The Invariance of the Index of Elliptic Operators

Constantine Caramanis*
Harvard University

April 5, 1999

Abstract

In 1963 Atiyah and Singer proved the famous Atiyah-Singer Index Theorem, which states, among other things, that the space of elliptic pseudodifferential operators is such that the collection of operators with any given index forms a connected subset. Contained in this statement is the somewhat more specialized claim that the index of an elliptic operator must be invariant under sufficiently small perturbations. By developing the machinery of distributions and in particular Sobolev spaces, this paper addresses this more specific part of the famous Theorem from a completely analytic approach. We first prove the regularity of elliptic operators, then the finite dimensionality of the kernel and cokernel, and finally the invariance of the index under small perturbations.

*cmcaram@fas.harvard.edu

Acknowledgements

I would like to express my thanks to a number of individuals for their contributions to this thesis, and to my development as a student of mathematics. First, I would like to thank Professor Clifford Taubes for advising my thesis, and for the many hours he spent providing both guidance and encouragement. I am also indebted to him for helping me realize that there is no analysis without geometry. I would also like to thank Spiro Karigiannis for his very helpful critical reading of the manuscript, and Samuel Grushevsky and Greg Landweber for insightful guidance along the way.

I would also like to thank Professor Kamal Khuri-Makdisi who instilled in me a love for mathematics. Studying with him has had a lasting influence on my thinking. If not for his guidance, I can hardly guess where in the Harvard world I would be today. Along those lines, I owe both Professor Dimitri Bertsekas and Professor Roger Brockett thanks for all their advice over the past 4 years.

Finally, but certainly not least of all, I would like to thank Nikhil Wagle, Allison Rumsey, Sanjay Menon, Michael Emanuel, Thomas Knox, Demian Ordway, and Benjamin Stephens for the help and support, mathematical or other, that they have provided during my tenure at Harvard in general, and during the researching and writing of this thesis in particular.

April 5th, 1999
Lowell House, I-31
Constantine Caramanis

Contents

1 Introduction

This paper defines, and then examines some properties of a certain class of linear differential operators known as elliptic operators. We investigate the behavior of this class of maps operating on the space of sections of a vector bundle over a compact manifold. The ultimate goal of the paper is to show that if an operator L is elliptic, then the **index** of the operator, given by

$$\mathrm{Index}(L) := \mathrm{dimKernel}(L) - \mathrm{dimCokernel}(L),$$

is invariant under sufficiently small perturbations of the operator L. This is one of the claims of the Atiyah-Singer Index Theorem, which in addition to the invariance of the index of elliptic operators under sufficiently small perturbation, asserts that in the space of elliptic pseudodifferential operators, operators with a given index form connected components. As this second part of the Theorem is beyond the scope of this paper, we restrict our attention to proving the invariance of the index.

Section 2 contains a discussion of the constructions on flat space, i.e. Euclidean space, that we use to prove the main Theorem. Section 2.1 develops the necessary theory of Sobolev spaces. These function spaces, as we will make precise, provide a convenient mechanism for measuring the “amount of derivative” a function or function-like object (a distribution) has. In addition, they help classify these functions and distributions in a very useful way, in regards to the proof of the Theorem. Finally, Sobolev spaces and Sobolev norms capture the essential properties of elliptic operators that ensure invariance of the index. Section 2.1.1 discusses a number of properties of these so-called Sobolev spaces. Section 2.1.2 states and proves the Rellich Lemma—a statement about compact imbeddings of one Sobolev space into another. Section 2.1.3 relates these Sobolev spaces to elliptic operators by proving the basic elliptic estimate, one of the keys to the proof of the invariance of the index. Section 2.2 applies the machinery developed in 2.1 to conclude that elements of the kernel of an elliptic operator are smooth (in fact we conclude the local regularity of elliptic operators), and that the kernel is finite dimensional. This finite dimensionality is especially important, as it ensures that the “index” makes sense as a quantity.

The discussion in section 2 deals only with bounded open sets $\Omega \subset \mathbb{R}^n$. Section 3 generalizes the results of section 2 to compact Riemannian manifolds. Section 3.1 patches up the local constructions using partitions of unity. Section 3.2 deals with the primary differences and complications introduced by the local nature of compact manifolds and sections of vector bundles: section 3.2.1 discusses connections and covariant derivatives, and section 3.2.2 discusses the Riemannian metric and inner products. Finally section 3.3 combines the results of sections 2 and 3 to conclude the proof of the invariance of the index of an elliptic operator. The paper concludes with section 4 which discusses a concrete example of an elliptic differential operator on a compact manifold. A short Appendix includes the connection between the Index Theorem and the Riemann-Roch Theorem, and gives an alternative proof of Elliptic Regularity.

Example 1 As an illustration of the index of a linear operator, consider any linear map $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. By the Rank-Nullity Theorem, we know that $\text{index}(T) = n - m$. This is a rather trivial example, as the index of T depends only on the dimension of the range and domain, both of which are finite.

However when we consider infinite dimensional function spaces, Rank-Nullity no longer applies, and we have to rely on particular properties of elliptic operators, to which we now turn.

The general form of a linear differential operator L of order k is

$$L = \sum_{|\alpha| \leq k} a_\alpha(x) \partial^\alpha,$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index, and $|\alpha| = \sum_i \alpha_i$. In this paper we consider elliptic operators with smooth coefficients, i.e. with $a_\alpha \in \mathcal{C}^\infty$.

Definition 1 *A linear differential operator L of degree K is* **elliptic** *at a point x_0 if the polynomial*

$$P_{x_0}(\xi) := \sum_{|\alpha| = k} a_\alpha(x_0) \xi^\alpha,$$

is invertible except when $\xi = 0$.

This polynomial is known as the principal symbol of the elliptic operator. When we consider scalar valued functions, the polynomial is scalar valued, and hence the criterion for ellipticity is that the homogeneous polynomial $P_{x_0}(\xi)$ be non-vanishing at $\xi \neq 0$. There are very many often encountered elliptic operators, such as the following:

(i) $\bar{\partial} = \frac{1}{2}(\partial_x + i\partial_y)$, the Dirac operator on $\mathbb{C}$, also known as the Cauchy-Riemann operator. This operator is elliptic on all of $\mathbb{C}$ since the associated polynomial is $P_{\bar{\partial}}(\xi_1, \xi_2) = \xi_1 - i\xi_2$ which of course is nonzero for $\xi \neq 0$.

(ii) The Cauchy-Riemann operator is an example of a Dirac operator. Dirac operators in general are elliptic.

(iii) $\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$, the Laplace operator, is also elliptic, since the associated polynomial $P_\triangle(\xi_1, \xi_2) = \xi_1^2 + \xi_2^2$ is nonzero for $\xi \neq 0$ (recall that $\xi \in \mathbb{R}^2$ here).

It is a consequence of the basic theory of complex analysis that both operators described above have smooth kernel elements. As this paper shows, this holds in general for all elliptic operators. The Index Theorem asserts that when applied to spaces of sections of vector bundles over compact manifolds, these operators have a finite dimensional kernel and cokernel, and furthermore the difference of these two quantities, their index, is invariant under sufficiently small perturbations.

We now move to a development of the tools we use to prove the main Theorem.

2 Euclidean Space

Much of the analysis of manifolds and associated objects occurs locally, i.e. open sets of the manifold are viewed locally as bounded open sets in $\mathbb{R}^n$ via the appropriate local homeomorphisms, or charts. Because of this fact, many of the tools and methods we use for the main Theorem are essentially local constructions. For this reason in this section we develop various tools, and also properties of elliptic operators on bounded open sets of Euclidean space. At the beginning of section 3 we show that in fact these constructions and tools make sense, and are useful when viewed on a compact manifold.

2.1 Sobolev Spaces

A preliminary goal of this paper is to show that elliptic operators have smooth kernel elements, that is, if L is an elliptic operator, then the solutions to

$$Lu = 0,$$

are $\mathcal{C}^\infty$ functions. In fact, something stronger is true: elliptic operators can be thought of as "smoothness preserving" operators because, as we will soon make precise, if u satisfies $Lu = f$ then u turns out to be smoother then *a priori* necessary.

Example 2 A famous example of this is the Laplacian operator introduced above;

$$\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

While f need only have its first two derivatives for $\triangle f$ to make sense, if f is in the kernel of the operator, it is harmonic, and hence in $\mathcal{C}^\infty$.

Example 3 Consider the wave operator,

$$\Box = \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}.$$

The principal symbol of the wave operator is $P_\Box(\xi) = \xi_1^2 - \xi_2^2$ which vanishes for $\xi_1 = \xi_2$. Hence the wave operator, $\Box$, is not elliptic. Consider solutions to

$$\Box f = 0.$$

If $f(x,y)$ is such that $f(x,y) = g(x+y)$ for some g, then f satisfies the wave equation, however it need not be smooth.

There are then two immediate issues to consider: first, what if f above does not happen to have two continuous derivatives? That is to say, in general, if L has order k, but $u \notin \mathcal{C}^k$, then viewing u as a distribution, $u \in \mathcal{C}^{-\infty}$ we can understand the equation $Lu = f$ in this distributional sense. However given $Lu = f$ understood in this sense, what can we conclude about u? Secondly, we need some more convenient way to detect, or measure, the presence of higher derivatives. Fortunately, both of these issues are answered by the same construction: that of Sobolev spaces.

2.1.1 Definition of Sobolev Spaces

The main idea behind these function spaces is the fact that the Fourier transform is a unitary isomorphism on L^2 and it carries differentiation into multiplication by polynomials. We first define the family of function spaces H_k for $k \in \mathbb{Z}_{\geq 0}$—Sobolev spaces of nonnegative integer order—and then we discuss Sobolev spaces of arbitrary order—the so-called distribution spaces.
Nonnegative integer order Sobolev spaces are proper subspaces of L^2, and are defined by:

$$H_k = \{f \in L^2 \,|\, \partial^\alpha f \in L^2, \text{ where by } \partial^\alpha f \text{ we mean}$$
$$\text{the distributional derivative of } f\}.$$

We now use the duality of differentiation and multiplication by a polynomial, under the Fourier transform, to arrive at a more convenient characterization of these spaces.

Theorem 1 *A function $f \in L^2$ is in $H_k \subset L^2$ iff $(1+|\xi|^2)^{k/2}\hat{f}(\xi) \in L^2$. Furthermore, the two norms:*

$$f \longrightarrow \left[\sum_{|\alpha|\leq k} \| \partial^\alpha f \|^2_{L^2}\right]^{1/2} \quad \textit{and} \quad f \longrightarrow \left[\int |\hat{f}(\xi)|^2 (1+|\xi|^2)^k \, d\xi\right]^{1/2}$$

are equivalent.

Proof. This Theorem follows from two inequalities. We have:

$$\begin{aligned} (1+|\xi|^2)^k &\leq 2^k \max(1, |\xi|^{2k}) \\ |\xi|^{2k} &\leq C \sum_{j=1}^{n} |\xi_j^k|^2 \end{aligned}$$

where C is the reciprocal of the minimum value of $\sum_{j=1}^{n} |\xi_j^k|^2$ on $|\xi| = 1$. Putting this all together we find:

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

This, together with the fact that

$$h(|\xi|) = \frac{(1+|\xi|^2)^k}{\sum_{|\alpha|\leq k} |\xi^\alpha|^2},$$

is continuous away from zero, and tends to a constant as $|\xi| \to \infty$ concludes the proof. $\square$

Under this second equivalent definition, the integer constraint naturally imposed by the first definition disappears. This allows us to define Sobolev spaces H_s where $s \in \mathbb{R}$, and whose elements satisfy:

$$u \in H_s \iff (1+|\xi|^2)^{s/2}\hat{u}(\xi) \in L^2.$$

The elements of H_s are not necessarily proper functions, unless $s \geq 0$. However, note that for an object u as above, we know that for any Schwartz-class function $\phi \in \mathcal{S}$, we have $\phi u \in L^1$. This follows, since

$$\begin{aligned}\int |\phi u| &= \int |\phi(1+|\xi|^2)^{-s/2}| \cdot |u(1+|\xi|^2)^{s/2}| \\ &\leq \| \phi(1+|\xi|^2)^{-s/2} \|_{L^2} \cdot \| u \|_s < \infty.\end{aligned}$$

By defining the linear functional $T_u : \mathcal{C} \to \mathbb{C}$ by $T_u(\phi) = \int u\phi$ we can view u as an element of $\mathcal{S}'$, the space of tempered distributions, the dual space of $\mathcal{S}$, the Schwartz-class functions. Recall that a primary motivation for tempered distributions is to have a subspace of $(\mathcal{C}_c^\infty)^* = \mathcal{C}^{-\infty}$ on which we can apply the Fourier transform. Indeed, $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$, and we can define the general space H_s as a subset of $\mathcal{S}'$ as follows:

$$H_s = \left\{ f \in \mathcal{S}' \;\middle|\; \| f \|_s^2 := \int |\hat{f}(\xi)|^2 (1+|\xi|^2)^s \, d\xi < \infty \right\}.$$

From this definition we immediately have: $t \leq t' \Rightarrow H_{t'} \subset H_t$ since we know $\| \cdot \|_t \leq \| \cdot \|_{t'}$. Note also that H_s can be easily made into a Hilbert space by defining the inner product:

$$\langle f \,|\, g \rangle_s := \int \hat{f}(\xi)\overline{\hat{g}(\xi)}(1+|\xi|^2)^s \, d\xi.$$

