Chapter 1. Introduction

Chapter 2. Preliminaries
2.1. Three Basic Theorems in Functional Analysis
2.2. Compact Operators
2.3. The Lax-Milgram Theorems
2.4. Fredholm Operators
2.5. Weak Convergence and Compactness
2.6. Classical Function Spaces
2.7. Mollifiers, Cutoff Functions and Partitions of Unity
2.8. Domains with $C^{k,\alpha}$ Boundary

Chapter 3. Sobolev spaces
3.1. Weak Derivatives
3.2. Sobolev Spaces
3.3. Approximation by Smooth Functions
3.4. Traces
3.5. Difference Quotients

Chapter 4. Second Order Linear Elliptic Boundary Value Problems
4.1. Variational Formulations
4.2. The Dirichlet Problem Sans Lower Order Terms

Bibliography

Contents
CHAPTER 1

Introduction

This course concerns boundary value problems for second order elliptic equations given on domains in Euclidean space. More specifically, we will discuss the variational formulations of such problems, which is essential material for anyone interested in studying analysis or partial differential equations.

My intention is to move quickly through preliminary material and get to the heart of the course — a discussion of the variational formulation of elliptic boundary value problems and a presentation of basic existence, uniqueness and regularity results for them — as soon as possible.

This is the first draft of the notes for this course, and they were written in some haste. No doubt there are many errors and inconsistencies. I ask for your patience, and that you bring any errors you find to my attention. I am also open to any suggestions you may have for their improvement.

I made extensive use of the following texts while preparing these notes, and suggest them as references.

(1) “Partial Differential Equations” by Lawrence Evans.


(3) “Sobolev Spaces” by Robert Adams and John Fournier.

I also highly recommend the following texts which cover material beyond the scope of the course, but may be of some use to you.

(1) “Non-homogeneous boundary value problems” by J.L. Lions and E. Magenes discusses boundary value problems for higher order elliptic operators.

(2) Gerald Folland’s “Partial Differential Equations” contains good introductions to layer potentials and pseudodifferential calculus.

(3) Pierre Grisvard’s “Elliptic Problems in Nonsmooth Domains” first gives an excellent (but fast-paced) review of the material presented here and then goes on to discuss boundary value problems under somewhat weaker regularity assumptions than we make here.

Much of the material in the preliminaries — with Section 2.4 a notable exception — can be found in “Real Analysis: Modern Techniques and Their Application” by Gerald Folland.
Preliminaries

In this chapter, we review a number of basic definitions and results which will be used throughout these notes. I do not suggest that you read through this material in its entirety at the beginning of the course. Rather, I recommend that you consult this section as needed. Many of the results discussed here were originally developed in order to analyze partial differential equations and without this context, it is difficult to appreciate the utility of much of this material.

Throughout this chapter and these notes, all normed linear spaces are vector spaces over the field of real numbers. Small modifications must be made if normed linear spaces over the complex numbers are considered instead.

2.1. Three Basic Theorems in Functional Analysis

You should already be familiar with the following three basic theorems regarding Banach and normed linear spaces. If not, I suggest you refer to [4] or [1].

**Theorem 1 (Open mapping theorem).** Suppose that $T : X \to Y$ is a continuous linear mapping between Banach spaces. Then $T$ is surjective if and only if it is an open mapping (that is, if it takes open sets in $X$ to open sets in $Y$).

**Theorem 2 (Uniform boundedness principle).** Suppose that $X$ is a Banach space, and that $Y$ is a normed linear space. Suppose also that $F$ is a collection of bounded linear operators $X \to Y$. If for each $x \in X$, 
\[
\sup_{T \in F} \|Tx\| < \infty
\] (1)

then
\[
\sup_{T \in F} \|T\| < \infty.
\] (2)

**Theorem 3 (Hahn-Banach theorem).** Suppose that $Y$ is a subspace of a normed linear space $X$, and that $T : Y \to \mathbb{R}$ is a bounded linear functional. Then there is a bounded linear functional $\tilde{T} : X \to \mathbb{R}$ which extends $T$ (that is, $\tilde{T}(y) = T(y)$ for all $y \in Y$) whose norm is equal to that of $T$.

These three basic theorems have a large number of useful consequences. For instance, the following results are immediate consequences of the open mapping theorem

**Theorem 4 (Bounded inverse theorem).** Suppose that $X$ and $Y$ are Banach spaces. The inverse of a bijective bounded linear mapping $T : X \to Y$ is bounded.
Theorem 5 (Closed graph theorem). Suppose that $X$ and $Y$ are Banach spaces, and that $T : X \to Y$ is a linear operator. Then $T$ is bounded if and only if the graph of $T$

$$\{(x, y) \in X \times Y : Tx = y\}$$

is closed.

Suppose that $Y$ is a subspace of the Banach space $X$. We denote by $X/Y$ the vector space of cosets of $Y$. That is, $X/Y$ consists of the equivalence classes of the relation

$$x_1 \sim x_2 \text{ if and only if } x_1 - x_2 \in Y.$$  

(4)

We will denote the equivalence class to which the element $x$ belongs by $x + Y$. Note that if $Y$ is closed, then $X/Y$ is a Banach space when endowed with the norm

$$\|x + Y\| = \inf_{y \in Y} \|x - y\|.$$  

(5)

If $Y$ is not closed, then (5) is no longer a norm (it is instead a seminorm). The following is another consequence of the open mapping theorem.

Theorem 6. Suppose that $X$ and $Y$ are Banach spaces, and that $T : X \to Y$ is a continuous linear operator. If $\text{im}(T)$ is closed, then $\text{im}(T)$ is isomorphic to $X/\ker(T)$.

Proof. We define $\tilde{T} : X/\ker(T) \to \text{im}(Y)$ via the formula

$$\tilde{T}(x + \ker(T)) = T(x).$$

(6)

We observe that $\ker(T)$ is closed since $T$ is continuous, and that $T$ is bijective and continuous. Since $\text{im}(Y)$ is a closed subset of the Banach space $Y$, it is a Banach space and we apply the open mapping theorem in order to conclude that $\tilde{T}$ is an isomorphism. \qed

Note that if $T : X \to Y$ is a continuous linear mapping and $\text{im}(T)$ is not closed, then $\text{im}(T)$ is not a Banach space and hence cannot be isomorphic to $X/\ker(T)$, which is a Banach space since $\ker(T)$ is closed if and only if $T$ is continuous.

2.2. Compact Operators

Suppose that $X$ is a topological space, and that $V$ is a subset of $X$. Then $V$ is compact if every covering of it by open sets admits a finite subcover. When $X$ is a Banach space, there are a number of other useful ways to characterize compact sets:

Theorem 7. Suppose that $X$ is a Banach space, and that $V$ is a subset of $X$. Then the following are equivalent:

1. The set $V$ is compact (i.e., every covering of $V$ by open sets admits a finite subcovering).

2. Every sequence contained in $V$ has a convergent subsequence whose limit is in $V$. 

2.2. Compact Operators

Suppose that $X$ is a topological space, and that $V$ is a subset of $X$. Then $V$ is compact if every covering of it by open sets admits a finite subcover. When $X$ is a Banach space, there are a number of other useful ways to characterize compact sets:

Theorem 7. Suppose that $X$ is a Banach space, and that $V$ is a subset of $X$. Then the following are equivalent:

1. The set $V$ is compact (i.e., every covering of $V$ by open sets admits a finite subcovering).

2. Every sequence contained in $V$ has a convergent subsequence whose limit is in $V$. 

2.2. COMPACT OPERATORS

(3) The set \( V \) is closed and for every \( \epsilon > 0 \) there exists a finite collection of points \( x_1, \ldots, x_n \in V \) such that

\[
V \subseteq \bigcup_{i=1}^{n} B_{\epsilon}(x_i).
\]

We say that a set \( V \) of a Banach space \( X \) is totally bounded if for every \( \epsilon > 0 \) there exists a finite collection of points \( x_1, \ldots, x_n \in V \) such that

\[
V \subseteq \bigcup_{i=1}^{n} B_{\epsilon}(x_i)
\]

so that the third criterion for compactness in Theorem 7 can be summarized by saying that \( V \) is compact if and only if it is closed and totally bounded.

The following result, due to F. Riesz, will be used frequently in the remainder of this section.

**Theorem 8 (Riesz’ lemma).** Suppose that \( X \) is a Banach space, that \( Y \) is a closed proper subspace of \( X \), and that \( 0 < \alpha < 1 \) is a real number. Then there exists \( x \in X \setminus Y \) such that

\[
\|x\| = 1
\]

and

\[
\inf_{y \in Y} \|x - y\| \geq \alpha.
\]

**Proof.** We choose \( x_1 \) in \( X \setminus Y \) and let

\[
r = \inf_{y \in Y} \|x_1 - y\|.
\]

Since \( Y \) is closed, \( r > 0 \). Suppose that \( \epsilon > 0 \). Then there exists \( y_1 \in Y \) such that

\[
r \leq \|x_1 - y_1\| < r + \epsilon.
\]

We set

\[
x = \frac{x_1 - y_1}{\|y_1 - x_1\|}
\]

so that

\[
\|x\| = 1
\]

and

\[
\inf_{y \in Y} \|y - x\| = \inf_{y \in Y} \left\| y - \frac{x_1}{\|x_1 - y_1\|} y_1 \right\|.
\]

Since

\[
\frac{y_1}{\|x_1 - y_1\|}
\]

is in \( Y \), we see from (14) that

\[
\inf_{y \in Y} \|x - y\| = \inf_{y \in Y} \left\| y - \frac{x_1}{\|x_1 - y_1\|} \right\| = \frac{r}{r + \epsilon}.
\]

Since \( r/(r + \epsilon) \) increases to 1 as \( \epsilon \to 0 \), it follows from (14) that

\[
\frac{y_1}{\|x_1 - y_1\|}
\]

is in \( Y \), we see from (14) that

\[
\inf_{y \in Y} \|x - y\| = \inf_{y \in Y} \left\| y - \frac{x_1}{\|x_1 - y_1\|} \right\| = \frac{r}{r + \epsilon}.
\]

Since \( r/(r + \epsilon) \) increases to 1 as \( \epsilon \to 0 \), it follows from (14) that

\[
\frac{y_1}{\|x_1 - y_1\|}
\]

is in \( Y \), we see from (14) that

\[
\inf_{y \in Y} \|x - y\| = \inf_{y \in Y} \left\| y - \frac{x_1}{\|x_1 - y_1\|} \right\| = \frac{r}{r + \epsilon}.
\]

Since \( r/(r + \epsilon) \) increases to 1 as \( \epsilon \to 0 \), it follows from (14) that
It follows immediately from Riesz’ lemma that if $Y$ is an infinite-dimensional subspace of a Banach space $X$, then there exists a sequence $\{y_j\}$ in $Y$ such that

$$\|y_i - y_j\| \geq \frac{1}{2}$$

for all positive integers $i$ and $j$ such that $i \neq j$. Sequences of this type are often used as a substitute for orthonormal bases in Hilbert spaces, as in the proof of the following theorem.

**Theorem 9.** The Banach space $X$ is finite-dimensional if and only if the closed unit ball in $X$ is compact.

**Proof.** If $X$ is finite-dimensional, it is isomorphic to $\mathbb{R}^n$ for some positive integer $n$. In this case, the closed unit ball of $X$ is identified with a closed, bounded subset of $\mathbb{R}^n$, and so it is compact.

Suppose now that $X$ is infinite-dimensional and that its closed unit ball is compact. We apply Riesz’ lemma in order to construct a sequence $\{x_j\}$ such that

$$\|x_j\| = 1$$

for all $j = 1, 2, \ldots$ and

$$\|x_i - x_j\| \geq \frac{1}{2}$$

whenever $i$ and $j$ are positive integers such that $i \neq j$. Since $\{x_j\}$ is contained is the closed unit ball of $X$, which we have assumed to be compact, it has a convergent subsequence. But this conclusion is contradicted by (19), which implies that no subsequence of $\{x_n\}$ can be Cauchy. We conclude that $X$ is finite-dimensional. □

We say that an operator $K : X \to Y$ between Banach spaces is compact if $\overline{\text{im}(K)}$ is compact. It follows from Theorem 9 that $K$ is compact if and only if whenever $\{x_n\}$ is a bounded sequence in $X$, $\{K(x_n)\}$ has a convergent subsequence.

**Exercise 1.** Suppose that $X$ and $Y$ are Banach spaces, and that $K : X \to Y$ is compact. Show that $K$ is bounded.

We denote by $T^*$ the adjoint of $T$, which is the bounded linear operator $Y^* \to X^*$ defined via the requirement that

$$\langle Tx, \varphi \rangle = \langle x, T^* \varphi \rangle$$

for all $x \in X$ and $\varphi \in Y^*$. By

$$\langle x, \varphi \rangle,$$

where $\varphi \in X^*$ and $x \in X$, we mean the value obtained by evaluating the linear functional $\varphi$ at the point $x$ — that is, $\varphi(x)$. We omit the proof of the following theorem, which can be found in most any functional analysis textbook.

**Theorem 10.** Suppose that $X$ and $Y$ are Banach spaces, and that $K : X \to Y$ is a compact operator. Then the adjoint $K^* : Y^* \to X^*$ is also compact.
2.2. COMPACT OPERATORS

We now use Riesz’s lemma to establish a basic result concerning the spectrum $\sigma(K)$ of a compact operator $K : X \to Y$ between Banach spaces. Recall that $\sigma(K)$ is the set of $\lambda \in \mathbb{R}$ such that the operator $\lambda I - K$ does not admit a continuous inverse.

**Theorem 11.** Suppose that $X$ and $Y$ are Banach spaces, and that $K : X \to Y$ is a compact operator. Then the spectrum $\sigma(K)$ of $K$ is either finite or a sequence which converges to 0. Moreover, for each $\lambda \in \sigma(K)$, kernel of $\lambda I - K$ is finite-dimensional and the image of $\lambda I - K$ is closed.

**Proof.** We will show that for any $\epsilon > 0$

$$Y = \{x : Kx = \lambda x \text{ with } |\lambda| \geq \epsilon\}$$

is finite-dimensional. Suppose that $x_1, x_2, \ldots$ are elements of $Y$ such that

$$\lim_{j \to \infty} x_j = x.$$  \hfill (23)

Then there exist real numbers $\lambda_1, \lambda_2, \ldots$ such that

$$\epsilon \leq |\lambda_j| \leq \|T\|$$  \hfill (24)

and

$$K(x_j) = \lambda_j x_j.$$  \hfill (25)

for all positive integers $j$. Since the $\lambda_j$ are contained in a compact subset of $\mathbb{R}$, there is a subsequence of $\{\lambda_j\}$ which converges to a real number $\lambda$ such that $\epsilon \leq |\lambda| \leq \|K\|$. By passing to a subsequence, we obtain

$$\lambda_j \to \lambda.$$  \hfill (26)

By combining (23), (25) and (26) we see that

$$\lim_{j \to \infty} K(x_j) = \lim_{j \to \infty} \lambda_j x_j = \lambda x.$$  \hfill (27)

Since $K$ is continuous,

$$\lim_{j \to \infty} K(x_j) = x.$$  \hfill (28)

We see from (27) and (28) that $K(x) = \lambda x$, which implies that $x \in Y$. We conclude that $Y$ is closed.

We will now suppose that $Y$ is infinite-dimensional and derive a contradiction. We apply Riesz’ lemma (which is applicable since $Y$ is closed) in order to construct $y_1, y_2, \ldots$ in $Y$ such that

$$\|y_j\| = 1$$  \hfill (29)

for all positive integers $j$ and

$$\|y_i - y_j\| \leq \frac{1}{2}$$  \hfill (30)

for all positive integers $i$ and $j$ such that $i \neq j$. By the definition of $Y$, there exist $\lambda_1, \lambda_2, \ldots$ in $\mathbb{R}$ such that

$$\epsilon \leq |\eta_j| \leq \|T\|$$  \hfill (31)
and
\[ Ky_j = \eta_j y_j \]
for all positive integers \( j \). We observe that
\[ \| \eta_i^{-1}K(y_i) - \eta_j^{-1}K(y_j) \| = \| y_i - y_j \| \geq \frac{1}{2}, \]
(33)
for all positive integer \( i \) and \( j \) such that \( i \neq j \). By manipulating (31), we see that
\[ \frac{1}{2} \leq |\eta_i^{-1} - \eta_j^{-1}| \| K(y_i) \| + |\eta_j^{-1}| \| K(y_i) - K(y_j) \| \]
for all positive integers \( i,j \) such that \( i \neq j \). Since \( K \) is compact and \( \{ \eta_j \} \) is contained in a compact subset of \( \mathbb{R} \), there are subsequence of \( \{ K(x_j) \} \) and \( \{ \eta_j \} \) which are convergent and hence Cauchy. But this contradicts (34). We conclude that \( Y \) is finite-dimensional.

We now suppose that for some \( \lambda \in \sigma(K) \), the kernel of \( \lambda I - K \) is infinite-dimensional and we will derive a contradiction. The kernel of any bounded linear mapping is closed, so we apply Riesz' lemma in order to obtain \( z_1, z_2, \ldots \) such that
\[ \| z_j \| = 1 \]
(35)
for all positive integers \( j \),
\[ \| z_i - z_j \| \geq \frac{1}{2} \]
(36)
for all pairs of positive integers \( i,j \) such that \( i \neq j \), and
\[ \lambda_j z_j = K(z_j) \]
(37)
for all positive integers \( j \). We observe that
\[ \| K(z_j) - K(z_i) \| = |\lambda| \| z_i - z_j \| \geq \frac{|\lambda|}{2}. \]
(38)
But (38) implies that no subsequence of \( \{ K(z_j) \} \) is Cauchy, which contradicts the the assumption that \( K \) is compact.

We suppose that \( y \) is the limit of a sequence in the image of \( \lambda I - K \). That is, we suppose that there exists a sequence \( \{ x_n \} \) such that
\[ \lim_{n \to \infty} (\lambda x_n - K x_n) = y. \]
(39)
Since \( K \) is compact, by passing to a sequence we can assume that \( K x_n \) is convergent. It is clear, then, from (39) that the sequence \( x_n \) converges to some \( x \in X \). The continuity of the operator \( \lambda I - K \) implies that
\[ y = \lim_{n \to \infty} (\lambda I - K) x_n = (\lambda I - K) \left( \lim_{n \to \infty} x_n \right) = (\lambda I - K) x, \]
from which we conclude that the image of \( T \) is closed.

Suppose that \( X \) and \( Y \) are Banach spaces. It is easy to see that the set of compact operators \( X \to Y \) is a closed subspace of the set of linear operators \( X \to Y \). In particular, if \( \{ K_n \} \) is a sequence of compact operators which converges to \( K \) in operator norm, then \( K \) is compact. Moreover, every operator of finite rank is compact, so that any operator which is the limit of finite rank operators is compact. We say that the Banach space \( Y \) has the approximation
property if the converse is true — that is, if every compact operator is the limit of a sequence of finite rank operators.

Not every Banach space has the approximation property \( \text{[2]} \); however, the following theorem gives a useful sufficient condition for a Banach space to have the approximation property. Before we state it, we require a further definition. A sequence \( \{x_n\} \) in a Banach space \( X \) is a Schauder basis for \( X \) if for every \( x \in X \) there exist real numbers \( \alpha_1, \alpha_2, \ldots \) such that

\[
\lim_{n \to \infty} \left\| \sum_{j=1}^{n} \alpha_j x_j - x \right\| = 0. \tag{41}
\]

**Theorem 12.** Suppose that \( X \) and \( Y \) are Banach spaces, and that \( Y \) has a Schauder basis. Then every compact operator \( X \to Y \) is the limit of a sequence of finite rank operators \( X \to Y \).

Clearly, every separable Hilbert space has a Schauder basis, as does the space \( L^p(\mathbb{R}^n) \) when \( 1 \leq p < \infty \), and also the space \( C(X) \) of continuous functions on a compact metric space. On the other hand, the space \( L^p(\mathbb{R}^n) \) does not admit a Schauder basis since it is not separable. Note, though, that not every separable Banach space has a Schauder basis (indeed, in \( \text{[2]} \), a separable Banach space which does not have a Schauder basis and does not have the approximation property is constructed).

