MAT218A NOTES

December 10, 2017

James Bremer

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 second 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.

(2) "Elliptic Partial Differential Equations of Second Order" by David Gilbarg and Neil Trudinger.

(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.

Much of the material in the preliminaries — with Sections 2.3 and 2.10 notable exceptions — can be found in "Real Analysis: Modern Techniques and Their Application" by Gerald Folland.

CHAPTER 2

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 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 [**8**] or [**4**].

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 $x \in X$,*

$$\sup_{T\in F} \|Tx\| < \infty \tag{1}$$

then

$$\sup_{T\in F} \|T\| < \infty. \tag{2}$$

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 (i.e., $\tilde{T}(y) = T(y)$ for all $y \in Y$) and 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*

$$\{(x, y) \in X \times Y : Tx = y\} \tag{3}$$

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

$$x_1 \sim x_2 \text{ if and only if } x_1 - x_2 \in Y. \tag{4}$$

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

$$\|x + Y\| = \inf_{y \in Y} \|x - y\|. \tag{5}$$

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* $\mathrm{im}(T)$ *is closed, then* $\mathrm{im}(T)$ *is isomorphic to* $X/\ker(T)$*.*

PROOF. We define $\widetilde{T} : X/\ker(T) \to \mathrm{im}(T)$ via the formula

$$\widetilde{T}(x + \ker(T)) = T(x). \tag{6}$$

We observe that $\ker(T)$ is closed since T is continuous, and that T is bijective and continuous. Since $\mathrm{im}(Y)$ 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 $\widetilde{T}$ is an isomorphism. □

Note that if $T : X \to Y$ is a continuous linear mapping and $\mathrm{im}(T)$ is not closed, then $\mathrm{im}(T)$ 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.

(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*

$$V \subset \bigcup_{i=1}^{n} B_\epsilon(x_i).$$

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

$$V \subset \bigcup_{i=1}^{n} B_\epsilon(x_i) \tag{7}$$

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*

$$\|x\| = 1 \tag{8}$$

and

$$\inf_{y \in Y} \|x - y\| \geq \alpha. \tag{9}$$

PROOF. We choose x_1 in $X \setminus Y$ and let

$$r = \inf_{y \in Y} \|x_1 - y\|. \tag{10}$$

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

$$r \leq \|x_1 - y_1\| < r + \epsilon. \tag{11}$$

We set

$$x = \frac{x_1 - y_1}{\|y_1 - x_1\|} \tag{12}$$

so that

$$\|x\| = 1 \tag{13}$$

and

$$\inf_{y \in Y} \|y - x\| = \inf_{y \in Y} \left\| y - \frac{x_1}{\|x_1 - y_1\|} + \frac{y_1}{\|x_1 - y_1\|} \right\|. \tag{14}$$

Since

$$\frac{y_1}{\|x_1 - y_1\|} \tag{15}$$

is in Y, we see from (14) that

$$\inf_{y \in Y} \|x - y\| = \inf_{y \in Y} \left\| y - \frac{x_1}{\|x_1 - y_1\|} \right\| = \frac{r}{r + \epsilon}. \tag{16}$$

Since $r/(r+\epsilon)$ increases to 1 as $\epsilon \to 0$, we can ensure the conclusion of the theorem holds by choosing ϵ sufficiently small. □

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

$$\|y_i - y_j\| \geq \frac{1}{2} \tag{17}$$

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

$$\|x_j\| = 1 \tag{18}$$

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

$$\|x_i - x_j\| \geq \frac{1}{2} \tag{19}$$

whenever i and j are positive integers such that $i \neq j$. Since $\{x_j\}$ is contained in 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 $\overline{\operatorname{im}(K)}$ is compact.It follows from Theorem 7 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

$$\langle Tx, \varphi \rangle = \langle x, T^*\varphi \rangle \tag{20}$$

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

$$\langle x, \varphi \rangle , \tag{21}$$

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

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.*

We now use Riesz's lemma to establish a basic result concerning the spectrum point

$$\sigma_p(K) = \{\lambda : \ker(\lambda I - K) \neq 0\} \tag{22}$$

of a compact operator $K : X \to Y$ between Banach spaces. The spectrum of K is

$$\sigma(K) = \{\lambda : \lambda I - K \text{ is not invertible}\}, \tag{23}$$

We will later see that $\sigma(K) \setminus \{0\} = \sigma_p(K) \setminus \{0\}$ (this depends on the compactness of K, though, and is not true for general linear operators).

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. Suppose that $\{\lambda_n\}$ is a sequence of *distinct* real numbers in $\sigma_p(K) \setminus \{0\}$. Then there exist $x_1, x_2, \ldots$ such that

$$Kx_n = \lambda_n x_n. \tag{24}$$

For each n, we let S_n be the subspace spanned by $x_1, \ldots, x_n$. It is easy to see that the vectors $x_1, \ldots, x_n$ must be linearly independent (since the λ_n are distinct). It follows that S_n is properly contained in S_{n+1} for all n. Moreover, each of these subspaces is finite-dimensional and hence closed. It follows from these two observations that we can apply Riesz' Lemma to construct a sequence $\{y_n\}$ such that $y_n \in S_n$, $\|y_n\| = 1$, and

$$\inf_{z \in S_{n-1}} \|y_n - z\| \geq \frac{1}{2}. \tag{25}$$

Then for any $m < n$ we have:

$$\|\lambda_n^{-1} K y_n - \lambda_m^{-1} K y_m\| = \left\|\lambda_n^{-1}\left(K y_n - \lambda_n y_n\right) - \lambda_m^{-1} K y_m + y_n\right\| \geq \frac{1}{2}. \tag{26}$$

This follows from (25) since Ky_m and $Ky_n - \lambda_n y_n$ are elements of S_{n-1}. To see that $Ky_n - \lambda_n y_n \in S_{n-1}$, we write

$$y_n = \sum_{j=1}^{n} \alpha_j x_j. \tag{27}$$

It follows easily that

$$K y_n = \sum_{j=1}^{n} \alpha_j \lambda_j x_j, \tag{28}$$

from which we conclude that

$$K y_n - \lambda_n y_n = \sum_{j=1}^{n} \left(\alpha_j \lambda_j - \alpha_j \lambda_n\right) x_j = \sum_{j=1}^{n-1} \left(\alpha_j \lambda_j - \alpha_j \lambda_n\right) x_j. \tag{29}$$

Now it follows from (26) that

$$\frac{1}{2} \leq \left|\lambda_n^{-1} - \lambda_m^{-1}\right| \|K(y_n)\| + \left|\lambda_m^{-1}\right| \|K(y_n) - K(y_m)\| \tag{30}$$

for all positive integers $m < n$. We may assume, without loss of generality, that the sequence $K(y_n)$ is convergent by the compactness of K. If $\{\lambda_n\}$ converges to any $\lambda \neq 0$, then we have a contradiction.

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

$$\|z_j\| = 1 \tag{31}$$

for all positive integers j,

$$\|z_i - z_j\| \geq \frac{1}{2} \tag{32}$$

for all pairs of positive integers i,j such that $i \neq j$, and

$$\lambda_j z_j = K(z_j) \tag{33}$$

for all positive integers j. We observe that

$$\|K(z_j) - K(z_i)\| = |\lambda| \, \|z_i - z_j\| \geq \frac{|\lambda|}{2}. \tag{34}$$

But (34) 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

$$\lim_{n\to\infty} (\lambda x_n - K x_n) = y. \tag{35}$$

Since K is compact, by passing to a sequence we can assume that $K x_n$ is convergent. It is clear, then, from (35) that the sequence x_n converges to some $x \in X$. The continuity of the operator $\lambda I - K$ implies that

$$y = \lim_{n\to\infty} (\lambda I - K)\, x_n = (\lambda I - K) = (\lambda I - K) \left(\lim_{n\to\infty} x_n \right) = (\lambda I - K)\, x, \tag{36}$$

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 [**5**]; however, the following theorem gives a useful sufficient condition for a Banach space 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

$$\lim_{n\to\infty} \left\| \sum_{j=1}^{n} \alpha_j x_j - x \right\| = 0. \tag{37}$$

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 $L^\infty(\mathbb{R}^n)$ 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 [**5**], a separable Banach space which does not have a Schauder basis and does not have the approximation property is constructed).

EXERCISE 2. *Suppose that (Ω, μ) 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*

$$T[f](x) = \int_\Omega k(x, y) f(y)\ dy. \tag{38}$$

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

2.3. 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

$$Ax = b \tag{39}$$

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

$$A^* z = 0 \tag{40}$$

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

$$Ax = 0 \tag{41}$$

admits only the trivial solution. In order 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 acting on infinite-dimensional vector spaces which have 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.

2.3.1. 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

$$\|(y, z)\| = \|y\| + \|z\|. \tag{42}$$

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

$$A(y, x) = y + z \tag{43}$$

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

$$P : Y \oplus Z \to Z \tag{44}$$

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 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 [**13**]. 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 13. *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 $\mathrm{im}(P) = Y$.

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

$$A(y, z) = y + z \tag{45}$$

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 $\mathrm{im}(P) = Y$.

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 how that the addition mpa $A : Y \oplus Z \to X$ defined by the formula

$$A(y, z) = y + z \tag{46}$$

is an isomorphism. The map A is plainly continuous since

$$\|y+z\| \leq \|y\|+\|z\|. \tag{47}$$

It is surjective since

$$x = Px + (I-P)x \tag{48}$$

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

$$y_1 + z_1 = y_2 + z_2 \tag{49}$$

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

$$Px_1 = y_1 \tag{50}$$

and

$$Px_2 = y_2. \tag{51}$$

We combine (49), (50) and (51) in order to conclude that

$$P(x_1 - x_2) + (z_1 - z_2) = 0. \tag{52}$$

By applying P to both sides of (52) and make use of the facts that $P^2 = P$ and $Z = \ker(P)$ we obtain

$$P(x_1 - x_2) = 0, \tag{53}$$

from which we conclude that $y_1 = y_2$. It follows from this and (49) 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 14. *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

$$\varphi_j(v_i) = \begin{cases} 1 & \text{if } \ i = j \\ 0 & \text{otherwise.} \end{cases} \tag{54}$$

Now we apply the Hahn-Banach theorem in order to extend each of the φ_j to mappings $Y \to \mathbb{R}$. We also let

$$Z = \bigcap_{j=1}^{n} \ker(\varphi_j), \tag{55}$$

and define the mapping $P : X \to X$ via

$$P(x) = \sum_{j=1}^{n} \varphi_j(x) v_j. \tag{56}$$

We observe that P is a continuous projection whose image is Y. It follows from Theorem 13 that Y is complemented. □

THEOREM 15. *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 φ is continuous. As is the restriction $\tilde{\varphi} : Z \to X/Y$ of φ 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 13 that Y is complemented in X. □

Note that the requirement that Y be closed in Theorem 15 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 14 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 16. *Suppose that X and Y are Banach spaces, and that $T : X \to Y$ is a continuous linear mapping. If $\operatorname{im}(T)$ 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 $Y/\operatorname{im}(T)$ 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

$$A(x, z) = Tx + z. \tag{57}$$

Since

$$\|Tx + z\| \leq \|Tx\| + \|z\| \leq \|T\| \, \|x\| + \|z\| \leq (1 + \|T\|) \left(\|x\| + \|z\|\right), \tag{58}$$

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

$$\sum_{j=1}^{n} \alpha_j v_j \tag{59}$$

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

$$y = \sum_{j=1}^{n} \alpha_j w_j + \sum_{j=1}^{n} \alpha_j u_j + u \tag{60}$$

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

$$T\left(x_1 - x_2\right) + z_1 - z_2 = 0. \tag{61}$$

The restriction of the canonical mapping $Y \to Y/\operatorname{im}(Z)$ to Z is clearly injective, and by applying that mapping to (61) we see that $z_1 = z_2$. It follows from this fact and (61) that

$$T\left(x_1 - x_2\right) = 0. \tag{62}$$

Since T is injective, (62) 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 $\mathrm{im}(T)$. We conclude that $\mathrm{im}(T)$ is closed. $\square$

2.3.2. 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

$$M^\perp = \{f \in X^* : f(x) = 0 \text{ for all } x \in M\}. \tag{63}$$

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

$$N_\perp = \{x \in X : f(x) = 0 \text{ for all } f \in N\}. \tag{64}$$

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 17. *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) $\mathrm{im}(T)^\perp = \ker(T^*)$

(2) $\overline{\mathrm{im}(T)} = \ker(T^*)_\perp$

(3) $\ker(T) = \mathrm{im}(T^*)_\perp$

(4) $\overline{\mathrm{im}(T^*)} \subset \ker(T)^\perp$

From Theorem 13, 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

$$Tx = y. \tag{65}$$

In particular, if the image of T is closed, then (65) admits a solution if and only if

$$\phi(y) = 0 \tag{66}$$

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

2.3.3. 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

$$\mathrm{ind}(T) = \dim(\ker(T)) - \dim(Y/\mathrm{im}(T)). \tag{67}$$

Fredholm operators of index 0 will play a particularly important role in what follows. They have many of the convenient properties possessed by linear mappings $\mathbb{R}^n \to \mathbb{R}^n$. For instance, a Fredholm operator of index 0 $T : X \to Y$ is surjective if and only it is injective.

We will now provide an alternative definition of Fredholm operator, which might be more familiar to the reader.

THEOREM 18. *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* $\mathrm{im}(T)$ *is closed. Then the dual space of* $Y/\mathrm{im}(T)$ *is isomorphic to* $\ker(T^*)$.

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

$$\langle Tx, \varphi \rangle = \langle x, T^*(\varphi) \rangle = 0 \tag{68}$$

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

$$\tilde{\varphi}(y + \mathrm{im}(T)) = \varphi(y) \tag{69}$$

is a well-defined linear functional in the dual of $Y/\mathrm{im}(T)$. We denote by Λ the map which takes $\varphi \in \ker(T^*)$ to the linear functional $\widetilde{\varphi}$ defined via (69). Since $\varphi(\mathrm{im}(T)) = 0$,

$$\|\widetilde{\varphi}\| = \sup_{\|y+\mathrm{im}(T)\|=1} |\widetilde{\varphi}(y)| \leq \sup_{\|y\|=1} |\varphi(y)| = \|\varphi\|, \tag{70}$$

from which we see that Λ is a bounded mapping. We also observe that Λ is bijective; indeed, its inverse is the map taking

$$\psi : Y/\mathrm{im}(T) \to \mathbb{R} \tag{71}$$

to the map $\widetilde{\psi} : Y \to \mathbb{R}$ defined via

$$\widetilde{\psi}(y) = \psi(y + \mathrm{im}(T)). \tag{72}$$

Since $\mathrm{im}(T)$ is closed, $Y/\mathrm{im}(T)$ 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 Λ is an isomorphism. □

THEOREM 19. *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*

$$\mathrm{ind}(T) = \dim(\ker(T)) - \dim(\ker(T^*)). \tag{73}$$

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 16). According to Theorem 18, that $\mathrm{im}(T)$ is closed implies that $(Y/\mathrm{im}(T))^*$ is isomorphic to $\ker(T^*)$. We conclude that

$$\dim(\ker(T^*)) = \dim((Y/\mathrm{im}(T))^*). \tag{74}$$

But $Y/\mathrm{im}(T)$ is finite-dimensional, so

$$\dim(Y/\mathrm{im}(T)) = \dim((Y/\mathrm{im}(T))^*). \tag{75}$$

We combine (74) and (75) in order to obtain

$$\dim(\ker(T^*)) = \dim(Y/\mathrm{im}(T)) < \infty, \tag{76}$$

which suffices to establish (73) and the assertion that $\ker(T^*)$ is finite-dimensional.

Now we suppose that $\ker(T)$, $\ker(T^*)$ are finite-dimensional and that $\operatorname{im}(T)$ is closed. By Theorem 18,

$$(Y/\operatorname{im}(T))^* \sim \ker(T^*). \tag{77}$$

We see from (77) that the codimension of $\operatorname{im}(T)$ 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. $\square$

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*

$$T[\phi_i] = \phi_{i+1} \quad \textit{for all} \;\; i = 1, 2, \dots. \tag{78}$$

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 20. *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

$$I + T^*, \tag{79}$$

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 19 in order to conclude that T is Fredholm. $\square$

We conclude from Theorems 14, 15 and 16 that a Fredholm operator induces the direct sum decompositions

$$X = \ker(T) \oplus X' \tag{80}$$

and

$$Y = \operatorname{im}(T) \oplus Y', \tag{81}$$

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 21. *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*

$$ST = I - K_1 \tag{82}$$

and

$$TS = I - K_2. \tag{83}$$

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 (82) and (83) hold. By Theorem 19, $I - K_1$ and $I - K_2$ are Fredholm, so $\dim(\ker(I - K_1))$ and $\dim(Y/\operatorname{im}(TS))$ are finite-dimensional. We observe that

$$\ker(T) \subset \ker(ST) = \ker(I - K_1) \tag{84}$$

and that

$$Y/\mathrm{im}(T) \subset Y/\mathrm{im}(TS) \tag{85}$$

since $\mathrm{im}(TS) \subset \mathrm{im}(T)$. We conclude that $\ker(T)$ and $Y/\mathrm{im}(T)$ 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

$$X = X' \oplus \ker(T) \tag{86}$$

and a finite dimensional subspace Y' of Y such that

$$Y = Y' \oplus \mathrm{im}(T). \tag{87}$$

The restriction of T to X' is a continuous bijective linear mapping $X' \to \mathrm{im}(T)$. We let $\tilde{S} : \mathrm{im}(T) \to X'$ 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

$$ST = ST(P + I - P) = \tilde{S}T(I - P) = I - P \tag{88}$$

and

$$TS = TS(Q + I - Q) = T\tilde{S}(I - Q) = I - Q. \tag{89}$$

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 21 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

$$ST = I - K_1 \tag{90}$$

and

$$TS = I - K_2. \tag{91}$$

We call any bounded linear operator S for which there exists compact operators K_1 and K_2 such that (82) and (83) 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 21 that

$$\mathrm{ind}(T) = -\,\mathrm{ind}(S) \tag{92}$$

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

Theorem 22. *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* $\mathrm{ind}(T + K) = \mathrm{ind}(T)$.

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

$$ST = I - K_1 \tag{93}$$

and

$$TS = I - K_2 \tag{94}$$

with K_1 and K_2 compact operators. We observe that

$$S\,(T + K) = I - K_1 + SK \tag{95}$$

and

$$(T + K)\,S = I - K_2 + KS. \tag{96}$$

Since $K_1 - KS$ and $K_2 - SK$ are compact operators, we conclude from (95) and (96) that $T + K$ is Fredholm and S is a parametrix for $T + K$. It follows that

$$\operatorname{ind}(S) = -\operatorname{ind}(T + K), \tag{97}$$

but we we also have

$$\operatorname{ind}(S) = -\operatorname{ind}(T) \tag{98}$$

since S is a parametrix for T. We conclude from (97) and (98) that

$$\operatorname{ind}(T) = -\operatorname{ind}(S) = \operatorname{ind}(T + K). \tag{99}$$

□

Any isomorphism $B : X \to Y$ between Banach spaces is Fredholm of index 0. Consequently, it follows from Theorem 22 that any operator of the form

$$B + K \tag{100}$$

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 23. *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

$$X = X' \oplus \ker(T) \tag{101}$$

and a finite dimensional subspace Y' of Y such that

$$Y = Y' \oplus \operatorname{im}(T). \tag{102}$$

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

$$B(x) = T(x') + S(z), \tag{103}$$

where

$$x = x' + z \tag{104}$$

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 can now very easily establish the following:

THEOREM 24. *If $K : X \to X$ is a compact linear mapping between Banach spaces, then $\sigma_p(K) \setminus \{0\} = \sigma(K) \setminus \{0\}$.*

PROOF. It is obvious that $\sigma_p(K) \subset \sigma(K)$. Now suppose that $\lambda \in \sigma(K) \setminus \{0\}$ but $\lambda \notin \sigma_p(K) \setminus \{0\}$. That is, suppose that $\lambda I - K$ is injective but not surjective. Since $\lambda I - K$ is Fredholm of index 0 and $\dim \ker(\lambda I - K) = 0$, the codimension of the image of $\lambda I - K$ must be 0. But that implies that $\lambda I - K$ is surjective, and hence an isomorphism, which contradicts our assumption that $\lambda \in \sigma(K)$. $\square$

We now turn to the solvability of the linear equation

$$Tx = y \tag{105}$$

when $T : X \to Y$ is a Fredholm operator of index 0. We have already observed that T is injective if and only if it is surjective. By combining this observation with the open mapping theorem we obtain the following well-known theorem:

THEOREM 25 (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*

$$Tx = y \tag{106}$$

is uniquely solvable for all $y \in Y$ or the corresponding homogeneous equation

$$Tx = 0 \tag{107}$$

admits nontrivial solutions. In the former case, the inverse of the operator T is bounded.

Since the image of T is closed, the preannihilator of the kernel of T^* coincides with the image of T. This together with the fact that $\ker(T^*)$ is finite-dimensional when T is Fredholm gives us:

THEOREM 26. *Suppose that X and Y are Banach spaces, and that $T : X \to Y$ is a Fredholm operator. Suppose also that*

$$\varphi_1, \ldots, \varphi_n \tag{108}$$

is a basis for $\ker(T^*)$, and that

$$x_1, \ldots, x_m \tag{109}$$

is a basis for $\ker(T)$. Then

$$Tx = y \tag{110}$$

admits a solution if and only if

$$\langle y, \varphi_j \rangle = 0 \quad \text{for all} \quad j = 1, \ldots, n. \tag{111}$$

In the event that $Tx = y$ does admit a solution x_0, then any solution is of the form

$$x_0 + \sum_{j=1}^{m} a_j x_j. \tag{112}$$

2.4. The Lax-Milgram Theorem

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

$$|L[x](x)| \geq \lambda \|x\|^2 \tag{113}$$

for all $x \in X$.

THEOREM 27 (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 (113) implies

$$\lambda\|x\|^2 \leq |\langle L[x], x\rangle| \leq \|L[x]\| \, \|x\| \tag{114}$$

for all $x \in X$. Dividing both sides of (114) by $\|x\|$ yields

$$\lambda\|x\| \leq \|L[x]\| . \tag{115}$$

The identity (115) 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 (115) 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 $\mathrm{im}(L)$. Since $\mathrm{im}(L)$ 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 $L : X \to \mathrm{im}(L)$ is an isomorphism.

We suppose now that $\mathrm{im}(L) \neq X^*$. Then there exists $f \in X^* \setminus \mathrm{im}(L)$. By the Hahn-Banach theorem, there exists ϕ in $(X^*)^*$ such that

$$\langle \phi, f\rangle = 1 \tag{116}$$

and

$$\phi|_{\mathrm{im}(L)} = 0. \tag{117}$$

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

$$f(u) = 1 \tag{118}$$

and

$$\langle L[x], u\rangle = 0 \tag{119}$$

for all $x \in X$. We combine (116) and (117) with the assumption that L is coercive in order to conclude that

$$\lambda\|u\|^2 \leq \langle L[u], u\rangle = 0, \tag{120}$$

from which we see that $u = 0$. However, this contradicts (118). We conclude that $\mathrm{im}(L) = V^*$. ∎

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

$$B[x,y] = L[x](y). \tag{121}$$

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

$$|B[x,y]\| \le C\|x\|\|y\| \tag{122}$$

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

$$|B[x,x]| \ge \lambda\|x\|^2 \tag{123}$$

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

Theorem 28. *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*

$$B[u,v] = f(v) \tag{124}$$

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.

Theorem 29 (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*

$$f(x) = (u,x) \tag{125}$$

for all $x \in X$

A slight modification of the argument we used to establish the Lax-Milgram theorem gives us the following theorem.

Theorem 30. *Suppose that X is a reflexive Banach space, and that $T : X \to X^*$ is a bounded linear mapping. Suppose also that there exists $\lambda > 0$ such that*

$$\|Lx\| \ge \lambda\|x\| \quad \textit{for all} \quad x \in X \tag{126}$$

and that

$$\sup_{x \in X} |\langle Lx, y\rangle| > 0 \quad \textit{for all} \quad y \in X \setminus \{0\}. \tag{127}$$

Then L is an isomorphism.

2.5. Weak Convergence and the Banach-Alaoglu theorem

Suppose that X is a Banach space, and that X^* is its dual space. Until now, we have considered only the strong topology on X^* and only the norm topology on X. There are, however, other toplogies on these spaces which are of use to us.

We say that a sequence $\{x_n\}$ in X converges to x weakly provided

$$\phi(x_n) \to \phi(x) \tag{128}$$

for all $\phi \in X^*$. We use the notation $x_n \rightharpoonup x$ to indicate that the sequence $\{x_n\}$ converges weakly to x.

We say that $\{\phi_n\} \subset X^*$ converges to $\phi \subset X^*$ in the weak-* topology provided

$$\phi_n(x) \to \phi(x) \text{ for all } x \in X. \tag{129}$$

The original version of the Banach-Alaoglu theorem is

THEOREM 31 (Banach-Alaoglu I). *If X is a Banach space, then the closed unit ball of X^* is compact in the weak-* topology.*

The weak-* topology on the unit ball of X^* is metrizable when X is separable (although the weak-* topology on the whole space X^* is metrizable if and only if X^* is finite-dimensional). In that case, we have the following version of the Banach-Alaoglu theorem:

THEOREM 32 (Banach-Alaoglu II). *If X is a separable Banach space, then the closed unit ball of X^* is sequentially compact in the weak-* topology.*

When X is reflexive, the weak-* toplogy on $(X^*)^*$ coincides with the weak topology on X. In this case, we have teh following version of the Banach-Alaoglu theorem. Note that we are not making the assumption that X is separable.

THEOREM 33 (Banach-Alaoglu III). *Suppose that X is a reflexive Banach space. Then the closed unit ball of X is sequentially compact in the weak topology.*

The following particular form of the preceding theorem is the one we will use most often:

THEOREM 34 (Banach-Alaoglu IV). *Suppose that Ω is an open subset of $\mathbb{R}^n$, and that $1 < p < \infty$ is a real number. Then any bounded sequence in $L^p(\Omega)$ has a weakly convergent subsequence.*

2.6. Galerkin Discretization

Suppose that X is a Banach space, that $X_1 \subset X_2 \subset X_3 \subset \cdots$ is a sequence of finite-dimensional subspaces of X, and that for each positive integer j, $P_j : X \to X_j$ is a projection operator. We say that the subspaces $\{X_j\}$ together with the projection operators $\{P_j\}$ form a projective approximation scheme for X if

$$\lim_{n\to\infty} \|P_j x - x\|_X = 0 \tag{130}$$

for all $x \in X$ and

$$P_n P_m = P_n \tag{131}$$

whenever $n \leq m$ are positive integers. We note that by the uniform boundedness principle, (130) implies that there exists a constant C such that

$$\|P_j\| \leq C \tag{132}$$

for all positive integers j. We also observe that (130) means that the union of the subspaces X_j must be dense in X; that is,

$$\overline{\cup_{j=1}^{\infty} X_j} = X. \tag{133}$$

In particular, the existence of a projective approximation scheme for X implies that X is separable.

In the event that X is a Hilbert space and $\{\varphi_j\}$ is an orthonormal basis for X, a projective approximation scheme for X can be obtained by letting

$$S_j = \text{span}\{\varphi_1, \ldots, \varphi_j\}, \tag{134}$$

and taking P_j to be the orthogonal projection operator

$$P_j\,[f] = \sum_{i=1}^{j} \langle f, \varphi_i \rangle\, \varphi_i. \tag{135}$$

Suppose now that X and Y are reflexive Banach spaces, and that $A : X \to Y$ is a bounded linear mapping. Suppose also that $\{X_j, P_j\}$ is a projective approximation scheme for X, and that $\{Y_j, Q_j\}$ is a projective approximation scheme for Y such that $\dim(Y_j) = \dim(X_j)$ for each $j = 1, 2 \ldots$. For each positive integer n, we let $A_n = Q_n A|_{X_n}$. By a slight abuse of terminology, we call each of the equations

$$A_n x_n = Q_n y \quad (x_n \in X_n) \tag{136}$$

a Galerkin discretization of

$$Ax = y. \tag{137}$$

We will shortly given conditions which guarantee that for for all sufficiently large n the equation (136) admits a solution x_n, and that the sqeuence $\{x_n\}$ converges to the solution x of (137). We observe first, though, that Equation (136) is equivalent to a linear system of equations. To see this, we let $\varphi_1, \ldots, \varphi_m$ be a basis for X_n and $\psi_1, \ldots, \psi_m$ be a basis for Y_n. Then we can represent x_n as

$$x_n = \sum_{j=1}^{m} a_j \varphi_j, \tag{138}$$

$Q_n y$ as

$$Q_n y = \sum_{i=1}^{m} b_i \psi_i, \tag{139}$$

and, for each $j = 1, \ldots, k$, we represent $Q_n A \varphi_j$ as

$$Q_n A \varphi_j = \sum_{i=1}^{m} c_{ij} \psi_i. \tag{140}$$

Then (136) is equivalent to requiring that

$$\sum_{j=1}^{m} c_{ij} a_j = b_i \;\; \text{for all} \quad i = 1, \ldots, m. \tag{141}$$

In the event that X and Y are Hilbert spaces, $\{\varphi_j\}$ is an orthonormal basis for X, $\{\psi_j\}$ is an orthonormal basis for Y, and P_j and Q_j are the corresponding orthogonal projection operators then

$$c_{ij} = \langle A\varphi_j, \psi_i \rangle\,, \tag{142}$$

and

$$b_i = \langle y, \psi_i \rangle . \tag{143}$$

THEOREM 35. *Suppose that X and Y are reflexive Banach spaces, that $A : X \to Y$ is a bounded linear operator, that $\{X_j, P_j\}$ is a projective approximation scheme for X, that $\{Y_j, Q_j\}$ is a projective approximation scheme for Y such that* $\dim(Y_j) = \dim(X_j)$ *for all $j = 1, 2, \ldots$, and that for each positive integer n, A_n is the operator $X_n \to Y_n$ defined via $A_n = Q_n A|_{X_n}$. Suppose also that there exist a positive integer N and a positive real number γ such that for all $n \geq N$ and all $x_n \in X_n$*

$$\|A_n x_n\|_Y \geq \gamma \|x_n\|_X . \tag{144}$$

Then for each $y \in Y$, there exists a unique solution x of the equation $Ax = y$, and for each $n \geq N$ the equation

$$A_n x_n = Q_n y \tag{145}$$

admits a unique solution x_n. Moreover, the sequence $\{x_n\}$ converges to x; in fact, if c is a constant such that $\|Q_n\| \leq c$ and $\|P_n\| \leq c$ for all positive integers n, then

$$\|x_n - x\|_X \leq \left(1 + \frac{c}{\gamma}\|A\|\right) \inf_{v \in X_n} \|x - v\|_X . \tag{146}$$

PROOF. We fix $y \in Y$ and let c be a constant such that $\|Q_n\| \leq c$ and $\|P_n\| \leq c$ for all positive integers n. Condition (144) implies that each $A_n, n \geq N$, is an isomorphism. In particular, the equation

$$A_n x_n = Q_n y \tag{147}$$

has a unique solution x_n for all $n \geq N$, and

$$\|x_n\| \leq \frac{1}{\gamma} \|Q_n y\|_Y \leq c \|y\| \tag{148}$$

for all $n \geq N$. By the Banach-Alaoglu theorem, there exists a subsequence of $\{x_n\}$ which converges weakly to some element x of X. Without loss of generality, we pass to this subsequence. It is not hard to establish that $Q_n y = Q_n A x_n$ converges weakly to Ax. On the other hand, $Q_n y$ converges strongly (and weakly) to y, so by the uniquness of weak limits we must have $Ax = y$.

Now we observe that for any $v \in X_n$,

$$\begin{aligned}
\|x_n - x\|_X &\leq \|x_n - v\|_X + \|v - x\|_X \\
&\leq \frac{1}{\gamma} \|Q_n A (x_n - v)\|_Y + \|v - x\|_X \\
&= \frac{1}{\gamma} \|Q_n A (x - v)\|_Y + \|v - x\|_X \\
&\leq \left(1 + \frac{1}{\gamma} \|Q_n\| \|A\|\right) \|v - x\|_X \\
&\leq \left(1 + \frac{c}{\gamma} \|A\|\right) \|v - x\|_X .
\end{aligned} \tag{149}$$

This suffices to establish (146). Note that

$$Q_n A x_n = Q_n A x \tag{150}$$

since $Au = y$ implies $Q_n A x = Q_n y$ and $Q_n A x_n = A_n x_n = Q_n y_n$ by construction.

We now need only establish that x is the unique solution of $Ax = y$. To that end, we suppose that $\tilde{x}$ is such that $A\tilde{x} = y$ and let $w = x - \tilde{x}$. Then $Aw = 0$. It is easy to verify that

$$A_n P_n x \to Ax \tag{151}$$

for any $x \in X$. Using this, we obtain

$$\|P_n w\| \leq \frac{1}{\lambda} \|A_n P_n w\| \to \|Aw\| = 0. \tag{152}$$

From this we conclude that $P_n w \to 0$ as $n \to \infty$, so $w = 0$. □

As is often the case, perturbing the operator A by a compact operator does not fundamentally change the situation. Indeed:

THEOREM 36. *Suppose that the hypotheses of Theorem 35 are satisfied. Suppose, in addition, that $B : X \to Y$ is compact operator, and that* $\ker(A + B) = \{0\}$. *Suppose also that for each positive integer n, $A_n = Q_n A|_{X_n}$ and $B_n = Q_n B|_{X_n}$. Then for any $y \in Y$, there exists a unique solution x to the equation*

$$Ax + Bx = y. \tag{153}$$

Moreover, for all n which are sufficently large the equation

$$Q_n (A_n + B_n) x_n = Q_n y \tag{154}$$

has a unique solution $x_n \in X_n$ and there exists a constant C such that

$$\|x - x_n\|_X \leq C \inf_{v \in X_n} \|x - v\|. \tag{155}$$

PROOF. We will show that $A + B$ satisfies condition (144). That is, there exists λ such that

$$\|(A_n + B_n) x_n\|_Y \geq \lambda \|x_n\|_X. \tag{156}$$

for all $x_n \in X_n$. Suppose this is not the case. Then there exists a sequence $\{x_j\}$ with $x_j \in X_{n_j}$ such that $\|x_j\| = 1$ and

$$\left\|Q_{n_j} (A + B) x_j\right\|_Y \to 0 \quad \text{as} \quad j \to \infty. \tag{157}$$

By the Banach-Alaoglu theorem, we may assume without loss of generality that x_j converges weakly to some $x \in X$. Then $Q_{n_j} (A + B) x_j$ converges weakly to $(A + B) x$. It follows from this and (157) that $(A + B) x = 0$. Since the kernel of $A+B$ is $\{0\}$, we must have $x = 0$. But B is compact, so $Bx_j \to 0$ (some subsequence of $\{Bx_j\}$ is convergent, and it must converge to $x = 0$). It follows that

$$\begin{aligned} \|x_j\|_X &\leq \|Q_{n_j} A x_j\|_Y \\ &\leq \|Q_{n_j}(A + B)x_j\|_Y + \frac{1}{\gamma} \|Q_{n_j} B x_j\|_Y \to 0 \quad \text{as} \quad j \to \infty, \end{aligned} \tag{158}$$

but this contradicts the fact that $\|x_j\| = 1$ for all $j = 1, 2 \ldots$. □

The preceding theorem can be easily applied to coercive operators. Suppose that X is a reflexive Hilbert space, and that that L is a linear mapping $X \to X^*$. We say that L is coercive provided there exists a constant $\lambda > 0$ such that

$$\langle Lx, x\rangle \geq \lambda \langle x, x\rangle \tag{159}$$

for all $x \in X$. We now let $\{X_j, P_j\}$ be a projective approximation scheme. Moreover, for each $j = 1, 2, \ldots$, we set $X_j^* = P_j^* X_j$. Then $\{X_j^*, P_j^*\}$ is a projective approximation scheme for X^*. Moreover, since

$$\langle Lx, x\rangle \geq \lambda \langle x, x\rangle = \lambda \|x\|^2 \tag{160}$$

for all $x \in X$, we have

$$\lambda \|x_n\|^2 \leq |\langle x_n, Lx_n\rangle| = |\langle P_n x_n, Lx_n\rangle| = |\langle x_n, P_n^* L x_n\rangle| \leq \|P_n^* L x_n\| \|x_n\| \tag{161}$$

for all $x_n \in X_n$. In particular, condition (144) of Theorem 35 is satisfied. If follows that for each $y \in X^*$, there exists a unique solution to the equation $Lx = y$, and that the approximations of x formed via Galerkin discretization converge. The open mapping theorem now implies that L has a continuous inverse.

Note that the existence of a projective approximation scheme for X implies that X is separable, so that the Lax-Milgram theorem applies in greater generality that the result just obtained. However we have just established that coercive operators $X \to X^*$ are susceptible to a particularly simply case of Galerkin discretization provided X is seperable, a stronger conclusion than that of the Lax-Milgram theorem.

2.7. 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.

2.7.1. Spaces of Continuous Functions. Suppose that Ω 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 $C\left(\overline{\Omega}\right)$ of all continuous functions $\Omega \to \mathbb{R}$ with compact support contained in Ω. 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 Ω.

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 $C\left(\overline{\Omega}\right)$ is not the same space as $C\left(\mathbb{R}^n\right)$, which means that the notation $C\left(\overline{\Omega}\right)$ is misleading when Ω is not bounded.*

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

$$\|f\| = \sup_{x \in \Omega} |f(x)|, \tag{162}$$

as is the vector space $C(\overline{\Omega})$. Neither $C(\Omega)$ nor $C_c(\Omega)$ are Banach spaces with respect to (162), 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

$$\|f\|_i = \sup_{x \in K_i} |f(x)|, \quad i = 1, 2, \ldots. \tag{163}$$

We will not make use of this last observation, but a further discussion can be found in, for instance, [**15**].

Exercise 6. *Suppose that Ω 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 Ω is not bounded, then there exist uniformly continuous functions $\Omega \to \mathbb{R}$ which are not bounded.*

Suppose that Ω is a bounded open subset of $\mathbb{R}^n$. The following theorem, a proof of which can be found in [**8**] (for instance), characterizes the compact subsets of $C\left(\overline{\Omega}\right)$. We say that a subset Φ of $C\left(\overline{\Omega}\right)$ is equicontinuous if for all $\epsilon > 0$ there exists $\delta > 0$ such that

$$|f(x) - f(y)| < \epsilon \tag{164}$$

for all x and y in $\overline{\Omega}$ such that

$$|x - y| < \delta. \tag{165}$$

Similarly, we say that Φ is uniformly bounded if there exists a $M > 0$ such that

$$\|f\| < M \tag{166}$$

for all $f \in \Phi$.

Theorem 37 (Arzelà-Ascoli). *Suppose that Ω is a bounded open subset of $\mathbb{R}^n$, and that Φ is a subset of $C\left(\overline{\Omega}\right)$. Then Φ is compact if and only if it is equicontinuous, closed and uniformly bounded.*

The following well-known theorem is also of some interest to us.