Sobolev spaces can be especially useful because they are precisely related to the spaces $\mathcal{C}^k$. This is the content of the so-called Sobolev Embedding Theorem, whose proof we omit (see, e.g. Rudin [9] or Adams [1]):

Theorem 2 (Sobolev Embedding Theorem) *If $s > k + \frac{1}{2}n$, where n is the dimension of the underlying space $\mathbb{R}^n$, then $H_s \subset \mathcal{C}^k$ and we can find a constant $C_{s,k}$ such that*

$$\sup_{|\alpha|\le k} \sup_{x\in\mathbb{R}^n} |\partial^\alpha f(x)| \le C_{s,k} \parallel f \parallel_s .$$

Corollary 1 *If $u \in H_s$ for every $s \in \mathbb{R}$, then it must be that $u \in \mathcal{C}^\infty$.*

The Sobolev Embedding Theorem also gives us the following chain of inclusions:

$$\mathcal{S}' \supset \cdots \supset H_{-|s|} \supset \cdots \supset H_0 = L^2 \supset \cdots \supset H_{|s|} \supset \cdots \supset \mathcal{C}^\infty.$$

We have the following generalization of Theorem 1 above, which will prove very useful in helping us measure the "amount of derivative" a particular function has:

Theorem 3 *For $k \in \mathbb{N}$, $s \in \mathbb{R}$, and $f \in \mathcal{S}'$, we have $f \in H_s$ iff $\partial^\alpha f \in H_{s-k}$ when $|\alpha| \le k$. Furthermore,*

$$\parallel f \parallel_s \quad \textit{and} \quad \left[\sum_{|\alpha|\le k} \parallel \partial^\alpha f \parallel_{s-k}^2 \right]^{1/2},$$

are equivalent norms, and $|\alpha| \le k$ implies that $\partial^\alpha : H_s \to H_{s-k}$ is a bounded operator.

Hence we can consider elliptic operators as continuous mappings, with $L : \mathcal{S}' \to \mathcal{S}'$ in general, and $L : H_s \to H_{s-k}$ in particular.

Corollary 2 *If $u \in \mathcal{C}^{-\infty}$ and has compact support, then $u \in \mathcal{S}'$, and moreover $u \in H_s$ for some s.*

PROOF. If a distribution u has compact support, it must have finite order, that is, $\exists\, C, N$ such that

$$|T_u \phi| \le C \parallel \phi \parallel_{\mathcal{C}^N}, \quad \forall \phi \in \mathcal{C}_c^\infty.$$

Then we can write (as in, e.g. Rudin [9])

$$u = \sum_\beta D^\beta f_\beta,$$

where β is a multi-index, and the $\{f_\beta\}$ are continuous functions with compact support. But then $f_\beta \in \mathcal{C}_c$ and thus $f_\beta \in L^2 = H_0$. Therefore by Theorem 3, u is at least in $H_{-|\beta|}$. $\square$

We now list some more technical Lemmas which we use:

Lemma 1 *In the negative order Sobolev spaces (the result is obvious for $s \ge 0$) convergence in $\| \cdot \|_s$ implies the usual weak* distributional convergence.*

PROOF. We show, equivalently, that convergence with respect to $\| \cdot \|_s$ implies so-called strong distributional convergence, i.e. uniform convergence on compact sets. For $u_n, u \in H_s$ and $\| u_n - u \|_s \to 0$, and $\forall f \in \mathcal{S}$,

$$\begin{aligned} \Big| \int (u_n - u) f \Big| &= \Big| \int (\hat{u}_n - \hat{u}) * \hat{f} \Big| \\ &\leq \int |\hat{u}_n - \hat{u}| |\hat{f}|, \end{aligned}$$

by Plancherel, and then by Young. This yields

$$\begin{aligned} \int |\hat{u}_n - \hat{u}||\hat{f}| &= \int |(1+|\xi|^2)^s (\hat{u}_n - \hat{u})| \cdot |\hat{f}(1+|\xi|^2)^{-s}| \\ &\leq \| (1+|\xi|^2)^s (\hat{u}_n - \hat{u}) \|_{L^2} \cdot \| \hat{f}(1+|\xi|^2)^{-s} \|_{L^2} \\ &= \| u_n - u \|_s \cdot \| f \|_{|s|} \leq \| u_n - u \|_s \cdot \| f \|_k \quad (k \geq |s|) \\ &= \| u_n - u \|_s \cdot C \| f \|_{\mathcal{C}^k} \leq \varepsilon_n \cdot \| f \|_{\mathcal{C}^k}, \end{aligned}$$

where the last equality follows from Theorem 3, and $\varepsilon_n \to 0$. That strong convergence implies weak* convergence is straightforward. □

Lemma 2 *For $s \in \mathbb{R}$ and $\sigma > \frac{1}{2}n$, we can find a constant C that depends only on σ and s such that if $\phi \in \mathcal{S}$ and $f \in H_s$, then*

$$\| \phi f \|_s \leq \Big[\sup_x |\phi(x)| \Big] \| f \|_s + C \| \phi \|_{|s-1|+1+\sigma} \| f \|_{s-1} .$$

The following Lemma says that the notion of a localized Sobolev space makes sense. This is important, as we use such local Sobolev spaces in the proof of the local regularity of elliptic operators in section 2.2.

Lemma 3 *Multiplication by a smooth, rapidly decreasing function, is bounded on every H_s, i.e. for $\phi \in \mathcal{S}$, the map $f \mapsto \phi f$ is bounded on H_s for all $s \in \mathbb{R}$.*

Let $\Omega \subset \mathbb{R}^n$ be any domain with boundary. The localized Sobolev spaces contain the proper Sobolev spaces. We say that $u \in H_s^{\text{loc}}$ if and only if $\phi u \in H_s(\Omega)$ for all $\phi \in \mathcal{C}_c^\infty(\Omega)$, which is to say that the restriction of u to any open ball $B \subset \Omega$ with closure $\bar{B}$ in the interior of Ω, is in $H_s(B)$.

The proofs of both of these Lemmas are rather technical. The idea is to use powers of the operator

$$\Lambda^s = [I - (2\pi)^{-2} \triangle]^{s/2} \hat{f}(\xi),$$

and the fact that under the Fourier transform, the above becomes

$$(\Lambda^s f)\hat{}(\xi) = (1+|\xi|^2)^{s/2} \hat{f}(\xi).$$

2.1.2 The Rellich Lemma

As we saw above, from the definition of the Sobolev spaces we have the automatic inclusion $H_{t'} \subset H_t$ whenever $t \leq t'$. In fact, a much stronger result holds. Recall that if $t \leq t'$, the norm $\| \cdot \|_t$ is weaker, and hence admits more compact sets. The Rellich Lemma makes this precise.

Theorem 4 (Rellich Lemma) *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with smooth boundary* [1]*. If $t' > t$ then the embedding by the inclusion map $H_{t'}(\Omega) \hookrightarrow H_t(\Omega)$ is compact, i.e. every bounded sequence in $H_{t'}(\Omega)$ has a convergent subsequence when viewed as a sequence in $H_t(\Omega)$.*

An operator is called compact if it sends bounded sets to precompact sets. This is precisely the content of the second part of the theorem.

Proof. Take any bounded sequence $\{f_n\}$ in $H_{t'}$. We want to show that there is a convergent subsequence that converges to $f \in H_t$ for any $t < t'$. In fact, since the Sobolev spaces are Banach spaces, we need only show the existence of a Cauchy subsequence. Again we exploit the properties of the Fourier transform. By assumption, our domain $\Omega \subset \mathbb{R}^n$ is bounded. Then we can find a function $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ with $\phi \equiv 1$ on a neighborhood of $\bar{\Omega}$. Since the f_n are all supported on Ω, we can write $f_n = \phi f_n$ and therefore

$$\hat{f}_n(\xi) = (\phi f_n)\hat{}(\xi) \Rightarrow \hat{f}_n = \hat{\phi} * \hat{f}_n.$$

But since the Fourier transform takes Schwartz-class functions to Schwartz-class functions, i.e. $\mathcal{F} : \mathcal{S} \to \mathcal{S}$, $\hat{\phi} \in \mathcal{S}$ and therefore $\hat{\phi} * \hat{f}_n$ must be in $\mathcal{C}^\infty$. Then by the Cauchy-Schwarz inequality and some algebra, we find

$$(1 + |\xi|^2)^{t'/2} |\hat{f}_n(\xi)| \leq 2^{|t'|/2} \| \phi \|_{|t'|} \| f_n \|_{t'} .$$

But since $\hat{\phi}(\xi) \in \mathcal{S}$ so is $P(\xi) \cdot \hat{\phi}(\xi)$ for any polynomial $P(\xi)$. In particular, similarly to the above inequality we easily find that for $j = 1, \dots, n$,

$$(1 + |\xi|^2)^{t'/2} |\partial_j \hat{f}_n(\xi)| \leq 2^{|t'|/2} \| 2\pi i x_j \phi \|_{|t'|} \| f_n \|_{t'} .$$

Now by our boundedness assumption, we must have $\| f_n \|_{t'} \leq C_{t'}$ for all f_n. But then by the two equations above, the family $\{\hat{f}_n\}$ is equicontinuous. Since we are on a complete metric space, we can apply the Arzela-Ascoli Theorem, which asserts the existence of a convergent subsequence $\hat{f}_{k_n}$ which we rename to $\hat{f}_n$. By the Theorem, this subsequence converges uniformly on compact sets. In fact, more is true: f_n converges in $H_t(\Omega)$ for $t < t'$. To see this, take any

[1] In fact this Theorem holds for more general conditions. In particular, Ω need only have the so-called segment property. See Adams [1] for a full discussion.

$M > 0$. Then,

$$\begin{aligned}
\| f_n - f_m \|_t^2 &= \int_{|\xi|\le M} (1+|\xi|^2)^t |\hat{f}_n - \hat{f}_m|^2(\xi)\, d\xi \\
&\quad + \int_{|\xi|\ge M} (1+|\xi|^2)^{t-t'} (1+|\xi|^2)^{t'} |\hat{f}_n - \hat{f}_m|^2(\xi)\, d\xi \\
&\le \left[\sup_{|\xi|\le M} |\hat{f}_n - \hat{f}_m|^2(\xi) \right] \int_{|\xi|\le M} (1+|\xi|^2)^t \, d\xi \\
&\quad + (1+M^2)^{t-t'} \int_{|\xi|\ge M} (1+|\xi|^2)^{t'} |\hat{f}_n - \hat{f}_m|^2(\xi)\, d\xi \\
&\le \left[\sup_{|\xi|\le M} |\hat{f}_n - \hat{f}_m|^2(\xi) \right] \int_{|\xi|\le M} (1+|\xi|^2)^t \, d\xi \\
&\quad + (1+M^2)^{t-t'} \| f_n - f_m \|_{t'}^2 .
\end{aligned}$$

Now $t' > t$ strictly, implies that $t - t' < 0$. Therefore since $\| f_n - f_m \|_{t'}$ is bounded by $2C_{t'}$, the second term in the final expression becomes arbitrarily small as we let M get very large. Now the first term may also be made arbitrarily small by choosing m, n sufficiently large, for we know from Arzela-Ascoli that since $\{|\xi| \le M\}$ is compact,

$$\sup_{|\xi|\le M} |\hat{f}_n - \hat{f}_m|^2(\xi) \longrightarrow 0 \quad \text{as } m, n \to \infty.$$

Since the expression $\int_{|\xi|\le M} (1+|\xi|^2)^t \, d\xi$ is finite and moreover independent of m, n, that f_n is a Cauchy sequence in $H_t(\Omega)$ follows, concluding the Rellich Lemma. $\square$

2.1.3 Basic Sobolev Elliptic Estimate

In this section we discuss the main inequality that elliptic differential operators satisfy, and which we use to prove the local regularity of elliptic operators in section 2.2.1, and then to prove key steps in the main Theorem in section 3.3.

Recall the definition of an elliptic operator: A differential operator

$$L = \sum_{|\alpha|\le k} a_\alpha(x) \partial^\alpha,$$

where $a_\alpha \in \mathcal{C}^\infty$, is elliptic at a point x_0 if the polynomial

$$P_{x_0}(\xi) = \sum_{|\alpha|=k} a_\alpha(x_0) \xi^\alpha,$$

is invertible except where $\xi = 0$. Note that the polynomial $P_{x_0}(\xi)$ is homogeneous of degree k and therefore letting $A_{x_0} = \min_{|\xi|=1} \left| \sum_{|\alpha|\le k} a_\alpha(x_0) \xi^\alpha \right|$, we have the inequality

$$\left| \sum_{|\alpha|\leq k} a_\alpha(x_0)\xi^\alpha \right| \geq A_{x_0}|\xi|^k.$$

We say that L is elliptic on $\Omega \subset \mathbb{R}^n$ if it is elliptic at every point there. Note further that since we have $a_\alpha \in \mathcal{C}^\infty$, if L is elliptic on a compact set, then there is a constant A satisfying the above inequality for all points x_0. We are now ready to prove the main estimate.

