on Ω for $0 \leq |\alpha| \leq m$. (This convenient abuse of notation leads to ambiguities if Ω is unbounded; e.g., $C^m(\overline{\mathbb{R}^n}) \neq C^m(\mathbb{R}^n)$ even though $\overline{\mathbb{R}^n} = \mathbb{R}^n$.) $C^m(\overline{\Omega})$ is a closed subspace of $C_B^m(\Omega)$, and therefore also a Banach space with the same norm

$$\left\| \phi \, ; C^m(\overline{\Omega}) \right\| = \max_{0 \leq \alpha \leq m} \sup_{x \in \Omega} |D^\alpha \phi(x)|.$$

1.29   (Spaces of Hölder Continuous Functions)   If $0 < \lambda \leq 1$, we define $C^{m,\lambda}(\overline{\Omega})$ to be the subspace of $C^m(\overline{\Omega})$ consisting of those functions ϕ for which, for $0 \leq \alpha \leq m$, $D^\alpha \phi$ satisfies in Ω a Hölder condition of exponent λ, that is, there exists a constant K such that

$$|D^\alpha \phi(x) - D^\alpha \phi(y)| \leq K|x - y|^\lambda, \qquad x, y \in \Omega.$$

$C^{m,\lambda}(\overline{\Omega})$ is a Banach space with norm given by

$$\left\| \phi \, ; C^{m,\lambda}(\overline{\Omega}) \right\| = \left\| \phi \, ; C^m(\overline{\Omega}) \right\| + \max_{0 \leq |\alpha| \leq m} \sup_{\substack{x, y \in \Omega \\ x \neq y}} \frac{|D^\alpha \phi(x) - D^\alpha \phi(y)|}{|x - y|^\lambda}.$$

It should be noted that for $0 < \nu < \lambda \leq 1$,

$$C^{m,\lambda}(\overline{\Omega}) \subsetneqq C^{m,\nu}(\overline{\Omega}) \subsetneqq C^m(\overline{\Omega}).$$

Since Lipschitz continuity (that is, Hölder continuity of exponent 1) does not imply everywhere differentiability, it is clear that $C^{m,1}(\overline{\Omega}) \not\subset C^{m+1}(\overline{\Omega})$. In general, $C^{m+1}(\overline{\Omega}) \not\subset C^{m,1}(\overline{\Omega})$ either, but the inclusion is possible for many domains Ω, for instance convex ones as can be seen by using the Mean-Value Theorem. (See Theorem 1.34.)

1.30 If Ω is bounded, the following two well-known theorems provide useful criteria for the denseness and compactness of subsets of $C(\overline{\Omega})$. If $\phi \in C(\overline{\Omega})$, we may regard ϕ as defined on $\overline{\Omega}$, that is, we identify ϕ with its unique continuous extension to the closure of Ω.

1.31   THEOREM   (The Stone-Weierstrass Theorem)   Let Ω be a bounded domain in $\mathbb{R}^n$. A subset $\mathscr{A}$ of $C(\overline{\Omega})$ is dense in $C(\overline{\Omega})$ if it has the following four properties:

  (i) If $\phi, \psi \in \mathscr{A}$ and $c \in \mathbb{C}$, then $\phi + \psi$, $\phi\psi$, and $c\phi$ all belong to $\mathscr{A}$.

  (ii) If $\phi \in \mathscr{A}$, then $\overline{\phi} \in \mathscr{A}$, where $\overline{\phi}$ is the complex conjugate of ϕ.

  (iii) If $x, y \in \overline{\Omega}$ and $x \neq y$, there exists $\phi \in \mathscr{A}$ such that $\phi(x) \neq \phi(y)$.

  (iv) If $x \in \overline{\Omega}$, there exists $\phi \in \mathscr{A}$ such that $\phi(x) \neq 0$. ∎

1.32   COROLLARY   If Ω is bounded in $\mathbb{R}^n$, then the set P of all polynomials in $x = (x_1, \ldots, x_n)$ having rational-complex coefficients is dense in $C(\overline{\Omega})$. (A *rational-complex* number is a number of the form $c_1 + ic_2$ where c_1 and c_2 are rational numbers.) Hence $C(\overline{\Omega})$ is separable.

Proof.   The set of all polynomials in x is dense in $C(\overline{\Omega})$ by the Stone-Weierstrass Theorem. Any polynomial can be uniformly approximated on the compact set $\overline{\Omega}$ by elements of the countable set P, which is therefore also dense in $C(\overline{\Omega})$. ∎

1.33   THEOREM   (The Ascoli-Arzela Theorem)   Let Ω be a bounded domain in $\mathbb{R}^n$. A subset K of $C(\overline{\Omega})$ is precompact in $C(\overline{\Omega})$ if the following two conditions hold:

  (i) There exists a constant M such that $|\phi(x)| \leq M$ holds for every $\phi \in K$ and $x \in \Omega$.