Theorem 38 (Stone-Weierstrauss). *Suppose that Ω is a bounded opne subset of $\mathbb{R}^n$. Then a subalgebra Φ of $C\left(\overline{\Omega}\right)$ which contains 1 is dense if and only if it separates points. That is, if for any x and y in Ω, there exists $f \in \Phi$ such that $f(x) \neq f(y)$.*

2.7.2. Multi-index notation for partial 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

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n, \tag{167}$$

we define the factorial of a multi-index α via

$$\alpha! = \alpha_1!\,\alpha_2! \cdots \alpha_n!\,, \tag{168}$$

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

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}. \tag{169}$$

Moreover, for each multi-index α, we denote by D^α the partial differential operator which acts on sufficiently smooth functions u defined on an open set $\Omega \subset \mathbb{R}^n$ via the formula

$$D^\alpha u = \left(\frac{\partial \varphi}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial \varphi}{\partial x_2}\right)^{\alpha_2} \cdots \left(\frac{\partial \varphi}{\partial x_n}\right)^{\alpha_n} u. \tag{170}$$

We will sometimes also use the alternative notation $D_i u$ to refer to the derivative of u with respect to the i^{th} coordinate; that is,

$$D_i u = \frac{\partial u}{\partial x_i}. \tag{171}$$

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

$$u(y) = \sum_{|\alpha|=0}^{k} \frac{\alpha!}{j!} D^\alpha u(x)(y-x)^\alpha + O\left(|y-x|^{k+1}\right) \tag{172}$$

and the product rule is

$$D^\alpha(uv) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta}(u) D^\beta(v), \tag{173}$$

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

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!\,(\alpha-\beta)!}. \tag{174}$$

We will sometimes use Du to refer to the gradient of the function u; that is,

$$Du(x) = \begin{pmatrix} D_1 u(x) \\ D_2 u(x) \\ \vdots \\ D_n u(x) \end{pmatrix}. \tag{175}$$

2.7.3. Spaces of Smooth Functions. Suppose that Ω 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 $C_b^k(\Omega)$ the subspace of $C^k(\Omega)$ consisting of functions whose derivatives through order k are continuous and bounded, by $C^k\left(\overline{\Omega}\right)$ the subspace of $C^k(\Omega)$ consisting of k-times differentiable functions whose derivatives through order k are bounded and uniformly continuous on Ω, and by $C_c^k(\Omega)$ the subspace of $C^k\left(\overline{\Omega}\right)$ consisting of k-times differentiable functions with compact support contained in Ω. The vector spaces $C_b^k(\Omega)$ and $C^k\left(\overline{\Omega}\right)$ are Banach spaces with respect to the norm

$$\|f\| = \sum_{|\beta| \le k} \sup_{x\in\Omega} \left|D^\beta f(x)\right|. \tag{176}$$

By $C^\infty(\Omega)$ we mean the vector space of functions infinitely differentiable functions $\Omega \to \mathbb{R}$. We denote by $C_b^\infty(\Omega)$ the subspace of $C^\infty(\Omega)$ consisting of infinitely differentiable functions $\Omega \to \mathbb{R}$ whose derivatives of all orders are bounded, and by $C_c^\infty(\Omega)$ the space of infinitely differentiable functions with compact support contained in Ω. All of these spaces are Fréchet spaces with respect to appropriately chosen families of seminorms (see, for instance, [**15**]); however, we will not make use of this fact.

EXERCISE 7. *Suppose that Ω is an open convex set in $\mathbb{R}^n$. Show that the norm (176) and*

$$\|f\|_0 = \sup_{x\in\Omega} |f(x)| + \sum_{|\beta|=k} \sup_{x\in\Omega} \left|D^\beta(x)\right| \tag{177}$$

are equivalent norms for $C^k\left(\overline{\Omega}\right)$ *(Hint: use some form of Taylor's theorem).*

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

$$\left|D^\beta f(x) - D^\beta f(y)\right| \leq C\left|x-y\right|^\alpha \tag{178}$$

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

$$\|f\| = \|f\|_{C^m\left(\overline{\Omega}\right)} + \sup_{|\beta|=k}\ \sup_{x,y\in\Omega,\ x\neq y} \frac{\left|D^\beta f(x) - D^\beta f(y)\right|}{|x-y|^\alpha}, \tag{179}$$

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

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

$$C^{k,\alpha}\left(\overline{\Omega}\right) \subsetneq C^{k,\lambda}\left(\overline{\Omega}\right) \subsetneq C^k\left(\overline{\Omega}\right). \tag{180}$$

Note carefully the assumption that Ω *is bounded.*

Suppose that Ω 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 $C^{k,\alpha}\left(\overline{\Omega'}\right)$. By $\Omega' \subset\subset \Omega$ we mean that Ω' is compactly supported in Ω; that is, there exists a compact set K such that $\overline{\Omega'} \subset K \subset \Omega$.

For the sake of convenience, we set

$$C^{k,0}\left(\overline{\Omega}\right) = C^k\left(\overline{\Omega}\right) \tag{181}$$

and

$$C^{k,0}(\Omega) = C^k(\Omega) \tag{182}$$

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 Ω 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

$$\psi(x_1, \ldots, x_n) = \begin{pmatrix} \psi_1(x_1, \ldots, x_n) \\ \vdots \\ \psi_m(x_1, \ldots, x_n) \end{pmatrix} \tag{183}$$

for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Then we say that ψ is a $C^{k,\alpha}$ mapping if each of the mappings ψ_j is an element of $C^{k,\alpha}(\overline{\Omega})$. 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

$$\|f(x) - f(y)\| \leq C\|x-y\| \tag{184}$$

for all $x, y \in \Omega$.

EXERCISE 9. *Suppose that Ω 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 $C^{k,\alpha}\left(\overline{\Omega}\right)$ is compactly embedded in $C^k\left(\overline{\Omega}\right)$. Hint: use the Arzelá-Ascoli theorem.*

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

2.7.5. 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 model-ling 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

$$\sum_{j=1}^{n} |f(b_j) - f(a_j)| \leq \epsilon \tag{185}$$

whenever $(a_1, b_1), \ldots, (a_n, b_n)$ is a finite collection of disjoint open intervals in $[a,b]$ such that

$$\sum_{j=1}^{n} (b_j - a_j) \leq \delta. \tag{186}$$

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 [**8**]).

THEOREM 39. *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 $x \in [a,b]$*

$$f(x) = f(a) + \int_a^x f'(y)\, dy. \tag{187}$$

That the usual integration by parts formula

$$f(b)g(b) - f(a)g(a) = \int_a^b f'(x)g(x)\, dx + \int_a^b f(x)g'(x)\, dx \tag{188}$$

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 [**8**]).

We conclude that a Lipschitz continuous function $f : [a,b] \to \mathbb{R}$ is differentiable almost everywhere, and that the integration by parts formula (188) holds for such functions. Also, from (187), we see that the derivative of f must be bounded almost everywhere.

Rademacher's theorem, which we state below, extends Theorem 39 to higher dimensions. A proof can be found in [**16**].

THEOREM 40 (Rademacher). *Suppose that Ω 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*

$$\|f(x) - f(y)\| \leq C\|x - y\| \tag{189}$$

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

It is important to understand that Theorem 40 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

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - Th\|}{\|h\|} = 0. \tag{190}$$

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 41. *Suppose that Ω is an open subset of $\mathbb{R}^n$, that Ω' is an open subset of $\mathbb{R}^m$, and that Ω'' 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 42. *Suppose that Ω 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*

$$(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x). \tag{191}$$

It is somewhat more difficult to establish the following change of variables formula for Lipschitz mappings. A proof can be found in [**7**].

THEOREM 43. *Suppose that Ω and is a open subset of $\mathbb{R}^n$, and that $\psi : \Omega \to \mathbb{R}^n$ is a bilipschitz mapping. Then*

$$\int_\Omega g(\psi(x)) \left|\det\left(\psi'(x)\right)\right| \, dx = \int_{\psi(\Omega)} g(y) \, dy \tag{192}$$

for all measurable $g : \Omega \to \mathbb{R}$. In particular, the measure of $\psi(\Omega)$ is

$$\int_\Omega \left|\det\left(\psi'(x)\right)\right| \, dx. \tag{193}$$

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.

2.8. Mollifiers, Cutoff Functions and Partitions of Unity

The function $\eta : \mathbb{R}^n \to \mathbb{R}$ defined via

$$\eta(x) = \begin{cases} \exp\left(-\frac{1}{1-|x|^2}\right) & \text{if } |x| \leq 1 \\ 0 & \text{if } |x| > 1, \end{cases} \tag{194}$$

is an element of $C_c^\infty\left(\mathbb{R}^n\right)$. For each $\epsilon > 0$, we define

$$\eta_\epsilon(x) = (\alpha)^{-1} \epsilon^{-n} \eta(x/\epsilon), \tag{195}$$

where

$$\alpha = \int_{\mathbb{R}^n} \eta(x)\ dx, \tag{196}$$

so that η_ϵ is supported on the ball $\{x : |x| \leq \epsilon\}$ and

$$\int \eta_\epsilon(x)\ dx = 1. \tag{197}$$

We call η_ϵ the standard mollifier on $\mathbb{R}^n$. The sequence of functions

$$\eta_\epsilon * u(x) = \int_{\mathbb{R}^n} \eta_\epsilon(x-y)u(y)\ dy \tag{198}$$

obtained by convolving u with η_ϵ is called mollification or regularization of f. The following theorem, whose proof can be easily found in the literature (for example, in [**8**]), enumerates some of the key properties of mollifiers.

THEOREM 44. *Suppose that $u \in L^1_{loc}(\mathbb{R}^n)$, and that η_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 Ω is an open set in $\mathbb{R}^n$ such that $0 < h < \operatorname{dist}(\operatorname{supp}(u), \partial\Omega)$, then the support of $\eta_h * u$ is contained in Ω.

EXERCISE 11. *Suppose that η_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 $h \to 0$.*

EXERCISE 12. *Suppose that Ω is a bounded open subset of $\mathbb{R}^n$, and that f and g are continuous functions $\Omega \to \mathbb{C}$ such that*

$$\int_\Omega f(x)\psi(x)\ dx = \int_\Omega g(x)\psi(x)\ dx \tag{199}$$

whenever ψ is a infinitely differentiable function whose support is contained in Ω. Show that $f(x) = g(x)$ for all $x \in \Omega$.

We will typically work with functions which are only defined on an open subset Ω of $\mathbb{R}^n$. If $u \in L^1_{\text{loc}}(\Omega)$, then the mollification

$$\eta_h * u(x) = \int_\Omega \eta_h(x-y)u(y)\ dy = \int_\Omega \eta_h(y)u(x-y)\ dy \tag{200}$$

is defined for all $x \in \Omega$ and $h > 0$ such that $0 < \operatorname{dist}(x, \partial\Omega) < h$.

The following definition is useful when mollifying functions defined on subsets of $\mathbb{R}^n$. If Ω is an open subset of $\mathbb{R}^n$ and Ω' is an open subset of Ω, then we say that Ω' is compactly embedded in Ω 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 45. *Suppose that Ω is an open subset of $\mathbb{R}^n$, that $\Omega' \subset\subset \Omega$, and that η_h denotes the standard mollifier. Then:*

(1) If $u \in L^p_{loc}(\Omega)$ *with* $1 \le p < \infty$*, then* $\eta_h * u \to u$ *in* $L^p(\Omega')$ *and* $\eta_h * u$ *converges to* u *almost everywhere in* Ω'*.*

(2) If $u \in C(\Omega)$*, then* $\eta_h * u$ *converges to* u *uniformly on* Ω'*.*

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 46. *Suppose that* U *is an open subset of* $\mathbb{R}^n$*, and that* $V \subset\subset U$*. Then there exists a nonnegative function* $\psi \in C_c^\infty(U)$ *which is identically* 1 *on* $\overline{V}$*.*

PROOF. We denote the standard mollifier by η_ϵ, let W be an open set such that

$$\overline{V} \subset W \subset \overline{W} \subset U, \tag{201}$$

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

$$\psi(x) = \chi_W * \eta_\epsilon \tag{202}$$

is the desired cutoff function. To see this, we observe that according Conclusion (1) of Theorem 44, $f_\epsilon \in C^\infty(\mathbb{R}^n)$, and that conclusion (4) of the same theorem implies that the support of ψ is contained in U. Moreover, for all $x \in V$ the support of the function $\eta_\epsilon(x - \cdot)$ is contained in W, so that

$$\psi(x) = \int_{\mathbb{R}^n} \eta_\epsilon(x-y)\chi_W(y)\ dy = \int_W \eta_\epsilon(x-y)\ dy = \int_{\mathbb{R}^n} \eta_\epsilon(x-y)\ dy = 1. \tag{203}$$

It is the case that $\psi(x) \ge 0$ for all $x \in \mathbb{R}^n$ since both η_ϵ and ξ_W are nonnegative functions. □

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

$$A \subset \bigcup_{\alpha \in \mathscr{O}} U_\alpha \tag{204}$$

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 (204) if

(1) each of the functions ψ_j is supported in one of the sets U_α;

(2) $0 \le \psi_j(x) \le 1$ 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 ψ_j vanish;

(4) $\displaystyle\sum_{j=1}^{\infty} \psi_j(x) = 1$ 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 47. *Suppose that A is a compact subset of $\mathbb{R}^n$, and that*

$$A \subset \bigcup_{\alpha \in \mathscr{O}} U_\alpha \tag{205}$$

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

(1) Each ψ_j is supported in one of the open sets U_α;

(2) $0 \le \psi_j(x) \le 1$ 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

$$A \subset \bigcup_{j=1}^m U_j \tag{206}$$

be a covering of A by open sets chosen from the collection $\{U_\alpha : \alpha \in \mathscr{O}\}$. Now we construct a collection of open sets $V_1, \ldots, V_m$ such that

$$A \subset \bigcup_{j=1}^m V_j \tag{207}$$

and

$$V_j \subset\subset U_j \tag{208}$$

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

$$V_{j,\epsilon} = \left\{x \in U_j \ : \ \operatorname{dist}\left(x, U_j^c\right) > \epsilon\right\}. \tag{209}$$

We claim that if $\epsilon > 0$ is sufficiently small, then

$$A \subset \bigcup_{j=1}^m V_{j,\epsilon}. \tag{210}$$

If not, then for each $n > 0$, there exists x_n such that

$$x_n \in A \setminus \bigcup_{j=1}^m V_{j,1/n} = A \cap \bigcap_{j=1}^m V_{j,1/n}^c \tag{211}$$

Since A is compact, the sequence $\{x_n\}$ has a subsequence converging to a point $x \in A$. From (211) we see that

$$x \notin \bigcup_{j=1}^m V_{j,1/n} \tag{212}$$

for all $j = 1, \ldots, m$ and all positive integers n. We conclude from (212) that $x \in U_j^c$ for all $j = 1, \ldots, m$. But this is a contradiction since the U_j cover A and $x \in A$. We conclude that (210) holds.

According to Theorem 46, for each $j = 1, \dots, m$, there exists a function $\varphi_j \in C_c^\infty(U_j)$ which is 1 on V_j. Then the functions $\psi_1, \dots, \psi_m$ defined via the formula

$$\psi_j(x) = \frac{\varphi_j(x)}{\sum_{i=1}^m \varphi_j(x)} \tag{213}$$

have the desired properties. ☐

THEOREM 48. *Suppose that A is an arbitrary subset of $\mathbb{R}^n$, and that*

$$A \subset \bigcup_{\alpha \in \mathscr{O}} U_\alpha \tag{214}$$

is a covering of A by open sets. Then there exists a smooth partition of unity $\psi_1, \psi_2, \dots$ subordinate to (214).

PROOF. We first suppose that A is open. Then

$$A = \bigcup_{k=1}^\infty A_k, \tag{215}$$

where the sets A_k are defined by the formula

$$A_k = \{x \in A : |x| \leq k \text{ and } \operatorname{dist}(x, \partial A) \geq 1/k\}. \tag{216}$$

Note that if A is not open, then (215) need not hold. For each $\alpha \in \mathscr{O}$ and each $k \geq 1$, we let

$$V_{\alpha,k} = U_\alpha \cap \operatorname{int}(A_{k+1} \setminus A_{k-2}), \tag{217}$$

where we set $A_0 = A_{-1} = \emptyset$. Then, for each fixed $k \geq 1$,

$$A_k \subset \bigcup_{\alpha \in \mathscr{O}} V_{\alpha,k} \tag{218}$$

is an open covering of the compact set A_k. For each $k \geq 1$, we invoke Theorem 47 in order to obtain a smooth partition of unity

$$\psi_1^k, \psi_2^k, \dots, \psi_{m_k}^k \tag{219}$$

subordinate to the covering (218). We define the function $\sigma : A \to \mathbb{R}$ via the formula

$$\sigma(x) = \sum_{k=1}^\infty \sum_{j=1}^{m_k} \psi_j^k(x); \tag{220}$$

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 $k = 1, 2, \dots$ and $1 \leq j \leq m_k$, we define

$$\varphi_{k,j}(x) = \frac{\psi_j^k(x)}{\sigma(x)}. \tag{221}$$

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 defined via

$$B = \bigcup_{\alpha \in \mathscr{O}} U_\alpha \tag{222}$$

subordinate to the open covering

$$B \subset \bigcup_{\alpha \in \mathscr{O}} U_\alpha. \tag{223}$$

We conclude the proof by observing that the resulting smooth partition of unity for B subordinate to the covering (223) is also a smooth partition of unity for A subordinate to the covering (214). □

Note carefully that Theorem 48 asserts that a finite partition of unity can be obtained when A is compact; however, Theorem 49, which applies in the general case, only asserts the existence of a countable partition of unity.

THEOREM 49. *Suppose that A is an arbitrary set in $\mathbb{R}^n$, and that*

$$A \subset \bigcup_{j=1}^{\infty} U_j \tag{224}$$

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 (224) such that

$$\operatorname{supp}(\psi_j) \subset U_j. \tag{225}$$

PROOF. We let $\varphi_1, \varphi_2, \ldots$ be a smooth partition of unity subordinate to the covering

$$A \subset \bigcup_{j=1}^{\infty} U_j \tag{226}$$

whose existence is ensured by Theorem 48. We define

$$I_1 = \{j \geq 1 : \operatorname{supp}(\varphi_j) \subset U_1\} \tag{227}$$

and, for $k > 1$,

$$I_k = \{j \geq 1 : \operatorname{supp}(\varphi_j) \subset U_k \ \text{ and } \ j \notin I_1 \cup I_2 \cup \cdots \cup I_{k-1}\}. \tag{228}$$

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

$$\psi_j(x) = \sum_{i \in I_j} \varphi_i(x). \tag{229}$$

□

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

Suppose that Ω 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 Ω 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

$$V = \{(y_1, \ldots, y_n) : -a_i < y_i < a_i \text{ for all } i = 1, 2, \ldots, n\} \tag{230}$$

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

$$V' = \{(y_1, \ldots, y_{n-1}) : -a_i < y_i < a_i \text{ for all } i = 1, 2, \ldots, n-1\}, \tag{231}$$

such that

$$-a_n < \psi(y_1, \ldots, y_{n-1}) < a_n \quad \text{for all} \quad (y_1, \ldots, y_{n-1}) \in V', \tag{232}$$

$$V \cap \Omega = \{(y_1, \ldots, y_n) : (y_1, \ldots, y_{n-1}) \in V' \text{ and } y_n < \psi(y_1, \ldots, y_{n-1})\}, \tag{233}$$

and

$$V \cap \partial\Omega = \{(y_1, \ldots, y_n) : (y_1, \ldots, y_{n-1}) \in V' \text{ and } y_n = \psi(y_1, \ldots, y_{n-1})\}. \tag{234}$$

A bounded open set Ω 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) ψ is a $C^{k,\alpha}$ mapping;

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

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

(4) $\partial\Omega \cap V = \{y \in V : \psi_n(y) = 0\}$, where ψ_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, [**10**] for an example of a domain which is a $C^{0,1}$ submanifold but not a $C^{0,1}$ domain.

EXERCISE 13. *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 Ω 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

$$V = \{(y_1, \ldots, y_n) : -a_i < y_i < a_i \text{ for all } i = 1, 2, \ldots, n\} \tag{235}$$

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

Similarly, Ω 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

$$V = \{(y_1, \ldots, y_n) : -a_i < y_i < a_i \text{ for all } i = 1, 2, \ldots, n\} \tag{236}$$

(2) $y - z$ is in Ω whenever $y \in \overline{\Omega} \cap V$ and z is contained in the cone

$$C = \{(t_1, \ldots, t_{n-1}, t_n) : \cot(\theta) \,|(t_1, \ldots, t_{n-1})| < t_n < h\}. \tag{237}$$

THEOREM 50. *Suppose that Ω is an open set in $\mathbb{R}^n$. Then Ω has the segment property if and only if its boundary is C^0.*

THEOREM 51. *Suppose that Ω is an open set in $\mathbb{R}^n$. Then Ω has the cone property if and only if its boundary is $C^{0,1}$.*

EXERCISE 14. *Prove Theorem 50.*

See [**10**] for a proof of Theorem 51. We will not make use of the following theorem, but it is a straightforward consequence of Theorem 51 that you might find useful.

THEOREM 52. *Suppose that Ω is a bounded open convex set in $\mathbb{R}^n$. Then Ω is a $C^{0,1}$ domain.*

Suppose that $\Omega \subset \mathbb{R}^n$ is either a $C^{k,0}$ domain with $k \geq 0$ an integer or a $C^{0,1}$ domain. Then we can define the outward-pointing normal derivative on the boundary $\partial\Omega$ as follows. Given $x \in \partial\Omega$, we let V be an open set containing x which has the properties (230) through (234). We let O be the $n \times n$ orthogonal matrix and ξ the vector in $\mathbb{R}^n$ such that

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = O \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + \xi. \tag{238}$$

The outward-pointing unit normal to the surface $y_n = \psi(y_1, \dots, y_{n-1})$ at the point

$$(y_1, \dots, y_{n-1}, \psi(y_1, \dots, y_{n-1}))$$

is

$$\tilde{\nu}(y_1, \dots, y_{n-1}) = \frac{1}{\sqrt{1 + |D\psi\,(y_1, \dots, y_{n-1})|^2}} \begin{pmatrix} -D\psi(y_1, \dots, y_{n-1}) \\ 1 \end{pmatrix}. \tag{239}$$

We define the outward-pointing unit normal to $\partial\Omega$ at the point $x_1, \dots, x_n$ to be

$$O^* \tilde{\nu}(y_1, \dots, y_{n-1}), \tag{240}$$

where $x_1, \dots, x_n$ and $y_1, \dots, y_n$ are related by (238).

We now define the surface integral

$$\int_{\partial\Omega} f(x)\; dS(x) \tag{241}$$

in the event that Ω is either a $C^{k,0}$ domain with $k \geq 1$ or a $C^{0,1}$ domain. We let $V_1, \dots, V_m$ be an open covering of $\partial\Omega$ each set of which has properties (230) through (234). We also let $\varphi_1, \dots, \varphi_m$ be a smooth partition of unity subordinate to this covering. Moreover, for each $j = 1, \dots, m$, we denote by $y_1^{(j)}, \dots, y_n^{(j)}$ the coordinate system associated with V_j, by ψ_j the function associated with V_j, and by $a_1^{(j)}, \dots, a_n^{(j)}$ the real numbers such that

$$V_j = \left\{ y_1^{(j)}, \dots, y_n^{(j)} : -a_1^{(j)} \leq y_1^{(j)} \leq a_1^{(j)}, \quad \dots, \quad -a_n^{(j)} \leq y_n^{(j)} \leq a_n^{(j)} \right\}. \tag{242}$$

Also, for each $j = 1, \dots, m$, we let O_j and ξ_j denote the orthogonal matrix and vector such that

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = O_j \begin{pmatrix} y_1^{(j)} \\ \vdots \\ y_n^{(j)} \end{pmatrix} + \xi_j, \tag{243}$$

and define $\tilde{f}_j$ via

$$\tilde{f}_j\left(y_1^{(j)}, \dots, y_n^{(j)}\right) = f(x_1 \dots, x_n)\varphi_j(x_1, \dots, x_n), \tag{244}$$

where $y_1^{(j)}, \dots, y_n^{(j)}$ and $x_1, \dots, x_n$ are related by (243). We say that $f : \partial\Omega \to \mathbb{C}$ is integrable if, for each $i = j, \dots, m$, the function

$$\tilde{f}_j\left(y_1^{(j)}, \dots, y_{n-1}^{(j)}, \psi_j(y_1^{(j)}, \dots, y_{n-1}^{(j)})\right) \tag{245}$$

is integrable on the set

$$V_j' = \left\{y_1^{(j)}, \dots, y_{n-1}^{(j)} : -a_1^{(j)} \le y_1^{(j)} \le a_1^{(j)}, \quad \dots, \quad -a_{n-1}^{(j)} \le y_{n-1}^{(j)} \le a_{n-1}^{(j)}\right\}. \tag{246}$$

In this event, we define the surface integral via

$$\begin{aligned} \int_{\partial\Omega} f(x)\ dS(x) = \sum_{j=1}^m \int_{V_j'} \Bigg(&\tilde{f}_j\left(y_1^{(j)}, \dots, y_{n-1}^{(j)}, \psi_j\left(y_1^{(j)}, \dots, y_{n-1}^{(j)}\right)\right) \times \\ &\sqrt{1 + \left|D\psi_j\left(y_1^{(j)}, \dots, y_{n-1}^{(j)}\right)\right|^2}\ \Bigg)\ dy_1^{(j)} dy_2^{(j)} \dots dy_{n-1}^{(j)}. \end{aligned} \tag{247}$$

It is tedious, but not difficult, to verify that this definition does not depend on the choice of the sets $V_1, \dots, V_m$ or the partition of unity $\varphi_1, \dots, \varphi_m$.

EXERCISE 15. *Write down suitable definitions for the outward-pointing unit normal for $\partial\Omega$ and the surface integral*

$$\int_{\partial\Omega} f(x)\ dS \tag{248}$$

in the event that Ω is a $C^{k,0}$ submanifold with boundary in $\mathbb{R}^n$.

2.10. Integration by Parts

We will make extensive use of the following two theorems.

THEOREM 53. *Suppose that Ω is a bounded $C^{0,1}$ domain, and that $u \in C^1\left(\overline{\Omega}\right)$. Then*

$$\int_\Omega D_i u(x)\ dx = \int_{\partial\Omega} u(x)\nu_i(x)\ dS(x), \tag{249}$$

where $\nu_i(x)$ denotes the i^{th} component of the outward-pointing unit normal at the point $x \in \partial\Omega$.

PROOF. We give a very careful proof of this theorem to illustrate certain techniques which are often used to prove results regarding $C^{k,\alpha}$ domains. In the future, similar proofs will be dealt with in a more cursory fashion.

For each x in $\partial\Omega$, let V_x be an open set containing x with the properties (230) through (234). For each x, let U_x be an open set containing x and compactly contained in V_x. Since $\partial\Omega$ is compact, we can choose a finite subcover $U_{x_1}, \ldots, U_{x_m}$ of $\partial\Omega$. For each $i = 1, \ldots, m$, we let $V_i = V_{x_i}$. Obviously, $V_1, \ldots, V_m$ also covers $\partial\Omega$. Now we let

$$V_0 = \Omega \setminus \{\overline{U_{x_1}} \cup \ldots \cup \overline{U_{x_m}}\}. \tag{250}$$

Then V_0 is an open set which is compactly contained in Ω, and $V_0, V_1, \ldots, V_m$ is an open covering of Ω. We now let $\varphi_0, \ldots, \varphi_m$ be a smooth partition of unity subordinate to the cover $V_0, \ldots, V_n$ of Ω.

We observe that

$$\int_\Omega D_i u(x)\ dx = \int_\Omega D_i \left(u \sum_{j=0}^n \varphi_j \right)(x)\ dx = \sum_{j=0}^n \int_\Omega D_i \left(\varphi_j u\right)(x)\ dx. \tag{251}$$

Since φ_0 is supported in $V_0 \subset \Omega$ and u is defined on all of V_0, we can extend $\varphi_0(x)u(x)$ to a function in $C^1(\mathbb{R}^n)$ by simply letting it be equal to 0 in the exterior of V_0. Now, we let

$$R = \{(x_1, x_2, \ldots, x_n) : -b_1 < x_1 < b_1, \quad -b_2 < x_2 < b_2, \quad \ldots, \quad -b_n < x_n < b_n\} \tag{252}$$

be a large cube containing V_0 whose boundary does not intersect V_0. Clearly, it is the case that

$$\begin{aligned} \int_\Omega D_i\left(\varphi_0 u\right)(x)\ dx &= \int_R D_i\left(\varphi_0 u\right)(x)\ dx \\ &= \int_{-b_1}^{b_1}\int_{-b_2}^{b_2}\cdots\int_{-b_n}^{b_n} D_i\left(\varphi_0 u\right)(x_1, \ldots, x_n)\ dx_1 dx_2 \cdots dx_n. \end{aligned} \tag{253}$$

By the fundamental theorem of calculus (see, for instance, Theorem 3.35 in [**8**]),

$$\begin{aligned} \int_{-b_i}^{b_i} D_i\left(\varphi_0 u\right)(x_1, \ldots, x_n)\ dx_i =& \varphi_0\left(x_1, \ldots, b_i, \ldots, x_n\right) u\left(x_1, \ldots, b_i, \ldots, x_n\right) \\ &- \varphi_0\left(x_1, \ldots, -b_i, \ldots, x_n\right) u\left(x_1, \ldots, -b_i, \ldots, x_n\right) \\ &= 0 - 0 = 0. \end{aligned} \tag{254}$$

It follows that from (253) and (254) that

$$\int_\Omega D_i\left(\varphi_0 u\right)(x)\ dx = 0. \tag{255}$$

We now fix $1 \leq j \leq n$ and recall how V_j was constructed. In particular, there exist a coordinate system $y_1, \ldots, y_n$ and a function ψ such that (230) through (234) hold. We let $O = (O_{ij})$ be an orthogonal matrix and ξ be a vector in $\mathbb{R}^n$ such that

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = O \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + \xi, \tag{256}$$

and define $\tilde{u}$ via the formula

$$\tilde{u}(y) = u\left(Oy+\xi\right)\varphi_j\left(Oy+\xi\right). \tag{257}$$

It follows that

$$\frac{\partial}{\partial x_i} = \sum_{k=1}^{n} \frac{\partial y_k}{\partial x_i}\frac{\partial}{\partial y_k} = \sum_{k=1}^{n} O_{ki}\frac{\partial}{\partial y_k}. \tag{258}$$

We observe that the integral

$$\int_\Omega D_i\left(\varphi_j(x)u(x)\right)\ dx \tag{259}$$

is equal to

$$\sum_{k=1}^{n} O_{ki}\ \int_{-a_1}^{a_1}\cdots\int_{-a_{n-1}}^{a_{n-1}}\int_{-a_n}^{\psi(y_1,\ldots,y_{n-1})} D_k\tilde{u}(y_1,\ldots,y_n)\ dy_1 dy_2\cdots dy_n. \tag{260}$$

Note that the Jacobian determinant of an orthogonal transformation is 1. For $k=1,\ldots,n-1$, it is the case that

$$\begin{aligned} D_k\left(\int_{-a_n}^{\psi(y_1,\ldots,y_{n-1})}\tilde{u}(y)\ dy_n\right) &= \int_{-a_n}^{\psi(y_1,\ldots,y_{n-1})} D_k\tilde{u}(y)\ dy_n \\ &\quad + D_k\psi(y_1\ldots,y_{n-1})\tilde{u}(y_1,\ldots,y_{n-1},\psi(y_1,\ldots,y_{n-1})). \end{aligned} \tag{261}$$

By integrating this expression first with respect to y_k, then with respect to the rest of the coordinates $y_1,\ldots,y_{n-1}$ and making use of the compact support of $\tilde{u}$, we see that

$$\begin{aligned} &\int_{-a_1}^{a_1}\cdots\int_{-a_{n-1}}^{a_{n-1}}\int_{-a_n}^{\psi(y_1,\ldots,y_{n-1})} D_k\tilde{u}(y)\ dy_1dy_2\cdots dy_n \\ &= -\int_{-a_1}^{a_1}\cdots\int_{-a_{n-1}}^{a_{n-1}}\tilde{u}\left(y_1\ldots,y_{n-1},\psi\left(y_1,\ldots,y_{n-1}\right)\right)D_k\psi(y_1,\ldots,y_{n-1})\ dy_1dy_2\cdots dy_{n-1} \end{aligned} \tag{262}$$

for $k=1,\ldots,n-1$. On the other hand, it is readily apparent that

$$\begin{aligned} &\int_{-a_1}^{a_1}\cdots\int_{-a_{n-1}}^{a_{n-1}}\int_{-a_n}^{\psi(y_1,\ldots,y_{n-1})} D_n\tilde{u}(y_1,\ldots,y_n)\ dy_1dy_2\cdots dy_n \\ &= \int_{-a_1}^{a_1}\cdots\int_{-a_{n-1}}^{a_{n-1}}\tilde{u}(y_1,\ldots,y_{n-1},\psi(y_1,\ldots,y_{n-1}))\ dy_1dy_2\cdots dy_{n-1}. \end{aligned} \tag{263}$$

Now, we let $\tilde{\nu}$ denote the outward pointing unit normal to the surface defined via $y_n = \psi(y_1,\ldots,y_{n-1})$; that is,

$$\tilde{\nu}(y_1,\ldots,y_{n-1}) = \frac{1}{\sqrt{1+|D\psi(y_1,\ldots,y_{n-1}|^2}}\begin{pmatrix} -D\psi(y_1,\ldots,y_{n-1}) \\ 1 \end{pmatrix} \tag{264}$$

(here $D\psi$ refers to the gradient of ψ). From (260), (262) and (263), we see that (259) is equal to

$$\int_{-a_1}^{a_1}\cdots\int_{-a_{n-1}}^{a_{n-1}} \tilde{u}\left(y_1\ldots,\psi(y_1,\ldots,y_{n-1})\right) \tag{265}$$
$$\left(O^*\tilde{\nu}\left(y_1,\ldots,y_{n-1}\right)\right)_i\sqrt{1+\left|D\psi\left(y_1,\ldots,y_{n-1}\right)\right|^2}\;dy_1dy_2\cdots dy_{n-1},$$

where $\left(O^*\tilde{\nu}\right)_i$ denotes the i^{th} component of the vector $O^*\tilde{\nu}$. Since

$$\left(O^*\tilde{\nu}(y_1,\ldots,y_{n-1})\right)_i = \nu_i(x_1,\ldots,x_n), \tag{266}$$

(263) is simply the expression obtained directly from the definition of surface integral

$$\int_{\partial\Omega} u(x)\varphi_j(x)\nu_i(x)\;dS(x). \tag{267}$$

In particular,

$$\int_\Omega D_i\left(u(x)\varphi_j(x)\right)\;dx = \int_{\partial\Omega} u(x)\varphi_j(x)\nu_i(x)\;dS(x). \tag{268}$$

We sum (268) over $j=1,\ldots,m$ to obtain the conclusion of the theorem. □

THEOREM 54. *(Integration by parts) Suppose that Ω is a bounded $C^{0,1}$ domain, and that $u,v\in C^1\left(\overline{\Omega}\right)$. Then*

$$\int_\Omega D_iu(x)v(x)\;dx = -\int_\Omega u(x)D_iv(x)\;dx + \int_{\partial\Omega} u(x)v(x)\nu_i(x)\;dS(x), \tag{269}$$

where $\nu_i(x)$ denotes the i^{th} componenet of the outward-pointing unit normal at the point $x\in\partial\Omega$.

PROOF. This follows immediately from the preceding theorem and the fact that $D_i(uv) = \left(D_iu\right)v + u\left(D_iv\right)$. □

Theorem 54 is a special case of the divergence theorem and we will, by a slight abuse of terminology, sometimes refer to it as the divergence theorem.

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

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

$$\int_\Omega u(x) D^\alpha \psi(x)\ dx = (-1)^{|\alpha|} \int_\Omega v(x)\psi(x)\ dx \tag{270}$$

for all ψ in $C_c^\infty(\Omega)$.

Exercise 16. *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 17. *Show that if $u \in C^k(\Omega)$ and $|\alpha| \le k$ is a multi-index, then u has weak derivatives of orders 0 through k, and the α^{th} weak derivative of u is the a^{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 α^{th} weak derivative of a function $u \in L^1_{\text{loc}}(\Omega)$ by $D^\alpha u$.

Exercise 18. *Suppose that Ω is an open subset of $\mathbb{R}^n$, and that α 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 19. *Suppose that Ω 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 20. *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 21. *Suppose that Ω is an open subset of $\mathbb{R}^n$, that α, β and γ are multi-indices such that $\alpha = \beta + \gamma$, and that $u \in L^1_{loc}(\Omega)$ such that D^α, D^β and D^γ exist. Show that $D^\alpha u = D^\beta D^\gamma u = D^\gamma D^\beta u$.*

Exercise 22. *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 2.7.*

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 α^{th} weak derivative, then both $\eta_h * u$ and $\eta_h * D^\alpha u$ are infinitely differentiable functions, and hence are pointwise defined.

THEOREM 55. *Suppose that Ω is an open set in $\mathbb{R}^n$, that u is an element of $L^1_{loc}(\Omega)$ whose α^{th} weak derivative exists. Then for all $x \in \Omega$ and $h > 0$ such that $\overline{B_h(x)} \subset \Omega$,*

$$D^\alpha \left(\eta_h * u\right)(x) = \eta_h * \left(D^\alpha u\right)(x). \tag{271}$$

PROOF. Suppose that $h > 0$, and let $\Omega' = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > h\}$. Then $\eta_h * u$ is defined for all $x \in \Omega'$; in fact, it is an element of $C^\infty(\Omega')$. Moreover, for all $x \in \Omega'$ we have

$$\begin{aligned} D^\alpha \left(\eta_h * u\right)(x) &= h^{-n} \int_\Omega D_x^\alpha \eta \left(\frac{x-y}{h}\right) u(y)\ dy \\ &= h^{-n}(-1)^{|\alpha|} \int_\Omega D_y^\alpha \eta \left(\frac{x-y}{h}\right) u(y)\ dy \\ &= h^{-n} \int_\Omega \eta \left(\frac{x-y}{h}\right) D^\alpha u(y)\ dy \\ &= \eta_h * D^\alpha u(x), \end{aligned} \tag{272}$$

which estalbishes the theorem. □

3.2. Sobolev Spaces

Suppose that Ω is an open set in $\mathbb{R}^n$, that $k \geq 0$ is an integer, and that $1 \leq p \leq \infty$ 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 α with $|\alpha| \leq k$, $D^\alpha u$ exists in the weak sense and belongs to $L^p(\Omega)$. For $1 \leq p < \infty$, the space $W^{k,p}(\Omega)$ is a Banach space with respect to the norm

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha|=0}^{k} \|D^\alpha u\|^p_{L^p(\Omega)} \right)^{1/p}, \tag{273}$$

and $W^{k,\infty}(\Omega)$ is a Banach space with respect to the norm

$$\|u\|_{W^{k,\infty}(\Omega)} = \sup_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)}. \tag{274}$$

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

$$(u, v) = \sum_{|\alpha|=0}^{k} \int_\Omega D^\alpha u(x)\ D^\alpha v(x)\ dx. \tag{275}$$

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 $H_0^k(\Omega)$. We will often encounter the dual space of $H_0^1(\Omega)$, which we denote by $H^{-1}(\Omega)$. Note that $H^{-1}(\Omega)$ need not coincide with $(H^1(\Omega))^*$.