**Exercise 2.** Suppose that \((\Omega, \mu)\) is a measure space, that \( k(x, y) \) is an element of \( L^2(\Omega \times \Omega) \), and that \( T : L^2(\Omega) \to L^2(\Omega) \) is the linear operator defined via the formula

\[
T[f](x) = \int_{\Omega} k(x, y)f(y) \, dy. \tag{42}
\]

Show that \( T \) is a compact operator. Operators of this type are called Hilbert-Schmidt operators and the function \( k \) is referred to as the kernel of \( T \).

### 2.3. The Lax-Milgram Theorems

Suppose that \( X \) is a reflexive Banach space, and that \( X^* \) is its dual. We say that a bounded linear mapping \( L : X \to X^* \) is coercive if there exists \( \lambda > 0 \) such that

\[
|L[x](x)| \geq \lambda \|x\|^2 \tag{43}
\]

for all \( x \in X \).

**Theorem 13 (Lax-Milgram).** Suppose that \( X \) is a reflexive Banach space, that \( X^* \) is its dual space, and that \( L : X \to X^* \) is a bounded linear map. If \( L \) is coercive, then it is an isomorphism (that is, it is invertible, and its inverse is also continuous).

**Proof.** We use \( \langle f, x \rangle \) to denote the duality pairing of \( X^* \) with \( X \) — in particular the value of the linear function \( L[x] \) at the point \( y \in X \) is \( \langle L[x], y \rangle \). We observe that (43) implies

\[
\lambda \|x\|^2 \leq |\langle L[x], x \rangle| \leq \|L[x]\| \|x\| \tag{44}
\]

for all \( x \in X \). Dividing both sides of (44) by \( \|x\| \) yields

\[
\lambda \|x\| \leq \|L[x]\|. \tag{45}
\]
The identity (45) implies that $L$ is injective. It also implies that the range of $L$ is closed. To see that, we suppose that $L[x_n] \to y$. Then $\{L[x_n]\}$ is a Cauchy sequence and we see from (45) that $\{x_n\}$ is Cauchy as well. We denote by $z$ the limit of $\{x_n\}$. The continuity of $L$ gives us $L[x] = \lim_n L[x_n] = y$, from which we conclude that $L$ has closed range. So $L$ is a continuous bijective mapping from $X$ to $\text{im}(L)$. Since $\text{im}(L)$ is a closed subset of $X^*$, it is a Banach space and we may apply the open mapping theorem. By doing so, we see that $L : X \to \text{im}(L)$ is an isomorphism.

We suppose now that $\text{im}(L) \neq X^*$. Then, by the Hahn-Banach theorem, there exists $\phi$ in $(X^*)^*$ such that

$$\langle \phi, f \rangle = 1$$

and

$$\phi|_{\text{im}(L)} = 0.$$  \hfill (47)

Since $V$ is reflexive, we may identify $\phi \in (X^*)^* = X$ with an element $u \in X$ such that

$$f(u) = 1$$

and

$$\langle L[u], u \rangle = 0$$

for all $x \in X$. We combine (46) and (47) with the assumption that $L$ is coercive in order to conclude that

$$\lambda \|u\|^2 \leq \langle L[u], u \rangle = 0,$$  \hfill (50)

from which we see that $u = 0$. However, this contradicts (48). We conclude that $\text{im}(L) = V^*$.

Clearly, we can identify the continuous linear mapping $L : X \to X^*$ with the bilinear form $B : X \times X \to \mathbb{R}$ defined via

$$B[x, y] = L[x](y).$$  \hfill (51)

We say that the bilinear form $B$ is bounded if there exists $C > 0$ such that

$$|B[x, y]| \leq C \|x\| \|y\|$$

for all $x, y \in X$. Obviously, $B$ is bounded if and only if $L$ is bounded. We say that $B$ is coercive if

$$|B[x, x]| \geq \lambda \|x\|^2$$

for all $x \in X$. The Lax-Milgram theorem can be rephrased as follows.

**Theorem 14.** If $X$ is a reflexive Banach space $B$ is a bounded, coercive bilinear form $X \times X \to \mathbb{R}$, then for each $f \in X^*$, there exists a unique $u$ such that

$$B[u, v] = f(v)$$

for all $v \in X$.

Since Hilbert spaces are reflexive and inner products are coercive bilinear forms, the Lax-Milgram theorem implies the Riesz representation theorem (with which you are hopefully familiar).
Theorem 15 (Riesz representation theorem). Suppose that $X$ is a Hilbert space, and that $f : X \to \mathbb{R}$ is a bounded linear functional on $X$. Then there exists a unique $u \in X$ such that $\|u\| = \|f\|$ and

$$f(x) = (u, x)$$

for all $x \in X$

2.4. Fredholm Operators

Suppose that $A$ is a real-valued $n \times n$ matrix, and that $A^*$ is its transpose. Then $\mathbb{R}^n$ is the orthogonal direct sum of the image of $A$ and the kernel of $A^*$, as well as the orthogonal direct sum of the image of $A^*$ and the kernel of $A$, and the dimension of $\text{im}(A)$ is equal to the dimension of $\text{im}(A^*)$. These elementary observations have a number of useful consequences. Among them, that the equation

$$Ax = b$$

is uniquely solvable for each $b \in \mathbb{R}^n$ if and only if

$$A^*z = 0$$

admits only the trivial solution, which is the case if and only if the equation

$$Ax = 0$$

admits only the trivial solution. In other words, the linear operator corresponding to the matrix $A$ is injective if and only if it is surjective.

In this section, we discuss a class of operators which generalize this rather useful property of square matrices. Before we give the principal definition, we review some of the basic properties of direct sum decompositions of Banach spaces.

2.4.1. Direct Sums and Complemented Subspaces. If $Y$ and $Z$ are Banach spaces, then the direct sum $Y \oplus Z$ is the Banach space obtained by endowing the vector space $Y \times Z$ with the norm

$$\|(y, z)\| = \|y\| + \|z\|.$$  \hspace{1cm} (59)

We say that a subspace $Y$ of a Banach space $X$ is complemented in $X$ if there exists a subspace $Z$ of $X$ such that the addition map $A : Y \oplus Z \to X$ defined via the formula

$$A(y, x) = y + z$$

is an isomorphism (meaning that it is a continuous bijective linear mapping whose inverse is also continuous). If $Y$ is complemented in $X$, then $Y$ is necessarily closed since the composition of the inverse of $A$ with the projection

$$P : Y \oplus Z \to Z$$

defined via $P(y, z) = z$ is a continuous linear mapping whose kernel is $Y$ (the kernel of a linear mapping is closed if and only if the mapping is continuous).

Any closed subspace of a Hilbert space is complemented: if $M$ is a closed subspace of a Hilbert space $X$, then $X = M \oplus M^\perp$, where $M^\perp$ denotes the orthogonal complement of
2.4. FREDHOLM OPERATORS

the space $X$. The same is not true of Banach spaces. In fact, if $X$ is a Banach space and every closed subspace of $X$ is complemented, then $X$ is isomorphic to a Hilbert space [7]. In general, it is difficult to determine whether or not a particular closed subspace $Y$ of a Banach space $X$ is complemented. However, as we will now show, subspaces of finite dimension and closed subspaces of finite codimension are complemented (the codimension of $Y$ in $X$ is the dimension of the quotient space $X/Y$).

Suppose that $X$ is a Banach space. A linear mapping $P : X \to X$ such that $P^2 = P$ is called a projection. The following theorem characterizes complemented subspaces as the kernels and images of continuous projections.

**Theorem 16.** Suppose that $X$ is a Banach space, and that $Y$ is a subspace of $X$. Then the following are equivalent:

1. The subspace $Y$ is complemented in $X$.
2. There is a continuous projection $P : X \to X$ such that $\ker(P) = Y$.
3. There is a continuous projection $P : X \to X$ such that $\text{im}(P) = Y$.

**Proof.** First, we show that (1) implies (2). To that end, we suppose that $Y$ is complemented in $X$ so that there exists a closed subspace $Z$ of $X$ such that the addition map

$$A : Y \oplus Z \to X$$

defined via the formula

$$A(y, z) = y + z$$

is a linear isomorphism. We observe that the composition $P$ of $A^{-1} : X \to Y \oplus Z$ with the mapping $Y \oplus Z \to X$ which takes map $(y, z)$ to $z$ is a continuous projection $X \to X$ whose kernel is $Y$.

To see that (2) implies (3), we observe that if $P : X \to X$ is a continuous projection such that $\ker(P) = Y$, then $I - P$ is a continuous projection such that $\text{im}(P) = Y$.

We now conclude the proof by showing that (3) implies (1). We suppose that $Y$ is the image of a continuous projection $P : X \to X$, and that $Z$ is the kernel of $P$. We will show that the addition map $A : Y \oplus Z \to X$ defined by the formula

$$A(y, z) = y + z$$

is an isomorphism. The map $A$ is plainly continuous since

$$\|y + z\| \leq \|y\| + \|z\|.$$  

It is surjective since

$$x = Px + (I - P)x$$

with $Px \in Y$ and $(I - P)x \in Z$ whenever $x \in X$. Suppose that

$$y_1 + z_1 = y_2 + z_2$$

where $y_1, y_2 \in Y$ and $z_1, z_2 \in Z$. Since $Y$ is the image of $P$, there exist $x_1, x_2$ such that

$$Px_1 = y_1$$

and

$$Px_2 = y_2.$$
We combine (66), (67) and (68) in order to conclude that
\[ P(x_1 - x_2) + (z_1 - z_2) = 0. \] (69)
By applying \( P \) to both sides of (69) and make use of the facts that \( P^2 = P \) and \( Z = \ker(P) \) we obtain
\[ P(x_1 - x_2) = 0, \] (70)
from which we conclude that \( y_1 = y_2 \). It follows from this and (66) that \( z_1 = z_2 \). We conclude that \( A \) is also injective. We now apply the open mapping theorem in order to see that the bijective continuous linear mapping \( A \) is an isomorphism. □

**Theorem 17.** Suppose that \( Y \) is a finite-dimensional subspace of the Banach space \( X \). Then there exists a closed subspace \( Z \) such that \( X = Y \oplus Z \).

**Proof.** We let \( \{v_1, v_2, \ldots, v_n\} \) be a basis for the subspace \( Y \). For each \( j = 1, \ldots, n \) we define the bounded linear function \( \varphi_j : Y \to \mathbb{R} \) via the formula
\[ \varphi_j(v_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise}. \end{cases} \] (71)
Now we apply the Hahn-Banach theorem in order to extend each of the \( \varphi_j \) to mappings \( Y \to \mathbb{R} \). We also let
\[ Z = \bigcap_{j=1}^{n} \ker(\varphi_j), \] (72)
and define the mapping \( P : X \to X \) via
\[ P(x) = \sum_{j=1}^{n} \varphi_j(x)v_j. \] (73)
We observe that \( P \) is a continuous projection whose kernel is \( Y \). It follows from Theorem 16 that \( Y \) is complemented. □

**Theorem 18.** Suppose that \( X \) is a Banach space, and that \( Y \) is a closed subspace of \( X \) of finite codimension \( n \). Then there exists an \( n \)-dimensional subspace \( Z \) of \( X \) such that \( X = Y \oplus Z \).

**Proof.** We let \( \{v_1, \ldots, v_n\} \) be a basis for the space \( X/Y \). For each \( j = 1, \ldots, n \), we choose an element \( w_j \) of \( X \) which is in the equivalence class \( v_j \), and we denote by \( Z \) be the subspace of \( X \) spanned by \( w_1, \ldots, w_n \). The kernel of the canonical projection \( \varphi : X \to X/Y \) is \( Y \), which is a closed subspace, so \( \varphi \) is continuous. As is the restriction \( \tilde{\varphi} : Z \to X/Y \) of \( \varphi \) to the finite dimensional (and hence closed) subspace \( Z \) of \( Y \). Moreover, \( \tilde{\varphi} \) is clearly bijective. It follows from the open mapping theorem that \( \tilde{\varphi}^{-1} : X/Y \to Z \) is continuous. We observe that the composition \( P = \tilde{\varphi}^{-1} \circ \psi \) is a continuous projection \( X \to X \) whose kernel is \( Y \). We conclude form this observation and Theorem 16 that \( Y \) is complemented in \( X \). □

Note that the requirement that \( Y \) be closed in Theorem 18 is essential since there are subspaces of finite codimension which are not closed, and hence cannot be complemented. Indeed, the kernel of any discontinuous linear functional \( T : X \to \mathbb{R} \) is a subspace of \( X \) of codimension 1 which is not closed. There is no need for such a requirement in Theorem 17.
since all finite-dimensional subspaces are necessarily closed. We now show that if a subspace of finite codimension is the image of a continuous linear operator, then it must be closed.

**Theorem 19.** Suppose that $X$ and $Y$ are Banach spaces, and that $T : X \to Y$ is a continuous linear mapping. If $\text{im}(T)$ is of finite codimension in $Y$, then it is closed (and hence complemented).

**Proof.** Without loss of generality, we assume that $T$ is injective (if not, then we may replace $T$ with the injective mapping $X/\ker(T) \to Y$ it induces). We choose a basis $\{v_1, \ldots, v_n\}$ for the space $Y/\text{im}(T)$ and for each $j = 1, \ldots, n$, we choose a representative $w_i$ in $Y$ of the equivalence class $v_i$. We denote by $Z$ the subspace of $Y$ spanned by $w_1, \ldots, w_n$. Since $Z$ is finite-dimensional (and hence closed), $Z$ is a Banach space. Now we define the mapping $A : X \oplus Z \to Y$ by the formula

$$A(x, z) = Tx + z. \quad (74)$$

Since

$$\|Tx + z\| \leq \|Tx\| + \|z\| \leq \|T\| \|x\| + \|z\| \leq (1 + \|T\|)(\|x\| + \|z\|), \quad (75)$$

$A$ is bounded. Suppose that $y \in Y$. Then there exist $\alpha_1, \ldots, \alpha_n$ such that

$$\sum_{j=1}^{n} \alpha_j v_j \quad (76)$$

is the equivalence class in $Y/\text{im}(T)$ containing $y$. Since $w_j$ is an element of the equivalence class containing $v_j$, we have

$$y = \sum_{j=1}^{n} \alpha_j w_j + \sum_{j=1}^{n} \alpha_j u_j + u \quad (77)$$

where $u_1, \ldots, u_n$ and $u$ are contained in $\text{im}(T)$. We conclude that $A$ is surjective. Now suppose that

$$T(x_1 - x_2) + z_1 - z_2 = 0. \quad (78)$$

The restriction of the canonical mapping $Y \to Y/\text{im}(Z)$ to $Z$ is clearly injective, and by applying that mapping to (78) we see that $z_1 = z_2$. It follows from this fact and (78) that

$$T(x_1 - x_2) = 0. \quad (79)$$

Since $T$ is injective, (79) implies that $x_1 = x_2$. We conclude that $A$ is injective as well as surjective and continuous. We now apply the open mapping theorem in order to conclude that it is an isomorphism. Consequently, $A$ carries closed subsets of $X \oplus Z$ to closed subsets of $Z$. The image of the closed subset $X \times \{0\}$ of $X \oplus Z$ under $A$ is $\text{im}(T)$. We conclude that $\text{im}(T)$ is closed. \(\square\)

**2.4.2. Annihilators and Preannihilators.** If $M$ is a subset of a Banach space $X$, then the annihilator $M^\perp$ of $M$ is the closed subspace of $X^*$ defined via

$$M^\perp = \{ f \in X^* : f(x) = 0 \text{ for all } x \in M \}. \quad (80)$$

Similarly, if $N$ is a subset of $X^*$, then the preannihilator $N_\perp$ of $N$ is the closed subspace of $X$ defined as follows:

$$N_\perp = \{ x \in X : f(x) = 0 \text{ for all } f \in N \}. \quad (81)$$
2.4. FREDHOLM OPERATORS

If $A$ is an $n \times m$ matrix, then the kernel of $A$ is the orthogonal complement of the image of $A^*$ and the kernel of $A^*$ is the orthogonal complement of the image of $A$. The following theorem generalizes these observations to the case of bounded linear mapping between Banach spaces.

**Theorem 20.** Suppose that $T : X \to Y$ is a continuous linear map between Banach spaces, and that $T^* : Y^* \to X^*$ is its adjoint. Then

1. $\text{im}(T)^\perp = \ker(T^*)$
2. $\overline{\text{im}(T)} = \ker(T^*)^\perp$
3. $\ker(T) = \text{im}(T^*)^\perp$
4. $\overline{\text{im}(T^*)} \subset \ker(T)^\perp$

From Theorem 16, we see that if the image of $T$ is closed, then the image of $T$ is the preannihilator of the kernel of the adjoint $T^*$. This gives us a solvability criterion for the equation

$$Tx = y. \quad (82)$$

In particular, if the image of $T$ is closed, then (82) admits a solution if and only if

$$\phi(y) = 0 \quad (83)$$

for all $\phi \in \ker(T^*)$.

### 2.4.3. Fredholm Operators

Suppose that $X$ and $Y$ are Banach spaces. We say that a continuous linear mapping $T : X \to Y$ is a Fredholm operator if the kernel of $T$ is of finite dimension and the image of $T$ is of finite codimension. The index of a Fredholm operator $T$ is defined to be

$$\text{ind}(T) = \dim(\ker(T)) - \dim(Y/\text{im}(T)). \quad (84)$$

We will now provide an alternative definition of Fredholm operator.

**Theorem 21.** Suppose that $X$ and $Y$ are Banach spaces, that $T : X \to Y$ is a continuous linear mapping, and that $T^* : Y^* \to X^*$ is its adjoint. Suppose further that $\text{im}(T)$ is closed. Then the dual space of $Y/\text{im}(T)$ is isomorphic to $\ker(T^*)$.

**Proof.** We observe that if $\varphi : Y \to \mathbb{R}$ is in $\ker(T^*)$ then

$$\langle Tx, \varphi \rangle = \langle x, T^*(\varphi) \rangle = 0$$

for all $x \in X$ — that is, $\varphi(\text{im}(T)) = 0$. It follows that the map $\tilde{\varphi} : (Y/\text{im}(T)) \to \mathbb{R}$ defined via

$$\tilde{\varphi}(y + \text{im}(T)) = \varphi(y) \quad (86)$$

is a well-defined linear functional in the dual of $Y/\text{im}(T)$. We denote by $\Lambda$ the map which takes $\varphi \in \ker(T^*)$ to the linear functional $\tilde{\varphi}$ defined via (86). Since $\varphi(\text{im}(T)) = 0$,

$$\|\tilde{\varphi}\| = \sup_{|y + \text{im}(T)| = 1} |\tilde{\varphi}(y)| \leq \sup_{|y| = 1} |\varphi(y)| = \|\varphi\|, \quad (87)$$
from which we see that $\Lambda$ is a bounded mapping. We also observe that $\Lambda$ is bijective; indeed, its inverse is the map taking
$$\psi : Y / \text{im}(T) \to \mathbb{R}$$
(88)
to the map $\tilde{\psi} : Y \to \mathbb{R}$ defined via
$$\tilde{\psi}(y) = \psi(y + \text{im}(T)).$$
(89)
Since $\text{im}(T)$ is closed, $Y / \text{im}(T)$ is a Banach space and its dual space is a Banach space. Consequently, the open mapping theorem applies and we invoke it in order to conclude that $\Lambda$ is an isomorphism. □

**Theorem 22.** Suppose that $X$ and $Y$ are Banach spaces, that $T : X \to Y$ is a continuous linear mapping, and that $T^* : Y^* \to X^*$ is its adjoint. Then $T$ is Fredholm if and only if its image is closed and both $\ker(T)$ and $\ker(T^*)$ are of finite dimension. Moreover, if $T$ is Fredholm then
$$\text{ind}(T) = \dim(\ker(T)) - \dim(\ker(T^*)).$$
(90)

**Proof.** We suppose first that $T$ is a Fredholm operator. Then the kernel of $T$ is finite-dimensional by definition and the image of $T$ is closed since it is of finite codimension in $Y$ (see Theorem 19). According to Theorem 21, that $\text{im}(T)$ is closed implies that $(Y / \text{im}(T))^*$ is isomorphic to $\ker(T^*)$. We conclude that
$$\dim(\ker(T^*)) = \dim((Y / \text{im}(T))^*).$$
(91)
But $Y / \text{im}(T)$ is finite-dimensional, so
$$\dim(Y / \text{im}(T)) = \dim((Y / \text{im}(T))^*).$$
(92)
We combine (91) and (92) in order to obtain
$$\dim(\ker(T^*)) = \dim(Y / \text{im}(T)) < \infty,$$
(93)
which suffices to establish (90) and the assertion that $\ker(T^*)$ is finite-dimensional.