  (ii) For every $\epsilon > 0$ there exists $\delta > 0$ such that if $\phi \in K$, $x, y \in \Omega$, and $|x - y| < \delta$, then $|\phi(x) - \phi(y)| < \epsilon$. ∎

The following is a straightforward imbedding theorem for the various continuous function spaces introduced above. It is a preview of the main attraction, the Sobolev imbedding theorem of Chapter 5.

1.34   THEOREM   Let m be a nonnegative integer and let $0 < \nu < \lambda \leq 1$. Then the following imbeddings exist:

$$C^{m+1}(\overline{\Omega}) \to C^m(\overline{\Omega}), \tag{3}$$

$$C^{m,\nu}(\overline{\Omega}) \to C^m(\overline{\Omega}), \tag{4}$$

$$C^{m,\lambda}(\overline{\Omega}) \to C^{m,\nu}(\overline{\Omega}). \tag{5}$$

If Ω is bounded, then imbeddings (4) and (5) are compact. If Ω is convex, we have the further imbeddings

$$C^{m+1}(\overline{\Omega}) \to C^{m,1}(\overline{\Omega}), \tag{6}$$

$$C^{m+1}(\overline{\Omega}) \to C^{m,\lambda}(\overline{\Omega}). \tag{7}$$

If Ω is convex and bounded, then imbeddings (3) is compact, and so is (7) if $\lambda < 1$.

Proof. The existence of imbeddings (3) and (4) follows from the obvious inequalities

$$\left\| \phi \,;\, C^m(\overline{\Omega}) \right\| \leq \left\| \phi \,;\, C^{m+1}(\overline{\Omega}) \right\|,$$

$$\left\| \phi \,;\, C^m(\overline{\Omega}) \right\| \leq \left\| \phi \,;\, C^{m,\lambda}(\overline{\Omega}) \right\|.$$

To establish (5) we note that for $|\alpha| \leq m$,

$$\sup_{\substack{x,y \in \Omega \\ 0 < |x-y| < 1}} \frac{|D^\alpha \phi(x) - D^\alpha \phi(y)|}{|x - y|^\nu} \leq \sup_{x,y \in \Omega} \frac{|D^\alpha \phi(x) - D^\alpha \phi(y)|}{|x - y|^\lambda}$$

and

$$\sup_{\substack{x,y \in \Omega \\ |x-y| \geq 1}} \frac{|D^\alpha \phi(x) - D^\alpha \phi(y)|}{|x - y|^\nu} \leq 2 \sup_{x \in \Omega} |D^\alpha \phi(x)|,$$

from which we conclude that

$$\left\| \phi \,;\, C^{m,\nu}(\overline{\Omega}) \right\| \leq 2 \left\| \phi \,;\, C^{m,\lambda}(\overline{\Omega}) \right\|.$$

If Ω is convex and $x, y \in \Omega$, then by the Mean-Value Theorem there is a point $z \in \Omega$ on the line segment joining x and y such that $D^\alpha \phi(x) - D^\alpha \phi(y)$ is given by $(x - y) \cdot \nabla D^\alpha \phi(z)$, where $\nabla u = (D_1 u, \ldots, D_n u)$. Thus

$$|D^\alpha \phi(x) - D^\alpha \phi(y)| \leq n|x - y| \left\| \phi \,;\, C^{m+1}(\overline{\Omega}) \right\|, \tag{8}$$

and so

$$\left\| \phi \,;\, C^{m,1}(\overline{\Omega}) \right\| \leq n \left\| \phi \,;\, C^{m+1}(\overline{\Omega}) \right\|.$$

Thus (6) is proved, and (7) follows from (5) and (6).

Now suppose that Ω is bounded. If A is a bounded set in $C^{0,\lambda}(\overline{\Omega})$, then there exists M such that $\left\| \phi \,;\, C^{0,\lambda}(\overline{\Omega}) \right\| \leq M$ for all $\phi \in A$. But then $|\phi(x) - \phi(y)| \leq M|x-y|^\lambda$ for all $\phi \in A$ and all $x, y \in \Omega$, whence A is precompact in $C(\overline{\Omega})$ by the Ascoli-Arzela Theorem 1.33. This proves the compactness of (4) for $m = 0$. If $m \geq 1$ and