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

$$Du(x) = \begin{pmatrix} D_1 u(x) \\ D_2 u(x) \\ \vdots \\ D_n u(x) \end{pmatrix} \tag{276}$$

Moreover, we define

$$\|Du\|_p = \left(\|D_1 u(x)\|_p^p + \cdots \|D_n u(x)\|_2^p\right)^{1/p}. \tag{277}$$

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

$$D_i D_j u(x) \tag{278}$$

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

$$\left(\sum_{i,j=1}^n \|D_i D_j u(x)\|_p^p\right)^{1/p}. \tag{279}$$

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 to establish, in Theorem 58, that $C^\infty(\Omega)$ functions are dense in $W^{k,p}(\Omega)$.

THEOREM 56. *Suppose that Ω 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 $\psi \in C_c^\infty(\Omega)$. Then the product ψu 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 $\varphi \in C_c^\infty(\Omega)$. Since $\psi\varphi$ is an element of $C_c^\infty(\Omega)$,

$$\int_\Omega D^\alpha u(x)\psi(x)\varphi(x)\ dx = (-1)^{|\alpha|}\int_\Omega u(x) D^\alpha\left(\psi\varphi\right)(x)\ dx. \tag{280}$$

We insert the identity

$$D^\alpha(\psi\varphi)(x)\ dx = \sum_{\beta\leq\alpha}\binom{\alpha}{\beta} D^{\alpha-\beta}\psi(x) D^\beta\varphi(x) \tag{281}$$

into (280) to obtain

$$\int_\Omega D^\alpha u(x)\psi(x)\varphi(x)\ dx = (-1)^{|\alpha|}\int_\Omega u(x)\left(\sum_{\beta\leq\alpha}\binom{\alpha}{\beta} D^{\alpha-\beta}\psi(x) D^\beta\varphi(x)\right)\ dx. \tag{282}$$

We rearrange (282) as

$$\begin{aligned}\int_\Omega u(x)\psi(x)D^\alpha\varphi(x)\ dx = (-1)^{|\alpha|}\int_\Omega D^\alpha u(x)\psi(x)\varphi(x)\ dx - \\ \sum_{\beta<\alpha}\binom{\alpha}{\beta}\int_\Omega u(x)D^{\alpha-\beta}\psi(x)D^\beta\varphi(x)\ dx.\end{aligned} \tag{283}$$

and apply the induction hypothesis to each of the terms in the sum on the right-hand side of (283) to obtain

$$\begin{aligned}\int_\Omega u(x)\psi(x)D^\alpha\varphi(x)\ dx =& (-1)^{|\alpha|}\int_\Omega D^\alpha u(x)\psi(x)\varphi(x)\ dx - \\ & (-1)^{|\alpha|}\sum_{\beta<\alpha}(-1)^{|\beta|+|\alpha|}\binom{\alpha}{\beta}\int_\Omega D^\beta\left(u(x)D^{\alpha-\beta}\psi(x)\right)\varphi(x)\ dx.\end{aligned} \tag{284}$$

We conclude that $D^\alpha(\psi u)$ exists and equals

$$D^\alpha u(x)\psi(x) - \sum_{\beta<\alpha}(-1)^{|\beta|+|\alpha|}\binom{\alpha}{\beta}D^\beta\left(u(x)D^{\alpha-\beta}\psi(x)\right). \tag{285}$$

By assumption $D^\alpha u \in L^p(\Omega)$ and our induction hypothesis implies that each of the functions

$$D^\beta\left(u(x)D^{\alpha-\beta}\psi(x)\right) \tag{286}$$

is an element of $L^p(\Omega)$. Since the products of $C_c^\infty(\Omega)$ functions with $L^p(\Omega)$ functions are in $L^p(\Omega)$, we conclude that $D^\alpha(\psi u)$ is an element of $L^p(\Omega)$. $\square$

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 $u \in W^{k,p}_{\text{loc}}$ if $u_n \to u$ in $W^{k,p}(\Omega')$ whenever $\Omega' \subset\subset \Omega$. In light of this definition, the conclusion of Exercise 22 can be rephrased as saying that $W^{k,\infty}_{\text{loc}}(\Omega)$ coincides with $C^{0,1}(\Omega)$ when Ω is an open set in $\mathbb{R}$. Later, will we see that

$$W^{1,\infty}_{\text{loc}}(\Omega) = C^{0,1}(\Omega) \tag{287}$$

for any open subset Ω of $\mathbb{R}^n$. Moreover, it follows from the standard embedding theorems for Sobolev spaces that

$$W^{k,\infty}_{\text{loc}}(\Omega) = C^{k-1,1}(\Omega) \tag{288}$$

for all positive integers k. It is not always the case that $W^{k,\infty}(\Omega)$ coincides with $C^{k-1,1}\left(\overline{\Omega}\right)$, although it does so under mild regularity assumptions on Ω. Not surprisingly, we will sometimes use the notation $H^k_{\text{loc}}(\Omega)$ to refer to $W^{k,2}_{\text{loc}}(\Omega)$.

3.3. Approximation by Smooth Functions

In this section, we will two key results (Theorems 58 and 62) 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 posses those properties and then appealing to the density of smooth functions in $W^{k,p}(\Omega)$.

THEOREM 57. *Suppose that* Ω *is an open set in* $\mathbb{R}^n$*, and that* η_h *is the standard mollifier. Suppose also that* $u \in W^{k,p}_{loc}(\Omega)$ *with* $1 \leq p < \infty$ *a real number and* $k \geq 0$ *an integer. Then the sequence of functions* $\eta_h * u$ *converges to* u *in* $W^{k,p}_{loc}(\Omega)$*.*

PROOF. Suppose that α is a multi-index such that $0 \leq |\alpha| \leq k$, and that $\Omega' \subset\subset \Omega$. According to Theorem 55, for all sufficiently small h and all $x \in \Omega'$

$$D^\alpha (\eta_h * u)(x) = \eta_h * D^\alpha u(x). \tag{289}$$

By integrating both sides of (289) over Ω', we obtain

$$\int_{\Omega'} |D^\alpha (\eta_h * u - u)(x)|^p \, dx = \int_{\Omega'} |\eta_h * D^\alpha u(x) - D^\alpha u(x)|^p \, dx. \tag{290}$$

Now we invoke Theorem 45 to conclude that

$$\lim_{h\to 0} \int_{\Omega'} |\eta_h * D^\alpha u(x) - D^\alpha u(x)|^p \, dx = 0. \tag{291}$$

We combine (290) with (291) to obtain

$$\lim_{h\to 0} \| D^\alpha (\eta_h * u) - D^\alpha u \|_{L^p(\Omega')} = 0, \tag{292}$$

from which we conclude that $\eta_h * u$ converges in $W^{k,p}(\Omega')$ to u. $\square$

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 to produce a sequence of smooth functions approximating an element of $W^{k,p}(\Omega)$.

EXERCISE 23. *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 58 (Meyers-Serrin). *Suppose that* Ω *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

$$\Omega = \bigcup_{k=1}^{\infty} \Omega_k \tag{293}$$

and we let $\{\psi_k : k = 1, 2, \ldots\}$ be a smooth partition of unity subordinate to the covering (293). That is, $\psi_1, \psi_2, \ldots$ is a sequence in $C_c^\infty(\mathbb{R}^n)$ with the following properties:

(1) $\operatorname{supp}(\psi_j) \subset \Omega_j$ 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 x 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

$$u_j(x) = u(x)\psi_j(x). \tag{294}$$

We now let $\epsilon > 0$. By Theorem 56, u_j is an element of $W^{k,p}(\Omega)$. Moreover, Ω_j is compact contained in Ω. So we can apply Theorem 57 to see that

$$\lim_{h\to 0} \|\eta_h * u_j - u_j\|_{W^{k,p}(\Omega_{j+1})} = 0. \tag{295}$$

Since u_j has compact support contained in Ω_j, for sufficiently small h the support of $\eta_h * u_j$ is also contained in Ω_j. So, for each $j \geq 1$, there exists a real number h_j such that

$$\operatorname{supp}(\eta_{h_j} * u_j) \subset \Omega_j \tag{296}$$

and

$$\left\|\eta_{h_j} * u_j - u_j\right\|_{W^{k,p}(\Omega_j)} < \frac{\epsilon}{2^j}. \tag{297}$$

Of course, since both u_h and $\eta_{h_j} * u_j$ have support contained in Ω_j, we have

$$\left\|\eta_{h_j} * u_j - u_j\right\|_{W^{k,p}(\Omega)} = \left\|\eta_{h_j} * u_j - u_j\right\|_{W^{k,p}(\Omega_j)} < \frac{\epsilon}{2^j}. \tag{298}$$

Since only a finite number of the functions $u_1, u_2, \ldots$ are nonzero in each of the sets Ω_j, only a finite number of the functions $\eta_{h_1} * u_1, \eta_{h_2} * u_2, \ldots$ are nonzero in each of the sets Ω_{j+1} and

$$v(x) = \sum_{j=1}^{\infty} \eta_{h_j} * u_j(x). \tag{299}$$

defines a $C^\infty(\Omega)$ function. We combine (298) and (299) to arrive at

$$\|v - u\|_{W^{k,p}(\Omega)} = \left\|\sum_{j=1}^{\infty} \eta_\epsilon * u_j - \sum_{j=1}^{\infty} u_j\right\|_{W^{k,p}(\Omega)} \leq \sum_{j=1}^{\infty} \|\eta_\epsilon * u_j - u_j\|_{W^{k,p}(\Omega)} \leq \sum_{j=1}^{\infty} \frac{\epsilon}{2^j} = \epsilon,$$

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 ϕ such that $\|\phi' - u'\|_\infty < 1/4$, so u cannot be the limit of a sequence of infinitely differentiable functions in $W^{k,p}(\Omega)$. Note that in many books (e.g., [**1**]), $W^{k,p}(\Omega)$ is defined for $k = 1, 2, \ldots$ and $1 \leq p < \infty$ as the completion of $C^\infty(\Omega)$ with respect to the norm

$$\|\psi\|_{k,p} = \left(\sum_{|\alpha|\leq k} \|D^\alpha \psi\|^p_{L^p(\Omega)}\right)^{1/p}, \tag{300}$$

and $W^{k,\infty}(\Omega)$ is defined as the completion of $C^\infty(\Omega)$ with respect to the norm

$$\|\psi\|_{k,\infty} = \sup_{|\alpha|\leq k} \|D^\alpha \psi\|_{L^\infty(\Omega)}. \tag{301}$$

The preceding discussion shows that this definition differs from ours in the case $p = \infty$ (see, for instance, Theorem 3.16 in [**1**]).

As an application Theorem 58, we now sharpen Theorem 56.

THEOREM 59. *Suppose that* Ω *is an open set in* $\mathbb{R}^n$, *and that* $1 \le p < \infty$ *and* $k \ge 1$. *Suppose also that* $\psi \in C^{k-1,1}(\overline{\Omega})$. *Then the mapping*

$$u(x) \to \psi(x)u(x) \tag{302}$$

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 α is a multi-index such that $|\alpha| \le k$. We apply the standard chain rule from multivariable calculus to conclude that

$$D^\alpha(\psi u)(x) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta}\psi(x) D^\beta u(x) \tag{303}$$

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

$$\left|D^\beta \psi(x)\right| \le C \tag{304}$$

for all $|\beta| \le k$ and almost all $x \in \Omega$. We combine (303) with (304) to conclude that

$$|D^\alpha(\psi u)(x)| \le C \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \left|D^\beta u(x)\right| \tag{305}$$

for almost all $x \in \Omega$. It clearly follows from(305) that $D^\alpha(\psi u) \in L^p(\Omega)$, and that there exists a constant C' which depends on α and ψ but not u such that

$$\|D^\alpha(\psi u)(x)\|_p \le C' \|u\|_{W^{k,p}(\Omega)}. \tag{306}$$

We conclude that there exists $C'' > 0$ such that

$$\|\psi u\|_{W^{k,p}(\Omega)} \le C'' \|u\|_{W^{k,p}(\Omega)} \tag{307}$$

for all $u \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$. It now follows from Theorem 58 that the mapping (302) 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 the mapping defined for $u \in W^{k,p}(\Omega)$ via the formula

$$u(x) \to \psi(x)u(x). \tag{308}$$

□

The following two theorems are established in much the same fashion as Theorem 59 — by applying the theorems of Section 2.7.5 and then appealing to Theorem 59.

THEOREM 60. *Suppose that* Ω *is a subset of* $\mathbb{R}^n$, *and that* Ω' *is a subset of* $\mathbb{R}^m$. *Suppose also that* $k \ge 1$ *is an integer, that* $1 \le p < \infty$ *is a real number, and that* $\psi : \Omega \to \Omega'$ *is a* $C^{k-1,1}$ *mapping. Then the map*

$$u(x) \to u(\psi(x)) \tag{309}$$

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

THEOREM 61. *Suppose that* Ω *is a subset of* $\mathbb{R}^n$, *and that* Ω' *is a subset of* $\mathbb{R}^m$. *Suppose also that* $k \ge 1$ *is an integer, that* $1 \le p < \infty$ *is a real number, and that* $\psi : \Omega \to \Omega'$ *is a bijective mapping such that* ψ *and* ψ^{-1} *are* $C^{k-1,1}$ *mappings. Then there exists a constant* $C > 0$ *such* *that*

$$C^{-1}\|u \circ \psi\|_{W^{k,p}(\Omega)} \leq \|u\|_{W^{k,p}(\Omega')} \leq C\|u \circ \psi\|_{W^{k,p}(\Omega)} \tag{310}$$

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 $C^\infty\left(\overline{\Omega}\right)$ is always contained in $W^{k,p}(\Omega)$ and in light of Theorem 58 it is natural to ask if $C^\infty\left(\overline{\Omega}\right)$ 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

$$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0. \end{cases} \tag{311}$$

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

Theorem 62. *Suppose that Ω 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 $C^\infty\left(\overline{\Omega}\right)$ is dense in $W^{k,p}(\Omega)$.*

Proof. We let $u \in W^{k,p}(\Omega)$. We observe that since the boundary of $\partial\Omega$ is continuous, Ω has the segment property (see Section 2.9). 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

$$\partial\Omega \subset \bigcup_{x \in \partial\Omega} 1/2U_x, \tag{312}$$

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

$$\partial\Omega \subset \bigcup_{j=1}^m 1/2U_{x_j}. \tag{313}$$

For each $j = 1, \ldots, m$, we set $U_j = U_{x_j}$ and $V_j = 1/2U_{x_j} \cap \Omega$ and let γ_j denote the vector γ_{x_j}. We also choose an open set V_0 such that $V_0 \subset\subset \Omega$ and

$$\Omega \subset \bigcup_{j=0}^m V_j. \tag{314}$$

Next, we choose a smooth partition of unity $\psi_0, \ldots, \psi_m$ subordinate to the covering

$$\Omega \subset V_0 \cup \bigcup_{j=1}^m 1/2U_j \tag{315}$$

so that

(1) $\psi_0 \in C_c^\infty(V_0)$;

(2) $\psi_j \in C_c^\infty(1/2U_j)$ for each $j = 1, \ldots, m$;

(3) $0 \leq \psi_j(x) \leq 1$ for all $j = 0, 1, \ldots, m$ and $x \in \mathbb{R}^n$;

(4) $\sum_{j=0}^{m} \psi_j(x) = 1$ for all $x \in \Omega$.

Moreover, for each $j = 0, 1, \ldots, m$ we define $u_j : V_j \to \mathbb{R}$ via the formula

$$u_j(x) = \psi_j(x)u(x). \tag{316}$$

Note that the support of ψ_j is not necessarily contained in $V_j = 1/2U_j \cap \Omega$.

We now fix $\epsilon > 0$; we will construct a function $w \in C^\infty\left(\overline{\Omega}\right)$ such that

$$\|u - w\|_{W^{k,p}(\Omega)} < \epsilon. \tag{317}$$

To that end, we define the function $w_0 : V_0 \to \mathbb{R}$ via the formula

$$w_0(x) = \eta_{h_0} * u_0(x), \tag{318}$$

where η_h denotes the standard mollifier and h_0 is chosen so that

$$\|w_0 - u_0\|_{W^{k,p}(V_0)} < \frac{\epsilon}{2}. \tag{319}$$

For each $j = 1, \ldots, m$, all sufficiently small $h > 0$ and all $x \in \overline{V_j}$, we define $u_{j,h} : \overline{V_j} \to \mathbb{R}$ by the formula

$$u_{j,h}(x) = u_j(x + h\gamma_j) \tag{320}$$

and $v_{j,h} \in C^\infty\left(\overline{V_j}\right)$ via the formula

$$v_{j,h}(x) = \eta_h * u_{j,h}(x). \tag{321}$$

We observe that

$$\|D^\alpha v_{j,h} - D^\alpha u_j\|_{L^p(V_j)} \leq \|D^\alpha v_{j,h} - D^\alpha u_{j,h}\|_{L^p(V_j)} + \|D^\alpha u_{j,h} - D^\alpha u_j\|_{L^p(V_j)} \tag{322}$$

for all multi-indices $|\alpha| \leq k$. Since

$$D^\alpha u_{j,h}(x) = (D^\alpha u_j)(x + h\gamma_j) \tag{323}$$

and translation is continuous in the L^p norm,

$$\|D^\alpha u_{j,h} - D^\alpha u_j\|_{L^p(V_j)} \to 0 \quad \text{as} \quad h \to 0 \tag{324}$$

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

$$\|D^\alpha v_{j,h} - D^\alpha u_{j,h}\|_{L^p(V_j)} \to 0 \quad \text{as} \quad h \to 0 \tag{325}$$

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

$$\|w_j - u_j\|_{W^{k,p}(V_j)} \leq \frac{\epsilon}{2^{j+1}}. \tag{326}$$

We now define the function $w \in C^\infty\left(\overline{\Omega}\right)$ via the formula

$$w(x) = \sum_{j=0}^{m} \psi_j(x)w_j(x). \tag{327}$$

We combine (316), (319) and (326) to conclude that

$$\begin{aligned}
\|u-w\|_{W^{k,p}(\Omega)} &= \left\|\sum_{j=0}^m \psi_j u - \sum_{j=0}^m \psi_j w\right\|_{W^{k,p}(\Omega)} \\
&\leq \sum_{j=0}^m \|\psi_j u - \psi_j w\|_{W^{k,p}(\Omega)} \\
&= \sum_{j=0}^m \|u_j - w_j\|_{W^{k,p}(V_j)} \\
&\leq \epsilon.
\end{aligned} \tag{328}$$

□

EXERCISE 24. *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 η_h is the standard mollifier. Show that*

$$\|\eta_h * u_h - u\|_p \to 0 \quad as \quad h \to 0. \tag{329}$$

We shall make frequent use of the following theorems, which are straightforward to prove using the results of this section.

THEOREM 63. *Suppose that $1 \leq p < \infty$ is a real number, that $k \geq 0$ is an integer, that Ω is an open set, that*

$$\Omega \subset \bigcup_{j=1}^N U_j \tag{330}$$

is a covering of Ω by open sets, and that $\{\psi_j\}$ is a smooth partition of unity subordinate to the covering (330). Then u is an element of $W^{k,p}(\Omega)$ if and only if each of the functions

$$\psi_j(x)u(x) \tag{331}$$

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

$$C^{-1}\|u\|_{W^{k,p}(\Omega)} \leq \sum_{j=1}^N \|\psi_j u\|_{W^{k,p}(U_j)} \leq C\|u\|_{W^{k,p}(\Omega)} \tag{332}$$

for all $u \in W^{k,p}(\Omega)$; that is, the $\|u\|_{W^{k,p}(\Omega)}$ norm is equivalent to the norm

$$\sum_{j=1}^N \|\psi_j u\|_{W^{k,p}(U_j)}. \tag{333}$$

THEOREM 64. *Suppose that $1 \leq p < \infty$ is a real number, that $k > 0$ is an integer, that Ω is a bounded open set in $\mathbb{R}^n$ which is $C^{k-1,1}$ domain.*

3.4. The Trace Operator

Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded open set. We are ultimately interested in solving boundary value problems given on Ω. 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 Ω. However, using Theorem 62 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 Ω is sufficiently regular.

THEOREM 65. *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*

$$\mathscr{T} : W^{1,p}(\Omega) \to L^p(\partial\Omega) \tag{334}$$

such that

$$\mathscr{T}[u] = u|_{\partial\Omega} \tag{335}$$

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

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

$$\partial\Omega \subset \bigcup_{j=1}^{m} U_j \tag{336}$$

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

(1) ψ_j and ψ_k^{-1} are Lipschitz mappings;

(2) $U_j \cap \Omega = \psi_j\left(V_j \cap \{(x_1, \ldots, x_n) : x_n > 0\}\right)$; and

(3) $U_j \cap \partial\Omega = \psi_j\left(V_j \cap \{(x_1, \ldots, x_n) : x_n = 0\}\right)$.

We choose a smooth partition of unity $\{\gamma_j\}$ subordinate to the covering (336) and define, for each $j = 1, \ldots, m$, the function $\eta_j : V_j \to \mathbb{R}$ via the formula

$$\eta_j(x) = \gamma_j(\psi_j(x)) \tag{337}$$

Suppose that $u \in C^1\left(\overline{\Omega}\right)$, and for each $j = 1, \ldots, m$, define the functions $u_j : V_j \to \mathbb{R}$ via the formula

$$u_j(x) = u(\psi_j(x))\eta_j(x) \tag{338}$$

We will show that there exists a positive constant C not depending on u such that

$$\|u_j\|_{L^p(V_j \cap \{x_n = 0\})} \leq C \|u_j\|_{W^{1,p}(V_j \cap \{x_n > 0\})}. \tag{339}$$

The well-known change of variables formula from measure theory implies that there exists a constant C_1 not depending on u such that

$$\|\gamma_j u\|_{L^p(\partial\Omega \cap U_j)} \leq C_1 \|u_j\|_{L^p(V_j \cap \{x_n = 0\})} \tag{340}$$

and Theorem 60 implies that there exists a constant C_2 not depending on u such that

$$\|u_j\|_{W^{1,p}(V_j \cap \{x_n > 0\})} \leq C_2 \|\gamma_j u\|_{W^{1,p}(\Omega \cap U_j)}. \tag{341}$$

So it will follow from (339) that

$$\begin{aligned}
\|u\|_{L^p(\partial\Omega)} &= \left\|\sum_{j=1}^m \gamma_j u\right\|_{L^p(\partial\Omega)} \\
&\leq \sum_{j=1}^m \|\gamma_j u\|_{L^p(\partial\Omega\cap U_j)} \\
&\leq C_1 \sum_{j=1}^m \|u_j\|_{L^p(V_j\cap\{x_n=0\})} \\
&\leq CC_1 \sum_{j=1}^m \|u_j\|_{W^{1,p}(V_j\cap\{x_n>0\})} \\
&\leq CC_1C_2 \sum_{j=1}^m \|\gamma_j u\|_{W^{1,p}(\Omega\cap U_j)} \\
&\leq mCC_1C_2\|u\|_{W^{1,p}(\Omega)}.
\end{aligned} \tag{342}$$

We will rarely make these kinds of arguments explicitly in the future; we will instead simply say that using a partition of unity and the properties of the domain Ω, we reduce the theorem to proving it for the case of a function given on V_j.

We observe that the support of η_j — and hence u_j — is compactly contained in V_j, so that when we apply the divergence theorem we obtain

$$\begin{aligned}
\int_{V_j\cap\{x_n=0\}} |u_j(x)|^p \ dx &= -\int_{V_j\cap\{x_n>0\}} \frac{\partial}{\partial x_n}\left(|u_j(x)|^p\right) \ dx \\
&= -\int_{V_j\cap\{x_n>0\}} p\,|u_j(x)|^{p-1}\,\mathrm{sign}(u_j(x))\frac{\partial u_j(x)}{\partial x_n} \ dx \\
&\leq \int_{V_j\cap\{x_n>0\}} p\,|u_j(x)|^{p-1}\left|\frac{\partial u_j(x)}{\partial x_n}\right| \ dx.
\end{aligned} \tag{343}$$

By letting

$$q = \frac{1}{1-\frac{1}{p}}, \quad a = \left|\frac{\partial u_j(x)}{\partial x_n}\right| \quad \text{and} \quad b = |u_j(x)|^{p-1} \tag{344}$$

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

$$\begin{aligned}
\left|\frac{\partial u_j(x)}{\partial x_n}\right| |u_j(x)|^{p-1} &\leq \frac{1}{p}\left|\frac{\partial u_j(x)}{\partial x_n}\right|^p + \frac{1}{q}|u_j(x)|^{q(p-1)} \\
&= \frac{1}{p}\left|\frac{\partial u_j(x)}{\partial x_n}\right|^p + \frac{1}{q}|u_j(x)|^p.
\end{aligned} \tag{345}$$

We conclude, by inserting (345) into (343), that for each $j = 1, \ldots, m$ there exists $C_j > 0$ such that

$$\int_{\{x_n=0\}\cap V_j} |u_j(x)|^p \ dx \leq C_j\|u_j\|_{W^{1,p}(V_j\cap\{x_n>0\})}. \tag{346}$$

It now follows from Theorem 62 that $\mathscr{T}$ extends by continuity to a bounded linear mapping $W^{1,p}(\Omega) \to W^{1,p}(\Omega)$. If $u \in C\left(\overline{\Omega}\right) \cap W^{1,p}(\Omega)$, then the sequence $\{u_m\}$ of $C^\infty\left(\overline{\Omega}\right)$ functions constructed in Theorem 62 converge uniformly on $\overline{\Omega}$ to u. We conclude from this observation that $u_m \to u$ on the boundary of $\partial\Omega$, so that $\mathscr{T}(u) = u|_{\partial\Omega}$. $\square$

The mapping $\mathscr{T}$ is called the trace operator, $\mathscr{T}[u]$ is known as the trace of the function u, and the image of $W^{k,p}(\Omega)$ under the mapping $\mathscr{T}$ is known as the trace space of $W^{k,p}(\Omega)$. The mapping $\mathscr{T}$ is not a surjection onto $L^p(\Omega)$. Later, we will characterize the trace space of $H^l(\Omega)$ when Ω is a Lipschitz domain.

THEOREM 66. *Suppose that Ω is a Lipschitz domain in $\mathbb{R}^n$, and that $1 \le p < \infty$. Then $u \in W_0^{1,p}(\Omega)$ if and only if $\mathscr{T}[u] = 0$.*

PROOF. If $u \in C_c^\infty(\Omega)$, then it is a consequence of (65) that $\mathscr{T}[u] = 0$. Since functions in $W_0^{k,p}(\Omega)$ are the limits in $W^{k,p}(\Omega)$ norm of $C_c^\infty(\Omega)$ functions and $\mathscr{T}$ is continuous, it follows that the trace of a function in $W_0^{1,p}(\Omega)$ is 0.

We will now show that if $u \in W^{1,p}(\Omega)$ such that $\mathscr{T}[u] = 0$, then u is the limit in $W^{1,p}(\Omega)$ of a sequence of $C_c^\infty(\Omega)$ functions. Via a localization argument virtually identical to that used in the proof Theorem 65, we see that it suffices to show that if U is an open ball in $\mathbb{R}^n$ centered at 0, $V = \mathbb{R}_+^n \cap U$ and u is an element of $W^{1,p}(V)$ whose support in bounded away from ∂U and whose trace on ∂V is 0, then u is the limit of a sequence of $C_c^\infty(V)$ functions.

Since the trace of u is 0, there exist functions $u_m \in C^1\left(\overline{V}\right)$ such that

$$\|u_m - u\|_{W^{1,p}(V)} \to 0 \tag{347}$$

and $\|u_m\|_{L^p(\partial V)} \to 0$. Since the u_m are continuous and ∂V is compact, we may assume by passing to a subsequence that u_m converges to 0 uniformly on ∂V. Indeed, we can ensure that

$$|u_m(x)| \le \frac{1}{m^2} \quad \text{for all} \quad x \in \partial V. \tag{348}$$

We now let G be an element of $C^1(\mathbb{R})$ such that

$$G(t) = \begin{cases} 0 & \text{if} \quad |t| \le 1 \\ t & \text{if} \quad |t| \ge 2 \end{cases} \tag{349}$$

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

$$v_m(x) = \frac{1}{m} G(m \cdot u_m(x)). \tag{350}$$

It is the case that $v_m \to u$ in $W^{1,k}(V)$. Since $v_m(x) = 0$ for all x such that $u_m(x) < 1/m$ and $|u_m(x)| < \frac{1}{m^2}$, each v_m has compact support bounded away from ∂V. It follows that we can mollify v_m in order to form a sequence of $C_c^\infty(V)$ functions which converge to u in $W^{1,p}(V)$. $\square$

Suppose that Ω is $C^{0,1}$ domain and that $u \in C^1\left(\overline{\Omega}\right)$. 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) ψ and ψ^{-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 ψ, we may assume that

$$\left|\frac{\partial \psi}{\partial x_n}(x)\right| = 1 \tag{351}$$

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

$$\frac{\partial}{\partial x_n} u \circ \psi \tag{352}$$

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

$$\frac{\partial u}{\partial \nu} \tag{353}$$

at the point y to be

$$\frac{\partial}{\partial x_n} u \circ \psi(x). \tag{354}$$

It is easy to verify that this definition is independent of the choice of U and ψ. 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 exists a function $u \in C^1\left(\overline{\Omega}\right)$ such that

$$u|_{\partial\Omega} = 0 \tag{355}$$

and

$$\frac{\partial u}{\partial \nu} \neq 0. \tag{356}$$

But, if $u \in W^{1,p}(\Omega)$, then (355) implies that $u \in W_0^{1,p}(\Omega)$ and hence is the limit of a sequence $\{\varphi_k\}$ of $C_c^\infty(\Omega)$ functions. But the normal derivative of each of the φ_k is 0, so we cannot have

$$\frac{\partial \phi_k}{\partial \nu} \to \frac{\partial u}{\partial \nu} \tag{357}$$

in any reasonable norm.

We close this section with the following generalization of the divergence theorem, which follows immediately from Theorem 65 and the classical divergence theorem.

THEOREM 67. *Suppose that Ω is a bounded Lipschitz domain in $\mathbb{R}^n$, that $u \in W^{1,p}(\Omega)$ for some $p \geq 1$, and that $\psi \in C^{0,1}\left(\overline{\Omega}\right)$. Then for all $i = 1, \ldots, n$,*

$$\int_\Omega D_i\psi(x)u(x)\; dx = -\int_\Omega \psi(x) D_i u(x)\; dx + \int_{\partial\Omega} \psi(x)\nu_i(x)\mathscr{T}u(x)\; dx, \tag{358}$$

where $\nu_i(x)$ is the i^{th} component of the outward-pointing unit normal vector at the point $x \in \partial\Omega$.

3.5. Extension Operators

Suppose that Ω is an open subset of $\mathbb{R}^n$, that $1 \leq p < \infty$ is a real number, and that $k \geq 0$ is an integer. We say that a linear mapping $E : W^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^n)$ is a simple (k,p)-extension operator if

(1) $E[u](x) = u(x)$ for almost all $x \in \Omega$;

(2) there exists a constant $C > 0$ such that $\|E[u]\|_{W^{k,p}(\mathbb{R}^n)} \leq C\|u\|_{W^{k,p}(\Omega)}$ for all $u \in W^{k,p}(\Omega)$.

A mapping E which takes functions defined almost everywhere in Ω to functions defined almost everywhere in Ω is a strong k-extension operator for Ω if for all $0 \leq m \leq k$ and $1 \leq p < \infty$, the restriction of E to $W^{m,p}(\Omega)$ is a simple (m,p)-extension operator. Finally, if E is a strong k-extension operator for all nonnegative integers k, then we call it a total extension operator for the domain Ω.

Exercise 25. *Suppose that $k \geq 0$ is an integer, and that $1 \leq p < \infty$ is a real number. Show that if E is a simple (k,p)-extension operator for Ω and that Ω' is an open set containing Ω, then there exists a bounded linear mapping $T : W^{k,p}(\Omega) \to W^{k,p}(\Omega')$ such that $T[u]|_{\Omega} = u$.*

Theorem 68. *Suppose that Ω is a Lipschitz domain (that is, a $C^{0,1}$ domain), and that $1 \leq p < \infty$. Then there exists a simple $(1,p)$-extension operator for Ω.*

Proof. Suppose that $u \in C^1(\{(x_1,\dots,x_n) : x_n \geq 0\})$ whose support in contained an open ball V centered at 0. It is easy to verify that the function ψ defined via

$$\psi(x_1,\dots,x_{n-1},x_n) = \begin{cases} u(x_1,\dots,x_{n-1},x_n) & x_n > 0 \\ -3u(x_1,\dots,x_{n-1},-x_n) + 4u(x_1,\dots,x_{n-1},-x_n/2) & x_n < 0 \end{cases} \tag{359}$$

is an element of $C^1(\mathbb{R}^n)$, and that there exists $C > 0$ which does not depend on u such that

$$\|\psi\|_{W^{1,p}(V)} \leq C\|u\|_{W^{1,p}(V\cap\{x_n>0\})}. \tag{360}$$

The general case now follows from Theorem 62 and a (by now) standard localization argument. Note that we are implicitly "straightening out" the boundary of $\partial\Omega$ here. This is why we cannot use this technique to obtain a (k,p) extension operator. □

Note that it follows easily from the proof of Theorem 68 that we can always assume that $E[u]$ has compact support. The argument of the preceding proof can be easily extended to yield the following theorem.

Theorem 69. *Suppose that Ω is a $C^{k-1,1}$ domain, that $1 \leq p < \infty$ is a real number, and that $k \geq 1$ is an integer. Then there exists a simple (k,p)-extension operator for Ω.*

However, the requirements placed on the boundary of Ω by Theorem 69 are much stronger than necessary. See Chapter 6 of [**16**] for a proof of the following much improved result.

THEOREM 70 (Stein). *Suppose that Ω is a Lipschitz domain in $\mathbb{R}^n$. Then there exists a total extension operator for Ω.*

Note too the article [**12**], which characterizes the domains which admit a simple (k,p)-extension operator for every $k \geq 0$ and every $1 \leq p \leq \infty$.

3.6. The Fourier Transform and Sobolev Spaces

Throughout this section we will use the convention

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} \exp(2\pi i x \cdot \xi) f(x)\, dx \tag{361}$$

for the Fourier transform and we will denote by $\|u\|_1$ the usual norm in $W^{k,2}(\mathbb{R}^n) = H^k(\mathbb{R}^n)$; that is,

$$\|u\|_1 = \sqrt{\sum_{|\alpha|\leq k} \|D^\alpha u\|^2_{L^2(\mathbb{R}^n)}}. \tag{362}$$

Since

$$\|u\|_{L^2(\mathbb{R}^n)} = \|\widehat{u}\|_{L^2(\mathbb{R}^n)} \tag{363}$$

and

$$\widehat{D^\alpha u}(\xi) = (2\pi i \xi)^\alpha \widehat{u}(\xi), \tag{364}$$

the $H^k(\mathbb{R}^n)$ norm of a function u can be bounded using the Fourier transform of u and vice-versa. More specifically:

THEOREM 71. *Suppose that $k \geq 0$ is an integer. Then*

$$\|u\|_2 = \sqrt{\int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \left(1 + |\xi|^2\right)^k \, d\xi} \tag{365}$$

is equivalent to the usual norm

$$\|u\|_1 = \sqrt{\sum_{|\alpha|\leq k} \|D^\alpha u\|^2_{L^2(\mathbb{R}^n)}}. \tag{366}$$

in $H^k(\mathbb{R}^n)$.

PROOF. Suppose that $u \in H^k(\mathbb{R}^n)$ so that $D^\alpha u \in L^2(\mathbb{R}^n)$ for all $|\alpha| \leq k$. We observe that

$$\widehat{D^\alpha u}(\xi) = (\xi)^\alpha \widehat{u}(\xi), \tag{367}$$

so that

$$\|D^\alpha u\|^2_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |(2\pi i \xi)^\alpha|^2 \, |u(\xi)|^2 \, d\xi = (2\pi)^{|\alpha|} \int_{\mathbb{R}^n} |\xi^\alpha|^2 \, |u(\xi)|^2 \, d\xi. \tag{368}$$

We observe that

$$|\xi^\alpha| \leq \begin{cases} |\xi|^k & \text{if } \ |\xi| \geq 1 \\ 1 & \text{if } |\xi| \leq 1. \end{cases} \tag{369}$$

when $|\alpha|\leq k$. Consequently, there exists $C_1 > 0$ such that

$$|\xi^\alpha|^2 \leq \max\{1, |\xi|^k\}^2 \leq (1 + |\xi|^k)^2 = 1 + 2|\xi|^k + |\xi|^{2k} \leq C_1(1 + |\xi|^{2k}) \tag{370}$$

for all $|\alpha|\leq k$. By summing (370) over α we see that there exists $C_2 > 0$ such that

$$\sum_{|\alpha|\leq k} |\xi^\alpha|^2 \leq C_2(1 + |\xi|^{2k}). \tag{371}$$

Consequently,

$$\begin{aligned}\sum_{|\alpha|\leq k} \|D^\alpha u\|^2_{L^2(\mathbb{R}^n)} &= \sum_{|\alpha|\leq k} \|\widehat{D^\alpha u}\|^2_{L^2(\mathbb{R}^n)} \\ &= \sum_{|\alpha|\leq k} (2\pi)^{|\alpha|} \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 |\xi^\alpha|^2 \, d\xi \\ &\leq (2\pi)^k \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \sum_{|\alpha|\leq k} |\xi^\alpha|^2 \, d\xi \\ &\leq C_2 (2\pi)^k \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \, (1 + |\xi|^2)^k \, d\xi.\end{aligned} \tag{372}$$

Now we define functions f_1 and f_2 via the formulas

$$f_1(\xi) = \sum_{j=1}^n |\xi_j^k|^2 \tag{373}$$

and

$$f_2(\xi) = |\xi|^{2k}. \tag{374}$$

We denote by M_1 the maximum of f_1 on the unit sphere in $\mathbb{R}^n$ and by M_2 the maximum of f_2 on the unit sphere in $\mathbb{R}^n$. Then

$$f_2(\xi) \leq M_2/M_1 f_1(\xi) \tag{375}$$

for all $|\xi|= 1$. Since both f_1 and f_2 are homogeneous of degree $2k$, the inequality (375) in fact holds for all $\xi \in \mathbb{R}^n$; in particular, if we set $C_3 = M_2/M_1$ then

$$|\xi|^{2k} \leq C_3 \sum_{j=1}^n |\xi_j^k|^2 \tag{376}$$

for all $\xi \in \mathbb{R}^n$. From the binomial theorem we have that

$$(1 + |\xi|^2)^k \leq 2^k \max\{1, |\xi|^{2k}\}, \tag{377}$$

and we combine (376), (377) to conclude that

$$\begin{aligned}(1 + |\xi|^2)^k &\leq 2^k \left(1 + |\xi|^{2k}\right) \\ &\leq 2^k \max\{1, C_3\} \left(1 + \sum_{j=1}^n |\xi_j^k|^2\right) \\ &\leq 2^k \max\{1, C_3\} \sum_{|\alpha|\leq k} |\xi^\alpha|^2\end{aligned} \tag{378}$$

for all $\xi \in \mathbb{R}^n$. Note that the last inequality follows since 1 and

$$\|\xi_j^k\|^2, \tag{379}$$

for each $j = 1, \ldots, n$, are terms in the sum

$$\sum_{|\alpha|\le k} |\xi^\alpha|^2. \tag{380}$$

It follows from (367) and (378) that there exists $C_4 > 0$ such that

$$\begin{aligned}
\int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 (1+|\xi|^2)^k \, d\xi &\le C_4 \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \left(\sum_{|\alpha|\le k} |\xi^\alpha|^2\right) d\xi \\
&= C_4 \sum_{|\alpha|\le k} \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \, |\xi^\alpha|^2 \, d\xi \\
&\le \frac{C_4}{(2\pi)^k} \sum_{|\alpha|\le k} \|\widehat{D^\alpha u}\|^2_{L^2(\mathbb{R}^n)}. \\
&= \frac{C_4}{(2\pi)^k} \sum_{|\alpha|\le k} \|D^\alpha u\|^2_{L^2(\mathbb{R}^n)}.
\end{aligned} \tag{381}$$

□

One important difference between the norms

$$\|u\|_1 = \sqrt{\sum_{|\alpha|\le k} \|D^\alpha u\|^2_{L^2(\mathbb{R}^n)}} \tag{382}$$

and

$$\|u\|_2 = \sqrt{\int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \left(1+|\xi|^2\right)^k \, d\xi} \tag{383}$$

is that the later generalizes to the case where k is no longer an integer; that is, it gives us a reasonable method for defining fractional derivatives and Sobolev spaces of fractional order. In particular, we define $H^s(\mathbb{R}^n)$ for real numbers $s \ge 0$ to be the space of functions $u \in L^2(\mathbb{R}^n)$ such that

$$\int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 (1+|\xi|^2)^s \, d\xi < \infty. \tag{384}$$

Clearly, $H^s(\mathbb{R}^n)$ is a Hilbert space with respect to the inner product

$$\langle u, v\rangle = \int_{\mathbb{R}^n} \widehat{u}(\xi)\overline{\widehat{v}(\xi)}(1+|\xi|^2)^s \, d\xi. \tag{385}$$

There is no real difficulty in allowing negative values of s. In this case, the elements of $H^s(\mathbb{R}^n)$ might no longer be functions, so we have to adapt our definition slightly. For $s < 0$, we let $H^s(\mathbb{R}^n)$ be the space of tempered distributions such that

$$\int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 (1+|\xi|^2)^s \, d\xi < \infty. \tag{386}$$

It is easy to verify that the space $H^{-s}(\mathbb{R}^n)$ is isomorphic to the dual space of $H^s(\mathbb{R}^n)$.