Now we suppose that $\ker(T)$, $\ker(T^*)$ are finite-dimensional and that $\text{im}(T)$ is closed. By Theorem 21,
$$(Y / \text{im}(T))^* \sim \ker(T^*).$$
(94)
We see from (94) that the codimension of $\text{im}(T)$ is equal to the (finite) dimension of $\ker(T^*)$. Since we have assumed that the dimension of the kernel of $T$ is finite, we that $T$ is Fredholm. □

**Exercise 3.** Suppose that $X$ is a Hilbert space, that $\{\phi_j\}_{j=1}^\infty$ is an orthonormal basis for $X$, and that $T$ is the linear mapping $X \to X$ defined via the formula
$$T[\phi_i] = \phi_{i+1} \quad \text{for all } i = 1, 2, \ldots$$
(95)
Suppose also that $k$ is a positive integer. What is the kernel of $T^k$? What is the cokernel of $T^k$? What is the index of $T^k$?

**Theorem 23.** Suppose that $X$ is a Banach spaces, and that $T : X \to X$ is compact. Then $I + T$ is a Fredholm operator.
Proof. From Theorem 11, we see that the kernel of $I + T$ is finite-dimensional, and that its image is closed. The adjoint of $I + T$ is

$$I + T^*,$$

where $T^*: Y^* \to X^*$ is the adjoint of $T$. Since $T^*$ is also compact (by Theorem 10), we see from Theorem 11 that the kernel of $I + T^*$ is finite-dimensional. We now apply Theorem 22 in order to conclude that $T$ is Fredholm. □

We conclude from Theorems 17, 18 and 19 that a Fredholm operator induces the direct sum decompositions

$$X = \ker(T) \oplus X'$$

and

$$Y = \text{im}(T) \oplus Y',$$

where $\ker(T)$ is a finite-dimensional subspace of $X$ and $Y'$ is a finite-dimensional subspace of $Y$. This direct sum decomposition is crucial in the proof of the next theorem, which characterizes Fredholm operators as those which are invertible “modulo compact operators.”

**Theorem 24.** An operator $T: X \to Y$ between Banach spaces is Fredholm if and only if there exist a bounded linear operator $S: Y \to X$ and a pair of compact operators $K_1: Y \to Y$, $K_2: X \to X$ such that

$$ST = I - K_1$$

and

$$TS = I - K_2.$$ 

Proof. We first suppose that there exist a bounded linear operator $S: Y \to X$ and compact operators $K_1: Y \to Y$, $K_2: X \to X$ such that (99) and (100) hold. By Theorem 22, $I - K_1$ and $I - K_2$ are Fredholm, so $\dim(\ker(I - K_1))$ and $\dim(Y/\text{im}(TS))$ are finite-dimensional. We observe that

$$\ker(T) \subseteq \ker(ST) = \ker(I - K_1)$$

and that

$$Y/\text{im}(T) \subset Y/\text{im}(TS)$$

since $\text{im}(TS) \subset \text{im}(T)$. We conclude that $\ker(T)$ and $Y/\text{im}(T)$ are finite-dimensional so that $T$ is Fredholm.

Now we suppose that $T$ is Fredholm. Then there exists a closed subspace $X'$ of $X$ such that

$$X = X' \oplus \ker(T)$$

and a finite dimensional subspace $Y'$ of $Y$ such that

$$Y = Y' \oplus \text{im}(T).$$

The restriction of $T$ to $X'$ is a continuous bijective linear mapping $X' \to \text{im}(T)$. We let $\tilde{S}: \text{im}(T) \to X'$ denote the inverse of this mapping, which is continuous by the open mapping theorem. We extend $\tilde{S}$ to a bounded linear mapping $S: Y \to X$ such that $S(Y') = 0$ by
linearity. If we let $P$ be the projection $X \to \ker(T)$ and let $Q$ be the projection $Y \to Y'$, then

$$ST = ST(P + I - P) = \tilde{S}T(I - P) = I - P$$  \hspace{1cm} (105)

and

$$TS = TS(Q + I - Q) = T\tilde{S}(I - Q) = I - Q.$$  \hspace{1cm} (106)

The projections $P$ and $Q$ are compact since $\ker(T)$ and $Y'$ are finite dimensional. □

In fact, it is clear from the proof of Theorem 24 that a continuous linear mapping $T : X \to Y$ is Fredholm if and only if there exists a continuous linear mapping $S : X \to Y$ and finite rank operators $K_1 : X \to X$ and $K_2 : Y \to Y$ such that

$$ST = I - K_1$$  \hspace{1cm} (107)

and

$$TS = I - K_2.$$  \hspace{1cm} (108)

We call any bounded linear operator $S$ for which there exists compact operators $K_1$ and $K_2$ such that (99) and (100) holds a parametrix for $T$. We note that that the relationship is symmetric: if $S$ is a parametrix for $T$ then $S$ is Fredholm and $T$ is a parametrix of $S$. Moreover, it is clear from the proof of Theorem 24 that

$$\text{ind}(T) = - \text{ind}(S)$$  \hspace{1cm} (109)

whenever $S$ is a parametrix for the operator $T$. This observation leads immediately to the following result.

**Theorem 25.** Suppose that $X$ and $Y$ are Banach spaces, that $T : X \to Y$ is Fredholm, and that $K : X \to Y$ is compact. Then $T + K$ is Fredholm and $\text{ind}(T + K) = \text{ind}(T)$.

**Proof.** Let $S$ be a parametrix for the operator $T$ so that

$$ST = I - K_1$$  \hspace{1cm} (110)

and

$$TS = I - K_2$$  \hspace{1cm} (111)

with $K_1$ and $K_2$ compact operators. We observe that

$$S(T + K) = I - K_1 + SK$$  \hspace{1cm} (112)

and

$$(T + K)S = I - K_2 + KS.$$  \hspace{1cm} (113)

Since $K_1 - KS$ and $K_2 - SK$ are compact operators, we conclude from (112) and (113) that $T + K$ is Fredholm and $S$ is a parametrix for $T + K$. It follows that

$$\text{ind}(S) = - \text{ind}(T + K),$$  \hspace{1cm} (114)

but we we also have

$$\text{ind}(S) = - \text{ind}(T)$$  \hspace{1cm} (115)

since $S$ is a parametrix for $T$. We conclude from (114) and (115) that

$$\text{ind}(T) = - \text{ind}(S) = \text{ind}(T + K).$$  \hspace{1cm} (116)
Any isomorphism \( B : X \to Y \) between Banach spaces is Fredholm of index 0. Consequently, it follows from Theorem 25 that any operator of the form
\[
B + K
\]
where \( K : X \to Y \) is compact is Fredholm of index 0. In fact, all Fredholm operators of index 0 are of this form:

**Theorem 26.** Suppose that \( X \) and \( Y \) are Banach spaces, and that \( T : X \to Y \) is a bounded linear operator. Then \( T \) is a Fredholm operator of index 0 if and only if there exist an isomorphism \( B : X \to Y \) and a finite rank operator \( F : X \to Y \) such that \( T = B + F \).

**Proof.** We have already seen that an operator of the form \( B + K \) with \( K \) compact is Fredholm of index 0. So we suppose that \( T : X \to Y \) is a Fredholm operator of index 0. Then there exists a closed subspace \( X' \) of \( X \) such that
\[
X = X' \oplus \ker(T)
\]
and a finite dimensional subspace \( Y' \) of \( Y \) such that
\[
Y = Y' \oplus \operatorname{im}(T).
\]
Moreover, since \( T \) is of index 0 the dimensions of \( \ker(T) \) and \( Y' \) are equal. We denote by \( S \) an isomorphism \( \ker(T) \to Y' \). Suppose that \( x \in X \). We define \( B : X \to Y \) via the formula
\[
B(x) = T(x') + S(z),
\]
where
\[
x = x' + z
\]
is the unique decomposition of \( x \) into \( x' \in X' \) and \( z \in \ker(T) \). It is easy to verify that \( B \) is an isomorphism, and that \( S \) extends to a finite rank linear operator \( X \to Y \).

We now turn to the solvability of the linear equation
\[
Tx = y
\]
when \( T : X \to Y \) is a Fredholm operator of index 0. Since the image of any Fredholm operator is closed, we always have the following solvability condition:

**Theorem 27.** Suppose that \( X \) and \( Y \) are Banach spaces, and that \( T : X \to Y \) is a Fredholm operator. Then the equation
\[
Tx = y
\]
has a solution if and only if
\[
\langle y, \varphi \rangle = 0
\]
for all \( y \) in \( \ker(T^*) \subset Y^* \).

**Theorem 28 (Fredholm Alternative).** Suppose that \( X \) and \( Y \) are Banach spaces, and that \( T : X \to Y \) is a Fredholm operator of index 0. Then either the equation
\[
Tx = y
\]
is uniquely solvable for each \( y \in Y \) or the corresponding homogeneous equation
\[
Tx = 0
\] (126)
admits nontrivial solutions. In the later case, the inverse of the operator \( T \) is bounded.

**Proof.** From Theorem 27, we see that
\[
Tx = y
\] (127)
is solvable for each \( y \in Y \) if and only if the equation
\[
T^*x = 0
\] (128)
admits only the trivial solution. But since \( T \) is a Fredholm operator of index 0,
\[
\dim(\ker(T)) = \dim(\ker(T^*)),
\] (129)
so (128) admits only the trivial solution if and only if \( \dim(\ker(T)) = 0 \) — that is, if and only if
\[
Tx = 0
\] (130)
admits only the trivial solution. If this is the case, then \( T \) is injective and surjective. Since it is bounded, it then follows from the open mapping theorem that \( T \) has a bounded inverse. \( \square \)

### 2.5. Weak Convergence and Compactness

Suppose that \( X \) is a Banach space, and that \( X^* \) is its dual space. We say that a sequence \( \{x_n\} \) in \( X \) converges to \( x \in X \) weakly if
\[
\langle \phi, x_n \rangle \to \langle \phi, x \rangle \quad \text{for all } \phi \in X^*.
\] (131)
We use the notation \( x_n \rightharpoonup x \) to indicate that the sequence \( \{x_n\} \) converges weakly to \( x \). We call the topology generated by the notion of weak convergence — which is the weakest topology with respect to which every bounded linear functions on \( X \) is continuous — the weak topology on \( X \). We say that a subset \( Y \) of \( X \) is weakly compact if it is compact in the weak topology. In this case, any bounded sequence in \( Y \) has a subsequence which converges weakly to some element of \( Y \).

The following is a consequence of the well-known Banach-Alaoglu theorem.

**Theorem 29.** Suppose that \( \Omega \) is an open subset of \( \mathbb{R}^n \), and that \( 1 < p < \infty \) is an integer. Then any bounded subset of \( L^p(\Omega) \) is weakly compact. In particular, any bounded sequence in \( L^p(\Omega) \) has a weakly convergent subsequence.

### 2.6. Classical Function Spaces

In this section, we define several classical spaces of continuous functions, smooth functions and Hölder continuous functions and review several important results relating to them.
2.6. CLASSICAL FUNCTION SPACES

2.6.1. Spaces of Continuous Functions. Suppose that \( \Omega \) is an open subset of \( \mathbb{R}^n \). We denote by \( C(\Omega) \) the vector space of all continuous functions \( f : \Omega \to \mathbb{R} \), by \( C_b(\Omega) \) the subspace of \( C(\Omega) \) consisting of continuous functions \( f : \Omega \to \mathbb{R} \) which are bounded, by \( C(\overline{\Omega}) \) the subspace of \( C_b(\Omega) \) of functions which are uniformly continuous in addition to being bounded, and by \( C_c(\Omega) \) the subspace of \( C(\overline{\Omega}) \) of all continuous functions \( \Omega \to \mathbb{R} \) with compact support contained in \( \Omega \). The notation \( C(\overline{\Omega}) \) is used for the space of bounded, uniformly continuous functions \( \Omega \to \mathbb{R} \) because any such function admits a unique continuous extension to the closure \( \overline{\Omega} \) of \( \Omega \).

Exercise 4. Show that any continuous function on a compact subset of \( \mathbb{R}^n \) is uniformly continuous.

Exercise 5. Suppose that \( \Omega = \mathbb{R}^n \). Show that \( C(\overline{\Omega}) \) is not the same space as \( C(\mathbb{R}^n) \), which means that the notation \( C(\overline{\Omega}) \) is misleading when \( \Omega \) is not bounded.

The vector space \( C_b(\Omega) \) is a Banach space when endowed with the uniform norm

\[
\|f\| = \sup_{x \in \Omega} |f(x)|,
\]  

(132)
as is the vector space \( C(\overline{\Omega}) \). Neither \( C(\Omega) \) nor \( C_c(\Omega) \) are Banach spaces with respect to (132), although if \( K_1 \subset K_2 \subset \ldots \) is an increasing sequence of compact sets such that \( \Omega = \bigcup K_j \), then they are Fréchet spaces with respect to the family of seminorms

\[
\|f\|_i = \sup_{x \in K_i} |f(x)|, \quad i = 1, 2, \ldots
\]  

(133)
We will not make use of this last observation, but a further discussion can be found in, for instance, [8].

Exercise 6. Suppose that \( \Omega \) is an open set in \( \mathbb{R}^n \). Show that any uniformly continuous function \( f : \Omega \to \mathbb{R} \) is bounded. Give an example to show that if \( \Omega \) is not bounded, then there exist uniformly continuous functions \( \Omega \to \mathbb{R} \) which are not bounded.

Suppose that \( \Omega \) is a bounded open subset of \( \mathbb{R}^n \). The following theorem, a proof of which can be found in [4] (for instance), characterizes the compact subsets of \( C(\overline{\Omega}) \). We say that a subset \( \Phi \) of \( C(\overline{\Omega}) \) is equicontinuous if for all \( \epsilon > 0 \) there exists \( \delta > 0 \) such that

\[
|f(x) - f(y)| < \epsilon
\]  

(134)for all \( x \) and \( y \) in \( \overline{\Omega} \) such that

\[
|x - y| < \delta.
\]  

(135)Similarly, we say that \( \Phi \) is uniformly bounded if there exists a \( M > 0 \) such that

\[
\|f\| < M
\]  

(136)for all \( f \in \Phi \).

Theorem 30 (Arzelà-Ascoli). Suppose that \( \Omega \) is a bounded open subset of \( \mathbb{R}^n \), and that \( \Phi \) is a subset of \( C(\overline{\Omega}) \). Then \( \Phi \) is compact if and only if it is equicontinuous, closed and uniformly bounded.
2.6.2. Spaces of Smooth Functions. Suppose that $\Omega$ is an open set in $\mathbb{R}^n$. We denote by $C^k(\Omega)$ the vector space of functions $\Omega \to \mathbb{R}$ whose derivatives through order $k$ are continuous, by $C^k_b(\Omega)$ the subspace of $C^k(\Omega)$ consisting of functions whose derivatives through order $k$ are continuous and bounded, by $C^k(\overline{\Omega})$ the subspace of $C^k(\Omega)$ consisting of $k$-times differentiable functions whose derivatives through order $k$ are bounded and uniformly continuous on $\Omega$, and by $C^k_c(\Omega)$ the subspace of $C^k(\overline{\Omega})$ consisting of $k$-times differentiable functions with compact support contained in $\Omega$. The vector spaces $C^k_b(\Omega)$ and $C^k(\overline{\Omega})$ are Banach spaces with respect to the norm

$$\|f\| = \sum_{|\beta| \leq k} \sup_{x \in \Omega} |D^\beta f(x)|.$$  \hspace{1cm} (137)

By $C^\infty(\Omega)$ we mean the vector space of functions infinitely differentiable functions $\Omega \to \mathbb{R}$. We denote by $C^\infty_b(\Omega)$ the subspace of $C^\infty(\Omega)$ consisting of infinitely differentiable functions $\Omega \to \mathbb{R}$ whose derivatives of all orders are bounded, and by $C^\infty_c(\Omega)$ the space of infinitely differentiable functions with compact support contained in $\Omega$. All of these spaces are Fréchet spaces with respect to appropriately chosen families of seminorms (see, for instance, [8]); however, we will not make use of this fact.

**Exercise 7.** Suppose that $\Omega$ is an open connected set in $\mathbb{R}^n$. Show that the norm (137) and

$$\|f\|_0 = \sup_{x \in \Omega} |f(x)| + \sum_{|\beta| = k} \sup_{x \in \Omega} |D^\beta(x)|$$  \hspace{1cm} (138)

are equivalent norms for $C^k(\overline{\Omega})$.

2.6.3. Hölder Spaces. Suppose that $\Omega$ is an open set in $\mathbb{R}^n$. For $k$ a nonnegative integer and $0 < \alpha \leq 1$ a real number we denote by $C^{k,\alpha}(\overline{\Omega})$ the subspace of $C^k(\overline{\Omega})$ consisting of functions whose derivatives of orders through $k$ satisfy a Hölder condition of exponent $\alpha$; that is, there exists a constant $C$ such that

$$|D^\beta f(x) - D^\beta f(y)| \leq C|x - y|^\alpha$$  \hspace{1cm} (139)

for all $x, y \in \Omega$ and all multi-indices $|\beta| \leq k$. When endowed with the norm

$$\|f\| = \|f\|_{C^m(\overline{\Omega})} + \sup_{0 \leq |\beta| \leq m} \sup_{x, y \in \Omega} \frac{|D^\beta f(x) - D^\beta f(y)|}{|x - y|^\alpha},$$  \hspace{1cm} (140)

$C^{k,\alpha}(\overline{\Omega})$ is a Banach space.

**Exercise 8.** Suppose that $\Omega$ is a bounded open set in $\mathbb{R}^n$. Show that for any positive integer $k$ and any real numbers $0 < \lambda < \alpha \leq 1$,

$$C^{k,\alpha}(\overline{\Omega}) \subset C^{k,\lambda}(\overline{\Omega}) \subset C^k(\overline{\Omega}).$$  \hspace{1cm} (141)

Note carefully the assumption that $\Omega$ is bounded.

Suppose that $\Omega$ is an open set in $\mathbb{R}^n$, $k$ is a nonnegative integer and $0 < \alpha \leq 1$ is a real number. We denote by $C^{k,\alpha}(\Omega)$ the subspace of $C^k(\Omega)$ consisting of functions whose restrictions to each bounded open subset $\Omega' \subset \subset \Omega$ is contained in $C^{k,\alpha}(\overline{\Omega'})$. By $\Omega' \subset \subset \Omega$ we mean that $\Omega'$ is compactly supported in $\Omega$; that is, there exists a compact set $K$ such that $\overline{\Omega'} \subset K \subset \Omega$. 

For the sake of convenience, we set
\[ C^k,0 (\overline{\Omega}) = C^k (\overline{\Omega}) \] (142)
and
\[ C^{k,0} (\Omega) = C^k (\Omega) \] (143)
for all nonnegative integers \( k \).

Suppose that \( k \geq 0 \) is an integer, that \( 0 \leq \alpha \leq 1 \) is a real number, and that \( \Omega \) is an open subset of \( \mathbb{R}^n \). Suppose also that \( \psi : \Omega \to \Omega', \) and that \( \psi_1, \ldots, \psi_m \) are mappings \( \Omega \to \mathbb{R} \) such that
\[ \psi(x_1, \ldots, x_n) = \begin{pmatrix} \psi_1(x_1, \ldots, x_n) \\ \vdots \\ \psi_m(x_1, \ldots, x_n) \end{pmatrix} \] (144)
for all \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \). Then we say that \( \psi \) is a \( C^{k,\alpha} \) mapping if each of the mappings \( \psi_j \) is an element of \( C^{k,\alpha}(\Omega) \). Obviously, the \( C^{0,1} \) mappings \( \Omega \to \mathbb{R}^m \) are the Lipschitz continuous functions \( \Omega \to \mathbb{R}^m \); that is, \( f : \Omega \to \mathbb{R}^m \) is a \( C^{0,1} \) mapping if and only if there exists \( C > 0 \) such that
\[ \| f(x) - f(y) \| \leq C \| x - y \| \] (145)
for all \( x, y \in \Omega \).