A is bounded in $C^{m,\lambda}(\overline{\Omega})$, then A is bounded in $C^{0,\lambda}(\overline{\Omega})$ and there is a sequence $\{\phi_j\} \subset A$ such that $\phi_j \to \phi$ in $C(\overline{\Omega})$. But $\{D_1\phi_j\}$ is also bounded in $C^{0,\lambda}(\overline{\Omega})$ so there exists a subsequence of $\{\phi_j\}$ which we again denote by $\{\phi_j\}$ such that $D_1\phi_j \to \psi_1$ in $C(\overline{\Omega})$. Convergence in $C(\overline{\Omega})$ being uniform convergence on Ω, we have $\psi_1 = D_1\phi$. We may continue to extract subsequences in this manner until we obtain one for which $D^\alpha\phi_j \to D^\alpha\phi$ in $C(\overline{\Omega})$ for each α satisfying $0 \le |\alpha| \le m$. This proves the compactness of (4). For (5) we argue as follows:

$$\frac{|D^\alpha\phi(x) - D^\alpha\phi(y)|}{|x-y|^\nu} = \left(\frac{|D^\alpha\phi(x) - D^\alpha\phi(y)|}{|x-y|^\lambda}\right)^{\nu/\lambda} |D^\alpha\phi(x) - D^\alpha\phi(y)|^{1-\nu/\lambda}$$
$$\le \text{const}\,|D^\alpha\phi(x) - D^\alpha\phi(y)|^{1-\nu/\lambda} \tag{9}$$

for all ϕ in a bounded subset of $C^{m,\lambda}(\overline{\Omega})$. Since (9) shows that any sequence bounded in $C^{m,\lambda}(\overline{\Omega})$ and converging in $C^m(\overline{\Omega})$ is Cauchy and so converges in $C^{m,\nu}(\overline{\Omega})$, the compactness of (5) follows from that of (4).

Finally, if Ω is both convex and bounded, the compactness of (3) and (7) follows from composing the continuous imbedding (6) with the compact imbeddings (4) and (5) for the case $\lambda = 1$. ∎

1.35   The existence of imbeddings (6) and (7), as well as the compactness of (3) and (7), can be obtained under less restrictive hypotheses than the convexity of Ω. For instance, if every pair of points $x, y \in \Omega$ can be joined by a rectifiable arc in Ω having length not exceeding some fixed multiple of $|x - y|$, then we can obtain an inequality similar to (8) and carry out the proof. We leave it to the reader to show that (6) is not compact.

The Lebesgue Measure in $\mathbb{R}^n$

1.36   Many of the vector spaces considered in this monograph consist of functions integrable in the Lebesgue sense over domains in $\mathbb{R}^n$. While we assume that most readers are familiar with Lebesgue measure and integration, we nevertheless include here a brief discussion of that theory, especially those aspects of it relevant to the study of the L^p spaces and Sobolev spaces considered hereafter. All proofs are omitted. For a more complete and systematic discussion of the Lebesgue theory, as well as more general measures and integrals, we refer the reader to any of the books [Fo], [Ro], [Ru2], and [Sx].

1.37   (Sigma Algebras)   A collection Σ of subsets of $\mathbb{R}^n$ is called a σ-*algebra* if the following conditions hold:

(i) $\mathbb{R}^n \in \Sigma$.

(ii) If $A \in \Sigma$, then its complement $A^c \in \Sigma$.

(iii) If $A_j \in \Sigma$, $j = 1, 2, \ldots$, then $\bigcup_{j=1}^{\infty} \in \Sigma$.

It follows from (i)–(iii) that:

(iv) The empty set $\emptyset \in \Sigma$.

(v) If $A_j \in \Sigma$, $j = 1, 2, \ldots$, then $\bigcap_{j=1}^{\infty} \in \Sigma$.

(vi) If A, $B \in \Sigma$, then $A - B = A \cap B^c \in \Sigma$.

1.38 (Measures) By a *measure* μ on a σ-algebra Σ we mean a function on Σ taking values in either $\mathbb{R} \cup \{+\infty\}$ (a *positive measure*) or $\mathbb{C}$ (a *complex measure*) which is *countably additive* in the sense that

$$\mu \left(\bigcup_{j=1}^{\infty} A_j \right) = \sum_{j=1}^{\infty} \mu(A_j)$$

whenever $A_j \in \Sigma$, $j = 1, 2, \ldots$ and the sets A_j are pairwise disjoint, that is, $A_j \cap A_k = \emptyset$ for $j \neq k$. For a complex measure the series on the right must converge to the same sum for all permutations of the indices in the sequence $\{A_j\}$, and so must be absolutely convergent. If μ is a positive measure and if A, $B \in \Sigma$ and $A \subset B$, then $\mu(A) \leq \mu(B)$. Also, if $A_j \in \Sigma$, $j = 1, 2, \ldots$ and $A_1 \subset A_2 \subset \cdots$, then $\mu \left(\bigcup_{j=1}^{\infty} A_j \right) = \lim_{j \to \infty} \mu(A_j)$.

1.39 THEOREM (Existence of Lebesgue Measure) There exists a σ-algebra Σ of subsets of $\mathbb{R}^n$ and a positive measure μ on Σ having the following properties:

(i) Every open set in $\mathbb{R}^n$ belongs to Σ.

(ii) If $A \subset B$, $B \in \Sigma$, and $\mu(B) = 0$, then $A \in \Sigma$ and $\mu(A) = 0$.