Theorem 5 *If L is a differential operator of degree k, with coefficients $a_\alpha \in \mathcal{C}^\infty$, and is elliptic on a neighborhood of the closure of an open bounded set that has smooth boundary, $\bar{\Omega} \subset \mathbb{R}^n$, then for all $s \in \mathbb{R}$ there exists a constant $C > 0$ such that for any element $u \in H_s(\Omega)$ with compact support, u satisfies:*

$$\| u \|_s \leq C(\| Lu \|_{s-k} + \| u \|_{s-1}).$$

Proof. Following Folland's development, we prove this Theorem in three steps:

(i) We assume that a_α are constant, and zero for $|\alpha| < k$;

(ii) We drop the assumption on the constant coefficients a_α;

(iii) Finally we prove the general case.

Thus first assume we have

$$Lu = \sum_{|\alpha|=k} a_\alpha \partial^\alpha u.$$

Taking the Fourier transform and using the duality of differentiation and multiplication by polynomials we have:

$$\widehat{(Lu)}(\xi) = (2\pi i)^k \sum_{|\alpha|=k} a_\alpha \xi^\alpha \hat{u}(\xi).$$

Then with some algebraic manipulation we have:

$$\begin{aligned}
(1+|\xi|^2)^s|\hat{u}(\xi)|^2 &= (1+|\xi|^2)^{s-k}(1+|\xi|^2)^k|\hat{u}(\xi)|^2 \\
&\leq 2^k((1+|\xi|^2)^{s-k}|\hat{u}(\xi)|^2 + 2^k|\xi|^{2k}(1+|\xi|^2)^{s-k}|\hat{u}(\xi)|^2 \\
&\leq 2^k((1+|\xi|^2)^{s-k}|\hat{u}(\xi)|^2 + 2^k A^{-2}(1+|\xi|^2)^{s-k}|\widehat{Lu}(\xi)|^2.
\end{aligned}$$

The second inequality follows because if the a_α are constant, surely we can choose some A independent of x_0 such that $\left|\sum_{|\alpha|\leq k} a_\alpha(x_0)\xi^\alpha\right| \geq A_{x_0}|\xi|^k$, i.e. such that the above holds. Now integrating both sides yields:

$$\begin{aligned}
\| u \|_s^2 &\leq 2^k \| u \|_{s-k}^2 + 2^k A^{-2} \| Lu \|_{s-k}^2 \\
&\leq 2^k(A^{-2} \| Lu \|_{s-k}^2 + \| u \|_{s-1}^2),
\end{aligned}$$

and finally for a proper choice of constant, $C_0 = 2^{k/2}\max(A^{-1}, 1)$, we have the desired inequality:

$$\| u \|_s \le C_0(\| Lu \|_{s-k} + \| u \|_{s-1}).$$

For the second step, we still assume that the lower order coefficients of the operator are zero, but the highest order terms are not restricted to be constants. The idea behind the proof is to first look at distributions u supported locally in a small δ-neighborhood of a point x_0, and to show that the desired inequality holds by comparing the operator L with the constant coefficient operators $L_{x_0} := \sum_{|\alpha|=k} a_\alpha(x_0)\partial^\alpha$, i.e. operators which satisfy the inequality of the Theorem by step 1 above. After this, we use the fact that closed and bounded implies compact in $\mathbb{R}^n$ (Heine-Borel) to choose a finite number of these δ-neighborhoods around points $\{x_1, \ldots, x_N\}$ to cover $\overline{\Omega}$. Finally, we use a partition of unity subordinate to this covering to show that in fact the inequality holds for a general $u \in H_s(\Omega)$. Now for the details. By step 1 above we have the inequality:

$$\| u \|_s \le C_0(\| L_{x_0}u \|_{s-k} + \| u \|_{s-1}),$$

for L_{x_0} as above. Since the coefficients are smooth, we expect that in a small neighborhood of any point x_0, the constant coefficient operator L_{x_0} does not differ much from the original operator L. If we write any distribution u as $u = \sum_{i=1}^N \zeta_i u$ for $\{\zeta_i\}$ a partition of unity subordinate to some finite open cover, we will be able to take advantage of this local "closeness" of L and L_{x_0}. We must first estimate this "closeness":

$$\| Lu - L_{x_0}u \|_{s-k} = \Big\| \sum_{|\alpha|=k} [a_\alpha(\cdot) - a_\alpha(x_0)]\partial^\alpha u \Big\|_{s-k}.$$

Note that since Ω is a bounded set, we can assume without loss of generality that the coefficient functions $a_\alpha(x)$ actually have compact support. Then there exists a constant $C_1 > 0$ such that

$$|a_\alpha(x) - a_\alpha(x_0)| \le C_1|x - x_0| \quad (|\alpha| = k,\ x \in \mathbb{R}^n,\ x_0 \in \Omega).$$

Choose $\delta = (4(2\pi n)^k C_0 C_1)^{-1}$, for C_0, C_1 as defined above. Also choose some $\phi \in \mathcal{C}_c^\infty(B_{2\delta}(0))$ such that $0 \le \phi \le 1$ and $\phi \equiv 1$ on $B_\delta(0)$, and some ζ supported on $B_\delta(x_0)$ for some $x_0 \in \Omega$. Using this, and the well chosen constant δ above, we have:

$$\sup_x |\phi(x - x_0)[a_\alpha(x) - a_\alpha(x_0)]| \le C_1(2\delta) = \frac{1}{2(2\pi n)^k C_0},$$

and hence using Lemma 2 and Theorem 3 above, we have for any x,

$$\begin{aligned}
\| [a_\alpha(x) - a_\alpha(x_0)]\partial^\alpha(\zeta u) \|_{s-k} &= \| \phi(x - x_0)[a_\alpha(x) - a_\alpha(x_0)]\partial^\alpha(\zeta u) \|_{s-k} \\
&\le \frac{1}{2(2\pi n)^k C_0} \| \partial^\alpha(\zeta u) \|_{s-k} + C_2 \| \partial^\alpha(\zeta u) \|_{s-k-1} \\
&\le \frac{1}{2n^k C_0} \| \zeta u \|_s + (2\pi)^k C_2 \| \zeta u \|_{s-1},
\end{aligned}$$

where C_2 depends only on $\| \phi(x - x_0)[a_\alpha(x) - a_\alpha(x_0)] \|_{|s-k-1|+n+1}$ and in particular, does not depend on x_0. Now since we are working in $\mathbb{R}^n$, and $|\alpha| = k$ there are at most n^k multi-indices α, and therefore we have,

$$\begin{aligned}
\| L(\zeta u) - L_{x_0}(\zeta u) \|_{s-k} &\leq \sum_{|\alpha|\leq k} \| [a_\alpha(x) - a_\alpha(x_0)]\partial^\alpha(\zeta u) \|_{s-k} \\
&\leq \frac{1}{2C_0} \| \zeta u \|_s + (2\pi n)^k C_2 \| \zeta u \|_{s-1} .
\end{aligned}$$

Then by the good old triangle inequality and also step 1, we have:

$$\begin{aligned}
\| \zeta u \|_s &\leq C_0(\| L(\zeta u) \|_{s-k} + \| L(\zeta u) - L_{x_0}(\zeta u) \|_{s-k} + \| \zeta u \|_{s-1}) \\
&\leq C_0 \| L(\zeta u) \|_{s-k} + \frac{1}{2} \| \zeta u \|_s + [(2\pi n)^k C_2 + 1]C_0 \| \zeta u \|_{s-1},
\end{aligned}$$

and then taking $C_3 = 2[(2\pi n)^k C_2 + 1]C_0$ (which thanks to the above development is independent of x_0) we have

$$\| \zeta u \|_s \leq C_3(\| L(\zeta u) \|_{s-k} + \| \zeta u \|_{s-1}).$$

But now we are almost done. For since $\overline{\Omega} \subset \mathbb{R}^n$ is compact, it is totally bounded, and hence can be covered by a finite number of δ-balls $B_\delta(x_1), \ldots, B_\delta(x_N)$ with $x_i \in \Omega$. Then if we take a partition of unity $\{\zeta_i\}$ subordinate to this cover, we have for any $u \in H_s(\Omega)$

$$\begin{aligned}
\| u \|_s &= \Big\| \sum_1^N \zeta_i u \Big\|_s \leq \sum_1^N \| \zeta_i u \|_s \\
&\leq C_3 \sum_1^N (\| L(\zeta_i u) \|_{s-k} + \| \zeta_i u \|_{s-1}) \\
&= C_3 \sum_1^N (\| \zeta_i L u \|_{s-k} + \| [L, \zeta_i] u \|_{s-k} + \| \zeta_i u \|_{s-1}) \\
&\leq C_4(\| L u \|_{s-k} + \| u \|_{s-1}),
\end{aligned}$$

as desired. Note that in the third line above $[\cdot\,,\cdot]$ denotes the usual commutator operator, defined by $[A, B] = AB - BA$. The final inequality follows from the fact that if L is a differential operator of order k, ζ_i a smooth function, then $[L, \zeta_i]$ is an operator of degree $k - 1$.

We are now finally ready to prove the general case. Then suppose L is an elliptic operator of degree k. We can write $L = L_0 + L_1$ where we have

$$L_0 = \sum_{|\alpha|=k} a_\alpha(x)\partial^\alpha, \quad L_1 = \sum_{|\alpha|<k} a_\alpha(x)\partial^\alpha.$$

Note that L_1, while it need not be elliptic, is an operator of degree at most $k - 1$. Then by assuming again that its coefficients have compact support, we can apply Lemma 3 and Theorem 3 from above, to get:

$$\| L_1 u \|_{s-k} \leq C_5 \| u \|_{s-1} .$$

Since step 2 applies to L_0, we have:

$$\begin{aligned}
\| u \|_s &\leq C_4(\| L_0 u + L_1 u - L_1 u \|_{s-k} + \| u \|_{s-1}) \\
&\leq C_4(\| (L_0 + L_1) u \|_{s-k} + \| L_1 u \|_{s-k} + \| u \|_{s-1}) \\
&\leq C_4(C_5 + 1)(\| Lu \|_{s-k} + \| u \|_{s-1}),
\end{aligned}$$

which completes the proof. □

2.2 Elliptic Operators

Armed with the above inequality, we are ready to prove some of the mapping properties of elliptic operators. In particular, we prove the local regularity of elliptic operators, and the the finite dimensionality of the kernel and cokernel of elliptic operators. First we prove local regularity.

2.2.1 Local Regularity of Elliptic Operators

The goal is to show that elliptic operators in general possess some "smoothness preserving" properties, as do the Laplace and Cauchy-Riemann operators which are elliptic. In this section we take a pointwise approach. For an alternative proof emphasizing the "smoothing" properties of elliptic operators, see section B in the Appendix. We prove this in two steps, proving first a Lemma and then the Theorem. This is where the Sobolev machinery is especially helpful, as we are exactly trying to "measure" the amount of derivative a function has. Before we go on to prove the regularity of elliptic operators, we need to define one more "derivative measuring" tool to go along with the Sobolev spaces: Difference Quotients (a method due to Nirenberg [7]). Difference quotients are essentially approximations to a function's partial derivatives, and they provide a mechanism for determining when $\partial f \in H_s$ when all we know *a priori* is that $f \in H_s$.

Definition 2 *If f is a distribution, we define the family of distributions $\triangle_h^i f$ by*

$$\triangle_h^i f = \frac{1}{h}(f_{he_i} - f),$$

where f_{he_i} is defined as the translation of f by he_i (and of course the translation is defined in the distributional sense: $\langle f_x , \phi \rangle = \langle f_x , \phi_{-x} \rangle$) where e_i denotes an element of the standard basis for $\mathbb{R}^n$. The following Theorem gives a necessary and sufficient condition for $\partial f \in H_s$.

Theorem 6 *Suppose $f \in H_s$ for some $s \in \mathbb{R}$. Then*

$$\| \partial_i f \|_s = \limsup_{h \to 0} \| \triangle_h^i f \|_s .$$

In particular, $\partial_i f \in H_s$ iff $\triangle_h^i f$ remains bounded as $h \to 0$.

