Introduction to Elliptic Operators and Index Theory

Michael Singer, University College London

Contents

CHAPTER 0

Introduction

These notes were produced to accompany an LTCC[1] course given in October 2016. The pre-requisites for the course were not particular clearly stated, but probably included

- Basic familiarity with manifolds and vector bundles, in particular the tangent bundle, differential forms and the de Rham complex;
- The simplest notions of functional analysis, in particular bounded operators on Hilbert spaces and the like.

Almost everything is rather classical (index theory goes back to the 1960s, after all) but I wanted to try to give a treatment which combines the classic papers with various insights that I've gleaned over the years.

0.1. Approach

The approach is to construct a class of linear operators, the pseudodifferential operators, which have very good properties and which are general enough to contain 'approximate inverses' of elliptic differential operators. We start with a black-box approach, based on extending the algebraic properties of the symbol sequence we say for differential operators.

[1]London Taught Course Centre

CHAPTER 1

Elliptic operators

1.1. Definitions

We begin in $\mathbb{R}^n$ (or in an open set X of $\mathbb{R}^n$). Let $p(x,\xi)$ be a function on $X \times \mathbb{R}^n$, smooth in x, polynomial in ξ, of degree $\leqslant k$. If necessary, we write this in terms of its coefficients

$$\sum_{|\alpha|\leqslant k} p_\alpha(x)\xi^\alpha$$

using multi-index notation $\xi^\alpha = \xi_1^{\alpha_1}\cdots\xi_n^{\alpha_n}$, where $\alpha = (\alpha_1,\ldots,\alpha_n)$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$.

A differential operator of order k on X is obtained by formally substituting

$$\xi_j \to D_j = -i\frac{\partial}{\partial x_j}$$

into this polynomial. It is denoted $P = p(x,D)$. The set of all differential operators on X of order $\leqslant k$ is denoted $\mathrm{Diff}^k(X)$. The set of all differential operators on X is denoted $\mathrm{Diff}^*(X)$: it is filtered by the order. Diff^* is a filtered algebra, in that the sum of two operators of order $\leqslant k$ is again of order $\leqslant k$, while under composition, the degrees add:

$$\mathrm{Diff}^k(X) \times \mathrm{Diff}^\ell(X) \longrightarrow \mathrm{Diff}^{k+\ell}(X), \quad (P,Q) \mapsto P\circ Q. \tag{1.1.1}$$

The commutator of operators is also important,

$$[P,Q]u = P(Qu) - Q(Pu);$$

if $\operatorname{ord} P = k$, $\operatorname{ord} Q = \ell$, then the order of $[P,Q]$ is $k+\ell-1$.

Remark 1.1.1. There are a number of reasons for the factor of $-i$ in the definition. It fits in very well with the Fourier transform, which will get much usage later; more fundamentally, D_j is a formally self-adjoint operator and so real polynomials correspond to formally self-adjoint operators (at least if p is independent of x).

Example 1.1.2. The laplacian is associated to the metric, $|\xi|^2$. For a variable metric, it is a little more complicated, namely the p in question is $|\xi|_g^2$ plus lower order times.

Note that $\mathrm{Diff}^0(X) = C^\infty(X)$, viewed as multiplication operators.

Note that if $P \in \mathrm{Diff}^*(X)$, then P defines linear maps

$$P : C^\infty(X) \to C^\infty(X) \tag{1.1.2}$$

and

$$P : C_0^\infty(X) \to C_0^\infty(X) \tag{1.1.3}$$

where the 0 denotes compact support. Differential operators are also *local* in the sense that

$$\operatorname{supp}(Pu) \subset \operatorname{supp}(u) \tag{1.1.4}$$

for any function u.

DEFINITION 1.1.3. If $P \in \operatorname{Diff}^k(X)$, the (principal) symbol $\sigma_k(P)$ is defined to be

$$\sigma_k(P)(x,\xi) = \sum_{|\alpha|=k} p_\alpha(x)\xi^\alpha. \tag{1.1.5}$$

$p(x,D)$ is said to be *elliptic* at x if $\sigma_k(P)(x,\xi) \neq 0$ for all (real) non-zero ξ.

EXAMPLE 1.1.4. The Laplacian is elliptic with symbol $|\xi|^2$.

THEOREM 1.1.5. *For each integer k and any given open subset $X \subset \mathbb{R}^n$, there is an exact sequence*

$$0 \to \operatorname{Diff}^{k-1}(X) \to \operatorname{Diff}^k(X) \xrightarrow{\sigma_k} S^k(X) \to 0. \tag{1.1.6}$$

Here $S^k(X)$ is the space of polynomials homogeneous of degree k in ξ, with coefficients smooth functions in X (the typical element of $S^k(X)$ looks like the RHS of (1.1.5)).

Furthermore, the symbol map is an algebra homomorphism: if

$$\operatorname{ord} P = k,\ \operatorname{ord} Q = \ell, \tag{1.1.7}$$

then

$$\sigma_{k+\ell}(P \circ Q) = \sigma_k(P)\sigma_\ell(Q). \tag{1.1.8}$$

PROOF. The exactness of (1.1.6) should be pretty clear. For (1.1.8) it is enough to consider the case of monomials

$$p(x,\xi) = p_\alpha \xi^\alpha,\ q(x,\xi) = q_\beta(x)\xi^\beta$$

where $|\alpha| = k$, $|\beta| = \ell$. Then if u is a smooth function,

$$\begin{aligned} P(x,D)(q(x,D)u) &= p_\alpha(x)D^\alpha(q_\beta(x)D^\beta u) && (1.1.9)\\ &= p_\alpha(x)q_\beta(x)D^{\alpha+\beta}u + r(x,D)u && (1.1.10)\end{aligned}$$

where r is of order $< k+\ell$.

EXERCISE 1.1.6. Show that the piece of $r(x,\xi)$ homogeneous of degree $k+\ell-1$ in ξ can be written

$$-i\partial_\xi p \partial_x q = -i\sum_{j=1}^n \frac{\partial p}{\partial \xi_j}\frac{\partial q}{\partial x_j}. \tag{1.1.11}$$

The assertion (1.1.8) follows by $\mathbb{C}$-bilinearity of composition. □

In algebraic terms, (1.1.8) has the interpretation that the algebra $S^*(X)$ of polynomials in ξ with coefficients smooth functions of x is the *associated graded algebra* of the filtered algebra Diff^*.

Note the interesting fact that σ_k is a homomorphism of algebras but the algebra structure of $S^*(X)$ is much simpler than that of $\operatorname{Diff}^*(X)$—in particular it is abelian (at least for this setting of scalar differential operators).

Obviously there is much more to a differential operator than its symbol, but we shall see that (once these notions have been suitably extended) basic questions about differential operators on compact manifolds are answered by knowledge of the principal symbol. For example, if the symbol is invertible, and the manifold is compact, the operator is also 'essentially invertible', that is, up to finite-dimensional errors.

Exercise 1.1.7. Show that the $\overline{\partial}$-operator

$$\overline{\partial} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)$$

is elliptic as an operator on complex-valued function in $\mathbb{R}^2$.

Exercise 1.1.8. Show that the grad-div-curl operator

$$(\mathbf{a}, \phi) \longmapsto (\operatorname{curl}\mathbf{a} + \nabla\phi, \operatorname{div}\mathbf{a}) \tag{1.1.12}$$

is elliptic acting on pairs $(\mathbf{a}, \phi)$ of vector fields and functions on $\mathbb{R}^3$. (For operators acting on vector-valued functions, the definition of ellipticity is that the symbol, which is now a matrix whose entries are functions of (x, ξ), should be invertible for all non-zero real ξ.)

Exercise 1.1.9. Given any operator $P \in \operatorname{Diff}^k(X)$ we can go back to $p(x, \xi)$ by 'oscillatory testing'. Show that

$$p(x, \xi) = e^{-i\langle x,\xi\rangle}P(e^{i\langle x,\xi\rangle}). \tag{1.1.13}$$

Show further that

$$\sigma_k(P)(x, \xi) = \lim_{|\xi|\to\infty} |\xi|^{-k}p(x, \xi) \tag{1.1.14}$$

Exercise 1.1.10. The point of view I have taken here[1] is strongly motivated by quantum mechanics. The function $p(x, \xi)$ is thought of as a 'classical observable' that is, by definition, a smooth function of position x and momentum ξ. Then $p(x, D)$ is the corresponding 'quantum observable' operating (unboundedly) on the quantum state space $L^2(\mathbb{R}^n)$.

Let

$$P = p(x, D), Q = q(x, D) \tag{1.1.15}$$

where $p = p_k(x; \xi) + \cdots$ and $q = q_\ell(x; \xi) + \cdots$, these being the terms homogeneous of degree m and m' respectively, and the $\cdots$ denoting terms of lower degree in ξ. Show that

$$[P, Q] = r(x, D) \tag{1.1.16}$$

where $r = r_{k+\ell-1} + \cdots$ and

$$r_{k+\ell-1} = -i\{p_k, q_\ell\} \tag{1.1.17}$$

is the *Poisson bracket* of p_m and $q_{m'}$, defined generally by

$$\{f, g\} = \sum_{j=1}^{n}\left(\frac{\partial f}{\partial \xi_j}\frac{\partial g}{\partial x_j} - \frac{\partial g}{\partial \xi_j}\frac{\partial f}{\partial x_j}\right) \tag{1.1.18}$$

[This follows immediately from Exercise 1.1.6.]

EXERCISE 1.1.11. For composition of differential operators $P = p(x,D)$ and $Q = q(x,D)$, of orders k and ℓ, obtain the following formula:

$$PQ = R = r(x,D), \tag{1.1.19}$$

where

$$r(x,\xi) = \exp\left(\frac{i}{2}\varpi(D_x, D_\xi, D_y, D_\eta)\right) p(x,\xi)q(y,\eta)|_{y=x,\eta=\xi}\,. \tag{1.1.20}$$

Here

$$\varpi(D_x, D_\xi, D_y, D_\eta) = \langle D_\xi, D_y\rangle - \langle D_\eta, D_x\rangle$$

and the formula is interpreted by (formal) expansion of the exponential in power series. The sum will be finite because p and q are polynomial in ξ.

1.2. Invariance properties

In exercise 1.1.10, it was noted that $p(x,\xi)$ should be viewed as a function of position x and momentum ξ. The geometry underlying the Poisson bracket is symplectic geometry. This strongly suggests that from a more invariant point of view, if $x \in V$ (a vector space) then ξ should be in the dual space V^*. On $V \times V^*$ there is a natural symplectic form and in coordinates the corresponding Poisson bracket is (1.1.18).

Let us explain why this is indeed the correct point of view. We shall do this in the context of manifolds and vector bundles.

Let M be a smooth connected oriented manifold and let E and F be complex vector bundles over M. There are many ways to define the space of differential operators acting between sections of E and sections of F. To make the invariance more obvious, we use connections.

Let

$$C^\infty(M,E) = \{\text{smooth sections of } E \text{ over } M\}. \tag{1.2.1}$$

If k is a positive integer, denote by $\Omega^k(M,E)$ the space of E-valued k-forms; we have

$$\Omega^k(M,E) = C^\infty(M, E\otimes\Lambda^k T^*M), \Omega^0(M,E) = C^\infty(M,E). \tag{1.2.2}$$

A connection ∇ in (or on) E is (for us) a linear operator

$$\nabla : \Omega^0(M,E) \longrightarrow \Omega^1(M,E) \tag{1.2.3}$$

which satisfies the Leibniz rule

$$\nabla(f\otimes s) = \mathrm{d}f\otimes s + f\otimes\nabla s \text{ for all } f\in C^\infty(M),\ s\in C^\infty(M,E). \tag{1.2.4}$$

Notice that this is a *global* definition, but a standard argument (using the Leibniz rule, and suitable choices of f) shows that it is really local. More precisely, the value of ∇s at a point x in M depends only upon s and its first derivatives at x (in any local trivialization near x). More precisely: choose a local trivialization of E in an open set U. If E is of rank N, then for each local section $s \in C^\infty(U,E)$, we have a $\mathbb{C}^N$-valued function $\widetilde{s}$ defined in U. Then there is an $N\times N$ matrix of 1-forms, A, defined over U, such that

$$\widetilde{\nabla s} = \mathrm{d}\widetilde{s} + A\widetilde{s} \tag{1.2.5}$$

Because ∇ is locally determined, for any open set $U \subset M$ (not necessarily a set on which E is trivial), there is an induced linear operator

$$\nabla : \Omega^0(U, E) \longrightarrow \Omega^1(U, E) \tag{1.2.6}$$

which we do not distinguish from the original one.

Definition 1.2.1. If E and F are complex vector bundles the space of symbols of order k from E to F is the space of sections of the vector bundle $S^kTM \otimes \mathrm{Hom}(E, F)$.

The space of all symbols of order k is denoted $S^k(X; E, F)$.

A symbol $\sigma \in S^k(M; E, F)$ can be viewed in various ways. We have

$$\sigma_k \in C^\infty(M, \mathrm{Hom}(S^kT^*M \otimes E, F)) = C^\infty(M, S^kTM \otimes \mathrm{Hom}(E, F). \tag{1.2.7}$$

Recall also that if V is any vector space, the k-fold symmetric tensor product S^kV is canonically identifiable with the set of polynomials on V^*, homogeneous of degree k. So at any given point x of M, σ gives a function

$$\sigma(x, \xi) \in \mathrm{Hom}(E_x, F_x), x \in M, \xi \in T^*M \tag{1.2.8}$$

homogeneous of degree k in ξ, that is,

$$\sigma(x, t\xi) = t^k\sigma(x, \xi) \text{ for all } t \in \mathbb{R}. \tag{1.2.9}$$

In local trivializations, $\sigma(x, \xi)$ is identified with a matrix, each of whose entries is smooth in x and a homogeneous polynomial of degree k in ξ.

A more invariant and global interpretation of $\sigma(x, \xi)$ is as follows:

$$\sigma \in C^\infty(T^*M, \mathrm{Hom}(\pi^*E, \pi^*F)),$$

where $\pi : T^*M \to M$ is the projection, with the additional condition that σ is homogeneous of degree k on the fibres of π.

Using connections and symbols, we extend the definitions in $\mathbb{R}^n$ to any manifold. Let's warm up with the case of first-order operators:

Definition 1.2.2. The space of first-order differential operators $\mathrm{Diff}^1(M; E, F)$ is defined to consist of all operators of the form

$$L = \sigma_1 \circ (-i\nabla) + \sigma_0 \tag{1.2.10}$$

where σ_j are symbols of order j from E to F and ∇ is an arbitrarily chosen connection in E.

The first term in (1.2.10) is the composite

$$C^\infty(M; E) \xrightarrow{-i\nabla} C^\infty(M; T^* \otimes E) \xrightarrow{\sigma_1} C^\infty(M; F) \tag{1.2.11}$$

where we think of σ_1 as a section of $\mathrm{Hom}(T^*M \otimes E, F)$ as above. Hence L, as defined in (1.2.10) does map sections of E to sections of F, as suggested by the notation!

Proposition 1.2.3. *The definition of* $\mathrm{Diff}^1(M; E, F)$ *is independent of the choice of connection* ∇*.*

With respect to local trivializations of E *and* F *over some open set* U*, every* $L \in \mathrm{Diff}^1(M; E, F)$ *has the form* $\ell(x, D)$*, where* ℓ *is an* $N \times N'$*-valued function of* (x, ξ)*, polynomial of degree* $\leqslant 1$ *in* ξ *and smooth in* x*. Here* N *and* N' *are the ranks of* E *and* F*.*

PROOF. The difference of two connections is algebraic, $\nabla' = \nabla + \tau$, where τ is a section of $\mathrm{End}(E) \otimes T^*$. Thus

$$-i\sigma_1 \circ \nabla + \sigma_0 = L = -i\sigma_1' \circ \nabla' + \sigma_0' \tag{1.2.12}$$

is solved by

$$\sigma_1' = \sigma_1,\ \sigma_0' - i\sigma_1' \circ \tau = \sigma_0. \tag{1.2.13}$$

So a change in connection can always be absorbed by a change in σ_0. It is clear from the expression for ∇ in local coordinates (see (1.2.5)) that locally

$$L = -i\sigma_1 \circ \mathrm{d} + (\sigma_1 \circ A + \sigma_0)$$

which is of the stated form. □

REMARK 1.2.4. The annoying factor of i in the definition is put in so that real symbols correspond to self-adjoint operators and for consistency. Other authors adopt different conventions and put the i elsewhere.

Notice that the first-order symbol σ_1 does not change under change of connection. We have also given it an invariant geometric interpretation as a section of a certain bundle.

We now extend the above definition from first order operators to operators of arbitrary order. For this, we need to choose a connection also on TM. It is natural to choose this to be *torsion-free*. Recall that there is then an induced connection on all 'tensor bundles', in particular the symmetric tensor products S^jT^*M. Combining this connection with the original one on E, we obtain a connection on $E \otimes S^JT^*$, for every j. All these connections will be denoted by ∇.

In particular, with this choice made, we obtain an operator $\nabla^{(2)}$

$$\nabla^{(2)} : C^\infty(M, E) \longrightarrow C^\infty(M, E \otimes S^2T^*M) \tag{1.2.14}$$

as the composite

$$C^\infty(M, E) \stackrel{\nabla}{\to} C^\infty(M, E \otimes T^*) \stackrel{\nabla}{\to} C^\infty(M, E \otimes T^* \otimes T^*) \to C^\infty(M, E \otimes S^2T^*) \tag{1.2.15}$$

where the last arrow is the algebraic operation of symmetrization. (If we take the skew part instead of the symmetric part, we simply get the curvature operator. The torsion-free condition is needed for this assertion to be correct.)

Iterating the above construction, we may define, for each j,

$$\nabla^{(j)} : C^\infty(M, E) \longrightarrow C^\infty(M, E \otimes S^jT^*). \tag{1.2.16}$$

Given a finite collection of bundle morphisms

$$\sigma_j : E \otimes S^jT^* \longrightarrow F, \tag{1.2.17}$$

in other words symbols from E to F of order j, we define a differential operator

$$P = \sum_{j=0}^{k} (-i)^j \sigma_j \circ \nabla^{(j)} \tag{1.2.18}$$

of order k.

NOTATION 1.2.5. The set of all differential operators of order k from E to F is denoted $\mathrm{Diff}^k(M; E, F)$, $\mathrm{Diff}^k(E, F)$ or even Diff^k if no confusion can result from the notational simplifications.

The set of all differential operators from E to F (of arbitrary but finite order) is denoted $\mathrm{Diff}^*(M; E, F)$.

It is left to the reader to check that $\mathrm{Diff}^k(M;E,F)$ is independent of the choice of connection used to define the $\nabla^{(j)}$.

Note that $\mathrm{Diff}^*(M;E,E)$ is a graded ring under composition, and that the corresponding Lie bracket given by commutator $[P_1,P_2]$ has order $\leqslant k_1+k_2$ if k_1 and k_2 are the orders of P_1 and P_2.

Exercise 1.2.6. Show that the set of differential operators $\mathrm{Diff}^*(M;E,F)$ is independent of the choice of connection used in (1.2.17).

Exercise 1.2.7. An alternative definition is that in any local coordinates and trivializing charts, L is given locally by a differential operator in the usual sense in an open subset of $\mathbb{R}^n$. Show that this is equivalent to the definition using a connection.

If $\widetilde{\nabla}$ is a new connection and we write

$$\sum_{j=0}^{k}(-i)^j\widetilde{\sigma}_j\circ\widetilde{\nabla}^{(j)}=P=\sum_{j=0}^{k}(-i)^j\sigma_j\circ\nabla^{(j)} \tag{1.2.19}$$

then it turns out that

$$\widetilde{\sigma}_k=\sigma_k. \tag{1.2.20}$$

This is called the *principal symbol* of L.