EXERCISE 26. *Prove that if $s > \frac{n}{2}$, then $H^s(\mathbb{R}^n)$ is a Banach algebra.*

The norm $\|u\|_1$ provides us with a definition of $H^k(\Omega)$ for Ω an open subset of $\mathbb{R}^n$, but only for nonnegative integer orders. On the other hand, $\|u\|_2$ gives a definition of $H^s(\mathbb{R}^n)$ for all real-valued s, but it is not readily applicable in the case of open subsets of $\mathbb{R}^n$. The following theorem will provide us with yet another equivalent norm, which we will use to define $H^s(\Omega)$ when s is a positive real number and Ω is an open subset of $\mathbb{R}^n$.

THEOREM 72. *Suppose that $0 < s < 1$ is a real number and $n > 0$ is an integer. Then there exists a constant $C > 0$ depending only on s and n such that*

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{2s+n}}\,dx\;dy = C\int_{\mathbb{R}^n} |\xi|^{2s}\,|\widehat{u}(\xi)|^2\;d\xi. \tag{387}$$

for all $u \in H^s(\mathbb{R}^n)$.

PROOF. We observe that the Fourier transform of the function

$$u(x+h)-u(x) \tag{388}$$

is

$$(\exp(2\pi i h\cdot\xi)-1)\,\widehat{u}(\xi). \tag{389}$$

From Planacherel's theorem we see that

$$\int_{\mathbb{R}^n} |u(x+h)-u(x)|^2\;dx = \int_{\mathbb{R}^n} |\exp(2\pi i h\cdot\xi)-1|^2\,|\widehat{u}(\xi)|^2\;d\xi. \tag{390}$$

We multiply both sides of (388) by $|h|^{-2s-n}$ and integrate with respect to h to obtain

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x+h)-u(x)|^2}{|h|^{2s+n}}\,dx\;dh = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|\exp(2\pi i h\cdot\xi)-1|^2}{|h|^{2s+n}}\,|\widehat{u}(\xi)|^2\;d\xi\;dh. \tag{391}$$

By changing the order of integration in the integral on the right-hand side of (391), we see that

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x+h)-u(x)|^2}{|h|^{2s+n}}\,dx\;dh = \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \left(\int_{\mathbb{R}^n} \frac{|\exp(2\pi i h\cdot\xi)-1|^2}{|h|^{2s+n}}\;dh\right) d\xi. \tag{392}$$

Using polar coordinates we see that

$$\begin{aligned}\int_{\mathbb{R}^n} \frac{|\exp(2\pi i h\cdot\xi)-1|^2}{|h|^{2s+n}}\;dh &= \int_0^\infty \frac{\rho^{n-1}}{\rho^{2s+n}} \int_{|s|=1} |\exp(2\pi i\rho s\cdot\xi)-1|^2\;ds d\rho \\ &= \int_0^\infty \rho^{-2s-1} \int_{|s|=1} |\exp(2\pi i\rho s\cdot\xi)-1|^2\;ds d\rho.\end{aligned} \tag{393}$$

Now we let $p = t|\xi|^{-1}$ in (393) to obtain

$$\int_{\mathbb{R}^n} \frac{|\exp(2\pi i h\cdot\xi)-1|^2}{|h|^{2s+n}}\;dh = |\xi|^{2s}\int_0^\infty t^{-2s-1}\int_{|s|=1} \left|\exp\left(2\pi i t s\cdot\frac{\xi}{|\xi|}\right)-1\right|^2\;ds\;dt. \tag{394}$$

We observe that

$$\int_{|s|=1} \left|\exp\left(2\pi i t s\cdot\frac{\xi}{|\xi|}\right)-1\right|^2\;dt \tag{395}$$

is $O(t^2)$ as $t \to 0$ and $O(1)$ as $t \to \infty$, so that the integral (394) converges. By radial symmetry its value does not depend on ξ. So there exists a constant $C > 0$ depending only on s and n such that

$$\int_{\mathbb{R}^n} \frac{|\exp(2\pi i h \cdot \xi) - 1|^2}{|h^{2s+n}|} \, dh = C|\xi|^{2s} \tag{396}$$

for all $\xi \in \mathbb{R}^n$. Inserting (396) into (391) yields

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x+h) - u(x)|^2}{|h|^{2s+n}} \, dx \, dh = C \int_{\mathbb{R}^n} |\xi|^{2s} |\widehat{u}(\xi)|^2 \, d\xi. \tag{397}$$

We let $h = y - x$ in (397) to conclude that

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x-y|^{2s+n}} \, dx \, dy = C \int_{\mathbb{R}^n} |\xi|^{2s} |\widehat{u}(\xi)|^2 \, d\xi. \tag{398}$$

□

If $s = k + \lambda$, where k is a nonnegative integer and $0 < \lambda < 1$, then we define another norm for $H^s(\mathbb{R}^n)$ by

$$\|u\|_3 = \sqrt{\sum_{|\alpha| \le k} \int_{\mathbb{R}^n} |D^\alpha u(x)|^2 \, dx + \sum_{|\alpha|=k} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{2\lambda+n}} \, dxdy}. \tag{399}$$

It is easy to verify via Theorem 72 that $\|u\|_3$ is equivalent to $\|u\|_1$ and $\|u\|_2$. For Ω an open subset of $\mathbb{R}^n$, we let $H^s(\Omega)$ be the set of functions $u \in H^k(\Omega)$ such that

$$\int_\Omega \int_\Omega \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{2\lambda+n}} \, dxdy < \infty \tag{400}$$

for all $|\alpha| = k$. This is a Banach space with respect to the norm

$$\|u\| = \sqrt{\sum_{|\alpha| \le k} \int_\Omega |D^\alpha u(x)|^2 \, dx + \sum_{|\alpha|=k} \int_\Omega \int_\Omega \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{2\lambda+n}} \, dxdy}. \tag{401}$$

As usual, we let $H_0^s(\Omega)$ denote the closure of $C_c^\infty(\Omega)$ with respect to the norm (401), and use $H^{-s}(\Omega)$ to denote the dual of $H_0^s(\Omega)$.

The norm (401) motivates the following definition of the fractional order space $W^{k,p}(\Omega)$ when Ω is an open subset of $\mathbb{R}^n$. Suppose that $s = k + \lambda$ with k a nonnegative integer and $0 < \lambda < 1$. We say that $u \in W^{k,p}(\Omega)$ is an element of the space provided $u \in W^{s,p}(\Omega)$ and

$$\int_\Omega \int_\Omega \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x-y|^{n+sp}} < \infty \tag{402}$$

for all $|\alpha| = k$. When endowed with the norm

$$\|u\| = \left(\sum_{|\alpha| \le k} \int_\Omega |D^\alpha u(x)|^p \, dx + \sum_{|\alpha|=k} \int_\Omega \int_\Omega \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x-y|^{n+\lambda p}} \, dxdy \right)^{1/p}, \tag{403}$$

$W^{s,p}(\Omega)$ becomes a Banach space. The closure of $C_c^\infty(\Omega)$ with respect to the norm (403) is typically denoted by $W_0^{s,p}(\Omega)$.

There are $m$any other methods for defining the fractional order Sobolev spaces $W^{k,p}(\Omega)$. For the most part, they coincide when the domain Ω is sufficiently regular. For instance,

$$\|u\| = \inf\left\{\|f\|_{H^s(\mathbb{R}^n)} : f \in H^s(\mathbb{R}^n) \text{ such that } f|_\Omega = u\right\}. \tag{404}$$

is equivalent to the norms for $H^s(\Omega)$ we have defined when Ω is Lipschitz. We note that this fact is strongly related to results showing existence of extension operators for fractional order Sobolev spaces. See also [**1**] and [**9**] for definitions of the Besov and Triebel-Lizorkin spaces which generalize the fractional order spaces we consider here.

3.7. Fractional Order Sobolev Spaces on the Boundary of a Domain

We have not yet introduced definitions of $H^s(\partial\Omega)$ or $W^{s,p}(\partial\Omega)$ when s is a real number and $\Omega \subset \mathbb{R}^n$ is a bounded open set in $\mathbb{R}^n$. In fact, we have not even treated the case when s is a positive integer. We now do so.

Our definition will require fairly strong regularity assumptions on the boundary of Ω. We suppose that $s = k + \lambda$ with k a positive integer and $0 < \lambda < 1$, and that Ω is a $C^{k-1,1}$ domain. Then there exist a covering

$$\partial\Omega \subset \bigcup_{j=1}^{m} U_j \tag{405}$$

of $\partial\Omega$ by open sets, a collection of open sets $V_1, \ldots, V_m$ in $\mathbb{R}^n$, and a collection of mappings $\psi_1, \ldots, \psi_m$ such that

(1) ψ_j is a bijective mapping $V_j \to U_j$;

(2) ψ_j and ψ_j^{-1} are $C^{k-1,1}$ mappings;

(3) $\psi_j(V_j \cap \{x_n < 0\}) = U_j \cap \Omega$;

(4) $\psi_j(V_j \cap \{x_n = 0\}) = U_j \cap \partial\Omega$.

Now we let $\eta_1, \ldots, \eta_m$ be a partition of unity subordinate to the covering (405). For each $j = 1, \ldots, m$, we let u_j be the restriction of $(\eta_j u) \circ \psi_j$ to $V_j \cap \{x_n = 0\}$. We can view u_j as a compactly supported function defined on $\mathbb{R}^{n-1}$. We let $W^{s,p}(\partial\Omega)$ consist of all functions u in $L^1_{\text{loc}}(\partial\Omega)$ such that

$$\|u\|_{W^{s,p}(\partial\Omega)} = \left(\sum_{j=1}^{m} \|u_j\|^p_{W^{s,p}(\mathbb{R}^{n-1})}\right)^{1/p} < \infty. \tag{406}$$

It is easy to verify that $W^{s,p}(\partial\Omega)$ is a Banach space with respect to the norm (406). A straightforward but tedious argument shows that alternate choices of the sets $U_1, \ldots, U_m$, $V_1, \ldots, V_m$, mappings $\psi_1, \ldots, \psi_m$ and partition of unity $\eta_1, \ldots, \eta_m$ lead to equivalent norms. We use $H^s(\partial\Omega)$ to denote $W^{s,2}(\partial\Omega)$ and we define $H^{-s}(\partial\Omega)$ to be the dual space of $H^s(\partial\Omega)$. Of course, $H^s(\Omega)$

When $0 < s < 1$, we can avoid the use of a partition of unity and local parameterizations of the boundary. Indeed,

$$\|u\|_{W^{s,p}(\partial\Omega)}= \left(\int_{\partial\Omega} |u(x)|^p \ dx + \int_{\partial\Omega}\int_{\partial\Omega} \frac{|u(x)-u(y)|^p}{|x-y|^{sp+n-1}} \ dx \ dy\right)^{1/p} \tag{407}$$

is equivalent to (406) when $0 < s < 1$. Note that $sp + n - 1$ is correct exponent in the denominator because $\partial\Omega$ is $n-1$ dimensional rather than n dimensional.

We will now show that the trace space of $H^k(\Omega)$ is $H^{k-1/2}(\partial\Omega)$ when Ω is a $C^{k-1,1}$ domain. We suppose that $U_1, \ldots, U_m$, $V_1, \ldots, V_m$, and $\psi_1, \ldots, \psi_m$ are as before, but now we asumme that U_0 an open set compactly contained in Ω such that

$$\Omega \subset \bigcup_{j=0}^{m} U_j, \tag{408}$$

and let $\eta_0, \eta_1, \ldots, \eta_m$ be a smooth partition of unity subordinate to the cover (408). We also let E be a total extension operator for Ω such that $E[u]$ is always compactly supported. For each $j = 1, \ldots, m$, we let u_j be the restriction of $(\eta_j E[u]) \circ \psi_j$ to $V_j \subset \mathbb{R}^n$ and v_j the restriction of u_j to $V_j \cap \{x_n = 0\} \subset \mathbb{R}^{n-1}$. Then

$$\left(\|\eta_0 u\|^p_{H^k(\mathbb{R}^n)} + \sum_{j=1}^{m} \|u_j\|^p_{H^k(\mathbb{R}^n)}\right)^{1/p} \tag{409}$$

is equivalent to the usual $H^k(\Omega)$ norm and

$$\left(\sum_{j=1}^{m} \|v_j\|^p_{H^{k-1/2}(\mathbb{R}^{n-1})}\right)^{1/p} \tag{410}$$

is, by definition, a norm for $H^{k-1/2}(\partial\Omega)$. Note that we use the extension operator because the function $(\eta_j u) \circ \psi_j$ is only defined for points inside $V_j \cap \{x_n < 0\}$ and its zero extension may not be an element of $H^k(\mathbb{R}^n)$, whereas u_j compactly supported in V_j and so its zero extension is an element of $H^k(\mathbb{R}^n)$. It follows that to study the trace of functions in $H^k(\Omega)$, it suffices to consider the restriction of compactly supported functions in $\mathbb{R}^n$ to the set $\{(x_1, \ldots, x_{n-1}, x_n) : x_n = 0\}$.

THEOREM 73. *Suppose that* $\frac{1}{2} < s \leq \infty$ *and that* $n \geq 1$ *is an integer. Then the operator*

$$R : C_c^\infty(\mathbb{R}^n) \to C_c^\infty\left(\mathbb{R}^{n-1}\right) \tag{411}$$

defined via

$$R[f](x_1, \ldots, x_{n-1}) = f(x_1, \ldots, x_{n-1}, 0) \tag{412}$$

admits an extension to a continuous linear mapping

$$H^s(\mathbb{R}^n) \to H^{s-1/2}\left(\mathbb{R}^{n-1}\right). \tag{413}$$

Moreover, the extended operator has a continuous right inverse.

PROOF. We observe that

$$\int_{\mathbb{R}^{n-1}} \exp(2\pi i\eta\cdot y)\widehat{R[f]}(\eta) \ d\eta = R[f](y) = f(y,0) = \int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}} \exp(2\pi i\eta\cdot y)\widehat{f}(\eta,\zeta) \ d\eta d\zeta \tag{414}$$

for all $y \in \mathbb{R}^{n-1}$. In particular,

$$R\left[f\right](y) = \int_{\mathbb{R}^{n-1}} \exp(2\pi i\eta\cdot y)\left(\int_{\mathbb{R}} \widehat{f}(\eta,\zeta)\ d\zeta\right) d\eta \tag{415}$$

for all $y \in \mathbb{R}^{n-1}$, from which we see that

$$\widehat{R\left[f\right]}(\eta) = \int_{\mathbb{R}} \widehat{f}(\eta,\zeta)\ d\zeta. \tag{416}$$

It follows from (416) and the Cauchy-Schwartz inequality that

$$\begin{aligned}\left|\widehat{R\left[f\right]}(\eta)\right|^2 &= \left|\int_{\mathbb{R}} \widehat{f}(\eta,\zeta)\left(1+|\eta|^2+|\zeta|^2\right)^{s/2}\left(1+|\eta|^2+|\zeta|^2\right)^{-s/2}\ d\zeta\right|^2 \\ &\leq \left(\int_{\mathbb{R}}\left|\widehat{f}(\eta,\zeta)\right|^2\left(1+|\eta|^2+|\zeta|^2\right)^{s}\ d\zeta\right)\left(\int_{\mathbb{R}}\left(1+|\eta|^2+|\zeta|^2\right)^{-s}\ d\zeta\right).\end{aligned} \tag{417}$$

We rewrite the second factor on the right in (417) as

$$2\int_0^\infty \left(\alpha^2+r^2\right)^{-s}\ dr, \tag{418}$$

where $\alpha^2 = 1+|\eta|^2$, and introduce the new variable $u = r/\alpha$ to obtain the integral

$$2\alpha^{1-2s}\int_0^\infty \left(1+u^2\right)^{-s}\ du. \tag{419}$$

We note that this integral is convergent since $s > \frac{1}{2}$ and set

$$C_s = 2\int_0^\infty \left(1+u^2\right)^{-s}\ du, \tag{420}$$

so that (417) may be rewritten as

$$\left|\widehat{R\left[f\right]}(\eta)\right|^2\left(1+|\eta|^2\right)^{s-1/2} \leq C_s\int_{\mathbb{R}}\left|\widehat{f}(\eta,\zeta)\right|^2\left(1+|\eta|^2+|\zeta|^2\right)\ d\zeta. \tag{421}$$

(Note that it is α^2 and not α which is equal to $1+|\eta|^2$). Integrating both sides of (421) with respect to η gives

$$\|R\left[f\right]\|^2_{H^{s-1/2}} \leq C_s\int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}}\left|\widehat{f}(\eta,\zeta)\right|^2\left(1+|\eta|^2+|\zeta|^2\right)^s\ d\zeta d\eta = C_s\|f\|^2_{H^s}. \tag{422}$$

It remains to show that the trace operator has a continous right inverse. To that end, we let φ be a $C_c^\infty\left(\mathbb{R}\right)$ function such that

$$\varphi(y) = 1 \ \text{ for all } \ |y|\leq 1. \tag{423}$$

Given $u \in C_c^\infty\left(\mathbb{R}^{n-1}\right)$, we define $U:\mathbb{R}^n\to\mathbb{R}$ via

$$U(x_1,\ldots,x_{n-1},x_n) = \int_{\mathbb{R}^{n-1}} \widehat{u}(\xi)\varphi\left((1+|\xi|^2)^{1/2}x_n\right)\exp\left(2\pi i\xi\cdot(x_1,\ldots,x_{n-1})\right)\ d\xi. \tag{424}$$

Obviously,

$$U(x_1,\ldots,x_{n-1},0) = \int_{\mathbb{R}^{n-1}} \widehat{u}(\xi)\exp\left(2\pi i\xi\cdot(x_1,\ldots,x_{n-1})\right)\ d\xi = u(x_1,\ldots,x_{n-1}). \tag{425}$$

Moreover,

$$\begin{aligned}\widehat{U}(\xi,\eta) &= \int_{\mathbb{R}}\int_{\mathbb{R}^{n-1}} \exp(-2\pi i\xi\cdot x')\exp(-2\pi i\eta\cdot x_n)U(x',x_n)\ dx'dx_n \\ &= \widehat{u}(\xi)\int_{\mathbb{R}} \exp(-2\pi i x_n\eta)\varphi\left((1+|\xi|^2)^{1/2}x_n\right)\ dx_n \\ &= \widehat{u}(\xi)\frac{\widehat{\varphi}\left((1+|\xi|^2)^{-1/2}\eta\right)}{\left(1+|\xi|^2\right)^{1/2}}.\end{aligned} \tag{426}$$

The last equality in (426) is obtained by introducing the new variable $u=(1+|\xi|^2)^{1/2}x_n$. Now (426) implies

$$\begin{aligned}&\int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}}\left|\widehat{U}(\xi,\eta)\right|^2\left(1+|\xi|^2+|\eta|^2\right)^s\ d\xi d\eta \\ &= \int_{\mathbb{R}^{n-1}}\frac{|\widehat{u}(\xi)|^2}{\left(1+|\xi|^2\right)}\left(\int_{\mathbb{R}}\left(1+|\xi|^2+|\eta|^2\right)^s\left|\widehat{\varphi}\left(1+|\xi|^2)^{-1/2}\eta\right)\right|^2\ d\eta\right)d\xi.\end{aligned} \tag{427}$$

We rewrite the integral with respect to η appearing on the right-hand side of (427) as

$$\int_{\mathbb{R}}\left(\alpha^2+|\eta|^2\right)^s\left|\widehat{\varphi}\left(\frac{\eta}{\alpha}\right)\right|^2\ d\eta \tag{428}$$

where $\alpha^2=1+|\xi|^2$. By introducing the new variable $u=\eta/\alpha$ we see that (428) is equal to

$$\alpha^{2s+1}\int_{\mathbb{R}}\left(1+u^2\right)^s\left|\widehat{\varphi}\left(u\right)\right|^2\ du. \tag{429}$$

We conclude that

$$\int_{\mathbb{R}}\left(1+|\xi|^2+|\eta|^2\right)^s\left|\widehat{\varphi}\left(1+|\xi|^2)^{-1/2}\eta\right)\right|^2\ d\eta = C_s\left(1+|\xi|^2\right)^{s+1/2}, \tag{430}$$

where

$$C_s=\int_{\mathbb{R}}\left(1+u^2\right)^s\left|\widehat{\varphi}\left(u\right)\right|^2\ du. \tag{431}$$

The integral defining C_s is convergent since $\varphi\in C_c^\infty(\mathbb{R})$. By inserting (430) into (427), we obtain

$$\int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}}\left|\widehat{U}(\xi,\eta)\right|^2\left(1+|\xi|^2+|\eta|^2\right)^s\ d\xi d\eta = C_s\int_{\mathbb{R}^{n-1}}|\widehat{u}(\xi)|^2\left(1+|\xi|^2\right)^{s-1/2}d\xi; \tag{432}$$

that is, $\|U\|_{H^s(\mathbb{R}^n)}\leq C_s\|u\|_{H^{s-1/2}(\mathbb{R}^{n-1})}$. The mapping which takes u to U extends by continuity to the desired continuous right inverse of the trace operator. □

EXERCISE 27. *Show that the restriction operator R does not extend to a bounded continuous mapping $L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^{n-1})$.*

EXERCISE 28. *Show that the restriction operator R does not extend to a bounded continuous mapping $H^{1/2}(\mathbb{R}^n)\to L^2(\mathbb{R}^{n-1})$.*

(Hint: Let $u : \mathbb{R} \to \mathbb{R}$ be the inverse Fourier transform of the function $f : \mathbb{R} \to \mathbb{R}$ defined by the formula

$$f(\xi) = \begin{cases} \frac{1}{\xi \log(\xi)} & \xi > 2 \\ 0 & \xi \leq 2. \end{cases} \tag{433}$$

Note that u and f are elements of $L^2(\mathbb{R})$. For $h > 0$, define u_h via $u_h(x) = \eta_h * u(x)$, where η_h is the standard mollifier. Moreover, fix an aritrary $\psi \in C_c^\infty(\mathbb{R})$ and let $v : \mathbb{R}^2 \to \mathbb{R}$ be defined by $v(x, y) = \psi(x)u(y)$ and, for each $h > 0$, define $v_h : \mathbb{R}^2 \to \mathbb{R}$ via $v_h(x, y) = \psi(x)u_h(y)$. Show that v is an element of $H^{1/2}(\mathbb{R}^2)$, that v_h converges to v in $H^{1/2}(\mathbb{R}^2)$, but that $\|R[v_h]\|_{L^2(\mathbb{R})} \to \infty$ as $h \to 0$.)

EXERCISE 29. *Show that for any integer $l \geq 0$ there exists a mapping $T_l : C_c^\infty(\mathbb{R}^{n-1}) \to C_c^\infty(\mathbb{R}^n)$ such that*

$$\frac{\partial^k T_l[f]}{\partial x_n^k}(x_1, \dots, x_{n-1}, 0) = 0 \tag{434}$$

for all $k = 0, 1 \dots, l-1$ and $(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$, and

$$\frac{\partial^l T_l[f]}{\partial x_n^l}(x_1, \dots, x_{n-1}, 0) = f(x_1, \dots, x_{n-1}) \tag{435}$$

for all $(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$. Conclude that for any integer $l \geq 0$, there exists an operator

$$R_l : \prod_{k=0}^{l} C_c^\infty\left(\mathbb{R}^{n-1}\right) \to C_c^\infty(\mathbb{R}^n) \tag{436}$$

such that if

$$f = R_l(g_1, \dots, g_l), \tag{437}$$

then

$$\frac{\partial^k f}{\partial x_n^k}(x_1, \dots, x_{n-1}, 0) = g_k(x_1, \dots, x_{n-1}) \tag{438}$$

for all $k = 0, 1, \dots, l$ and $(x_1, \dots, x_{n-1})$.

Hint: Proceed as in Theorem 73 but replace φ with the function $\psi(y) = y^l/l!$ for all $|y| \leq 1$ in order to define T_l. Then define R_l via

$$R_l[f] = T_0[g_0] + T_1\left[g_1 - \frac{\partial}{\partial x_n}T_0[g_0]\right] + T_2\left[g_2 - \frac{\partial^2}{\partial x_n^2}T_0[g_0] - \frac{\partial^2}{\partial x_n^2}T_1\left[g_1 - \frac{\partial}{\partial x_n}T_0[g_0]\right]\right] + \cdots. \tag{439}$$

The following theorem is an immediate consequence of Theorem 73 and the discussion preceding it.

THEOREM 74. *Suppose that k is a positive integer, Ω is a $C^{k-1,1}$ domain in $\mathbb{R}^n$, and that $1 \leq p < \infty$. Then the trace operator*

$$\mathscr{T} : C^\infty\left(\overline{\Omega}\right) \to C(\partial\Omega) \tag{440}$$

defined via

$$\mathscr{T}[u] = u|_{\partial\Omega} \tag{441}$$

admits an extension to a continuous linear mapping

$$H^k(\Omega) \to H^{k-1/2}(\partial\Omega). \tag{442}$$

Moreover, the extended operator has a continuous right inverse.

In fact, using Exercise 29 we can easily establish the following:

THEOREM 75. *Suppose that k is a positive integer, and that Ω is a $C^{k-1,1}$ domain in $\mathbb{R}^n$. Then the trace operator*

$$\mathscr{T} : C^\infty\left(\overline{\Omega}\right) \to \prod_{j=0}^{k-1} C(\partial\Omega) \tag{443}$$

defined via

$$\mathscr{T}[u] = \left(u, \frac{\partial u}{\partial \nu}, \frac{\partial^2 u}{\partial \nu^2}, \dots, \frac{\partial^{k-1} u}{\partial \nu^{k-1}},\right) \tag{444}$$

admits an extension to a continuous linear mapping

$$H^k(\Omega) \to \prod_{j=0}^{k-1} H^{k-j-1/2}(\partial\Omega). \tag{445}$$

Moreover, the extended operator has a continuous right inverse.

When Ω is Lipschitz, the preceding results do not allow us to characterize the trace of $H^k(\Omega)$ for $k > 1$. However, we have the following theorem characterizing the trace of $H^2(\Omega)$ when Ω is Lipschitz. A proof can be found in [**14, 11**]; see also [**3**].

THEOREM 76. *Suppose that Ω is Lipschitz domain in $\mathbb{R}^n$, and that $\mathscr{T}$ is the mapping defined for $u \in C^\infty\left(\overline{\Omega}\right)$ via*

$$u \to \left(u|_{\partial\Omega}, \left.\frac{\partial u}{\partial x_1}\right|_{\partial\Omega} \left.\frac{\partial u}{\partial x_2}\right|_{\partial\Omega}, \dots, \left.\frac{\partial u}{\partial x_n}\right|_{\partial\Omega}\right). \tag{446}$$

Suppose also that $V^{3/2}(\partial\Omega)$ is the closure of the image of $\mathscr{T}$ with respect to the norm in the space

$$H^1(\partial\Omega) \oplus H^{1/2}(\partial\Omega) \oplus H^{1/2}(\partial\Omega) \oplus \dots \oplus H^{1/2}(\partial\Omega). \tag{447}$$

Then $\mathscr{T}$ extends a bounded linear mapping $H^2(\Omega) \to V^{3/2}(\partial\Omega)$, and that mapping admits a continuous right inverse.

See [**10**] and its references for characterizations of the trace spaces of $W^{k,p}(\Omega)$ in the general case.

3.8. Sobolev Inequalities and Embeddings

We will now establish two inequalities which will be used shortly to show that Sobolev spaces can be embedded in various Lebesgue and Hölder spaces. See Chapter 4 of [**1**] for a much more thorough discussion of this topic.

THEOREM 77 (Gagliardo-Nirenberg-Sobolev). *Suppose that* $1 \leq p < n$, *and that*

$$p^* = \frac{np}{n-p}. \tag{448}$$

Then there exists $C > 0$ *such that*

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C\|Du\|_{L^p(\mathbb{R}^n)} \tag{449}$$

for all $u \in C_c^1(\mathbb{R}^n)$.

PROOF. For each $1 \leq i \leq n$,

$$u(x_1, \ldots, x_i, \ldots, x_n) = \int_{-\infty}^{x_i} \frac{\partial u}{\partial x_i}(x_1, \ldots, t_i, \ldots x_n)\ dt_i, \tag{450}$$

and so

$$|u(x_1, \ldots, x_i, \ldots, x_n)| \leq \int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)|\ dt_i. \tag{451}$$

It follows that

$$|u(x)|^n \leq \prod_{i=1}^{n} \int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)|\ dt_i. \tag{452}$$

We take the $1/(n-1)^{th}$ power of each side of (452) to obtain

$$|u(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)|\ dt_i \right)^{\frac{1}{n-1}}. \tag{453}$$

We integrate (453) with respect to x_1 to obtain

$$\begin{aligned} \int |u(x)|^{\frac{n}{n-1}}\, dx_1 &\leq \int_{-\infty}^{\infty} \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)|\ dt_i \right)^{\frac{1}{n-1}}\ dx_1 \\ &= \left(\int_{-\infty}^{\infty} |Du(t_1, x_2, \ldots, x_n)|\ dt_1 \right)^{\frac{1}{n-1}} \\ &\cdot \int_{-\infty}^{\infty} \prod_{i=2}^{n} \left(\int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)|\ dt_i \right)^{\frac{1}{n-1}}\ dx_1. \end{aligned} \tag{454}$$

Repeated application of Hölder's inequality shows that

$$\int |f_1(x) f_2(x) \cdots f_n(x)| dx \leq \prod_{j=1}^{n} \|f_j\|_{p_j} \tag{455}$$

whenever $1 \leq p_1, p_2, \ldots, p_n \leq \infty$ satisfy

$$\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_n} = 1. \tag{456}$$

We call (455) the generalized Hölder inequality and we invoke it with $p_1 = p_2 = \ldots = p_{n-1} = n-1$ to obtain

$$\begin{aligned} &\int_{-\infty}^{\infty} \prod_{i=2}^{n} \left(\int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)| \ dt_i \right)^{\frac{1}{n-1}} \ dx_1 \\ &\leq \left(\prod_{i=2}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)| \ dx_1 \ dt_i \right)^{\frac{1}{n-1}} . \end{aligned} \tag{457}$$

Inserting (457) into (452) yields

$$\begin{aligned} \int |u(x)|^{\frac{n}{n-1}} \, dx_1 \leq &\left(\int_{-\infty}^{\infty} |Du(t_1, x_2, \ldots, x_n)| \ dt_1 \right)^{\frac{1}{n-1}} \\ &\cdot \left(\prod_{i=2}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)| \ dx_1 \ dt_i \right)^{\frac{1}{n-1}} . \end{aligned} \tag{458}$$

We integrate both sides of (458) with respect to x_2 to obtain

$$\begin{aligned} &\int \int |u(x)|^{\frac{n}{n-1}} \, dx_1 dx_2 \leq \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du(x_1, \ldots, t_2, \ldots x_n)| \ dx_1 dt_2 \right)^{\frac{1}{n-1}} \\ &\cdot \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |Du(t_1, x_2, \ldots, x_n)| \ dt_1 \cdot \prod_{i=3}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)| \ dx_1 \ dt_i \right)^{\frac{1}{n-1}} \ dx_2 \end{aligned} \tag{459}$$

We apply the generalized Hölder inequality once again to obtain

$$\begin{aligned} &\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |Du(t_1, x_2, \ldots, x_n)| \ dt_1 \cdot \prod_{i=3}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)| \ dx_1 \ dt_i \right)^{\frac{1}{n-1}} \ dx_2 \\ &\leq \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du(t_1, x_2, \ldots, x_n)| \ dt_1 dx_2 \right)^{\frac{1}{n-1}} \cdot \\ &\prod_{i=3}^{n} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)| \ dx_1 \ dx_2 \ dt_i \ \right)^{\frac{1}{n-1}} \end{aligned} \tag{460}$$

Inserting (460) into (459) yields

$$\begin{aligned} &\int \int |u(x)|^{\frac{n}{n-1}} \, dx_1 dx_2 \\ &\leq \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du(x_1, t_2, x_3, \ldots x_n)| \ dx_1 dt_2 \right)^{\frac{1}{n-1}} \cdot \left(\int_{-\infty}^{\infty} |Du(t_1, x_2, x_3, \ldots x_n)| \ dt_1 dx_2 \right)^{\frac{1}{n-1}} \cdot \\ &\prod_{i=3}^{n} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du(x_1, \ldots, t_i, \ldots x_n)| \ dx_1 \ dx_2 \ dt_i \ \right)^{\frac{1}{n-1}} . \end{aligned} \tag{461}$$

By repeatedly applying this procedure we see that

$$\begin{aligned}\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}}\,dx &\le \prod_{i=1}^{n}\left(\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} Du(x_1,\ldots,t_i,\ldots,x_n)dx_1\cdots dt_i\cdots dx_n\right)^{\frac{1}{n-1}} \\ &= \left(\int_{\mathbb{R}^n} |Du(x)|\ dx\right)^{\frac{n}{n-1}},\end{aligned} \tag{462}$$

which is the inequality in the case $p=1$.

We now suppose that $1<p<n$. We apply (462) to $v=|u|^{\gamma}$, where

$$\gamma = \frac{p(n-1)}{n-p}, \tag{463}$$

to obtain

$$\left(\int_{\mathbb{R}^n} |u(x)|^{\frac{\gamma n}{n-1}}\ dx\right)^{\frac{n-1}{n}} \le \gamma\int_{\mathbb{R}^n} |u(x)|^{\gamma-1}|Du|\ dx. \tag{464}$$

We apply Hölder inequality to the integral on the right-hand side of (462) to obtain

$$\left(\int_{\mathbb{R}^n} |u(x)|^{\frac{\gamma n}{n-1}}\ dx\right)^{\frac{n-1}{n}} \le \gamma\left(\int_{\mathbb{R}^n} |u(x)|^{(\gamma-1)\frac{p}{p-1}}\ dx\right)^{\frac{p-1}{p}}\left(\int_{\mathbb{R}^n} |Du(x)|^{p}\ dx\right)^{\frac{1}{p}}. \tag{465}$$

Now we observe that

$$(\gamma-1)\frac{p}{p-1} = \frac{\gamma n}{n-1} = \frac{np}{n-p} = p^*. \tag{466}$$

In light of (466), (465) is equivalent to

$$\left(\int_{\mathbb{R}^n} |u(x)|^{p^*}\ dx\right)^{\frac{1}{p^*}} \le \gamma\left(\int_{\mathbb{R}^n} |Du(x)|^{p}\ dx\right)^{\frac{1}{p}}, \tag{467}$$

which establishes the theorem in the case when $u\in C_c^1(\mathbb{R}^n)$.

□

We call the number p^* appearing in Theorem 77 the Sobolev conjugate of p. We note that

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}. \tag{468}$$

THEOREM 78 (Morrey's inequality). *Suppose that $n<p\le\infty$, and that $\gamma = 1-n/p$. Then there exists a constant C such that*

$$\|u\|_{C^{0,\gamma}\left(\overline{\mathbb{R}^n}\right)} \le C\|u\|_{W^{1,p}(\mathbb{R}^n)} \tag{469}$$

for all $u\in C^1(\mathbb{R}^n)$.

PROOF. First, we will show that there exists a constant $C>0$ depending only on n such that for all $r>0$,

$$\int_{B_r(x)} |u(y)-u(x)|\ dy \le Cr^n\int_{B_r(x)} \frac{|Du(y)|}{|x-y|^{n-1}}\ dy. \tag{470}$$

To see this, we fix w such that $|w|= 1$. Then if $0 < s < r$,

$$\begin{aligned} |u(x+sw)-u(x)| &= \left|\int_0^s \frac{d}{dt}u(x+tw)\ dt\right| \\ &= \left|\int_0^s Du(x+tw)\cdot w\ dt\right| \\ &\leq \int_0^s |Du(x+tw)|\,dt. \end{aligned} \tag{471}$$

It follows that

$$\begin{aligned} \int_{\partial B_1(0)} |u(x+sw)-u(x)|\ dS(x) &\leq \int_0^s \int_{\partial B_1(0)} |Du(x+tw)|\ dS(x)dt \\ &= \int_0^s \int_{\partial B_1(0)} |Du(x+tw)|\,\frac{t^{n-1}}{t^{n-1}}\ dS(x)dt \end{aligned} \tag{472}$$

We let $y = x + tw$ (so that $t = |tw|= |x - y|$) in the preceding integral to obtain

$$\begin{aligned} \int_{\partial B_1(0)} |u(x+sw)-u(x)|\ dS(x) &\leq \int_{B_s(x)} \frac{|Du(y)|}{|x-y|^{n-1}}\ dy \\ &\leq \int_{B_r(x)} \frac{|Du(y)|}{|x-y|^{n-1}}\ dy. \end{aligned} \tag{473}$$

Multiplying both sides of this equation by s^{n-1} and integrating from 0 to r with respect to s yields

$$\int_{\partial B_r(x)} |u(y)-u(x)|\ dS(x) \leq \frac{r^n}{n}\int_{B_r(x)} \frac{Du(y)}{|x-y|^{n-1}}\ dy; \tag{474}$$

this is (470).