(iii) If $A = \{x \in \mathbb{R}^n : a_j \leq x_j \leq b_j, \ j = 1, 2, \ldots, n\}$, then $A \in \Sigma$ and $\mu(A) = \prod_{j=1}^{n}(b_j - a_j)$.

(iv) μ is translation invariant. This means that if $x \in \mathbb{R}^n$ and $A \in \Sigma$, then $x + A = \{x + y : y \in A\} \in \Sigma$, and $\mu(x + A) = \mu(A)$. ∎

The elements of Σ are called *(Lebesgue) measurable subsets* of $\mathbb{R}^n$, and μ is called the *(Lebesgue) measure* in $\mathbb{R}^n$. (We normally suppress the word "Lebesgue" in these terms as it is the measure on $\mathbb{R}^n$ we mainly use.) For $A \in \Sigma$ we call $\mu(A)$ the *measure of A* or the *volume of A*, since Lebesgue measure is the natural extension of volume in $\mathbb{R}^3$. While we make no formal distinction between "measure" and "volume" for sets that are easily visualized geometrically, such as balls, cubes, and domains, and we write vol(A) in place of $\mu(A)$ in these cases. Of course the terms *length* and *area* are more appropriate in $\mathbb{R}^1$ and $\mathbb{R}^2$.

The reader may wonder whether in fact all subsets of $\mathbb{R}^n$ are Lebesgue measurable. The answer depends on the axioms of one's set theory. Under the most common axioms the answer is no; it is possible using the Axiom of Choice to construct a

nonmeasurable set. There is a version of set theory where every subset of $\mathbb{R}^n$ is measurable, but the Hahn-Banach theorem 1.13 becomes false in that version.

1.40   (Almost Everywhere)   If $B \subset A \subset \mathbb{R}^n$ and $\mu(B) = 0$, then any condition that holds on the set $A - B$ is said to hold *almost everywhere* (abbreviated a.e.) in A. It is easily seen that any countable set in $\mathbb{R}^n$ has measure zero. The converse is, however, not true.

1.41   (Measurable Functions)   A function f defined on a measurable set and having values in $\mathbb{R} \cup \{-\infty, +\infty\}$ is itself called *measurable* if the set

$$\{x \; : \; f(x) > a\}$$

is measurable for every real a. Some of the more important aspects of this definition are listed in the following theorem.

1.42   THEOREM   (a) If f is measurable, so is $|f|$.

  (b) If f and g are measurable and real-valued, so are $f + g$ and fg.

  (c) If $\{f_j\}$ is a sequence of measurable functions, then $\sup_j f_j$, $\inf_j f_j$, $\limsup_{j \to \infty} f_j$, and $\liminf_{j \to \infty} f_j$ are measurable.

  (d) If f is continuous and defined on a measurable set, then f is measurable.

  (e) If f is continuous on $\mathbb{R}$ into $\mathbb{R}$ and g is measurable and real-valued, then the composition $f \circ g$ defined by $f \circ g(x) = f\big((g(x)\big)$ is measurable.

  (f) **(Lusin's Theorem)** If f is measurable and $f(x) = 0$ for $x \in A^c$ where $\mu(A) < \infty$, and if $\epsilon > 0$, then there exists a function $g \in C_0(\mathbb{R}^n)$ such that $\sup_{x \in \mathbb{R}^n} g(x) \leq \sup_{x \in \mathbb{R}^n} f(x)$ and $\mu\left(\{x \in \mathbb{R}^n \; : \; f(x) \neq g(x)\}\right) < \epsilon.$ ∎

1.43   (Characteristic and Simple Functions)   Let $A \subset \mathbb{R}^n$. The function χ_A defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

is called the *characteristic function* of A. A real-valued function s on $\mathbb{R}^n$ is called a *simple function* if its range is a finite set of real numbers. If for every x, we have $s(x) \in \{a_1, \ldots, a_n\}$, then $s = \sum_{j=1}^{m} \chi_{A_j}(x)$, where $A_j = \{x \in \mathbb{R}^n \; : \; s(x) = a_j\}$, and s is measurable if and only if $A_1, A_2, \ldots, A_m$ are all measurable. Because of the following approximation theorem, simple functions are a very useful tool in integration theory.