Proof. Recall that multiplication by a rotation is the dual of translation under Fourier transform. Therefore we have

$$\begin{aligned}(\triangle_h^i f)\hat{}(\xi) &= \frac{1}{h}(\hat{f}_{he_i} - \hat{f}) \\ &= \frac{1}{h}(e^{2\pi i h\xi_i}\hat{f} - \hat{f}) \\ &= 2ie^{\pi i h\xi_i}\frac{\sin \pi h\xi_i}{h}\hat{f}(\xi).\end{aligned}$$

Now recall that $\| u \|_s^2 = \int |\hat{u}|^2(1+|\xi|^2)^s \, d\xi$, and therefore we have

$$\begin{aligned}\| \triangle_h^i f \|_s &= \int (1+|\xi|^2)^s (|\widehat{\triangle_h^i f}(\xi)|^2)\, d\xi \\ &\leq \int (1+|\xi|^2)^s |2\pi\xi_i \hat{f}(\xi)|^2 \, d\xi \\ &= \| \partial_i f \|_s^2 \,.\end{aligned}$$

The last inequality comes from the fact that $|\sin x| \leq |x|$. Note that if $\| \partial_i f \|$ is finite, then we can apply Lebesgue's Dominated Convergence Theorem (using $\sin ax/a \to x$) to get equality in the last inequality above, therefore yielding

$$\limsup_{h\to 0} \| \triangle_h^i f \|_s^2 \leq \| \partial_i f \|_s^2,$$

with equality if $\| \partial_i f \|_s^2 < \infty$.
Conversely, suppose $\| \partial_i f \|_s^2 = \infty$. Then for any N, we can find some M such that the truncated integral over $[-M, M]$ is greater than $2N$. But since

$$\frac{\sin ax}{a} \longrightarrow x \quad \text{as } a \to 0$$

we can find some h sufficiently small such that $h' < h$ implies that $\| \triangle_h^i f \|_s^2 > N$, and is hence unbounded as h goes to zero, completing the proof. $\square$

This is the main Theorem about difference quotients, which explains why they are useful for our present needs. We state without proof two other results about these difference quotients:

Lemma 4 *If $s \in \mathbb{R}$ and $\phi \in \mathcal{S}$, then the operator $[\triangle_h^i, \phi]$, defined by the usual commutator operation $[A, B] := AB - BA$, is bounded from $H_s \to H_s$ with bound independent of h.*

Corollary 3 *If L is a linear differential operator of order k, then $[\triangle_h^i, L]$ is a bounded operator from $H_s \to H_{s-k}$, with bound independent of h.*

Now we are ready to prove the regularity of elliptic operators.

Theorem 7 *If $\Omega \subset \mathbb{R}^n$ is an open bounded set, L is an elliptic differential operator of order k with $\mathcal{C}^\infty$ coefficients, and if $u \in H_s^{loc}(\Omega)$ and $Lu \in H_{s-k+1}^{loc}(\Omega)$, then $u \in H_{s+1}^{loc}(\Omega)$.*

PROOF. From our definition of the spaces $H_s^{\text{loc}}(\Omega)$, we know that $u \in H_{s+1}^{\text{loc}}(\Omega)$ iff $\phi u \in H_{s+1}$ for all $\phi \in \mathcal{C}_c^\infty(\Omega)$. By assumption, $u \in H_s^{\text{loc}}(\Omega)$ and $Lu \in H_{s-k+1}^{\text{loc}}(\Omega)$, and therefore we must have

$$L(\phi u) = \phi L u + [L, \phi] u \in H_{s-k+1},$$

because as we have already seen, $[L, \phi]$ is an operator of degree at most $k-1$, and hence we can apply Theorem 1 and Lemma 3 above. Then by Corollary 3 above, and the basic Sobolev elliptic estimate (Theorem 5), we have:

$$\begin{aligned}
\| \triangle_h^i(\phi u) \|_s &\leq C(\| L\triangle_h^i(\phi u) \|_{s-k} + \| \triangle_h^i(\phi u) \|_{s-1}) \\
&\leq C(\| \triangle_h^i L(\phi u) \|_{s-k} + \| [L, \triangle_h^i](\phi u) \|_{s-k} + \| \triangle_h^i(\phi u) \|_{s-1}) \\
&\leq C(\| \triangle_h^i L(\phi u) \|_{s-k} + C' \| \phi u \|_s + \| \triangle_h^i(\phi u) \|_{s-1}),
\end{aligned}$$

where the second inequality above follows by the triangle inequality, and the third by Corollary 3. Now note that since we already established $L(\phi u) \in H_{s-k+1}$, and $\phi u \in H_s$ by assumption, their respective Sobolev norms are finite. Then by Theorem 3, $\| \partial_i L(\phi u) \|_{s-k} < \infty$ and $\| \partial_i(\phi u) \|_{s-1} < \infty$. But then by Theorem 6, the right hand side of the last inequality above must be bounded independently of h as $h \to 0$, and therefore the lefthand side is bounded as $h \to 0$. Applying Theorem 6 again, we find that $\| \partial_j(\phi u) \|_s$ must be bounded, and hence $\phi u \in H_{s+1}$. Since ϕ was arbitrary, we have $u \in H_{s+1}^{\text{loc}}$ as required. $\square$

Theorem 8 *Suppose Ω, L are as above, and u, f are distributions such that $Lu = f$. If $f \in H_s^{loc}(\Omega)$ for some $s \in \mathbb{R}$, then $u \in H_{s+k}^{loc}(\Omega)$.*

PROOF. This proof is essentially a repeated application of the previous Theorem. Again, to conclude that $u \in H_{s+k}^{\text{loc}}(\Omega)$ we must show that $\forall \phi \in \mathcal{C}_c^\infty$, we have $\phi u \in H_{s+k}$. Then choose some $\phi \in \mathcal{C}_c^\infty$. Now choose a function $\phi_0 \in \mathcal{C}_c^\infty$ such that $\phi_0 \equiv 1$ on a neighorhood of $\text{supp}(\phi)$. As a Corollary to the Sobolev Embedding Theorem (Theorem 2) and Theorem 3, we know that any distribution with compact support is an element of $H_{t'}$ for some $t' \in \mathbb{R}$. Then $\phi_0 u \in H_{t'}$ for some t'. Since $H_t \supset H_{t'}$ for every $t \leq t'$, we can find some $t \leq t'$ such that $\phi_0 u \in H_t$ and $N = s + k - t \in \mathbb{N}$. We have chosen ϕ_0. We now choose $\phi_1, \ldots, \phi_N$. Note that $\text{supp}(\phi) \subsetneq \text{supp}(\phi_0)$. Then, we set $\phi_N = \phi$. We define the other functions as follows: take $\phi_1 \in \mathcal{C}_c^\infty$ such that $\phi_1 \equiv 1$ on a neighborhood of $\text{supp}(\phi)$, and such that ϕ_1 is supported in the set where $\phi_0 \equiv 1$. Similarly, take $\phi_i \in \mathcal{C}_c^\infty$ such that $\phi_i \equiv 1$ on a neighborhood of $\text{supp}(\phi)$, and $\text{supp}(\phi_i) \subset \{x \mid \phi_{i-1}(x) = 1\}$. We will show that $\phi_j u \in H_{t+j}$, and hence that $\phi u = \phi_N u \in H_{t+N} = H_{s+k}$ as required.

The proof of this is by induction. The base case is trivial since $\phi_0 u \in H_t$ by assumption. Then assume that $\phi_j u \in H_{t+j}$. Consider $\phi_{j+1} u$. Since $\phi_j \equiv 1$ on the support of ϕ_{j+1}, we have

$$\phi_{j+1} u = \phi_{j+1} \phi_j u,$$

and since $\phi_j u \in H_{t+j}$ by inductive assumption, we must also have $\phi_{j+1} u \in H_{t+j}$. Furthermore, we must also have

$$L(\phi_j u) = Lu = f \quad \text{on the support of } \phi_{j+1}.$$

This yields:

$$\begin{aligned} L(\phi_{j+1}u) &= L(\phi_{j+1}\phi_j u) \\ &= \phi_{j+1} L(\phi_j u) + [L, \phi_{j+1}](\phi_j u) \\ &= \phi_{j+1} f + [L, \phi_{j+1}](\phi_j u). \end{aligned}$$

Now, $[L, \phi_{j+1}](\phi_j u) \in H_{t+j-k+1}$ because $[L, \phi_{j+1}]$ is an operator of order at most $k-1$. Meanwhile, $\phi_{j+1} f \in H_s$ by assumption. But then we have

$$L(\phi_{j+1}u) \in H_{t+j-k+1}, \quad \text{and} \quad \phi_{j+1}u \in H_{t+j}.$$

But now we can apply the previous Theorem to conclude that in fact we must have:

$$\phi_{j+1}u \in H_{t+j+1} \Rightarrow \phi_N u = \phi u \in H_{s+k} \Rightarrow u \in H_{s+k}^{\text{loc}},$$

concluding the proof. $\square$

We have proved something considerably stronger than the fact that the elements of the kernel of an elliptic operator are smooth. In fact, our result quickly implies the smoothness of the elements of the kernel. For if u is in the kernel, it satisfies $Lu = 0$. Since $0 \in \mathcal{C}^k$ for any k, then we also have $u \in H_s$ for all s, which implies that $u \in \mathcal{C}^\infty$, as claimed.

2.2.2 Kernel and Cokernel of Elliptic Operators

In this section we show that essentially as a consequence of the basic Sobolev elliptic estimate, elliptic operators on compact spaces must have finite dimensional kernel and cokernel, and also have closed range, i.e. they are Fredholm. While we have not yet discussed compact manifolds, we see in section 3 that while the work done in section 2 carries over easily, the global versus local nature of the manifold and the individual choices of coordinate neighborhood introduce various complications. We postpone the discussion to section 3, and we prove the above statements for compact sets in $\mathbb{R}^n$.

As a preliminary step, we verify that the notion of kernel makes sense independently of the Sobolev norm being used.

Proposition 1 *If $f \in H_s$ and $\| f \|_{L^2} = 0$, then $\| f \|_s = 0$.*

This is an immediate consequence of the definition of $\| \cdot \|_s$:

$$\begin{aligned}
\| f \|_{L^2} = 0 \quad &\Rightarrow \quad \int |f|^2 = 0 \Rightarrow \int |f| = 0 \\
&\Rightarrow \quad |\hat{f}(\xi)| = \Big| \int f(x) e^{ix\xi}\, dx \Big| \le \int |f(x)|\, dx = 0 \\
&\Rightarrow \quad \| f \|_s = \int (1+|\xi|^2)^s |\hat{f}(\xi)|^2\, d\xi \; = \; 0.
\end{aligned}$$

Theorem 9 *If L is an elliptic operator on a compact set $\bar{\Omega} \subset \mathbb{R}^n$, then the dimension of the space of distributions in the kernel of L is finite.*

Proof. Recall the basic Sobolev elliptic estimate of section 2.1.3:

$$\| u \|_s \le C(\| Lu \|_{s-k} + \| u \|_{s-1}).$$

Note that by the Regularity Theorem, we are considering positive order Sobolev spaces, which are subsets of L^2. Since L^2 is a Hilbert space, if kernel(L) is infinite dimensional, we can take an infinite family of orthonormal functions in the kernel, say

$$S = \{u_1, u_2, \dots\}.$$

For $u \in \mathrm{Span}(S)$ we have $Lu = 0$ and hence the elliptic inequality above becomes

$$\| u \|_s \le \; C \; \| u \|_{s-1} \; .$$

But this means that if the $\{u_n\}$ are normalized in $L^2 = H_0$, then they are bounded in H_k for any $k \in \mathbb{N}$, and in particular they are bounded in H_s for some $s > 0$. But then by the Rellich Lemma, the infinite sequence is compact in L^2, and therefore contains a convergent subsequence, contradicting the assumed orthonormality of the sequence. Alternatively, by the basic Sobolev elliptic estimate and the Rellich Lemma, the kernel of L is locally compact, and hence finite dimensional. But in fact we do not have to rely on something as powerful as the Rellich Lemma. For the inequality $\| u \|_s \le C \; \| u \|_{s-1}$ combined with Theorem 3 asserts that

$$\| \nabla u_i \|_{L^2} \le M, \; \forall n,$$

for some M, hence the family is equicontinuous and we can apply the Ascoli-Arzela Theorem to conclude the same contradiction. In either case the contradiction proves that the kernel of the elliptic operator is finite dimensional. $\square$

We now would like to prove a similar fact about the cokernel of any elliptic operator L. The first result proved below gives a convenient representation of the cokernel of L in terms of the kernel of the adjoint. Implicit in any discussion about cokernel and adjoint, lies the issue of which inner product to choose. For a general elliptic operator L of degree N, we have $L : H_{s+N} \to H_s$, while its adjoint maps $L^* : H_s \to H_{s+N}$. Then the adjoint operator would be defined by the relation:

$$\langle \, Lf \, , \, g \, \rangle_{H_s} = \langle \, f \, , \, L^* g \, \rangle_{H_{s+N}}.$$

We denote this adjoint by L^*, and the adjoint defined by the usual L^2 inner product by $L^\dagger$. Since we care only about the kernel of the adjoint, we can avoid such formalism and use the L^2 inner-product and hence the L^2 adjoint $L^\dagger$, throughout, because by the Elliptic Regularity Theorem (Theorem 8) the elements of the kernel are smooth, and they have compact support. This is the content of the following Proposition.

Proposition 2 *If $\eta \in L^2$ and $L^*\eta = 0$ then $\eta \in \mathcal{C}^\infty$ and $L^\dagger\eta = 0$, and conversely.*

Proof. The adjoints L^* and $L^\dagger$ are both defined distributionally. Therefore it does not make sense, *a priori* to use the L^2 adjoint $L^\dagger$ on the entire domain of L^*. However, if L is an operator of degree k, the two adjoints are related by

$$L^* = \frac{1}{(1-\triangle)^k} L^\dagger.$$

Therefore L^* is elliptic iff $L^\dagger$ is. Therefore by elliptic regularity the L^2 adjoint is defined on any element of the kernel of L^*. Moreover, taking Fourier transforms we have

$$\begin{aligned} L^*\eta = 0 \quad &\Rightarrow \quad \frac{1}{(1-\triangle)^k} L^\dagger \eta = 0 \\ &\Rightarrow \quad \frac{1}{(1+|\xi|^2)^k} \hat{L^\dagger}\hat{\eta} = 0 \\ &\Rightarrow \quad L^\dagger \eta = 0, \end{aligned}$$

and therefore η is in the kernel of $L^\dagger$ if it is in the kernel of L^*. The converse holds similarly. □

Therefore we are justified in using the L^2 adjoint throughout. From now on we use $*$ to denote an operator's adjoint. This having been said, however, we prove the next result in the most general context of a linear operator mapping between Hilbert spaces $\mathcal{D}$ and $\mathcal{R}$.