Definition 1.2.8. The operator $P\in\mathrm{Diff}^k(M;E,F)$ is said to be *elliptic* if and only if $\sigma_k(P)$ is an invertible endomorphism $\pi^*E\to\pi^*F$ over $T^*M\setminus\{0\}$. By the homogeneity of σ_k, it is equivalent to demand invertibility over the cosphere bundle inside T^*M, or for all ξ sufficiently large.

.

Proposition 1.2.9. *There is an exact sequence*

$$0\longrightarrow\mathrm{Diff}^{k-1}(M;E,F)\longrightarrow\mathrm{Diff}^k(M;E,F)\longrightarrow S^k(M;E,F)\to 0 \tag{1.2.21}$$

If E and F are equipped with hermitian structures and M is equipped with a smooth measure so that the formal adjoint L^* of L is defined, then $\sigma_k(L^*)=\sigma_k(L)^*$ (for all real ξ).

If G is a third vector bundle and

$$P\in\mathrm{Diff}^k(M;F,G),\ \ Q\in\mathrm{Diff}^\ell(M;E,F) \tag{1.2.22}$$

so that $PQ\in Diff^{k+\ell}(M;E,G)$, then

$$\sigma_{k+\ell}(PQ)=\sigma_k(P)\sigma_\ell(Q). \tag{1.2.23}$$

Exercise 1.2.10. Let Δ_g be the Laplacian acting on functions on M, defined by some riemannian metric g. Show that $\sigma_2(\Delta_g)=|\xi|_g^2$, where this is the dual metric on T^*M.

Exercise 1.2.11. Suppose that E and F are equipped with hermitian structures (a hermitian metric in each fibre, smoothly varying with the fibre). Given also a smooth volume-form $\mathrm{d}\mu$ on M, we define L^2 inner products on $C^\infty(M;E)$ and similarly $C^\infty(M;F)$ by

$$\langle s, s' \rangle = \int_M (s, s')\,\mathrm{d}\mu \tag{1.2.24}$$

where the round brackets denote the pointwise inner product, and we assume either that M is compact or if not that at least one of s and s' has compact support. The *formal adjoint* of $P \in \mathrm{Diff}^k(M;E,F)$ is defined by the equation

$$\langle Ps, t \rangle = \langle s, P^* t \rangle \tag{1.2.25}$$

for all $s \in C_0^\infty(M,E)$, $t \in C_0^\infty(M,F)$.

Show that P^* is a differential operator in $\mathrm{Diff}^k(M;F,E)$ and that $\sigma_k(P^*) = \sigma_k(P)^*$.

Deduce that the 'rough Laplacian' $\nabla^*\nabla \in \mathrm{Diff}^2(M;E,E)$ is elliptic.

CHAPTER 2

Distributions and the Fourier Transform

2.1. Motivation for distributions

2.1.1. The Dirichlet Problem. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary. Suppose we wish to solve the problem

$$\Delta u = f \text{ in } \Omega, \quad u|\partial\Omega = 0. \tag{2.1.1}$$

with the boundary condition

$$u|\partial\Omega = 0. \tag{2.1.2}$$

Here f is say a given smooth function defined in Ω. There is a classical approach (the Dirichlet principle) along the following lines. Let

$$E(u) = \int_\Omega (\frac{1}{2}|\nabla u|^2 - fu)\,\mathrm{d}x \tag{2.1.3}$$

where u is (in the first instance) in $C_0^1(\Omega)$, the space of continuously differentiable functions which satisfy the boundary condition (2.1.2). If u_0 minimizes $E(u)$ over all $u \in C_0^1(\Omega)$, then for any smooth $\phi \in C_0^1(\Omega)$,

$$\frac{\mathrm{d}}{\mathrm{d}t}E(u_0 + t\phi)\Big|_{t=0} = 0. \tag{2.1.4}$$

This implies that

$$\int_\Omega ((\nabla u, \nabla\phi) - (f, \phi))\;\mathrm{d}x = 0. \tag{2.1.5}$$

If we knew that u were C^2, we could integrate by parts to obtain

$$\int_\Omega (\Delta u - f)\phi\,\mathrm{d}x = 0 \text{ for all } \phi \text{ with } \phi|\partial\Omega = 0. \tag{2.1.6}$$

If $\Delta u - f \in C^0$, this implies that $\Delta u = f$ in Ω.

We would like to use this idea to prove the existence of a solution u of the problem (2.1.3). To push it through, however (i.e. to find a minimizing u_0) we need to make sense of the idea that ∇u can be defined even if u is not differentiable: we somehow need to work with $u \in L^2(\Omega)$ such that $\nabla u \in L^2(\Omega)$. In fact, we do this exactly as suggested in the above calculation. Namely if $u \in L^2(\Omega)$, we define the derivatives $\partial_j u$ as a *functional* on $C_0^\infty(\Omega$ namely

$$\partial_j u[\phi] = -\int_\Omega u\partial_j\phi. \tag{2.1.7}$$

By Cauchy–Schwartz this is well-defined if $u \in L^2(\Omega)$. We can then say that $\partial_j u \in L^2(\Omega)$ if this functional is given by integration of ϕ against an L^2 function.

2.1.2. Green's Functions. It's been known for a long time that if $f \in C_0^\infty(\mathbb{R}^n)$, then the solution u which decays at ∞ to *Poisson's equation*

$$\Delta u = f \tag{2.1.8}$$

is given by

$$u(x) = \int G_n(x-y) f(y)\, \mathrm{d}y \tag{2.1.9}$$

where

$$G_n(x) = c_n |x|^{2-n}, \quad \text{for } n \geqslant 3, \tag{2.1.10}$$

and c_n is some constant. The point about G_n (the Green's function for the Laplacian in $\mathbb{R}^n$) is that it solves $\Delta G_n = 0$ away from $x = 0$ and has 'the right singularity' at $x = 0$.

In fact, the reason why (2.1.9) holds is the formula

$$\Delta_x G_n(x-y) = \delta(x-y) \tag{2.1.11}$$

where $\delta(x-y)$ is the Dirac δ-function, which has the property

$$\int \delta(x-y)\phi(y)\, \mathrm{d}y = \phi(x) \tag{2.1.12}$$

for all smooth functions ϕ. There is of course no function δ which has this property but that is the only problem: for each $y \in \mathbb{R}^n$ we have the evaluation map

$$\delta_y : C^\infty(\mathbb{R}^n) \to \mathbb{R}, \delta_y(\phi) = \phi(y). \tag{2.1.13}$$

Moreover, (2.1.11) makes perfect sense if we evaluate or 'test' against a function $\phi \in C_0^\infty(\mathbb{R}^n)$ and integrate by parts:

$$\int G_n(x-y)\Delta_x \phi(x)\, \mathrm{d}x = \phi(y) \tag{2.1.14}$$

for the singularity of $G_n(x-y)$ along the diagonal is integrable.

These two examples motivate the need to be able to differentiate objects that are more general than classically differentiable functions. The trick is to think in terms of duality, as already hinted at in (2.1.14) and (2.1.7).

2.1.3. A quick introduction to distributions. The theory of *distributions*, in its rigorous and present-day form, was systematically developed by Laurent Schwartz in the late 1940s (though has roots stretching back go George Green in the 1830s). A systematic treatment can be found in Volume I of Hörmander's series on the analysis of partial differential operators.

The space of distributions $C^{-\infty}(\mathbb{R}^n)$ is defined as the dual space of $C_0^\infty(\mathbb{R}^n)$. It is common to refer to elements of C_0^∞ as 'test functions'. There is a continuity condition involved in the definition of distribution which we shall not dwell upon (though it is of course important).

Example 2.1.1. We note that $C^\infty(\mathbb{R}^n) \subset C^{-\infty}(\mathbb{R}^n)$ for if $f \in C^\infty$, it defines a functional by

$$f[\phi] = \langle f, \phi \rangle = \int f(x)\phi(x)\, \mathrm{d}x. \tag{2.1.15}$$

Example 2.1.2. Evaluation maps. We have already seen the Dirac δ-function δ_y which is the function of evaluation $\phi \mapsto \phi(y)$. More generally, we have evaluation functionals of the form

$$\phi \mapsto \delta_y(p(D)\phi) \tag{2.1.16}$$

where $p(D)$ is a differential operator as before.

2.1.4. Differentiation of distributions. In 1 dimension, the derivative of a distribution u is defined by the formula

$$\langle u', \phi\rangle := -\langle u, \phi'\rangle. \tag{2.1.17}$$

since this is the correct formula if u is (continuously) differentiable and $\langle\cdot,\cdot\rangle$ is integration.

EXERCISE 2.1.3. Show that if u is a distribution on $\mathbb{R}$ and $u' = 0$ in the sense of (2.1.17), then u is a constant.

Similarly in $\mathbb{R}^n$,

$$\partial_j u(\phi) := -u(\partial_j\phi) \tag{2.1.18}$$

defines the derivative $\partial_j u$ of any distribution u.

As well as being able to differentiate distributions, one can also multiply by smooth functions: if $f \in C^\infty(\mathbb{R}^n)$, the distribution fu is defined by the formula

$$\langle fu, \phi\rangle = \langle u, f\phi\rangle. \tag{2.1.19}$$

for any test-function $\phi \in C_0^\infty$. In particular, $C^{-\infty}(\mathbb{R}^n)$ is a *module* over $C^\infty(\mathbb{R}^n)$.

Note also that there is a definition of $C^{-\infty}(X)$, where X is an open subset of $\mathbb{R}^n$—the space of continuous linear functionals on $C_0^\infty(X)$. Just like functions, distributions can be patched together in the following sense. If X is covered by open sets X_j, and a collection of distributions $u_j \in C^{-\infty}(X_j)$ has the property that $u_j = u_k$ as elements of $C^{-\infty}(X_j \cap X_k)$, then there is a unique distribution u in X with $u|X_j = u_j$.

EXERCISE 2.1.4. Prove the equation in $\mathbb{R}^3$

$$\Delta G_3 = \delta_0, G_3(x) = \frac{1}{4\pi|x|}. \tag{2.1.20}$$

EXERCISE 2.1.5. If T is an isometry of $\mathbb{R}^n$ and u is a distribution, suggest a definition of T^*u which extends pull-back on functions.

2.1.5. L^p and Sobolev spaces inside $C^{-\infty}$. Note that we can regard $C^0(\mathbb{R}^n)$ as a subspace of $C^{-\infty}(\mathbb{R}^n)$. If $f \in C^0$, the definition we want to make is

$$\langle f, \phi\rangle = \int f(x)\phi(x)\,\mathrm{d}x, \phi \in C_0^\infty. \tag{2.1.21}$$

The integral is certainly convergent. Moreover, if the RHS is known for every ϕ then $f \in C^0$ is uniquely determined. (In other words, the linear functional $\phi \mapsto \langle f, \phi\rangle$ uniquely determines $f \in C^0$.) To see this, suppose if possible that $f \in C^0$, $f(x_0) \neq 0$ but $\langle f, \phi\rangle = 0$ for all ϕ. By picking $\phi(x_0) = 1$, everywhere non-negative, and supported in a tiny ball $B(x_0, \delta)$ on which f doesn't change sign, we see that $\langle f, \phi\rangle \neq 0$ a contradiction. The existence of such a tiny ball follows as in elementary analysis courses from the continuity of f at x_0.

Thus we can unambiguously write $C^0 \subset C^{-\infty}$.

A harder result of the same kind is

PROPOSITION 2.1.6. *The space L^1_{loc} of locally L^1 functions is, in a natural way, a subspace of $C^{-\infty}$.*

PROOF. (Cf. Hörmander, Linear Partial Differential Operators I, Theorem 1.2.5). The point is to show that if $f \in L^1_{\mathrm{loc}}$ and

$$\int f(x)\phi(x)\,\mathrm{d}x = 0 \text{ for all } \phi \in C_0^\infty \tag{2.1.22}$$

then $f = 0$ almost everywhere. The proof uses the result that

$$\lim_{t\to 0}\frac{1}{t^n}\int_{|x-x_0|<t}|f(x)-f(x_0)|\,\mathrm{d}x = 0 \tag{2.1.23}$$

for almost all x_0 in place of the continuity in the previous discussion. □

More generally, $L^p(\mathbb{R}^n)$, $p \geqslant 1$, may be viewed as a subspace of $C^{-\infty}(\mathbb{R}^n)$. By Hölder's inequality implies that $L^p(\mathbb{R}^n) \subset L^1_{\mathrm{loc}}$ and since we have already agreed that $L^1_{\mathrm{loc}} \subset C^{-\infty}$, it follows that L^p is a well-defined subspace of $C^{-\infty}$.

By the device of convolution, one can show that any $f \in L^p(\mathbb{R}^n)$ can be approximated by $f_\varepsilon \in C^\infty$, i.e.

$$\|f_\varepsilon - f\|_p \to 0 \text{ as } \varepsilon \to 0.$$

In fact, the space $L^p(\mathbb{R}^n)$ may be viewed as the metric-space completion of C_0^∞ with respect to the p-norm inside the space of distributions $C^{-\infty}$.

Now that we understand $L^2(\mathbb{R}^n) \subset C^{-\infty}$, we may define the Sobolev space $H^s(\mathbb{R}^n)$ where initially s is a non-negative integer, to be the set of distributions u in $\mathbb{R}^n$ such that

$$D^\alpha u \in L^2(\mathbb{R}^n) \text{ for all } |\alpha| \leqslant k.$$

From its definition, H^s is a Hilbert space with the inner product

$$\langle u, v\rangle_k = \int \sum_{|\alpha|\leqslant k}(D^\alpha u, D^\alpha v)\,\mathrm{d}x$$

(Hermitian pointwise innner product in the integrand.) Note that there are many equivalent inner products on H^s all of which give the same topological vector space—we shall see this when we look at the Fourier transform of H^s below.

We can define H^{-s} to be the set of all distributions which are finite sums of the form $D^\alpha f$, $f \in L^2$, with $|\alpha| \leqslant s$. Then the pairing $(\phi, \psi) \mapsto \int \phi(x)\psi(x)$ extends to identify H^s and H^{-s} as dual spaces.

These spaces can be localized (i.e. defined for open sets of $\mathbb{R}^n$). We shall also define them using the Fourier transform in the next section.

Although smooth functions are very nice (infinitely nice!) it is very useful to be able to view PDEs as mapping between Sobolev spaces. Not only does it allow Hilbert Space theory to be brought to bear in linear problems, it allows the implicit and inverse function theorems to be applied in nonlinear ones.

It is clear from the definition that if $P = p(x, D)$ is a differential operator of order $\leqslant k$, then by duality P defines a mappings

$$P : C^{-\infty}(\mathbb{R}^n) \longrightarrow C^{-\infty}(\mathbb{R}^n) \text{ and } P : C_0^{-\infty}(\mathbb{R}^n) \longrightarrow C_0^{-\infty}(\mathbb{R}^n). \tag{2.1.24}$$

Also for any integer s,

$$P : H^s(\mathbb{R}^n) \longrightarrow H^{s-k}(\mathbb{R}^n). \tag{2.1.25}$$

This may be thought of as an abuse of notation. After all, the definition of a mapping should include its domain. In (2.1.25) we are really speaking of the *restriction* of the first of the maps in (2.1.24) to the subspace H^s, and are claiming that this restriction is continuous (i.e. bounded) between these normed linear spaces.

2.2. Support and singular support

If $u \in C^{-\infty}$ and $\Omega \subset \mathbb{R}^n$ is open, then we define $u|\Omega$, the restriction of u to Ω simply by restricting u to act on $C_0^\infty(\Omega)$. We say that u vanishes in Ω if $u|\Omega$ is identically 0 (as a functional).

Definition 2.2.1. If $u \in C^{-\infty}$, $\operatorname{supp}(u)$, the *support* of u is defined as follows: $x_0 \in \operatorname{supp}(u)$ if there is no open neighbourhood of x_0 to which the restriction of u is zero.

In other words, $a \notin \operatorname{supp}(u)$ if there is an open neighbourhood V of a such that $u|V = 0$. It follows from the definition that $\operatorname{supp}(u)$ is a closed set.

Similarly, we define *singular support*:

Definition 2.2.2. If $u \in C^{-\infty}$, then x_0 is in the singular support of u if and only if there is no open neighbourhood of x_0 to which the restriction of u is smooth. The set of all such points is *closed* and is denoted $\operatorname{sing-supp}(u)$.

Example 2.2.3. The δ-function δ_0 is supported at 0. And this is also its singular support.

2.3. The Fourier Transform

2.3.1. The Schwartz space and its dual. Recall the Schwartz space $\mathscr{S}(\mathbb{R}^n)$ consisting of functions $f \in C^\infty$, all of whose derivatives are rapidly decreasing at ∞. The Schwartz space is intermediate between C_0^∞ and C^∞,

$$C_0^\infty(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n) \subset C^\infty(\mathbb{R}^n). \tag{2.3.1}$$

The topological dual, $\mathscr{S}'(\mathbb{R}^n)$, called the space of tempered distributions, is accordingly intermediate between the distributions and those with compact support:

$$C_0^{-\infty}(\mathbb{R}^n) \subset \mathscr{S}'(\mathbb{R}^n) \subset C^{-\infty}(\mathbb{R}^n). \tag{2.3.2}$$

Polynomials lie in $\mathscr{S}'$ but not $\mathscr{S}$: if p is a polynomial, then

$$\phi \longmapsto \int p(x)\phi(x)\,\mathrm{d}x$$

is a continuous linear functional on $\mathscr{S}$ because of the rapid decrease of f.

We didn't go in to the definition of the topology on C_0^∞ but we shall say a few words about that of $\mathscr{S}$. For each non-negative integer N, define, for example,

$$\|f\|_N = \max_{|\alpha|+|\beta|\leqslant N} \sup |x^\alpha D^\beta f| \tag{2.3.3}$$

This is a countable family of norms and, in the usual way, we can turn them into a metric,

$$d(f,g) = \sum_{N=0}^{\infty} \frac{1}{2^N} \frac{\|f-g\|_N}{1+\|f-g\|_N}.$$

The Schwartz space is defined to be the set of all f for which each of the $\|f\|_N$ are finite, and it is made into a metric space with the above metric. Equivalently, $f_n \to f$ in $\mathscr{S}$ if for each N,

$$\|f_n - f\|_N \to 0 \text{ as } n \to \infty. \tag{2.3.4}$$

The space of tempered distributions is given the weak topology: so a sequence $u_n \to 0$ in $\mathscr{S}'$ if and only if $u_n[\phi] \to 0$ as $n \to \infty$ for each fixed $\phi \in \mathscr{S}$.

2.3.2. The Fourier Transform. If $f \in \mathscr{S}(\mathbb{R}^n)$, define

$$\widehat{f}(\xi) = \int e^{-i\langle \xi, x\rangle} f(x)\, \mathrm{d}x, \tag{2.3.5}$$

the Fourier transform in $\mathscr{S}$. This maps $\mathscr{S} \to \mathscr{S}$ because the Fourier transform interchanges differentiation and multiplication:

LEMMA 2.3.1. *If $f \in \mathscr{S}(\mathbb{R}^n)$ then the Fourier transform of $D_j f$ is $\xi_j \widehat{f}(\xi)$ and the Fourier transform of $x_j f(x)$ is $-D_j \widehat{f}(\xi)$.*

We also denote (2.3.5) by $\mathscr{F}$. From the lemma it follows that $\mathscr{F}$ is a mapping from $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}(\mathbb{R}^n)$. The second $\mathbb{R}^n$ is really dual to the first one—it is the space of 'frequencies'.

EXERCISE 2.3.2. Show that if $f \in \mathscr{S}$, then given multi-indices α and β, the semi-norm $\sup |\xi^\alpha D_\xi^\beta \widehat{f}(\xi)|$ can be estimated in terms of the norm $\|f\|_N$ from (2.3.3) for some N.