**Exercise 9.** Suppose that \( \Omega \) is a bounded open set in \( \mathbb{R}^n \), that \( k \) is a positive integer, and that \( 0 < \alpha \leq 1 \) is a real number. Show that \( C^{k,\alpha}(\overline{\Omega}) \) is compactly embedded in \( C^k (\overline{\Omega}) \). Hint: use the Arzelà-Ascoli theorem.

**Exercise 10.** Why do we only consider Hölder exponents which are less than or equal to 1?

**2.6.4. Lipschitz continuous functions.** The space \( C^{0,1} (\Omega) \), whose elements are known as Lipschitz continuous functions, will play an important role in this course. They are sufficiently smooth to take the place of differentiable functions much of the time, but they offer more flexibility in modeling physical problems than differentiable functions (this is particularly important when it comes to modeling the domains in which boundary value problems are given — the boundary of a square can be described using Lipschitz functions but not with differentiable functions).

Suppose that \( f : [a, b] \subset \mathbb{R} \to \mathbb{R} \) is a Lipschitz continuous function. It is easy to verify that \( f \) is absolutely continuous; that is, for all \( \epsilon > 0 \) there exists \( \delta > 0 \) such that
\[ \sum_{j=1}^n |f(b_j) - f(a_j)| \leq \epsilon \] (146)
whenever \( (a_1, b_1), \ldots, (a_n, b_n) \) is a finite collection of disjoint open intervals in \( [a, b] \) such that
\[ \sum_{j=1}^n (b_j - a_j) \leq \delta. \] (147)
Absolutely continuous functions are characterized by the following theorem, which is a standard result in measure theory (for a proof, see, for instance, Chapter 3 of [4]).
Theorem 31. A function $f : [a, b] \to \mathbb{R}$ is absolutely continuous if and only if the derivative $f'$ exists almost everywhere in $[a, b]$, the derivative $f'$ is integrable, and for every $x \in [a, b]$

$$f(x) = f(a) + \int_a^x f'(y)\, dy. \quad (148)$$

That the usual integration by parts formula

$$f(b)g(b) - f(a)g(a) = \int_a^b f'(x)g(x)\, dx + \int_a^b f(x)g'(x)\, dx \quad (149)$$

holds when $f$ and $g$ are absolutely continuous functions is also a standard result in measure theory (see, for instance, Theorem 3.36 in Chapter 3 of [4]).

We conclude that a Lipschitz continuous function $f : [a, b] \to \mathbb{R}$ is differentiable almost everywhere, and that the integration by parts formula (149) holds for such functions. Also, from (148), we see that the derivative of $f$ must be bounded almost everywhere.

Rademacher’s theorem, which we state below, extends Theorem 31 to higher dimensions. A proof can be found in [9].

Theorem 32 (Rademacher). Suppose that $\Omega$ is an open set in $\mathbb{R}^n$, and that $f : \Omega \to \mathbb{R}^m$ is Lipschitz continuous with Lipschitz constant $C$. That is, suppose that

$$\|f(x) - f(y)\| \leq C\|x - y\| \quad (150)$$

for all $x, y \in \Omega$. Then $f$ is differentiable almost everywhere in $\Omega$ and the operator norm of $f'(x)$ is bounded by $C$ for almost all $x \in \Omega$.

It is important to understand that Theorem 32 asserts the almost everywhere existence of the “total derivative” of $f$. That is, if $f : \Omega \to \mathbb{R}^m$ is a Lipschitz mapping, then for almost all $x \in \Omega$ there exists a linear mapping $T : \mathbb{R}^n \to \mathbb{R}$ such that

$$\lim_{h \to 0} \frac{\|f(x + h) - f(x) - Th\|}{\|h\|} = 0. \quad (151)$$

Many of the standard results of multivariable calculus require only the pointwise existence of total derivatives, and hence apply to Lipschitz continuous functions without significant modification. The multivariable chain rule and product rule are examples:

Theorem 33. Suppose that $\Omega$ is an open subset of $\mathbb{R}^n$, that $\Omega'$ is an open subset of $\mathbb{R}^m$, and that $\Omega''$ is an open subset of $\mathbb{R}^k$. Suppose also that $f : \Omega \to \Omega'$ and that $g : \Omega' \to \Omega''$. If $g$ is differentiable at $x$ and $f$ is differentiable at $g(x)$, then the composition $f \circ g$ is differentiable at $x$ and $(f \circ g)'(x) = f'(g(x))g'(x)$.

Theorem 34. Suppose that $\Omega$ is an open subset of $\mathbb{R}^n$, and that $f, g : \Omega \to \mathbb{R}^m$ are differentiable at $x \in \mathbb{R}$. Then

$$(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x). \quad (152)$$

It is somewhat more difficult to establish the following change of variables formula for Lipschitz mappings. A proof can be found in [3].
Theorem 35. Suppose that $\Omega$ and is a open subset of $\mathbb{R}^n$, and that $\psi : \Omega \to \mathbb{R}^n$ is a bilipschitz mapping. Then
\[
\int_\Omega g(\psi(x)) \det (\psi'(x)) \, dx = \int_{\psi(\Omega)} g(y) \, dy
\]
for all measurable $g : \Omega \to \mathbb{R}$. In particular, the measure of $\psi(\Omega)$ is
\[
\int_\Omega \det (\psi'(x)) \, dx.
\]
(153)

The implicit function theorem is an example of a result in multivariable calculus which requires $C^1$ differentiability and does not extend to Lipschitz continuous functions.

2.7. Mollifiers, Cutoff Functions and Partitions of Unity

The function $\eta : \mathbb{R}^n \to \mathbb{R}$ defined via
\[
\eta(x) = \begin{cases} 
\exp \left( -\frac{1}{1-|x|^2} \right) & \text{if } |x| \leq 1 \\
0 & \text{if } |x| > 1,
\end{cases}
\]
is an element of $C^\infty_c(\mathbb{R}^n)$. For each $\epsilon > 0$, we define
\[
\eta_\epsilon(x) = (\alpha)^{-1} \epsilon^{-n} \eta(x/\epsilon),
\]
where
\[
\alpha = \int_{\mathbb{R}^n} \eta(x) \, dx,
\]
so that $\eta_\epsilon$ is supported on the ball $\{ x : |x| \leq \epsilon \}$ and
\[
\int \eta_\epsilon(x) \, dx = 1.
\]
(157)

We call $\eta_\epsilon$ the standard mollifier on $\mathbb{R}^n$. The sequence of functions
\[
\eta_\epsilon * u(x) = \int_{\mathbb{R}^n} \eta_\epsilon(x-y)u(y) \, dy
\]
(159)
obtained by convolving $u$ with $\eta_\epsilon$ is called mollification or regularization of $f$. The following theorem, whose proof can be easily found in the literature (for example, in [4]), enumerates some of the key properties of mollifiers.

Theorem 36. Suppose that $u \in L^1_{\text{loc}}(\mathbb{R}^n)$, and that $\eta_h$ is the standard mollifier. Then:

1. For each $h > 0$, $\eta_h * u$ is an element of $C^\infty(\mathbb{R}^n)$.
2. If $u \in L^p(\mathbb{R}^n)$, where $1 \leq p < \infty$, then $\| \eta_h * u - u \|_p \to 0$ as $h \to 0$.
3. If $u \in C(\mathbb{R}^n)$, then $\eta_h * u$ converges to $u$ uniformly on compact subsets of $\mathbb{R}^n$.
4. If $u$ is compactly supported, and $\Omega$ is an open set in $\mathbb{R}^n$ such that $0 < h < \text{dist}(\text{supp}(u), \partial \Omega)$, then the support of $\eta_h * u$ is contained in $\Omega$.

Exercise 11. Suppose that $\eta_h$ is the standard mollifier. Show that there exists a function $u$ in $L^\infty(\mathbb{R}^n)$ such that $\eta_h * u$ does not converge to $u$ in $L^\infty(\mathbb{R}^n)$ norm as $h \to 0$. 
We will typically work with functions which are only defined on an open subset \( \Omega \) of \( \mathbb{R}^n \). If \( u \in L^1_{\text{loc}}(\Omega) \), then the mollification

\[
\eta_h * u(x) = \int_{\Omega} \eta_h(x - y)u(y) \, dy = \int_{\Omega} \eta_h(y)u(x - y) \, dy
\]

is defined for all \( x \in \Omega \) and \( h > 0 \) such that \( 0 < \text{dist}(x, \partial \Omega) < h \) (in which case the support of the function \( v(y) = \eta_h(x - y) \) is a compact subset contained in \( \Omega \) so that the integral in (160) converges).

The following definition is useful when mollifying functions defined on subsets of \( \mathbb{R}^n \). If \( \Omega \) is an open subset of \( \mathbb{R}^n \) and \( \Omega' \) is an open subset of \( \Omega \), then we say that \( \Omega' \) is compactly embedded in \( \Omega \) and write \( \Omega' \subset\subset \Omega \) if there exists a compact set \( K \) such that \( \Omega' \subset K \subset \Omega \). If \( u \in L^1_{\text{loc}}(\Omega) \) and \( \Omega' \subset\subset \Omega \), then for sufficiently small \( h \), \( \eta_h * u \) is defined and so it is reasonable to speak of the limit of \( \eta_h * u \) as \( h \to 0 \).

**Theorem 37.** Suppose that \( \Omega \) is an open subset of \( \mathbb{R}^n \), that \( \Omega' \subset\subset \Omega \), and that \( \eta_h \) denotes the standard mollifier. Then:

1. If \( u \in L^p_{\text{loc}}(\Omega) \) with \( 1 \leq p < \infty \), then \( \eta_h * u \to u \) in \( L^p(\Omega') \) and \( \eta_h * u \) converges to \( u \) almost everywhere in \( \Omega' \).
2. If \( u \in C(\Omega) \), then \( \eta_h * u \) converges to \( u \) uniformly on \( \Omega' \).
3. If \( u \in C(\overline{\Omega}) \), then \( \eta_h * u \) converges to \( u \) uniformly on \( \Omega \).

If \( u \) is integrable on arbitrary open sets of \( \Omega \) (which is not always the case — e.g., when \( \Omega = (0, 1) \) and \( u(x) = 1/x \)), then Formula (160) defines \( \eta_h * u \) for all \( x \in \Omega \). More specifically, if we extend \( u \) to all of \( \mathbb{R}^n \) by letting \( u(x) = 0 \) for all \( x \notin \Omega \) then

\[
\eta_h * u(x) = \int_{\Omega} \eta_h(x - y)u(y) \, dy = \int_{\Omega} \eta_h(y)u(x - y) \, dy
\]

is defined for all \( x \in \Omega \).

**Theorem 38.** Suppose that \( \Omega \) is an open set in \( \mathbb{R}^n \), that \( 1 \leq p < \infty \) is an integer, and that \( u \in L^p(\Omega) \). Then

\[
\lim_{h \to 0} \| \eta_h * u - u \|_{L^p(\Omega)} \to 0.
\]

We now use mollifiers to establish the existence of smooth cutoff functions; that is, functions which are exactly equal to 1 on a specified compact set and which decay smoothly to 0 outside of that set.

**Theorem 39.** Suppose that \( U \) is an open subset of \( \mathbb{R}^n \), and that \( V \subset\subset U \). Then there exists a nonnegative function \( \psi \in C^\infty_c(U) \) which is identically 1 on \( V \).

**Proof.** We denote the standard mollifier by \( \eta_\epsilon \), let \( W \) be an open set such that

\[
V \subset W \subset \overline{W} \subset U,
\]

(163)
and choose \( \epsilon < \min\{\text{dist}(\partial U, \partial W), \text{dist}(\partial V, \partial W)\} \). We claim that the function \( \psi \) defined via the formula

\[
\psi(x) = \chi_W * \eta
\]  

is the desired cutoff function. To see this, we observe that according Conclusion (1) of Theorem 36, \( f \in C^\infty(\mathbb{R}^n) \), and that conclusion (4) of the same theorem implies that the support of \( \psi \) is contained in \( U \). Moreover, for all \( x \in V \) the support of the function \( \eta(x - \cdot) \) is contained in \( W \), so that

\[
\psi(x) = \int_{\mathbb{R}^n} \eta(x - y) \chi_W(y) \, dy = \int_{W} \eta(x - y) \, dy = \int_{\mathbb{R}^n} \eta(x - y) \, dy = 1. 
\]  

It is the case that \( \psi(x) \geq 0 \) for all \( x \in \mathbb{R}^n \) since both \( \eta \) and \( \xi_W \) are nonnegative functions.

Suppose that \( A \) is an arbitrary subset of \( \mathbb{R}^n \), and that

\[
A \subseteq \bigcup_{\alpha \in \mathcal{O}} U_\alpha
\]

is an open covering of \( A \). We say that a sequence of \( C^\infty(\mathbb{R}^n) \) functions \( \psi_1, \psi_2, \ldots \) is a smooth partition of unity subordinate to the cover (166) if

1. each of the functions \( \psi_j \) is supported in one of the sets \( U_\alpha \);
2. \( 0 \leq \psi_j(x) \leq 1 \) for all \( x \in A \) and \( j = 1, 2, \ldots \);
3. there exists a neighborhood of each \( x \in A \) on which all but a finite number of the functions \( \psi_j \) vanish;
4. \( \sum_{j=1}^{\infty} \psi_j(x) = 1 \) for all \( x \in A \).

We now establish the existence of smooth partitions of unity, starting with the special case in which the set \( A \) is compact.

**Theorem 40.** Suppose that \( A \) is a compact subset of \( \mathbb{R}^n \), and that

\[
A \subseteq \bigcup_{\alpha \in \mathcal{O}} U_\alpha
\]

is a covering of \( A \) by open sets. Then there exist \( C^\infty_c(\mathbb{R}^n) \) functions \( \psi_1, \psi_2, \ldots, \psi_m \) such that

1. Each \( \psi_j \) is supported in one of the open sets \( U_\alpha \);
2. \( 0 \leq \psi_j(x) \leq 1 \) for all \( x \in \mathbb{R}^n \) and all \( j = 1, 2, \ldots, m \);
3. for each \( x \in A \), there exists an open neighborhood of \( x \) in which \( \sum_{j=1}^{m} \psi_j(x) = 1 \).

**Proof.** We let

\[
A \subseteq \bigcup_{j=1}^{m} U_j
\]
be a covering of $A$ by open sets chosen from the collection $\{U_\alpha : \alpha \in \mathcal{O}\}$. Now we construct a collection of open sets $V_1, \ldots, V_m$ such that
\[
A \subset \bigcup_{j=1}^m V_j
\] (169)
and
\[
V_j \subset U_j
\] (170)
for each $j = 1, 2, \ldots, m$. To that end, for each $\epsilon > 0$ and $j = 1, \ldots, m$, we let $V_{j,\epsilon}$ be the open subset of $U_j$ defined via
\[
V_{j,\epsilon} = \{ x \in U_j : \text{dist} (x, U^c_j) > \epsilon \}. \tag{171}
\]
We claim that if $\epsilon > 0$ is sufficiently small, then
\[
A \subset \bigcup_{j=1}^m V_{j,\epsilon}
\] (172)
If not, then for each $n > 0$, there exists $x_n$ such that
\[
x_n \in A \setminus \bigcup_{j=1}^m V_{j,1/n} = A \cap \bigcap_{j=1}^m V^c_{j,1/n}, \tag{173}
\]
Since $A$ is compact, the sequence $\{x_n\}$ has a subsequence converging to a point $x \in A$. From (173) we see that
\[
x \notin \bigcup_{j=1}^m V_{j,1/n}
\] (174)
for all $j = 1, \ldots, m$ and all positive integers $n$. We conclude from (174) that $x \in U^c_j$ for all $j = 1, \ldots, m$. But this is a contradiction since the $U_j$ cover $A$ and $x \in A$. We conclude that (172) holds.

According to Theorem 39, for each $j = 1, \ldots, m$, there exists a function $\varphi_j \in C_c^\infty (U_j)$ which is 1 on $V_j$. Then the functions $\psi_1, \ldots, \psi_m$ defined via the formula
\[
\psi_j(x) = \frac{\varphi_j(x)}{\sum_{i=1}^m \varphi_j(x)} \tag{175}
\]
have the desired properties.

**Theorem 41.** Suppose that $A$ is an arbitrary subset of $\mathbb{R}^n$, and that
\[
A \subset \bigcup_{\alpha \in \mathcal{O}} U_\alpha \tag{176}
\]
is a covering of $A$ by open sets. Then there exists a smooth partition of unity $\psi_1, \psi_2, \ldots$ subordinate to (176).

**Proof.** We first suppose that $A$ is open. Then
\[
A = \bigcup_{k=1}^\infty A_k, \tag{177}
\]
where the sets \( A_k \) are defined by the formula
\[
A_k = \{ x \in A : |x| \leq k \text{ and } \text{dist}(x, \partial A) \geq 1/k \}. 
\] (178)
Note that if \( A \) is not open, then (177) need not hold. For each \( \alpha \in \mathcal{O} \) and each \( k \geq 1 \), we let
\[
V_{\alpha,k} = U_\alpha \cap \text{int} (A_{k+1} \setminus A_{k-2}), 
\] (179)
where we set \( A_0 = A_{-1} = \emptyset \). Then, for each fixed \( k \geq 1 \),
\[
A_k \subset \bigcup_{\alpha \in \mathcal{O}} V_{\alpha,k} 
\] (180)
is an open covering of the compact set \( A_k \). For each \( k \geq 1 \), we invoke Theorem 40 in order to obtain a smooth partition of unity
\[
\psi_1^k, \psi_2^k, \ldots, \psi_{m_k}^k 
\] (181)
subordinate to the covering (180). We define the function \( \sigma : A \to \mathbb{R} \) via the formula
\[
\sigma(x) = \sum_{k=1}^{\infty} \sum_{j=1}^{m_k} \psi_j^k(x); 
\] (182)
note that only a finite number of terms are nonzero for each point \( x \) and that \( \sigma(x) > 0 \) for all \( x \in A \). For each pair \( k = 1, 2, \ldots \) and \( 1 \leq j \leq m_k \), we define
\[
\varphi_{k,j}(x) = \frac{\psi_j^k(x)}{\sigma(x)}. 
\] (183)
The collection of functions \( \{ \varphi_{k,j} \} \) is the desired smooth partition of unity.