Proposition 3 *If $L : \mathcal{D} \longrightarrow \mathcal{R}$ is a linear differential operator with closed range, then cokernel$(L) \cong$ kernel(L^*).*

Proof. We will show that

$$\mathrm{kernel}(L^*) = L(\mathcal{D})^\perp \cong \mathcal{R}/L(\mathcal{D}) = \mathrm{cokernel}(L).$$

We know that if $S \subset \mathcal{R}$ then

$$S^\perp := \{v \in \mathcal{R} \,|\, \langle\, s\, ,\, v\, \rangle = 0, \quad \forall s \in S\},$$

is a closed linear manifold. Furthermore, $\left(S^\perp\right)^\perp$ is the smallest closed linear manifold containing S. Then since $L(\mathcal{D})$ is closed by assumption, $\left(L(\mathcal{D})^\perp\right)^\perp = L(\mathcal{D})$. In particular, we have

$$\mathcal{R} = L(\mathcal{D}) \oplus L(\mathcal{D})^\perp.$$

This implies that the projection

$$\pi : L(\mathcal{D})^{\perp} \longrightarrow \mathcal{R}/L(\mathcal{D}),$$

is surjective. Since $L(\mathcal{D}) \cap L(\mathcal{D})^{\perp} = \{0\}$, the projection is also injective, and therefore it is an isomorphism. To show kernel$(L^*) = L(\mathcal{D})^{\perp}$, take $v \in L(\mathcal{D})^{\perp}$. Then by definition of perpendicular space, $\langle\, Lu\,,\, v\,\rangle = 0$ for all $u \in \mathcal{D}$, in other words $\langle\, u\,,\, L^*v\,\rangle = 0$ for all $u \in \mathcal{D}$. But $L^*v \in \mathcal{D}$, and the only element of $\mathcal{D}$ that is orthogonal to everything is 0. On the other hand, if we have $v \in$ kernel(L^*), then $L^*v = 0$, and hence

$$\langle\, u\,,\, L^*v\,\rangle = \langle\, Lu\,,\, v\,\rangle = 0, \quad \forall u\, \in \mathcal{D},$$

which implies $v \in L(\mathcal{D})^{\perp}$, concluding the proof. □

We now prove the missing link to the Proposition above.

Lemma 5 *If L is an elliptic operator, $L : H_s \longrightarrow H_{s-k}$, then it has closed range.*

We specify the Sobolev spaces in order to fix the norms, and hence the notion of convergence and closure.

PROOF. We want to show that if $\{g_i = Lu_i\}$ is a convergent sequence, then it converges to some $g = Lu$. Since the kernel is a closed linear manifold, we can write

$$\mathcal{D} = \text{kernel}(L)^{\perp} \oplus \text{kernel}(L) = V^{\perp} \oplus V.$$

Then for any element $u \in \mathcal{D}$ we can write $u = \pi_{V^\perp}u + \pi_V u$ uniquely. Therefore $Lu = L(\pi_{V^\perp}u + \pi_V u)$. We need to show that if Lu_n is a convergent sequence, then $L(\pi_{V^\perp}u_n)$ and $L(\pi_V u_n)$ both converge to some $L(\pi_{V^\perp}u)$ and $L(\pi_V u)$ respectively. The second is clear, for $L(\pi_V u_n)$ converges to $L(\pi_V u) = 0$ since for any $u \in \mathcal{D}$, we have $L(\pi_V u) = 0$. Showing the first is somewhat more tricky. Since H_s is a Hilbert space, it is in particular, a Banach space, and thus it is enough to show that if $\{Lu_i\}$ is Cauchy, then so is $\{u_i\}$. This will follow if we can show that

$$\| \, Lu \, \|_{s-k} \geq K \, \| \, u \, \|_s \, .$$

Suppose the inequality does not hold for any constant K. Then we can find a sequence $\{u_n\} \in H_s$ such that

$$\begin{aligned} & \| \, Lu_n \, \|_{s-k} < \frac{1}{n} \, \| \, u_n \, \|_s \\ \Leftrightarrow \quad & \| \, L\tilde{u}_n \, \|_{s-k} < \frac{1}{n}, \end{aligned}$$

where

$$\tilde{u}_n = \frac{u_n}{\| \, u_n \, \|_s} \in S = \{u \, \in \, H_s \mid \, \| \, u \, \|_s = 1\}.$$

Thus we can find a sequence $\{u_n\}$ in S with $\| Lu_n \|_{s-k} \to 0$. Now $S \subset H_s$ is bounded, and therefore by the Rellich Theorem (Theorem 4) $S \subset H_{s-1}$ must be compact. Then we can find some subsequence u_{n_j} that converges to some $u \in H_{s-1}$. Then by the basic elliptic estimate we have

$$1 = \| u_{n_j} \|_s \leq C(\| Lu_{n_j} \|_{s-k} + \| u_{n_j} \|_{s-1}).$$

But the right hand side converges to $\| u \|_{s-1}$ and therefore $\| u \|_{s-1} > C^{-1}$ and therefore $u \neq 0$, contradicting the assumption that the kernel is trivial. Therefore there must indeed be some constant K for which the inequality

$$\| Lu \|_{s-k} \geq K \| u \|_s,$$

holds, completing the proof. □

This concludes the proof of finite dimensionality of the cokernel as well as the kernel. For while we have not discussed in detail the adjoint operator, the next section shows that if L is elliptic, then so is L^*. The idea is that ellipticity is only a condition on the highest order terms of the operator, and the adjoint of these highest order terms is a nonvanishing multiple of them, and hence L^* is elliptic iff L is. This all depends upon the analogue of the above Theorems and Definitions to compact manifolds with coordinate charts. To this we now turn.

3 Compact Manifolds

We remark that already there appears a deficiency in the discussion up to this point. For the *compactness* of the space where the functions are defined is crucial for the proof to work—indeed consider the harmonic functions on $\mathbb{C}$. They are the kernel of an elliptic operator, and are certainly not finite dimensional. However, the majority of our discussion has been about bounded open sets in $\mathbb{R}^n$. The point is that we want to apply the above theorems to compact manifolds, not just to bounded sets in Euclidean space.

Then in this section, we discuss the application of the above techniques to spaces of sections of vector bundles over compact manifolds. In addition, we must discuss the analogues of various concepts from Euclidean space, in the context of compact manifolds. These are, in particular, differentiation, which gives the appropriate form of the differential operator, and integration, which provides a norm, an inner product, and hence an operator's adjoint.

3.1 Patching Up the Local Constructions

Our definition of elliptic operators is pointwise, and thus immediately carries over to compact manifolds. Our definition of Sobolev spaces depends upon the definition of integration and differentiation, however as soon as these are defined, there is nothing local about the definition of Sobolev spaces. Then we need only show that the two tools we developed in section 2, namely the Rellich Lemma and the basic Sobolev elliptic estimate, hold for compact manifolds.

The Rellich Lemma is the easier of the two to adapt. The Rellich Lemma is a theorem about the compactness of the embedding operator: $H_{t'} \to H_t$ for $t' > t$. Consider any sequence of functions $\{f_n\} \in H_{t'}$, and any partition of unity $\{\zeta_i\}_{i=1}^m$ subordinate to a finite cover of bounded sets $\{U_i\}$ of the manifold M^n. Then the Rellich Lemma as proved in section 2.1.2 above holds for $\zeta_i f_n$ for each i. Therefore by passing to a subsequence once for each i we conclude that it holds for compact manifolds.

We now prove that the basic Sobolev elliptic estimate holds for functions $f \in H_s(M^n)$. Let $\{U_i\}$ and $\{\zeta_i\}$ be as above, and for convenience write $f_i = \zeta_i f$. By the result of section 2.1.3 we have, for L an elliptic operator of degree k,

$$\| f_i \|_s \leq C_i(\| Lf_i \|_{s-k} + \| f_i \|_{s-1}).$$

Then we have,

$$\begin{aligned}
\| f \|_s = \left\| \sum_{i=1}^m f_i \right\|_s &\leq \sum_{i=1}^m \| f_i \|_s \\
&\leq \sum_{i=1}^m C_i(\| Lf_i \|_{s-k} + \| f_i \|_{s-1}) \\
&\leq mC_{j^*}(\| Lf_{j^*} \|_{s-k} + \| f_{j^*} \|_{s-1}) \\
&= mC_{j^*}(\| [L, \zeta_{j^*}]f + \zeta_{j^*} Lf \|_{s-k} + \| \zeta_{j^*} f \|_{s-1}),
\end{aligned}$$

where j^* indicates the index of the largest term in the sum. Now using the fact that $0 \leq \zeta_i \leq 1$ and that $[L, \zeta_{j^*}]$ is a differential operator of degree at most $k-1$, we have:

$$\begin{aligned}
\| f \|_s &\leq mC_{j^*}(\| [L, \zeta_{j^*}]f + \zeta_{j^*} Lf \|_{s-k} + \| \zeta_{j^*} f \|_{s-1}) \\
&\leq mC_{j^*}(\| [L, \zeta_{j^*}]f \|_{s-k} + \| \zeta_{j^*} Lf \|_{s-k} + \| \zeta_{j^*} f \|_{s-1}) \\
&\leq K(\| Lf \|_{s-k} + \| f \|_{s-1}),
\end{aligned}$$

and therefore this estimate holds for functions defined over the entire manifold.

3.2 Differences from Euclidean Space

We are now interested in applying our linear differential operators to spaces of sections—continuous maps from a compact manifold to a vector bundle $\pi : V \to M^n$ such that when composed with π equal the identity. We understand and manipulate sections by examining them locally via the vector bundle's local trivializations, regarding the sections as functions, i.e. by studying a section's representations with respect to a local trivialization. Suppose that $\{U_i\}$ is a covering of our manifold M^n by local coordinate neighborhoods, and $\{h_i\}$ are local trivializations of the bundle, i.e.

$$h_i(\pi^{-1}(U_i)) \cong U_i \times \mathbb{C}^m.$$

Recall then, that the so-called transition functions of the atlas (h_i) are a set of functions

$$g_{ij} : U_i \cap U_j \longrightarrow \mathrm{GL}(n, \mathbb{C}),$$

given by

$$h_j \circ h_i^{-1}(x, y) \;=\; (x, g_{ij}(x)y),$$

that satisfy a cocycle relation: $g_{ij}g_{jk} = g_{ik}$ on $U_i \cap U_j \cap U_k$. It is through these cocycles that the local representations of sections are related. For f a section, the local representation on U_i is defined as the function $f_i : M^n \cap U_i \longrightarrow \mathbb{C}^n$ that satisfies

$$h_i \circ f(x) = (x, f_i(x)) \in U_i \times \mathbb{C}^n.$$

These local representations are related by the cocycles as follows:

$$f_i = g_{ij} f_j \;\text{ in } U_i \cap U_j,$$

where recall that by f_i we really mean the n-tuple of functions $\{f_i^k\}_{k=1}^n$. Now the trouble (or some might say the fun) begins.

3.2.1 Connections and the Covariant Derivative

We would like differential operators, in particular the single derivative D_j to behave in a similar fashion as it behaves in Euclidean space—by sending functions to functions, and section representations to section representations. However, under the usual definition of derivation, if we let superscripts denote the particular coordinates of a function, and subscripts denote coordinate neighborhood, we have:

$$\begin{aligned}
\frac{\partial f_i^\alpha}{\partial x^1} &= \frac{\partial}{\partial x^1}(g_{ij}^{\alpha\beta} f_j^\beta) \\
&= \frac{\partial g_{ij}^{\alpha\beta}}{\partial x^1} f_j^\beta + g_{ij}^{\alpha\beta} \frac{\partial f_j^\beta}{\partial x^1} \\
&= g_{ij}^{\alpha\beta}\left(\frac{\partial f_j^\beta}{\partial x^1} + \left(\left(g_{ij}^{\alpha\beta}\right)^{-1} \frac{\partial g_{ij}^{\alpha\beta}}{\partial x^1}\right) f_j^\beta\right) \\
&= g_{ij}^{\alpha\beta}\left(\frac{\partial f_j^\beta}{\partial x^1}\right) + \Lambda,
\end{aligned}$$

for Λ some nonzero term, demonstrating that a section's derivatives do not transform via the cocycles $\{g_{ij}\}$. Therefore if we understand derivatives in the same sense as on Euclidean space, the derivative of a section's local representation no longer transforms like a local representation. The solution to this problem is to adopt a slightly different notion of a derivative that behaves sufficiently like the derivative on Euclidean space, e.g. satisfies Liebnitz's product differentiation rule etc. If we consider the tangent bundle over a manifold, then we can think of sections as vector fields. Then asking for a "derivative" that sends sections to sections is like asking for a relation between the tangent spaces at different points of the manifold. For this reason, even on arbitrary vector bundles this new "derivative", called a covariant derivative, which is compatible with the cocycle relation, is defined by a local tensor on the manifold, called a **linear connection**. Suppose that for each of the local neighborhoods of the manifold we have a matrix of 1-forms with elements:

$$a^{\alpha\beta} := a_1^{\alpha\beta}(x)dx^1 + a_2^{\alpha\beta}(x)dx^2 + \cdots a_n^{\alpha\beta}(x)dx^n,$$

where α, β range over $1, \ldots, m$ for m the dimension of the fibres of the bundle over M^n. Suppose further that these matrices transform on the intersection of two local coordinate neighborhoods $U, \hat{U}$ according to the rule:

$$\hat{a}^{\alpha\beta} = g^{\alpha\sigma}(a^{\sigma\gamma} + \left(g^{-1}\right)^{\sigma\delta} dg^{\delta\gamma}) \left(g^{-1}\right)^{\gamma\beta}.$$

Straight computation verifies that the covariant derivative, defined by the rules:

$$\begin{aligned}
(\nabla_1 f)^\alpha &:= \frac{\partial f^\alpha}{\partial x^1} + a_1^{\alpha\beta}(x) f^\beta \\
(\nabla_2 f)^\alpha &:= \frac{\partial f^\alpha}{\partial x^2} + a_2^{\alpha\beta}(x) f^\beta \\
&\vdots \\
(\nabla_n f)^\alpha &:= \frac{\partial f^\alpha}{\partial x^n} + a_n^{\alpha\beta}(x) f^\beta,
\end{aligned}$$

transforms by the transition functions $g^{\alpha\beta}$:

$$\left(\hat{\nabla}_j f\right)^\alpha = g^{\alpha\beta} \left(\nabla_j f\right)^\beta.$$

(Note that in the above to simplify notation and make the result more transparent we use the convention of implied summation on repeated indices). Furthermore, it is straightforward to verify that the covariant derivative defined in this manner, satisfies the rules of calculus which the usual derivative satisfies (e.g. Liebniz's product rule). The next Theorem guarantees the existence of connections.