THEOREM 2.3.3. *The Fourier transform $\mathscr{F}$ is an isomorphism $\mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ with inverse*

$$\mathscr{F}^{-1} : g \longmapsto \frac{1}{(2\pi)^n} \int e^{i\langle x, \xi\rangle} g(\xi)\, d\xi. \tag{2.3.6}$$

.

PROOF. Let $\mathscr{G}$ be the operator defined in (2.3.6). By the same argument as for $\mathscr{F}$, $\mathscr{G}$ maps $\mathscr{S}$ to $\mathscr{S}$, and $\mathscr{G}$ applied to $D_j g(\xi)$ is equal to $-x_j \mathscr{G}[g](x)$, while $\mathscr{G}$ applied to $\xi_j g(\xi)$ is equal to $\mathscr{G}[D_j g](x)$. Consider $T = \mathscr{G} \circ \mathscr{F}$. Then

$$T(x_j \phi(-)) = \mathscr{G}[-D_j \widehat{\phi}(-)] = x_j T\phi(x)$$

and

$$T(D_j \phi(-)) = \mathscr{G}[\xi_j \widehat{\phi}(-)] = D_j T\phi(x).$$

Thus T commutes with all multiplication and differentiation operators, and it follows from this (exercise) that T must be a multiple of the identity.

The multiple can be determined by evaluation of one particular example and the Gaussian $\phi(x) = e^{-|x|^2/2}$ is a good one to choose: this is because

$$(\partial_j + x_j)\phi(x) = 0 \text{ for each } j$$

so the Fourier transform $\widehat{\phi}(\xi)$ satisfies the same equations and so is a fixed multiple of $e^{-|\xi|^2/2}$. Explicit calculation with ϕ shows that $T[\phi] = \phi$, so T is equal to the identity. □

By duality, we extend the Fourier transform to $\mathscr{S}'$, the tempered distributions, through the formula

$$\widehat{u}[\phi] = u[\widehat{\phi}], \phi \in \mathscr{S}. \tag{2.3.7}$$

This is justified because if $u \in \mathscr{S}$, then the LHS is equal to

$$\int u(x) e^{-i\langle x,\xi\rangle} \phi(\xi)\, \mathrm{d}x \mathrm{d}\xi \tag{2.3.8}$$

and we can interchange the order of integration here: when the ξ-integral is performed first, we obtain

$$\int u(x)\widehat{\phi}(x)\, \mathrm{d}x, \tag{2.3.9}$$

which is the RHS of (2.3.7).

Example 2.3.4. The Fourier transform of δ_0 is equal to 1. Proof:

$$\widehat{\delta_0}[\phi] = \delta_0[\widehat{\phi}] = \widehat{\phi}(0) = \int \phi(x)\, \mathrm{d}x. \tag{2.3.10}$$

More generally, the Fourier transform of $D^\alpha \delta_0$ (which evaluates $(-D)^\alpha \phi$ at $x = 0$) is equal to the monomial ξ^α.

Exercise 2.3.5. Calculate the Fourier transform of δ_y.

2.3.3. Sobolev spaces via the Fourier Transform. Parseval's Formula states that the Fourier transform is (up to a factor $(2\pi)^n$) an isometry in L^2: for functions f and g in $\mathscr{S}$,

$$\langle f, g\rangle = (2\pi)^{-n}\langle \widehat{f}, \widehat{g}\rangle \tag{2.3.11}$$

where $\langle -, -\rangle$ is the hermitian inner product on functions on $\mathbb{R}^n$.

Exercise 2.3.6. Prove this result.

Strictly speaking, to prove that this is an isometry in L^2 requires an approximation argument of L^2 functions by functions in $\mathscr{S}$. We shall ignore this.

Note that (2.3.11) can be applied to $D^\alpha f, D^\alpha g$ and gives

$$\langle D^\alpha f, D^\alpha g\rangle = (2\pi)^{-n} \int \xi^{2\alpha} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\, \mathrm{d}\xi. \tag{2.3.12}$$

Thus for the Sobolev H^s-norm, we have

$$\|f\|_k = (2\pi)^{-n} \int \sum_{|\alpha|\leqslant k} \xi^{2\alpha} |\widehat{f}(\xi)|^2\, \mathrm{d}\xi. \tag{2.3.13}$$

It is easy to see that the norm on the RHS is equivalent to either of the simpler norms

$$\int (1 + |\xi|^{2k})|\widehat{f}(\xi)|^2\, \mathrm{d}\xi \text{ or } \int (1 + |\xi|^2)^k |\widehat{f}(\xi)|^2\, \mathrm{d}\xi \text{ or} \tag{2.3.14}$$

where of course

$$|\xi|^2 = \sum \xi_j^2. \tag{2.3.15}$$

Thus *the Fourier transform takes a function in H^s into a function in L^2 with respect to the measures in* (2.3.14)*.*

NOTATION 2.3.7. The notation

$$\langle x \rangle = \sqrt{1+|x|^2}, \ \ \langle \xi \rangle = \sqrt{1+|\xi|^2},$$

will sometimes be used as an abbreviation in what follows. This is known as the 'Japanese bracket'.

To summarize,

$$f \in H^s(\mathbb{R}^n) \Leftrightarrow \langle \xi \rangle^s \widehat{f}(\xi) \in L^2(\mathbb{R}^n). \tag{2.3.16}$$

2.3.4. Convolution. Multiplication of Fourier transforms corresponds to convolution of the original functions. If f and g are in $\mathscr{S}$

$$f * g(x) = \int f(y)g(x-y)\,\mathrm{d}y. \tag{2.3.17}$$

and

$$\widehat{f*g} = \widehat{f}(\xi)\widehat{g}(\xi). \tag{2.3.18}$$

This operation extends to distributions.

2.4. Parametrices for constant-coefficient operators via the Fourier Transform

Let $p(\xi)$ be an elliptic polynomial (with constant coefficients, as the notation suggests), of degree k. Recall that elliptic means that $p_k(\xi)$, the sum of terms homogeneous of degree k, is non-zero for all real $\xi \neq 0$—or in the case of 'systems' that $p_k(\xi)$ is an invertible matrix for all $\xi \neq 0$.

Let $P = p(D)$ be the corresponding constant-coefficient differential operator. The equation

$$p(D)u = f \tag{2.4.1}$$

if u and f are at least in $\mathscr{S}'(\mathbb{R}^n)$ is then equivalent in frequency-land to the equation

$$p(\xi)\widehat{u}(\xi) = \widehat{f}(\xi). \tag{2.4.2}$$

Now if $p(\xi) \neq 0$ for *all* ξ we could invert this to obtain

$$\widehat{u}(\xi) = p(\xi)^{-1}\widehat{f}(\xi). \tag{2.4.3}$$

We could then transform back to get a 'formula' for the solution, but already here, note what we can read from this equation.

- If $f \in C^\infty \cap \mathscr{S}'$ then $\widehat{f}$ is rapidly decreasing, and so is the RHS of (2.4.3). In particular, $\widehat{u}$ is also rapidly decreasing, and the solution u is also smooth. This is a result about *regularity of solutions* because *a priori*, u was only a tempered distribution.
- If $f \in \mathscr{S}$, then so is u.

REMARK 2.4.1. This may appear obvious, but for the wave equation, for example, (which is not elliptic) (weak) solutions of the homogenous equation need not be smooth.

Theorem 2.4.2. *Suppose that* $p(\xi)$ *is elliptic of order* k *and that* $u \in L^2(\mathbb{R}^n)$ *satisfies the equation*

$$p(D)u = f, \tag{2.4.4}$$

where $f \in H^s$. *Then automatically* $u \in H^{s+k}$. *In particular if* $f \in C^\infty$ *with all derivatives in* L^2 *(this* L^2 *condition just imposes some mild decay at* ∞*) then* $u \in L^2$ *implies* $u \in C^\infty$.

Remark 2.4.3. The proof will yield a good deal more, to be expanded upon in the next Chapter.

Proof. For simplicity suppose that we are dealing with a scalar problem so $p(\xi)$ is an honest polynomial, not matrix-valued. Since $\mathscr{F}$ is an isomorphism, we may take the Fourier transform of (2.4.4) to obtain

$$p(\xi)\widehat{u}(\xi) = \widehat{f}(\xi) \tag{2.4.5}$$

The assumptions $u \in L^2$, $f \in H^s$ translate to

$$\widehat{u} \in L^2, \int (1+|\xi|)^{2s}\widehat{f}(\xi)^2\,\mathrm{d}\xi < \infty. \tag{2.4.6}$$

The ellipticity of p implies that for sufficiently large ξ, $p(\xi) \neq 0$.

In more detail, the fact that $p_k(\xi) \neq 0$ for all ξ in the unit sphere $|\xi| = 1$ implies, by compactness of this sphere, that $|p_k(\xi)| > \delta > 0$ for all $|\xi| = 1$. If $q(\xi) = p(\xi) - p_k(\xi)$ is the sum of the 'lower-order terms' then $|q(\xi)| \leqslant C|\xi|^{k-1}$ for all $|\xi| \geqslant 1$, say, and some constant C depending on the coefficients of q. By homogeneity,

$$|p_k(\xi)| = |\xi|^k |p_k(\xi/|\xi|) > \delta|\xi|^k.$$

Combining this with the bound on q, we see that

$$|\xi| > R := C/\delta \Longrightarrow p(\xi) \neq 0.$$

Now let $\chi \in C^\infty(\mathbb{R})$ be a standard cut-off function,

$$0 \leqslant \chi \leqslant 1, \chi(t) = 1 \text{ for } t \leqslant 1, \chi(t) = 0 \text{ for } t \geqslant 2.$$

Then

$$g(\xi) = (1 - \chi(|\xi|/R))p(\xi)^{-1} \tag{2.4.7}$$

is well-defined, because $1 - \chi$ is non-zero only for $|\xi| > R$ and p is invertible there.

Multiply (2.4.5) by g. We obtain

$$(1 - \chi(|\xi|/R))\widehat{u}(\xi) = g(\xi)f(\xi) \Rightarrow \widehat{u}(\xi) = \chi(|\xi|/R)\widehat{u}(\xi) + g(\xi)\widehat{f}(\xi). \tag{2.4.8}$$

Since $\widehat{u} \in L^2$, the first term on the right lies in $L^2(\mathbb{R}^n, (1+|\xi|^2)^s\mathrm{d}\mu)$ for every s (because χ cuts it off to have compact support). By construction of g, there exists $C > 0$ such that

$$|g(\xi)| \leqslant C(1+|\xi|)^{-k} \tag{2.4.9}$$

for all ξ. Multiplying (2.4.8) by $(1+|\xi|)^{s+k}$ and applying the triangle inequality and (2.4.9),

$$\begin{aligned} |(1+|\xi|)^{s+k}\widehat{u}(\xi)| &\leqslant |(1+|\xi|)^{s+k}\chi(|\xi|/R)\widehat{u}(\xi)| + (1+|\xi|)^{s+k}|g(\xi)|\widehat{f}(\xi)| \\ &\leqslant |(1+|\xi|)^{s+k}\chi(|\xi|/R)\widehat{u}(\xi)| + C(1+|\xi|)^s|\widehat{f}(\xi)|. \end{aligned} \tag{2.4.10}$$

Taking the L^2 norms of each side gives

$$\|(1+|\xi|)^{s+k}\widehat{u}(\xi)\| \leqslant C_0\|u\| + C\|(1+|\xi|)^s\widehat{f}(\xi)\|. \tag{2.4.11}$$

Translating back to the x variables we see that the H^{k+s}- norm of u is bounded in terms of $\|u\|$ and the H^s-norm of f, as required. □

REMARK 2.4.4. Here we used

$$c(1+|\xi|) \leqslant \sqrt{1+|\xi|^2} \leqslant C(1+|\xi|)$$

for positive constants c and C to replace the weight $(1+|\xi|^2)^s$ in the Fourier definition of H^s with the weight $(1+|\xi|)^{2s}$ to simplify the notation.

CHAPTER 3

Pseudodifferential operators: the definitions

3.1. Introduction

Equation (2.4.8) from Chapter 2 can be read as

$$\widehat{u}(\xi) = \chi(\xi/R)\widehat{u}(\xi) + g(\xi)\widehat{p(D)u} \tag{3.1.1}$$

provided that u is Fourier transformable. Let

$$G(x) = \frac{1}{(2\pi)^n}\int e^{ix\xi}g(\xi)\,\mathrm{d}\xi, \;\; E(x) = \frac{1}{(2\pi)^n}\int e^{ix\xi}\chi(|\xi|/R)\,\mathrm{d}\xi. \tag{3.1.2}$$

Because $g(\xi)$ is smooth and decreasing like $|\xi|^{-k}$ for large $|\xi|$, $G(x) \in \mathscr{S}'$ is not necessarily smooth. By contrast, $E(x) \in \mathscr{S}$.

Let us try to understand G better. We can do so by noticing that

$$(1 + \xi D_x)e^{ix\cdot\xi} = (1 + |\xi|^2)e^{ix\cdot\xi} = \langle\xi\rangle^2 e^{ix\cdot\xi} \tag{3.1.3}$$

and

$$D_\xi^\alpha e^{ix\cdot\xi} = x^\alpha e^{ix\cdot\xi} \tag{3.1.4}$$

Recall also that if a function of ξ is bounded by a multiple of $\langle\xi\rangle^{-n-1}$ (Japanese bracket) then the integral of this function over $\mathbb{R}^n$ is uniformly and absolutely convergent. So we have for each α, $k + |\alpha| = n + 1$, that

$$x^\alpha G(x) = \frac{1}{(2\pi)^n}\int (D_\xi^\alpha e^{ix\xi})g(\xi)\phi(x)\,\mathrm{d}\xi = \frac{1}{(2\pi)^n}\int e^{ix\xi}(-D_\xi)^\alpha g(\xi)\phi(x)\,\mathrm{d}\xi \tag{3.1.5}$$

so

$$|x^\alpha G(x)| \leqslant C_\alpha \text{ if } |\alpha| > n - k. \tag{3.1.6}$$

Hence

$$|G(x)| \leqslant \frac{C}{|x|^{n+1-k}}. \tag{3.1.7}$$

EXERCISE 3.1.1. Show similarly that $|D^\beta G| \leqslant C_\beta/|x|^{n+1+|\beta|-k}$ and that if V_j are vector fields which vanish at 0, then

$$|V_1\cdots V_n G| \leqslant C/|x|^{n+1-k}.$$

(where the constant will depend upon the V_j).

In particular G is smooth on $\mathbb{R}^n \setminus 0$ and there it blows up no worse than like a negative power of $|x|$.

We can also describe G as a derivative of a continuous (decaying) function. For this, use the identity (3.1.3). Picking $N \geqslant n+1-k$ so that $N+k \geqslant n+1$, and writing

$$e^{ix\cdot\xi} = \left(\frac{1+\xi D_x}{1+|\xi|^2}\right)^N e^{ix\cdot\xi} \tag{3.1.8}$$

Inserting this we obtain the formula

$$G = \frac{1}{(2\pi)^n}\sum_{r=0}^{N}\binom{N}{r}\int (D_x\cdot\xi)^r\left(\frac{g(\xi)e^{ix\cdot\xi}}{(1+|\xi|^2)^N}\right) \tag{3.1.9}$$

If the rth term in the sum is expanded, the result is a linear combination of terms of the form $D_x^\alpha g_\alpha$ for multi-indices with $|\alpha| = r$. Here

$$g_\alpha(x) = \frac{1}{(2\pi)^n}\int \frac{\xi^\alpha g(\xi)}{(1+|\xi|^2)^N} e^{ix\xi}\,\mathrm{d}\xi \tag{3.1.10}$$

is absolutely convergent because of the choice of N, and so defines a bounded continuous (in fact rapidly decreasing) function of x. This proves that G can be written as a sum of derivatives of order $\leqslant n+1-k$ of bounded continuous functions of x.

3.2. Schwartz Kernels

The Schwartz kernel theorem is a very reassuring general result about continuous linear operators from $\mathscr{S}(\mathbb{R}^n)$ into $\mathscr{S}'(\mathbb{R}^n)$. We shall not prove it, but it motivates the construction and description of inverses and generalized inverses of differential operators that will be given below.

Suppose that $u \in \mathscr{S}'(\mathbb{R}^{2n})$. It is clear that if ϕ and ψ are Schwartz functions then

$$\phi\boxtimes\psi := \mathrm{pr}_1^*(\phi)\,\mathrm{pr}_2^*(\psi) \in \mathscr{S}(\mathbb{R}^{2n}) \tag{3.2.1}$$

and so $u[\phi\boxtimes\psi]$ is well-defined. If u were a function, this would simply be

$$\int u(x,y)\phi(x)\psi(y)\,\mathrm{d}x\mathrm{d}y. \tag{3.2.2}$$

But we can read this a different way: fixing ϕ, we get the linear form

$$\psi \longmapsto u[\phi\boxtimes\psi] \tag{3.2.3}$$

on $\mathscr{S}(\mathbb{R}^n)$. In other words, $u[\phi\boxtimes -] \in \mathscr{S}'$, and depends linearly on ϕ, so defines a map $T_u : \mathscr{S} \to \mathscr{S}'$, and this can be checked to be continuous with respect to the topologies on $\mathscr{S}$ and $\mathscr{S}'$.

The Schwartz kernel theorem states that conversely, every continous operator $A : \mathscr{S} \to \mathscr{S}'$ arises in this way, that is $A = T_u$ for some distribution u. u is called the Schwartz kernel of A and we may non-systematially write K_A for this.

Exercise 3.2.1. Write down the Schwartz kernels of the identity, of $p(D)$, and of $p(x,D)$.

Exercise 3.2.2. If f is say a smoothly bounded function (for each α, $|D^\alpha f(x)| \leqslant C_\alpha$) and M_f is the multiplication operator $\phi \mapsto f\phi$, identify the Schwartz kernels of $M_f \circ A$ and $A\circ M_f$ in terms of the Schwartz kernel of A. Same problem for the operators $D_j\circ A$ and $A\circ D_j$.

Definition 3.2.3. An operator R is called *smoothing* if its Schwartz kernel $K_R \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ and is smoothly bounded.

If P is an operator then a *parametrix* for P is by definition an operator A which inverts P modulo smoothing operators:

$$AP = 1 + R_1, \quad PA = 1 + R_2. \tag{3.2.4}$$

As we shall see, elliptic (differential) operators admit parametrices, and a motivation for the development of the theory of pseudodifferential operators is the systematic construction of such parametrices. Immediate consequences of the existence of a parametrix are 'elliptic regularity' and 'elliptic estimates'. Moreover, when transferred to a compact manifold, we shall obtain the full 'Fredholm package' for elliptic differential operators on a compact manifold.

3.3. Symbol classes and pseudodifferential operators

Having a parametrix for a differential operator in $\mathbb{R}^n$ is clearly a useful thing. From the Schwartz kernel theorem, a possible approach is to make an inspired guess of a class of operators, or equivalently Schwartz kernels, which is simple enough[1] to be able to work with and general enough to contain differential operators, smoothing operators, and parametrices for elliptic differential operators. We shall introduce the class of pseudodifferential operators to fit these requirements.