1.44   THEOREM   Given a real-valued function f with domain $A \subset \mathbb{R}^n$ there is a sequence $\{s_j\}$ of simple functions converging pointwise to f on A. If f is bounded, $\{s_j\}$ may be chosen so that the convergence is uniform. If f is measurable, each s_j may be chosen measurable. If f is nonnegative-valued, the sequence $\{s_j\}$ may be chosen to be monotonically increasing at each point. ∎

The Lebesgue Integral

1.45   We are now in a position to define the *(Lebesgue) integral* of a measurable, real-valued function defined on a measurable subset $A \subset \mathbb{R}^n$. For a simple function $s = \sum_{j=1}^{m} a_j \chi_{A_j}$, where $A_j \subset A$, A_j measurable, we define

$$\int_A s(x)\, dx = \sum_{j=1}^{m} a_j \mu(A_j). \tag{10}$$

If f is measurable and nonnegative-valued on A, we define

$$\int_A f(x)\, dx = \sup \int_A s(x)\, dx, \tag{11}$$

where the supremum is taken over measurable, simple functions s vanishing outside A and satisfying $0 \le s(x) \le f(x)$ in A. If f is a nonnegative simple function, then the two definitions of $\int_A f(x)\, dx$ given by (10) and (11) coincide. Note that the integral of a nonnegative function may be $+\infty$.

If f is measurable and real-valued, we set $f = f^+ - f^-$, where $f^+ = \max(f, 0)$ and $f^- = -\min(f, 0)$ are both measurable and nonnegative. We define

$$\int_A f(x)\, dx = \int_A f^+(x)\, dx - \int_A f^-(x)\, dx$$

provided at least one of the integrals on the right is finite. If *both integrals* are finite, we say that f is (Lebesgue) *integrable* on A. The class of integrable functions on A is denoted $L^1(A)$.

1.46 THEOREM   Assume all of the functions and sets appearing below are measurable.

(a) If f is bounded on A and $\mu(A) < \infty$, then $f \in L^1(A)$.

(b) If $a \le f(x) \le b$ for all $x \in A$ and if $\mu(A) < \infty$, then

$$a\, \mu(A) \le \int_A f(x)\, dx \le b\, \mu(A).$$

(c) If $f(x) \le g(x)$ for all $x \in A$, and if both integrals exist, then

$$\int_A f(x)\, dx \le \int_A g(x)\, dx.$$

(d) If $f, g \in L^1(A)$, then $f + g \in L^1(A)$ and

$$\int_A (f + g)(x)\, dx = \int_A f(x)\, dx + \int_A g(x)\, dx.$$

(e) If $f \in L^1(A)$ and $c \in \mathbb{R}$, then $cf \in L^1(A)$ and

$$\int_a (cf)(x)\, dx = c \int_A f(x)\, dx.$$

(f) If $f \in L^1(A)$, then $|f| \in L^1(A)$ and

$$\left| \int_A f(x)\, dx \right| \leq \int_A |f(x)|\, dx.$$

(g) If $f \in L^1(A)$ and $B \subset A$, then $f \in L^1(B)$. If, in addition, $f(x) \geq 0$ for all $x \in A$, then

$$\int_B f(x)\, dx \leq \int_A f(x)\, dx.$$

(h) If $\mu(A) = 0$, then $\int_A f(x)\, dx = 0$.

(i) If $f \in L^1(A)$ and $\int_B f(x) = 0$ for every $B \subset A$, then $f(x) = 0$ a.e. on A. ∎

One consequence of part (i) and the additivity of the integral is that sets of measure zero may be ignored for purposes of integration. That is, if f and g are measurable on A and if $f(x) = g(x)$ a.e. on A, then $\int_A f(x)\, dx = \int_A g(x)\, dx$. Accordingly, two elements of $L^1(A)$ are considered identical if they are equal almost everywhere. Thus the elements of $L_1(A)$ are actually not functions but equivalence classes of functions; two functions belong to the same element of $L_1(A)$ if they are equal a.e. on A. Nevertheless, we will continue to refer (loosely) to the elements of $L_1(A)$ as functions on A.

1.47 THEOREM If f is either an element of $L^1(\mathbb{R}^n)$ or measurable and nonnegative on $\mathbb{R}^n$, then the set function λ defined by

$$\lambda(A) = \int_A f(x)\, dx$$

is countably additive, and hence a measure on the σ-algebra of Lebesgue measurable subsets of $\mathbb{R}^n$. ∎

The following three theorems are concerned with the interchange of integration and limit processes.

1.48 THEOREM (The Monotone Convergence Theorem) Let $A \subset \mathbb{R}^n$ be measurable and let $\{f_j\}$ be a sequence of measurable functions satisfying $0 \leq f_1(x) \leq f_2(x) \leq \cdots$ for every $x \in A$. Then

$$\lim_{j \to \infty} \int_A f_j(x)\, dx = \int_A \left(\lim_{j \to \infty} f_j(x) \right) dx. \quad ∎$$

1.49 THEOREM (Fatou's Lemma)   Let $A \subset \mathbb{R}^n$ be measurable and let $\{f_j\}$ be a sequence of nonnegative measurable functions. Then

$$\int_A \left(\liminf_{j \to \infty} \right) dx \le \liminf_{j \to \infty} \int_A f_j(x) \, dx. \blacksquare$$