Suppose now that \( A \) is an arbitrary subset of \( \mathbb{R}^n \). Then we construct a smooth partition of unity for the set
\[
B = \bigcup_{\alpha \in \mathcal{O}} U_\alpha 
\] (184)
subordinate to the open covering
\[
B \subset \bigcup_{\alpha \in \mathcal{O}} U_\alpha. 
\] (185)
We conclude the proof by observing that the resulting smooth partition of unity for \( B \) subordinate to the covering (185) is also a smooth partition of unity for \( A \) subordinate to the covering (176).
\[ \square \]

**Theorem 42.** Suppose that \( A \) is an arbitrary set in \( \mathbb{R}^n \), and that
\[
A \subset \bigcup_{j=1}^{\infty} U_j. 
\] (186)
is a covering of \( A \) by open sets. Then there exists a smooth partition of unity \( \psi_1, \psi_2, \ldots \) subordinate to the covering (186) such that
\[
\text{supp}(\psi_j) \subset U_j. 
\] (187)
Proof. We let $\varphi_1, \varphi_2, \ldots$ be a smooth partition of unity subordinate to the covering

$$A \subseteq \bigcup_{j=1}^{\infty} U_j$$

whose existence is ensured by Theorem 39. We define

$$I_1 = \{j \geq 1 : \text{supp}(\varphi_j) \subset U_1\}$$

and, for $k > 1$,

$$I_k = \{j \geq 1 : \text{supp}(\varphi_j) \subset U_k \text{ and } j \notin I_1 \cup I_2 \cup \cdots \cup I_{k-1}\}.$$

We then define a new partition of unity $\psi_j : j = 1, 2, \ldots$ via the formula

$$\psi_j(x) = \sum_{i \in I_j} \varphi_i(x).$$

\[\square\]

2.8. Domains with $C^{k,\alpha}$ Boundary

Suppose that $\Omega$ is a bounded open set in $\mathbb{R}^n$, that $k$ is a nonnegative integer, and that $0 \leq \alpha \leq 1$ is a real number. The domain $\Omega$ is of class $C^{k,\alpha}$ if $\partial \Omega$ is locally the graph of a $C^{k,\alpha}$ function. That is, if for each $x \in \partial \Omega$ there exists an open set $V$ containing $x$ and a new orthogonal coordinate system $y_1, \ldots, y_n$ such that

1. $V$ is a hypercube in the new coordinate system; i.e., there exist $a_1, \ldots, a_n$ such that

$$V = \{(y_1, \ldots, y_n) : -a_i < y_i < a_i \text{ for all } i = 1, 2, \ldots, n\}$$

2. there exists a function $\psi$ in $C^{k,\alpha}(\overline{V'})$, where $V'$ is defined via

$$V' = \{(y_1, \ldots, y_{n-1}) : -a_i < y_i < a_i \text{ for all } i = 1, 2, \ldots, n-1\},$$

such that

$$-a_n < \psi(y_1, \ldots, y_{n-1}) < a_n \text{ for all } (y_1, \ldots, y_{n-1}) \in V';$$

$$V \cap \Omega = \{(y_1, \ldots, y_n) : y_n < \psi(y_1, \ldots, y_{n-1})\},$$

and

$$V \cap \partial \Omega = \{(y_1, \ldots, y_n) : y_n = \psi(y_1, \ldots, y_{n-1})\}.$$

A bounded open set $\Omega$ in $\mathbb{R}^n$ is a $C^{k,\alpha}$ submanifold with boundary in $\mathbb{R}^n$ if each $x \in \partial \Omega$ there exist an open set $V$ containing $x$ and an injective mapping $\psi : V \to \mathbb{R}^n$ such that

1. $\psi$ is a $C^{k,\alpha}$ mapping;

2. $\psi^{-1} : \psi(V) \to V$ is a $C^{k,\alpha}$ mapping;

3. $\Omega \cap V = \{y : \psi_n(y) < 0\}$, where $\psi_n$ denotes the $n^{th}$ component of $\psi(y) \in \mathbb{R}^n$;

4. $\partial \Omega \cap V = \{y : \psi_n(y) = 0\}$, where $\psi_n$ denotes the $n^{th}$ component of $\psi(y) \in \mathbb{R}^n$. 
It is easy to verify that a $C^{k,\alpha}$ domain in $\mathbb{R}^n$ is always a $C^{k,\alpha}$ submanifold with boundary in $\mathbb{R}^n$, but the converse is not always true. See, for instance, [5] for an example of a domain which is a $C^{0,1}$ submanifold but not a $C^{0,1}$ domain.

**Exercise 12.** Use the implicit function theorem to show that a $C^{1,0}$ submanifold with boundary in $\mathbb{R}^n$ is necessarily a $C^{1,0}$ domain.

There are two common geometric assumptions which are equivalent to requirements that $\Omega$ is a $C^{k,\alpha}$ domain of the appropriate type. The bounded open set $\Omega \subset \mathbb{R}^n$ has the segment property if for every $x \in \partial \Omega$ there exist a neighborhood $V$ of $x$, a new orthogonal coordinate system $(y_1,\ldots,y_n)$, and a real number $h > 0$ such that

1. $V$ is a hypercube in the new coordinate system; i.e., there exist $a_1,\ldots,a_n$ such that
   \[ V = \{(y_1,\ldots,y_n) : -a_i < y_i < a_i \text{ for all } i = 1,2,\ldots,n\} \]  
   \hspace{-1cm} (197)

2. $y - t(0,0,\ldots,0,1)$ is in $\Omega$ whenever $0 < t < h$ and $y \in \overline{\Omega} \cap V$.

Similarly, $\Omega$ has the cone property if for every $x \in \partial \Omega$ there exist a neighborhood $V$ of $x$, a new orthogonal coordinate system $(y_1,\ldots,y_n)$, and constants $h > 0$ and $0 < \theta \leq \pi/2$ such that

1. $V$ is a hypercube in the new coordinate system; i.e., there exist $a_1,\ldots,a_n$ such that
   \[ V = \{(y_1,\ldots,y_n) : -a_i < y_i < a_i \text{ for all } i = 1,2,\ldots,n\} \]  
   \hspace{-1cm} (198)

2. $y - z$ is in $\Omega$ whenever $y \in \overline{\Omega} \cap V$ and $z$ is contained in the cone
   \[ C = \{(t_1,\ldots,t_{n-1},t_n) : \cot(\theta) |(t_1,\ldots,t_{n-1})| < t_n < h\}. \]  
   \hspace{-1cm} (199)

**Theorem 43.** Suppose that $\Omega$ is an open set in $\mathbb{R}^n$. Then $\Omega$ has the segment property if and only if its boundary is $C^0$.

**Theorem 44.** Suppose that $\Omega$ is an open set in $\mathbb{R}^n$. Then $\Omega$ has the cone property if and only if its boundary is $C^{0,1}$.

See [5] for a proof of Theorem 44; prove Theorem 43 as an exercise. We will not make use of the following theorem, but it is a straightforward consequence of Theorem 44 that you might find useful.

**Theorem 45.** Suppose that $\Omega$ is a bounded open convex set in $\mathbb{R}^n$. Then the boundary of $\Omega$ is $C^{0,1}$. 

CHAPTER 3

Sobolev spaces

In this chapter, we discuss the elementary properties of Sobolev spaces, a family of function spaces which serve as the principle setting for the variational theory of elliptic partial differential equations.

3.1. Weak Derivatives

We now introduce notation which, inter alia, simplifies expressions involving partial derivatives. We call an n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers a multi-index. The absolute value of $\alpha = (\alpha_1, \ldots, \alpha_n)$ is

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n,$$

we define the factorial of a multi-index $\alpha$ via

$$\alpha! = \alpha_1! \alpha_2! \cdots \alpha_n!,$$

and for $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ we define

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$$

Moreover, for each multi-index $\alpha$, we denote by $D^\alpha$ the partial differential operator which acts on functions $u \in C^{[\alpha]}(\Omega)$ via the formula

$$D^\alpha u = \left( \frac{\partial \varphi}{\partial x_1} \right)^{\alpha_1} \left( \frac{\partial \varphi}{\partial x_2} \right)^{\alpha_2} \cdots \left( \frac{\partial \varphi}{\partial x_n} \right)^{\alpha_n} u.$$

Using multi-index notation, the Taylor expansion of a function $u$ about a point $x$ becomes

$$u(y) = \sum_{|\alpha|=0}^k \frac{\alpha!}{j!} D^\alpha u(x)(y-x)^\alpha + O\left(|y-x|^{k+1}\right)$$

and the product rule is

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

where $\alpha \leq \beta$ if and only $\alpha_i \leq \beta_i$ for all $i = 1, \ldots, n$ and

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha - \beta)!}.$$
Suppose that $\Omega$ is an open set in $\mathbb{R}^n$, and that $\alpha = (\alpha_1, \ldots, \alpha_n)$ is an n-tuple. We say that $v \in L^1_{loc}(\Omega)$ is the $\alpha^{th}$ weak derivative of $u \in L^1_{loc}(\Omega)$ if

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

(208)

for all $\psi$ in $C_c^\infty(\Omega)$.

**Exercise 13.** What is the first weak derivative of the function $f : \mathbb{R} \to \mathbb{R}$ defined via $f(x) = |x|$? Is $f$ twice weakly differentiable?

**Exercise 14.** Show that if $u \in C^k(\Omega)$ and $|\alpha| \leq k$ is a multi-index, then $u$ has weak derivatives of orders 0 through $k$, and the $\alpha^{th}$ weak derivative of $u$ is the $\alpha^{th}$ classical derivative of $u$.

Since the notions of classical and weak differentiability coincide when $u$ is classically differentiable, there is no harm in denoting the $\alpha^{th}$ weak derivative of a function $u \in L^1_{loc}(\Omega)$ by $D^\alpha u$.

**Exercise 15.** Suppose that $\Omega$ is an open subset of $\mathbb{R}^n$, and that $\alpha$ is a multi-index. Show that if $u, v_1$ and $v_2$ are elements of $L^1_{loc}(\Omega)$ such that $D^\alpha u = v_1$ and $D^\alpha u = v_2$, then $v_1 = v_2$ almost everywhere.

**Exercise 16.** Suppose that $\Omega$ is an open subset of $\mathbb{R}$, and that $u, v$ and $w$ are $L^1_{loc}(\Omega)$ functions such that the first weak derivative of both $u$ and $v$ is $w$. Show that there exists a constant $C > 0$ such that $u(x) - v(x) = C$ almost everywhere.

**Exercise 17.** What is the first weak derivative of the function $f : \mathbb{R} \to \mathbb{R}$ defined via the formula $f(x) = \sin(1/x)$?

**Exercise 18.** Suppose that $\Omega$ is an open subset of $\mathbb{R}^n$, that $\alpha$, $\beta$ and $\gamma$ are multi-indices such that $\alpha = \beta + \gamma$, and that $u \in L^1_{loc}(\Omega)$ such that $D^\alpha$, $D^\beta$ and $D^\gamma$ exist. Show that $D^\alpha u = D^\beta D^\gamma u = D^\gamma D^\beta u$.

**Exercise 19.** Suppose that $\Omega = (a, b)$ is an open interval in $\mathbb{R}$. Show that $f \in C^{0,1}(\Omega)$ if and only if $f$ is weakly differentiable and its weak derivative is an element of $L^\infty_{loc}(\Omega)$. You are free to make use the theorems of Section 2.6.

We will frequently rely on the following theorem that asserts that weak differentiation commutes with mollification. We note that if $u$ is locally integrable and admits an $\alpha^{th}$ weak derivative, then both $\eta_h \ast u$ and $\eta_h \ast D^\alpha u$ are infinitely differentiable functions, and hence are pointwise defined.

**Theorem 46.** Suppose that $\Omega$ is an open set in $\mathbb{R}^n$, that $u \in L^1_{loc}(\Omega)$ whose $\alpha^{th}$ weak derivative exists. Then for all $x \in \Omega$ and $h > 0$ such that $\overline{B_h(x)} \subset \Omega$,

$$D^\alpha (\eta_h \ast u)(x) = \eta_h \ast (D^\alpha u)(x).$$

(209)

**Proof.** We observe that for all $x$ and $h$ which satisfy the hypotheses of the theorem, the function

$$f_h(y) = \eta_h(x - y) = h^{-n} \eta \left( \frac{x - y}{h} \right)$$

(210)
has compact support contained in $\Omega$. Consequently,

$$D^\alpha (\eta_h * u)(x) = h^{-n} \int_\Omega D_x^\alpha \eta \left( \frac{x-y}{h} \right) u(y) \, dy$$

$$= h^{-n} (-1)^{|\alpha|} \int_\Omega D_y^\alpha \eta \left( \frac{x-y}{h} \right) u(y) \, dy$$

$$= h^{-n} \int_\Omega \eta \left( \frac{x-y}{h} \right) D^\alpha u(y) \, dy$$

$$= \eta_h * D^\alpha u(x), \quad (211)$$

from which (209) follows. $\square$

### 3.2. Sobolev Spaces

Suppose that $k \geq 0$ is an integer and $1 \leq p \leq \infty$ is a real number. The Sobolev space $W^{k,p}(\Omega)$ consists of all $L^1_{\text{loc}}(\Omega)$ functions $u : \Omega \to \mathbb{R}$ such that for each multi-index $\alpha$ with $|\alpha| \leq k$, $D^\alpha u$ exists in the weak sense and belongs to $L^p(\Omega)$. The space $W^{k,p}(\Omega)$ is a Banach space with respect to the norm

$$\|u\|_{W^{k,p}(\Omega)} = \left( \sum_{|\alpha|=0}^k \|D^\alpha u\|_p^p \right)^{1/p}, \quad (212)$$

where $\| \cdot \|_p$ refers to the $L^p(\Omega)$ norm

$$\|u\|_p = \left( \int_\Omega u(x) \, dx \right)^{1/p}. \quad (213)$$

The spaces $W^{k,2}(\Omega)$ are of particular importance because they are Hilbert spaces with respect to the inner product

$$(u, v) = \sum_{|\alpha|=0}^k \int_\Omega D^\alpha u(x) \, D^\alpha v(x) \, dx. \quad (214)$$

We will use the special notation $H^k(\Omega)$ to refer to $W^{k,2}(\Omega)$.

We observe that $C_c^\infty(\Omega)$ is a subspace of $W^{k,p}(\Omega)$, and we let $W_0^{k,p}(\Omega)$ be the completion of $C_c^\infty(\Omega)$ with respect to the $W^{k,p}(\Omega)$ norm. Plainly, $W_0^{k,p}(\Omega)$ is a closed subspace of $W^{k,p}(\Omega)$ and hence is a Banach space. We also denote the space $W_0^{k,2}(\Omega)$ by $H_0^k(\Omega)$.

For a function $u \in W^{1,p}(\Omega)$, we use the notation $Du$ to denote the gradient of $u$; that is,

$$Du(x) = \begin{pmatrix} D_1 u(x) \\ D_2 u(x) \\ \vdots \\ D_n u(x) \end{pmatrix}. \quad (215)$$

Moreover, we define

$$\|Du\|_p = \left( \|D_1 u(x)\|_p^p + \cdots + \|D_n u(x)\|_2^p \right)^{1/p}. \quad (216)$$
In a similar fashion, for \( u \in W^{2,p}(\Omega) \) we define \( D^2 u(x) \) to be the Hessian matrix whose \((i,j)\) entry is

\[
D_i D_j u(x)
\]

and denote by \( \|D^2 u\|_p \) the sum

\[
\left( \sum_{i,j=1}^n \|D_i D_j u(x)\|_p^p \right)^{1/p}.
\]

The following theorem shows that the product of a function in \( W^{k,p}(\Omega) \) with a smooth compactly supported function is an element of \( W^{k,p}(\Omega) \). We need this result in order to establish, in Theorem 49, that \( C^\infty_c(\Omega) \) functions are dense in \( W^{k,p}(\Omega) \).

**Theorem 47.** Suppose that \( \Omega \) is an open subset of \( \mathbb{R}^n \), that \( k \geq 0 \) is an integer and \( 1 \leq p \leq \infty \) is a real number, and that \( u \in W^{k,p}(\Omega) \), and that \( \psi \in C^\infty_c(\Omega) \). Then the product \( \psi u \) is an element of \( W^{k,p}(\Omega) \).

**Proof.** We prove the theorem by induction on \( k \). The result obviously holds when \( k = 0 \). We will show that if it holds for \( 0 \leq k < l \), then it holds when \( k = l \). Suppose that \( \varphi \in C^\infty_c(\Omega) \). Since \( \psi \varphi \) is an element of \( C^\infty_c(\Omega) \),

\[
\int_\Omega D^\alpha u(x) \psi(x) \varphi(x) \, dx = (-1)^{|\alpha|} \int_\Omega u(x) D^\alpha (\psi \varphi) (x) \, dx.
\]

We insert the identity

\[
D^\alpha (\psi \varphi)(x) \, dx = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} \psi(x) D^\beta \varphi(x)
\]

into (219) in order to obtain

\[
\int_\Omega D^\alpha u(x) \psi(x) \varphi(x) \, dx = (-1)^{|\alpha|} \int_\Omega u(x) \left( \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} \psi(x) D^\beta \varphi(x) \right) \, dx.
\]

We rearrange (221) as

\[
(-1)^{|\alpha|} \int_\Omega u(x) \psi(x) D^\alpha \varphi(x) \, dx = - \int_\Omega D^\alpha u(x) \psi(x) \varphi(x) \, dx + \sum_{\beta < \alpha} \binom{\alpha}{\beta} \int_\Omega u(x) D^{\alpha-\beta} \psi(x) D^\beta \varphi(x) \, dx.
\]

and apply the induction hypothesis to each of the terms in the sum on the right-hand side of (222) in order to obtain

\[
(-1)^{|\alpha|} \int_\Omega u(x) \psi(x) D^\alpha \varphi(x) \, dx = - \int_\Omega D^\alpha u(x) \psi(x) \varphi(x) \, dx + \sum_{\beta < \alpha} (-1)^{|\beta|} \binom{\alpha}{\beta} \int_\Omega D^\beta \left( u(x) D^{\alpha-\beta} \psi(x) \right) \varphi(x) \, dx.
\]
We conclude that $D^\alpha(\psi u)$ exists and equals
\[-D^\alpha u(x)\psi(x) + \sum_{\beta<\alpha} (-1)^{|\beta|} \binom{\alpha}{\beta} D^\beta \left( u(x)D^{\alpha-\beta} \psi(x) \right). \tag{224}\]

By assumption $D^\alpha u \in L^p(\Omega)$ and our induction hypothesis implies that each of the functions
\[D^\beta \left( u(x)D^{\alpha-\beta} \psi(x) \right) \tag{225}\]
is an element of $L^p(\Omega)$. Since the products of $C^\infty_c(\Omega)$ functions with $L^p(\Omega)$ functions are in $L^p(\Omega)$, we conclude that $D^\alpha(\psi u)$ is an element of $L^p(\Omega)$. \hfill \Box

For $1 \leq p \leq \infty$ a real number and $k \geq 0$ an integer, We denote by $W^{k,p}_{\text{loc}}(\Omega)$ the vector space of functions whose restrictions to any open subset $\Omega' \subset \subset \Omega$ are in $W^{k,p}(\Omega')$. We say that a sequence $\{u_n\} \subset W^{k,p}_{\text{loc}}(\Omega)$ converges to $u \in W^{k,p}_{\text{loc}}(\Omega)$ if $u_n \to u$ in $W^{k,p}(\Omega')$ whenever $\Omega' \subset \subset \Omega$.

### 3.3. Approximation by Smooth Functions

In this section, we will two key results (Theorems 49 and 53) on the approximation of elements of $W^{k,p}(\Omega)$ by smooth functions. They are the mechanism by which we establish many of the basic properties of functions in Sobolev spaces — by first demonstrating that sufficiently smooth functions possess those properties and then appealing to the density of smooth functions in $W^{k,p}(\Omega)$.

**Theorem 48.** Suppose that $\Omega$ is an open set in $\mathbb{R}^n$, and that $\eta_h$ is the standard mollifier. Suppose also that $u \in W^{k,p}_{\text{loc}}(\Omega)$ with $1 \leq p < \infty$ a real number and $k \geq 0$ an integer. Then the sequence of functions $\eta_h \ast u$ converges to $u$ in $W^{k,p}_{\text{loc}}(\Omega)$.

**Proof.** Suppose that $\alpha$ is a multi-index such that $0 \leq |\alpha| \leq k$, and that $\Omega' \subset \subset \Omega$. According to Theorem 46, for all sufficiently small $h$ and all $x \in \Omega'$
\[D^\alpha (\eta_h \ast u) (x) = \eta_h \ast D^\alpha u(x). \tag{226}\]

By integrating both sides of (226) over $\Omega'$, we obtain
\[
\int_{\Omega'} |D^\alpha (\eta_h \ast u - u) (x)|^p \, dx = \int_{\Omega'} |\eta_h \ast D^\alpha u(x) - D^\alpha u(x)|^p \, dx. \tag{227}\]

Now we invoke Theorem 37 in order to conclude that
\[
\lim_{h \to 0} \int_{\Omega'} |\eta_h \ast D^\alpha u(x) - D^\alpha u(x)|^p \, dx = 0. \tag{228}\]

We combine (227) with (228) in order to obtain
\[
\lim_{h \to 0} \|D^\alpha (\eta_h \ast u) - D^\alpha u\|_{L^p(\Omega')} = 0, \tag{229}\]
from which we conclude that $\eta_h \ast u$ converges in $W^{k,p}(\Omega')$ to $u$. \hfill \Box

If $u \in L^p(\Omega)$, then the sequence of $C^\infty(\Omega)$ functions obtained by mollifying the zero extension of $u$ converges to $u$ in $L^p(\Omega)$ norm. The same is not true for $W^{k,p}(\Omega)$ — as the following exercise shows — and we will need to use a somewhat more complicated construction in order to produce a sequence of smooth functions approximating an element of $W^{k,p}(\Omega)$. 
Exercise 20. Suppose that $\Omega = (0, 1)$, that $u : \Omega \to \mathbb{R}$ is the function defined via $u(x) = 1$, and that $\tilde{u} : \mathbb{R} \to \mathbb{R}$ is the zero extension of $u$. Show that $\eta_h * \tilde{u}$ does not converge to $u$ in $W^{1,1}(\Omega)$.