To continue the motivation, note that the Fourier transform gives a unifying way to represent kernels which are smooth outside the diagonal of $\mathbb{R}^n \times \mathbb{R}^n$. For example

$$\text{Identity operator}: \qquad \delta(x-y) = \frac{1}{(2\pi)^n}\int e^{i\xi\cdot(x-y)}\,\mathrm{d}\xi; \tag{3.3.1}$$

$$p(D): \qquad \frac{1}{(2\pi)^n}\int p(\xi)e^{i\xi\cdot(x-y)}\,\mathrm{d}\xi; \tag{3.3.2}$$

$$p(x,D): \qquad \frac{1}{(2\pi)^n}\int p(x,\xi)e^{i\xi\cdot(x-y)}\,\mathrm{d}\xi. \tag{3.3.3}$$

Remark 3.3.1. We did not emphasise it at the time but given a polynomial $p(x,\xi) = \sum_{|\alpha|\leqslant k} p_\alpha(x)\xi^\alpha$ there is more than one way to define an associated differential operator, precisely because D does not commute with x, (even though ξ does commute with x). The option we took before was to put all the D_j to the right of the $p_\alpha(x)$, so we defined

$$p(x,D) = p_L(x,D) = \sum p_\alpha(x)D^\alpha.$$

We can also define 'right quantization' of $p(x,\xi)$ by putting all the D_j on the left:

$$p_R(x,D) = \sum D^\alpha \circ p_\alpha(x).$$

If the coefficients of p are real (and scalar), then formal adjoint switches p_L and p_R:

$$p_L(x,D)^* = p_R(x,D) \text{ if } p \text{ is real.}$$

[1]Simplicity is, of course, in the eye of the beholder: perhaps tractable would be a better word

Exercise 3.3.2. Find a formula for $p_R(x,D)$ in terms of $p_L(x,D)$. (There is a formula involving $\exp(iD_y\cdot D_\xi)$).)

Exercise 3.3.3. Find the kernel representations of $p_L(x,D)$ and $p_R(x,D)$.

Exercise 3.3.4. Extend the discussion of adjoints to the case of systems, i.e. where $p(x,\xi)$ takes values in $N\times M$ matrices, and $p(x,D)$ maps $C^\infty(\mathbb{R}^n,\mathbb{C}^N)\to C^\infty(\mathbb{R}^n,\mathbb{C}^M)$.

We define symbol classes as smooth functions of (x,ξ) which generalize the idea of being a sum of homogeneous functions of ξ for large ξ. There are in fact several possible definitions. We take one of the simpler ones with uniform behaviour in x.

Definition 3.3.5. A function $a\in C^\infty(\mathbb{R}^p\times\mathbb{R}^n)$, with variables z and ξ in the two factors, is called a symbol of order k if for all multi-indices α,β there is an estimate of the form

$$\sup|D_z^\alpha D_\xi^\beta a(z,\xi)|\leqslant C_{\alpha\beta}\langle\xi\rangle^{k-|\beta|}. \tag{3.3.4}$$

The space of all symbols of order k is denoted $S^k(\mathbb{R}^p;\mathbb{R}^n)$. Here k can be any real number, though in practice it will usually be an integer in our applications.

Example 3.3.6. Any polynomial $p(x,\xi)=\sum_{|\alpha|\leqslant k}p_\alpha(x)\xi^\alpha$ is a symbol of order k provided that the coefficients p_α, and all their derivatives, are bounded.

More generally, if $h(z,\xi)$ is homogeneous of degree k in ξ for $|\xi|\geqslant R$, then h is a symbol of order k (provided the behaviour in z is suitably uniform). A finite sum of such homogeneous symbols is still a symbol. This observation motivates the following refinement of the class S^k:

Definition 3.3.7. If $k\in\mathbb{Z}$ then $a\in S^k$ is called a *polyhomogeneous ('phg') symbol* if there exists a sequence $a_j\in S^{k-j}$, where a_{k-j} is homogeneous of degree $k-j$ in ξ for all $|\xi|\geqslant 1$ such that

$$a-\sum_{j=0}^{N}a_j\in S^{k-N-1} \tag{3.3.5}$$

for every N. The space of all phg symbols is written S^k_{phg}.

Exercise 3.3.8. Find an example of a symbol of order 0, say, which is not phg.

The space S^k has a topology defined similarly to that of $\mathscr{S}$. For each positive integer N we may introduce the norm

$$\|a\|_N=\max_{|\alpha|+|\beta|\leqslant N}\sup\langle\xi\rangle^{|\beta|-k}|D_z^\alpha D_\xi^\beta a| \tag{3.3.6}$$

in S^k. This countable collection of norms gives S^k the structure of a Frechet space for every k.

Note that S^k increases with increasing k. The intersection of all the S^k is non-zero. It is denoted $S^{-\infty}$ and we have

$$a \in S^{-\infty} \iff \text{for every } \alpha, \beta, N, \text{ there exists } C_{N,\alpha,\beta} \text{ so that } |D_z^\alpha D_\xi^\beta a| \leqslant C_{N,\alpha,\beta} \langle \xi \rangle^{-N}. \tag{3.3.7}$$

Thus we may think of such a as being smooth in z and 'Schwartz in ξ'; more precisely every z-derivative of $a(z, \xi)$ is Schwartz in ξ, with all estimates uniform in z.

Some basic results about the symbol spaces can be proved in straightforward fashion by a combination of induction and Leibniz's rule for differentiation of products. For example, product $(a, b) \mapsto ab$ restricts from C^∞ to define a continuous map $S^k \times S^\ell \to S^{k+\ell}$ for all k and ℓ.

Similarly elliptic symbols have 'asymptotic inverses' that are again symbols.

Definition 3.3.9. The symbol $a \in S^k$ is called (uniformly) elliptic if there exist $R > 0$ and $\delta > 0$ such that

$$|a(z, \xi)| \geqslant \delta |\xi|^k \text{ for all } |\xi| \geqslant R \tag{3.3.8}$$

Proposition 3.3.10. *If $a \in S^k$ is elliptic, then there exists $b \in S^{-k}$ with $1 - ab \in S^{-\infty}$. Moreover, if a is phg, then b is also phg.*

Proof. The definition of b is straightforward, and parallels our construction of a parametrix for constant coefficient elliptic differential operators. Pick a standard smooth cut-off function equal to 1 for $|\xi| \leqslant R$ and vanishing for $|\xi| \geqslant 2R$. Define

$$b(z, \xi) = (1 - \chi(\xi)) a(z, \xi)^{-1}; \tag{3.3.9}$$

This precisely makes sense because a^{-1} exists where $1 - \chi \neq 0$. The elliptic estimate for a gives that the first symbol estimate (in S^{-k}) of b. The others follow by repeated differentiation of the identity $ab = 1 - \chi$ and induction. □

Remark 3.3.11. For 'systems' i.e. when our symbols take values in $\mathrm{Hom}(V, V')$, for two finite-dimensional complex vector spaces V, V', this proposition remains true. The ellipticity condition is now that

$$a(x; \xi)^{-1} \text{ exists for all } |\xi| > R \tag{3.3.10}$$

and its operator norm is bounded by $\delta |\xi|^{-k}$ for $|\xi| > R$ (independent of x).

The reason for introducing these spaces is the following:

Definition 3.3.12. A *pseudodifferential operator* A in $\mathbb{R}^n$ of order k is defined through its Schwartz kernel

$$K_A(x, y) = \frac{1}{(2\pi)^n} \int a(x, y; \xi) e^{i\xi \cdot (x - y)} \, \mathrm{d}\xi \tag{3.3.11}$$

where $a \in S^k(\mathbb{R}^{2n}; \mathbb{R}^n)$. The set of all pseudodifferential operators of order k is denoted $\Psi^k(\mathbb{R}^n)$. If $a \in S^k_{\mathrm{phg}}$, then we have a polyhomogeneous pseudodifferential operator of order k. The space of all pseudodifferential operators of order k in $\mathbb{R}^n$ is denoted Ψ^k (or $\Psi^k(\mathbb{R}^n)$), and Ψ^k_{phg} for the subspace of polyhomogeneous pseudodifferential operators.

For the avoidance of doubt, $\Psi^{-\infty}$ is defined by (3.3.11) with $a \in S^{-\infty}$.

In other words, if $u \in \mathscr{S}(\mathbb{R}^n)$,

$$Au(x) = \frac{1}{(2\pi)^n}\int a(x,y;\xi)e^{i\xi\cdot(x-y)}u(y)\,\mathrm{d}\xi\mathrm{d}y \tag{3.3.12}$$

$$= \frac{1}{(2\pi)^n}\int a(x,y;\xi)e^{i\xi\cdot x}\widehat{u}(\xi)\,\mathrm{d}\xi \tag{3.3.13}$$

where $\widehat{u}$ is the Fourier transform of u. Notice that in the second representation $\widehat{u} \in \mathscr{S}$ so the integral is absolutely and uniformly convergent, so Au is certainly continuous. Operating on both sides with $x^\alpha D_x^\beta$, integrating by parts and using the symbol estimates shows that in fact $Au \in \mathscr{S}$. Later we shall prove more precisely that

THEOREM 3.3.13. *If $A \in \Psi^k(\mathbb{R}^n)$, then A is a continuous linear operator $\mathscr{S} \to \mathscr{S}$.*

NOTATION 3.3.14. If $a \in S^k(\mathbb{R}^n;\mathbb{R}^n)$ is a symbol, it is customary to write $\mathrm{Op}(a)$ or $a(x,D)$ for the pseudodifferential operator

$$Au(x) = \frac{1}{(2\pi)^n}\int a(x;\xi)e^{i\xi\cdot(x-y)}u(y)\,\mathrm{d}y.$$

3.4. Some results about symbols

We shall often have occasion to write down expressions of the form

$$I(a) = \int a(x,\xi)e^{i\xi\cdot(x-y)}\,\mathrm{d}\xi \tag{3.4.1}$$

where a is a symbol in S^k. If $k < -n$, then this is absolutely and uniformly convergent and so defines a bounded continuous function of x. (The basic point here is that $\langle\xi\rangle^{-n-\delta}$ is integrable in n dimensions if $\delta > 0$. And the definition of S^k gives a bound on $|a|$ by a multiple of $\langle\xi\rangle^{-k}$.)

However, we want to be able to handle such integrals when k does not satisfy this condition. In this case, $I(a)$ still defines a distribution in $\mathscr{S}'(\mathbb{R}^{2n})$. One approach to deriving results about $I(a)$ might be to replace a by

$$a_j(z,\xi) = \chi(\xi/j)a(z,\xi),\ j = 1,2,\ldots \tag{3.4.2}$$

where χ is a standard cut-off function equal to 1 in a neighbourhood of $|\xi| \leqslant 1$, say, but vanishing for $|\xi| \geqslant 2$.

Then $a_j \to a$ uniformly and with all derivatives on sets of the form $\mathbb{R}^n \times \{|\xi| \leqslant R\}$, for any given R. Moreover a_j is compactly supported with respect to the ξ variable, so $I(a_j)$ is a very nice integral to which can apply all the usual tricks (integration by parts, differentiation under the integral sign) with impunity.

To obtain results about $I(a)$, however, we need convergence with respect to the topology of the symbol spaces. Sadly, a_j does not converge to a in the topology of S^k. However, it does converge in the topology of S^ℓ for every $\ell > k$, and this is sufficient for our applications.

PROPOSITION 3.4.1.

(a) *Let $\chi \in C_0^\infty(\{|\xi| < 1\}$ with $\chi(\xi) = 1$ in a neighbourhood of $\xi = 0$. Then $\chi_\varepsilon(\xi) := \chi(\varepsilon\xi)$ is uniformly bounded in S^0 and $\chi_\varepsilon \to 1$ in S^k for every $k > 0$.*

(b) *If* $a \in S^k$ *and* $b \in S^\ell$, *then* $ab \in S^{k+\ell}$ *and the map* $(a,b) \mapsto ab$ *is continuous with respect to the obvious topologies.*

(c) *If* $a \in S^k$, *then* $D_x^\alpha D_\xi^\beta a \in S^{k-|\beta|}$.

Proof. We give the proof of (i). The proof of (ii) follows by systematic use of the Leibniz rule and applying the known symbol estimates of a and b. The proof of (iii) is obvious.

For (i), we start by noting that since $\chi(\xi)$ is smooth and has support in the unit ball, it is certainly a symbol, and so for each β,

$$|D_\xi^\beta \chi(\xi)| \leqslant C_\beta \langle \xi \rangle^{-|\beta|}. \tag{3.4.3}$$

for some constant C_β.

From the definition of χ_ε, we see by the chain rule

$$D_\xi^\beta \chi_\varepsilon(\xi) = \varepsilon^{|\beta|}(D_\xi^\beta \chi)(\varepsilon\xi) \tag{3.4.4}$$

and so

$$|D^\beta \chi_\varepsilon| \leqslant C_\beta \varepsilon^{|\beta|} \langle \varepsilon\xi \rangle^{-|\beta|}. \tag{3.4.5}$$

If $|\beta| > 0$, then for $k \geqslant 0$, then we have

$$\langle \xi \rangle^{|\beta|-k} |D^\beta \chi_\varepsilon| \leqslant C_\beta \varepsilon^{|\beta|} \frac{\langle \xi \rangle^{|\beta|-k}}{\langle \varepsilon\xi \rangle^{|\beta|}} \tag{3.4.6}$$

Now

$$\begin{aligned} \frac{\langle \xi \rangle^{|\beta|-k}}{\langle \varepsilon\xi \rangle^{|\beta|}} &= \frac{(1+|\xi|^2)^{(|\beta|-k)/2}}{(1+\varepsilon^2|\xi|^2)^{|\beta|/2}} \\ &= \frac{(1+|\xi|^2)^{(|\beta|-k)/2}}{\varepsilon^{|\beta|}(\varepsilon^{-2}+|\xi|^2)^{|\beta|/2}} \\ &\leqslant \frac{1}{\varepsilon^{|\beta|}(\varepsilon^{-2}+|\xi|^2)^{k/2}} \leqslant \varepsilon^{k-|\beta|}, \end{aligned} \tag{3.4.7}$$

using the fact that

$$\frac{1+|\xi|^2}{\varepsilon^{-2}+|\xi|^2} \leqslant 1 \text{ if } \varepsilon \leqslant 1.$$

Substituting this into (3.4.6), we obtain the uniform estimate

$$\langle \xi \rangle^{|\beta|-k} |D^\beta \chi_\varepsilon| \leqslant C_\beta \varepsilon^k \tag{3.4.8}$$

which proves the result unless $|\beta| = 0$.

In this case, Taylor's formula gives

$$|1-\chi(\varepsilon\xi)| \leqslant \begin{cases} C\varepsilon|\xi| \text{ if } \varepsilon|\xi| \leqslant 1; \\ 1 \text{ if } \varepsilon|\xi| \geqslant 1. \end{cases} \tag{3.4.9}$$

This can clearly be replaced by the statement

$$|1-\chi(\varepsilon\xi)| \leqslant \begin{cases} C\varepsilon^\delta|\xi|^\delta \text{ if } \varepsilon|\xi| \leqslant 1; \\ 1 \text{ if } \varepsilon|\xi| \geqslant 1. \end{cases} \tag{3.4.10}$$

where δ is any number in $(0,1)$. For δ in this range, multiply by $\langle \xi \rangle^{-\delta}$. We obtain

$$\langle \xi \rangle^{-\delta}|1-\chi(\varepsilon\xi)| \leqslant \begin{cases} C\varepsilon^\delta \text{ if } \varepsilon|\xi| \leqslant 1; \\ \langle \xi \rangle^{-\delta} \text{ if } \varepsilon|\xi| \geqslant 1. \end{cases} \tag{3.4.11}$$

This shows that

$$\sup \langle \xi \rangle^{-\delta}|1-\chi(\varepsilon\xi)| = O(\varepsilon^\delta)$$

for $\delta \in [0,1]$ and completes the proof. □

Corollary 3.4.2. *If $a \in S^k$, then $a_\varepsilon = \chi_\varepsilon(\xi)a(x;\xi) \in S^{-\infty}$ is bounded in S^k and $a_\varepsilon \to a$ in S^ℓ for every $\ell > k$.*

Proof. Let $a_\varepsilon(x;\xi) = \chi_\varepsilon(\xi)a(x;\xi)$ where χ_ε is as in the proposition. We have seen that $\chi_\varepsilon \to \chi$ in S^δ for every positive δ. Since multiplication is continuous (part (ii) of the Proposition) it follows that $\chi_\varepsilon a \to a$ in $S^{k+\delta}$ for every δ. □

Theorem 3.4.3. *For $j = 0, 1, 2, \ldots$ suppose that $A_j \in \Psi^{k-j}(\mathbb{R}^n)$. Then there exists a pseudodifferential operator $A \in \Psi^k(\mathbb{R}^n)$ such that, for each N,*

$$A = \sum_{j=0}^{N} A_j \bmod \Psi^{k-N-1}(\mathbb{R}^n). \tag{3.4.12}$$

Proof. Since A_j is determined by its amplitude $a_j \in S^{k-j}$, it is sufficient to prove the analogue of (3.4.12) for a given sequence of symbols, that is: given $a_j \in S^{k-j}$, there exists $a \in S^k$ with the property that

$$a - \sum_{j=0}^{N} a_j \in S^{k-N-1} \tag{3.4.13}$$

for every N. Then $A = \mathrm{Op}(a, D)$ will have the required property.

Let χ be a standard cut-off function equal to 1 for $t \leqslant 1$ equal to 0 for $t \geqslant 2$, $0 \leqslant \chi(t) \leqslant 1$ for all t. Set

$$\widetilde{a}_j(x;\xi) = (1 - \chi(\varepsilon_j|\xi|))a_j(x;\xi). \tag{3.4.14}$$

The idea is to choose $\varepsilon_j \to 0$ so fast that

$$a(x;\xi) := \sum_{j=0}^{\infty} \widetilde{a}(x;\xi) \tag{3.4.15}$$

does the job. First note that the assumption that $\varepsilon_j \to 0$ implies that the sum (3.4.15) is locally finite: when ξ is restricted to a ball $\{|\xi| \leqslant R\}$, all but finitely many terms vanish. So $a \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$. However, this is not yet enough to prove (3.4.13).

Let us temporarily denote by $\|f\|_{\mu,\nu}$ the μ-th norm used to define the topology of S^ν, that is

$$\|f\|_{\mu,\nu} = \sup \langle\xi\rangle^{-\nu} \sum_{|\alpha|+|\beta|\leqslant\mu} \langle\xi\rangle^{|\beta|} |D_x^\alpha D_\xi^\beta f| \tag{3.4.16}$$

From this definition it follows at once that

$$\|f\|_{\mu,\nu} \leqslant \|f\|_{\mu',\nu'} \text{ whenever } \mu \leqslant \mu' \text{ and } \nu \geqslant \nu'. \tag{3.4.17}$$

We know that $(1 - \chi(\varepsilon|\xi|)a_j(x;\xi)$ tends to zero in S^{k-j+1} as $\varepsilon \to 0$ for fixed j. In particular, for each j, we can choose ε_j so small that

$$\|\widetilde{a}_j\|_{j,k-j+1} \leqslant 2^{-j}. \tag{3.4.18}$$

Now fix N and consider the difference

$$a - \sum_{j=0}^{N} \widetilde{a}_j = \sum_{j \geqslant N+1} \widetilde{a}_j = r_N, \text{ say.} \tag{3.4.19}$$

We need to show that $r_N \in S^{k-N-1}$, in other words

$$\|r_N\|_{\mu,k-N-1} < \infty \text{ for all } \mu. \tag{3.4.20}$$

For fixed μ, (3.4.17) says that

$$\|\widetilde{a}_j\}_{\mu,k-N-1} \leqslant \|wta_j\}_{j,k-j+1} \leqslant 2^{-j} \text{ if } \mu \leqslant j,\ k-j+1 \leqslant k-N-1 \text{ i.e. } j \geqslant N+2. \tag{3.4.21}$$

Thus if $\mu \leqslant N+2$, we write

$$R_N = \widetilde{a}_{N+1} + \sum_{j \geqslant N+2} \widetilde{a}_j. \tag{3.4.22}$$

The first term is in S^{k-N-1} anyhow, and

$$\| \sum_{j\geqslant N+2} \widetilde{a}_j \|_\mu \leqslant \sum_{j\geqslant N+2} \|\widetilde{a}_j\|_\mu \leqslant \sum_{j\geqslant N+2} 2^{-j} < \infty. \tag{3.4.23}$$

If $\mu > N+2$, we split the sum differently,

$$R_N = \sum_{j=N+1}^{\mu-1} \widetilde{a}_j + \sum_{j\geqslant \mu} \widetilde{a}_j. \tag{3.4.24}$$

Again, the finite sum lies in S^{k-N-1} and the $(\mu, k-N-1)$-norm of the infinite sum is estimated using (3.4.21). This completes the proof. □

Remark 3.4.4. Given $a \in S^k$ and a sequence $a_j \in S^{k-j}$, if (3.4.13) holds for every N, we write

$$a \sim \sum_{j=0}^{\infty} a_j. \tag{3.4.25}$$

This should be understood as an asymptotic expansion of a for large ξ. For the corresponding operators A and A_j in (3.4.12), we also write

$$A \sim \sum_{j=0}^{\infty} A_j. \tag{3.4.26}$$

Remark 3.4.5. It is pretty clear from the proof that if a' is another symbol in S^k satisfying (3.4.25), then $a - a' \in S^{-\infty}$, because this is by definition the intersection of all the S^k. Similarly, if A' is another pseudodifferential operator satisfying then $A - A' \in \bigcap_k \Psi^k$. We shall see later (Theorem 3.8.1) that this latter intersection is precisely $\Psi^{-\infty}$ (i.e. operators of the form $\mathrm{Op}(r, D)$, where $r \in S^{-\infty}$.