Theorem 10 *If $\pi : V \to M$ is an arbitrary vector bundle over a compact Riemannian manifold, then there exists a connection on V by which we can define the covariant derivatives of smooth sections.*

PROOF. We prove this by a naive approach. Since M is compact, we can choose a finite subcover $U_1, \ldots, U_N$ of coordinate neighborhoods for M. The idea is to fix a well behaved connection on one of the neighborhoods, and then use the transformation laws above to find the connection in the other neighborhoods, multiplying by bump functions whenever necessary, to remove singularities. Such a procedure is justified by the fact that differentiation is a pointwise operation. Then consider U_1. On this neighborhood, define the covariant derivative in the direction of the i^{th} basis vector to be the same form as the usual derivative for all i, that is, let the connection matrix be zero in this neighborhood. Now consider any other coordinate neighborhood U_j such that $U_1 \cap U_j \neq \emptyset$. Choose a bump function f_j that is supported on a neighborhood of $M - U_j$, and identically 1 on $M - U_j$. Then multiplying the transformation of the connection by this bump function, we obtain a well behaved connection on U_j that transforms consistently with the transformation laws with the connection on U_1. We can continue this process as long as we do not encounter any singularities that happen to lie on the boundaries of the coordinate neighborhoods. We can avoid this case by slightly deforming the boundaries of the coordinate neighborhoods. Then continuing this process of transforming and then bumping, we obtain a connection in a finite number of steps, thus proving the Theorem. □

3.2.2 The Riemannian Metric and Inner Products

Having understood the derivative in a manner consistent with the transition functions of a particular vector bundle, we have a reasonable notion of the meaning of a differential operator that operates on a section space. The natural question then is how can we understand the adjoint operator in this context. There are several other notions from analysis on Euclidean space that must be generalized and made sense of globally on a compact manifold before we can continue our analysis. In particular, in $\mathbb{R}^n$ we have a well defined, that is, a globally defined, definition of distance—we have the usual Euclidean metric. In addition, we have the usual inner product on $\mathbb{R}^n$ (which of course yields the metric). This inner product on $\mathbb{R}^n$ gives us the inner product on spaces of functions mapping to $\mathbb{R}^n$. Indeed for f, g square integrable functions, we have

$$\langle\, f,\, g\,\rangle_{\mathcal{F}} = \int \langle\, f,\, g\,\rangle_{\mathbb{R}^n}\, dx.$$

And it is in this context that we can define the adjoint operator. Then we have two tasks: to find a suitable notion of a metric on a manifold, and also of an inner product for sections of an arbitrary vector bundle. It turns out that both these problems are related; in fact, finding a suitable metric on a manifold reduces to the more general problem of defining an inner product on some vector bundle, even when the manifold does not naturally have the structure of a vector space. Consider the problem of computing the arclength of a curve on a manifold M^n. Certainly if we can do this, we have a metric. In $\mathbb{R}^n$ if we are given the curve

$\alpha(t)$, and say $t \in [0,1]$ we can define,

$$\begin{aligned} \text{arclength}(\alpha) &= \int_0^1 \langle\, \alpha'(t)\,, \alpha'(t)\,\rangle^{1/2} \\ &= \int_0^1 \left[\sum_{i=1}^n \left(\frac{dx_i(t)}{dt} \right)^2 \right]^{1/2}, \end{aligned}$$

where $\alpha'(t)$ is the velocity vector. But on any manifold (including $\mathbb{R}^n$) these velocity vectors are elements of the tangent bundle. In Euclidean space, the fact that we are integrating sections is obscured because the tangent bundle is trivial and isomorphic to the manifold itself, $\mathbb{R}^n$. Nevertheless, we see that as soon as we know how to define a suitable inner product on a vector bundle $\pi : V \to M$, we also have a metric on M. Significantly, it is this metric, called the **Riemannian metric**, that yields the volume form and therefore allows us to integrate functions over the manifold, and hence determine membership in the various spaces, $L^p(M)$ and $H_s(M)$.

If f, g are sections of V, then for $x \in M^n$, $f(x), g(x)$ are vectors in the fibre V_x above $\{x\}$, which we realize as elements of $\mathbb{R}^m$ via the local trivializations h_{U_i} $(x \in U_i)$. While the h_{U_i} locally are isomorphisms, there is nothing inherent in their definition that ensures that transformation between the local trivializations h_{U_i} and h_{U_j} respects the inner product. Therefore the expression $\langle\, f(x),\, g(x)\,\rangle_{\mathbb{R}^n}$ really makes no sense unless a particular coordinate neighborhood and associated local trivialization are specified.

The usual inner product on $\mathbb{R}^n$ is defined as a symmetric bilinear form; with respect to the standard basis, this takes the form of the identity matrix. Therefore we need to define another such symmetric, positive definite, bilinear form for vectors in the vector bundle V, that is independent of the choice of local coordinate patch. Recall that a tensor $b(\cdot\,,\cdot)$ of second order is a bilinear form, and we write

$$b(v,w) = b(\sum_1^n v_i e_i\,, \sum_1^n w_j e_j) = \sum_{i,j} b_{ij} v_i w_j,$$

where the $\{e_i\}$ denote the basis vectors of the fibre containing v, w. Further recall that if $\{\hat{e}_i\}$ forms another basis, and the two bases are related by $\hat{e}_j = \sum_i a_{ij} e_i$, then

$$\hat{b}_{ij} = b(\hat{e_i}, \hat{e_j}) = \sum_{k,l} a_{ik} a_{jl} b_{kl}.$$

Then, if we let the a_{ik} be the transition relations on the vector bundle, the expression $b(v,w)$ is independent of local representation for $v, w \in V$. Therefore in the intersection of two neighborhoods $U, \hat{U}$ we have $\hat{g}_{ij} \hat{v} \hat{w} = g_{ij} v w$. If the tensor is also symmetric and positive definite, this gives a globally defined "inner product" on the fibres of the bundle V. If the bundle V is the tangent bundle, the tensor, suggestively denoted $\langle \cdot, \cdot \rangle_p$ for $p \in M$, yields the Riemannian metric, as discussed above. For a general bundle V, we have defined an inner product on the space of square integrable sections, for the expression

$$\langle f, g \rangle_{\mathcal{F}} = \int b(f(x), g(x)) d(vol),$$

is a positive definite symmetric, bilinear form on the sections of the vector bundle $\pi : V \longrightarrow M^n$, that is independent of the local representations chosen. With this inner product, we can find the adjoint of any linear operator, as it is defined by the equation:

$$\langle Tv, w \rangle_{\mathcal{F}} = \langle v, T^* w \rangle_{\mathcal{F}}.$$

To appreciate the differences between general compact manifolds and Euclidean space, consider the adjoint operator of D_l, a simple directional derivative. On Euclidean space, the adjoint operator would be $-D_l$, as given by integration by parts. On a general compact manifold, we also need to consider the tensor and the volume form. Therefore we have

$$\begin{aligned}
\langle D_l f, g \rangle_{\mathcal{F}} &= \int \left(\frac{\partial f}{\partial x_l} \right) gb \, d(vol) \\
&= -\int f \left(\frac{\partial}{\partial x_l} (gb \, d(vol)) \right) \\
&= \int f \cdot (-1) \frac{1}{d(vol)} \left[b \, d(vol) \frac{\partial}{\partial x_l} + \left(\frac{\partial}{\partial x_l} b \, d(vol) \right) \right] g \, d(vol) \\
&= \langle f, (D_l)^* g \rangle_{\mathcal{F}}.
\end{aligned}$$

Therefore the adjoint operator can be rather more complicated than the usual adjoint operator on Euclidean space. However note that the degree of the operator remains the same. Moreover, a general elliptic operator's highest order terms agree with its adjoint's highest order terms up to multiplication by the same nonvanishing factor. Therefore an operator is elliptic if and only if its adjoint operator is as well.

Example 4 To fix ideas, we consider the particular example of the cotangent bundle over a Riemann surface X. The sections are the 1-forms. Locally, these look like $\frac{\partial f}{\partial z_i} dz_i$ where z_i is the local coordinate on the neighborhood $U_i \subset X$. The cocycles of this line bundle are of the form

$$g^{ik} = \frac{\partial z_i}{\partial z_j},$$

Then a tensor transforms like

$$\begin{aligned}
\hat{b}_{ij} = b(\hat{e}_i, \hat{e}_j) &= \sum_{k,l} a_{ik} a_{jl} b_{kl} \\
&= \sum_{k,l} \frac{\partial z_i}{\partial z_k} \frac{\partial z_j}{\partial z_l} b_{kl}.
\end{aligned}$$

So in particular, the expression

$$\sum_{i,j} \frac{\partial f}{\partial z_i} \frac{\partial g}{\partial z_j} b_{ij},$$

is invariant under change of local representation, and hence the inner product

$$\langle\, df,\, dg\, \rangle_{\mathcal{F}} = \int \sum_{ij} \frac{\partial f}{\partial z_i} \frac{\partial g}{\partial z_j} b_{ij} \sqrt{\det(b_{ij})}\, dz,$$

is well defined, with the required properties.

The following Theorem guarantees that any vector bundle $\pi : V \longrightarrow M^n$, for M^n a compact Riemannian manifold, admits such positive definite symmetric tensors. In particular then, the Theorem states that any compact Riemannian manifold admits a Riemannian metric.

Theorem 11 *If M^n is a compact Riemannian manifold, and $\pi : V \longrightarrow M^n$ is any vector bundle over M^n, we can find a positive definite, symmetric, bilinear tensor of order two on V. In other words, any compact Riemannian manifold admits a Riemannian metric.*

PROOF. Let $\{U_\alpha\}$ be any covering of our manifold M^n by coordinate neighborhoods, and V as in the statement of the Theorem. By the compactness of M^n, we can choose a finite subcover, call it $U_1, \ldots, U_N$. Denoting by V_{U_i} the preimage of U_i under the projection map π, we note that by definition, we have local trivializations $\{h_i\}$ such that

$$V_{U_i} \cong U_i \times \mathbb{C}^m,$$

where m is the dimension of the individual fibres V_p of V. Certainly then, there are functions $b_i : V \times V \longrightarrow \mathbb{C}$ such that b_i looks like a symmetric, bilinear, positive definite tensor *on* V_{U_i}. Let $\{b_i\}$ be such a family, and denote by b_i^j the representation of b_i in the neighborhood U_j of M^n. Now take $\{f_i\}$ to be smooth functions on M^n, such that $f_i > 0$ on U_i and $f_i \equiv 0$ on $M^n - U_i$. Then consider the tensor:

$$b = f_1 b_1 + \cdots + f_N b_N.$$

On U_i this looks like

$$f_i b_i + \sum_{j \neq i} f_j b_j^i.$$

Consider b_j^i on $U_i \cap U_j$. By definition of the transformations defined above, b_j^i is positive definite, symmetric, and above all, well defined, because b_j is. However, b_j^i may go to infinity, or to zero, or may not be defined, at some point $p \in U_i - U_j$. Thanks to the bump functions $\{f_i\}$, $f_i b_i + \sum_{j \neq i} f_j b_j^i$ can neither explode, nor vanish, and is everywhere well defined. We have left to show that any convex combination of positive definite symmetric tensors is again a positive definite symmetric tensor. But this is clear from the linearity of the tensor. Therefore finding a positive definite symmetric tensor on M^n amounts to finding the family of tensors $\{b_i\}$, and the family of smooth bump functions $\{f_i\}$. $\square$

Corollary 4 *Any vector bundle $\pi : V \to M$ over a compact manifold admits a Riemannian metric.*

PROOF. Any vector bundle is a manifold in its own right. However we cannot quite apply the Theorem directly as V need not be compact. However, by the compactness of M, we can write V as the union of a finite number of locally trivial sets:

$$V = \bigcup_{1}^{N} V_{U_i},$$

where the U_i are a finite subcover of M by coordinate neighborhoods, and $V_{U_i} = \pi^{-1}(U_i)$. Since $V_{U_i} \cong U_i \times \mathbb{C}^m$ its tangent bundle is itself, and from here the Corollary follows from the proof of the Theorem above. $\square$

So far we have seen that the differential operator itself depends upon the chosen connection, and its adjoint depends on the choice of inner product on the vector bundle. The next standard Theorem of Levi-Civita states that if the vector bundle is the tangent bundle, then for any metric tensor (b_{ik}), there exists exactly one linear connection that is compatible with the metric and satisfies $[\nabla_{e_i}, \nabla_{e_j}] - \nabla_{[e_i,e_j]} = 0$, i.e. is torsion free. By compatible with the metric given by a tensor b, we mean, as usual,

$$d(b(X,Y)) = b(\nabla X, Y) + b(X, \nabla Y).$$

The proof involves pushing through the algebra from the definitions, and can be found in, e.g. Laugwitz [5] or do Carmo [2].