We now fix $x \in \mathbb{R}^n$ and observe that (470) implies

$$\begin{aligned} |u(x)| &= \frac{1}{|B_1(x)|}\int_{B_1(x)} |u(x)|\ dy \\ &\leq \frac{1}{|B_1(x)|}\int_{B_1(x)} |u(x)-u(y)|\ dy + \frac{1}{|B_1(x)|}\int_{B_1(x)} |u(y)|\ dy \\ &\leq C\int_{B_1(x)} \frac{|Du(y)|}{|x-y|^{n-1}}\ dy + C\|u\|_{L^p(B_1(x))}. \end{aligned} \tag{475}$$

We now apply Hölder's inequality to obtain

$$\int_{B_1(x)} \frac{|Du(y)|}{|x-y|^{n-1}}\ dy \leq \left(\int_{B_1(x)} |Du(y)|^p\ dy\right)^{1/p} \left(\int_{B_1(x)} \frac{dy}{|x-y|^{(n-1)\frac{p-1}{p}}}\ dy\right)^{\frac{p-1}{p}}. \tag{476}$$

Since $p > n$, $(n-1)\frac{p}{p-1} < n$ so

$$\left(\int_{B_1(x)} \frac{dy}{|x-y|^{(n-1)\frac{p-1}{p}}}\ dy\right)^{\frac{p-1}{p}} < \infty. \tag{477}$$

We combine (475), (476) and (477) to obtain

$$\sup_x |u(x)| \leq C\|u\|_{W^{1,p}(\mathbb{R}^n)}. \tag{478}$$

We let $x, y \in \mathbb{R}^n$, let $r = |x-y|$, and let $W = B_r(x) \cap B_r(y)$. Then

$$|u(x) - u(y)| \leq \frac{1}{|W|}\int_W |u(x) - u(z)| \; dz + \frac{1}{|W|}\int_W |u(y) - u(z)| \; dz. \tag{479}$$

From (470), we have

$$\begin{aligned} \frac{1}{|W|}\int_W |u(x) - u(z)| \; dz &\leq C\frac{1}{|B_r(x)|}\int_{B_r(x)} |u(x) - u(z)| \; dz \\ &\leq \left(\int_{B_r(x)} |Du(z)|^p \; dz\right)^{1/p}\left(\int_{B_r(x)} \frac{dy}{|x-z|^{(n-1)\frac{p-1}{p}}} \; dz\right)^{\frac{p-1}{p}} \\ &\leq Cr^{1-\frac{n}{p}}\|Du\|_{L^p(\mathbb{R}^n)}. \end{aligned} \tag{480}$$

Similarly,

$$\frac{1}{|W|}\int_W |u(y) - u(z)| \; dz \leq Cr^{1-\frac{n}{p}}\|Du\|_{L^p(\mathbb{R}^n)}. \tag{481}$$

Inserting (481) and (480) into (479) yields

$$\frac{1}{|W|}\int_W |u(x) - u(z)| \; dz \leq Cr^{1-\frac{n}{p}}\|Du\|_{L^p(\mathbb{R}^n)} = C|x-y|^{1-\frac{n}{p}}\|Du\|_{L^p(\mathbb{R}^n)}. \tag{482}$$

It follows that

$$\sup_{x \neq y} \frac{|u(x) - u(y)|}{|x-y|^{1-\frac{n}{p}}} \leq C\|Du\|_{L^p(\mathbb{R}^n)}. \tag{483}$$

The result follows from (483) together with (478). □

It follows immediately from Theorems 77 and 78 that when $1 \leq p < n$, $W^{1,p}(\mathbb{R}^n)$ is continuously embedded in $L^{p^*}(\mathbb{R}^n)$ and when $p > n$, $W^{1,p}(\mathbb{R}^n)$ is continuously embedded in $C^{0,\gamma}\left(\overline{\mathbb{R}^n}\right)$ with $\gamma = 1 - n/p$.

Using standard interpolation results for Lebesgue spaces (see, for instance, Proposition 6.10 in Chapter 6 of [**8**]), we can say a bit more. In fact, when $1 \leq p < n$ and $p \leq q \leq p^*$, $W^{1,p}(\mathbb{R}^n)$ is continuously embedded in $L^q(\mathbb{R}^n)$.

From (468), one might expect that when $n = p$, $W^{1,p}(\mathbb{R}^n)$ is embedded in $L^\infty(\mathbb{R}^n)$. This is not the case when $n > 1$. We will not use the fact there, but $W^{1,p}(\mathbb{R}^n)$ is continuously embedded in the Banach space of functions of bounded mean oscillation (see, for instance, [**6**] for a proof). We now summarize the preceding discussion and prove a result for the case $n = p$.

THEOREM 79. *If* $1 \leq p < n$ *and*

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}, \tag{484}$$

then $W^{1,p}(\mathbb{R}^n)$ is continuously embedded in $L^q(\mathbb{R}^n)$ for all $q \in [p, p^*]$. Moreover, $W^{1,n}(\mathbb{R}^n)$ is continuously embedded in $L^q(\mathbb{R}^n)$ for all $n \leq q < \infty$. If $p > n$, then $W^{1,p}(\mathbb{R}^n)$ is continuously embedded in the Hölder space $C^{0,\gamma}\left(\overline{\mathbb{R}^n}\right)$ with $\gamma = 1 - n/p$.

PROOF. It only remains to prove the statement regarding $W^{1,n}(\mathbb{R}^n)$. Arguing as in the proof of Theorem 77, we see that

$$\|u\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq \prod_{i=1}^{N} \left\| \frac{\partial u}{\partial x_i} \right\|_{L^1(\mathbb{R}^n)}^{\frac{1}{n}} \tag{485}$$

when $u \in C_c^1(\mathbb{R}^n)$. We apply (485) to $u(x)|u(x)|^{m-1}$ with $m \geq 1$ and proceed as in the proof of Theorem 77 to obtain

$$\|u\|^m_{L^{\frac{mn}{n-1}}(\mathbb{R}^n)} \leq m \|u\|^{m-1}_{L^{\frac{(m-1)n}{n-1}}(\mathbb{R}^n)} \|Du\|_{L^n(\mathbb{R}^n)}. \tag{486}$$

Applying Young's inequality yields

$$\|u\|_{L^{\frac{mn}{n-1}}(\mathbb{R}^n)} \leq C \left(\|u\|_{L^{\frac{(m-1)n}{n-1}}(\mathbb{R}^n)} + \|Du\|_{L^n(\mathbb{R}^n)} \right). \tag{487}$$

Now we let $m = n$ to obtain

$$\|u\|_{L^{\frac{n^2}{n-1}}(\mathbb{R}^n)} \leq C \|u\|_{W^{1,n}(\mathbb{R}^n)}. \tag{488}$$

It follows that $W^{1,p}(\mathbb{R}^n)$ is continuously embedded in $L^q(\mathbb{R}^n)$ for all $n \leq q \leq n^2/(n-1)$. By repeating this argument with $m = n+1$, $m = n+2$, etc., we obtain the desired result. $\square$

Since $C_c^1(\Omega)$ is dense in $W_0^{1,p}(\Omega)$, we have the following analogous results for $W_0^{1,p}(\Omega)$.

THEOREM 80. *Suppose that Ω is an open set in $\mathbb{R}^n$. If $1 \leq p < n$ and*

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}, \tag{489}$$

then $W_0^{1,p}(\Omega)$ is continuously embedded in $L^q(\Omega)$ for all $q \in [p, p^*]$. The space $W_0^{1,n}(\Omega)$ is continuously embedded in $L^q(\Omega)$ for all $n \leq q < \infty$. Also, if $p > n$, then $W_0^{1,p}(\Omega)$ is continuously embedded in $C^{0,\gamma}\left(\overline{\Omega}\right)$ with $\gamma = 1 - \frac{n}{p}$.

If Ω is a bounded open set in $\mathbb{R}^n$, then Hölder's inequality implies that $L^q(\Omega)$ is continuously embedded in $L^p(\Omega)$ provided $q > p$. Moreover, in this case $C\left(\overline{\Omega}\right)$ is continuously embedded in $L^p(\Omega)$ for all $1 \leq p \leq \infty$. So we have the following:

THEOREM 81. *Suppose that $\Omega \subset \mathbb{R}^n$ is bounded. If $1 \leq p < n$ and*

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}, \tag{490}$$

then $W_0^{1,p}(\Omega)$ is continuously embedded in $L^q(\Omega)$ for all $1 \leq q \leq p^*$. The space $W_0^{1,n}(\Omega)$ is continuously embedded in $L^q(\Omega)$ for all $1 \leq q < \infty$. Finally, if $p > n$, then $W_0^{1,p}(\Omega)$ is continuously embedded in $L^q(\Omega)$ for all $1 \leq q \leq \infty$.

If Ω is a Lipschitz domain, then we may use the extension theorem together with Theorem 79 to establish the following.

THEOREM 82. *Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain. If $1 \leq p < n$,*

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n} \tag{491}$$

and $1 \leq q \leq p^*$, then there is a constant C such that

$$\|u\|_{L^q(\Omega)} \leq C \|u\|_{W^{1,p}(\Omega)} \tag{492}$$

for all $u \in W^{1,p}(\Omega)$. For all $p \leq q < \infty$, there exists a constant C' such that

$$\|u\|_{L^q(\Omega)} \leq C' \|u\|_{W^{1,n}(\Omega)} \tag{493}$$

for all $u \in W^{1,p}(\Omega)$. If $p > n$ and $\gamma = 1 - n/p$, then every $u \in W^{1,p}(\Omega)$ is equal almost everywhere equal to an element of $C^{0,\gamma}\left(\overline{\Omega}\right)$, and there is a constant C'' such that

$$\|u\|_{C^{0,\gamma}\left(\overline{\Omega}\right)} \leq C'' \|u\|_{W^{1,p}(\Omega)} \tag{494}$$

for all $u \in W^{1,p}(\Omega)$.

$$\tag{495}$$

PROOF. We prove only the first statement; the rest follow in a similar fashion. Since the boundary of Ω is Lipschitz, there exists a simple $(1, p)$ extension operator $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$. Let C be a constant such that

$$\|E[u]\|_{W^{1,p}(\mathbb{R}^n)} \leq C \|u\|_{W^{1,p}(\Omega)} \tag{496}$$

for all $u \in W^{1,p}(\Omega)$ and let C' be a constant such that

$$\|u\|_{L^{p*}(\mathbb{R}^n)} \leq C' \|u\|_{W^{1,p}(\mathbb{R}^n)} \tag{497}$$

for all $u \in W^{1,p}(\mathbb{R}^n)$. Then for all $u \in W^{1,p}(\Omega)$, we have

$$\|u\|_{L^{p^*}(\Omega)} \leq \|E[u]\|_{L^{p^*}(\mathbb{R}^n)} \leq C' \|E[u]\|_{W^{1,p}(\mathbb{R}^n)} \leq CC' \|u\|_{W^{1,p}(\Omega)}, \tag{498}$$

which shows that $W^{1,p}(\Omega)$ is continuously embedded in $L^{p^*}(\Omega)$. By Hölder's inquality, $L^{p^*}(\Omega)$ is continuously embedded in $L^q(\Omega)$ for all $1 \leq q \leq p^*$. The first conclusion of the theorem is now established. $\square$

It is a consequence of the following theorem that if Ω is bounded then the $W_0^{1,p}(\Omega)$ norm is equivalent to the norm $\|u\| = \|Du\|_{L^p(\Omega)}$. This will play an important role in Chapter 4.

THEOREM 83 (Poincare's Inequality). *Suppose that Ω is a bounded open subset in $\mathbb{R}^n$, and that $1 \leq p < \infty$. Then there is a constant C such that*

$$\|u\|_{L^p(\Omega)} \leq C \|Du\|_{L^p(\Omega)} \tag{499}$$

for all $u \in W_0^{1,p'}(\Omega)$.

PROOF. Suppose first that $1 \leq p < n$, and let

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}. \tag{500}$$

Then $W^{1,p}(\Omega)$ is continuously embedded in $L^{p^*}(\Omega)$. Since Ω is bounded, Hölder's inequality implies that $L^{p^*}(\Omega)$ is continuously embedded in $L^q(\Omega)$ for all $1 \leq q \leq p^*$. In particular, for all $1 \leq q \leq p^*$, there exists $C > 0$ such that

$$\|u\|_{L^q(\Omega)} \leq C\|Du\|_{L^p(\Omega)} \tag{501}$$

whenever $u \in W_0^{1,p}(\Omega)$. Since $p^* > p$, the theorem is proved in this case.

Suppose now that $p \geq n$. Since

$$p^* = \frac{1}{p'} - \frac{1}{n} = \frac{np}{n-p'} \to \infty \quad \text{as} \quad p' \to n, \tag{502}$$

we can choose $1 \leq p' < n$ such that $p^* > p$. It follows from the above discussion that there exists a constant C_1 such that

$$\|u\|_{L^{p*}(\Omega)} \leq C_1\|Du\|_{L^{p'}(\Omega)}. \tag{503}$$

for all $u \in W_0^{1,p}(\Omega)$. Since Ω is bounded and $p^* > p \geq n > p'$, Hölder's inequality implies that there exist constants C_2 and C_3 such that

$$\|u\|_{L^p(\Omega)} \leq C_2\|u\|_{L^{p^*}(\Omega)} \leq C_1C_2\|Du\|_{L^{p'}(\Omega)} \leq C_1C_2C_3\|Du\|_{L^p(\Omega)}. \tag{504}$$

□

By iterating the preceding results, the following theorems can be easily obtained.

Theorem 84. *Suppose that $\Omega \subset \mathbb{R}^n$ is Lipschitz domain, that $k < \frac{n}{p}$, and that*

$$\frac{1}{q} = \frac{1}{p} - \frac{k}{n}. \tag{505}$$

Then there exists a constant C such that

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{W^{k,p}(\Omega)} \tag{506}$$

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

Theorem 85. *Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain, that $k > \frac{n}{p}$. Then every $u \in W^{k,p}(\Omega)$ is almost everywhere equal to a function in $C^{k-\lfloor \frac{n}{p} \rfloor -1,\gamma}\left(\overline{\Omega}\right)$, where γ is any positive real number less than 1 if n/p is an integer, and*

$$\gamma = \left\lfloor \frac{n}{p} \right\rfloor + 1 - \frac{n}{p} \tag{507}$$

otherwise. Moreover, there exists a constant C such that

$$\|u\|_{C^{k-\lfloor \frac{n}{p} \rfloor -1,\gamma}\left(\overline{\Omega}\right)} \leq C\|u\|_{W^{k,p}(\Omega)} \tag{508}$$

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

3.9. Compact Embeddings

In this section, we will establish that the embeddings discussed in the preceding section are, in fact, compact. We say a Banach space X is compactly embedded in the Banach space Y if there exists a compact injective map $X \to Y$. Usually X is a subset of Y and the mapping under consideration is the inclusion map.

THEOREM 86. *Suppose that Ω is a bounded open set in $\mathbb{R}^n$. If $1 \leq p < n$ then $W_0^{1,p}(\Omega)$ is compactly embedded in $L^q(\Omega)$ for all $1 \leq q < p^*$, where*

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}. \tag{509}$$

If $p > n$, then $W_0^{1,p}(\Omega)$ is compactly embedded in $C\left(\overline{\Omega}\right)$. Moreover, $W_0^{1,n}(\Omega)$ is compactly embedded in $L^q(\Omega)$ for all $n \leq q < \infty$. In particular, $W_0^{1,p}(\Omega)$ is compactly embedded in $L^p(\Omega)$ for all $1 \leq p < \infty$.

PROOF. When $p > n$, the theorem follows immediately from the the Arzelá-Ascoli theorem and the fact that $W_0^{1,p}(\Omega)$ is continuously embedded in $C^{0,\gamma}\left(\overline{\Omega}\right)$ for some $\gamma > 0$. When $p = n$, the theorem reduces to the case $p < n$ since $W_0^{1,n}(\Omega)$ is continuously embedded in $W_0^{1,p}(\Omega)$ for all $p < n$ (since Ω is bounded) and $p^* \to \infty$ as $p \to n$. So we need to prove the theorem only in the case $1 \leq p < n$.

We will first show that $W_0^{1,p}(\Omega)$ is compactly embedded in $L^1(\Omega)$ by showing that if A is a bounded set in $W_0^{1,p}(\Omega)$, then A is totally bounded in $L^1(\Omega)$. Recall that a set is totally bounded if given $\epsilon > 0$ then exist a covering of A by a finite collection of open balls of radius ϵ.

We claim that it suffices to consider the case in which $A \subset C_c^1(\Omega)$ since $C_c^1(\Omega)$ is dense in $W_0^{1,p}(\Omega)$. To see this, we first let C' be a constant such that

$$\|f\|_{L^1(\Omega)} \leq C' \|f\|_{W^{1,p}(\Omega)} \tag{510}$$

for all $f \in W^{1,p}(\Omega)$. Note that $W_0^{1,p}(\Omega)$ is embedded in $L^1(\Omega)$ since Ω is bounded. If A is a bounded subset of $W_0^{1,p}(\Omega)$ and $\epsilon > 0$, then for each $u \in A$, we choose $v \in C_c^1\left(\overline{\Omega}\right)$ such that $\|u - v\|_{W_0^{1,p}(\Omega)} < \frac{\epsilon}{2C'}$. The set A' of functions v formed in this fashion is bounded in $W_0^{1,p}(\Omega)$. If it is totally bounded in $L^1(\Omega)$, then we may choose a finite collection $B_1, \ldots, B_m$ of open balls of radius $\frac{\epsilon}{2}$ in $L^1(\Omega)$ with centers $c_1, \ldots, c_m$ which cover A'. In this event, if $u \in A$, v is the corresponding element of A', and c_j is the center of a ball in $L^1(\Omega)$ of radius $\frac{\epsilon}{2}$ which contains v, then

$$\begin{aligned} \|u - c_j\|_{L^1(\Omega)} &\leq \|u - v\|_{L^1(\Omega)} + \|v - c_j\|_{L^1(\Omega)} \\ &< C' \|u - v\|_{W_0^{1,p}(\Omega)} + \frac{\epsilon}{2} \\ &< \epsilon. \end{aligned} \tag{511}$$

So the open balls of radius ϵ centered at the points $c_1, \ldots, c_m$ cover A.

To reiterate, we assume without loss of generality that $A \subset C_c^1\left(\overline{\Omega}\right)$ such that $\|u\|_{W_0^{1,p}(\Omega)} \leq 1$ for all $u \in A$. For each $h > 0$, we define

$$A_h = \{\eta_h * u : u \in A\}. \tag{512}$$

Then A_h is bounded in $C\left(\overline{\Omega}\right)$ for

$$\begin{aligned} |\eta_h * u(x)| &= h^{-n}\left|\int_{|y|\le h}\eta\left(\frac{y}{h}\right)u\left(x-y\right)\ dy\right| \\ &= \left|\int_{|z|\le 1}\eta\left(z\right)u\left(x-hz\right)\ dz\right| \\ &\le C\left|\int_{|z|\le 1}u\left(x-hz\right)\ dz\right| \\ &\le C\|u\|_{L^1(\Omega)}. \end{aligned} \tag{513}$$

We observe that

$$\begin{aligned} |D_i\left(\eta_h * u(x)\right)| &= h^{-n}\left|\int_{|x-y|\le h}D_i\eta\left(\frac{x-y}{h}\right)u(y)\ dy\right| \\ &= \left|\int_{|z|\le 1}D_i\eta\left(z\right)u(x-hz)\ dz\right| \\ &\le \sup_{x\in\Omega}|D\eta(x)|\|u\|_{L^1(\Omega)} \end{aligned} \tag{514}$$

for all $i = 1, \ldots, n$. Thus

$$|Du_h(x)| \le C\|u\|_{L^1(\Omega)} \tag{515}$$

for all $x \in \Omega$, where $u_h = \eta_h * u$. It follows that

$$\begin{aligned} |u_h(x) - u_h(x+\delta)| &= \left|\int_0^1 \frac{d}{dt}u_h(x+t\delta)\ dt\right| \\ &\le \left|\int_0^1 Du_h(x+t\delta)\cdot\delta\ dt\right| \\ &\le C|\delta|\|u\|_{L^1(\Omega)}. \end{aligned} \tag{516}$$

In particular, A_h is equicontinuous.

It follows from the Arzelá-Ascoli theorem that A_h is precompact in $C\left(\overline{\Omega}\right)$. In particular, A_h is totally bounded in $C\left(\overline{\Omega}\right)$. Since $C\left(\overline{\Omega}\right)$ is continuously embedded in $L^1\left(\Omega\right)$, it follows that A_h is totally bounded in $L^1\left(\Omega\right)$.

Now we observe that

$$\begin{aligned} |u(x) - \eta_h * u(x)| &\le \int_{|z|\le 1}\eta(z)\left|u(x) - u(x-hz)\right|\ dz \\ &= \int_{|z|\le 1}\eta(z)\left|\int_0^h \frac{d}{dt}\left(u(x-tz)\right)dt\right|\ dz \\ &\le \int_{|z|\le 1}\eta(z)\left|\int_0^h Du(x-tz)\cdot z\ dt\right|\ dz \\ &\le \int_{|z|\le 1}\eta(z)\int_0^h |Du(x-tz)|\ dt\ dz. \end{aligned} \tag{517}$$

and we integrate over x to obtain

$$\begin{aligned}\int_{\Omega} |u(x) - \eta_h * u(x)| \ dx &\leq \int_{\Omega} \int_{|z|\leq 1} \eta(z) \int_0^h |Du(x - tz)| \ dt \ dz \ dx \\ &\leq \|Du\|_{L^1(\Omega)} \int_{|z|\leq 1} \eta(z) \int_0^h \ dt \ dz \\ &\leq Ch\|Du\|_{L^1(\Omega)} \\ &\leq Ch.\end{aligned} \tag{518}$$

(we note that u has compact support in Ω, so that this procedure is sensible for sufficiently small h). It follows from this estimate that A is totally bounded in $L^1(\Omega)$ (since u_h is uniformly close to u in the L^1 norm).

Now we suppose that $q < p^* = 1/p - 1/n$. We will use the fact that if $0 < p < q < r \leq \infty$ and

$$\frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{r}, \tag{519}$$

then

$$\|f\|_q \leq \|f\|_p^{\lambda} \|f\|_r^{1-\lambda} \tag{520}$$

(see, for instance, Chapter 6 of [**8**]). We apply this identity to obtain

$$\|u\|_{L^q(\Omega)} \leq \|u\|_{L^1(\Omega)}^{\lambda} \|u\|_{L^{p^*}(\Omega)}^{1-\lambda}. \tag{521}$$

Now it follows from the embedding theorem that there is a constant C such that

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{L^1(\Omega)}^{\lambda} \|u\|_{W_0^{1,p}(\Omega)}^{1-\lambda} \tag{522}$$

Since A is precompact in $L^1(\Omega)$ and bounded in $W_0^{1,p}(\Omega)$, it follows easily from this inequality that it is precompact in $L^q(\Omega)$. □

By combining the previous theorem with the Sobolev extension theorem, we easily obtain the following result.

THEOREM 87. *Suppose that Ω is a bounded Lipschitz domain in $\mathbb{R}^n$. If $1 \leq p < n$ then $W^{1,p}(\Omega)$ is compactly embedded in $L^q(\Omega)$ for all $1 \leq q \leq p^*$, where*

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}. \tag{523}$$

If $p > n$, then $W^{1,p}(\Omega)$ is compactly embedded in $C\left(\overline{\Omega}\right)$. Moreover, $W^{1,n}(\Omega)$ is compactly embedded in $L^q(\Omega)$ for all $n \leq q < \infty$. In particular, $W^{1,p}(\Omega)$ is compactly embedded in $L^p(\Omega)$ for all $1 \leq p < \infty$.

It is straightforward to extend these theorems to the spaces $W^{k,p}(\Omega)$ by iterating them. We observe, in particular, that $H^{k+1}(\Omega)$ is compactly embedded in $H^k(\Omega)$ whenever Ω is a Lipschitz domain and $H_0^{k+1}(\Omega)$ is compactly embedded in $H_0^k(\Omega)$ when Ω is a bounded open set in $\mathbb{R}^n$.

EXERCISE 30. *Suppose that Ω a bounded open set in $\mathbb{R}^n$, and that $s > t > 0$ are real numbers. Show that $H_0^s(\Omega)$ is compactly embedded in $H_0^t(\Omega)$. (Hint: use the "Fourier" Sobolev norm for fractional order spaces.)*

3.10. 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 5.

Suppose that Ω 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

$$\Delta_i^h u(x) = \frac{u(x + he_i) - u(x)}{h}. \tag{524}$$

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

$$\Delta^h u(x) = \begin{pmatrix} \Delta_1^h u(x) \\ \Delta_2^h u(x) \\ \vdots \\ \Delta_n^h u(x) \end{pmatrix} \tag{525}$$

and define

$$\left\|\Delta^h u\right\|_p = \left(\|\Delta_1^h u(x)\|_p^p + \cdots + \|\Delta_n^h u(x)\|_p^p\right)^{1/p} \tag{526}$$

for $1 \leq p < \infty$ and

$$\left\|\Delta^h u\right\|_\infty = \|\Delta_1^h u(x)\|_\infty + \cdots + \|\Delta_n^h u(x)\|_\infty. \tag{527}$$

THEOREM 88. *Suppose that Ω 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 Ω' is an open set such that $\Omega' \subset\subset \Omega$. Then*

$$\left\|\Delta_i^h u(x)\right\|_{L^p(\Omega')} \leq \|D_i u\|_{L^p(\Omega)} \tag{528}$$

whenever $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$.

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

$$\Delta_i^h u(x) = \frac{u(x + he_i) - u(x)}{h} = \frac{1}{h}\int_0^h D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)\ dt \tag{529}$$

for all $x \in \Omega'$. If $p = \infty$, then by taking absolute values on both sides of (529) we obtain

$$\left|\Delta_i^h u(x)\right| \leq \sup_{x \in \Omega', 0 < t < \operatorname{dist}(\Omega', \partial\Omega)} |D_i u(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)| \leq \sup_{x \in \Omega} |D_i u(x)|, \tag{530}$$

from which we conclude that (528) holds when $p = \infty$ and u is infinitely differentiable. Otherwise, we take the p^{th} power of both sides of (529) to obtain

$$\begin{aligned} \left|\Delta_i^h u(x)\right|^p &= \frac{1}{h^p}\left|\int_0^h D_i u(x_1,\dots,x_{i-1},x_i+t,x_{i+1},\dots,x_n)\ dt\right|^p \\ &\leq \frac{1}{h^p}\left(\int_0^h \left|D_i u(x_1,\dots,x_{i-1},x_i+t,x_{i+1},\dots,x_n)\right|\ dt\right)^p \end{aligned} \tag{531}$$

We see from Hölder's inequality that

$$\begin{aligned} &\int_0^h \left|D_i u(x_1,\dots,x_{i-1},\ x_i+t,x_{i+1},\dots,x_n)\right|\ dt \\ &\qquad\leq h^{1/q}\left(\int_0^h \left|D_i u(x_1,\dots,x_{i-1},x_i+t,x_{i+1},\dots,x_n)\right|^p\ dt\right)^{1/p}, \end{aligned} \tag{532}$$

where

$$\frac{1}{p}+\frac{1}{q}=1. \tag{533}$$

By inserting (532) into (531), we obtain

$$\begin{aligned} \left|\Delta_i^h u(x)\right|^p &\leq h^{p/q-p}\int_0^h \left|D_i u(x_1,\dots,x_{i-1},x_i+t,x_{i+1},\dots,x_n)\right|^p\ dt \\ &= \frac{1}{h}\int_0^h \left|D_i u(x_1,\dots,x_{i-1},x_i+t,x_{i+1},\dots,x_n)\right|^p\ dt. \end{aligned} \tag{534}$$

Note that $p/q - p = -1$ follows from (533). 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

$$\begin{aligned} \frac{1}{h}\int_0^h \left|D_i u(x_1,\dots,x_{i-1},x_i+t,x_{i+1},\dots,x_n)\right|^p\ dt &= \left|D_i u(x_1,\dots,x_{i-1},x_i+\xi_x,x_{i+1},\dots,x_n)\right|^p \\ &= \left|D_i u(x+\xi_x e_i)\right|^p. \end{aligned} \tag{535}$$

We insert (535) into (534) and integrate over Ω' in order to obtain

$$\begin{aligned} \int_{\Omega'} \left|\Delta_i^h u(x)\right|^p\ dx &= \int_{\Omega'} \left|D_i u(x+\xi_x e_i)\right|^p\ dx \\ &\leq \int_{\Omega} \left|D_i u(x))\right|^p\ dx, \end{aligned} \tag{536}$$

from which we conclude that (528) 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 (528) holds for arbitrary $u \in W^{k,p}(\Omega)$ now follows from Theorem 57 — that is, the observation that $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$. $\square$

Note that $\Delta_i^h u$ is an element of $L^p(\Omega')$ when $u \in L^p(\Omega)$ whenever $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$; in fact,

$$\left\|\frac{u(\cdot + he_i) - u(\cdot)}{h}\right\|_{W^{k,p}(\Omega')} \leq \frac{2}{h}\|u\|_{W^{k,p}(\Omega)}. \tag{537}$$

Theorem 88 is useful because it gives us a bound on the $L^p(\Omega')$ norm of $\Delta^h u$ which is independent of h. We now establish that the converse also holds; that is, if the L^p norm of $\Delta_i^h u$ is bounded independently of h, then the weak derivative $D_i u$ exists and satisfies the same bound.

THEOREM 89. *Suppose that Ω 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 Ω' is an open set with $\Omega' \subset\subset \Omega$ and $0 < h < \operatorname{dist}(\Omega', \partial\Omega)$,*

$$\left\|\Delta_i^h u\right\|_{L^p(\Omega')} \leq C \tag{538}$$

Then the weak derivative $D_i u$ exists in Ω and

$$\|D_i u\|_{L^p(\Omega)} \leq C. \tag{539}$$

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 Ω such that

$$\Omega = \bigcup_{j=1}^{\infty} \Omega_j. \tag{540}$$

Since bounded sets in $L^p(\Omega_1)$ are weakly compact (see Theorem 34 in Section 2.5) and

$$\|\Delta_i^h u\|_{L^p(\Omega_1)} \leq C \tag{541}$$

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

$$\lim_{j\to\infty} \int_{\Omega_1} \Delta_i^{h_1(j)} u(x)\psi(x)\ dx = \int_{\Omega_1} v_1(x)\psi(x)\ dx \tag{542}$$

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

$$\lim_{j\to\infty} \int_{\Omega_2} \Delta_i^{h_2(j)} u(x)\psi(x)\ dx = \int_{\Omega_2} v_2(x)\psi(x)\ dx \tag{543}$$

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

$$\lim_{j\to\infty} \int_{\Omega_k} \Delta_i^{h_k(j)} u(x)\psi(x)\ dx = \int_{\Omega_k} v(x)\psi(x)\ dx \tag{544}$$

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

$$s_k = h_k(k); \tag{545}$$

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 $s_k < \operatorname{dist}(\operatorname{supp}(\psi), \partial\Omega)$ for all $k \geq l$. Consequently,

$$\int_\Omega \Delta_i^{s_k} u(x)\psi(x)\ dx = \int_{\operatorname{supp}(\psi)} \Delta_i^{s_k} u(x)\psi(x)\ dx \tag{546}$$

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

$$\lim_{k\to\infty} \int_\Omega \Delta_i^{s_k} u(x)\psi(x)\ dx = \int_\Omega v(x)\psi(x)\ dx. \tag{547}$$

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

$$\lim_{k\to\infty} \int_\Omega \Delta_i^{s_k} u(x)\psi(x)\ dx = \lim_{k\to\infty} \int_\Omega u(x)\Delta_i^{-s_k}\psi(x)\ dx = -\int_\Omega u(x)D_i\psi(x)\ dx. \tag{548}$$

From (547) and (548) we obtain

$$\int_\Omega u(x)D_i\psi(x)\ dx = -\int_\Omega v(x)\psi(x)\ dx. \tag{549}$$

Since ψ 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 Ω is a bounded Lipschitz domain.

THEOREM 90. *Suppose that Ω is an open subset of $\mathbb{R}^n$. Then $W^{1,\infty}_{loc}(\Omega) = C^{0,1}(\Omega)$.*

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 $0 < h < \text{dist}(\Omega', \partial\Omega)$

$$\|\Delta_i^h u\|_{L^\infty(\Omega')} \leq C, \tag{550}$$

where C is the Lipschitz constant for u in Ω'. Since Ω' is bounded, (550) 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,

$$\int_{\Omega'} \Delta_i^{h_j} u(x)\varphi(x)\ dx \to \int_{\Omega'} v(x)\varphi(x)\ dx \tag{551}$$

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

$$\int_{\Omega'} \Delta_i^{h_j} u(x)\varphi(x)\ dx = -\int_{\Omega'} u(x)\Delta_i^{-h_j}\varphi(x)\ dx \to -\int_{\Omega'} u(x)D_i\varphi(x)\ dx. \tag{552}$$

We combine (551) and (552) in order to obtain

$$\int_{\Omega'} u(x)D_i\varphi(x)\ dx = -\int_{\Omega'} v(x)\varphi(x)\ dx, \tag{553}$$

from which we conclude that $D_i u = v$. From (550) and (551) we see that

$$\left|\int_{\Omega'} v(x)\varphi(x)\ dx\right| = \left|\lim_{j\to\infty} \int_{\Omega'} \Delta_i^{h_j} u(x)\varphi(x)\ dx\right| \leq = C\|\varphi\|_{L^1(\Omega')} \tag{554}$$

for all $\varphi \in L^1(\Omega')$. We conclude that $v \in L^\infty(\Omega')$ (see, for instance, Theorem 6.13 in [**8**]).

We now suppose that $u \in W^{1,\infty}_{\text{loc}}(\Omega)$, and that Ω' is an open ball contained in Ω. For each $0 < h < \text{dist}(\Omega', \partial\Omega)$, we define u_h via the formula

$$u_h(x) = \eta_h * u(x), \tag{555}$$

where η_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 55 in order to see that

$$\|Du_h\|_{L^\infty(\Omega')} = \|D(\eta_h * u)\|_{L^\infty(\Omega')} \|\eta_h * Du\|_{L^\infty(\Omega')} \leq \|\eta_h\|_1 \|Du\|_{L^\infty(\Omega')} = \|Du\|_{L^\infty(\Omega')} < \infty \tag{556}$$

for all $0 < h < \operatorname{dist}(\Omega', \partial\Omega)$. Note that in (556) we use Du to refer to the weak gradient of u. Since $u_h \in C^\infty(\Omega')$,

$$u_h(x) - u_h(y) = \int_0^1 Du_h(y + t(x-y))\ dt \cdot (x-y) \tag{557}$$

for all $x, y \in \Omega'$. We conclude from (557) that

$$|u_h(x) - u_h(y)| \leq \|Du\|_{L^\infty(\Omega')} |x-y| \tag{558}$$

for all $x, y \in \Omega'$ and all sufficiently small h. By taking the limit as $h \to 0$ in (558), we see that

$$|u(x) - u(y)| \leq \|Du\|_{L^\infty(\Omega')} |x-y| \tag{559}$$

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 Ω'. 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 (559), 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 $C^{0,1}(\Omega')$. $\square$

By combining Theorem 90 with Theorem 68, we obtain the following:

THEOREM 91. *Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain. Then $W^{1,\infty}(\Omega) = C^{0,1}\left(\overline{\Omega}\right)$.*

CHAPTER 4

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

$$L\left[u\right](x) = -a^{ij}(x)D_iD_ju(x) + b^i(x)D_iu(x) + c(x)u(x), \tag{560}$$

and that u is a classical solution of the equation

$$L\left[u\right](x) = f(x) \tag{561}$$

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

$$\int_\Omega D_ju(x)D_i\left(a^{ij}(x)v(x)\right) + b_i(x)D^iu(x)v(x) + c(x)u(x)v(x)\ dx = \int_\Omega f(x)v(x)\ dx \tag{562}$$

for all sufficiently smooth functions v. Note that we are assuming that u vanishes on $\partial\Omega$, so that no boundary terms emerge in (562). The central observation of the variational theory of partial differential equations is that (562) is often sufficient to characterize the solution u of the partial differential equation (561). More specifically, under mild assumptions on the operator L, the forcing term f and the domain Ω, if (562) holds for all v in a suitable space of test functions then u solves (561). The great advantage of the variational formulation (562) over the equation (561) is that it requires less of u. In particular, (562) is sensible when u has only one *weak* derivative whereas u must have two classical derivatives in order for (561) to be meaningful.

A weakly differentiable function u which satisfies (562) is known as a *weak solution* of the equation (561). A twice weakly differentiable function u such that (560) holds almost everywhere is called a *strong solution* of (561). If u is twice differentiable and satisfies (560) everywhere, then it is a *classical solution* of (561). 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, partial differential operators of the form

$$L\left[u\right](x) = -D_i\left(a^{ij}(x)D_ju(x) + b^i(x)u(x)\right) + c^i(x)D_iu(x) + d(x)u(x), \tag{563}$$

where Ω 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

$$v \to \int_\Omega a^{ij}(x)D_iu(x)D_jv(x) + b^i(x)u(x)D_iv(x) + c^i(x)D_iu(x)v(x) + d(x)u(x)v(x)\ dx. \tag{564}$$

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 $H_0^1(\Omega)$. This ensures, among other things, that if $L\left[u\right] = f$ in the sense that

$$\left\langle L\left[u\right], v\right\rangle = \langle f, v\rangle \tag{565}$$

for all $v \in V$, then (562) holds for all test functions v in $C_c^\infty(\Omega)$ and it is reasonable to say that $L\left[u\right] = f$ "in the interior of Ω." 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 (564) 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

$$\sum a^{ij}\xi_i\xi_j \geq \lambda|\xi|^2 \tag{566}$$

for all $\xi \in \mathbb{R}^n$.