3.5. The symbol and principal symbol of a pseudodifferential operator

The class of operators Ψ^k is supposed to be a natural generalization of the Diff^k (at least if k is a positive integer!). In this section we define the principal symbol of $A \in \Psi^k$. For operators associated to general amplitudes $a(x, y; \xi)$ as in Definition 3.3.12, we also define the 'full symbol' $\sigma_A(x;\xi)$ which gives a canonical representation of A.

The input in the Definition 3.3.12 is an 'amplitude' $a(x, y; \xi)$ depending on $3n$ variables, whereas the output $K_A(x, y)$ depends only upon $2n$ variables, see (3.3.11).

One should expect therefore, that the map {amplitudes} → {kernels} will have a big null-space. This is indeed the case. What is true is that there is a bijective map

$$\Psi^k(\mathbb{R}^n) \longrightarrow S^k(\mathbb{R}^n;\mathbb{R}^n), \quad A \longmapsto \sigma_A \tag{3.5.1}$$

defined as follows. Let $e_\eta(x) = e^{i\eta\cdot x}$, and set

$$\sigma_A(x;\eta) = e_{-\eta}(x)A[e_\eta](x). \tag{3.5.2}$$

Then we have the formula

$$\sigma_A(x;\eta) = \frac{1}{(2\pi)^n}\int a(x,y;\xi)e^{i\xi\cdot(x-y)-i\eta\cdot(x-y)}\,\mathrm{d}\xi\mathrm{d}y \tag{3.5.3}$$

and after replacing y by $y-x$ and ξ by $\xi-\eta$ in this integral, we have

$$\sigma_A(x;\eta) = \frac{1}{(2\pi)^n}\int a(x,x+y;\xi+\eta)e^{i\xi\cdot y}\,\mathrm{d}\xi\,\mathrm{d}y \tag{3.5.4}$$

If the amplitude a in (3.3.11) happens to be independent of y, then this reduces to

$$\sigma_A(x;\eta) = \frac{1}{(2\pi)^n}\int a(x;\xi+\eta)e^{i\xi\cdot y}\,\mathrm{d}\eta\mathrm{d}y = \int a(x;\xi+\eta)\delta(\xi)\,\mathrm{d}\xi = a(x;\eta) \tag{3.5.5}$$

so that $\sigma_A(x;\xi) = a(x;\xi)$ in this case.

On the other hand, brushing some technicalities under the carpet, we have

$$Au = \sigma_A(x,D)u, \tag{3.5.6}$$

as follows by writing $u \in \mathscr{S}$ in terms of its Fourier transform. Indeed, applying A to

$$u(y) = \frac{1}{(2\pi)^n}\int e^{iy\cdot\xi}\widehat{u}(\xi)\,\mathrm{d}\xi \tag{3.5.7}$$

we obtain

$$\begin{aligned} Au(x) &= \frac{1}{(2\pi)^n}\int [Ae_\xi](x)\widehat{u}(\xi)\,\mathrm{d}\xi \\ &= \frac{1}{(2\pi)^n}\int \sigma_A(x;\xi)e^{i\xi\cdot x}\widehat{u}(\xi)\,\mathrm{d}\xi \\ &= \frac{1}{(2\pi)^n}\int \sigma_A(x;\xi)e^{i\xi\cdot(x-y)}u(y)\,\mathrm{d}\xi\mathrm{d}y \end{aligned} \tag{3.5.8}$$

as claimed.

This argument shows how to write the pseudodifferential operator A in (3.3.11) in 'reduced form' using $\sigma_A(x;\xi)$. What we have not verified is that $\sigma_A \in S^k$. To summarize:

THEOREM 3.5.1. *Let A be a pseudodifferential operator of order k. Then there exists a unique (reduced, full) symbol $\sigma_A \in S^k(\mathbb{R}^n;\mathbb{R}^n)$ such that*

$$K_A(x,y) = \frac{1}{(2\pi)^n}\int \sigma_A(x;\xi)e^{i\xi\cdot(x-y)}\,d\xi. \tag{3.5.9}$$

Moreover, we have the asymptotic expansion

$$\sigma_A(x;\xi) \sim \sum_\alpha \frac{i^{|\alpha|}}{\alpha!}\,D_y^\alpha D_\xi^\alpha a(x,y;\xi)\big|_{y=x}. \tag{3.5.10}$$

If σ_A is the 'full symbol' of A, then it is also useful to have a definition of the leading order term, generalizing the *principal symbol* of a differential operator.

DEFINITION 3.5.2. Define $\mathrm{Sym}^k = S^k/S^{k-1}$. For polyhomogeneous symbols, define $\mathrm{Sym}^k_{\mathrm{phg}}$ to be the space of functions $a \in C^\infty(\mathbb{R}^n \times (\mathbb{R}^n \setminus 0))$ (positively) homogeneous of degree k in the second variable,

$$a(x;t\xi) = t^k a(x;\xi) \text{ for all } t > 0.$$

REMARK 3.5.3. The difference between the definitions of Sym^k and $\mathrm{Sym}^k_{\mathrm{phg}}$ is not as great as it looks. If $a \in S^k_{\mathrm{phg}}$, then $a = a_k \bmod S^{k-1}_{\mathrm{phg}}$ for large ξ where $a_k(x;\xi)$ is homogeneous of degree k in ξ where defined. But such a function extends canonically as a homogeneous function for all $\xi \neq 0$. We are thus identifying $\mathrm{Sym}^k_{\mathrm{phg}}$ with the space of these canonically extended homogeneous functions.

DEFINITION 3.5.4. If A is a pseudodifferential operator of order k defined by the amplitude $a(x, y; \xi)$, then

$$\sigma_k(A) = [a(x, x; \xi)] \tag{3.5.11}$$

where the square brackets denote the equivalence class in Sym^k or in $\mathrm{Sym}^k_{\mathrm{phg}}$ if A is polyhomogeneous.

Note that if A is in 'reduced form', i.e. defined by the amplitude $\sigma_A(x;\xi)$, then $\sigma_k(A) = [\sigma_A]$.

3.6. Fundamental properties of pseudodifferential operators

In this section we gather, without proofs, the main properties of pseudodifferential operators in $\mathbb{R}^n$ that make them so useful in the study of differential operators. The first result concerns the formal 'algebraic' properties, including the (principal-) symbol exact sequence (3.6.1). Next we have a 'residuality result' which shows in particular that $\Psi^{-\infty} = \bigcap_k \Psi^k$. Then we summarize the mapping properties of pseudodifferential operators acting first on $\mathscr{S}$ and then on Sobolev spaces.

The following three-part Theorem shows that Ψ^* enjoys properties very similar to Diff^*:

THEOREM 3.6.1.

(i) *Symbol sequence: for every $k \in \mathbb{R}$, there is an exact sequence*

$$0 \longrightarrow \Psi^{k-1} \longrightarrow \Psi^k \xrightarrow{\sigma_k} \mathrm{Sym}_k \to 0. \tag{3.6.1}$$

(ii) *Composition: if $A \in \Psi^k$ and $B \in \Psi^\ell$ then $AB \in \Psi^{k+\ell}$ (with obvious interpretations if either k or ℓ is equal to $-\infty$). The principal symbol is multiplicative in the sense that*

$$\sigma_{k+\ell}(AB) = \sigma_k(A)\sigma_\ell(B). \tag{3.6.2}$$

(iii) *The principal symbol map is a $*$-homomorphism in the sense that $\sigma_k(A^*) = \sigma_k(A)^*$, where A^* is the L^2-adjoint of A.*

There is a parallel statement for the subclass of polyhomogeneous operators, in which everything is adorned with the subscript phg.

Since differential operators are polynomials, the entire information of a differential operator is captured by a finite number of symbols. This is no longer true of pseudodifferential operators. The next result identifies the 'residual space' of operators in $\bigcap_k \Psi^k$.

THEOREM 3.6.2. *The space $\bigcap_k \Psi^k$ is equal to $\Psi^{-\infty}$, the latter being defined through symbols of order $-\infty$. It consists of operators with smooth Schwartz kernels rapidly decreasing away from the diagonal of $\mathbb{R}^n \times \mathbb{R}^n$. If $A \in \Psi^{-\infty}$ there is a unique $a \in S^{-\infty}$ such that $A = \sigma_A(x, D)$.*

Now we move on to mapping properties.

PROPOSITION 3.6.3. *If $A \in \Psi^k$, then A defines a continuous linear map $\mathscr{S} \to \mathscr{S}$.*

We prove similarly that

PROPOSITION 3.6.4. *If $A \in \Psi^k$, then* sing-supp$(k_A) = \Delta$, *the diagonal of $\mathbb{R}^n \times \mathbb{R}^n$. In other words, the restriction of the distribution to $\mathbb{R}^n \times \mathbb{R}^n \setminus \Delta$ is smooth.*

It follows directly from this result that the singular support of Au cannot exceed the singular support of u if $A \in \Psi^k$. This is often called the 'pseudolocality property' of pseudodifferential operators.

PROPOSITION 3.6.5. *If $A \in \Psi^k(\mathbb{R}^n)$ and $u \in \mathscr{S}'(\mathbb{R}^n)$, then*

$$\text{sing-supp}(Au) \subset \text{sing-supp}(u) \tag{3.6.3}$$

PROOF. Let x_0 be a point outside of the singular support of f and let χ' be a cut-off function which is identically equal to 1 in a neighbourhood of x_0. Further, choose another cut-off function χ with the property that χ is identically equal to 1 on $\text{supp}(\chi')$ and with support so small that $\text{supp}(\chi) \cap \text{sing-supp}(f) = \emptyset$. (We'll see very soon why we need these *two* cut-off functions. We want to prove that if $A \in \Psi^k$, then Af is smooth near x_0, or equivalently $\chi' Af$ is smooth. If we write

$$f = \chi f + (1 - \chi)f, \tag{3.6.4}$$

then

$$\chi' Af = (\chi' A\chi)f + (\chi' A(1 - \chi))f \tag{3.6.5}$$

where the products are to be understood as operator composition. By definition, χf is smooth and we have seen that A maps smooth to smooth, so the first term is smooth. As for the second term, the kernel of the operator is

$$\chi'(x)(1 - \chi(y))K_A(x, y). \tag{3.6.6}$$

Now this is supported away from the diagonal for if we put $x = y$ in the χ factors, we obtain $\chi'(x)(1 - \chi(x) = \chi'(x) - \chi(x)\chi'(x) = 0$ since χ is identically 1 on $\text{supp}(\chi')$. Remark that if we don't have χ and χ' as defined, we cannot conclude that $\chi(x)(1 - \chi(y))$ is supported away from the diagonal. □

THEOREM 3.6.6 (Boundedness in Sobolev spaces). *If $A \in \Psi^k$, then A defines a bounded linear map $H^s \to H^{s-k}$ for every $s \in \mathbb{R}$.*

PROOF. To be supplied. Cf. Melrose's notes, §§2.13–2.16. □

Discussion of compactness also to be supplied.

3.7. Parametrix construction

Let $P = p(x, D)$ be an elliptic differential operator in $\mathbb{R}^n$. A parametrix is an operator A which is an inverse mod $\Psi^{-\infty}$, that is

$$AP - 1 \in \Psi^{-\infty}, \quad PA - 1 \in \Psi^{-\infty}. \tag{3.7.1}$$

The goal of this section is to show that the existence of such A is essentially reduced to algebra, given the results summarized in §3.6.

We shall present two slightly different arguments, though there are common elements to both.

We shall give more details in the next chapter, but let's consider the case of composition of PA, where $A \in \Psi^\ell$ and $P \in \mathrm{Diff}^k$. So suppose that

$$K_A(x, y) = \frac{1}{(2\pi)^n} \int a(x, y; \xi) e^{i\xi\cdot(x-y)}\, d\xi \tag{3.7.2}$$

Then

$$D_{x_j} K_A = \frac{1}{(2\pi)^n} \int (\xi_j + D_{x_j}) a(x, y; \xi) e^{i\xi\cdot(x-y)}\, d\xi \tag{3.7.3}$$

It follows that if $P = p(x, D)$, then

$$K_{PA} = \frac{1}{(2\pi)^n} \int p(x, \xi + D_x) a(x, y; \xi) e^{i\xi\cdot(x-y)}\, d\xi \tag{3.7.4}$$

Since p is smooth in the first variable and polynomial in the second variable it follows that

$$b(x, y; \xi) := p(x, \xi + D_x) a(x, y; \xi) \tag{3.7.5}$$

is as symbol of order $k+\ell$, so $B = PA$ is a pseudodifferential operator of order $k + \ell$ and

$$\sigma_{k+\ell}(PA) = \sigma_k(P)\sigma_\ell(A). \tag{3.7.6}$$

Proposition 3.7.1. *Let the notation be as above, with $P \in \mathrm{Diff}^k$, $A \in \Psi^\ell$ and $B = PA$. Then the amplitude b in (3.7.5) which gives B as a pseudodifferential operator has the formula*

$$b(x, y; \xi) = \sum_\alpha \frac{1}{\alpha!} \partial_\xi^\alpha p(x; \xi) D_x^\alpha a(x, y; \xi). \tag{3.7.7}$$

Note that the sum here is finite, extending only over α with $|\alpha| \leqslant k$.

Proof. If h is any vector, then by Taylor's theorem,

$$p(x; \xi + h) = \sum_\alpha \frac{1}{\alpha!} \partial_\xi^\alpha (x; \xi) h^\alpha. \tag{3.7.8}$$

This is a finite sum, so a little thought shows that we can substitute $h = D$ here, formally, and obtain the formula

$$p(x; \xi + D_x) = \sum_\alpha \frac{1}{\alpha!} \partial_\xi^\alpha (x; \xi) D_x^\alpha \tag{3.7.9}$$

The result now follows immediately from (3.7.5). □

Exercise 3.7.2. Obtain a formula for the composite AP where A is a pseudo-differential operator and P is a differential operator (without using the results of §3.6.

It turns out that this formula extends, in the sense of asymptotic expansion of symbols, for composition of pseudodifferential operators. See Theorem **??** below.

THEOREM 3.7.3. *Let $P = p(x, D)$ be an elliptic differential operator of order k on $\mathbb{R}^n$. (Recall that this means that the part p_k of p homogeneous of degree k in ξ satisfies*

$$|p_k(x;\xi)| \geqslant \delta|\xi|^k \text{ for } |\xi| \neq 0 \tag{3.7.10}$$

where $\delta > 0$.) Then P admits a parametrix *in Ψ^{-k}, that is to say an operator inverting P mod $\Psi^{-\infty}$:*

$$AP = 1 + R, \quad PA = 1 + R', \quad R, R' \in \Psi^{-\infty} \tag{3.7.11}$$

3.7.1. Proof via Neumann series. Let us show first that there exists $A_0 \in \Psi^{-k}$ such that $PA_0 - 1 \in \Psi^{-1}$. First of all, we know that $\sigma_k(P)$ is invertible. This means that there is a symbol $a_0 \in S^{-k}$ such that $\sigma_k(P)a_0 = 1$. By the exactness, there exists $A_0 \in \Psi^{-k}_{\text{phg}}$ with symbol equal to a_0, and $\sigma_0(PA_0) = \sigma_0(1)$. So the difference $R_1 := 1 - PA_0 \in \Psi^{-1}$ as required. We would like to define $A = A_0(1 - R_1)^{-1}$, for then $PA = 1$. We don't know that $1 - R_1$ is invertible as an operator so this argument does not quite work. However, we can work *formally*, though this requires some more theory which we'll come to in the next Chapter.

The usual proof for ordinary geometric progressions shows that

$$(1 - R_1)(1 + R_1 + \cdots + R_1^N) = 1 - R_1^{N+1} \tag{3.7.12}$$

and the RHS is 1 mod Ψ^{-N-1}. So if we replace A_0 by

$$A_N = A_0(1 + R_1 + \cdots + R_1^N) \tag{3.7.13}$$

(which is well-defined by the composition theorem),

$$PA_N = 1 - R_1^{N+1} \tag{3.7.14}$$

so we have inverted P mod Ψ^{-N-1}.

We shall see below (see the section on completeness and residuality) that in fact one can find $S \in \Psi^0$ such that $S \sim \sum_{j=0}^{\infty} R_1^j$. Precisely, this implies that

$$R := (1 - R_1)S \in \Psi^{-\infty} \tag{3.7.15}$$

Hence if we define $A \in A_1 \circ S$ we satisfy the first condition of (3.7.11).

Similarly, we find a 'left parametrix' B and a smoothing operator T such that $B \circ P = 1 - T$. Now the usual argument in the ring $\Psi^*/\Psi^{-\infty}$ that multiplicative inverses of invertible elements are unique shows that $B = A$ mod Ψ^{-infty}. In particular $B = A + U$, where $U \in \Psi^{-\infty}$ and it follows from $B \circ P = 1 - T$ that $A \circ P = 1 - T + U(1 - T)$ so A is also a left parametrix (with $R' = T + U(1 - T)$.

COROLLARY 3.7.4 (Elliptic Regularity). *Suppose that P is elliptic of order k in $\mathbb{R}^n$. Suppose that u and f satisfy*

$$Pu = f \tag{3.7.16}$$

where $u \in \mathscr{S}'$ and $f \in \mathscr{S}' \cap C^\infty$. (This means that f is smooth and doesn't grow to fast at ∞). Then u is in fact smooth.

Proof. If $Pu = f$, and A is a parametrix, then left-multiplying by A gives $u = Af - Ru$. Since A is a pseudodifferential operator, it maps C^∞ to C^∞. Since R is smoothing it maps $\mathscr{S}'$ to $\mathscr{S}$. Hence u is indeed smooth. □

Remark 3.7.5. Little needs to be changed if P is a 'system', i.e. if the symbols are matrix-valued. The ellipticity condition is that $p_k(x;\xi)$ is invertible for all $x \in \mathbb{R}^n$, $\xi \neq 0$, and the norm of $p_k^{-1}(x;\xi)$ is bounded by a multiple of $|\xi|^{-k}$ for all $\xi \neq 0$.

3.7.2. Alternative approach. We used the full force of the formal properties of pseudodifferential operators to prove Theorem 3.7.3. We can also make the construction, only assuming the composition formula of a differential operator with a pseudodifferential operator as follows.