Theorem 12 *If the inner product on the tangent bundle $\pi : TM^n \longrightarrow M^n$ is given by the positive definite symmetric metric tensor (b_{ik}), there exists exactly one torsion free connection that is compatible with the metric, and it is given by*

$$\alpha_i^{jk} := \frac{1}{2} b_{ir} \sum_r \left(\frac{\partial b_{rj}}{\partial x_k} + \frac{\partial b_{kr}}{\partial x_j} - \frac{\partial b_{jk}}{\partial x_r} \right).$$

The converse of this Theorem does not hold, that is, not every torsion free connection is derived from a particular Riemannian metric.

While a Riemannian metric automatically specifies the connection on the tangent bundle, the metric itself is not uniquely determined *a priori* by properties of the manifold. Therefore specifying the metric, and the inner products to be used on the particular vector bundles must be part of the given information in the problem of computing the index of an elliptic operator. As the next section shows, a choice of a different metric is equivalent to some perturbation of the operator. If the perturbation is sufficiently small therefore, i.e. if the new metric is sufficiently close to the old, then the index of any given elliptic operator is preserved.

3.3 Proof of the Invariance of the Index

We have finally developed the machinery necessary to prove the main Theorem.

Theorem 13 *If M^n is a compact Riemannian manifold, $\pi : V \longrightarrow M^n$ is a vector bundle over M^n, and L is an elliptic operator on the space of sections of V, then $\exists\, \varepsilon > 0$ such that if r is any operator that satisfies*

$$\| rf \|_{L^2} \leq \varepsilon(\| \nabla f \|_{L^2}^2 + \| f \|_{L^2})^{1/2}, \;\; \forall f \in H_1,$$

then the perturbed operator $L' := L + r$ has the same index as L:

$$\textit{Index}(L) = \textit{Index}(L').$$

This Theorem also addresses the issue of choosing different Riemannian metrics. For two distinct metrics coming from metric tensors b, b', we say that b, b' are close if the resulting changes in the connection and the volume form are small. For this we need $|b - b'| + |\partial_k b - \partial_k b'| < \varepsilon$.

We prove this Theorem in three steps:

(i) We show that the dimension of the kernel of an elliptic operator, locally, can only decrease.

(ii) We then show that the dimension of the cokernel must also decrease.

(iii) Finally we prove that the difference of these dimensions remains the same for sufficiently small perturbations.

Lemma 6 *The dimension of the kernel of an elliptic operator L can only decrease under perturbation.*

Proof. For notational simplicity, we assume L has order 1. This assumption is made without loss of generality and the proof below goes through without change in the general case. First, recall the basic Sobolev elliptic estimate, and observe its equivalent expression:

$$\| u \|_{H_1}^2 \leq C(\| Lu \|_{L^2}^2 + \| u \|_{L^2}^2) \iff \| Lu \|_{L^2}^2 \geq \| \nabla u \|_{L^2}^2 - C \| u \|_{L^2}^2 .$$

Suppose first that L has trivial kernel. We show that $L' = L + r$ must also have trivial kernel. Now, if kernel$(L) = \{0\}$ then $\exists\, \lambda > 0$ that satisfies

$$\| Lu \|_{L^2}^2 \geq \lambda^2(\| \nabla u \|_{L^2}^2 + \| u \|_{L^2}^2), \;\; \forall u \in H_1.$$

For if this is not the case, then for any $\lambda > 0$ we can find some u such that

$$\| Lu \|_{L^2}^2 \leq \lambda^2(\| \nabla u \|_{L^2}^2 + \| u \|_{L^2}^2).$$

But then by the homogeneity of L, ∇, and $\| \cdot \|_{L^2}$ we can rescale u so that the right hand side above equals λ^2. Therefore we can find a sequence u_n such that

$$\begin{aligned} \| \nabla u_n \|_{L^2}^2 + \| u_n \|_{L^2}^2 &= 1 \ \forall n \in \mathbb{N} \\ \| L u_n \|_{L^2}^2 &\to 0 \text{ as } n \to \infty. \end{aligned}$$

By the Rellich Lemma, since the sequence $\{u_n\}$ is bounded in H_1, it is compact in $H_0 = L^2$, and therefore contains a convergent subsequence. By reindexing if necessary, again call this sequence $\{u_n\}$. Then by the basic estimate, we have:

$$\| u_n - u_m \|_{H_1}^2 \leq C(\| L(u_n - u_m) \|_{L^2}^2 + \| u_n - u_m \|_{L^2}^2),$$

which goes to zero by the above. But then u_n converges strongly in H_1, which violates our assumption that L has a trivial kernel since the H_1 norm of $u = \lim_n u_n$ is 1, hence $u \neq 0$.

Then choose such a $\lambda > 0$, and let $\varepsilon < \frac{\lambda}{2}$. We have

$$\begin{aligned} \| L'u \|_{L^2} &\geq \| Lu \|_{L^2} - \| ru \|_{L^2} \\ &\geq \lambda \| u \|_{H_1} - \varepsilon \| u \|_{H_1} \\ &= (\lambda - \varepsilon) \| u \|_{H_1} \geq \frac{\lambda}{2} \| u \|_{H_1} > 0, \end{aligned}$$

as long as $u \neq 0$. Therefore L' must also have a trivial kernel. Now suppose L has a nontrivial kernel W. Since W is closed, we can decompose H_1 as $W \oplus W^{\perp}$. By the above, there exists some $\lambda > 0$ such that if $u \in W^{\perp}$,

$$\| Lu \|_{L^2}^2 \geq \lambda^2(\| \nabla u \|_{L^2}^2 + \| u \|_{L^2}^2).$$

Take some $u \in \text{kernel}(L')$. By the decomposition above we can write $u = u_0 + u_1$, for $u_0 \in W$, $u_1 \in W^{\perp}$. This yields

$$\begin{aligned} & \| Lu_1 \|_{L^2} \leq \| ru_1 \|_{L^2} + \| ru_0 \|_{L^2} \\ \Rightarrow\ & \lambda(\| \nabla u_1 \|_{L^2}^2 + \| u_1 \|_{L^2}^2)^{1/2} \leq \varepsilon(\| \nabla u_1 \|_{L^2}^2 + \| u_1 \|_{L^2}^2)^{1/2} + \| ru_0 \|_{L^2} \\ \Rightarrow\ & \| \nabla u_1 \|_{L^2}^2 + \| u_1 \|_{L^2}^2 \leq \frac{4}{\lambda^2} \| ru_0 \|_{L^2}^2, \end{aligned}$$

where $\varepsilon < \frac{\lambda}{2}$ as before. We have proved that the projection

$$\Pi : \text{kernel}(L') \longrightarrow \text{kernel}(L),$$

is injective. $\square$

Therefore we see that the dimension of the kernel can only decrease under small perturbation, and that in fact we must have $\text{kernel}(L') \subset \text{kernel}(L)$. This fact seems less peculiar, perhaps, when we observe that λ which essentially determines an upper bound on ε, the size of the permissible perturbation, depends entirely on the operator in question. It does not follow, therefore, that subtracting the perturbation r from the operator L' would be a "legal", i.e. sufficiently small, perturbation.

For the next step of the proof, we construct a map $u_1 : W \to W^\perp$ such that for any $u_0 \in W$, we have $u_0 + u_1(u_0) \in \mathrm{kernel}(L')$. Choose some $u_0 \in W$. We seek some $u_1 \in W^\perp$ such that

$$\begin{aligned} & L'(u_1 + u_0) = 0 \\ \Leftrightarrow \quad & Lu_1 + r(u_1 + u_0) = 0 \\ \Leftrightarrow \quad & Lu_1 = -r(u_1 + u_0). \end{aligned}$$

Let Π_1 denote the projection map from the range of L onto the perpendicular of the cokernal space. Note that we can view the cokernel and its perpendicular as subspaces, since we showed in Lemma 5 that the image of an elliptic operator must be closed in the range. Then, since L is invertible on the space cokernel(L) we can write:

$$u_1 = L^{-1} \circ \Pi_1(-r(u_1 + u_0)) = T_{u_0}(u_1)$$

where we have defined the map T_{u_0} by:

$$T_{u_0}(x) := -L^{-1} \circ \Pi_1(rx + ru_0).$$

If we can show that T_{u_0} is a contraction mapping, we will have shown that it has a fixed point, and hence that the equation above is satisfied by some u_1. To this end, we have:

Lemma 7 *The map T_{u_0} as defined above is a contraction mapping, and in particular, it has a fixed point.*

Proof. We verify this directly.

$$\begin{aligned} \| L(T_{u_0}(x) - T_{u_0}(y)) \|_{L^2}^2 &= \| \Pi_1 r(x + u_0) - \Pi_1 r(y + u_0) \|_{L^2}^2 \\ &\leq \| r(x + u_0) - r(y + u_0) \|_{L^2}^2 \\ &= \| r(x - y) \|_{L^2}^2 \\ &\leq \varepsilon(\| \nabla(x - y) \|_{L^2}^2 + \| x - y \|_{L^2}^2)^{1/2} \\ &\leq \frac{\varepsilon}{\lambda} \| L(x - y) \|_{L^2}^2 \\ &\leq \frac{1}{2} \| L(x - y) \|_{L^2}^2, \end{aligned}$$

for $\varepsilon < \frac{\lambda}{2}$, as usual. Now if we regard L^{-1} as a map from image(L) to $W^\perp$, then $\| L \| = \| L^{-1} \|^{-1}$, and therefore T_{u_0} is indeed a contraction mapping, and hence has a unique fixed point u_1. $\square$

Therefore given any $u_0 \in W$ we have a map to a unique element $u_1(u_0) \in W^\perp$ that satisfies

$$Lu_1 + \Pi_1(r(u_1(u_0) + u_0) = 0,$$

whence

$$u_1(u_0) + u_0 \in \text{kernel}(L') \iff \Pi_1(r(u_1(u_0) + u_0)) = r(u_1(u_0) + u_0).$$

Now define the map

$$\begin{aligned} F : \text{kernel}(L) &\longrightarrow \text{cokernel}(L) \quad \text{by} \\ u_0 &\longmapsto (I - \Pi_1)(r(u_1(u_0) + u_0)) \in \text{cokernel}(L). \end{aligned}$$

This immediately yields the following.

Lemma 8 *kernel*$(L') \cong$ *kernel*(F)*.*

A similar statement holds for the cokernel of L' and F.

Lemma 9 *For the map F defined as above, we have:*

$$\textit{cokernel}(L') \cong \textit{cokernel}(F).$$

Proof. By Lemma 5, L has closed range, and hence we can decompose its range as $V \oplus V^\perp$ where $V = \text{cokernel}(L)$. Then for any v we can write $v = w_0 + w_1$ for $w_0 \in V$, $w_1 \in V^\perp$. We show that

$$v \in \text{cokernel}(L') \Leftrightarrow w_0 \in \text{cokernel}(F).$$

From Proposition 3 and Lemma 6 we have that $\text{cokernel}(L') \hookrightarrow \text{cokernel}(L)$ is an injection. Therefore if $v = w_0 + w_1 \in \text{cokernel}(L')$ we must have $w_1 = 0$. Therefore

$$v \in \text{cokernel}(L') \Rightarrow w_0 \in \text{cokernel}(L').$$

But again using the decomposition into $(\text{kernel}(L))$ and $(\text{kernel}(L))^\perp$ we have

$$\begin{aligned} L'(u) &:= L'(u_0 + u_1) = L(u_1) + \Pi_1(r(u_1 + u_0)) + F(u_0) \\ &\Rightarrow F(u_0) \perp w_0 \ \forall u_0 \in \text{kernel}(L) \\ &\Rightarrow w_0 \in \text{cokernel}(F), \end{aligned}$$

where the first implication follows because the first two terms, $L(u_1)$ and $\Pi_1(r(u_1 + u_0))$, are by definition in $V^\perp$.