Note that in (564) we have implicitly embedded the space V in the dual space V^*. In particular, we have embedded V into V^* through the composition map

$$V \xrightarrow{\ \iota\ } L^2(\Omega) \xrightarrow{\ \varphi\ } \left(L^2(\Omega)\right)^* \xrightarrow{\ T\ } V^*, \tag{567}$$

where ι is the inclusion map

$$\iota : V \to L^2(\Omega), \tag{568}$$

φ 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

$$f_u(v) = \int_\Omega u(x)v(x)\ dx, \tag{569}$$

and $T : \left(L^2(\Omega)\right)^* \to V^*$ is the linear map defined by

$$T\left[\phi\right] = \phi|_V. \tag{570}$$

Note that the map T is bounded (obviously), injective (because V contains $C_c^\infty(\Omega)$ 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:

$$\langle v, u\rangle_{V\times V^*} = \langle v, u\rangle_{L^2(\Omega)} = \int_\Omega v(x)u(x)\ dx. \tag{571}$$

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

$$L\left[u\right](x) = -D_i\left(a^{ij}(x)D_ju(x)\right) \tag{572}$$

with a^{ij} are bounded measurable functions $\Omega \to \mathbb{R}$ (i.e. elements of $L^\infty\left(\Omega\right)$), and that there exists a real number $\lambda > 0$ such that

$$a^{ij}(x)\xi_i\xi_j \geq \lambda \left|\xi\right|^2 \tag{573}$$

for all $x \in \Omega$ and $\xi \in \mathbb{R}^n$ (that is, we are assuming that L is strongly elliptic in Ω). The operator L is a mapping from $H_0^1(\Omega) \to H^{-1}(\Omega)$; in particular, $L\left[u\right]$ is the mapping which takes $v \in H_0^1(\Omega)$ to

$$\int_\Omega a^{ij}(x)D_ju(x)D_iv(x)\ dx. \tag{574}$$

A function $u \in H_0^1\left(\Omega\right)$ is a weak solution of the Dirichlet boundary value problem

$$\begin{cases} L\left[u\right](x) = f(x) \ \text{ in } \ \Omega \\ \quad\ u(x) = 0 \ \text{ on } \ \partial\Omega, \end{cases} \tag{575}$$

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

$$L\left[u\right](v) = \langle f, v\rangle \tag{576}$$

for all $v \in H_0^1\left(\Omega\right)$. In (576), $\langle f, v\rangle$ refers to the duality pairing of V^* and V.

THEOREM 92. *The mapping L defined via formula (572) is bounded and coercive.*

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

$$\left|a^{ij}(x)\right| \leq \eta \tag{577}$$

for all $i, j = 1, \ldots, n$ and $x \in \Omega$. Using (577) and Hölder's inequality, we see that

$$\begin{aligned} \left|\langle L\left[u\right], v\rangle\right| &= \left|\int_\Omega a^{ij}(x)D_ju(x)D_iv(x)\ dx\right| \\ &\leq \eta \sum_{i,j=1}^n \int_\Omega \left|D_ju(x)D_iv(x)\right| dx \\ &\leq \eta \sum_{i,j=1}^n \|D_ju\|_2\ \|D_iv\|_2 \\ &\leq \eta \sum_{i,j=1}^n \|Du\|_2\ \|Dv\|_2 \\ &\leq n^2\eta\|u\|_{H_0^1(\Omega)}\ \|v\|_{H_0^1(\Omega)} \end{aligned} \tag{578}$$

for all u and v in $H_0^1\left(\Omega\right)$. We conclude that L is bounded.

The strong ellipticity of L implies that

$$a^{ij}(x)D_ju(x)D_iu(x) \geq \lambda\left|Du(x)\right|^2 \tag{579}$$

for all $u \in H_0^1(\Omega)$. From (579) we conclude that

$$\langle L[u], u\rangle = \int_\Omega a^{ij}(x)D_j u(x)D_i u(x)\, dx \geq \lambda \int_\Omega |Du(x)|^2\, dx = \lambda \|Du\|_2^2. \tag{580}$$

According to Poincaré's inequality, there exists a real number $\beta > 0$ such that

$$\|u\|_2 \leq \beta \|Du\|_2 \tag{581}$$

for all $u \in H_0^1(\Omega)$. By combining (580) and (581) we see that

$$\begin{aligned} \langle L[u], u\rangle &= \int_\Omega a^{ij}(x)D_j u(x)D_i u(x)\, dx \\ &\geq \lambda \int_\Omega |Du(x)|^2\, dx \\ &= \lambda \|Du\|_2^2 \\ &\geq \frac{\lambda}{2}\|Du\|_2^2 + \frac{\lambda}{2\beta}\|u\|_2^2 \\ &\geq \min\left\{\frac{\lambda}{2}, \frac{\lambda}{2\beta}\right\} \|u\|_{H_0^1(\Omega)}^2 \end{aligned} \tag{582}$$

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 92, we can apply the Lax-Milgram theorem (Theorem 27 in Section 2.4) in order to conclude that (575) admits a *unique* weak solution u, and that there exists a constant C (depending on L and Ω but not f) such that

$$\|u\|_{H_0^1(\Omega)} \leq C \|f\|_{H^{-1}(\Omega)}. \tag{583}$$

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

$$\langle u, v\rangle = \langle L[u], v\rangle \tag{584}$$

and the Riesz representation theorem suffices to establish the existence of weak solutions of (575).

We now reduce the inhomogeneous boundary value problem

$$\begin{cases} L[u](x) = f(x) & \text{in } \Omega \\ u(x) = g & \text{on } \partial\Omega \end{cases} \tag{585}$$

to a homogeneous problem of the form (575). As before, we assume that $f \in H^{-1}(\Omega)$ and, in addition, we assume that g is the trace of a function ψ in $H^1(\Omega)$. In order for the trace operator to be defined, we will need to make some assumptions on the regularity of the boundary of Ω. We will assume that Ω is a bounded Lipschitz domain. We let w be a weak solution of the boundary value problem

$$\begin{cases} L[w](x) = f(x) - L[\psi](x) = f(x) + D_i\left(a^{ij}(x)D_j\psi(x)\right) & \text{in } \Omega \\ w(x) = 0 & \text{on } \partial\Omega. \end{cases} \tag{586}$$

By $f(x) + D_i\left(a^{ij}(x)D_j\psi(x)\right)$, we mean the bounded linear functional on $H_0^1(\Omega)$ defined via

$$v \to \langle f, v\rangle + \int_\Omega a^{ij} D_j\psi(x) D_i v(x)\ dx. \tag{587}$$

We observe that

$$\left|\int_\Omega a^{ij} D_j\psi(x) D_i v(x)\ dx\right| \le C\|\psi\|_{H^1(\Omega)}\ \|v\|_{H_0^1(\Omega)} \tag{588}$$

(the argument is identical to that used in the proof of Theorem 92 to show that B is bounded) so that

$$\|f + D_i\left(a^{ij} D_j\psi\right)\|_{H^{-1}(\Omega)} \le C\left(\|f\|_{H^{-1}(\Omega)} + \|\psi\|_{H^1(\Omega)}\right). \tag{589}$$

From (586), we see that

$$\int_\Omega L\left[w+\psi\right](x) v(x)\ dx = \langle f, v\rangle \tag{590}$$

for all $v \in H_0^1(\Omega)$, and that the trace of $w+\psi$ is g. In other words, $w+\psi$ is a weak solution of boundary value problem (585). Moreover, the Lax-Milgram theorem together with (589) implies that

$$\|w\|_{H^1(\Omega)} \le C\left(\|f\|_{H^{-1}(\Omega)} + \|\psi\|_{H^1(\Omega)}\right), \tag{591}$$

from which we obtain the bound

$$\|w+\psi\|_{H^1(\Omega)} \le C\left(\|f\|_{H^{-1}(\Omega)} + \|\psi\|_{H^1(\Omega)}\right) \tag{592}$$

for the solution $w+\psi$ of (585). We summarize our conclusions in the following theorem.

Theorem 93. *Suppose that Ω is a bounded open set in $\mathbb{R}^n$ with Lipschitz boundary, that $\mathscr{T} : H^1(\Omega) \to L^2(\partial\Omega)$ denotes the trace operator, and that L is a strongly elliptic operator of the form*

$$L\left[u\right](x) = -D_i\left(a^{ij} D_j u\right) \tag{593}$$

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

$$\begin{cases} L\left[u\right](x) = f(x) & \text{in } \Omega \\ \quad u(x) = \mathscr{T}\left[\psi\right](x) & \text{on } \partial\Omega. \end{cases} \tag{594}$$

Moreover, there is a constant $C > 0$ depending on Ω and L such that

$$\|u\|_{H^1(\Omega)} \le C\left(\|f\|_{H^{-1}(\Omega)} + \|\psi\|_{H_0^1(\Omega)}\right) \tag{595}$$

whenever u is the weak solution of (594).

Suppose that $\mathscr{T}\left[\tilde\psi\right] = \mathscr{T}\left[\psi\right]$, that u is the solution of (594) constructed in the preceding theorem, and that $\tilde u$ is the solution of

$$\begin{cases} L\left[\tilde u\right](x) = f(x) & \text{in } \Omega \\ \quad \tilde u(x) = \mathscr{T}\left[\tilde\psi\right](x) & \text{on } \partial\Omega \end{cases} \tag{596}$$

constructed in the same fashion. Then $u - \tilde{u}$ solves

$$\begin{cases} L\left[\tilde{u}-u\right](x) = 0 & \text{in } \Omega \\ u(x) - \tilde{u}(x) = 0 & \text{on } \partial\Omega. \end{cases} \tag{597}$$

It obviously follows that $u - \tilde{u} = 0$; i.e., $u = \tilde{u}$. In particular, the solution of the boundary value problem depends only on the trace of of ψ.

Theorem 74 characterizes the traces of functions in $H^1(\Omega)$ when Ω 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

$$\|\psi\|_{H^1(\Omega)} \leq C\|g\|_{H^{1/2}(\partial\Omega)}. \tag{598}$$

We obtain the following theorem by combining this observation with Theorem 93.

Theorem 94. *Suppose that Ω is a Lipschitz domain in $\mathbb{R}^n$, and that L is a strongly elliptic operator of the form*

$$L[u](x) = -D_i\left(a^{ij}D_j u\right) \tag{599}$$

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

$$\begin{cases} L[u](x) = f(x) & \textit{in } \Omega \\ u(x) = g(x) & \textit{on } \partial\Omega. \end{cases} \tag{600}$$

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

$$\|u\|_{H^1(\Omega)} \leq C\left(\|f\|_{H^{-1}(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)}\right) \tag{601}$$

whenever u is a weak solution of the boundary value problem (600). In other words, the operator $L \oplus \mathscr{T}$ is an isomorphism

$$H^1(\Omega) \to H^{-1}(\Omega) \oplus H^{1/2}(\partial\Omega). \tag{602}$$

4.3. The Dirichlet Problem for General Second Order Operators

In this section, we treat the Dirichlet boundary value problem

$$\begin{cases} L[u](x) = f(x) & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega \end{cases} \tag{603}$$

for a more general class of strongly elliptic second order operators. More specifically, we suppose that L is of the form

$$L[u](x) = -D_i\left(a^{ij}(x)D_j u(x) + b^i(x)u(x)\right) + c^i(x)D_i u(x) + d(x)u(x) \tag{604}$$

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

$$a^{ij}(x)\xi_i\xi_j \geq \lambda|\xi|^2 \tag{605}$$

for all $\xi \in \mathbb{R}^n$. In order to develop existence and uniqueness results for the boundary value problem (603), we will need to use elementary results from the theory of Fredholm operators in addition to the Lax-Milgram theorem.

Since the coefficients of L are bounded, there exists a real number $\eta > 0$ such that

$$\|a^{ij}\|_{L^\infty(\Omega)} \leq \eta \tag{606}$$

for all $i, j = 1, \ldots, n$,

$$\|b^i\|_{L^\infty(\Omega)} + \|c^i\|_{L^\infty(\Omega)} \leq \eta \tag{607}$$

for all $i = 1, \ldots, n$, and $\|d\|_{L^\infty(\Omega)} \leq \eta$. We observe that

$$\begin{aligned}
&|\langle L[u], v\rangle| \\
&\leq \left| \int_\Omega a^{ij}(x) D_j u(x) D_i v(x) + b^i(x) u(x) D_i v(x) + c^i(x) D_i u(x) v(x) + d(x) u(x) v(x)\ dx \right| \\
&\leq \sum_{i,j=1}^n \eta \|D_j u\|_2 \|D_i v\|_2 + \sum_{i=1}^n \eta \|u\|_2 \|D_i v\|_2 + \sum_{i=1}^n \eta \|D_i u\|_2 \|v\|_2 + \eta \|u\|_2 \|v\|_2 \\
&\leq (n^2 + 2n + 1) \eta \|u\|_{H_0^1(\Omega)} \ \|v\|_{H_0^1(\Omega)},
\end{aligned}$$

from which we conclude that the operator L is bounded. We cannot apply the Lax-Milgram theorem to L directly since it is no longer necessarily coercive. We will proceed by combining the Lax-Milgram theorem with the Fredholm alternative. In particular, we show that the operator $L + \sigma I$ is invertible when σ is a sufficiently large real number.

For each real number $\sigma > 0$, we define a new linear partial differential operator L_σ via the formula

$$L_\sigma[u](x) = L[u](x) + \sigma I[u](x), \tag{608}$$

where I denotes the embedding of $H_0^1(\Omega)$ into $H^{-1}(\Omega)$. That is, I is the mapping which takes $u \in H_0^1(\Omega)$ to the mapping $\varphi_u \in H^{-1}(\Omega)$ defined via the formula

$$\varphi_u(v) = \int_\Omega u(x) v(x)\ dx. \tag{609}$$

THEOREM 95. *The linear operator $I : H_0^1(\Omega) \to H^{-1}(\Omega)$ is compact.*

PROOF. We factor I as $I = I_2 I_1$, where

$$I_1 : H_0^1(\Omega) \to L^2(\Omega) \tag{610}$$

is the inclusion map and

$$I_2 : L^2(\Omega) \to H^{-1}(\Omega) \tag{611}$$

is the embedding of $L^2(\Omega)$ into $H^{-1}(\Omega)$. According to the Rellich-Kondrachov theorem, I_1 is compact. Consequently, the composition $I_2 I_1$ is as well. □

From (4.3), it is clear that L_σ is bounded. Moreover:

THEOREM 96. *For all sufficiently large σ, the operator L_σ defined in (608) is coercive.*

PROOF. We let u be an arbitrary element of $H_0^1(\Omega)$. Using the assumption that L is strongly elliptic, we obtain

$$\begin{aligned}\lambda \int_\Omega |Du(x)|^2 \, dx &\leq \int_\Omega a^{ij}(x) D_j u(x) D_i u(x) \, dx \\ &= \langle L_\sigma[u], u\rangle - \int_\Omega \left(b^i(x) + c^i(x)\right) u(x) D_i u(x) \, dx + d(x) \left(u(x)\right)^2 \, dx \\ &\quad - \sigma \int_\Omega (u(x))^2 \, dx.\end{aligned} \tag{612}$$

By letting

$$a = (2\epsilon)^{-1/2} u(x) \tag{613}$$

and

$$b = (2\epsilon)^{1/2} D_i u(x), \tag{614}$$

where ϵ is a positive real number which we will choose shortly, in Cauchy's inequality

$$2ab \leq a^2 + b^2, \tag{615}$$

we obtain

$$u(x) D_i u(x) \leq \frac{1}{4\epsilon} \left(u(x)\right)^2 + \epsilon \left(D_i u(x)\right)^2. \tag{616}$$

We insert (616) into (612) in order to conclude that

$$\begin{aligned}\lambda \int_\Omega |Du(x)|^2 \, dx &\leq \langle L_\sigma[u], u\rangle + 2\eta \sum_{i=1}^n \left(\frac{1}{4\epsilon}\int_\Omega (u(x))^2 \, dx + \epsilon \int_\Omega (D_i u(x))^2 \, dx\right) \\ &\quad + (\eta - \sigma) \int_\Omega (u(x))^2 \, dx \\ &= \langle L_\sigma[u], u\rangle + \left(\frac{2\eta n}{4\epsilon} + \eta - \sigma\right) \int_\Omega |u(x)|^2 \, dx + 2\eta\epsilon \int_\Omega |Du|^2 \, dx,\end{aligned} \tag{617}$$

where η is as in (607). We rearrange (617) as

$$(\lambda - 2\eta\epsilon) \int_\Omega |Du(x)|^2 \, dx + \left(\sigma - \frac{2\eta n}{4\epsilon} - \eta\right) \int_\Omega |u(x)|^2 \, dx \leq \langle L_\sigma[u], u\rangle \tag{618}$$

and let $\epsilon = \frac{\lambda}{4\eta}$ in (618) in order to obtain

$$\frac{\lambda}{2} \int_\Omega |Du(x)|^2 \, dx + \left(\sigma - \frac{8\eta^2 n + 4\lambda\eta}{4\lambda}\right) \int_\Omega |u(x)|^2 \, dx \leq \langle L_\sigma[u], u\rangle, \tag{619}$$

from which conclude that L_σ is coercive when σ is sufficiently large. $\square$

We now choose $\sigma > 0$ so as to ensure that L_σ is coercive. Since L_σ is bounded and coercive, the Lax-Milgram theorem implies that for each $f \in H_0^1(\Omega)$ there is a unique element u of $H_0^1(\Omega)$ such that

$$\langle L_\sigma[u], v\rangle = (f, v) \tag{620}$$

for all $v \in H_0^1(\Omega)$, and that there exists a constant $C > 0$ such that

$$\|u\| \leq C\|f\|. \tag{621}$$

In other words, the operator $L+\sigma I : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isomorphism. It follows that the operator L is a Fredholm operator of index 0 since it is the sum of the isomorphism $(L+\sigma I)$ and the compact operator $-\sigma I$. We conclude from this observation and Theorem 25 that the boundary value problem (603) is uniquely solvable for each right-hand side f if and only if the homogeneous equation

$$\begin{cases} L\left[u\right](x) = 0 & \text{in} \quad \Omega \\ \quad u(x) = 0 & \text{on} \quad \partial\Omega \end{cases} \tag{622}$$

admits only the trivial solution. In the following section, we will give conditions in the operator L which ensure that this is the case. For now, we use the Fredholm theory to derive solvability conditions for the problem (603) which involve the adjoint L^* of L.

We recall that $L^* : H_0^1(\Omega) \to H^{-1}(\Omega)$ is defined via the relation

$$\langle L\left[u\right], v\rangle = \langle u, L^*\left[v\right]\rangle \tag{623}$$

for all $v \in H_0^1(\Omega)$, and that the closure of the image of T is the preannihilator of the kernel of T^* (see Section 2.3). From (623) it is easy to see that L^* is given by

$$L^*\left[v\right](x) = -D_i\left(a^{ji}(x)D_jv(x) + c^iv(x)\right) + b^i(x)D_iv(x) + d(x)v(x). \tag{624}$$

We let $p = \dim(\ker(L))$. Since L is Fredholm of index 0, the dimension of $\ker(L^*)$ is also p, and we let $v_1^*, \ldots, v_p^*$ be a basis for $\ker(L^*) \subset H_0^1(\Omega)$. That is, $v_1^*, \ldots, v_p^*$ is a basis in the space of solutions of the homogeneous adjoint problem

$$\begin{cases} L^*\left[u\right](x) = 0 & \text{in} \quad \Omega \\ \quad u(x) = 0 & \text{on} \quad \partial\Omega. \end{cases} \tag{625}$$

Then f is in the image of L if and only if

$$f\left(v_j^*\right) = 0 \quad \text{for all} \quad j = 1, \ldots, p. \tag{626}$$

In the case where $f \in L^2(\Omega)$ and we identify it with the mapping which takes $v \in H_0^1(\Omega)$ to

$$\int_\Omega f(x)v(x)\,dx, \tag{627}$$

(626) is equivalent to

$$\int_\Omega f(x)v_l^*(x)\,dx = 0 \quad \text{for all} \quad l = 1, \ldots, p. \tag{628}$$

In the event that (626) holds, the set of solutions of (603) is

$$\left\{u + v : v \in \ker(L)\right\}, \tag{629}$$

where u is any particular solution. Since L induces an isomoprhism

$$H_0^1(\Omega)/\ker(L) \to \operatorname{im}(L) \subset H^{-1}(\Omega), \tag{630}$$

there exists a constant $C > 0$ such that

$$\|u\|_{H_0^1(\Omega)/\ker(L)} \leq C\|f\|_{H^{-1}(\Omega)} \tag{631}$$

whenever u is a weak solution of (603).

There is an alternative estimate for the Sobolev norms of the solutions of (603) which is sometimes more useful. Suppose that u is a weak solution of (603) so that

$$\int_\Omega \big(a^{ij}(x)D_ju(x)D_iv(x) + b^i(x)u(x)D_iv(x) + c^i(x)D_iu(x)v(x) + d(x)u(x)v(x)\big)\ dx = \langle f, v\rangle \tag{632}$$

for all $v \in H^1_0(\Omega)$. By taking $v = u$ and rearranging terms, we obtain

$$\int_\Omega a^{ij}(x)D_ju(x)D_iu(x)\ dx = -\int_\Omega \big(b^i(x) + c^i(x))u(x)D_iu(x) + d(x)u(x)u(x)\big)\ dx + \langle f, u\rangle . \tag{633}$$

Now using the boundedness of f, the strong ellipticity of L and Cauchy's inequality, we see that

$$\lambda\|Du\|^2_{L^2(\Omega)}\leq 2\eta\|u\|_{L^2(\Omega)}\|Du\|_{L^2(\Omega)}+\eta\|u\|^2_{L^2(\Omega)}+C_1\|f\|_{H^{-1}(\Omega)}\|u\|_{H^1(\Omega)}, \tag{634}$$

where C_1 is the operator norm of f. We now apply the inequality $ab \leq 1/2(a^2 + b^2)$ with

$$a = \frac{2\eta}{\sqrt{\lambda}}\|u\|_{L^2(\Omega)} \ \text{ and } \ b = \sqrt{\lambda}\|Du\|_{L^2(\Omega)} \tag{635}$$

to obtain

$$\lambda\|Du\|^2_{L^2(\Omega)}\leq \frac{2\eta^2}{\lambda}\|u\|^2_{L^2(\Omega)}+\frac{\lambda}{2}\|Du\|^2_{L^2(\Omega)}+\eta\|u\|^2_{L^2(\Omega)}+C_1\|f\|_{H^{-1}(\Omega)}\|u\|_{H^1(\Omega)}. \tag{636}$$

We rearrange (636) as

$$\frac{\lambda}{2}\|Du\|^2_{L^2(\Omega)}\leq \frac{2\eta^2}{\lambda}\|u\|^2_{L^2(\Omega)}+\eta\|u\|^2_{L^2(\Omega)}+C_1\|f\|_{H^{-1}(\Omega)}\|u\|_{H^1(\Omega)} \tag{637}$$

and use Poincaré's inequality to conclude that there exists a constant $C_2 > 0$ such that

$$\|u\|^2_{H^1(\Omega)}\leq C_2\left(\|u\|^2_{L^2(\Omega)}+\|f\|_{H^{-1}(\Omega)}\|u\|_{H^1(\Omega)}.\right) \tag{638}$$

whenever u is a weak solution of (603). We now apply the inequality $ab \leq 1/2(a^2 + b^2)$ with

$$a = \sqrt{C_2}\|f\|_{H^{-1}(\Omega)} \ \text{ and } \ b = \frac{1}{\sqrt{C_2}}\|u\|_{H^1(\Omega)} \tag{639}$$

to see that

$$\|u\|^2_{H^1(\Omega)}\leq C_2\left(\|u\|^2_{L^2(\Omega)}+\frac{C_2}{2}\|f\|^2_{H^{-1}(\Omega)}+\frac{1}{2C_2}\|u\|^2_{H^1(\Omega)}\right). \tag{640}$$

It follows that there exists $C > 0$ such that

$$\|u\|_{H^1(\Omega)}\leq C\left(\|u\|_{L^2(\Omega)}+\|f\|_{H^{-1}(\Omega)}\right) \tag{641}$$

whenever u is a weak solution of (603). We summarize our conclusions as follows:

Theorem 97. *Suppose that Ω is a bounded open set in $\mathbb{R}^n$, that the operator $L : H^1_0(\Omega) \to H^{-1}(\Omega)$ defined via*

$$L[u](x) = -D_i\left(a^{ij}(x)D_ju(x) + b^i(x)u(x)\right) + c^i(x)D_iu(x) + d(x)u(x) \tag{642}$$

is strongly elliptic on Ω with coefficients in $L^\infty(\Omega)$. Then the operator $L^* : H^1_0(\Omega) \to H^{-1}(\Omega)$ defined via

$$L^*[u](x) = -D_i\left(a^{ji}(x)D_ju(x) + c^i(x)u\right) + b^i(x)D_iu(x) + d(x)u(x) \tag{643}$$

is the adjoint of L. Moreover, if $f \in H^{-1}(\Omega)$ and $v_1^*, \ldots, v_p^*$ is a basis in the space of weak solutions of the problem

$$\begin{cases} L^*[v](x) = 0 & \text{in } \Omega \\ v(x) = 0 & \text{on } \partial\Omega, \end{cases} \tag{644}$$

then the elliptic boundary value problem

$$\begin{cases} L[u](x) = f & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega, \end{cases} \tag{645}$$

has a weak solution if and only if

$$f(v_l^*) = 0 \quad \text{for all } \; l = 1, \ldots, p. \tag{646}$$

Moreover, there exist constants $C_1 > 0$ and $C_2 > 0$ such that

$$\|u\|_{H_0^1(\Omega)/\ker(L)} \leq C_1 \|f\|_{H^{-1}(\Omega)} \tag{647}$$

and

$$\|u\|_{H_1(\Omega)} \leq C_2 \left(\|f\|_{H^{-1}(\Omega)} + \|u\|_{L^2(\Omega)} \right) \tag{648}$$

whenever u is a weak solution of (645).

We now suppose that Ω is a Lipschitz domain and consider the inhomogeneous problem

$$\begin{cases} L[u](x) = f & \text{in } \Omega \\ u(x) = g(x) & \text{on } \partial\Omega, \end{cases} \tag{649}$$

where $f \in H^{-1}(\Omega)$ and $g \in H^{1/2}(\partial\Omega)$. We reduce (649) to the homogenous problem

$$\begin{cases} L[v](x) = f - L[\varphi] & \text{in } \Omega \\ v(x) = 0 & \text{on } \partial\Omega \end{cases} \tag{650}$$

as before — that is, by letting $\mathscr{T}^{-1}$ denote a continuous right inverse of the trace operator and setting $\varphi = \mathscr{T}^{-1}[g]$. Note that, as before, there is no difficulty in defining $f - L[\varphi] \in H^{-1}(\Omega)$ even though φ is not necessarily in $H_0^1(\Omega)$. From our previous discussion, we see that (650) admits a solution if and only if

$$\langle f - L[\varphi], v_l^* \rangle = 0 \quad \text{for all } \; l = 1, \ldots, p. \tag{651}$$

We note that $\langle L[\varphi], v_l^* \rangle$ need not be equal to $\langle \varphi, L^*[v_l^*] \rangle$ when $\varphi \notin H_0^1(\Omega)$. In fact, under additional regularity conditions, (651) is equivalent to

$$\langle f, v_l^* \rangle = \int_{\partial\Omega} a^{ij} \nu_j D_i v_l^*(x) g(x) \, dS(x) \quad \text{for all } \; l = 1, \ldots, p. \tag{652}$$

where ν_j denotes the j^{th} component of the outward-pointing unit normal vector. To establish this, we will need to use a regularity result we will prove in Chapter 5. It implies that the weak solutions $v_1^*, \ldots, v_p^*$ of the adjoint boundary value problem

$$\begin{cases} L^*[v] = 0 & \text{in } \Omega \\ v = 0 & \text{on } \partial\Omega \end{cases} \tag{653}$$

are elements of $H^2_{\text{loc}}(\Omega) \cap H^1_0(\Omega)$, provided that a^{ij} and b^i are Lipschitz continuous. We will assume also the the coefficients c^i are Lipschitz continuous. Since $L^*[v_l^*] = 0$,

$$\begin{aligned}\int_\Omega &\big(a^{ji}(x)D_jv_l^*(x)D_iu(x) + c^i(x)v_l^*(x)D_iu(x) + \\ &b^i(x)D_iv_l^*(x)u(x) + d(x)u(x)v_l^*(x)\big) \ dx = 0\end{aligned} \tag{654}$$

for all $u \in C_c^\infty(\Omega)$. Since v_l^* is in $H^2_{\text{loc}}(\Omega)$ and a^{ij}, c^i are Lipschitz continuous and we can integrate by parts in (654) to see that

$$\int_\Omega \left(-D_i\left(a^{ji}(x)D_jv_l^*(x)\right) - D_i\left(c^i(x)v_l^*(x)\right) + b^i(x)D_iv_l^*(x) + d(x)v_l^*(x)\right)u(x)\ dx = 0 \tag{655}$$

for all $u \in C_c^\infty(\Omega)$. We conclude that

$$-D_i\left(a^{ji}(x)D_jv_l^*(x)\right) - D_i\left(c^i(x)v_l^*(x)\right) + b^i(x)D_iv_l^*(x) + d(x)v_l^*(x) = 0 \tag{656}$$

for almost all $x \in \Omega$. It follows from (656) that

$$\int_\Omega \left(-D_i\left(a^{ji}(x)D_jv_l^*(x)\right) - D_i\left(c^i(x)v_l^*(x)\right) + b^i(x)D_iv_l^*(x) + d(x)v_l^*(x)\right)\varphi(x)\ dx = 0. \tag{657}$$

Now integrating by parts in (657) yields

$$\begin{aligned}&\int_\Omega \left(a^{ji}(x)D_jv_l^*(x)D_i\varphi(x) + c^i(x)v_l^*(x)D_i\varphi(x) + b^i(x)D_iv_l^*(x)\varphi(x) + d(x)v_l^*(x)\varphi(x)\right)\ dx \\ &= \int_{\partial\Omega} a^{ji}(x)\nu_iD_jv_l^*(x)g(x)\ dS(x),\end{aligned} \tag{658}$$

where ν_i denotes the i^{th} component of the outward-pointing unit normal vector. Note that we have made use of the fact that the trace of v_l^* is 0, and that the trace of φ is g. It follows from (658) that

$$\begin{aligned}\langle L[\varphi], v_l^*\rangle &= \int_\Omega \left(a^{ij}D_j\varphi(x)D_iv_l^*(x) + b_i\varphi(x)D_iv_l^*(x) + c^iD_i\varphi(x)v_l^*(x) + d\varphi(x)v_l^*(x)\right)\ dx \\ &= \int_{\partial\Omega} a^{ji}(x)\nu_iD_jv_l^*(x)g(x)\ dS(x).\end{aligned} \tag{659}$$

That (651) is equivalent to (652) follows immediately from (659). We summarize our conclusions as follows:

THEOREM 98. *Suppose that Ω is a bounded Lipschitz domain in $\mathbb{R}^n$, that the operator $L : H^1(\Omega) \to H^{-1}(\Omega)$ defined via*

$$L[u](x) = -D_i\left(a^{ij}(x)D_ju + b^i(x)u\right) + c^i(x)D_iu + d(x)u(x) \tag{660}$$

is strongly elliptic on Ω, that a^{ij}, b^i and c^i are Lipschitz continuous, and that $d \in L^\infty(\Omega)$. Suppose also that $f \in H^{-1}(\Omega)$, $g \in H^{1/2}(\partial\Omega)$ and $v_1^*, \ldots, v_p^*$ is a basis in the space of weak solutions of the boundary value problem

$$\begin{cases} L^*[v](x) = 0 & \text{in } \Omega \\ v(x) = 0 & \text{on } \partial\Omega. \end{cases} \tag{661}$$

Then the boundary value problem

$$\begin{cases} L\left[u\right](x) = f & \text{in } \Omega \\ u(x) = g(x) & \text{on } \partial\Omega, \end{cases} \tag{662}$$

has a solution if and only if

$$\langle f, v_l^* \rangle = \int_{\partial\Omega} a^{ij}\nu_j D_i v_l^*(x) g(x)\, dS(x), \quad \text{for all } l = 1, \ldots, p. \tag{663}$$

Moreover, there exist constants $C_1 > 0$ and $C_2 > 0$ such that

$$\|u\|_{H^1(\Omega)/\ker(L)} \leq C_1 \left(\|f\|_{H^{-1}(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)} \right) \tag{664}$$

and

$$\|u\|_{H^1(\Omega)} \leq C_2 \left(\|f\|_{H^{-1}(\Omega)} + \|u\|_{L^2(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)} \right) \tag{665}$$

whenever u is a weak solution of (662).

Exercise 31. *Suppose that Ω is the unit disk in $\mathbb{R}^2$. Show that there exist $\lambda \in \mathbb{R}$ such that the boundary value problem*

$$\begin{cases} \Delta u(x) + \lambda^2 u(x) = 0 & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega \end{cases} \tag{666}$$

admits nontrivial (classical) solutions $u \in C^2(\Omega)$.

4.4. The Weak Maximum Principle

The weak maximum principle can be used to show that the boundary value problem

$$\begin{cases} L\left[u\right] = 0 & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{667}$$

has a unique solution under certain conditions on the operator L. In order to state it, a new definition is required. Suppose that $u \in H^1(\Omega)$. Then we say that $u \leq 0$ on $\partial\Omega$ provided the function

$$u^+ = \max\{u, 0\} \tag{668}$$

is an element of $H_0^1(\Omega)$. Similarly, we say that $u \leq r$ on Ω provided $(u - r) \leq 0$ on $\partial\Omega$ and we define the supremem of u on $\partial\Omega$ as follows

$$\sup_{x\in\partial\Omega} u(x) = \inf\{r \in \mathbb{R} : u \leq r \text{ on } \partial\Omega\}. \tag{669}$$

$$\tag{670}$$

Theorem 99. *Suppose that Ω is a bounded open set in $\mathbb{R}^n$, that the operator $L : H^1(\Omega) \to H^{-1}(\Omega)$ defined via*

$$L\left[u\right] = -D_i \left(a^{ij}(x) D_j u + b^i(x) u \right) + c^i(x) D_i u + d(x) u \tag{671}$$

is strongly elliptic on Ω with coefficients in $L^\infty(\Omega)$. Suppose also that that $u \in H^1(\Omega)$ such that $L\left[u\right] \leq 0$ in the sense that

$$\langle L[u], v \rangle \leq 0 \tag{672}$$

for all $v \geq 0$ in $C_c^1(\Omega)$, and that the coefficients d and c^i are such that

$$\int_\Omega b^i(x) D_i v(x) + d(x) v(x) \; dx \geq 0 \tag{673}$$

for all $v \geq 0$ in $C_c^1(\Omega)$. Then

$$u(x) \leq \sup_{x \in \partial\Omega} u^+(x) \tag{674}$$

for almost all $x \in \Omega$.

Proof. We observe that if $u \in H^1(\Omega)$ and $v \in H_0^1(\Omega)$, then $uv \in W_0^{1,1}(\Omega)$ and

$$D(uv) = vDu + uDv. \tag{675}$$

We now manipulate (672) in order to see that

$$\begin{aligned} &\int_\Omega a^{ij}(x) D_j u(x) D_i v(x) + b^i(x) u(x) D_i v(x) + c^i(x) v(x) D_i u(x) \; dx \\ \leq &-\int_\Omega d(x) u(x) v(x) dx \end{aligned} \tag{676}$$

for all $v \in H_0^1(\Omega)$ such that $v \geq 0$. By subtracting

$$\int_\Omega b^i(x) D_i(uv)(x) \; dx = \int_\Omega b^i(x) v(x) D_i u(x) + b^i(x) u(x) D_i v(x) \; dx \tag{677}$$

from both sides of (676) and invoking (673) we conclude that

$$\begin{aligned} &\int_\Omega a^{ij}(x) D_j u(x) D_i v(x) + (c^i(x) - b^i(x)) v(x) D_i u(x) \; dx \\ \leq &-\int_\Omega \left(d(x) u(x) v(x) + b^i(x) D_i(uv)(x) \right) \; dx \leq 0 \end{aligned} \tag{678}$$

for all v in $H_0^1(\Omega)$ such that $uv \geq 0$. Since the coefficients of L are bounded, it follows from (678) that there exists $C > 0$ such that

$$\int_\Omega a^{ij}(x) D_j u(x) D_i v(x) \; dx \leq C \int_\Omega v(x) |Du(x)| \; dx \tag{679}$$

for all $v \geq 0$ in $H_0^1(\Omega)$ such that $uv \geq 0$.

Now we suppose that $r > 0$ is such that

$$\sup_{x \in \partial\Omega} u(x) < r < \|u\|_{L^\infty(\Omega)}, \tag{680}$$

and we set $v = \max\{u - r, 0\}$. Then $v \in H_0^1(\Omega)$ and $uv \geq 0$ since $v(x) = 0$ for any $u(x) < 0$. Moreover, v is weakly differentiable and

$$Dv(x) = \begin{cases} Du(x) & u(x) > r \\ 0 & u(x) \leq r. \end{cases} \tag{681}$$

We let Γ denote the support of Dv and note that $\|v\|_{L^2(\Gamma)}$ must be positive since, otherwise, $u \leq r$ almost everywhere, which contradicts (680).

From (679) we conclude that

$$\int_\Omega a^{ij}(x) D_j u(x) D_i v(x)\ dx \leq C \int_\Omega v(x) |Du(x)|\ dx \tag{682}$$

for some constant C which does not depend on r. Since $Dv(x) = Du(x)$ for any x in Γ, (682) implies that

$$\int_\Omega a^{ij}(x) D_j v(x) D_i v(x)\ dx \leq C \int_\Gamma v(x) |Dv(x)|\ dx. \tag{683}$$

From (683) and the strong ellipiticty of L, we see that there exists $C' > 0$ which does not depend on r and such that

$$\|Dv\|^2_{L^2(\Omega)} \leq C' \int_\Gamma v(x) |Dv(x)|\ dx \leq C' \|v\|_{L^2(\Gamma)} \|Dv\|_{L^2(\Omega)}. \tag{684}$$

If $\|Dv\|_{L^2(\Omega)} > 0$, then we divide both sides of (684) by $\|Dv\|_{L^2(\Omega)}$ to obtain

$$\|Dv\|_{L^2(\Omega)} \leq C' \|v\|_{L^2(\Gamma)}. \tag{685}$$

If $\|Dv\|_{L^2(\Omega)} = 0$, then (685) automatically holds. In the case when $n \geq 3$, the Sobolev conjugate p^* of 2 satisfies

$$\frac{1}{p^*} = \frac{1}{2} - \frac{1}{n} = \frac{n-2}{2n}. \tag{686}$$

We apply the Sobolev imbedding theorem to conclude that

$$\|v\|_{L^{2n/(n-2)}(\Omega)} \leq \|Dv\|_{L^2(\Omega)} \leq C' \|v\|_{L^2(\Gamma)}. \tag{687}$$

Now we observe that

$$\frac{2}{n} + \frac{n-2}{n} = 1 \tag{688}$$

and invoke Hölder's inequality with $p = n/2$ and $q = n/(n-2)$ in order to obtain

$$\int_\Gamma v(x)^2\ dx \leq \left(\int_\Gamma |v|^{\frac{2n}{n-2}}\ dx \right)^{\frac{n-2}{n}} |\Gamma|^{2/n} \leq \left(\int_\Omega |v|^{\frac{2n}{n-2}}\ dx \right)^{\frac{n-2}{n}} |\Gamma|^{2/n} \tag{689}$$

so that

$$\|v\|_{L^2(\Gamma)} \leq \|v\|_{L^{2n/(n-2)}(\Omega)} |\Gamma|^{1/n}. \tag{690}$$

We combine (690) and (687) to obtain

$$\|v\|_{L^2(\Gamma)} \leq |\Gamma|^{1/n} \|v\|_{L^{2n/(n-2)}(\Omega)} \leq |\Gamma|^{1/n} C' \|v\|_{L^2(\Gamma)}. \tag{691}$$

From this and the fact that $\|v\|_{L^2(\Gamma)} \neq 0$, we see that

$$|\Gamma| \geq (C')^{-n} \tag{692}$$

In particular, since the support of Dv is contained in the set $\{x : u(x) > r\}$ and $Du = Dv$ there, it must be the case that the set $\{x : u(x) > r\}$ contains a subset of measure greater than or equal to $(C')^{-n}$ on which $Du \neq 0$.