We start as before with the construction of A_0,

$$PA_0 = 1 + R_1, \quad R_1 \in \Psi^{-1}.$$

From here, we can find $A_1 \in \Psi^{-k-1}$ such that $PA_1 = R_1 + R_2$, where $R_2 \in \Psi^{-2}$. All we ever use is the ellipticity of P and the composition rules for PA where P is a differential operator and $A \in \Psi^*$. In this way, we obtain for each $N > 0$, $A_j \in \Psi^{-k-j}$ and $R_j \in \Psi^{-j}$ such that

$$P(A_0 + \cdots + A_N) = 1 + R_{N+1}. \tag{3.7.17}$$

To go all the way to an error term in $\Psi^{-\infty}$ we need to be able to take the sum to infinity of the A_j just as we needed to sum to ∞ the geometric projection using the other method.

3.8. Proofs and sketch-proofs

We shall prove the results of §3.6 in the following order. First we shall prove the 'residuality theorem', Theorem 3.6.2. Next we shall prove the 'reduction theorem', Theorem 3.5.1. From these two results, the rest of Theorem 3.6.1 follows fairly easily.

Then we discuss the mapping properties—the proof of boundedness and compactness in Soboleve spaces will be supplied later.

3.8.1. Proof of the residuality theorem, Theorem 3.6.2.

Theorem 3.8.1. *A continuous linear operator $A : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ lies in $\bigcap_N \Psi^N$ if and only if the kernel K_A is smooth and rapidly decaying away from the diagonal in the sense that for all N and multi-indices α and β, there exists a constant $C_{N,\alpha,\beta}$ such that*

$$|D_x^\alpha D_y^\beta K_A(x,y)| \leqslant C_{N,\alpha,\beta}\langle x-y\rangle^{-N}. \tag{3.8.1}$$

More over, any such operator is the quantization of a symbol of order $S^{-\infty}$.

Proof. See, for example, Melrose's notes, Proposition 2.4.

Note that the last part follows from the Fourier Transform. For we see that, given $K_A(x,y)$, we need to find an amplitude a such that

$$K_A(x,y) = \frac{1}{(2\pi)^n}\int a(x,\xi)e^{i\xi\cdot(x-y)}\,\mathrm{d}\xi. \tag{3.8.2}$$

But if we define $g(x,z) = K_A(x,x-z)$, (3.8.2) becomes

$$g(x,z) = K_A(x,x-z) = \frac{1}{(2\pi)^n}\int a(x,\xi)e^{i\xi\cdot(z)}\,\mathrm{d}\xi \tag{3.8.3}$$

so that for fixed x, $a(x,\xi)$ is the Fourier transform of $z \mapsto g(x,z)$. Unravelling the definitions, we see that the decay conditions on K_A translate directly into the conditions ensuring $a \in S^{-\infty}$. □

3.8.2. Proof of the reduction theorem, Theorem 3.5.1. The following simple calculation shows that if $a(x,y;\xi)$ is a symbol of order k *which vanishes on the diagonal* $x = y$ then the operator with amplitude $a(x,y;\xi)$ is actually of order $\leqslant k-1$.

To see this, for any given j, consider the kernel with amplitude $(x_j - y_j)a(x,y;\xi)$. The corresponding operator has kernel

$$\begin{aligned} K_A(x,y) &= \frac{1}{(2\pi)^n}\int (x_j-y_j)a(x,y;\xi)e^{i\xi\cdot(x-y)}\,\mathrm{d}\xi \\ &= \frac{1}{(2\pi)^n}\int a(x,y;\xi)D_{\xi_j}e^{i\xi\cdot(x-y)}\,\mathrm{d}\xi \\ &= \frac{1}{(2\pi)^n}\int D_{\xi_j}a(x,y;\xi)e^{i\xi\cdot(x-y)}\,\mathrm{d}\xi. \end{aligned} \tag{3.8.4}$$

By the symbol estimates, we see that the operator is of order $\leqslant k-1$.

There is a caveat. The amplitude $(x_j - y_j)a(x,y;\xi)$ is only in S^k if $a(x,y;\xi)$ has some decay for $|x-y| \to \infty$ (i.e. away from the diagonal). Our definition of S^k requires uniform boundedness (and of all derivatives) in (x,y). This turns out not to be a serious problem here, for if

$$b(x,y;\xi) \in \langle x-y\rangle^{\mu} S^k, \text{ i.e. } \langle x-y\rangle^{-\mu} b(x,y;\xi) \in S^k \tag{3.8.5}$$

pick a cut-off function χ and split b,

$$b(x,y;\xi) = \chi(|x-y|)b(x,y;\xi) + (1-\chi(|x-y|))b(x,y;\xi). \tag{3.8.6}$$

The first term on the RHS lies in S^k and is supported near the diagonal. The second term, grows away from the diagonal, but is identically zero in a neighbourhood of it. One can show, by similar arguments to those elsewhere in this chapter, that the associated pseudodifferential operator has a smooth kernel, rapidly decaying away from the diagonal, that is, satisfying all estimates (3.8.1).

The above discussion suggests the following iterative process. Given $a(x,y;\xi)$, let $a_0(x;\xi) = a(x,x;\xi)$. Then $a(x,y;\xi) - a_0(x;\xi)$ vanishes on the diagonal and so is really a symbol $b_1(x,y;\xi)$ of order $\leqslant k-1$. Then define $a_1(x;\xi) = b_1(x,x;\xi)$ and continue.

This is the reason why we can get rid of the y-dependence in the more general class of symbols. Here is the exact result:

THEOREM 3.8.2. *If $a(x,y;\xi) \in S^k(\mathbb{R}^{2n};\mathbb{R}^n)$. Then there exists $\sigma_A(x;\xi)$,*

$$\sigma_A(x;\xi) \sim \sum_{\alpha} \frac{i^{|\alpha|}}{\alpha!} D_y^{\alpha} D_{\xi}^{\alpha} a(x,y;\xi)\,|_{y=x} \tag{3.8.7}$$

such that

$$\frac{1}{(2\pi)^n}\int a(x,y;\xi)e^{i\xi\cdot(x-y)}\,d\xi = \frac{1}{(2\pi)^n}\int \sigma_A(x;\xi)e^{i\xi\cdot(x-y)}\,d\xi. \tag{3.8.8}$$

In words, the pseudodifferential operators defined by a and by σ_A are equal.

PROOF. Use Taylor's Theorem with remainder in the form

$$a(x,y;\xi) = \sum_{|\alpha|<N} \frac{(-i)^{|\alpha|}}{\alpha!}(x-y)^\alpha D_y^\alpha a(x,x;\xi) + \sum_{|\alpha|=N} \frac{(-i)^{|\alpha|}}{\alpha!}(x-y)^\alpha R_{N,\alpha}(x,y;\xi) \tag{3.8.9}$$

and

$$R_{N,\alpha}(x,y;\xi) = \int_0^1 (1-t)^{N1} D_y^\alpha a(x,(1-t)x+ty;\xi), \mathrm{d}t. \tag{3.8.10}$$

Note that the sum of terms with $|\alpha| = j$ defines a symbol of order $k-j$ and the remainder term vanishes to order $k-N$ on the diagonal and so the corresponding pseudodifferential operator is of order $k-N$.

More precisely, the same integration by parts argument shows that the operator with amplitude

$$\frac{(-i)^{|\alpha|}}{\alpha!}(x-y)^\alpha D_y^\alpha a(x,x;\xi) \tag{3.8.11}$$

is the same as the operator with amplitude

$$\frac{i^{|\alpha|}}{\alpha!} D_\xi^\alpha D_y^\alpha a(x,x;\xi) \tag{3.8.12}$$

Choose an asymptotic sum $b(x;\xi)$ of the symbols (3.8.11). If B is the corresponding operator, then we have, for every N,

$$A - B = \sum_{j=0}^{N-1} A_j + R_N - B \text{ so } B = \sum_{j=0}^{N-1} A_j + S_N \tag{3.8.13}$$

where $S_N \in \Psi^{k-N}$. It follows that $A - B \in \Psi^{-\infty}$ and we may invoke the residuality characterization to complete the proof. □

3.8.3. Formal properties continued: adjoints. The L^2 inner product for functions on $\mathbb{R}^n$ is

$$(u,v) = \int v^*(x)u(x)\,\mathrm{d}x. \tag{3.8.14}$$

If A is a pseudodifferential operator with amplitude $a(x,y;\xi)$, then the (formal) adjoint A^* is defined by requiring

$$(u, Av) = (A^*u, v) \text{ for all } u, v \in \mathscr{S}. \tag{3.8.15}$$

The following is a simple computation, but is very important:

PROPOSITION 3.8.3. *The formal adjoint A^* of A is a pseudodifferential operator with amplitude*

$$(x,y;\xi) \longmapsto a^*(y,x;\xi).$$

In particular, if $A \in \Psi^k$, then $A^* \in \Psi^k$ and

$$\sigma_k(A^*) = \sigma_k(A)^*. \tag{3.8.16}$$

The proof is left to the reader. The result remains true for 'systems' if a^* is interpreted as the adjoint (conjugate-transpose) of the matrix a.

Less trivial is the following:

THEOREM 3.8.4. *Let A be a pseudodifferential operator with (full) symbol σ_A. Then*

$$\sigma_{A^*} = \sum_\alpha \frac{i^{|\alpha|}}{\alpha!} D_x^\alpha D_\xi^\alpha \sigma_A^*(x;\xi) \tag{3.8.17}$$

We defer the proof.

A corollary of the above considerations is the following:

PROPOSITION 3.8.5. *Let A be a pseudodifferential operator with amplitude $a(x,y;\xi)$. Then there is a unique 'left symbol' $\widetilde{\sigma}_A(y;\xi)$ such that*

$$K_A(x,y) = \frac{1}{(2\pi)^n} \int \widetilde{\sigma}_A(y;\xi) e^{i\xi\cdot(x-y)}\, d\xi. \tag{3.8.18}$$

Moreover,

$$\widetilde{\sigma}_A(y;\xi) = \sum \frac{(-i)^{|\alpha|}}{\alpha!} D_y^\alpha D_\xi^\alpha \sigma(y;\xi). \tag{3.8.19}$$

PROOF. Start from the representation of A^* in terms of its full symbol $\sigma_A^*(x;\xi)$. By preceding remarks, it follows that A is also represented by the operator (3.8.18) where

$$\widetilde{\sigma}_A(y;\xi) = \sigma_{A^*}(y;\xi)^*. \tag{3.8.20}$$

Now apply Theorem 3.8.4 to obtain (3.8.19). □

3.8.4. Proof of the Theorem 3.6.1. This is now straightforward. The exactness of the symbol sequence follows immediately by use of the representation $A = \mathrm{Op}(\sigma_A, D)$.

As for composition, we prove the following more precise version of (3.6.2).

THEOREM 3.8.6. *Suppose that A and B are pseudodifferential operators with (full) symbols σ_A and σ_B respectively. Then*

$$\sigma_{AB} \sim \sum_\alpha \frac{i^{|\alpha|}}{\alpha!} D_\xi^\alpha \sigma_A(x;\xi) D_x^\alpha \sigma_B(x;\xi) \tag{3.8.21}$$

PROOF. This is now very easy. Write A interms of its symbol $\sigma_A(x;\xi)$ and B in terms of its symbol $\widetilde{\sigma}_B(y;\xi)$. Then

$$Bu(y) = \frac{1}{(2\pi)^n} \int \widetilde{\sigma}_B(z;\eta) e^{i\eta\cdot(y-z)} u(z)\,\mathrm{d}z\mathrm{d}\eta \tag{3.8.22}$$

and so

$$ABu(x) = \frac{1}{(2\pi)^{2n}} \int \sigma_A(x;\xi)\widetilde{\sigma}_B(z;\eta) e^{i\xi\cdot(x-y)+i\eta\cdot(y-z)} u(z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}\xi\mathrm{d}\eta. \tag{3.8.23}$$

Doing the y integral first produces $(2\pi)^n\delta(\xi-\eta)$, so then doing the η integral yields

$$ABu(x) = \frac{1}{(2\pi)^n} \int \sigma_A(x;\xi)\widetilde{\sigma}_B(z;\xi) e^{i\xi\cdot(x-z)} u(z)\,\mathrm{d}x\mathrm{d}\xi. \tag{3.8.24}$$

Since $\sigma_A(x;\xi)\widetilde{\sigma}_B(y;\xi)$ is a symbol given that σ_A and $\widetilde{\sigma}_B$ are, it follows that AB is a pseudodifferential operator of order $k+\ell$. The formula for σ_{AB} from Theorem 3.8.6 follows from the formulae for writing a general symbol in reduced form. □

3.8.5. Mapping properties. We shall sketch the proof that if $A \in \Psi^k$, then A is a continuous linear map $\mathscr{S} \to \mathscr{S}$. By Corollary 3.4.2, it is sufficient to show that if $a \in S^{-\infty}$ and $f \in \mathscr{S}$, then given any μ,

$$\|Af\|_\mu \leqslant C_{\mu\nu}\|f\|_\nu \tag{3.8.25}$$

for some ν, where the constant $C_{\mu\nu}$ is controlled by one of the norms on S^ℓ. Here the subscripts refer to the norms on $\mathscr{S}$ (not Sobolev norms!), i.e.

$$\|f\|_\mu = \sum_{|\alpha|+|\beta|\leqslant\mu} \sup|x^\alpha D^\beta f|. \tag{3.8.26}$$

We start with the base case $\mu = 0$. We use essentially the same device as at the beginning of the chapter, the identity

$$\frac{1+\xi\cdot D_y}{1+|\xi|^2}e^{i\xi\cdot(x-y)} = e^{i\xi\cdot(x-y)}. \tag{3.8.27}$$

Choose L so large that

$$L \geqslant \ell+n+1 \tag{3.8.28}$$

and recall the notation

$$\langle\xi\rangle = \sqrt{1+|\xi|^2}. \tag{3.8.29}$$

We may assume that $a = a(y;\xi)$. Then

$$Af(x) = \frac{1}{(2\pi)^n}\int e^{i\xi\cdot(x-y)}a(y;\xi)f(y)\,\mathrm{d}\xi\mathrm{d}y. \tag{3.8.30}$$

Since a has compact support we may integrate by parts L times using the identity (3.8.27), obtaining

$$\begin{aligned} Af(x) &= \frac{1}{(2\pi)^n}\int e^{i\xi\cdot(x-y)}\left(\frac{1-\xi\cdot D_x}{1+|\xi|^2}\right)^L (a(y;\xi)f(y))\,\mathrm{d}\xi\mathrm{d}y \\ &= \frac{1}{(2\pi)^n}\int e^{i\xi\cdot(x-y)}\left(\frac{1}{1+|\xi|^2}\right)^L (a(y;\xi)f(y))\,\mathrm{d}\xi\mathrm{d}y \end{aligned} \tag{3.8.31}$$

Since $\langle\xi\rangle^{-L}a(y;\xi)$ is bounded by a multiple of $\langle\xi\rangle^{-n-1}$,

$$|Af(x)| \leqslant C\left(\int\frac{\mathrm{d}\xi}{\langle\xi\rangle^{n+1}}\right)\|f\|_L \leqslant C\|f\|_L. \tag{3.8.32}$$

To proceed further we use induction and some estimates of $[A,D]$ and $[A,M_j]$, where M_j is the operation of multiplication by x_j. The idea is as follows:

$$D_jAf = AD_jf - [A,D_j]f. \tag{3.8.33}$$

The first term on the RHS is controlled by the L-Schwartz norm of D_jf, i.e. by the $(L+1)$-Schwartz norm of f. The second term on the RHS is controlled provided that $[A,D_j]$ satisfies an estimate similar to (3.8.32). An estimate similar to (3.8.32) with A replaced by $[A,M_j]$, will give a bound on $\sup|x_jAf|$.

Now the kernel of $[A,D_j]$ is

$$(D_{x_j}+D_{y_j})K_A \tag{3.8.34}$$

which is of the form (3.3.11) but with

$$a(x,y;\xi) \text{ replaced by } (D_{x_j}+D_{y_j})a(x,y;\xi), \tag{3.8.35}$$

since $(D_{x_j} + D_{y_j})$ annihilates the exponential factor. This is again a symbol, and in particular we can bound a finite number of its derivatives with respect to (x, y) by one of the norms on S^ℓ.

Similarly, the kernel of $[A, M_j]$ is equal to

$$(x_j - y_j)K_A(x, y) \tag{3.8.36}$$

and an integration by parts shows that this is again of the form (3.3.11), where a is replaced by $D_{\xi_j}a(x, y; \xi)$. If a is a symbol of order ℓ then $D_\xi a$ is a symbol of order $\ell - 1$ by definition.

Repeated use of these commutator calculations allows us to prove (3.8.25) for every μ.

3.8.6. Proof of Proposition 3.6.4. The proof is made by copying the argument starting from (3.1.4). We must replace this equation by

$$(x - y)^\alpha e^{i\xi\cdot(x-y)} = D_\xi^\alpha e^{i\xi\cdot(x-y)} \tag{3.8.37}$$

Applying this in the formula for K_A and integrating by parts,

$$(x - y)^\alpha K_A(x, y) = \frac{1}{(2\pi)^n} D_\xi^\alpha a(x, y; \xi) e^{i\xi\cdot(x-y)}\, \mathrm{d}\xi \tag{3.8.38}$$

where again we may assume $a \in S^{-\infty}$ and argue by density provided we make estimates with respect to the norms defining the topology of S^ℓ. If $|\alpha| \geqslant \ell+n+1$ then as before the integral is absolutely and uniformly convergent and bounded by a multiple of the $|\alpha|$-norm on S^ℓ. Hence there is a constant C_α such that

$$(x - y)^\alpha K_A(x, y) \leqslant C_\alpha. \tag{3.8.39}$$

From this it follows that $|K_A(x, y)| \leqslant C|x - y|^{-\ell-n-1}$ and in particular K_A is continuous away from the diagonal. (This is not a sharp bound). A similar argument applies to estimate

$$(x - y)^\alpha D_x^\beta D_y^\gamma K_A(x, y) \tag{3.8.40}$$

where the choice of α (or rather a lower bound for α) is dictated by the choices of β and γ.

CHAPTER 4

Pseudodifferential operators on manifolds

4.1. Introduction

In the previous Chapter we moved from using the Fourier transform to defining a parametrix for constant-coefficient elliptic operators in $\mathbb{R}^n$ to the definition of pseudodifferential operators in $\mathbb{R}^n$.

Recall that a pseudodifferential operator of order k is an operator $A : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ whose Schwartz kernel has the form

$$K_A(x, y) = \frac{1}{(2\pi)^n} \int a(x, y; \xi) e^{i\xi\cdot(x-y)}\, \mathrm{d}\xi, \tag{4.1.1}$$

where the *amplitude* a lies in the class of symbols of order k. The set of all pseudodifferential operators of order k is denoted Ψ^k or $\Psi^k(\mathbb{R}^n)$. The subspace of polyhomogeneous pseudodifferential operators arises by taking $a \in S^k_{\mathrm{phg}}$ and is denoted Ψ^k_{phg}. Recall that the integral defining K_A is only absolutely and uniformly convergent if $a \in S^{-n-1}$. Otherwise some fancy footwork is required. One trick is the observation from the last chapter that $S^{-\infty}$ is dense in S^k in the topology of S^ℓ, for every $\ell > k$. This means that we can always think of a formula like (4.1.1) as a limit in which a is replaced by $\chi(\varepsilon|\xi|)a(x;\xi)$, χ being a standard cut-off function as in Chapter 3, §3.4.

In any case, the integral defining K_A is a tempered distribution on $\mathbb{R}^{2n}$ which is smooth away from the diagonal.

4.2. Pseudodifferential operators on manifolds

Let X be a compact manifold without boundary (connected and orientable, for the sake of argument).

- Extension of distributions to manifolds;
- Schwartz kernel theorem for manifolds

Remarks about the need for a density.

In particular any linear map $A : C^\infty(X) \to C^{-\infty}(X)$ has a Schwartz kernel $K_A(x, y)$ which is a section of $1 \boxtimes \Omega$, in other words it is a density 'in the y variables'.