Theorem 49 (Meyers-Serrin). Suppose that $\Omega$ is an open set in $\mathbb{R}^n$, and that $1 \leq p < \infty$ is a real number. Then $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.

Proof. We choose a sequence of open sets $\Omega_1 \subset \subset \Omega_2 \subset \subset \Omega_3 \subset \subset$ such that
\[
\Omega = \bigcup_{k=1}^{\infty} \Omega_k
\] (230)
and we let $\{\psi_k : k = 1, 2, \ldots\}$ be a smooth partition of unity subordinate to the covering (230). That is, $\psi_1, \psi_2, \ldots$ is a sequence in $C^\infty_c(\mathbb{R}^n)$ with the following properties:

1. $\text{supp}(\psi_j) \subset \Omega_j$ for each $j = 1, 2, \ldots$;
2. $0 \leq \psi_j(x) \leq 1$ for all $x \in \mathbb{R}^n$ and $j = 1, 2, \ldots$;
3. for each $x \in \Omega$, there exists compact set $K$ containing on which only a finite number of the functions $\psi_1, \psi_2, \ldots$ are nonzero;
4. $\sum_{j=1}^{\infty} \psi_j(x) = 1$ for all $x \in \Omega$.

We now suppose that $u \in W^{k,p}(\Omega)$ and for each $j = 1, 2, \ldots$ we define $u_j$ by the formula
\[
u_j(x) = u(x)\psi_j(x).
\] (231)
According to Theorem 47, each of the functions $u_j$ is contained in $W^{k,p}(\Omega)$. Moreover,
\[
u(x) = \sum_{j=1}^{\infty} u_j(x) \text{ for all } x \in \Omega
\] (232)
and
\[
\text{supp}(u_j) \subset \Omega_j \text{ for all } j = 1, 2, \ldots
\] (233)
We let $\epsilon > 0$ be an arbitrarily small real number, and denote by $\eta_h$ the standard mollifier. Since $\Omega_{j+1} \subset \subset \Omega$, according to Theorem 48
\[
\lim_{h \to 0} \|\eta_h * u_j - u_j\|_{W^{k,p}(\Omega_{j+1})} = 0
\] (234)
for each $j = 1, 2, \ldots$. Moreover, since $u_j$ is supported in $\Omega_j$, for sufficiently small $h$ the function $\eta_h * u_j$ is supported in $\Omega_{j+1}$. Accordingly, for each $j \geq 1$, there exists a real number $h_j$ such that
\[
\text{supp}(\eta_{h_j} * u_j) \subset \Omega_{j+1}
\] (235)
and
\[
\|\eta_{h_j} * u_j - u_j\|_{W^{k,p}(\Omega_{j+1})} < \frac{\epsilon}{2^j}
\] (236)
The support of $\eta_{h_j} * u_j - u_j$ is also contained in $\Omega_{j+1}$ (the support of $u_j = \psi_j u$ is in $\Omega_j$). It follows from this observation and (236) that
\[
\|\eta_{h_j} * u_j - u_j\|_{W^{k,p}(\Omega_j)} \leq \|\eta_{h_j} * u_j - u_j\|_{W^{k,p}(\Omega_{j+1})} < \frac{\epsilon}{2^j}
\] (237)
Since only a finite number of the functions \( u_1, u_2, \ldots \) are nonzero in each of the sets \( \Omega_j \), only a finite number of the functions \( \eta_{h_1} \ast u_1, \eta_{h_2} \ast u_2, \ldots \) are nonzero in each of the sets \( \Omega_{j+1} \) and

\[
v(x) = \sum_{j=1}^{\infty} \eta_{h_j} \ast u_j(x).
\]

defines a \( C^\infty(\Omega) \) function. We combine (232), (236) and (238) in order to arrive at

\[
\|v - u\|_{W^{k,p}(\Omega)} = \left\| \sum_{j=1}^{\infty} \eta_j \ast u_j - \sum_{j=1}^{\infty} u_j \right\|_{W^{k,p}(\Omega)} \leq \sum_{j=1}^{\infty} \|\eta_j \ast u_j - u_j\|_{W^{k,p}(\Omega)} \leq \sum_{j=1}^{\infty} \frac{\epsilon}{2^j} = \epsilon, \tag{239}
\]

from which the conclusion of the theorem follows.

The Meyers-Serrin theorem cannot be extended to the case \( p = \infty \). To see this, we let \( \Omega = \{ x \in \mathbb{R} : -1 < x < 1 \} \) and \( u(x) = |x| \). We observe that \( u'(x) = x/|x| \) for all \( x \neq 0 \), so \( u \in W^{1,\infty}(\Omega) \). But there is no \( C^1(\Omega) \) function \( \phi \) such that \( \|\phi' - u'\|_\infty < 1/4 \), so \( u \) cannot be the limit of a sequence of infinitely differentiable functions in \( W^{k,p}(\Omega) \).

As an application Theorem 49, we now sharpen Theorem 47.

**Theorem 50.** Suppose that \( \Omega \) is an open set in \( \mathbb{R}^n \), and that \( 1 \leq p < \infty \) and \( k \geq 1 \). Suppose also that \( \psi \in C^{k-1,1}(\Omega) \). Then the mapping

\[
u(x) \to \psi(x) u(x) \tag{240}
\]

is a bounded linear mapping \( W^{k,p}(\Omega) \to W^{k,p}(\Omega) \).

**Proof.** We suppose that \( u \in C^\infty(\Omega) \cap W^{k,p}(\Omega) \), and that \( \alpha \) is a multi-index such that \( |\alpha| \leq k \). We apply the standard chain rule from multivariable calculus in order to conclude that

\[
D^\alpha(\psi u)(x) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} \psi(x) D^\beta u(x) \tag{241}
\]

for almost all \( x \in \Omega \). We observe that since \( \psi \in C^{k-1,1}(\Omega) \), there exists a constant \( C > 0 \) such that

\[
|D^\beta \psi(x)| \leq C \tag{242}
\]

for almost all \( x \in \Omega \). We combine (241) with (242) in order to conclude that

\[
|D^\alpha(\psi u)(x)| \leq C \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} |D^\beta u(x)| \tag{243}
\]

for almost all \( x \in \Omega \). It clearly follows from (243) that \( D^\alpha(\psi u) \in L^p(\Omega) \), and that there exists a constant \( C' \) which depends on \( \alpha \) and \( \psi \) but not \( u \) such that

\[
\|D^\alpha(\psi u)(x)\|_p \leq C' \|u\|_{W^{k,p}(\Omega)}. \tag{244}
\]

We conclude that there exists \( C'' > 0 \) such that

\[
\|\psi u\|_{W^{k,p}(\Omega)} \leq C'' \|u\|_{W^{k,p}(\Omega)} \tag{245}
\]

for all \( u \in C^\infty(\Omega) \cap W^{k,p}(\Omega) \). It now follows from Theorem 49 that the mapping (240) extends to a bounded linear mapping \( W^{k,p}(\Omega) \to W^{k,p}(\Omega) \). Since convergence in \( L^p(\Omega) \) implies pointwise almost everywhere convergence, the mapping obtained by extension coincides with
3.3. APPROXIMATION BY SMOOTH FUNCTIONS

The mapping defined for \( u \in W^{k,p}(\Omega) \) via the formula
\[
  u(x) \to \psi(x)u(x).
\]

(246)

The following two theorems are established in much the same fashion as Theorem 50 — by applying the theorems of Section 2.6.4 and then appealing to Theorem 50.

**Theorem 51.** Suppose that \( \Omega \) is a subset of \( \mathbb{R}^n \), and that \( \Omega' \) is a subset of \( \mathbb{R}^m \). Suppose also that \( k \geq 1 \) is an integer, that \( 1 \leq p < \infty \) is a real number, and that \( \psi : \Omega \to \Omega' \) is a \( C^{k-1,1} \) mapping. Then the map
\[
  u(x) \to u(\psi(x))
\]

(247) is a bounded linear mapping \( W^{k,p}(\Omega') \to W^{k,p}(\Omega) \).

**Theorem 52.** Suppose that \( \Omega \) is a subset of \( \mathbb{R}^n \), and that \( \Omega' \) is a subset of \( \mathbb{R}^m \). Suppose also that \( k \geq 1 \) is an integer, that \( 1 \leq p < \infty \) is a real number, and that \( \psi : \Omega \to \Omega' \) is a bijective mapping such that \( \psi \) and \( \psi^{-1} \) are \( C^{k-1,1} \) mappings. Then there exists a constant \( C > 0 \) such that
\[
  C^{-1}\|u \circ \psi\|_{W^{k,p}(\Omega')} \leq \|u\|_{W^{k,p}(\Omega')} \leq C\|u \circ \psi\|_{W^{k,p}(\Omega)}
\]

(248) for all \( u \in W^{k,p}(\Omega') \). That is, the \( W^{k,p}(\Omega') \) norm of \( u(x) \) is equivalent to the \( W^{k,p}(\Omega) \) norm of the composition \( u(\psi(x)) \).

The space \( C^\infty(\overline{\Omega}) \) is always contained in \( W^{k,p}(\Omega) \) and in light of Theorem 49 it is natural to ask if \( C^\infty(\overline{\Omega}) \) is dense in \( W^{k,p}(\Omega) \). To see that this is not always the case, we let \( \Omega = \{(x,y) \in \mathbb{R}^2 : 0 < |x| < 1 \text{ and } 0 < y < 1\} \) and define the function \( f : \Omega \to \mathbb{R} \) via the formula
\[
  f(x) = \begin{cases} 
  1 & \text{if } x > 0 \\
  0 & \text{if } x < 0.
  \end{cases}
\]

(249)

Then \( f \in W^{1,p}(\Omega) \) for all integers \( p \geq 1 \), but there is plainly no sequence of functions in \( C^1(\overline{\Omega}) \) which converges to \( f \). However, as we now show, it is the case under mild regularity assumptions on the boundary of \( \Omega \).

**Theorem 53.** Suppose that \( \Omega \) is a bounded open set with continuous boundary. Suppose also that \( 1 \leq p < \infty \) is a real number and \( k \geq 0 \) is an integer. Then \( C^\infty(\overline{\Omega}) \) is dense in \( W^{k,p}(\Omega) \).

**Proof.** We observe that since the boundary of \( \partial \Omega \) is continuous, \( \Omega \) has the segment property (see Section 2.8). In particular, for each \( x \in \partial \Omega \), there exists an open ball \( U_x \) centered at \( x \) and a vector \( \gamma_x \in \mathbb{R}^n \) such that \( y + t\gamma_x \in \Omega \) for all \( y \in \overline{\Omega} \cap U_x \) and all \( 0 < t < 1 \). Then
\[
  \partial \Omega \subseteq \bigcup_{x \in \partial \Omega} 1/2U_x,
\]

(250) where \( 1/2U_x \) denotes the open ball centered at \( x \) whose radius is half that of the open ball \( U_x \), us an open covering of the compact set \( \partial \Omega \). Consequently, there exists a finite collection
of the balls \( U_{x_1}, \ldots, U_{x_m} \) such that
\[
\partial \Omega \subset \bigcup_{j=1}^{m} 1/2U_{x_j}.
\] (251)
For each \( j = 1, \ldots, m \), we set \( U_j = U_{x_j} \) and \( V_j = 1/2U_{x_j} \cap \Omega \) and let \( \gamma_j \) denote the vector \( \gamma_{x_j} \). We also choose an open set \( V_0 \) such that \( V_0 \subset \subset \Omega \) and
\[
\Omega \subset \bigcup_{j=0}^{m} V_j.
\] (252)
Next, we choose a smooth partition of unity \( \psi_0, \ldots, \psi_m \) subordinate to the covering
\[
\Omega \subset V_0 \cup \bigcup_{j=1}^{m} 1/2U_j
\] (253)
so that
\[
\begin{align*}
(1) & \quad \psi_0 \in C_c^\infty (V_0); \\
(2) & \quad \psi_j \in C_c^\infty (1/2U_j) \text{ for each } j = 1, \ldots, m; \\
(3) & \quad 0 \leq \psi_j (x) \leq 1 \text{ for all } j = 0, 1, \ldots, m \text{ and } x \in \mathbb{R}^n; \\
(4) & \quad \sum_{j=0}^{m} \psi_j (x) = 1 \text{ for all } x \in \Omega.
\end{align*}
\]
Moreover, for each \( j = 0, 1, \ldots, m \) we define \( u_j : V_j \to \mathbb{R} \) via the formula
\[
u_j(x) = \psi_j(x) u(x).
\] (254)
Note that the support of \( \psi_j \) is not necessarily contained in \( V_j = 1/2U_j \cap \Omega \).

We now fix \( \epsilon > 0 \); we will construct a function \( w \in C^\infty (\bar{\Omega}) \) such that
\[
\|u - w\|_{W^{k,p}(\Omega)} < \epsilon.
\] (255)
To that end, we define the function \( w_0 : V_0 \to \mathbb{R} \) via the formula
\[
w_0(x) = \eta_{h_0} \ast u_0 (x),
\] (256)
where \( \eta_h \) denotes the standard mollifier and \( h_0 \) is chosen so that
\[
\|w_0 - u_0\|_{W^{k,p}(V_0)} < \frac{\epsilon}{2}.
\] (257)
For each \( j = 1, \ldots, m \), all sufficiently small \( h > 0 \) and all \( x \in \nabla_j \), we define \( u_{j,h} : \nabla_j \to \mathbb{R} \) by the formula
\[
u_{j,h}(x) = u_j (x + 2h\gamma_j)
\] (258)
and \( v_{j,h} \in C^\infty (\overline{V_j}) \) via the formula
\[
v_{j,h}(x) = \eta_h \ast u_{j,h} (x).
\] (259)
We observe that
\[
\|D^\alpha v_{j,h} - D^\alpha u_j\|_{L^p(V_j)} \leq \|D^\alpha v_{j,h} - D^\alpha u_{j,h}\|_{L^p(V_j)} + \|D^\alpha u_{j,h} - D^\alpha u_j\|_{L^p(V_j)}
\] (260)
for all multi-indices \( |\alpha| \leq k \). Since
\[
D^\alpha u_{j,h}(x) = (D^\alpha u_j)(x + 2h\gamma_j)
\] (261)
and translation is continuous in the $L^p$ norm,

$$\|D^\alpha u_{j,h} - D^\alpha u_j\|_{L^p(V_j)} \to 0 \text{ as } h \to 0$$

(262) whenever $|\alpha| \leq k$. Moreover, a simple modification of the standard argument showing that the mollification of a function converges in $L^p$ norm shows that

$$\|D^\alpha v_{j,h} - D^\alpha u_{j,h}\|_{L^p(V_j)} \to 0 \text{ as } h \to 0$$

(263) for all $|\alpha| \leq k$. We conclude that for each $j = 1, \ldots, m$ there exists a function $w_j$ in $C^\infty(V_j)$ such that

$$\|w_j - u_j\|_{W^{k,p}(V_j)} \leq \frac{\epsilon}{2j+1}.$$  

(264)

We now define the function $w \in C^\infty(\Omega)$ via the formula

$$w(x) = \sum_{j=0}^m \psi_j(x)w_j(x).$$  

(265)

We combine (254), (257) and (264) in order to conclude that

$$\|u - w\|_{W^{k,p}(\Omega)} = \left\| \sum_{j=0}^m \psi_j u - \sum_{j=0}^m \psi_j w \right\|_{W^{k,p}(\Omega)}$$

\begin{align*}
&\leq \sum_{j=0}^m \|\psi_j u - \psi_j w\|_{W^{k,p}(\Omega)} \\
&= \sum_{j=0}^m \|u_j - w_j\|_{W^{k,p}(V_j)} \\
&\leq \epsilon.
\end{align*}

(266)

**Exercise 21.** Suppose that $u \in L^p(\mathbb{R}^n)$, where $1 \leq p < \infty$. Suppose also that $u_h$ is defined via the formula $u_h(x) = u(x + h)$, and that $\eta_h$ is the standard mollifier. Show that

$$\|\eta_h * u_h - u\|_p \to 0 \text{ as } h \to 0.$$  

(267)

## 3.4. Traces

Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded open set. We are ultimately interested in solving boundary value problems given on $\Omega$. However, it is not a priori clear that the notion of the “boundary values on $\partial \Omega$” of a function $u$ in $W^{k,p}(\Omega)$ is well-defined. After all, $\partial \Omega$ is typically a set of measure 0 in $\mathbb{R}^n$ (although there are open sets in $\mathbb{R}^n$ whose boundaries are of positive measure in $\mathbb{R}^n$) and $u$, as an element of $L^p(\Omega)$, is only defined almost everywhere in $\Omega$. However, using Theorem 53 it is easy to establish that there is a reasonable notion of “boundary values on $\partial \Omega$” for functions in $W^{1,p}(\Omega)$, assuming that the boundary of $\Omega$ is sufficiently regular.
Theorem 54. Suppose that $1 \leq p < \infty$ is a real number, that $k \geq 0$ is an integer, that $\Omega$ is an open set, that

$$\Omega \subset \bigcup_{j=1}^{\infty} U_j$$

is a covering of $\Omega$ by open sets, and that $\{\psi_j\}$ is a smooth partition of unity subordinate to the covering (268). Then $u$ is an element of $W^{k,p}(\Omega)$ if and only if the function

$$\sum_{j=1}^{\infty} \psi_j(x) u(x)$$

is an element of $W^{k,p}(\Omega)$. Moreover, then there exists a constant $C > 0$ such that

$$C^{-1} \|u\|_{W^{k,p}(\Omega)} \leq \left\| \sum_{j=1}^{\infty} \psi_j u \right\|_{W^{k,p}(\Omega)} \leq C \|u\|_{W^{k,p}(\Omega)}$$

for all $u \in W^{k,p}(\Omega)$.

Theorem 55. Suppose that $\Omega \subset \mathbb{R}^n$ is a $C^{0,1}$ domain, and that $1 \leq p < \infty$ is a real number. Then there exists a continuous linear mapping

$$\mathcal{T} : W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

such that

$$\mathcal{T}[u] = u|_{\partial\Omega}$$

for all $u \in C(\overline{\Omega})$.

Proof. Since $\Omega$ is a $C^{0,1}$ domain, there exists a covering of $\partial\Omega$ by open sets

$$\partial\Omega \subset \bigcup_{j=1}^{m} U_j$$

with the property that for each $j = 1, \ldots, m$ there exist an open set $V_j \subset \mathbb{R}^n$ and a mapping $\psi_j : V_j \to U_j$ such that

1. $\psi_j$ and $\psi_k^{-1}$ are Lipschitz mappings;
2. $U_j \cap \Omega = \psi_j(V_j \cap \{(x_1, \ldots, x_n) : x_n > 0\})$; and
3. $U_j \cap \partial\Omega = \psi_j(V_j \cap \{(x_1, \ldots, x_n) : x_n = 0\})$.