1.50 THEOREM (The Dominated Convergence Theorem)   Let $A \subset \mathbb{R}^n$ be measurable and let $\{f_j\}$ be a sequence of measurable functions converging to a limit pointwise on A. If there exists a function $g \in L^1(A)$ such that $|f_j(x)| \le g(x)$ for every j and all $x \in A$, then

$$\lim_{j \to \infty} \int_A f_j(x) \, dx = \int_A \left(\lim_{j \to \infty} f_j(x) \right) dx. \blacksquare$$

1.51 (Integrals of Complex-Valued Functions)   The integral of a complex-valued function over a measurable set $A \subset \mathbb{R}^n$ is defined as follows. Set $f = i + iv$, where u and v are real-valued and call f measurable if and only if u and v are measurable. We say f is integrable over A, and write $f \in L^1(A)$, provided $|f| = (u^2 + v^2)^{1/2}$ belongs to $L^1(A)$ in the sense described in Paragraph 1.45. For $f \in L^1(A)$, and only for such f, the integral is defined by

$$\int_A f(x) \, dx = \int_A u(x) \, dx + i \int_A v(x) \, dx.$$

It is easily checked that $f \in L^1(A)$ if and only if $u, v \in L^1(A)$. Theorem 1.42(a,b,d–f), Theorem 1.46(a,d–i), Theorem 1.47 (assuming $f \in L^1(\mathbb{R}^n)$), and Theorem 1.50 all extend to cover the case of complex f.

The following theorem enables us to express certain complex measures in terms of Lebesgue measure μ. It is the converse of Theorem 1.47.

1.52 THEOREM (The Radon-Nikodym Theorem)   Let λ be a complex measure defined on the σ-algebra Σ of Lebesgue measurable subsets of $\mathbb{R}^n$. Suppose that $\lambda(A) = 0$ for every $A \in \Sigma$ for which $\mu(A) = 0$. Then there exists $f \in L^1(\mathbb{R}^n)$ such that for every $A \in \Sigma$

$$\lambda(A) = \int_A f(x) \, dx.$$

The function f is uniquely determined by λ up to sets of measure zero. $\blacksquare$

1.53   If f is a function defined on a subset A of $\mathbb{R}^{n+m}$, we may regard f as depending on the pair of variables (x, y) with $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. The integral of f over A is then denoted by

$$\int_A f(x, y) \, dx \, dy$$

or, if it is desired to have the integral extend over all of $\mathbb{R}^{n+m}$,

$$\int_{\mathbb{R}^{n+m}} f(x, y) \chi_A(x, y) \, dx \, dy,$$

where χ_A is the characteristic function of A. In particular, if $A \subset \mathbb{R}^n$, we may write

$$\int_A f(x) \, dx = \int_A f(x_1, \ldots, x_n) \, dx_1 \cdots dx_n.$$

1.54 THEOREM (Fubini's Theorem) Let f be a measurable function on $\mathbb{R}^{m+n}$ and suppose that at least one of the integrals

$$I_1 = \int_{\mathbb{R}^{n+m}} |f(x, y)| \, dx, dy,$$

$$I_2 = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} |f(x, y)| \, dx \right) dy, \qquad (12)$$

$$I_3 = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} |f(x, y)| \, dy \right) dx$$

exists and is finite. For I_2, we mean by this that there is an integrable function g on $\mathbb{R}^n$ such that $g(y)$ is equal to the inner integral for almost all y, and similarly for I_3. Then

    (a) $f(\cdot, y) \in L^1(\mathbb{R}^n)$ for almost all $y \in \mathbb{R}^m$.

    (b) $f(x, \cdot) \in L^1(\mathbb{R}^m)$ for almost all $x \in \mathbb{R}^n$.

    (c) $\int_{R^m} f(\cdot, y) \, dy \in L^1(\mathbb{R}^n)$.

    (d) $\int_{R^n} f(x, \cdot) \, dx \in L^1(\mathbb{R}^m)$.

    (e) $I_1 = I_2 = I_3$.

Distributions and Weak Derivatives

1.55 We require in subsequent chapters some of the basic concepts and techniques of the Schwartz theory of distributions [Sch], and we present here a brief description of those aspects of the theory that are relevant for our purposes. Of special importance is the notion of weak or distributional derivative of an integrable function. One of the standard definitions of Sobolev spaces is phrased in terms of such derivatives. (See Paragraph 3.2.) Besides [Sch], the reader is referred to [Ru1] and [Y] for more complete treatments of the spaces $\mathscr{D}(\Omega)$ and $\mathscr{D}'(\Omega)$ introduced below, as well as useful generalizations of these spaces.

1.56 (Test Functions) Let Ω be a domain in $\mathbb{R}^n$. A sequence $\{\phi_j\}$ of functions belonging to $C_0^\infty(\Omega)$ is said to *converge in the sense of the space* $\mathscr{D}(\Omega)$ to the function $\phi \in C_0^\infty(\Omega)$ provided the following conditions are satisfied:

(i) there exists $K \Subset \Omega$ such that supp $(\phi_j - \phi) \subset K$ for every j, and

(ii) $\lim_{j\to\infty} D^\alpha \phi_j(x) = D^\alpha \phi(x)$ uniformly on K for each multi-index α.