- Vector bundle extension

Definition 4.2.1. The linear operator A is a pseudodifferential operator of order k on X if for smooth functions...to be added

By patching arguments, one proves the analogues of the main theorems stated in §3.6, more precisely:

Theorem 4.2.2. *Let X be a compact manifold and let E and F be complex vector bundles over X.*

(i) *Symbol sequence: for every* $k \in \mathbb{R}$*, there is an exact sequence*

$$0 \longrightarrow \Psi^{k-1}(X;E,F) \longrightarrow \Psi^k(X;E,F) \xrightarrow{\sigma_k} \mathrm{Sym}_k(X;E,F) \to 0. \tag{4.2.1}$$

(ii) *Composition: if* $A \in \Psi^k(X;F,G)$ *and* $B \in \Psi^\ell(X;E,F)$ *then* $AB \in \Psi^{k+\ell}(X;E,G)$ *(with obvious interpretations if either* k *or* ℓ *is equal to* $-\infty$*). The principal symbol is multiplicative in the sense that*

$$\sigma_{k+\ell}(AB) = \sigma_k(A)\sigma_\ell(B). \tag{4.2.2}$$

(iii) *The principal symbol map is a* $*$*-homomorphism in the sense that* $\sigma_k(A^*) = \sigma_k(A)^*$*, where* A^* *is the* L^2*-adjoint of* A *(for chosen metrics on the bundles and a volume form on* X*).*

There is a parallel statement for the subclass of polyhomogeneous operators, in which everything is adorned with the subscript phg.

Patching arguments also prove that if $A \in \Psi^k(X;E,F)$, then A is bounded as a map between Sobolev spaces $H^s(X;E) \to H^{s-k}(X;F)$.

4.3. The elliptic package for differential operators on compact manifolds

We start by noting that the results of the previous section imply the existence of a parametrix for any elliptic operator on X:

THEOREM 4.3.1. *Let* $P : C^\infty(X;E) \to C^\infty(X;F)$ *be an elliptic operator of order* k*, where* X *is compact and* E *and* F *are complex vector bundles. Then there exists a parametrix* $A \in \Psi^{-k}(X;F,E)$ *for* P*, that is an operator such that*

$$PA = 1 - R_1, \quad AP = 1 - R_2 \tag{4.3.1}$$

where R_1 *and* R_2 *are smoothing operators.*

PROOF. Given Theorem 4.2.2, either of the iterative arguments in §3.7 can be used without essential change. □

Combining this with a little functional analysis, we obtain the following fundamental theorem.

THEOREM 4.3.2. *Let* $P : C^\infty(X;E) \to C^\infty(X;F)$ *be an elliptic operator of order* k*, where* X *is compact and* E *and* F *are complex vector bundles. Then for every* s*,*

$$P : H^s(X;E) \to H^{s-k}(X;F) \tag{4.3.2}$$

is bounded and Fredholm*. The kernel consists of smooth sections of* E *and the range, for any* s*, can be complemented by smooth sections of* F*. (One may choose metrics and take these smooth sections to be a basis for the kernel of the formal adjoint operator* P^**.*

Moreover, the generalized inverse G *defined by inverting* P *on its range, picking the solution orthogonal to* $\ker P$*, and to be zero on the orthogonal complement of* $\operatorname{im} P$ *lies in* $\Psi^{-k}(X;F,E)$ *and differs from any parametrix by smoothing operators.*

The facts we needd from functional analysis are:

- The unit ball B in a Hilbert space H is compact if and only if H is finite-dimensional.

- A smoothing operator R on a compact manifold is compact in L^2 in the sense: if $\|f_j\| \leqslant 1$ in L^2 for all j, then Rf_j has an L^2-convergent subsequence.

Proof. That the kernels of P and of P^* consist of smooth sections follows as in $\mathbb{R}^n$: if $Pu = 0$, then

$$APu = u - R_1 u = 0 \tag{4.3.3}$$

and so $u = R_1 u$. Since R_1 is smoothing, u is smooth. First of all, the existence of a parametrix guarantees that $\ker P$ consists of smooth sections of E. Consider $B = \ker P \cap \{u : \|u\|_{L^2} \leqslant 1\}$.

Because $B = R_1(B)$ the second of our 'standard facts' shows that B is precompact in L^2. Since it is also closed, it is compact. Hence $\ker P$ is finite-dimensional by the first of our standard facts.

We move on to the proof that the range of $P : H^s \to H^{s-k}$ is closed.

Let π_1 be the orthogonal projection from $L^2(E)$ onto $\ker P$ and let $W = \ker(P)^\perp \cap H^s$. We claim that there is an estimate of the form:

$$\|Pu\|_{s-k} \geqslant C\|u\|_s \text{ for all } u \in W. \tag{4.3.4}$$

where $C > 0$ and the subscripts denote Sobolev norms.

If (4.3.4) fails, there is a sequence u_n with

$$\|u_n\|_s = 1, \quad Pu_n \to 0. \tag{4.3.5}$$

Applying A, we obtain

$$u_n - R_1 u_n \to 0. \tag{4.3.6}$$

Using that R is compact again, it follows that u_n has a convergent subsequence, which we may assume is the original one. The limit, u_∞ must have norm 1. Then $Pu_\infty = 0$, yet u_∞ is orthogonal to $\ker P$, contradiction.

Now let us show that $\operatorname{im} P$ is closed. Suppose that f_n is a sequence in $\operatorname{Im} P$. In particular there exist $u_n \in L^2(X; E)$ with $Pu_n = f_n$. We assume that $f_n \to f$, and need to find u such that $Pu = f$. By replacing u_n with $(1 - \pi_1)u_n$, we may suppose that $u_n \in W$. Then (4.3.4) gives

$$\|u_n\| \leqslant \frac{1}{C}\|f_n\| \tag{4.3.7}$$

and so the u_n are uniformly bounded. Applying the parametrix to the equation $Pu_n = f_n$ yields

$$u_n = R_1 u_n + Af_n. \tag{4.3.8}$$

Now $Af_n \to Af$ and since we now know that $\|u_n\|$ is uniformly bounded, passage to a subsequence gives, again by the compactness of the operator R_1, that u_n is convergent. The limit must satisfy $Pu = f$, showing that the range is closed.

Since the range is closed, the orthogonal complement of $PH^s(X; E)$, which is isomorphic to $\ker P^*$ can be identified with the cokernel of P. This shows everything except that the generalized inverse is a pseudodifferential operator.

We have

$$GP = 1 - \pi_1, \quad PG = 1 - \pi_2 \tag{4.3.9}$$

and

$$AP = 1 - R, \quad PA = 1 - R'. \tag{4.3.10}$$

Hence

$$GPA = (1 - \pi_1)A = G(1 + R') \text{ so } G = A - \pi_1 A - GR'. \tag{4.3.11}$$

Similarly,

$$APG = A - A\pi_2 = G + RG \text{ so } G = A - A\pi_2 - RG. \tag{4.3.12}$$

Notice that neither of (4.3.11) or (4.3.12) implies that G is a pseudodifferential operator, because it is not clear that RG or GR' is a pseudodifferential operator. However, we can combine them so that G appears sandwiched between the smoothing operators R and R':

$$G = A - \pi_1 A + A\pi_2 - AR_1 + R_1 G R_1. \tag{4.3.13}$$

Now we are in good shape because the first four terms are all pseudodifferential operators, and the last is smoothing (being a composite of smoothing $\times$ bounded $\times$ smoothing). □

Remark 4.3.3. Another characterization of the generalized inverse is

$$GP = 1 - \pi_1, \quad PG = 1 - \pi_2 \tag{4.3.14}$$

where π_1 is the L^2 projection onto $\ker P$ and π_2 is the orthogonal projection onto $\operatorname{im} P^{\perp}$.

Remark 4.3.4. The existence of the parametrix and its boundedness in Sobolev spaces implies the standard elliptic estimates

$$\|u\|_s \leqslant C(\|Pu\|_{s-k} + \|u\|_0) \tag{4.3.15}$$

if $s > 0$. (Here $\|\cdot\|_0$ stands for the H^0 i.e. L^2-norm.)

Corollary 4.3.5 (The Fredholm Alternative). *Let $P : C^\infty(X;E) \to C^\infty(X;F)$ be an elliptic operator of order k and suppose that we have metrics on E and F and a given volume element on X, giving everything L^2 metrics. Let P^* denote the formal adjoint of P. Then given $f \in H^s(X;F)$, the equation*

$$Pu = f \tag{4.3.16}$$

is solvable for $u \in H^{s+k}(X;E)$ if and only if $f \perp \ker(P^*)$.

Proof. Follows at once from the above. □

Example 4.3.6. On a compact connected riemannian manifold X, the equation $\Delta u = f$ if and only if $\int_X f \, \mathrm{d}\mu = 0$. This follows from the Fredholm alternative, for Δ is formally self-adjoint. So the equation is solvable if and only if $f \perp \ker \Delta$. But the kernel consists just of the constants:

$$\Delta u = 0 \Rightarrow \int u \Delta u = 0 \Leftrightarrow \int |\mathrm{d}u|^2 = 0. \tag{4.3.17}$$

On a connected manifold, $\mathrm{d}u = 0$ implies u is a constant.

4.4. Trace and index—topological properties of index

Apart from allowing us to prove the basic properties of elliptic operators in a considerable degree of generality, one can also obtain some general qualitative results about the index of elliptic operators.

We have now seen that if X is a compact manifold and P is an elliptic differential operator of order k between complex vector bundles E and F, then P is a Fredholm operator $H^s(X;E) \longrightarrow H^{s-k}(F)$. For a Fredholm operator P, the *index* is defined as

$$\operatorname{Ind}(P) = \dim \ker P - \dim \operatorname{Coker} P. \tag{4.4.1}$$

The celebrated Index Theorem of Atiyah and Singer gives a formula for (4.4.1) in terms of topological data (the topology of the bundles E and F and of the principal symbol $\sigma_k(P)$).

We shall not say more about this formula, but we shall describe an approach to its study, via parametrices and *traces*. In particular, we shall obtain some qualitative information about the index (in particular some of its stability properties—i.e. its invariance under deformation of the operator P).

4.4.1. Trace of a smoothing operator. As usual I am missing out some details here, but what follows should be nearly self-contained. Let X be a compact manifold, let $E \to X$ be a complex vector bundle, and let $R : C^\infty(X;E) \to C^\infty(X;E)$ be a smoothing operator. Recall that this means that $K_R \in C^\infty(X \times X; E \boxtimes (E^* \otimes \Lambda^n))$ where the second factor is the tensor product of the dual of E with the n-forms on X, so that if $f \in C^\infty(X;E)$, $R\,\mathrm{pr}_1^*(f)$ lies in $\mathrm{pr}_1^*(E) \otimes \mathrm{pr}_2^* \Lambda^n$ so that the push-forward or integral over the fibres of pr_1 is well-defined.

Definition 4.4.1. For such a smoothing operator R, the trace of R, $\mathrm{Tr}(R)$ is defined to be

$$\mathrm{Tr}\, R = \int_\Delta \mathrm{tr}(K_R). \tag{4.4.2}$$

In local coordinates, near the diagonal, K_R has the form $e(x,y)\,\mathrm{d}y$, where $e(x,y) \in E_x \otimes E_y^*$. Then the pull-back to the diagonal is equal to $e(x,x)\,\mathrm{d}x$, and since $e(x,x) \in \mathrm{End}(E_x)$, $\mathrm{tr}\, e(x,x)$ is well-defined. This is the meaning of (4.4.2)

Example 4.4.2. Let E be a vector bundle over a compact manifold X and suppose that E is equipped with a fibre metric h and volume element $\mathrm{d}\mu$. Let V be a finite-dimensional subspace of $C^\infty(X;E)$. We are going to write down the Schwartz kernel K_R of the orthogonal projection operator R onto V.

Let $\dim V = d$ and let $\sigma_1, \dots, \sigma_d$ be an orthonormal basis of V, so that

$$\int_X h(\sigma_i, \sigma_j)\,\mathrm{d}\mu = \delta_{ij}. \tag{4.4.3}$$

(We assume that h is antilinear in the second variable.)

Now for any section f of E, we define a section f^* of E^* as follows: for each $x \in X$,

$$f_x^* : E_x \to \mathbb{C},\, f_x^*[u_x] = h_x(u_x, f_x) \text{ for all } u_x \in E_x. \tag{4.4.4}$$

We claim that

$$K_R(x,y) = \sum_{j=1}^d e_j(x) \boxtimes e_j^*(y). \tag{4.4.5}$$

Then as a section of $\mathrm{pr}_1^*(E)$ over $X \times X$,

$$K_R(x,y)f(y) = \sum_{j=1}^d e_i(x) h_y(f_y, e_{j,y}). \tag{4.4.6}$$

Integration with respect to y now gives

$$\int_X K_R(x,y) f(y)\,\mathrm{d}\mu_y = \sum_{j=1}^d e_i(x) \left(\int_X h(f, e_j)\,\mathrm{d}\mu \right). \tag{4.4.7}$$

In particular if f is in the L^2-orthogonal complement of V then $Rf = 0$ (since f is orthogonal to each of the e_j). And similarly, $Re_j = e_j$ for each j, as required.

REMARK 4.4.3. It is possible to extend the definition of Tr considerably: for example to operators whose kernels are merely continuous, but we shall not need this.

It should also be noted that there is a theory of traces for bounded linear operators on a Hilbert space H. If $\dim H = \infty$, not every operator has a trace. However, there is a set of operator which *do* have traces, and this is an ideal in the algebra of all bounded operators. (Possible references are the little book on trace ideals by Barry Simon, or Hörmander, Analysis of partial differential operators, Chapter 19, Volume III.)

From the above considerations, it follows if R is orthogonal projection onto V, then

$$\dim V = \operatorname{Tr}(R). \tag{4.4.8}$$

Now let P be an elliptic operator between bundles E and F over X. Let G be the generalized inverse as in Theorem 4.3.1 and let π_1 be the projection onto $\ker P$, π_2 the orthogonal projection onto $\ker P^*$,

$$GP = 1 - \pi_1, \quad PG = 1 - \pi_2. \tag{4.4.9}$$

Hence

$$\operatorname{Ind}(P) = \dim\ker P - \dim\ker P^* = \operatorname{Tr}\pi_1 - \operatorname{Tr}\pi_2. \tag{4.4.10}$$

Combining this essentially trivial result with formal properties of the trace, we obtain the much more useful

THEOREM 4.4.4. *Let P be elliptic as above and let A be a parametrix, so*

$$AP = 1 - R_1, \quad PA = 1 - R_2. \tag{4.4.11}$$

Then

$$\operatorname{Ind}(P) = \operatorname{Tr} R_1 - \operatorname{Tr} R_2 \tag{4.4.12}$$

The key to proving this is the commutator property for the trace:

PROPOSITION 4.4.5. *Let A be a pseudodifferential operator and let R be smoothing. Then*

$$\operatorname{Tr}(AR) = \operatorname{Tr}(RA). \tag{4.4.13}$$

PROOF. By partitions of unity subordinate to trivializing coordinate charts, we reduce the result to proving that if A is pseudodifferential operator in $\mathbb{R}^n$ and and R is a smoothing operator supported in a 'box' $\{|x| < r\} \times \{|y| < r\}$, then (4.4.13) holds for such A and R. If A were smoothing the result follows from the formulae for composing smoothing operators, for

$$K_{AR}(x,z) = \int A(x,y)R(y,z)\,\mathrm{d}y, \quad K_{RA}(x,z) = \int R(x,y)A(y,z)\,\mathrm{d}y, \tag{4.4.14}$$

Hence

$$\operatorname{Tr} AR = \int \operatorname{tr}(A(x,y)R(y,x))\,\mathrm{d}x\mathrm{d}y, \quad \operatorname{Tr} RA = \int \operatorname{tr}(R(x,y)A(y,x))\,\mathrm{d}x\mathrm{d}y. \tag{4.4.15}$$

That these are equal follows from interchanging x and y in the second, and using the fact that pointwise,

$$\operatorname{tr} A(x,y)R(y,x) = \operatorname{tr} R(y,x)A(x,y). \tag{4.4.16}$$

(All integrals and manoeuvres are well-defined because R is compactly supported in a box.) The general result follows from this by approximating the kernel of A by smoothing operators. We know that this can be done by Corollary 3.4.2 from Chapter 3. □

Proof. (of Theorem 4.4.1). Let A be any parametrix, G the generalized inverse. We have seen that G is a pseudodifferential operator and differs from A by a smoothing operator, S, say, so

$$G = A + S. \tag{4.4.17}$$

Then

$$1 - \pi_1 = GP = (A+S)P = 1 - R_1 + SP \text{ so that } R_1 = \pi_1 + SP. \tag{4.4.18}$$

Similarly,

$$R_2 = \pi_2 + PS. \tag{4.4.19}$$

Hence

$$\operatorname{Tr} R_1 = \operatorname{Tr} \pi_1 + \operatorname{Tr} SP, \;\; \operatorname{Tr} R_2 = \operatorname{Tr} \pi_2 + \operatorname{Tr} PS = \operatorname{Tr} \pi_2 + \operatorname{Tr} SP. \tag{4.4.20}$$

since S is smoothing. Subtracting,

$$\operatorname{Tr} R_1 - \operatorname{Tr} R_2 = \operatorname{Tr} \pi_1 - \operatorname{Tr} \pi_2 = \operatorname{Ind} P. \tag{4.4.21}$$

□

From this result, we obtain the following stability result for the index:

Theorem 4.4.6. *Let $P_t : C^\infty(X;E) \to C^\infty(X;F)$ for $t \in [0,1]$ be a continuous family of elliptic operators of order k. Then*

$$\operatorname{Ind} P_0 = \operatorname{Ind} P_1. \tag{4.4.22}$$

Remark 4.4.7. The continuity is easily defined either by working locally and demanding that the coefficients, with respect to trivializations which are independent of t, depend continuously on t, or by us of a fixed connection to define P_t and demanding that the various symbol maps (see Chapter 1) are all continuous in t.

Example 4.4.8. As a particular case, suppose that P and Q are two elliptic operators with $\sigma_k(P) = \sigma_k(Q)$. Then $P - Q = L$, say, is of order $\leqslant k-1$. Then if we set

$$P_t = (1-t)P + tQ = P - tL \tag{4.4.23}$$

we have $\sigma_k(P_t) = \sigma_k(P)$ for all t. The theorem applies and gives

$$\operatorname{Ind} P = \operatorname{Ind} Q. \tag{4.4.24}$$

To be colloquial, the index doesn't depend on the 'lower order terms' of an elliptic operator.

PROOF. Since X is compact, we may choose a cover of X by trivializing coordinate charts, and also fix once and for all a partition of unity subordinate to this cover. We can use these data to define a definite 'quantization' or 'lifting' map

$$S^k(X) \to \Psi^k(X). \tag{4.4.25}$$

Using this, given a continuous family of elliptic operators P_t, one can show that there exists a continuous family of parametrices A_t. Then automatically, the error terms R_t and R'_t will depend continuously on t since

$$R_t = 1 - A_t P_t, \;\; R'_t = 1 - P_t A_t. \tag{4.4.26}$$

Hence $t \mapsto \operatorname{Tr} R_t$ and $t \mapsto \operatorname{Tr} R'_t$ are continuous in t and so is their difference, which, for each t, is equal to $\operatorname{Ind} P_t$. However this is an integer, so must be independent of t. The result follows. □

This result shows two highly significant properties of the index. From the topological point of view, it shows that one can define a map from homotopy classes of elliptic symbols to $\mathbb{Z}$: pick a symbol in the homotopy class, lift it to an operator then map to the index. By the above results, the integer is independent of all choices. This is usually called the analytic index a-Ind of an elliptic symbol. In the early 1960's Atiyah and Singer defined another map, t-Ind from elliptic symbols (better interpreted in terms of K-theory). t-Ind is defined in purely topological terms. They proved that $t\text{-Ind} = a\text{-Ind}$ giving a 'formula' for a-Ind in purely topological terms.