We choose a smooth partition of unity $\{\gamma_j\}$ subordinate to the covering (273) and define, for each $j = 1, \ldots, m$, the function $\eta_j : V_j \to \mathbb{R}$ via the formula

$$\eta_j(x) = \gamma_j(\psi_j(x))$$

and the function $u_j : V_j \to \mathbb{R}$ via the formula

$$u_j(x) = u(\psi_j(x)).$$

From Theorems 54 and 52 we conclude that if there exists $C > 0$ such that

$$\|u_j\|_{L^p(V_j \cap \{x_n = 0\})} \leq C \|u_j\|_{W^{k,p}(V_j \cap \{x_n > 0\})}$$

for all $j = 1, \ldots, m$.
then there exists $C' > 0$ such that
\[ \|u\|_{L^p(\partial \Omega)} \leq C \|u\|_{W^{k,p}(\Omega)}. \] (277)
That is, through the assumption that $\partial \Omega$ is Lipschitz and with the help of the partition of unity $\{\psi_j\}$ we have reduced the theorem to a local calculation in $\mathbb{R}^n$.

We observe that the support of $u_j$ is contained in $V_j$, so that when we apply the divergence theorem we obtain
\[ \int_{V_j \cap \{x_n = 0\}} |u_j(x)|^p \, dx = -\int_{V_j \cap \{x_n > 0\}} \frac{\partial}{\partial x_n} (|u_j(x)|^p) \, dx \]
\[ = -\int_{V_j \cap \{x_n > 0\}} p |u_j(x)|^{p-1} \text{sign}(u_j(x)) \frac{\partial u_j(x)}{\partial x_n} + |u_j(x)|^p \, dx. \] (278)
By letting
\[ q = \frac{1}{1 - \frac{1}{p}}, \quad a = \left| \frac{\partial u_j(x)}{\partial x_n} \right| \quad \text{and} \quad b = |u_j(x)|^{p-1} \] (279)
in the inequality $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$, which holds for all $a, b > 0$ and all $1 < p, q < \infty$ such that $p^{-1} + q^{-1} = 1$, we obtain
\[ \left| \frac{\partial u_j(x)}{\partial x_n} \right| |u_j(x)|^{p-1} \leq \frac{1}{p} \left| \frac{\partial u_j(x)}{\partial x_n} \right|^p + \frac{1}{q} |u_j(x)|^{q(p-1)} \]
\[ = \frac{1}{p} \left| \frac{\partial u_j(x)}{\partial x_n} \right|^p + \frac{1}{q} |u_j(x)|^p. \] (280)
We conclude, by inserting (280) into (278), that for each $j = 1, \ldots, m$ there exists $C_j > 0$ such that
\[ \int_{\{x_n = 0\} \cap V_j} |u_j(x)|^p \, dx \leq C_j \|u_j\|_{W^{1,p}(V_j \cap \{x_n > 0\})}. \] (281)

It now follows from Theorem 53 that $\mathcal{T}$ extends by continuity to a bounded linear mapping $W^{1,p}(\Omega) \rightarrow W^{1,p}(\Omega)$. If $u \in C(\overline{\Omega}) \cap W^{1,p}(\Omega)$, then the sequence $\{u_m\}$ of $C^\infty(\overline{\Omega})$ functions constructed in Theorem 53 converge uniformly on $\overline{\Omega}$ to $u$. We conclude from this observation that $u_m \rightarrow u$ on the boundary of $\partial \Omega$, so that $\mathcal{T}(u) = u |_{\partial \Omega}$. \hfill \square

The mapping $\mathcal{T}$ is called the trace operator, $\mathcal{T}[u]$ is known as the trace of the function $u$, and the image of $W^{k,p}(\Omega)$ under the mapping $\mathcal{T}$ is known as the trace space of $W^{k,p}(\Omega)$. The mapping $\mathcal{T}$ is not a surjection onto $L^p(\Omega)$. We will later characterize the trace space of $H^1(\Omega)$ when $\Omega$ is a Lipschitz domain and we will state without proof a much more general theorem which characterizes the trace space of $W^{k,p}(\Omega)$ under suitable smoothness assumptions on $\partial \Omega$.

**Theorem 56.** Suppose that $\Omega$ is a Lipschitz domain in $\mathbb{R}^n$, and that $1 \leq p < \infty$. Then $u \in W^{1,p}_0(\Omega)$ if and only if $\mathcal{T}[u] = 0$.

**Proof.** If $u \in C^\infty_c(\Omega)$, then it is a consequence of (55) that $\mathcal{T}[u] = 0$. Since functions in $W^{k,p}_0(\Omega)$ are the limits in $W^{k,p}(\Omega)$ norm of $C^\infty_c(\Omega)$ functions and $\mathcal{T}$ is continuous, it follows that the trace of a function in $W^{1,p}_0(\Omega)$ is 0.
By Theorems 50 and 55, in order to show that if $u$ is an element of such that $\mathcal{T}[u] = 0$, then $u \in W^{1,p}_0(\Omega)$, it suffices to show that if $u \in C^1(\bar{\Omega})$ such that $\mathcal{T}[u] = 0$ then $u \in W^{1,p}_0(\Omega)$. Moreover, via a localization argument which is virtually identical to that used in the proof Theorem 55, in order to establish this it suffices to show that if $u$ is an element of $C^1(\bar{V})$ such that $\text{supp}(u) \subset 1/2U$ and $\mathcal{T}[u] = 0$, then $u \in W^{1,p}_0(\Omega \cap \{x_n < 0\})$.

To this end, we let $G$ be a $C^1(\mathbb{R})$ function such that

$$G(t) = \begin{cases} 0 & \text{if } |t| \leq 1 \\ t & \text{if } |t| \geq 2 \end{cases}$$

and define a sequence $\{u_m\}$ of functions via the formula

$$u_m(x) = \frac{1}{m} G(m \cdot u(x)).$$

We observe that each of the $u_m$ is an element of $C^1(V)$, and that

$$\text{supp}(u_m) = \{x \in 1/2V \cap \{x_n > 0\} : u(x) > 1/m\}.$$  \hspace{1cm} (284)

Because $u(x) = 0$ for all $x \in V \cap \{x_n = 0\}$ and $\text{supp}(u) \subset 1/2U$, the support of each $u_m$ is contained in the interior of $V$. Moreover, it is easy to verify that $u_m \to u$ in $W^{1,p}(V)$. So $u$ is the limit of a sequence of functions $\{u_m\}$ which are in $C^1_c(V)$. By mollifying the $u_m$, we obtain a sequence of $C^\infty_c(V)$ functions which converge to $u$ in $W^{1,p}(V)$. \hspace{1cm} \Box

Suppose that $\Omega$ is $C^{0,1}$ domain and that $u \in C^1(\bar{\Omega})$. Then not only does $u$ have a well-defined trace on $\partial \Omega$, it admits a derivative with respect to the outward-pointing unit normal vector on $\partial \Omega$. Let $U$ be an open ball $U$ which intersects $\partial \Omega$ such that there exist an open set $V$ in $\mathbb{R}^n$ and a bijective mapping $\psi : V \to U$ with the following properties:

1. $\psi$ and $\psi^{-1}$ are $C^{k-1,1}$ mappings;
2. $\Omega \cap U = \psi(V \cap \{x_n > 0\})$; and
3. $\partial \Omega \cap U = \psi(V \cap \{x_n = 0\})$.

By normalizing the map $\psi$, we may assume that

$$\left\| \frac{\partial \psi}{\partial x_n}(x) \right\| = 1$$

for all $x \in V$. Since the composition $u \circ \psi$ is an element of $C^{0,1}(\bar{V} \cap \{x_n > 0\})$, the derivative

$$\frac{\partial}{\partial x_n} u \circ \psi$$

extends to the set $V \cap \{x_n = 0\}$. For each point $y \in U \cap \partial \Omega$, there exists $x \in V \cap \{x_n = 0\}$ such that $\psi(x) = y$. We take the value of

$$\frac{\partial u}{\partial \nu}$$

(287)
3.5. Difference Quotients

In this section, we discuss a mechanism for establishing the weak differentiability of a function and for estimating the $L^p$ norms of its derivatives. The results of this section will be used to establish the regularity of weak solutions of elliptic boundary value problems in Chapter ??.

Suppose that $\Omega$ is a bounded open set in $\mathbb{R}^n$. We denote by $e_i$ the vector whose $i$th component is 1 and whose remaining components are 0. For each $i = 1, \ldots, n$, we define the difference quotient of $u$ in the direction $e_i$ via the formula

$$
\Delta^h_i u(x) = \frac{u(x + he_i) - u(x)}{h}.
$$

We denote by $\Delta^h u(x)$ the vector

$$
\Delta^h u(x) = \left( \begin{array}{c} 
\Delta^h_1 u(x) \\
\Delta^h_2 u(x) \\
\vdots \\
\Delta^h_n u(x)
\end{array} \right)
$$

and define

$$
\|\Delta^h u\|_p = \left( \|\Delta^h_1 u(x)\|_p^p + \cdots + \|\Delta^h_n u(x)\|_p^p \right)^{1/p}
$$

for $1 \leq p < \infty$ and

$$
\|\Delta^h u\|_\infty = \|\Delta^h_1 u(x)\|_\infty + \cdots + \|\Delta^h_n u(x)\|_\infty.
$$

**Theorem 57.** Suppose that $\Omega$ is a bounded open set in $\mathbb{R}^n$, that $p \geq 1$ is a real number, and that $1 \leq i \leq n$ is an integer. Suppose also that $\Omega'$ is an open set such that $\Omega' \subset \subset \Omega$. Then

$$
\|\Delta^h_i u(x)\|_{L^p(\Omega')} \leq \|D_i u\|_{L^p(\Omega)}
$$

at the point $y$ to be

$$
\frac{\partial}{\partial x_n} u \circ \psi(x).
$$

It is easy to verify that this definition is independent of the choice of $U$ and $\psi$. Note, though, that the map $\frac{\partial}{\partial \nu}$ does not extend to a mapping from $W^{1,p}(\Omega)$ into any reasonable class of functions on $\partial \Omega$. To see this, we observe that there are functions $u \in C^1(\overline{\Omega})$ such that

$$
u u = 0
$$

and

$$\frac{\partial u}{\partial \nu} \neq 0.
$$

But, if $u \in W^{1,p}(\Omega)$, then (289) implies that $u \in W^{1,p}_0(\Omega)$ and hence is the limit of a sequence $\{\varphi_k\}$ of $C^\infty_c(\Omega)$ functions. But the normal derivative of each of the $\varphi_k$ is 0, so we cannot have

$$
\frac{\partial \varphi_k}{\partial \nu} \to \frac{\partial u}{\partial \nu}
$$

in any reasonable norm. It is easy to verify, however, that $\frac{\partial}{\partial \nu}$ extends to a bounded linear mapping $W^{2,p}(\Omega) \to L^p(\partial \Omega)$. 

3.5. Difference Quotients

In this section, we discuss a mechanism for establishing the weak differentiability of a function and for estimating the $L^p$ norms of its derivatives. The results of this section will be used to establish the regularity of weak solutions of elliptic boundary value problems in Chapter ??.
otherwise we take the $p$ from which we conclude that (296) holds when for all $x$ theorem implies that for each $\Omega$ whenever $0 < h < \text{dist}(\Omega', \partial \Omega)$.

**Proof.** We first suppose that $u \in W^{k,p}(\Omega) \cap C^\infty(\Omega)$. Then

$$
\Delta_i^h u(x) = \frac{u(x + he_i) - u(x)}{h} = \frac{1}{h} \int_0^h D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n) \, dt \tag{297}
$$

for all $x \in \Omega'$. If $p = \infty$, then by taking absolute values on both sides of (297) we obtain

$$
|\Delta_i^h u(x)| \leq \sup_{x \in \Omega', 0 < t < \text{dist}(\Omega', \partial \Omega)} |D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)| \leq \sup_{x \in \Omega} |D_i u(x)|, \tag{298}
$$

from which we conclude that (296) holds when $p = \infty$ and $u$ is infinitely differentiable. Otherwise, we take the $p^{th}$ power of both sides of (297) in order to obtain

$$
|\Delta_i^h u(x)|^p = \frac{1}{h^p} \left| \int_0^h D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n) \, dt \right|^p \leq \frac{1}{h^p} \left( \int_0^h |D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)| \, dt \right)^p, \tag{299}
$$

We see from Hölder’s inequality that

$$
\int_0^h |D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)| \, dt \leq h^{1/q} \left( \int_0^h |D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)|^q \, dt \right)^{1/p}, \tag{300}
$$

where

$$
\frac{1}{p} + \frac{1}{q} = 1. \tag{301}
$$

By inserting (300) into (299), we obtain

$$
|\Delta_i^h u(x)|^p \leq h^{p/q - p} \int_0^h |D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)|^p \, dt \tag{302}
$$

$$
\leq \frac{1}{h} \int_0^h |D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)|^p \, dt.
$$

Note that $p/q - p = -1$ follows from (301). Since $D_i u$ is continuous, the integral mean value theorem implies that for each $x \in \Omega'$, there exists $0 \leq \xi_x \leq h$ such that

$$
\frac{1}{h} \int_0^h |D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)|^p \, dt = |D_i u(x_1, \ldots, x_{i-1}, x_i + \xi_x, x_{i+1}, \ldots, x_n)|^p \tag{303}
$$

$$
= |D_i u(x + \xi_x e_i)|^p.
$$
We insert (303) into (302) and integrate over $\Omega'$ in order to obtain
\[
\int_{\Omega'} |\Delta^h_i u(x)|^p \, dx = \int_{\Omega'} |D_i u(x + \xi_i e_i)|^p \, dx \leq \int_{\Omega} |D_i u(x)|^p \, dx,
\]
(304)
from which we conclude that (296) holds for all $1 \leq p < \infty$ as well as $p = \infty$ when $u$ is in $C^\infty(\Omega) \cap W^{k,p}(\Omega)$. That (296) holds for arbitrary $u \in W^{k,p}(\Omega)$ now follows from Theorem 48 — that is, the observation that $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.

Note that $\Delta^h_i u$ is an element of $L^p(\Omega')$ when $u \in L^p(\Omega)$ whenever $0 < h < \text{dist}(\Omega', \partial \Omega)$; in fact,
\[
\left\| \frac{u(\cdot + he_i) - u(\cdot)}{h} \right\|_{W^{k,p}(\Omega')} \leq \frac{2}{h} \|u\|_{W^{k,p}(\Omega)}.
\]
(305)

Theorem 57 is useful because it gives us a bound on the $L^p(\Omega')$ norm of $\Delta^h_i u$ which is independent of $h$. We now establish that the converse also holds; that is, if the $L^p$ norm of $\Delta^h_i u$ is bounded independently of $h$, then the weak derivative $D_i u$ exists and satisfies the same bound.

**Theorem 58.** Suppose that $\Omega$ is a bounded open set in $\mathbb{R}^n$, that $1 < p < \infty$ is a real number, that $1 \leq i \leq n$ is an integer, and that $u \in L^p(\Omega)$. Suppose also that there exists a constant $C > 0$ such that whenever $\Omega'$ is an open set with $\Omega' \subset \subset \Omega$ and $0 < h < \text{dist}(\Omega', \partial \Omega)$,
\[
\|\Delta^h_i u\|_{L^p(\Omega')} \leq C.
\]
(306)
Then the weak derivative $D_i u$ exists in $\Omega$ and
\[
\|D_i u\|_{L^p(\Omega)} \leq C.
\]
(307)

**Proof.** We let $i$ be any integer $1 \leq i \leq n$ and choose a sequence $\Omega_1 \subset \Omega_2 \subset \cdots$ of open sets contained in $\Omega$ such that
\[
\Omega = \bigcup_{j=1}^{\infty} \Omega_j.
\]
(308)
Since bounded sets in $L^p(\Omega_1)$ are weakly compact (see Theorem 29 in Section 2.5) and
\[
\|\Delta^h_i u\|_{L^p(\Omega_1)} \leq C
\]
(309)
for all sufficiently small $h$, there exists a sequence $\{h_1(j)\}_{j=1}^{\infty}$ converging to 0 and a function $v_1 \in L^p(\Omega_1)$ such that $\|v_1\|_{L^p(\Omega_1)} \leq C$, and
\[
\lim_{j \to \infty} \int_{\Omega_1} \Delta^h_i h_1(j) u(x) \psi(x) \, dx = \int_{\Omega_1} v_1(x) \psi(x) \, dx
\]
(310)
for all $\psi \in C_c^\infty(\Omega_1)$. We apply the same logic in order to conclude that there is a subsequence $\{h_2(j)\}$ of $\{h_1(j)\}$ — that is, there exists a function $i : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that $h_2(j) = h_1(i(j))$ for all $j \geq 1$ — and a function $v_2 \in L^2(\Omega_2)$ such that $\|v_2\|_{L^p(\Omega_2)} \leq C$ and
\[
\lim_{j \to \infty} \int_{\Omega_2} \Delta^h_i h_2(j) u(x) \psi(x) \, dx = \int_{\Omega_2} v_2(x) \psi(x) \, dx
\]
(311)
for all $\psi \in C_c^\infty(\Omega_2)$. By the uniqueness of weak limits, $v_2(x) = v_1(x)$ for almost all $x \in \Omega_1$. Consequently, we may replace $v_1$ and $v_2$ by a single function $v \in L^p(\Omega_2)$. Continuing in this fashion, we obtain functions $v \in L^p(\Omega)$ and $h_k(j)$ such that $\|v\|_{L^p(\Omega)} \leq C$, $h_{k+1}(j)$ is a subsequence of $h_k(j)$ and

$$
\lim_{j \to \infty} \int_{\Omega_k} \Delta^{h_k(j)} u(x) \psi(x) \, dx = \int_{\Omega_k} v(x) \psi(x) \, dx
$$

for all $\psi \in C_c^\infty(\Omega_k)$. We now define a sequence $\{s_k\}$ via the formula

$$
s_k = h_k(k);
$$

that is, $s_k$ is the diagonalization of $h_k(j)$.

Now suppose that $\psi \in C_c^\infty(\Omega)$. Since $s_k \to 0$, there exists $l$ such that $s_k < \text{dist}(\text{supp}(\psi), \partial \Omega)$ for all $k \geq l$. Consequently,

$$
\int_{\Omega} \Delta_s^{s_k} u(x) \psi(x) \, dx = \int_{\text{supp}(\psi)} \Delta_s^{s_k} u(x) \psi(x) \, dx
$$

is well-defined for $k \geq l$ and

$$
\lim_{k \to \infty} \int_{\Omega} \Delta_s^{s_k} u(x) \psi(x) \, dx = \int_{\Omega} v(x) \psi(x) \, dx.
$$

But for $l \geq k$, we also have

$$
\lim_{k \to \infty} \int_{\Omega} \Delta_s^{s_k} u(x) \psi(x) \, dx = \lim_{k \to \infty} \int_{\Omega} u(x) \Delta_s^{-s_k} \psi(x) \, dx = -\int_{\Omega} u(x) D_i \psi(x) \, dx.
$$

From (315) and (316) we obtain

$$
\int_{\Omega} u(x) D_i \psi(x) \, dx = -\int_{\Omega} v(x) \psi(x) \, dx.
$$

Since $\psi$ is an arbitrary element of $C_c^\infty(\Omega)$, we conclude $v$ is the $i^{th}$ weak derivative of $u$.

We close this section by characterizing the spaces $W^{1,\infty}_{\text{loc}}(\Omega)$ for arbitrary open sets in $\mathbb{R}^n$ and $W^{1,\infty}(\Omega)$ in the event that $\Omega$ is a Lipschitz domain.

**Theorem 59.** Suppose that $\Omega$ is an open subset of $\mathbb{R}^n$. Then $W^{1,\infty}_{\text{loc}}(\Omega) = C^{0,1}(\Omega)$.