There is a locally convex topology on the vector space $C_0^\infty(\Omega)$ which respect to which a linear functional T is continuous if and only if $T(\phi_j) \to T(\phi)$ in $\mathbb{C}$ whenever $\phi_j \to \phi$ in the sense of the space $\mathscr{D}(\Omega)$. Equipped with this topology, $C_0^\infty(\Omega)$ becomes a TVS called $\mathscr{D}(\Omega)$ whose elements are called *test functions*. $\mathscr{D}(\Omega)$ is not a normable space. (We ignore the question of uniqueness of the topology asserted above. It uniquely determines the dual of $\mathscr{D}(\Omega)$ which is sufficient for our purposes.)

1.57 (Schwartz Distributions)   The dual space $\mathscr{D}'(\Omega)$ of $\mathscr{D}(\Omega)$ is called the *space of (Schwartz) distributions* on Ω. $\mathscr{D}'(\Omega)$ is given the weak-star topology as the dual of $\mathscr{D}(\Omega)$, and is a locally convex TVS with that topology. We summarize the vector space and convergence operations in $\mathscr{D}'(\Omega)$ as follows: if S, T, T_j belong to $\mathscr{D}'(\Omega)$ and $c \in \mathbb{C}$, then

$$(S + T)(\phi) = S(\phi) + T(\phi), \qquad \phi \in \mathscr{D}(\Omega),$$
$$(cT)(\phi) = c\, T(\phi), \qquad \phi \in \mathscr{D}(\Omega),$$

$T_j \to T$ in $\mathscr{D}'(\Omega)$ if and only if $T_j(\phi) \to T(\phi)$ in $\mathbb{C}$ for every $\phi \in \mathscr{D}(\Omega)$.

1.58 (Locally Integrable Functions)   A function u defined almost everywhere on Ω is said to be *locally integrable* on Ω provided $u \in L^1(U)$ for every open $U \Subset \Omega$. In this case we write $u \in L_{\mathrm{loc}}^1(\Omega)$. Corresponding to every $u \in L_{\mathrm{loc}}^1(\Omega)$ there is a distribution $T_u \in \mathscr{D}'(\Omega)$ defined by

$$T_u(\phi) = \int_\Omega u(x)\phi(x)\, dx, \qquad \phi \in \mathscr{D}(\Omega). \tag{13}$$

Evidently T_u, thus defined, is a linear functional on $\mathscr{D}(\Omega)$. To see that it is continuous, suppose that $\phi_j \to \phi$ in $\mathscr{D}(\Omega)$. Then there exists $K \Subset \Omega$ such that supp $(\phi_j - \phi) \subset K$ for all j. Thus

$$|T_u(\phi_j) - T_u(\phi)| \leq \sup_{x\in K} |\phi_j(x) - \phi(x)| \int_K |u(x)|\, dx.$$

The right side of the above inequality tends to zero as $j \to \infty$ since $\phi_j \to \phi$ uniformly on K.

1.59   Not every distribution $T \in \mathscr{D}'(\Omega)$ is of the form T_u defined by (13) for some $u \in L_{\mathrm{loc}}^1(\Omega)$. Indeed, if $0 \in \Omega$, there can be no locally integrable function δ on Ω such that for every $\phi \in \mathscr{D}(\Omega)$

$$\int_\Omega \delta(x)\phi(x)\, dx = \phi(0).$$

However, the linear functional δ defined on $\mathscr{D}(\Omega)$ by

$$\delta(\phi) = \phi(0) \tag{14}$$

is easily seen to be continuous and hence a distribution on Ω. It is called a *Dirac distribution*.

1.60   (Derivatives of Distributions)   Let $u \in C^1(\Omega)$ and $\phi \in \mathscr{D}(\Omega)$. Since ϕ vanishes outside some compact subset of Ω, we obtain by integration by parts in the variable x_j

$$\int_\Omega \left(\frac{\partial}{\partial x_j} u(x)\right) \phi(x)\, dx = -\int_\Omega u(x) \left(\frac{\partial}{\partial x_j}\phi(x)\right) dx.$$

Similarly, if $u \in C^{|\alpha|}(\Omega)$, then integration by parts $|\alpha|$ times leads to

$$\int_\Omega \left(D^\alpha u(x)\right)\phi(x)\, dx = (-1)^{|\alpha|} \int_\Omega u(x) D^\alpha \phi(x)\, dx.$$