For all $m > 0$, we let

$$\Omega_m = \left\{ x \in \Omega : u(x) \geq \|u\|_\infty - \frac{1}{m} \right\} \bigcap \operatorname{supp}(Du) \tag{693}$$

and we set

$$\Omega' = \bigcap_{m>0} \Omega_m. \tag{694}$$

Since u obtains its essential supremum on Ω', and is therefore constant on that set, $Du = 0$ almost everywhere on Ω'. In particular, Ω' must be of measure 0. But the Ω_m are a sequence of decreasing sets of finite measure, so

$$|\Omega'| = \lim_{m\to\infty} |\Omega_m| \geq (C')^{-n}. \tag{695}$$

We conclude that $u \leq \sup_{\partial\Omega} u^+$ from this contradiction

We leave the modifications necessary to establish the result in the event that $n = 1$ or $n = 2$ to the reader. □

The following follows easily from Theorem 99.

THEOREM 100. *Suppose that Ω is a bounded open set in $\mathbb{R}^n$, that the operator $L : H^1(\Omega) \to H^{-1}(\Omega)$ defined via*

$$L[u] = -D_i\left(a^{ij}(x)D_j u + b^i(x)u\right) + c^i(x)D_i u + d(x)u \tag{696}$$

is strongly elliptic on Ω with coefficients in $L^\infty(\Omega)$. Suppose also that that $u \in H^1(\Omega)$ such that $L[u] \geq 0$ in the sense that

$$\langle L[u], v\rangle \geq 0 \tag{697}$$

for all $v \geq 0$ in $C^1_c(\Omega)$, and that the coefficients d and c^i are such that

$$\int_\Omega b^i(x)D_i v(x) + d(x)v(x)\ dx \geq 0 \tag{698}$$

for all $v \geq 0$ in $C^1_c(\Omega)$. Then

$$u(x) \geq \inf_{x\in\partial\Omega} u^-(x) \tag{699}$$

for almost all $x \in \Omega$.

The following theorem is an immediate consequence of Theorems 99 and 100.

THEOREM 101. *Suppose that Ω is a bounded open set in $\mathbb{R}^n$, that*

$$L[u] = -D_i\left(a^{ij}(x)D_j u(x) + b^i(x)u(x)\right) + c^i(x)D_i u(x) + d(x)u(x) \tag{700}$$

is a strongly elliptic operator $H^1_0(\Omega) \to H^{-1}(\Omega)$ with essentially bounded coefficients, and that

$$\int_\Omega b^i(x)D_i v(x) + d(x)v(x)\ dx \geq 0 \tag{701}$$

for all $v \in H^1_0(\Omega)$ such that $v \geq 0$. Then the boundary value problem

$$\begin{cases} L[u](x) = 0 & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega \end{cases} \tag{702}$$

admits only the trivial solution $u = 0$. This is the case, in particular, if $b^i = 0$ for all i and $d \geq 0$.

We are now prepared to state our main result regarding strongly elliptic second order operators which satisfy the hypotheses of the weak maximal theorem.

THEOREM 102. *Suppose that* Ω *is a Lipschitz domain in* $\mathbb{R}^n$, *that*

$$L\left[u\right](x) = -D_i\left(a^{ij}(x)D_j u(x) + b^i(x)u(x)\right) + c^i(x)D_i u(x) + d(x)u(x) \tag{703}$$

is strongly elliptic on Ω *with coefficients in* $L^\infty(\Omega)$. *Suppose also that*

$$\int_\Omega b^i(x)D_i v(x) + d(x)v(x)\ dx \geq 0 \tag{704}$$

for all $v \in H_0^1(\Omega)$ *such that* $v \geq 0$. *This is the case, in particular, if* L *is of the form*

$$L\left[u\right](x) = -D_i\left(a^{ij}(x)D_j u(x)\right) + c^i(x)D_i u(x) + d(x)u(x) \tag{705}$$

with $d(x) \geq 0$ *for almost all* x *in* Ω. *Then for each* $f \in H^{-1}(\Omega)$ *and* $g \in H^{1/2}(\partial\Omega)$ *the boundary value problem*

$$\begin{cases} L\left[u\right](x) = f(x) & \text{in } \Omega \\ u(x) = g & \text{on } \partial\Omega \end{cases} \tag{706}$$

admits a unique weak solution u. *Moreover, there exists a constant* $C > 0$ *such that*

$$\|u\|_{H^1(\Omega)} \leq C\left(\|f\|_{H^{-1}(\Omega)} + \|g\|_{H^{1/2}(\partial\Omega)}\right). \tag{707}$$

whenever u *is the weak solution of (706) In other words the operator*

$$L \oplus \mathscr{T} : H^1(\Omega) \to H^{-1}(\Omega) \oplus H^{1/2}(\partial\Omega) \tag{708}$$

is an isomorphism.

PROOF. The trace operator $\mathscr{T} : H^1(\Omega) \to H^{1/2}(\partial\Omega)$ has a continuous right inverse, which we denote by $\mathscr{T}^{-1}$ (this choice is not unique). We let $C_1 > 0$ be such that

$$\|\mathscr{T}^{-1}[h]\|_{H^1(\Omega)} \leq C_1 \|h\|_{H^{1/2}(\partial\Omega)} \tag{709}$$

for all $h \in H^{1/2}(\partial\Omega)$.

We will now analyze the boundary value problem

$$\begin{cases} L\left[w\right](x) = \tilde{f}(x) & \text{in } \Omega \\ w(x) = 0 & \text{on } \partial\Omega, \end{cases} \tag{710}$$

According to Theorem 96, there exists $\sigma > 0$ such that $K = L + \sigma I : H_0^1(\Omega) \to H^{-1}(\Omega)$ is coercive and hence invertible. It follows that $L = K - \sigma I$ is Fredholm of index 0 since σI is compact by Theorem 95. Theorem 101 implies that the dimension of the kernel of L is 0; it follows that the dimension of the cokernel of $\mathrm{Im}(L)$ is 0. In particular, the exists a unique solution w of (710) for any given $\tilde{f}$. Consequently, we can view the operator L as defining a continuous bijective mapping $H_0^1(\Omega) \to H^{-1}(\Omega)$. By the bounded inverse theorem, this bijection has a continuous inverse. That is, there exists a constant C_2 such that

$$\|w\|_{H_0^1(\Omega)} \leq C_2 \left\|\tilde{f}\right\|_{H^{-1}(\Omega)} \tag{711}$$

whenever w is the solution of (710).

We now let $g \in H^{1/2}(\partial\Omega)$ and $f \in H^{-1}(\Omega)$ be given. We let $\varphi = \mathscr{T}^{-1}[g]$ so that

$$\|\varphi\|_{H^1(\Omega)} \leq C_1 \|g\|_{H^{1/2}(\partial\Omega)}, \tag{712}$$

and we let $\tilde{f} = f - L[\varphi]$. Obviously,

$$\left\|\tilde{f}\right\|_{H^{-1}(\Omega)} \leq \|f\|_{H^{-1}(\Omega)} + \|L\| \|\varphi\|_{H^{-1}(\Omega)} \leq \|f\|_{H^{-1}(\Omega)} + C_1 \|L\| \|g\|_{H_{1/2}(\partial\Omega)}. \tag{713}$$

Now we let w be the solution of (710). Then $w + \varphi$ solves (706) and

$$\begin{aligned} \|w+\varphi\|_{H^1(\Omega)} &\leq \|w\|_{H^1(\Omega)} + \|\varphi\|_{H^{-1}(\Omega)} \\ &\leq C_2 \left\|\tilde{f}\right\|_{H^{-1}(\Omega)} + C_1 \|g\|_{H^{1/2}(\partial\Omega)} \\ &\leq C_2 \left(\|f\|_{H^{-1}(\Omega)} + C_1 \|L\| \|g\|_{H_{1/2}(\partial\Omega)}\right) + C_1 \|g\|_{H^{1/2}(\partial\Omega)} \\ &\leq C \left(\|f\|_{H^{-1}(\Omega)} + \|g\|_{H_{1/2}(\partial\Omega)}\right), \end{aligned} \tag{714}$$

where $C = C_2 + C_1 (C_2 \|L\| + 1)$. We observe that this constant C does not depend on f or g. Among other things, the bound (714) implies that the solution of (706) is unique. □

4.5. The Neumann Problem

We now develop a weak formulation of the Neumann boundary value problem for an operator of the form

$$L[u](x) = -D_i \left(a^{ij}(x) D_j u(x) + b^i(x)\right) + c^i(x) D_i u(x) + d(x) u(x). \tag{715}$$

We assume, as usual, that L is strongly elliptic, and that the coefficients of L are bounded measurable functions. In this section, we view L as a mapping $H^1(\Omega) \to (H^1(\Omega))^*$. Moreover, we will assume that Ω is a bounded, connected Lipschitz domain. Among other things, this implies that $H^1(\Omega)$ is compactly contained in $L^2(\Omega)$.

We take f in to be an element of $(H^1(\Omega))^*$ and $g \in \left(H^{1/2}(\partial\Omega)\right)^*$. We say that $u \in H^1(\Omega)$ is a weak solution of the Neumann problem for L provided

$$\begin{aligned} &\int_\Omega \left(a^{ij} \; D_j u(x) D_i v(x) + b^i(x) u(x) D_i v(x) + c^i(x) D_i u(x) v(x) + d(x) u(x) v(x)\right) \, dx \\ &\qquad = \langle f, v\rangle + \langle g, \mathscr{T}(v)\rangle \end{aligned} \tag{716}$$

for all $v \in H^1(\Omega)$. Given f and g, we let $F_{f,g}$ denote that bounded linear functional $H^1(\Omega) \to \mathbb{R}$ defined via

$$\langle F_{f,g}, v\rangle = \langle f, v\rangle + \langle g, \mathscr{T}(v)\rangle. \tag{717}$$

Then (716) is equivalent to requiring that $L[u] = F_{f,g}$. In the (fairly typical) case that $f \in L^2(\Omega)$ and $g \in L^2(\partial\Omega)$, (716) becomes

$$\begin{aligned} &\int_\Omega \left(a^{ij} D_j u(x) D_i v(x) + b^i(x) u(x) D_i v(x) + c^i(x) D_i u(x) v(x) + d(x) u(x) v(x)\right) \, dx \\ &= \int_\Omega f(x) v(x) \, dx + \int_{\partial\Omega} g(x) \mathscr{T}[v](x) \, dS(x). \end{aligned} \tag{718}$$

To understand the nature of the boundary conditions imposed by (716), we will momentarily assume that $u \in H^2(\Omega)$ which satisfies (718), and that the coefficients a^{ij}, b^i are Lipschitz continuous. Then we can integrate by parts in (718) to see that

$$\int_\Omega \left(-D_i\left(a^{ij}D_ju(x)+b_i(x)u(x)\right)v(x)+c^i(x)D_iu(x)v(x)+d(x)u(x)v(x)\right)\ dx \tag{719}$$

$$+\int_{\partial\Omega}\left(a^{ij}\nu_iD_ju(x)+b^i\nu_iu(x)\right)v(x)\ dS(x)=\int_\Omega f(x)v(x)\ dx+\int_{\partial\Omega}g(x)v(x)\ dS(x)$$

for all $v \in C^\infty\left(\overline{\Omega}\right)$. If $v \in C_c^\infty(\Omega)$, then the boundary terms in (719) vanish, leaving us with

$$\int_\Omega \left(-D_i\left(a^{ij}D_ju(x)+b^iu(x)\right)v(x)+c^i(x)D_iu(x)v(x)+d(x)u(x)v(x)\ \right)dx \tag{720}$$

$$=\int_\Omega f(x)v(x)\ dx.$$

We conclude that

$$-D_i(a^{ij}D_ju(x)+b^i\nu_iu(x))+c^i(x)D_iu(x)+d(x)u(x)=f(x) \tag{721}$$

almost everywhere in Ω. By combining (721) and (719), we see that (718) implies that

$$\int_{\partial\Omega}\left(a^{ij}\nu_iD_ju(x)+b^i\nu_iu(x)\right)v(x)\ dS(x)=\int_{\partial\Omega}g(x)v(x)\ dS(x) \tag{722}$$

for all $v \in C^\infty\left(\overline{\Omega}\right)$. Clearly, (722) implies that

$$a^{ij}(x)\nu_iD_ju(x)+b^i(x)\nu_iu(x)=g(x) \tag{723}$$

for almost all $x \in \partial\Omega$. In light of (721) and (723), it is reasonable to call (716) a weak formulation of the boundary value problem

$$\begin{cases} L[u](x)=f(x) \text{ in } \Omega \\ a^{ij}\nu_iD_ju(x)+b^i\nu_iu(x)=g(x) \text{ on } \partial\Omega, \end{cases} \tag{724}$$

We can now proceed just as we did in the case of the Dirichlet problem. The proof that L is Fredholm of index 0 is essentially identical. The adjoint $L^*: H^1(\Omega)\to\left(H^1(\Omega)\right)^*$ is given by

$$L^*[u](x)=-D_i\left(a^{ji}(x)D_ju(x)+c^i(x)u(x)\right)+b^i(x)D_iu(x)+d(x)u(x). \tag{725}$$

If we let $v_1^*,\ldots,v_p^*$ denote a basis for $\ker(L^*)$, then (724) is solvable if and only if

$$F_{f,g}\left(v_l^*\right)=0 \text{ for all } l=1,\ldots,p. \tag{726}$$

We note that (726) imposes conditions on both f and g and it is equivalent to

$$\langle f,v_l^*\rangle+\langle g,\mathscr{T}\left[v_l^*\right]\rangle=0 \text{ for all } l=1,\ldots,p. \tag{727}$$

If $v_1, \ldots, v_p$ is a basis for $\ker(L)$ and u is any weak solution of (724), then the set of solutions of (724) consists of any function of the form

$$u + \sum_{;=1}^{n} a_; v_l. \tag{728}$$

Moreover, since $L : H^1(\Omega)/\ker(L) \to \Im(L)$ is a continous bijection, the bounded inverse theorem implies that there exists a constant $C > 0$ such that

$$\|u\|_{H^1(\Omega)/\ker(L)} \leq C \|F_{f,g}\|_{(H^1(\Omega))^*} \tag{729}$$

whever u is a weak solution of (724). Moreover, just as in the case of the Dirichlet problem, we can obtain an alternative bound of the form

$$\|u\|_{H^1(\Omega)} \leq C \left(\|F_{f,g}\|_{(H^1(\Omega))^*} + \|u\|_{L^2(\Omega)} \right) \tag{730}$$

on weak solutions of (724).

In the particular case of operators of the form

$$L[u](x) = -D_i \left(a^{ij} D_j u(x) \right), \tag{731}$$

the solvability condition is quite simple. In this case, the adjoint of L is

$$L^*[u](x) = -D_i \left(a^{ji} D_j u(x) \right), \tag{732}$$

We claim the the kernel of both L and L^* consists of the constant functions on Ω. To see this we observe that if $L[u](x) = 0$, then

$$0 = \langle L[u], u \rangle = \int_\Omega a^{ij} D_j u(x) D_i u(x)\ dx \geq \lambda |\Omega| \|Du\|^2_{H^1(\Omega)}, \tag{733}$$

which implies that $Du = 0$. The same argument applies to L^*. If $f \in L^2(\Omega)$ and $g \in L^2(\partial\Omega)$, then (726) is equivalent to

$$\int_\Omega f(x)\ dx + \int_{\partial\Omega} g(x) = 0. \tag{734}$$

Moreover, if u is a particular solution of (724), then every solution of (724) is of the form

$$u(x) + C \tag{735}$$

with C a constant.

4.6. Mixed boundary conditions

Once again, we let Ω be a bounded, connected Lipschitz domain in $\mathbb{R}^n$, and let L be a strongly elliptic operator of the form

$$L[u](x) = -D_i \left(a^{ij} D_j u(x) + b^i(x) u(x) \right) + c^i(x) D_i u(x) + d(x) u(x) \tag{736}$$

whose coefficients are $L^\infty(\Omega)$ functions. In the preceding section, we saw that the space of test functions V in the weak formulation

$$\langle L[u], v \rangle = \langle f, v \rangle \quad \text{for all} \quad v \in V \tag{737}$$

has a strong effect on the boundary conditions of the boundary value problem corresponding to it. In particular, choosing $V = H^1_0(\Omega)$ lead to Dirichlet boundary conditions, while

$V = H^1(\Omega)$ gave us the Neumann problem. In this section, we show that modifying the space of testing functons can give rise to a mixed Dircihlet/Neumann boundary value problem.

Suppose that $\partial\Omega$ is decomposed as

$$\partial\Omega = \Gamma_D \cup \Gamma_N \tag{738}$$

with Γ_D and Γ_N Lipschitz curves such that there exist bounded extension operators $H^{1/2}(\Gamma_D) \to H^{1/2}(\partial\Omega)$ and $H^{1/2}(\Gamma_N) \to H^{1/2}(\partial\Omega)$. We wish to devise a weak formulation for the boundary value problem

$$\begin{cases} L[u](x) = f(x) \ \text{ in } \ \Omega \\ u(x) = g_D(x) \ \text{ on } \ \Gamma_D \\ a^{ij}\nu_i D_j u(x) + b^i \nu_i u(x) = g_N(x) \ \text{ on } \ \Gamma_N. \end{cases} \tag{739}$$

Here, $f \in (H^1(\Omega))^*$, $g_D \in H^{1/2}(\Gamma_D)$, and $g_N \in \left(H^{1/2}(\Gamma_N)\right)^*$. We say that u is a weak solution of (739) provided $\mathscr{T}[u](x) = g_D(x)$ for almost all $x \in \Gamma_D$, and

$$\langle L[u], v\rangle = \langle f, v\rangle + \langle g_N, \mathscr{T}[v]\rangle \tag{740}$$

for all v in the space

$$H_D^1(\Omega) = \left\{v \in H^1(\Omega) : \mathscr{T}[v](x) = 0 \ \text{ for almost all } \ x \in \Gamma_D\right\}. \tag{741}$$

We first reduce (736) to the homogeneous boundary value problem

$$\begin{cases} L[w](x) = \tilde{f}(x) \ \text{ in } \ \Omega \\ w(x) = 0 \ \text{ on } \ \Gamma_D \\ a^{ij}\nu_i D_j w(x) + b^i(x)\nu_i w(x) = g_N(x) \ \text{ on } \ \Gamma_N. \end{cases} \tag{742}$$

To be entirely clear, we say that $w \in H_D^1(\Omega)$ is a weak solution of (742) provided

$$\langle L[u], v\rangle = \langle f, v\rangle + \langle g_N, \mathscr{T}[v]\rangle \tag{743}$$

for all v in the space $H_D^1(\Omega)$. If φ is an element of $H^1(\Omega)$ such that

$$\mathscr{T}(\varphi) = E[g_D], \tag{744}$$

where $E : H^{1/2}(\Gamma_D) \to H^{1/2}(\partial\Omega)$ is an extension operator, and

$$\tilde{f}(x) = f(x) - L[\varphi](x), \tag{745}$$

then is easy to verify that $w + \varphi$ is a weak solution of (736) if and only if w is a weak solution of (742).

It is easy to establish that the operator L is Fredholm of index 0, and the adjoint of L is the operator $L^* : H_D^1(\Omega) \to (H_D^1(\Omega))^*$ defined via

$$L^*[u](x) = -D_i\left(a^{ji}(x)D_j u(x) + c^i(x)u(x)\right) + b^i(x)D_i u(x) + d(x)u(x). \tag{746}$$

It follows that if $v_1^*, \ldots, v_p^*$ is a basis in the kernel of L^*, then (742) is solvable if and only if

$$\langle f, v_l^*\rangle - \langle L[\varphi], v_l^*\rangle + \langle g_N, \mathscr{T}[v_l^*]\rangle = 0 \ \text{ for all } \ l = 1, \ldots, p. \tag{747}$$

The functions $v_1^*, \ldots, v_p^*$ are weak solutions of the boundary value problem

$$\begin{cases} L^*[w](x) = 0 \text{ in } \Omega \\ w(x) = 0 \text{ on } \Gamma_D \\ a^{ji}\nu_i D_j w(x) + c^i(x)\nu_i w(x) = 0 \text{ on } \Gamma_N. \end{cases} \tag{748}$$

As in the case of the Dirichlet problem, we can derive a more satisfying solvability criterion under slightly stronger regularity assumptions. We suppose that a^{ij}, b^i and c^i are Lipschitz continuous. Then, it follows from regularity results of Chapter 5 that the functions $v_1^*, \ldots, v_p^*$ are elements of $H^2(\Omega)$. Integrating by parts, we find that for each $p = 1, \ldots, l$,

$$-D_i\left(a^{ji}(x)D_j v_l^*(x)\right) - D_i\left(c^i(x)v_l^*(x)\right) + b^i(x)D_i v_l^*(x) + d(x)v_l^*(x) = 0 \tag{749}$$

almost everywhere in Ω. Multiplying both sides of (749) by φ, integrating over Ω, and integrating by parts gives

$$\langle L[\varphi], v_l^*\rangle = \int_{\Gamma_D} a^{ji}(x)\nu_i D_j v_l^*(x) g_D(x)\, dS(x). \tag{750}$$

By combining (750) and (747), we see that (736) is solvable if and only if

$$\langle f, v_l^*\rangle + \langle g_N, \mathscr{T}[v_l^*]\rangle_{\Gamma_N} = \int_{\Gamma_D} a^{ji}(x)\nu_i D_j v_l^*(x) g_D(x)\, dS(x) = 0 \text{ for all } l = 1, \ldots, p. \tag{751}$$

It is also straightfoward to show that there exists $C > 0$ such that

$$\|u\|_{H^1(\Omega)} \leq C\left(\|u\|_{L^2(\Omega)} + \|f\|_{H^{-1}(\Omega)} + \|g_D\|_{H^{1/2}(\Gamma_D)} + \|g_N\|_{\left(H^{-1/2}(\Gamma_N)\right)^*}\right) \tag{752}$$

whenever u is a weak solution of (736).

4.7. The Robin problem

One of the nice features of the Fredholm theory is that the addition of a compact operator to L will result in yet another Fredholm operator of index 0. In this section, we exploit this fact to develop a weak formulation of the Robin boundary value problem

$$\begin{cases} L[u](x) = f(x) \text{ for all in } \Omega \\ a^{ij}(x)\nu_i D_j u(x) + b^i(x)\nu_i(x)u(x) + B(x)u(x) = g(x) \text{ on } \partial\Omega. \end{cases} \tag{753}$$

Here, we will assume that Ω is a bounded, connected Lipschitz domain, and that L is a strongly elliptic operator of the form

$$L[u](x) = -D_i\left(a^{ij}(x)D_j u(x) + b^i(x)u(x)\right) + c^i(x)D_i u(x) + d(x)u(x) \tag{754}$$

with $L^\infty(\Omega)$ coefficients. We take $f \in H^{-1}(\Omega)$, $B \in L^\infty(\partial\Omega)$ and $g \in \left(H^{1/2}(\partial\Omega)\right)^*$. We define the operator $K : H^1(\Omega) \to (H^1(\Omega))^*$ via

$$\langle K[u], v\rangle = \int_\Omega B(x)\mathscr{T}[u](x)\mathscr{T}[v](x)\, dS(x). \tag{755}$$

Again, we view L as the operator $H^1(\Omega) \to (H^{-1}(\Omega))^*$ defined via

$$\langle L[u], v\rangle = \int_\Omega \big(a^{ij}(x)D_ju(x)\ D_iu(x) + b^i(x)u(x)D_iv(x) + c^i(x)D_iu(x)v(x) + d(x)u(x)v(x)\big)\ dx. \tag{756}$$

We say that $u \in H^1(\Omega)$ is a weak solution of (753) if

$$\langle L[u] + K[u]\rangle = \langle f, v\rangle + \langle g, \mathscr{T}[v]\rangle \tag{757}$$

for all $v \in H^1(\Omega)$. The operator K is compact since K is bounded $H^{1/2+\epsilon}(\Omega) \to (H^1(\Omega))^*$ for any $\epsilon > 0$ and $H^1(\Omega)$ is compactly embedded in $H^{1/2+\epsilon}(\Omega)$. That is, K can be factored as an operator

$$H^1(\Omega) \to H^{1/2+\epsilon}(\Omega) \to \big(H^1(\Omega)\big)^*. \tag{758}$$

The operator L is Fredholm of index 0 and K is compact, so $L + K$ is Fredholm of index 0. If we let $F_{f,g}$ denote the element of $(H^1(\Omega))^*$ defined via

$$\langle F_{f,g}, v\rangle = \langle f, v\rangle + \langle g, \mathscr{T}[v]\rangle, \tag{759}$$

then (753) admits weak solution if and only if

$$F_{f,g}(v_l^*) = 0 \quad \text{for all} \quad l = 1, \ldots, m, \tag{760}$$

where $v_1^*, \ldots, v_m^*$ is a basis for $\ker(L^* + K^*)$. Moreover, if u is a weak solution of (753), then

$$\|u\|_{H^1(\Omega)/\ker(L+K)} \leq C\left(\|f\|_{(H^1(\Omega))^*} + \|g\|_{\left(H^{1/2}(\partial\Omega)\right)^*}\right). \tag{761}$$

CHAPTER 5

Regularity of Solutions of Elliptic Boundary Value Problems

In the preceding chapter, we studied the existence of weak solutions of elliptic boundary value problems under fairly mild hypotheses. We now study these problems under stronger regularity assumptions. Although we focus on the Dirichlet problem, similar results can be obtained for other boundary value problems.

5.1. Interior Regularity of Solutions of the Dirichlet Problem

Suppose that Ω is a Lipschitz domain in $\mathbb{R}^n$, and that

$$L\left[u\right](x) = -D_i\left(a^{ij}(x)D_ju(x) + b^i(x)u(x)\right) + c^i(x)D_iu(x) + d(x)u(x) \tag{762}$$

is a strongly elliptic partial differential operator with bounded coefficients. In the previous chapter, we showed that the operator

$$L \oplus \mathscr{T} : H^1(\Omega) \to H^{-1}(\Omega) \oplus H^{1/2}(\partial\Omega) \tag{763}$$

associated with the Dirichlet boundary value problem

$$\begin{cases} L\left[u\right](x) = f(x) & \text{in } \Omega \\ \quad u(x) = g(x) & \text{on } \partial\Omega \end{cases} \tag{764}$$

is a Fredholm operator of index 0 and gave sufficient (but not necessary) conditions under which it is an isomorphism. We will now show that when f is in $L^2(\Omega)$ and the coefficients of L are Lipschitz continuous, the solutions of (764) are also more regular on the interior of Ω.

Theorem 103. *Suppose that Ω is a bounded open set in $\mathbb{R}^n$, that*

$$L\left[u\right](x) = -D_i\left(a^{ij}(x)D_ju(x) + b^i(x)u(x)\right) + c^i(x)D_iu(x) + d(x)u(x) \tag{765}$$

is strongly elliptic in Ω, that $a^{ij}, b^i \in C^{0,1}\left(\overline{\Omega}\right)$, and that $c^i, d \in L^\infty(\Omega)$. Suppose also that $f \in L^2(\Omega)$, and that $u \in H^1(\Omega)$ is a weak solution of the problem $L\left[u\right](x) = f(x)$ — that is,

$$\begin{aligned}&\int_\Omega a^{ij}(x)D_ju(x)D_iv(x) + b^i(x)u(x)D_iv(x) + c^i(x)v(x)D_iu(x) + d(x)u(x)v(x)\ dx \\ &= \int_\Omega v(x)f(x)\ dx\end{aligned} \tag{766}$$

for all $v \in H^1_0(\Omega)$. Then $u \in H^2_{loc}(\Omega)$ and u satisfies the equation

$$-a^{ij}(x)D_iD_ju(x) + (-D_ja^{ji}(x) + c^i(x) - b^i(x))D_iu(x) + \left(d(x) - D_ib^i(x)\right)u(x) = f(x) \tag{767}$$

for almost all $x \in \Omega$. Moreover, for each open set $\Omega' \subset\subset \Omega$, there exists a real number $C > 0$ such that

$$\|u\|_{H^2(\Omega')} \leq C \left(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right). \tag{768}$$

Proof. We define the function $g \in L^2(\Omega)$ via the formula

$$g(x) = f(x) + (b^i(x) - c^i(x))D_i u(x) + (D_i b^i(x) - d(x))u(x) \tag{769}$$

so that

$$\int_\Omega a^{ij}(x) D_j u(x) D_i v(x)\ dx = \int_\Omega g(x) v(x)\ dx \tag{770}$$

for all $v \in H_0^1(\Omega)$. We now assume that v is compactly supported in $H_0^1(\Omega)$. Then we may replace v with $\Delta_k^{-h} v$ in (770) to obtain

$$\begin{aligned} \int_\Omega \Delta_k^h \left(a^{ij}(x) D_j u(x)\right) D_i v(x)\ dx &= -\int_\Omega a^{ij}(x) D_j u(x) D_i \left(\Delta_k^{-h} v(x)\right)\ dx \\ &= -\int_\Omega g(x) \Delta_k^{-h} v(x)\ dx. \end{aligned} \tag{771}$$

Note that difference quotients commute with differential operators. We observe that

$$\begin{aligned} \Delta_k^h \left(a^{ij} D_j u\right)(x) &= \frac{a^{ij}(x + he_k) D_j u(x + he_k) - a^{ij}(x) D_j u(x)}{h} \\ &= a^{ij}(x + he_k) \Delta_k^h D_j u(x) + \left(\Delta_k^h a^{ij}(x)\right) D_j u(x). \end{aligned} \tag{772}$$

Inserting (772) into (771) yields

$$\int_\Omega \left(a^{ij}(x + he_k) D_j \Delta_k^h u(x) + \left(\Delta_k^h a^{ij}(x)\right) D_j u(x)\right) D_i v(x)\ dx = -\int_\Omega g(x) \Delta_k^{-h} v(x)\ dx. \tag{773}$$

We let

$$\overline{g} = \begin{pmatrix} \left(\Delta_k^h a^{1j}(x)\right) D_j u(x) \\ \left(\Delta_k^h a^{2j}(x)\right) D_j u(x) \\ \vdots \\ \left(\Delta_k^h a^{nj}(x)\right) D_j u(x) \end{pmatrix} \tag{774}$$

and rearrange (773) as

$$\int_\Omega a^{ij}(x + he_k) D_j \Delta_k^h u(x) D_i v(x)\ dx = -\int_\Omega \overline{g} \cdot Dv(x) + g(x) \Delta_k^{-h} v(x)\ dx. \tag{775}$$

We apply Hölder's inequality to obtain

$$\begin{aligned} \int_\Omega a^{ij}(x + he_k) D_j \Delta_k^h u(x) D_i v(x)\ dx &\leq \left(\|\overline{g}\|_2 + \|g\|_2\right) \|Dv\|_2 \\ &\leq C \left(\|u\|_{H^1(\Omega)} + \|f\|_{L^2(\Omega)}\right) \|Dv\|_2. \end{aligned} \tag{776}$$

We note that the preceding argument applies when $v \in H_0^1(\Omega)$ which is not compactly supported in Ω as long as the difference quotients $\Delta_k^h v$ and those in (774) are defined. We will make use of this fact in the following section.

Now we suppose that $\Omega' \subset\subset \Omega$ and let $\eta \in C_c^\infty(\Omega)$ such that $\eta(x) = 1$ for all $x \in \Omega'$ and $0 \le \eta \le 1$. Moreover, we define v via the formula

$$v(x) = \eta^2(x)\Delta_k^h u(x). \tag{777}$$

The strong ellipticy of L implies that

$$\lambda \int_\Omega \eta^2(x) \left|D\Delta_k^h u(x)\right|^2 \, dx \le \int_\Omega \eta^2(x) a^{ij}(x+he_k) D_i\Delta_k^h u(x) D_j \Delta_k^h u(x) \, dx. \tag{778}$$

We observe that

$$D_i v(x) = 2\eta(x)D_i\eta(x)\Delta_k^h u(x) + \eta^2(x)\Delta_k^h D_i u(x) \tag{779}$$

so that

$$D_i v(x) - 2\eta(x)D_i\eta(x)\Delta_k^h u(x) = \eta^2(x)\Delta_k^h D_i u(x). \tag{780}$$

By inserting (780) into (778) we obtain

$$\begin{aligned}\lambda \int_\Omega \eta^2(x) \left|\Delta_k^h Du \ (x)\right|^2 \, dx &\le \int_\Omega \eta^2(x) a^{ij}(x+he_k) D_i\Delta_k^h u(x) D_j \Delta_k^h u(x) \, dx \\ &= \int_\Omega a^{ij}(x+he_k)\left(D_i v(x) - 2\eta(x)D_i\eta(x)\Delta_k^h u(x)\right) D_j\Delta_k^h u(x) \, dx.\end{aligned} \tag{781}$$

From the inequality (776) we see that

$$\int_\Omega a^{ij}(x+he_k) D_i v(x) D_j\Delta_k^h u(x) \, dx \le C(\|u\|_{H^1(\Omega)} + \|f\|_{L^2(\Omega)})\|Dv\|_2. \tag{782}$$

Since $0 \le \eta \le 1$,

$$\begin{aligned}\|Dv\|_2 &= \|2\eta D\eta \Delta_k^h u + \eta^2 \Delta_k^h Du\|_2 \\ &\le 2\|\Delta_k^h u D\eta\|_2 + \|\eta D\Delta_k^h u\|_2\end{aligned} \tag{783}$$

Moreover, since the a^{ij} are bounded and $0 \le \eta \le 1$, there exists C' such that

$$\int_\Omega a^{ij}(x+he_k)\left(2\eta(x)D_i\eta(x)\Delta_k^h u(x)\right) D_j\Delta_k^h u(x) \, dx \le C'\|\eta D\Delta_k^h u\|_2 \|\Delta_k^h u D\eta\|_2. \tag{784}$$

We combine (784), (783) and (781) to see that

$$\begin{aligned}\lambda \int_\Omega \eta^2(x) \left|\Delta_k^h Du(x)\right|^2 \, dx \le{}& C''(\|u\|_{H^1(\Omega)} + \|f\|_{L^2(\Omega)})\left(\|\Delta_k^h u D\eta\|_2 + \|\eta D\Delta_k^h u\|_2\right) \\ &+ C''\|\eta D\Delta_k^h u\|_2 \|\Delta_k^h u D\eta\|_2\end{aligned} \tag{785}$$

for some constant $C'' > 0$. We apply Young's inequality —

$$ab \le \frac{1}{\epsilon}a^2 + \epsilon b^2 \tag{786}$$

— repeatedly to obtain

$$\begin{aligned}&\lambda\|\eta\Delta_k^h Du\|_2^2 \le C''\left(\frac{1}{\epsilon}\|u\|_{H^1(\Omega)}^2 + \epsilon\|\Delta_k^h u D\eta\|_2^2 + \frac{1}{\epsilon}\|u\|_{H^1(\Omega)}^2 + \epsilon\|\eta D\Delta_k^h u\|_2^2\right. \\ &+ \frac{1}{\epsilon}\|f\|_{L^2(\Omega)}^2 + \epsilon\|\Delta_k^h u D\eta\|_2^2 + \frac{1}{\epsilon}\|f\|_{L^2(\Omega)}^2 + \epsilon\|\eta D\Delta_k^h u\|_2^2 \\ &\left. + \epsilon'\|\eta D\Delta_k^h u\|_2^2 + \frac{1}{\epsilon'}\|\Delta_k^h u D\eta\|_2^2\right),\end{aligned} \tag{787}$$

which we rearrange as

$$\lambda\|\eta\Delta_k^h Du\|_2^2 \leq C''\left(\frac{2}{\epsilon}\|u\|_{H^1(\Omega)}^2 + \frac{2}{\epsilon}\|f\|_{L^2(\Omega)}^2 + \left(2\epsilon + \frac{1}{\epsilon}\right)\|\Delta_k^h u D\eta\|_2^2 + 3\epsilon\|\eta\Delta_k^h Du\|_2^2\right). \quad (788)$$

This implies

$$\left(\lambda - 3C''\epsilon\right)\|\eta\Delta_k^h Du\|_2^2 \leq C''\left(\frac{2}{\epsilon}\|u\|_{H^1(\Omega)}^2 + \frac{2}{\epsilon}\|f\|_{L^2(\Omega)}^2 + \left(2\epsilon + \frac{1}{\epsilon}\right)\|\Delta_k^h u D\eta\|_2^2\right). \quad (789)$$

By choosing ϵ sufficiently small and noting that there exists a constant C''' such that $\|\Delta_k^h u D\eta\|_2^2 \leq C'''\|u\|_{H^1(\Omega)}^2$, we obtain

$$\|\eta\Delta_k^h Du\|_2^2 \leq C''''\left(\|u\|_{H^1(\Omega)}^2 + \|f\|_{L^2(\Omega)}^2\right) \quad (790)$$

for some suitably chosen constant C''''. Since $\eta = 1$ on Ω', we can apply Theorem 89 to see that $u \in H^2(\Omega')$.

By letting $v = \zeta^2 u$ in the identity (770), where ζ is a cutoff function which is 1 on a set Ω'' such that $\Omega' \subset\subset \Omega'' \subset\subset \Omega$, we obtain

$$\lambda^2 \int_\Omega \zeta^2(x)\left|Du(x)\right|^2\, dx \leq \int_\Omega \zeta^2(x) a^{ij}(x) D_j u(x) D_i u(x)\, dx = \int_\Omega \zeta^2(x) g(x) u(x)\, dx. \quad (791)$$

Using Young's inequality and the technique we used above, we can easily show that (791) implies that there exists C'''' such that

$$\|u\|_{H^1(\Omega)} \leq C''''\left(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}\right). \quad (792)$$

Since u is in $H^2(\Omega')$, we can integrate by parts in (766) with v chosen to be smooth with compact support in Ω' in order to see that (767) must be satisfied for almost all $x \in \Omega'$. Since $\Omega' \subset\subset \Omega$ is arbitrary, it follows that this identity holds for almost all $x \in \Omega$. □

Note that we assumed $u \in H^1(\Omega)$ in Theorem 103, so that if the coefficients of L are Lipschitz and u solves (764) with $f \in L^2(\Omega)$, then no additional regularity condition on g is needed to ensure that $u \in H^2_{\text{loc}}(\Omega)$.