Conversely, suppose $w_0 \in \text{cokernel}(F)$. Then $w_0 \in \text{cokernel}(L)$ since $\text{cokernel}(F) \subset \text{cokernel}(L)$. Now, $w_1 \in (\text{cokernel}(L))^\perp$. By Proposition 3, and again by Lemma 6 we have that the projection $\text{cokernel}(L') \to \text{cokernel}(L)$ is injective, and therefore $v = w_0 + w_1$ must be in the cokernel of L' if w_0 is in the cokernel of F. □

Finally, we observe that we established in section 2.2.2 that the kernel and cokernel are finite dimensional subspaces. Therefore

$$F : \text{kernel}(L) \longrightarrow \text{cokernel}(L),$$

is a linear transformation between finite dimensional vector subspaces. Therefore by the Rank-Nullity Theorem, its index must be the difference in dimension between its domain and range spaces:

$$\mathrm{Index}(F) = \mathrm{dimKernel}(L) - \mathrm{dimCokernel}(L) = \mathrm{Index}(L).$$

By Lemma 8 and Lemma 9 we conclude the Theorem:

$$\mathrm{Index}(L) = \mathrm{Index}(L').$$

□

4 Example: The Torus

Consider the torus $T = \mathbb{R}^2/\mathbb{Z}^2$; that is, the torus with transition functions $y \mapsto y+1$, and $x \mapsto x+1$. Consider the connection $A = i(\alpha\, dx + \beta\, dy) = i \cdot a$ where $\alpha, \beta \in \mathbb{R}$. We then define the covariant derivative

$$\nabla_{\bar\partial} = \bar\partial + i \cdot a.$$

This operator is elliptic because $\bar\partial$ is, and ellipticity depends only on the highest order terms. First note that since the transformation functions for the tensors are the identity, then the naive choice for an inner product, namely the identity, indeed provides us with a positive definite symmetric form independent of coordinate chart. But viewing $\mathbb{R}^2 \cong \mathbb{C}$ we see that this 2×2 tensor that looks like the identity corresponds to the convenient metric $|\cdot|$. Then to check compatibility with this metric we have:

$$\begin{aligned}
\bar\partial\langle f, g\rangle = \bar\partial(f\bar g) &= \bar g\bar\partial f + f\bar\partial\bar g \\
&= \bar g(\bar\partial + i\cdot a)f + f\bar\partial\bar g - i\cdot af\bar g \\
&= \bar g\nabla_{\bar\partial} f + f\overline{(\nabla_{\bar\partial} g)},
\end{aligned}$$

which is what we needed to show.

Now consider any (smooth) section $\psi(x,y)$ on T. Since this is periodic in both x and y we can expand it in a Fourier series in both variables to get:

$$\psi(x,y) = \sum_{k,l} c_{kl} e^{ikx} e^{ily}.$$

Since $\bar\partial = \frac{1}{2}(\partial_x + i\partial_y)$, we can easily compute its covariant derivative

$$\nabla_{\bar\partial}(\psi) = \sum_{k,l} \left(\frac{ik - l}{2} + ia\right) c_{kl} e^{ikx} e^{ily}.$$

Now, since $k, l \in \mathbb{Z}$, the operator $\nabla_{\bar\partial}$ will have a nontrivial kernel (in particular a 1 dimensional kernel) if and only if $\alpha, \beta \in \mathbb{Z}$. Now from the given metric on the bundle, it is clear that its adjoint behaves similarly with respect to its kernel, as the adjoint is nothing more than a constant multiple of the original operator, due to the fact that the transition functions of tensors are the identity. Therefore while $\alpha, \beta \notin \mathbb{Z}$ we have a trivial kernel and cokernel, thus the index is 0. However as soon as we have $\alpha, \beta \in \mathbb{Z}$, the kernel and cokernel both become 1 complex dimensional. Again, the index is 0, as expected.

A Elliptic Operators and Riemann-Roch

The Riemann-Roch Theorem states that if M is a compact Riemannian manifold of genus g, the alternating sum of the dimensions of the cohomology groups H^k is related to the genus in a very simple way:

$$\dim H^0 - \dim H^1 + \dim H^2 - \cdots + (-1)^n H^n = 1 - g,$$

where n is the dimension of the manifold. We show that this alternating sum is the index of an elliptic operator.

To fix ideas we consider the case $n = 2$. Consider the complex

$$0 \longrightarrow \mathcal{O} \xrightarrow{\bar{\partial}_1} \mathcal{E}^{0,1} \xrightarrow{\bar{\partial}_2} \mathcal{E}^{0,2} \xrightarrow{\bar{\partial}_3} 0,$$

where $\mathcal{E}^{0,k}$ are the $(0,k)$-forms on the manifold. Since $\bar{\partial}^2 = 0$, the sequence is indeed a copmlex. We also have the mapping $\bar{\partial}_2^* : \mathcal{E}^{0,2} \longrightarrow \mathcal{E}^{0,1}$. Now extend $\bar{\partial}_1$ by defining it as identically zero on $\mathcal{E}^{0,1}$, and similarly for $\bar{\partial}_2^*$. Now consider the operator $D = \bar{\partial}_1 + \bar{\partial}_2^*$. We have

$$D : \mathcal{O} \oplus \mathcal{E}^{0,2} \longrightarrow \mathcal{E}^{0,1},$$

and D is elliptic because $\bar{\partial}_1$ and $\bar{\partial}_2^*$ are. Furthermore, we have

$$\begin{aligned}
\text{kernel}(D) &= \text{kernel}(\bar{\partial}_1) \cup \text{kernel}(\bar{\partial}_2^*) \cong \text{kernel}(\bar{\partial}_1) \cup \text{cokernel}(\bar{\partial}_2), \\
\text{cokernel}(D) &= \text{cokernel}(\bar{\partial}_1) \cap \text{cokernel}(\bar{\partial}_2^*) \\
&\cong \text{kernel}(\bar{\partial}_2) \cap \text{cokernel}(\bar{\partial}_1).
\end{aligned}$$

Meanwhile we have:

$$\begin{aligned}
H_0 &= \text{kernel}(\bar{\partial}_1)/\text{Image}(0) \cong \text{kernel}(\bar{\partial}_1), \\
H_1 &= \text{kernel}(\bar{\partial}_2)/\text{Image}(\bar{\partial}_1) \cong \text{kernel}(\bar{\partial}_2) \cap \text{cokernel}(\bar{\partial}_1), \\
H_2 &= \text{kernel}(\bar{\partial}_3)/\text{Image}(\bar{\partial}_2) \cong \text{cokernel}(\bar{\partial}_2).
\end{aligned}$$

Putting the above together we find:

$$\begin{aligned}
\dim \text{kernel}(D) &= \dim \text{kernel}(\bar{\partial}_1) + \dim \text{cokernel}(\bar{\partial}_2) \\
&= \dim H^0 + \dim H^2; \\
\dim \text{cokernel}(D) &= \dim \text{cokernel}(\bar{\partial}_1) \cap \text{cokernel}(\bar{\partial}_2^*) \\
&= \dim H^1,
\end{aligned}$$

and therefore we conclude

$$\text{Index}(D) = \dim H^0 - \dim H^1 + \dim H^2 = 1 - g.$$

B An Alternate Proof of Elliptic Regularity

The proof of elliptic regularity given in section 2.2.1 relies on Nirenberg's difference quotients, which are essentially pointwise approximations to the partial derivatives of a distribution. In this section we present a proof which more clearly illustrates the smoothing nature of elliptic operators.

Theorem 14 *If L is an elliptic operator with constant coefficients and of the form*

$$P_N(D) = \sum_{|\alpha|=N} a_\alpha D^\alpha,$$

and if $P_N(D)u = v$ for $u, v \in \mathcal{C}_c^{-\infty}(\Omega)$, i.e. u, v are distributions with compact support, then $v \in H_s$ implies that $u \in H_{s+N}$.

Proof. Since u, v both have compact support, in particular they are tempered, and hence their Fourier transform exists (distributionally defined). Taking the Fourier transform of the equation $P_N(D)u = v$ we get $\widehat{P_N(D)}\hat{u} = \hat{v}$. Consider now the function $\hat{Q}(t) = \hat{P}_N \cdot |\hat{P}_N|^{-1}$. Now Q always has modulus 1, $\arg \hat{Q} = \arg \hat{P}_N$, and $\hat{Q}, \hat{Q}^{-1}$ are both isometries on L^2. Therefore the operator defined by $\hat{Q}$, given by $u \mapsto Qu$, and $\hat{u} \mapsto \hat{Q}\hat{u}$ is of order 0. This yields

$$\begin{aligned} P_N u = v \quad &\Rightarrow \quad (Q + P_N)u = Qu + v \\ &\Rightarrow \quad u = (Q + P_N)^{-1}(Qu + v) \\ &\Rightarrow \quad \hat{u} = \frac{|\hat{P}_N|}{\hat{P}_N(1 + |\hat{P}_N|)} \cdot (\widehat{Qu} + \hat{v}). \end{aligned}$$

Now by the ellipticity of P_N, and by the fact that it has no lower order terms and that its highest order terms have constant coefficients, we can find a constant A such that the symbol of P_N satisfies

$$|P_N(x, \xi)| = \left| \sum_{|\alpha|=N} a_\alpha \xi^\alpha \right| \geq A|\xi|^N.$$

Therefore

$$\frac{|\hat{P}_N|}{\hat{P}_N(1 + |\hat{P}_N|)} \sim \frac{1}{(1 + |\xi|)^N},$$

and hence $(Q + P_N)^{-1}$ is a smoothing operator of degree N, i.e. an operator of degree $-N$. Now, since u has compact support, it must lie in H_t for some t. Q is an order 0 operator, and hence Qu is also in H_t. Then $Qu + v \in H_{\min(s,t)}$. But then since $u = (Q + P_N)^{-1}(Qu + v)$, we must have that $u \in H_{\min(s+N, t+N)}$ and thus $t = \min(s + N, t + N)$ hence implying that $t = s + N$ and thus $u \in H_{s+N}$ as claimed. $\square$

The next Theorem generalizes the above to operators with lower order terms, and distributions whose support is not necessarily compact.

Theorem 15 *Suppose P is a differential operator of degree N, of the form $P(D) = P_N(D) + R(D)$ where P_N is an operator of the above form, and R is any differential operator of order at most $N - 1$. If $P(D)u = v$ for $u, v \in \mathcal{C}^{-\infty}(\Omega)$, then $v \in H_s^{loc}$ implies $u \in H_{s+N}^{loc}$.*

PROOF. The main idea behind the proof is to use the fact that if $\phi \in \mathcal{C}_c^\infty(\Omega)$ then $[P(D), \phi]$ is an operator of degree $N - 1$, as we have seen before. Using this fact we can apply the previous Theorem. Recall that $u \in H_s^{\text{loc}}$ iff for any $f \in \mathcal{C}_c^\infty(\Omega)$, $fu \in H_s$. Now, by slightly shrinking Ω, taking $\Omega' \subset\subset \Omega$, we can assume without loss of generality that $u \in H_t^{\text{loc}}$ for some t. Then pick any $f \in \mathcal{C}_c^\infty(\Omega)$. By definition, we have $fu \in H_s$. Writing $P_N = P - R$ we have

$$\begin{aligned} P_N(D)(fu) &= (P(D) - R(D))(fu) \\ &= fP(D)u + [P(D), f]u - R(D)(fu) \\ &= fP(D)u + L(D)u, \end{aligned}$$

where L is an operator of degree $N - 1$ with compact support. That L has compact support follows from the fact that $P(D)f$ and $fP(D)$ both have compact support. Therefore we conclude that $L(D)u \in H_{t-N+1}$. Since $v \in H_s$ by assumption, this yields

$$P_N(D)(fu) = fv + L(D)u \in H_{\min(s,t-N+1)}.$$

But then from Theorem 14 above we have $fu \in H_{\min(s+N,t+1)}$ and hence $u \in H_{s+N}^{\text{loc}}$, as desired. □

Note that we maintain the condition that the highest order terms have constant coefficients. To remove this condition we use the fact that a general elliptic operator L is locally "close" to an elliptic operator whose highest order coefficients are constant, and argue as in Theorem 5 in section 2.1.3.

References

[1] Adams, R., "Sobolev Spaces," New York: Academic Press, 1975.

[2] do Carmo, M., "Riemannian Geometry," Boston MA: Birkhauser, 1993.

[3] Folland, G., "Introduction to Partial Differential Equations," Princeton, NJ: Princeton University Press, 1995.

[4] Forster, O., "Lectures on Riemann Surfaces," New York: Springer-Verlag, 1993.

[5] Laugwitz, D., "Differential and Riemannian Geometry," New York: Academic Press, 1965.

[6] McMullen, C., "Class Notes for Math 212b," Cambridge, MA: Harvard University, 1999.

[7] Nirenberg, L., "On Elliptic Partial Differential Equations," Italy, Pisa: Ann. Scuola Norm. Sup. Pisa, 13 (1959) 115-162.

[8] Royden, H., "Real Analysis," New York: Macmillan Publishing Company, 1988.

[9] Rudin, W., "Functional Analysis," Boston, MA: McGraw Hill, 1991.

[10] Spivak, M., "A Comprehensive Introduction to Differential Geometry," Wilmington, DE: Publish or Perish Inc., 1979.