**Proof.** We suppose first that $u \in C^{0,1}(\Omega)$, and that $1 \leq i \leq n$ is an integer. If $\Omega' \subset \subset \Omega$, then for all $0 < h < \text{dist}(\Omega', \partial \Omega)$

$$
\|\Delta_i^h u\|_{L^\infty(\Omega')} \leq C,
$$

where $C$ is the Lipschitz constant for $u$ in $\Omega'$. Since $\Omega'$ is bounded, (318) implies that the sequence $\{\Delta_i^h u\}$ is bounded in $L^2(\Omega')$. Consequently, there is a sequence $h_j \to 0$ and a function $v \in L^2(\Omega')$ such that $\Delta_i^{h_j} u \rightharpoonup v$ weakly in $L^2(\Omega')$. In particular,

$$
\int_{\Omega'} \Delta_i^{h_j} u(x) \varphi(x) \, dx \to \int_{\Omega'} v(x) \varphi(x) \, dx
$$

for all $\varphi \in C_c^\infty(\Omega')$. We observe that

$$
\int_{\Omega'} \Delta_i^{h_j} u(x) \varphi(x) \, dx = -\int_{\Omega'} u(x) \Delta_i^{-h_j} \varphi(x) \, dx \to -\int_{\Omega'} u(x) D_i \varphi(x) \, dx.
$$
3.5. Difference Quotients

We combine (319) and (320) in order to obtain
\[ \int_{\Omega'} u(x)D_i \varphi(x) \, dx = - \int_{\Omega'} v(x) \varphi(x) \, dx, \]  \hspace{1cm} (321)
from which we conclude that \( D_i u = v \). From (318) and (319) we see that
\[ \left| \int_{\Omega'} v(x) \varphi(x) \, dx \right| = \left| \lim_{j \to \infty} \int_{\Omega'} \Delta_{h_j} u(x) \varphi(x) \, dx \right| \leq C \| \varphi \|_{L^1(\Omega')} \]  \hspace{1cm} (322)
for all \( \varphi \in L^1(\Omega') \). We conclude that \( v \in L^\infty(\Omega') \) (see, for instance, Theorem 6.13 in \[4\]).

We now suppose that \( u \in W^{1,\infty}_{\text{loc}}(\Omega) \), and that \( \Omega' \) is an open ball contained in \( \Omega \). For each \( 0 < h < \text{dist}(\Omega', \partial \Omega) \), we define \( u_h \) via the formula
\[ u_h(x) = \eta_h * u(x), \]  \hspace{1cm} (323)
where \( \eta_h \) denotes the standard mollifier (as usual). Since \( u \in L^p(\Omega') \), the sequence \( u_h \) converges to \( u \) for almost all \( x \in \Omega' \); in fact, in converges at every point \( x \) in the Lebesgue set \( L(f) \) of \( f \). We apply Theorem 46 in order to see that
\[ \| D u_h \|_{L^\infty(\Omega')} = \| D (\eta_h * u) \|_{L^\infty(\Omega')} \leq \| \eta_h \|_1 \| D u \|_{L^\infty(\Omega')} = \| D u \|_{L^\infty(\Omega')} < \infty \]  \hspace{1cm} (324)
for all \( 0 < h < \text{dist}(\Omega', \partial \Omega) \). Note that in (324) we use \( D u \) to refer to the weak gradient of \( u \). Since \( u_h \in C^\infty(\Omega') \),
\[ u_h(x) - u_h(y) = \int_0^1 D u_h(y + t(x - y)) \, dt \cdot (x - y) \]  \hspace{1cm} (325)
for all \( x, y \in \Omega' \). We conclude from (325) that
\[ |u_h(x) - u_h(y)| \leq \| D u \|_{L^\infty(\Omega')} |x - y| \]  \hspace{1cm} (326)
for all \( x, y \in \Omega' \) and all sufficiently small \( h \). By taking the limit as \( h \to 0 \) in (326), we see that
\[ |u(x) - u(y)| \leq \| D u \|_{L^\infty(\Omega')} |x - y| \]  \hspace{1cm} (327)
for all \( x, y \) in the Lebesgue set of \( f \). It follows that \( u \) agrees almost everywhere with a function \( u^* \) which is Lipschitz continuous in \( \Omega' \). Note that we define \( u^* \) as follows. For each \( x \notin L(f) \), we choose a sequence \( \{x_n\} \) in \( L(f) \) such that \( x_n \to x \). From (327), we conclude that \( \{u(x_n)\} \) is Cauchy and has a limit. We set \( u^*(x) = \lim_n u(x_n) \). This uniquely defines a representation of \( u \) in \( C^{0,1}(\Omega') \).

By combining Theorem 59 with Theorem 59, we obtain the following:

**Theorem 60.** Suppose that \( \Omega \subset \mathbb{R}^n \) is a Lipschitz domain. Then \( W^{1,\infty}(\Omega) = C^{0,1}(\Omega) \).
Second Order Linear Elliptic Boundary Value Problems

In this chapter, we introduce variational formulations of certain second order linear elliptic boundary value problems and discuss their solvability.

4.1. Variational Formulations

Suppose that $L$ is a differential operator of the form

$$L[u](x) = -a^{ij}(x)D_iD_ju(x) + b^i(x)D_iu(x) + c(x)u(x),$$

and that $u$ is a classical solution of the equation

$$L[u](x) = f(x)$$

in the domain $\Omega \subset \mathbb{R}^n$ which vanishes on the boundary of $\Omega$. By applying the divergence theorem (i.e., integrating by parts) we see that

$$\int_{\Omega} D_ju(x)D_i\left(a^{ij}(x)v(x)\right) + b_i(x)D^iu(x)v(x) + c(x)u(x)v(x)\, dx = \int_{\Omega} f(x)v(x)\, dx$$

for all sufficiently smooth functions $v$. Note that we are assuming that $u$ vanishes on $\partial \Omega$, so that no boundary terms emerge in (330). The central observation of the variational theory of partial differential equations is that (330) is often sufficient to characterize the solution $u$ of the partial differential equation (329). More specifically, under mild assumptions on the operator $L$, the forcing term $f$ and the domain $\Omega$, if (330) holds for all $v$ in a suitable space of test functions then $u$ solves (329). The great advantage of the variational formulation (330) over the equation (329) is that it requires less of $u$. In particular, (330) is sensible when $u$ has only one weak derivative whereas $u$ must be twice weakly differentiable in order for (329) to be meaningful.

A weak differentiable function $u$ which satisfies (330) is known as a \textit{weak solution} of the equation (329). A twice weakly differentiable function $u$ such that (328) holds almost everywhere is called a \textit{strong solution} of (328). If $u$ is twice differentiable and satisfies (328) everywhere, then it is a \textit{classical solution} of (328). A common approach to the analysis of a partial differential equation — one which we will take in this chapter and the next — is to first establish the existence of weak solutions under minimal regularity assumptions and then go on to prove that under slightly stronger conditions, weak solutions are in fact strong or classical solutions.

In the interests of imposing the weakest possible regularity conditions on the operator $L$, we will consider second order linear partial differential equations in divergence form. That is,
4.1. VARIATIONAL FORMULATIONS

Partial differential operators of the form
\[ L[u](x) = -D_i \left( a^{ij}(x) D_j u(x) + b^i(x) u(x) \right) + c^i(x) D_i u(x) + d(x) u(x), \]  
(331)
where \( \Omega \) is a bounded open set in \( \mathbb{R}^n \). The operator \( L \) is no longer a mapping which takes functions to functions; instead, it is a map from an appropriately chosen closed subspace \( V \) of \( H^1(\Omega) \) into its dual space \( V^* \). In particular, it maps the function \( u \) to the mapping \( V \to \mathbb{R} \) defined via
\[ v \to \int_\Omega a^{ij}(x) D_i u(x) D_j v(x) + b^i(x) u(x) D_i v(x) + c^i(x) D_i u(x) v(x) + d(x) u(x) v(x) \, dx. \]  
(332)

The choice of the subspace \( V \) will depend on the boundary conditions being imposed on the solution; however, we will always require that \( V \) contain \( H^1_0(\Omega) \). This ensures, among other things, that if \( L[u] = f \) in the sense that
\[ \langle L[u], v \rangle = \langle f, v \rangle \]  
(333)
for all \( v \in V \), then (330) holds for all test functions \( v \) in \( C^\infty_c(\Omega) \) and it is reasonable to say that \( L[u] = f \) “in the interior of \( \Omega \).” We will suppose that the coefficients \( a^{ij}, b^i, c^i \) and \( d \) are bounded, measurable functions \( \Omega \to \mathbb{R} \). This last assumption is sufficient since the expression (332) does not involve any derivatives of the coefficients \( a^{ij} \). Moreover, we will assume that \( L \) is strongly elliptic; that is, we suppose that there exists a real number \( \lambda > 0 \) such that
\[ \sum a^{ij} \xi_i \xi_j \geq \lambda |\xi|^2 \]  
(334)
for all \( \xi \in \mathbb{R}^n \).

Note that in (332) we have implicitly embedded the space \( V \) in the dual space \( V^* \). In particular, we have embedded \( V \) into \( V^* \) through the composition map
\[ \begin{array}{cccc}
V & \xrightarrow{\iota} & L^2(\Omega) & \xrightarrow{\varphi} & (L^2(\Omega))^* & \xrightarrow{T} & V^*, \\
\end{array} \]  
(335)
where \( \iota \) is the inclusion map
\[ \iota : V \to L^2(\Omega), \]  
(336)
\( \varphi \) is the isometric isomorphism which takes \( u \in L^2(\Omega) \) to the bounded linear functional \( f_u : L^2(\Omega) \to \mathbb{R} \) defined via
\[ f_u(v) = \int_\Omega u(x) v(x) \, dx, \]  
(337)
and \( T : (L^2(\Omega))^* \to V^* \) is the linear map defined by
\[ T[\phi] = \phi|_V. \]  
(338)

Note that the map \( T \) is bounded (obviously), injective (because \( V \) contains \( C^\infty_c(\Omega) \) and is therefore dense in \( L^2(\Omega) \)), and has dense range (since \( V \) is reflexive). When \( u \in L^2(\Omega) \) and \( v \in V \), the duality pairing between \( V \) and \( V^* \) agrees with the \( L^2(\Omega) \) norm:
\[ \langle v, u \rangle_{V \times V^*} = \langle v, u \rangle_{L^2(\Omega)} = \int_\Omega v(x) u(x) \, dx. \]  
(339)

Note also that this embedding of \( V \) into \( V^* \) is plainly not compatible with the usual identification of the Hilbert space \( V \) with its dual space.
4.2. The Dirichlet Problem Sans Lower Order Terms

Suppose that $\Omega$ is a bounded open set in $\mathbb{R}^n$, that $L$ is a second order partial differential operator of the form

$$L[u](x) = -D_i \left( a^{ij}(x) D_j u(x) \right)$$

(340)

with $a^{ij}$ are bounded measurable functions $\Omega \to \mathbb{R}$, and that there exists a real number $\lambda > 0$ such that

$$a^{ij}(x) \xi_i \xi_j \geq \lambda |\xi|^2$$

(341)

for all $x \in \Omega$ and $\xi \in \mathbb{R}^n$ (that is, we are assuming that $L$ is strongly elliptic in $\Omega$). The operator $L$ is a mapping from $H^1_0(\Omega) \to H^{-1}(\Omega)$; in particular, $L[u]$ is the mapping which takes $v \in H^1_0(\Omega)$ to

$$\int_\Omega a^{ij}(x) D_j u(x) D_i v(x) \, dx.$$  

(342)

A function $u \in H^1_0(\Omega)$ is a weak solution of the Dirichlet boundary value problem

$$\begin{cases} 
L[u](x) = f(x) \quad \text{in} \quad \Omega \\
u(x) = 0 \quad \text{on} \quad \partial \Omega,
\end{cases}$$

(343)

where $f \in H^{-1}(\Omega)$, if

$$L[u](v) = \langle f, v \rangle$$

(344)

for all $v \in H^1_0(\Omega)$. In (344), $\langle f, v \rangle$ refers to the duality pairing of $V^*$ and $V$.

**Theorem 61.** The mapping $L$ defined via formula (340) is bounded and coercive.

**Proof.** The $a^{ij}$ are bounded, so there exists a real number $\eta > 0$ such that

$$|a^{ij}(x)| \leq \eta$$

(345)

for all $i, j = 1, \ldots, n$ and $x \in \Omega$. Using (345) and Hölder’s inequality, we see that

$$|\langle L[u], v \rangle| = \left| \int_\Omega a^{ij}(x) D_j u(x) D_i v(x) \, dx \right|$$

$$\leq \eta \sum_{i,j=1}^n \int_\Omega |D_j u(x) D_i v(x)| \, dx$$

$$\leq \eta \sum_{i,j=1}^n \|D_j u\|_2 \|D_i v\|_2$$

$$\leq \eta \sum_{i,j=1}^n \|D u\|_2 \|D v\|_2$$

$$\leq n^2 \eta \|u\|_{H^1_0(\Omega)} \|v\|_{H^1_0(\Omega)}$$

(346)

for all $u$ and $v$ in $H^1_0(\Omega)$. We conclude that $L$ is bounded.

The strong ellipticity of $L$ implies that

$$a^{ij}(x) D_j u(x) D_i u(x) \geq \lambda |D u(x)|^2$$

(347)
for all \( u \in H_0^1(\Omega) \). From (347) we conclude that
\[
\langle L[u], u \rangle = \int_{\Omega} a^{ij}(x) D_j u(x) D_i u(x) \, dx \geq \lambda \int_{\Omega} |D u(x)|^2 \, dx = \lambda \| Du \|_2^2.
\] (348)
According to Poincaré’s inequality, there exists a real number \( \beta > 0 \) such that
\[
\| u \|_2 \leq \beta \| Du \|_2.
\] (349)
for all \( u \in H_0^1(\Omega) \). By combining (348) and (349) we see that
\[
\langle L[u], u \rangle \geq \lambda \int_{\Omega} |D u(x)|^2 \, dx = \lambda \| Du \|_2^2
\] (350)
\[
\geq \frac{\lambda}{2} \| Du \|_2^2 + \frac{\lambda}{2\beta} \| u \|_2^2
\] \[
\geq \min \left\{ \frac{\lambda}{2}, \frac{\lambda}{2\beta} \right\} \| u \|_{H_0^1(\Omega)}^2
\]
for all \( u \in H_0^1(\Omega) \), where \( C \) is an appropriately chosen constant. We conclude that \( L \) is coercive. \( \square \)

In light of Theorem 61, we can apply the Lax-Milgram theorem (Theorem 13 in Section 2.3) in order to conclude that (343) admits a unique weak solution \( u \), and that there exists a constant \( C \) (depending on \( L \) and \( \Omega \) but not \( f \)) such that
\[
\| u \|_{H_0^1(\Omega)} \leq C \| f \|_{H^{-1}(\Omega)}.
\] (351)
Note that if \( a^{ij} = a^{ji} \) for all \( i, j = 1, \ldots, n \), then \( L \) defines an inner product on \( H_0^1(\Omega) \) through the formula
\[
\langle u, v \rangle = \langle L[u], v \rangle
\] (352)
and the Riesz representation theorem suffices to establish the existence of weak solutions of (343).

We now reduce the inhomogeneous boundary problem
\[
\begin{cases}
L[u](x) = f(x) \quad \text{in} \quad \Omega \\
u(x) = g \quad \text{on} \quad \partial \Omega
\end{cases}
\] (353)
to a homogeneous problem of the form (343). As before, we assume that \( f \in H^{-1}(\Omega) \) and, in addition, we assume that \( g \) is the trace of a function \( \psi \) in \( H^1(\Omega) \). Note that this is the weakest assumption we can make on the function \( g \) since the existence of a solution of (351) in \( H^1(\Omega) \) implies that \( g \) is the trace of a function in \( H^1(\Omega) \). We let \( w \) be a weak solution of the boundary value problem
\[
\begin{cases}
L[w](x) = f(x) + D_i \left( a^{ij}(x) D_j \psi(x) \right) \quad \text{in} \quad \Omega \\
w(x) = 0 \quad \text{on} \quad \partial \Omega.
\end{cases}
\] (354)
By \( f(x) + D_i (a^{ij}(x) D_j \psi(x)) \), we mean the bounded linear functional on \( H^1_0(\Omega) \) defined via
\[
v \to \langle f, v \rangle + \int_\Omega a^{ij} D_j \psi(x) D_i v(x) \, dx. \tag{355}\]
We observe that
\[
\left| \int_\Omega a^{ij} D_j \psi(x) D_i v(x) \, dx \right| \leq C \| \psi \|_{H^1_0(\Omega)} \| v \|_{H^1_0(\Omega)} \tag{356}
\]
(the argument is identical to that used in the proof of Theorem 61 to show that \( B \) is bounded) so that
\[
\| f + D_i (a^{ij} D_j \psi) \|_{H^{-1}(\Omega)} \leq C \left( \| f \|_{H^{-1}(\Omega)} + \| \psi \|_{H^1_0(\Omega)} \right). \tag{357}
\]
From (354), we conclude that
\[
\int_\Omega L \left[ w + \psi \right] (x) v(x) \, dx = \langle f, v \rangle \tag{358}
\]
for all \( v \in H^1_0(\Omega) \), and that the trace of \( w + \psi \) is \( g \). In other words, \( w + \psi \) is a weak solution of boundary value problem (353). Moreover, the Lax-Milgram theorem together with (357) implies that
\[
\| w \|_{H^1_0(\Omega)} \leq C \left( \| f \|_{H^{-1}(\Omega)} + \| \psi \|_{H^1_0(\Omega)} \right), \tag{359}
\]
from which we obtain the bound
\[
\| w + \psi \|_{H^1_0(\Omega)} \leq C \left( \| f \|_{H^{-1}(\Omega)} + \| \psi \|_{H^1_0(\Omega)} \right) \tag{360}
\]
for the solution \( w + \psi \) of (353). We summarize our conclusions in the following theorem.

**Theorem 62.** Suppose that \( \Omega \) is a bounded open set in \( \mathbb{R}^n \), that \( \mathcal{T} : H^1(\Omega) \to L^2(\partial \Omega) \) denotes the trace operator, and that \( L \) is a strongly elliptic operator of the form
\[
L [u] (x) = -D_i \left( a^{ij} D_j u \right) \tag{361}
\]
with \( a^{ij} \) bounded, measurable functions. Then for each \( f \in H^{-1}(\Omega) \) and \( \psi \in H^1(\Omega) \) there is a unique weak solution \( u \) of the Dirichlet problem
\[
\begin{cases} 
L [u] (x) = f(x) \quad &\text{in } \Omega \\
u(x) = \mathcal{T} [\psi] (x) &\text{on } \partial \Omega.
\end{cases} \tag{362}
\]
Moreover, there is a constant \( C > 0 \) depending on \( \Omega \) and \( L \) such that
\[
\| u \|_{H^1_0(\Omega)} \leq C \left( \| f \|_{H^{-1}(\Omega)} + \| \psi \|_{H^1_0(\Omega)} \right) \tag{363}
\]
whenever \( u \) is the weak solution of (362).

Theorem 62 characterizes the traces of functions in \( H^1(\Omega) \) when \( \Omega \) is a Lipschitz domain. In particular, it asserts that for every \( g \in H^{1/2}(\partial \Omega) \), there exists a function \( \psi \in H^1(\Omega) \) whose trace is \( g \) and such that
\[
\| \psi \|_{H^1(\Omega)} \leq C \| g \|_{H^{1/2}(\partial \Omega)}. \tag{364}
\]
We obtain the following theorem by combining this observation with Theorem 62.
4.2. THE DIRICHLET PROBLEM SANS LOWER ORDER TERMS

**Theorem 63.** Suppose that $\Omega$ is a Lipschitz domain in $\mathbb{R}^n$, and that $L$ is a strongly elliptic operator of the form

$$L[u](x) = -D_i(a^{ij}D_ju)$$

with $a^{ij}$ bounded, measurable functions. Then for every $f \in H^{-1}(\Omega)$ and every $g \in H^{1/2}(\partial\Omega)$ there exists a unique weak solution $u$ of the Dirichlet boundary value problem

$$\begin{cases}
L[u](x) = f(x) & \text{in } \Omega \\
u(x) = g(x) & \text{on } \partial\Omega.
\end{cases}$$

Moreover, there exists a constant $C > 0$ which depends on $L$ and $\Omega$ such that

$$\|u\|_{H^1_0(\Omega)} \leq C \left( \|f\|_{H^{-1}(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)} \right)$$

whenever $u$ is a weak solution of the boundary value problem (366).

In other words, the operator $L \oplus \mathcal{T}$ is an isomorphism

$$H^1(\Omega) \to H^{-1}(\Omega) \oplus H^{1/2}(\partial\Omega).$$
Bibliography