We note too that in the event that u is the unique weak solution of the boundary value problem

$$\begin{cases} L[u] = f & \text{in } \Omega \\ u = g & \text{on } \partial\Omega, \end{cases} \quad (793)$$

then we can bound the term $\|u\|_{L^2(\Omega)}$ on the right-hand side of (768) via a multiple of $\|f\|_{L^2(\Omega)}$ and $\|g\|_{H^{1/2}(\partial\Omega)}$ (since we already have a bound on $\|u\|_{H^1(\Omega)}$ in terms of these two norms in that case).

5.2. Global Regularity of Solutions of the Dirichlet Problem

We now establish that the solutions of the homogeneous boundary value problem

$$\begin{cases} L\left[u\right] = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega; \end{cases} \tag{794}$$

are elements of $H^2(\Omega)$ — as opposed to merely being elements of $H^2_{\mathrm{loc}}(\Omega)$ — under certain regularity assumptions on L and Ω.

THEOREM 104. *Suppose that $\Omega \subset \mathbb{R}^n$ is a $C^{1,1}$ domain, that*

$$L\left[u\right](x) = -D_i\left(a^{ij}(x)D_ju(x) + b^i(x)u(x)\right) + c^i(x)D_iu(x) + d(x)u(x) \tag{795}$$

is strongly elliptic in Ω, that $a^{ij}, b^i \in C^{0,1}\left(\overline{\Omega}\right)$, and that $c^i, d \in L^\infty(\Omega)$. Suppose also that $f \in L^2(\Omega)$, and that $u \in H^1(\Omega)$ is a weak solution of the problem

$$\begin{cases} L\left[u\right] = f & \textit{in } \Omega \\ u = 0 & \textit{on } \partial\Omega; \end{cases} \tag{796}$$

that is, $u \in H^1_0(\Omega)$ and

$$\begin{aligned} &\int_\Omega a^{ij}(x)D_ju(x)D_iv(x) + b^i(x)u(x)D_iv(x) + c^i(x)v(x)D_iu(x) + d(x)u(x)v(x)\ dx \\ &= \int_\Omega v(x)f(x)\ dx \end{aligned} \tag{797}$$

for all $v \in H^1_0(\Omega)$. Then $u \in H^2(\Omega)$, u satisfies the equation

$$-a^{ij}(x)D_iD_ju(x) + (-D_ja^{ji}(x) + c^i(x) - b^i(x))D_iu(x) + \left(d(x) - D_ib^i(x)\right)u(x) = f(x) \tag{798}$$

for almost all $x \in \Omega$, and there exists a real number $C > 0$ such that

$$\|u\|_{H^2(\Omega)} \leq C\left(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right) \tag{799}$$

PROOF. Since Ω is a $C^{1,1}$ domain, for each point x on the boundary of Ω, there is a ball B containing x, an open neighborhood N of 0 in $\mathbb{R}^n$ and a bijective $C^{1,1}$ mapping $\psi : N \to B$ such that $\psi(N \cap \mathbb{R}^n_+) = B \cap \Omega$ and $\psi(N \cap \partial\mathbb{R}^n_+) = \partial\Omega \cap B$. This mapping takes the equation under consideration to one of the same form, and the resulting solution of the transformed can be transformed back to a solution of the original problem. Consequently, we need only consider the case of a domain Ω which is a bounded open set in the upper half space $\mathbb{R}^n_+$ with boundary $\partial\mathbb{R}^n_+ = \{(x_1, \ldots, x_{n-1}, 0) : x_1, \ldots, x_{n-1} \in \mathbb{R}\}$.

If we let η be an element of $C^\infty_c(N)$, then for sufficiently small h and $k = 1, \ldots, n-1$,

$$v = \eta^2 \Delta^h_k u \in H^1_0\left(N \cap \mathbb{R}^n_+\right) \tag{800}$$

since u is supported on $N \cap \mathbb{R}^n_+$ and its trace is 0 on $N \cap \partial\mathbb{R}^n_+$ by assumption. Moreover, $\eta\Delta^h_k f$ is well-defined whenever $f \in N \cap \mathbb{R}^n_+$ and $k = 1, \ldots, n-1$. There were the requirements for applying the argument of the preceding section, By applying it to v, we see that D_kD_ju exists and that there is a constant C' such that

$$\|D_kD_ju\|_{L^2(\Omega)} \leq C'\left(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right) \tag{801}$$

for all $j = 1, \ldots, n$ and $k = 1, \ldots, n-1$.

We cannot use the same method to prove a bound on $D_n D_n u$. However, we can use the assumption that L is strongly elliptic to estimate $D_n D_n u$. In particular, by letting $\xi = (0, 0, \ldots, 0, 1)$ in the inequality

$$a^{ij}\xi_i\xi_j \geq \lambda|\xi|^2, \tag{802}$$

we see that $a^{n,n}(x) \geq \lambda$ for all $x \in \Omega$. From this and the fact that

$$-a^{ij}(x)D_iD_ju(x) + (-D_ja^{ji}(x) + c^i(x) - b^i(x))D_iu(x) + \left(d(x) - D_ib^i(x)\right)u(x) = f(x), \tag{803}$$

which is a consequence of Theorem 103, we see that

$$\begin{aligned} D_nD_nu(x) = \frac{1}{\lambda}\Bigg(&\sum_{i=1}^{n-1}\sum_{j=1}^{n} a^{ij}(x)D_iD_ju(x) + \sum_{i=1}^{n}\sum_{j=1}^{n} D_ja^{ji}(x)D_iu(x) \\ &- \sum_{i=1}^{n}\left(c^i(x) - b^i(x)\right)D_iu(x) - \left(d(x) + \sum_{i=1}^{n} D_ib^i(x)\right)u(x) + f(x)\Bigg). \end{aligned} \tag{804}$$

The desired estimate on D_nD_nu follows easily. □

If $g \in H^{3/2}(\partial\Omega)$, then there exists $\varphi \in H^2(\Omega)$ such that $\mathscr{T}[\varphi] = g$ and the $H^2(\Omega)$ norm of φ is a multiple of the $H^{3/2}(\partial\Omega)$ norm of g. If u is a weak solution of the homogeneous boundary value problem

$$\begin{cases} L[u] = f - L[\varphi] & \text{in } \Omega \\ \qquad\quad u = 0 & \text{on } \partial\Omega, \end{cases} \tag{805}$$

then $v = u + \varphi$ is a weak solution of

$$\begin{cases} L[v] = f & \text{in } \Omega \\ \quad v = g & \text{on } \partial\Omega. \end{cases} \tag{806}$$

From Theorem 104, we see that

$$\begin{aligned} \|u\|_{H^2(\Omega)} \leq C\left(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right) &\leq C\left(\|v - \varphi\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right) \\ &\leq C\left(\|v\|_{L^2(\Omega)} + \|\varphi\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right). \end{aligned} \tag{807}$$

From this and the fact that $u = v - \varphi$ we easily obtain

$$\begin{aligned} \|v\|_{H^2(\Omega)} &\leq C'\left(\|v\|_{L^2(\Omega)} + \|\varphi\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right) \\ &\leq C''\left(\|v\|_{L^2(\Omega)} + \|g\|_{H^{3/2}(\partial\Omega)} + \|f\|_{L^2(\Omega)}\right). \end{aligned} \tag{808}$$

If v is the unique solution of (806) then we have the estimate

$$\|v\|_{H^2(\Omega)} \leq C'''\left(\|g\|_{H^{3/2}(\partial\Omega)} + \|f\|_{L^2(\Omega)}\right). \tag{809}$$

since the $H^1(\Omega)$ norm of v is bounded in terms of norms of f and g already appearing.

Theorem 105. *Suppose that Ω is a $C^{1,1}$ domain in $\mathbb{R}^n$, that*

$$L[u](x) = -D_i\left(a^{ij}(x)D_ju(x)\right) + c^i(x)D_iu(x) + d(x)u(x) \tag{810}$$

is a strongly elliptic partial differential operator on Ω, that $a^{ij} \in C^{0,1}\left(\overline{\Omega}\right)$ and that c^i and d are elements of $L^\infty(\Omega)$. Suppose also that $d(x) \geq 0$ for almost all $x \in \Omega$. Then for every

$f \in L^2(\Omega)$ and $g \in H^{3/2}(\partial\Omega)$, the boundary value problem

$$\begin{cases} L[u](x) = f(x) & \text{in } \Omega \\ \quad u(x) = g(x) & \text{on } \partial\Omega. \end{cases} \tag{811}$$

admits a unique solution $u \in H^2(\Omega)$. Moreover, there is a constant C not depending on f or g such that

$$\|u\|_{H^2(\Omega)} \leq C\left(\|f\|_{H^2(\Omega)} + \|g\|_{H^{3/2}(\partial\Omega)}\right) \tag{812}$$

whenever u solves (811). In other words, the operator $L \oplus \mathscr{T}$ is an isomorphism $H^2(\Omega) \to L^2(\Omega) \oplus H^{3/2}(\partial\Omega)$.

It is an obvious consequence of this theorem that the general elliptic operator

$$L[u](x) = -D_i\left(a^{ij}(x)D_j u(x) + b^i(x)u(x)\right) + c^i(x)D_i u(x) + d(x)u(x) \tag{813}$$

is a Fredholm operator of index 0 which maps $H^2(\Omega)$ into $L^2(\Omega) \oplus H^{3/2}(\partial\Omega)$, assuming that the appropriate regularity conditions on Ω and the coefficients of L are met (since the principal part of L is an isomorphism and the residual operator is compact).

5.3. Higher Order Local and Global Regularity for the Dirichlet Problem

Suppose that

$$L[u](x) = -D_i\left(a^{ij}(x)D_j u(x) + b^i(x)u(x)\right) + c^i(x)D_i u(x) + d(x)u(x) \tag{814}$$

is a strongly elliptic operator on a domain Ω such that a^{ij} and b^i are Lipschitz continuous and c^i, d are essentially bounded. Then, if u satisfies

$$L[u](x) = f(x) \quad \text{in } \Omega \tag{815}$$

with $f \in L^2(\Omega)$, we have $u \in H^2_{\text{loc}}(\Omega)$. If $k \geq 1$, $f \in H^k(\Omega)$, $a^{ij}, b^i \in C^{k,1}\left(\overline{\Omega}\right)$ and $c^i, d \in C^{k-1}\left(\overline{\Omega}\right)$, then we can apply a differential operator D^α of order less than or equal to k to both sides of (815) to obtain

$$D^\alpha L[u](x) = D^\alpha f(x) \quad \text{in } \Omega. \tag{816}$$

A somewhat tedius calculation shows that we can rearrange the equation (816) as

$$L[D^\alpha u](x) = \tilde{f}(x) \quad \text{in } \Omega, \tag{817}$$

with $\tilde{f} \in L^2(\Omega)$. Note that applying the differential operator D^α to $L[u]$ in (816) involves differentiating the coefficients of L. We can now apply Theorem 103 to see that $D^\alpha u \in H^2_{\text{loc}}(\Omega)$ for any $|\alpha| \leq k$. In particular, $u \in H^{k+2}_{\text{loc}}(\Omega)$.

Assuming, in addition, that Ω is a $C^{k+1,1}$ domain, we can "straigten the boundary of Ω" and apply a similar argument to conclude that solutions of the boundary value problem

$$\begin{cases} L[u](x) = f(x) & \text{in } \Omega \\ \quad u(x) = g(x) & \text{on } \partial\Omega \end{cases} \tag{818}$$

are elements of $H^{k+2}(\Omega)$ when $f \in H^k(\Omega)$ and $g \in H^{k+\frac{3}{2}}(\partial\Omega)$. In the event that (818) admits unique solutions, we can rephrase this in the usual way: $L \oplus \mathscr{T}$ is an isomorphism $H^{k+2}(\Omega) \to H^k(\Omega) \oplus H^{k+3/2}(\partial\Omega)$.

One further remark is in order. From the Sobolev embedding theorem, we know that if $f \in H^k(\Omega)$ for all nonnegative integers k, then $f \in C^\infty(\Omega)$. It follows that when Ω is a C^∞ domain, $f \in C^\infty(\Omega)$ and $g \in C^\infty(\partial\Omega)$, then the solutions of

$$\begin{cases} L[u](x) = f(x) & \text{in } \Omega \\ u(x) = g(x) & \text{on } \partial\Omega \end{cases} \tag{819}$$

are also elements of $C^\infty(\Omega)$. Other, similar, results can obtained by combining the basic facts of elliptic regularity with Sobolev embedding theorems.

5.4. A Regularity Result for Convex Domains

When Ω is Lipschitz (which is the typical case in applications, particularly numerical analysis), our definition of the space $H^\alpha(\partial\Omega)$ is only valid when $\alpha \leq 1$ since local parameterizations of the boundary only admit one derivative. Nonethessless, we have characterized the trace of $H^2(\Omega)$ as the space $V^{3/2}(\partial\Omega)$ (see Theorem 76), and it is reasonable to ask whether or not solutions of the boundary value problem

$$\begin{cases} L[u](x) = f(x) & \text{in } \Omega \\ u(x) = g(x) & \text{on } \partial\Omega \end{cases} \tag{820}$$

are elements of $H^2(\Omega)$ when $f \in L^2(\Omega)$ and $g \in V^{3/2}(\partial\Omega)$.

In general, this is not the case. To see this, we suppose that Ω is a bouned domain in $\mathbb{R}^2$ with a single corner point of angle ω at the origin, and that in a neighborhood of the origin, the boundary of Ω coincides with the curve Γ parameterized via

$$\begin{aligned} x(r) &= r\cos(\omega) \\ y(r) &= r\sin(\omega). \end{aligned} \tag{821}$$

It is easy to verify that when $k\pi/\omega$ is not an integer, the function

$$u_k(r,\theta) = r^{\frac{k\pi}{\omega}} \sin\left(\frac{k\pi}{\omega}\theta\right) \tag{822}$$

is harmonic and that its restriction to Γ is 0. But when $\omega > \pi$, the second derivative of u_k with respect to r is not square integrable. It follows that there exists a solution of (820) with $f = 0$ and g smooth which is not an element of $H^2(\Omega)$.

When Ω is convex, however, solutions of (820) are necessarily elements of $H^2(\Omega)$ under mild conditions on L. We now briefly outline the proof of this result.

A proof of the following theorem can be found in Chapter 3 of [**10**]. It can be viewed as a consequence of the fact that the trace of the second fundamental form

$$B(\xi,\eta) = -\frac{\partial\nu}{\partial\xi}\cdot\eta \tag{823}$$

for the boundary of a convex domain is nonpositive.

Theorem 106. *Suppose that Ω is a convex, bounded C^2 domain in $\mathbb{R}^n$, and that*

$$L[u](x) = -D_i\left(a^{ij}(x)D_j u(x)\right) \tag{824}$$

is a strongly elliptic operator whose coefficients are Lipschitz. Then there exists a constant C depending only on the diameter *of Ω and the Lipschitz norms of the coefficients of L such that*

$$\|u\|_{H^2(\Omega)} \leq C \|L\left[u\right]\|_{L^2(\Omega)} \tag{825}$$

whenever $u \in H^2\left(\Omega\right) \cap H_0^1\left(\Omega\right)$.

In order to establish the existence of solutions of the homogeneous problem

$$\begin{cases} L\left[u\right](x) = f(x) & \text{in } \Omega \\ \quad u(x) = 0 & \text{on } \partial\Omega \end{cases} \tag{826}$$

for Ω convex and with L satisfying the hypotheses of the preceding theorem, we first construct a sequence of convex open sets $\{\Omega_m\}$ such that $\Omega_m \subset \Omega$ and $\mathrm{dist}(\partial\Omega_m, \partial\Omega) \to 0$ as $m \to \infty$. Next, we solve the obvious Dirichlet boundary value problem on each of the domains. We can extend each of the resulting functions to $\mathbb{R}^n$ since the Sobolev extension theorem applies to Lipschitz domains. We view the resulting functions as a sequence $\{u_m\}$ of functions defined on Ω. Theorem 106 implies that this sequence is uniformly bounded in $H^2\left(\Omega\right)$, so that the Banach-Alaoglu theorem implies that there is a weakly convergence subsequence of $\{u_m\}$. The weak limit is a solution of the boundary value problem (826). The key to this approach is the uniform bound provided by Theorem 106.

The usual procedures can be applied in order to study linear elliptic operators with lower order terms and the inhomogeneous boundary value problem.

CHAPTER 6

Elementary Results from the Calculus of Variations

Solving certain partial differential equation can be shown to be equivalent to minimizing an "energy" functional given on a Banach space. The study of such problems is called the calculus of variations, and we will now give a treatment of certain elementary techiniques in the field.

6.1. Fréchet Derivatives

We say that a mapping $T : X \to Y$ between Banach spaces is Fréchet differentiable at the point $x \in X$ provided there exists a linear mapping $L : X \to Y$ such that

$$\lim_{h \to 0} \frac{\|T[x+h] - T[x] - L[h]\|}{\|h\|} = 0. \tag{827}$$

It is easy to verify that Fréchet derivatives are unique, and we call the linear mapping L the Fréchet derivative of T at the point x. We say that T is Fréchet differentiable on an open set U in X if it is Fréchet differentiable at every point of U. When this is the case, we denote the mapping $U \to \mathscr{L}(X,Y)$ which takes x to the Fréchet derivative of T at x by T'. We will depart from our usual practice and use the notation T'_x to denote the image of x under T', which is the Fréchet derivative of T at the point x. In other words, $T'_x[h] = L[h]$, where L is as in (827).

Exercise 32. *Suppose that Ω is a bounded, open set in $\mathbb{R}^n$. Show that the map $T : C\left(\overline{\Omega}\right) \to C\left(\overline{\Omega}\right)$ defined via $T[f](x) = (f(x))^2$ is Fréchet differentiable at every point of $C\left(\overline{\Omega}\right)$, and that*

$$T'_f[g](x) = 2f(x)g(x). \tag{828}$$

Exercise 33. *Suppose that Ω is a bounded, open set in $\mathbb{R}^n$. Show that the map $T : C\left(\overline{\Omega}\right) \to C\left(\overline{\Omega}\right)$ defined via $T[f](x) = \exp(f(x))$ is Fréchet differentiable at every point of $C\left(\overline{\Omega}\right)$ and compute its Fréchet derivative.*

Fréchet derivatives generalize the Jacobian or "total" derivative of vector calculus, and can be used in much the same way. For instance, the iteration

$$x_{n+1} = x_n - \left(T'_{x_n}\right)^{-1}[T[x_n]], \tag{829}$$

generalizes Newton's method for the solution of $T[x] = 0$. The Newton-Kantorovich Theorem gives conditions under which (829) converges (see, for instance, [**2**]).

6.2. Nonlinear Functionals

Here, we will mostly be concerned with real-valued nonlinear functionals acting on Banach spaces; that is, mappings of the form $T : X \to \mathbb{R}$. It is easy to verify that many of the elementary result of calculus generalize to this setting, including those regarding local extrema. We say that a functional $T : X \to \mathbb{R}$ has a critical point at x if T is Fréchet differentiable in an open set containing x and $T'_x = 0$.

THEOREM 107. *If the functional $T : X \to \mathbb{R}$ is Fréchet differentiable on an open set U and T has a local minimum at the point $x \in U$, then $T'_x = 0$. That is, a local minumum of T is a critical point.*

PROOF. Let $y \in X$ and define $f : \mathbb{R} \to \mathbb{R}$ via $f(t) = T\left[x + ty\right]$. It is easy to verify that f is differentiable, and that $f'(t) = T'_{x+ty}\left[y\right]$. If T has a local minimum at x, then f has a local minimum at 0, so that $0 = f'(0) = T'_x\left[y\right]$. Since this is true for all $y \in X$, we have $T'_x = 0$, as desired. □

We say that a linear function $T : X \to \mathbb{R}$ is convex if

$$T\left[(1-t)x + ty\right] \le (1-t)T\left[x\right] + tT\left[y\right] \tag{830}$$

for all x, y in X and $0 \le t \le 1$. It is strictly convex provided

$$T\left[(1-t)x + ty\right] < (1-t)T\left[x\right] + tT\left[y\right] \tag{831}$$

for all x, y in X and $0 < t < 1$. Any norm $\|\cdot\|$ on X is a convex linear functional. A functional $T : X \to \mathbb{R}$ is weakly lower semicontinuous if whenever $\{x_k\}$ is a sequence which converges weakly to x,

$$T\left[x\right] \le \liminf_{k\to\infty} I\left[x_k\right]. \tag{832}$$

We omit a proof of the following theorem, which can be found in many references.

THEOREM 108. *If $T : X \to \mathbb{R}$ is a continuous, convex functional on the Banach space X, then T is weakly lower semicontinuous.*

One important consequence of Theorem 108 is that any norm given on a Banach space X is weakly lower semicontinuous. This follows because norms are convex and continuous. We leave it to the reader to verify that if $\|\cdot\|$ is a norm for X, then $f(x) = \|x\|^2$ is also weakly lower semicontinuous.

We say that a functional $T : X \to \mathbb{R}$ is coercive provided whenever $\{x_j\}$ is a sequence such that $\|x_j\| \to \infty$, $T\left[x_j\right] \to \infty$. The following theorem regarding coercive functionals is of fundamental importance.

THEOREM 109. *Suppose that X is a reflexive Banach space, and that $T : X \to \mathbb{R}$ is continuous, coercive and weakly lower semicontinuous. Then T has a global minimum.*

PROOF. Let

$$m = \inf_{x\in X} T\left[x\right], \tag{833}$$

and let $\{x_k\}$ be a sequence such that $T[x_k]$ converges monotonically from above to m. Since $T[x_k]$ does not converge to ∞, the coercivity of T implies that $\{x_k\}$ is bounded (note that this is the case even if $T[x_k]$ converges to $-\infty$, although we will shortly show that $m \neq -\infty$). Since X is reflexive, the Banach-Alaoglu theorem (which holds that bounded sets in reflexive Banach spaces are weakly relatively compact) implies that there is a subsequence of $\{x_k\}$ which converges weakly to some $x \in X$. Without loss of generality, we may assume that $x_k \rightharpoonup x$. Now, by the weak lower semicontinuity of T, we obtain

$$T[x] \leq \liminf_{k\to\infty} I[x_k] = m, \tag{834}$$

from which we conclude that x is a global minimizer of T. Note that since T is a mapping into $\mathbb{R}$ and not the extended reals, it cannot be the case that $m = -\infty$. □

It is, of course, a corollary of Theorem 109 that a continuous, convex, coercive linear function as a global minimum. Strict convexity can often be used to obtain the uniqueness results, as in the following theorems.

THEOREM 110. *If $T : X \to \mathbb{R}$ is strictly convex, then T has at most one global minimum in X.*

PROOF. Suppose that x_1 and x_2 are distinct global minima of T. Then, by strict convexity, we have

$$\inf_{x\in X} T[x] \leq T\left[\frac{x_1+x_2}{2}\right] < \frac{1}{2}T[x_1] + \frac{1}{2}T[x_2] = \inf_{x\in X} T[x], \tag{835}$$

which is a contradiction. □

THEOREM 111. *Suppose that $T : X \to \mathbb{R}$ is strictly convex and Fréchet differentiable on X. Then T has at most one critical point in X.*

PROOF. If $f : \mathbb{R} \to \mathbb{R}$ is differentiable and strictly convex, then f' must be strictly increasing. Since $f'(x) = 0$ if x is a critical point, it follows that there can be at most one critical point.

To prove the general case, suppose that u is a critical point of T, fix $v \in X$, and define $f : \mathbb{R} \to \mathbb{R}$ via

$$f(t) = T[u + tv]. \tag{836}$$

Then f is differentiable and

$$f'(t) = T'_{u+tv}[v]. \tag{837}$$

It is easy to verify that f' is strictly convex and $f'(0) = T'_u[v] = 0$ (since u is a critical point). Then $f'(t) \neq 0$ for all $t\ neq 0$, from which we conclude that

$$T'_{u+tv}[v] \neq \tag{838}$$

for all $t \neq 0$. Since this holds for all v, we conclude that u is the only critical point of T. □

THEOREM 112. *Suppose that $T : X \to \mathbb{R}$ is a linear functional on the Banach space X, and that T is Fréchet differentiable on X. Suppose also that*

$$(T'_u - T'_v)[u - v] \geq 0 \tag{839}$$

for all u, v in X. Then T is convex. If strict inequality holds in (839) when $u \neq v$, then T is strictly convex.

Proof. We fix u and v, and define $f : \mathbb{R} \to \mathbb{R}$ via

$$f(t) = T\left[u + t(v-u)\right]. \tag{840}$$

Then f is differentiable and

$$f'(t) = T'_{u+t(v-u)}\left[v - u\right] \tag{841}$$

If $s < t$ then

$$\begin{aligned} f'(t) - f'(s) &= \left(T'_{u+t(v-u)} - T'_{u+s(v-u)}\right)\left[v-u\right] \\ &= \frac{1}{t-s}\left(T'_{u+t(v-u)} - T'_{u+s(v-u)}\right)\left[(u+t(v-u)) - (u+s(v-u))\right] \\ &\geq 0. \end{aligned} \tag{842}$$

We conclude that f is nondecreasing, which implies that it is convex ($f''(x) \geq 0$). In particular,

$$f(t) = f(1t + 0(1-t)) \leq tf(1) + (1-t)f(0), \tag{843}$$

which is equivalent to

$$T\left[tv + (1-t)u\right] \leq tI\left[v\right] + (1-t)I\left[u\right]. \tag{844}$$

If the inequality is strict, then f is strictly increasing, and we get strict inequality in (844). □

Exercise 34. *Is the functional $T : \mathbb{R} \to \mathbb{R}$ defined via $T\left[x\right] = \exp(x)$ coercive?*

6.3. Application to a Linear Elliptic Partial Differential Equation

We will suppose that Ω is an open, bounded set in $\mathbb{R}^n$, that $q \in L^\infty(\Omega)$ such that $q(x) \geq 0$ almost everywhere in Ω, and that $h \in L^2(\Omega)$. It is easy to verify that the Fréchet derivative of the linear functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined via

$$I\left[u\right] = \frac{1}{2}\int_\Omega \left|Du(x)\right|^2\, dx + \frac{1}{2}\int_\Omega q(x)\left(u(x)\right)^2\, dx - \int_\Omega h(x)u(x)\, dx \tag{845}$$

is

$$I'_u\left[v\right] = \int_\Omega Du(x)\cdot Dv(x)\, dx + \int_\Omega q(x)u(x)v(x)\, dx - \int_\Omega h(x)v(x)\, dx. \tag{846}$$

If $u \in H_0^1(\Omega)$ is a critical point for I, then

$$I'_u = 0; \tag{847}$$

that is,

$$\int_\Omega Du(x)\cdot Dv(x)\, dx + \int_\Omega q(x)u(x)v(x)\, dx - \int_\Omega h(x)v(x)\, dx = 0 \text{ for all } v \in H_0^1(\Omega). \tag{848}$$

In other words, u is a critical point for I if and only if u is a weak solution of the boundary value problem

$$\begin{cases} -\Delta u(x) + q(x)u(x) = h(x) & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega. \end{cases} \tag{849}$$

We will now show that I is strictly convex and coercive. It then follows that I has a global minimum u, which must be its unique critical point. In particular, u is the unique weak solution of (849). We first show that I is coercive. Since $q(x)(u(x))^2$ is nonnegative, we have

$$\begin{aligned} I\left[u\right] &\geq \frac{1}{2}\|Du\|^2_{L^2(\Omega)} - \int_\Omega h(x)u(x)\ dx \\ &\geq \frac{1}{2}\|Du\|^2_{L^2(\Omega)} - \|h\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)} \end{aligned} \tag{850}$$

Now we apply Poincaré's inequality to see that

$$I[u] \geq C\|u\|^2_{H^1_0(\Omega)} - \|h\|_{L^2(\Omega)}\|u\|_{H^1_0(\Omega)}. \tag{851}$$

for some $C > 0$. It follows that $I\left[u_k\right] \to \infty$ if $\|u_k\|_{H^1_0(\Omega)} \to \infty$. The strict convexity of I follows from Theorem 111 since for $u \neq v$,

$$\begin{aligned} \left(I'_u - I'_v\right)\left[u - v\right] &= \int_\Omega (Du(x) - Dv(x)) \cdot (Du(x) - Dv(x))\ dx + \int_\Omega q(x)(u(x) - v(x))^2\ dx \\ &\geq \int_\Omega D(u - v)(x) \cdot D(u - v)(x)\ dx > 0. \end{aligned} \tag{852}$$

6.4. Application to a Semilinear Elliptic Partial Differential Equation

We will now apply the machinary of this chapter to the boundary value problem

$$\begin{cases} -\Delta u(x) + q(x)u(x) = h(x) + f(u) & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega \end{cases} \tag{853}$$

under the assumptions that Ω is an open, bounded set in $\mathbb{R}^n$, that $q \in L^\infty(\Omega)$ such that $q(x) \geq 0$ almost everywhere in Ω, that $h \in L^2(\Omega)$, and that $f : \mathbb{R} \to \mathbb{R}$ is continuous and bounded.

We first define the function $F : \mathbb{R} \to \mathbb{R}$ via the formula

$$F(t) = \int_0^t f(s)\ ds. \tag{854}$$

Then it is easy to see that $u \in H^1_0(\Omega)$ is a critical point of

$$I\left[u\right] = \frac{1}{2}\int_\Omega Du(x) \cdot Du(x)\ dx + \frac{1}{2}\int_\Omega q(x)(u(x))^2\ dx - \int_\Omega h(x)u(x)\ dx - \int_\Omega F(u(x))\ dx. \tag{855}$$

if and only if u is a weak solution of (853). From Poincaré's inequality, we know that the norm $\|u\|_*$ defined via

$$\|u\|^2_* = \int_\Omega Du(x) \cdot Du(x)\ dx + \int_\Omega q(x)(u(x))^2\ dx \tag{856}$$

is equivalent to the $H_0^1(\Omega)$ norm (note that there is nothing to preclude q from being equal to 0 everywhere in Ω). We will show that I is coercive and weakly lower semicontinuous with respect to $\|\cdot\|_*$. It will then follow that I has a global minimum u which is a weak solution of (853).

First, we address the coercivity of I. Since f is bounded, there exists $M > 0$ such that $|f(t)| \leq M$ for all t. It follows that

$$|F(t)| \leq \int_0^t f(s)\ ds \leq M\,|t|\,. \tag{857}$$

From this and the facts that $H_0^1(\Omega)$ is continuously embedded in $L^1(\Omega)$ and the $\|{*}\|_*$ norm is equivalent to the $H_0^1(\Omega)$ norm, we see that

$$\left|\int_\Omega F(u(x))\ dx\right| \leq M\int_\Omega |u(x)|\ \ dx = M\,\|u\|_{L^1(\Omega)} \leq C\,\|u\|_{H_0^1(\Omega)} \leq C\,\|u\|_* \tag{858}$$

for all $u \in H_0^1(\Omega)$. Using this, we have

$$\begin{aligned} I\,[u] &= \frac{1}{2}\|u\|_*^2 - \int_\Omega h(x)u(x)\ dx - \int_\Omega F(u(x))\ dx \\ &\geq \frac{1}{2}\|u\|_*^2 - C\|u\|_{L^2(\Omega)} - C\|u\|_* \\ &\geq \frac{1}{2}\|u\|_*^2 - C\|Du\|_{L^2(\Omega)} - C\|u\|_* \\ &\geq \frac{1}{2}\|u\|_*^2 - C\|u\|_*. \end{aligned} \tag{859}$$

In the second to last line, we used Poincaré's inequality. Inequality (859) shows that I is coercive.

Now we will show that I is weakly lower semicontinuous. Suppose that $\{u_k\}$ is a sequence in $H_0^1(\Omega)$ which converges weakly to u. The image of a weakly convergent sequence under compact mappings is strongly convergent and the embedding of $H_0^1(\Omega)$ into $L^2(\Omega)$ is compact. It follows that some subsequence of u_k converges to u in $L^2(\Omega)$. So, without loss of generality, we can assume that $\{u_k\}$ converges to u weakly in $H_0^1(\Omega)$, strongly in $L^2(\Omega)$, and pointwise almost everywhere.

Since $\{u_k\}$ converges pointwise almost everywhere to u, which is in square integrable, there exists a function $g \in L^2(\Omega)$ such that $|u_k(x)| \leq g(x)$ for almost all $x \in \Omega$. We observe that

$$F(u_k(x)) \leq M|u_k(x)| \leq Mg(x) \tag{860}$$

and note that since $L^2(\Omega)$ is continously embedded in $L^1(\Omega)$, we can apply the dominated convergence theorem to obtain

$$\lim_k \int_\Omega F(u_k(x))\ dx = \int_\Omega F(u(x)). \tag{861}$$

Obviously,

$$\lim_k \int_\Omega h(x)u_k(x)\ dx = \int_\Omega h(x)u(x)\ dx. \tag{862}$$

Since norms and their squares are weakly lower semicontinuous, we have

$$\|u\|_*^2 \leq \liminf_k \|u_k\|_*^2. \tag{863}$$

Now combining (861), (862) and (863) gives

$$\begin{aligned} I[u] &= \frac{1}{2}\|u\|_*^2 - \int_\Omega F(u(x))\ dx - \int_\Omega u(x)h(x)\ dx \\ &= \frac{1}{2}\|u\|_*^2 - \liminf_{k\to\infty} \int_\Omega F(u_k(x))\ dx - \liminf_{k\to\infty} \int_\Omega u_k(x)h(x)\ dx \\ &\leq \liminf_{k\to\infty} \frac{1}{2}\|u_k\|_*^2 - \liminf_{k\to\infty} \int_\Omega F(u_k(x))\ dx - \liminf_{k\to\infty} \int_\Omega u_k(x)h(x)\ dx \\ &= \liminf_{k\to\infty} I[u_k], \end{aligned} \tag{864}$$

which establishes the weak lower semicontinuity of I. We conclude that I has a global minimum u.

6.5. A Less Restrictive Condition on the Nonlinear Term

We will now consider the boundary value problem

$$\begin{cases} -\Delta u(x) + q(x)u(x) = h(x) + f(u) & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega \end{cases} \tag{865}$$

under somewhat less restrictive conditions on the function f. More specifically, we assume that Ω is an open, bounded set in $\mathbb{R}^n$, that $q \in L^\infty(\Omega)$ such that $q(x) \geq 0$ almost everywhere in Ω, and that $h \in L^2(\Omega)$. We will also suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuous function such that

$$|f(t)| \leq M + b|t|. \tag{866}$$

Here, we will assume that $b < \lambda_1$, where λ_1 is the smallest eigenvalue of the operator $-\Delta + q$. The variational characterization of λ_1 is

$$\lambda_1 = \inf_{u \in H_0^1(\Omega)\setminus\{0\}} \frac{\langle(-\Delta+q)[u], u\rangle_{L^2(\Omega)}}{\langle u, u\rangle_{L^2(\Omega)}} = \inf_{u \in H_0^1(\Omega)\setminus\{0\}} \frac{\int_\Omega Du(x)\cdot Du(x)\ dx + \int q(x)(u(x))^2\ dx}{\int_\Omega (u(x))^2\ dx} \tag{867}$$

As before, we define the function $F : \mathbb{R} \to \mathbb{R}$ via the formula

$$F(t) = \int_0^t f(s)\ ds \tag{868}$$

and $I : H_0^1(\Omega) \to \mathbb{R}$ via

$$I[u] = \frac{1}{2}\int_\Omega Du(x)\cdot Du(x)\ dx + \frac{1}{2}\int_\Omega q(x)(u(x))^2\ dx. - \int_\Omega h(x)u(x)\ dx - \int_\Omega F(u(x))\ dx, \tag{869}$$

and we let $\|\cdot\|_*$ denote the norm defined via (856). From (868), we see that

$$|F(t)| \leq M|t| + \frac{b}{2}|t|^2. \tag{870}$$

The weakly lower semincontinuity of I follows as before, so in order to establish the existence of a global minumum of I, we need only prove that I is coercive. To that end, we observe that Formula (867) implies that

$$\|u\|^2_{L^2(\Omega)} \leq \frac{1}{\lambda_1}\|u\|^2_* \tag{871}$$

for all $u \in H^1_0(\Omega)$. We combine (870) and (871) to obtain

$$\begin{aligned} \left|\int_\Omega F(u(x))\ dx\right| &\leq M\int_\Omega |u(x)|\ \ dx + \frac{b}{2}\int_\Omega |u(x)|^2 \\ &\leq C\|u\|_* + \frac{b}{2\lambda_1}\|u\|^2_*. \end{aligned} \tag{872}$$

It follows that

$$\begin{aligned} I[u] &= \frac{1}{2}\|u\|^2_* - \int_\Omega h(x)u(x)\ dx - \int_\Omega F(u(x))\ dx \\ &\geq \frac{1}{2}\|u\|^2_* - C\|u\|_{L^2(\Omega)} - C\|u\|_* - \frac{b}{2\lambda_1}\|u\|^2_* \\ &\geq \left(\frac{1}{2} - \frac{b}{2\lambda_1}\right)\|u\|^2_* - C\|Du\|_{L^2(\Omega)} - C\|u\|_* \\ &\geq \left(\frac{1}{2} - \frac{b}{2\lambda_1}\right)\|u\|^2_* - C\|u\|_*. \end{aligned} \tag{873}$$

Since $\frac{1}{2} - \frac{b}{2\lambda_1} > 0$, (873) implies that I is coercive.

Bibliography

[1] Adams, R., and Fournier, J. *Sobolev Spaces*, second ed. Elsevier, 2003.

[2] Argyros, I. *Convergence and Applications of Newton-Type Methods*. Springer, 2008.

[3] Besov, O., Il'in, V., and Nikol'skiĭ. *Integral representations of functions and imbedding theorems*, vol. 2. V. H. Winston and Sons, Wasington, D.C., 1979.

[4] Conway, J. B. *A Course in Functional Analysis*, fourth ed. Springer, 1997.

[5] Enflo, P. A counterexample to the approximation property in Banach spaces. *Acta Mathematica 130* (1973), 309–317.

[6] Evans, L. *Partial Differential Equations*, second ed. American Mathematical Society, 2010.

[7] Evans, L., and Gariepy, R. *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.

[8] Folland, G. B. *Real Analysis: Modern Techniques and Their Application*, second ed. Wiley-Interscience, New York, 1999.

[9] Grafakos, L. *Modern Fourier Analysis*. Springer, 2009.

[10] Grisvard, P. *Elliptic Problems in Nonsmooth Domains*. Pitman, 1985.

[11] Johnsson, A., and Wallin, H. *Function spaces on subsets of* $\mathbb{R}^n$. Harwood Academic Publishers, 1984.

[12] Jones, P. W. Quasiconformal mappings and extendability of functions in Sobolev spaces. *Acta Mathematica 147* (1981).

[13] Lindenstrauss, J., and Tzafiri, L. On the complemented subspace problem. *Israel Journal of Mathematics 9* (1971), 263–269.

[14] Marschall, J. The trace of Sobolev-Slobodeckij spaces on Lipschitz domains. *Manuscripta Mathematica 58* (1987), 47–65.

[15] Rudin, W. *Functional Analysis*, second ed. Wiley, 1991.

[16] Stein, E. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970.