This motivates the following definition of the derivative $D^\alpha T$ of a distribution $T \in \mathscr{D}'(\Omega)$:

$$(D^\alpha T)(\phi) = (-1)^{|\alpha|} T(D^\alpha \phi). \tag{15}$$

Since $D^\alpha \phi \in \mathscr{D}(\Omega)$ whenever $\phi \in \mathscr{D}(\Omega)$, $D^\alpha T$ is a functional on $\mathscr{D}(\Omega)$, and it is clearly linear. We show that it is continuous, and hence a distribution on Ω. To this end suppose $\phi_j \to \phi$ in $\mathscr{D}(\Omega)$. Then

$$\operatorname{supp}\left(D^\alpha(\phi_j - \phi)\right) \subset \operatorname{supp}\left(\phi_j - \phi\right) \subset K$$

for some $K \Subset \Omega$. Moreover,

$$D^\beta\left(D^\alpha(\phi_j - \phi)\right) = D^{\beta+\alpha}(\phi_j - \phi)$$

converges to zero uniformly on K as $j \to \infty$ for each multi-index β. Hence $D^\alpha \phi_j \to D^\alpha \phi$ in $\mathscr{D}(\Omega)$. Since $T \in \mathscr{D}'(\Omega)$ it follows that

$$D^\alpha T(\phi_j) = (-1)^{|\alpha|} T(D^\alpha \phi_j) \to (-1)^{|\alpha|} T(D^\alpha \phi) = D^\alpha T(\phi)$$

in $\mathbb{C}$. Thus $D^\alpha T \in \mathscr{D}'(\Omega)$.

We have shown that every distribution in $\mathscr{D}'(\Omega)$ possesses derivatives of all orders in $\mathscr{D}'(\Omega)$ in the sense of definition (15). Furthermore, the mapping D^α from $\mathscr{D}'(\Omega)$ into $\mathscr{D}'(\Omega)$ is continuous; if $T_j \to T$ in $\mathscr{D}'(\Omega)$ and $\phi \in \mathscr{D}(\Omega)$, then

$$D^\alpha T_j(\phi) = (-1)^{|\alpha|} T_j(D^\alpha \phi) \to (-1)^{|\alpha|} T(D^\alpha \phi) = D^\alpha T(\phi).$$

1.61   EXAMPLES

1. If $0 \in \Omega$ and $\delta \in \mathscr{D}'(\Omega)$ is the Dirac distribution defined by (14), then $D^\alpha \delta$ is given by

$$D^\alpha \delta(\phi) = (-1)^{|\alpha|} D^\alpha \phi(0).$$

2. If $\Omega = \mathbb{R}$ (i.e., $n = 1$) and $H \in L^1_{\text{loc}}(\mathbb{R})$ is the Heaviside function defined by

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0, \end{cases}$$

then the derivative $(T_H)'$ of the corresponding distribution T_H is δ. To see this, let $\phi \in \mathscr{D}(\mathbb{R})$ have support in the interval $[-a, a]$. Then

$$(T_H)'(\phi) = -T_H(\phi') = -\int_0^a \phi'(x)\, dx = \phi(0) = \delta(\phi).$$

1.62   (Weak Derivatives)   We now define the concept of a function being the weak derivative of another function. Let $u \in L^1_{\text{loc}}(\Omega)$. There may or may not exist a function $v_\alpha \in L^1_{\text{loc}}(\Omega)$ such that $T_{v_\alpha} = D^\alpha T_u$ in $\mathscr{D}'(\Omega)$. If such a v_α exists, it is unique up to sets of measure zero and is called the *weak* or *distributional* partial derivative of u, and is denoted by $D^\alpha u$. Thus $D^\alpha u = v_\alpha$ in the weak (or distributional) sense provided $v_\alpha \in L^1_{\text{loc}}(\Omega)$ satisfies

$$\int_\Omega u(x) D^\alpha \phi(x)\, dx = (-1)^{|\alpha|} \int_\Omega v_\alpha(x)\phi(x)\, dx$$

for every $\phi \in \mathscr{D}(\Omega)$.

If u is sufficiently smooth to have a continuous partial derivative $D^\alpha u$ in the usual (classical) sense, then $D^\alpha u$ is also a weak partial derivative of u. Of course, $D^\alpha u$ may exist in the weak sense without existing in the classical sense. We shall show in Theorem 3.17 that certain functions having weak derivatives (those in Sobolev spaces) can be suitably approximated by smooth functions.

1.63   Let us note in conclusion that distributions in Ω can be multiplied by smooth functions. If $T \in \mathscr{D}'(\Omega)$ and $\omega \in C^\infty(\Omega)$, the product $\omega T \in \mathscr{D}'(\Omega)$ is defined by

$$(\omega T)(\phi) = T(\omega\phi), \qquad \phi \in \mathscr{D}(\Omega).$$

If $T = T_u$ for some $u \in L^1_{\text{loc}}(\Omega)$, then $\omega T = T_{\omega u}$. The Leibniz rule (see Paragraph 1.2) is easily checked to hold for $D^\alpha(\omega T)$.