On the other hand, another aspect of Theorem 4.4.1 is that it shows that the the index of an elliptic operator is expressible, at least in principle, in terms of traces and hence as the integral of a top-degree differential form over X. Note that the smoothing operators we need the traces of are in principle determined algorithmically by the coefficients of our operator P.

In the second part of this course, this idea will be pursued for elliptic operators of 'Dirac type'. This will yield the index of operators of Dirac type as the integral of an explicit differential form on X: moreover, this differential form has a natural interpretation in terms of characteristic classes of the bundles involved.

CHAPTER 5

Elliptic complexes and Dirac operators

5.1. Introduction

5.2. Elliptic complexes

A complex $(E^{\cdot}, D^{\cdot})$ is a sequence of bundles and differential operators

$$0 \longrightarrow C^{\infty}(E^0) \xrightarrow{D^0} C^{\infty}(E^1) \xrightarrow{D^1} C^{\infty}(E^2) \cdots C^{\infty}(E^N) \to 0. \tag{5.2.1}$$

with the property $D^j D^{j-1} = 0$ for all j. The canonical example is the de Rham complex in which $E^j = \Lambda^j T^* X$ and $D^j = \mathrm{d}$. In order to simplify things, we shall assume that in general all the D^j are of order 1 and we shall abuse notation by denoting them all by D. We tacitly assume that all the E^j are genuine non-zero vector bundles, but it is sometimes convenient to augment (5.2.1) by adjoining $E^{-1} = 0$ and $E^{N+1} = 0$.

The condition $D^2 = 0$ means that the (generalized) cohomology groups are defined,

$$H^i(E^{\cdot}, D^{\cdot}) = \frac{\ker D : C^{\infty}(X, E^i) \to C^{\infty}(X, E^{i+1})}{D C^{\infty}(X, E^{i-1})} \tag{5.2.2}$$

These spaces can be quite bad in general, but not if the complex is elliptic:

Passing to symbols, (5.2.1) gives rise to a sequence of bundle maps

$$0 \to E^0_x \xrightarrow{\sigma_1(D)_{x,\xi}} E^1_x \xrightarrow{\sigma_1(D)_{x,\xi}} E^2_x \to \cdots E^N_x \to 0 \tag{5.2.3}$$

for each (x, ξ) in the total space of T^*X. $D^2 = 0$ implies that $\sigma(D)^2 = 0$ by the formal properties of the symbol.

Definition 5.2.1. The complex (5.2.1) is *elliptic* if the symbol sequence (5.2.3) is exact for all (x, ξ) with $\xi \neq 0$.

We shall prove a generalized Hodge theorem for elliptic complexes, to the effect that the cohomology spaces are finite-dimensional and that each cohomology class has a unique harmonic representative. But first some examples.

Example 5.2.2. De Rham complex Let $E^r = \Lambda^r T^* X$. Then $C^{\infty}(X; E^r) = \Omega^r(X)$ is the space of differential r-forms on X, and we have the exterior derivative d giving us a complex. The symbol map is

$$\sigma_{x,\xi}(\alpha) = i\xi \wedge \alpha \tag{5.2.4}$$

It is an interesting to check that the complex is elliptic, i.e. that if $\xi \wedge \beta = 0$, there exists α such that $\beta = \xi \wedge \alpha$.

Example 5.2.3. If X is a complex manifold, of complex dimension m, say, there is a bigrading of the (complex-valued) differential forms, where

$\Omega^{p,q}(X)$ consists of the forms with 'p dz_i and q d$\overline{z}_i$' in their local expressions, $(z_1, \ldots, z_m)$ being local holomorphic coordinates. The exterior derivative breaks as a sum

$$\mathrm{d} = \partial + \overline{\partial},\ \partial : \Omega^{p,q} \to \Omega^{p+1,q},\ \ \overline{\partial} : \Omega^{p,q} \to \Omega^{p,q+1}. \tag{5.2.5}$$

These satisfy

$$\partial^2 = 0,\ \overline{\partial}^2 = 0, \partial\overline{\partial} + \overline{\partial}\partial = 0 \tag{5.2.6}$$

by virtue of $\mathrm{d}^2 = 0$. In particular, for fixed p, we have the complex

$$0 \to \Omega^{p,0}(X) \xrightarrow{\overline{\partial}} \Omega^{p,1}(X) \xrightarrow{\overline{\partial}} \cdots \to \Omega^{p,m}(X) \to 0 \tag{5.2.7}$$

with cohomology $H^{p,*}(X)$ (the Dolbeault cohomology groups).

There is a generalization: if $E \to X$ is any *holomorphic* vector bundle, one can define a 'twisted' $\overline{\partial}$ operator, denoted $\overline{\partial}_E$, such that

$$\overline{\partial}_E : \Omega^0(X; E) \to \Omega^{0,1}(X; E) \tag{5.2.8}$$

which satisfies the Leibniz rule relative to $\overline{\partial}$ ($\overline{\partial}_E(fs) = \overline{\partial} f \otimes s + f \otimes \overline{\partial}_E s$ for every function f and section s) and such that $\overline{\partial}_E s = 0$ in $U \subset X$ if and only if s is a holomorphic section of E in U.

Then there is a unique extension of (5.2.8)

$$\Omega^0(X; E) \to \Omega^{0,1}(X; E) \to \Omega^{0,2}(X; E) \to \cdots \to \Omega^{0,m}(X; E) \tag{5.2.9}$$

defining a complex. The p-th cohomology group is denoted $H^p(X; \mathscr{O}(E))$ and is isomorphic to the sheaf cohomology of the sheaf $\mathscr{O}(E)$ of holomorphic sections of E. Such groups are very important basic invariants of complex manifolds. If we take E to be the p-th exterior power of the holomorphic cotangent bundle, (5.2.9) reduces to (5.2.7) so $H^{p,q}$ is the q-th sheaf cohomology group of $\Lambda^p(T^{1,0})^*$.

Example 5.2.4. Deformation complex. Elliptic complexes often arise in the study of moduli spaces. For example, suppose that X is a compact riemannian 4-manifold and $E \to X$ is a bundle with hermitian metric. If A is a unitary connection on E, the curvature $F(A)$ is a 2-form with values in $\mathrm{End}(E)$ (more specifically, the skew-adjoint endomorphisms of E). In 4 dimensions, there is the decomposition $\Omega^2 = \Omega^2_+ \oplus \Omega^2_-$ of the space of 2-forms, the $+1$ and -1 eigenspaces of the Hodge $*$ opearator. The connection A is called anti-self dual (ASD) if the component $F(A)^+$ of $F(A)$ in $\Omega^2_+(X; \mathrm{End}(E))$ is zero.

The moduli space of all such instantons

$$\mathscr{M} = \{A : F(A)^+ = 0\}/\operatorname{Aut}(X : E) \tag{5.2.10}$$

has been much studied and led Donaldson and others to revolutionary results about the differential topology of 4-manifolds.

We state without proof that the tangent space $T_[A]\mathscr{M}$ at a (gauge-equivalence class of) the connection A is the cohomology of the complex

$$0 \to \Omega^0(X; \mathfrak{u}(E)) \to \Omega^1(X; \mathfrak{u}(E)) \to \Omega^2_+(X; \mathfrak{u}(E)) \to 0. \tag{5.2.11}$$

Here $\mathfrak{u}(E)$ is the bundle of skew-adjoing endomorphisms of E. The middle space represents an infinitesimal variation a in A and the map to $\Omega^2_+(X : \mathfrak{u}(E)$ is the linearization $a \longmapsto \mathrm{d}^+_A a$ of the ASD equations. The first map $\phi \longmapsto \mathrm{d}_A\phi$ is the infinitesimal action of the group of gauge transformations $\mathrm{Aut}(E)$ on the space of connections. One can check that this is an elliptic complex provided that $F(A)^+ = 0$. Under good conditions (see the book of Donaldson and

Kronheimer, for example) the moduli space is smooth at $[A]$ with tangent space equal to the first cohomology of this complex.

5.2.1. Generalized Hodge Theorem. Return to a general elliptic complex (E, D). Choose hermitian metrics on the bundles E^j and a volume element on X. Then for each j,

$$D : C^\infty(X; E^j) \to C^\infty(X; E^{j+1}) \tag{5.2.12}$$

has an L^2 adjoint

$$D^* : C^\infty(X; E^{j+1}) \to C^\infty(X; E^j) \tag{5.2.13}$$

We now make the following construction. Define

$$W_0 = \bigoplus_j E^{2j}, \quad W_1 = \bigoplus_j E^{2j+1} \tag{5.2.14}$$

and $L_0 = D + D^*$ acting from W_0 to W_1 and $L_1 = D + D^*$ acting from W_1 to W_0. More precisely L_0 involves only the D^{2j} and the adjoints $(D^{2j+1})^*$, while for L_1 it is the other way around.

PROPOSITION 5.2.5. *The complex (E, D) is elliptic if and only if*

$$L_0 : C^\infty(X; W_0) \to C^\infty(X; W_1) \tag{5.2.15}$$

is elliptic (if and only if L_1 is elliptic). In this case the two 'Laplacians' L_1L_0 and L_0L_1 are also self-adjoint elliptic operators.

REMARK 5.2.6. Both L_1L_0 and L_0L_1 are operators of Laplace type and preserve the degree. More precisely, acting on $s \in C^\infty(X; E^{2j})$,

$$L_1L_0 s = (D^{2j})^* D^{2j} + D^{2j-1}(D^{2j-1})^*. \tag{5.2.16}$$

PROOF. We show first that L_1L_0 is elliptic if and only if the complex (E, D) is elliptic. Let us simplify notation even further by just writing $\sigma = \sigma(D)$ and $\sigma^* = \sigma(D)^* = \sigma(D^*)$. (It is to be understood that the symbols are all evaluated at some fixed $\xi \neq 0$.) Then

$$\sigma(L_1L_0) = \sigma\sigma^* + \sigma^*\sigma. \tag{5.2.17}$$

We have seen in (5.2.16) that L_1L_0 preserves degree and maps $E2j$ into itself. Thus to check that (5.2.17) is an isomorphism it is enough, by elementary linear algebra, to check that it is injective on E^{2j} for each j.

So suppose that the complex is elliptic and that $e \in E^{2j}$ is annihilated by (5.2.17),

$$(\sigma\sigma^* + \sigma^*\sigma)e = 0. \tag{5.2.18}$$

Taking the inner product with e we learn that

$$|\sigma^* e|^2 + |\sigma e|^2 = 0 \text{ and hence } \sigma e = 0, \sigma^* e = 0. \tag{5.2.19}$$

Since the complex is elliptic, there exists $\eta \in E^{2j-1}$ such that $\sigma\eta = e$. But then the last of (5.2.18) gives $\sigma^*\sigma\eta = 0$, hence $\sigma\eta = 0$. Hence $e = 0$ as required.

Conversely, suppose that L_1L_0 is elliptic so that (5.2.17) is invertible. Given $e \in E^{2j}$ with $\sigma(e) = 0$, we must find $\eta \in E^{2j-1}$ such that $\sigma(\eta) = e$. For this, let A be the inverse of (5.2.17) and set

$$\eta = \sigma^* Ae. \tag{5.2.20}$$

Then η does the job. The crucial observation is that the endomorphism $\sigma\sigma^*$ of E^{2j} commutes with A. This follows at once because $\sigma\sigma^*$ commutes with $\sigma\sigma^* + \sigma^*\sigma$. Hence

$$\sigma(\eta) = \sigma\sigma^* A e = A \circ (\sigma\sigma^*) e = A A^{-1} e = e \tag{5.2.21}$$

where in the third equality we used $\sigma(e) = 0$.

We finish off the proof by noting that the ellipticity of $L_1 L_0$ is equivalent to that of L_0 (or L_1). This uses exactly the same kind of positivity arguments and $D^2 = 0$ that we've seen before. Indeed, since $\sigma(L_1) = \sigma(L_0)^*$ it is elementary linear algebra that $\sigma(L_0)$ is an isomorphism if and only if $\sigma(L_0)^*\sigma(L_0)$ is an isomorphism. □

It is customary to denote $L_1 L_0$ by Δ, the generalized Hodge Laplacian associated to the elliptic complex.

THEOREM 5.2.7. *Let X be a compact manifold and let (E, D) be an elliptic complex over X. Then the cohomology groups $H^j(X; E^\cdot)$ are finite-dimensional for each j.*

Moreover, for each j, we have the decomposition

$$L^2_s(M, E^j) = \ker(D + D^*) \oplus \operatorname{Im} D \oplus \operatorname{Im} D^*. \tag{5.2.22}$$

and hence every cohomology class in H^j has a unique harmonic representative. The index of $D + D^*$ is equal to the Euler characteristic of the complex.

THEOREM 5.2.8. *If (E, D) is an elliptic complex over a compact manifold, then the cohomology groups are finite-dimensional. The direct sum of the even cohomology groups is identifiable with the null-space of L, the direct sum of the odd cohomology groups is identifiable with its cokernel.*

REMARK 5.2.9. This result includes the important facts that the de Rham groups of a compact manifold are finite dimensional and also that the Dolbeault groups of a compact complex manifold are finite-dimensional.

5.3. Operators of Dirac type

The operator L_0 associated to an elliptic complex (first-order, as always here) has the property that $\sigma(L_0)$ is an isomorphism, hence the second-order symbol $\sigma(L_0)^*\sigma(L_0)$ is an isomorphism. For the de Rham and Dolbeault complexes, this second-order symbol turns out to be $|\xi|^2$ times the identity, where the length-squared of ξ is computed using the same metric used to define the adjoint.

Indeed, if α is a form of degree k, then

$$\sigma(d)_\xi(\alpha) = i\xi \wedge \alpha, \;\; \sigma(d^*)_\xi(\alpha) = -iiota_\xi \alpha \tag{5.3.1}$$

where in the second we really have $\xi^\sharp$ the vector associated by the metric with the cotangent vector ξ.

To figure out $\sigma(d + \mathrm{d}^*)^2$, we may suppose that $\xi = |\xi| e_1$ and consider separately the cases $\alpha = e_1 \wedge \beta$ on the one hand and that α contains no e_1-term on the other. If $\alpha = e_1 \wedge \beta$, then

$$i|\xi| e_1 \wedge \alpha = 0, \; -i|\xi|\iota_{e_1} = -i|\xi|\beta, \tag{5.3.2}$$

where β has no e_1-term. Hence for such α,

$$\sigma(d + \mathrm{d}^*)^2_\xi(\alpha) = i|\xi_1| e_1 \wedge (-i|\xi|\beta) = |\xi|^2 \alpha. \tag{5.3.3}$$

The computation is essentially the same in the case that α is a form with no e_1 term.

Different authors use slightly different terminology, but let us say that a *first-order*, (formally) self-adjoint operator

$$P : C^\infty(X;S) \longrightarrow C^\infty(X;S) \tag{5.3.4}$$

is of *Dirac type* if and only if

$$\sigma(P)^2_\xi = |\xi|^2 \otimes 1 \tag{5.3.5}$$

Here S is a hermitian vector bundle.

Example 5.3.1. Suppose $\dim X$ is odd. Consider the de Rham complex, equipped with a metric coming from a riemannian metric on X. Let W_0 and W_1 be, as before, the sum of the even and the odd form bundles over X. Although these are not literally the same, they are isomorphic by the $*$ operator. In odd dimensions, the $*$-operator maps even forms to odd forms and vice versa. So the operator

$$*\,(\mathrm{d} + \mathrm{d}^*) = *\mathrm{d} + \mathrm{d}* \tag{5.3.6}$$

is a self-adjoint operator from W_0 to itself. By the above computations, it is an operator of Dirac type.

Example 5.3.2. Suppose that $\dim X$ is even and consider the de Rham complex as in the previous example. Now W_0 and W_1 are not in general isomorphic (consider for example $X = S^2$). However, if we combine them and the first-order operators L_0 and L_1 as follows:

$$L = \begin{bmatrix} 0 & L_1 \\ L_0 & 0 \end{bmatrix}, L : C^\infty(X; W_0 \oplus W_1) \to C^\infty(X; W_0 \oplus W_1) \tag{5.3.7}$$

then L is self-adjoint and

$$L^2 = \begin{bmatrix} L_1 L_0 & 0 \\ 0 & L_0 L_1 \end{bmatrix} \tag{5.3.8}$$

so that L is of Dirac type.

Example 5.3.3. There is a similar discussion for the Dolbeault complex and $\overline{\partial} + \overline{\partial}^*$.

5.3.1. Clifford algebras. Let V be a real euclidean vector space. The Clifford algebra of V is the polynomial algebra on V subject to the relations

$$wv + vw = -2\langle w, v\rangle. \tag{5.3.9}$$

It is denoted by $\mathrm{Cl}(V)$.

Proposition 5.3.4. *Let V be a real euclidean vector space of dimension n. As a vector space, $\mathrm{Cl}(V)$ is isomorphic to the exterior algebra $\Lambda^* V$. In particular it has dimension 2^n.*

Proof. If we choose an orthonormal basis $(e_1, \dots, e_n)$ of V, then the Clifford algebra relations become

$$e_i^2 = -1, e_i e_j = -e_j e_i \text{ if } i \neq j. \tag{5.3.10}$$

The tensor algebra on V consists of linear combinations of elements of the form

$$e_\alpha = e_{\alpha_1} \cdots e_{\alpha_k} \tag{5.3.11}$$

We claim that in $\mathrm{Cl}(V)$ this is equal to an element of the form

$$\pm\, e_{\beta_1} e_{\beta_2} \cdots e_{\beta_\ell} \text{ where } \beta_1 < \beta_2 < \cdots < \beta_\ell. \tag{5.3.12}$$

To see this, suppose that $\alpha_1 < \cdots < \alpha_p$ but $\alpha_{p+1} \leqslant \alpha_p$. If $\alpha_{p+1} = \alpha_p = i$, say, then $e_i e_i$ appears and this is equal to -1. Thus in $\mathrm{Cl}(V)$, $e_\alpha = e_{\alpha'}$ where $|\alpha'| = |\alpha| - 2$. We may assume by induction that we have already dealt with such terms.

If $i = \alpha_{p+1} < \alpha_p = j$ then either $i \in \{\alpha_1, \ldots, \alpha_{p-1}\}$ or not. In the first case, we can keep moving the e_i to the left, changing signs each time, until it meets the e_i in $e_{\alpha_1} \cdots e_{\alpha_{p-1}}$ at which point it explodes to give -1. Once again we have reduced to a monomial with 2 fewer terms which we suppose we already know how to deal with. In the opposite case, we move e_i to the left, changing signs each time, until the first $p+1$ of the e_{α_r} are in the right order. This gives an inductive method to start from any 'word' in the e_i and turn it into a correctly ordered word in a finite number of steps.

The result follows since the e_α with $\alpha_1 < \cdots < \alpha_k$ form a basis for the full exterior algebra on V. □

The proposition shows that the algebra structure of $\mathrm{Cl}(V)$ gives another algebra structure on $\Lambda^* V$. It is not hard to describe this structure.

PROPOSITION 5.3.5. *When* $\mathrm{Cl}(V)$ *is identified with* $\Lambda^* V$*, we have*

$$v \cdot \alpha = v \wedge \alpha + \iota_v \alpha, \quad (v \in V, \alpha \in \Lambda^* V) \tag{5.3.13}$$

where ι_v *is interior product with the element of* V^* *dual to* v *by the metric. The product of two arbitrary elements of* $\Lambda^* V$ *is uniquely determined by this formula.*

REMARK 5.3.6. One can consider more generally $\mathrm{Cl}(V, Q)$ where Q is a (non-degenerate) quadratic form on any finite-dimensional real vector space. There is an elaborate theory of the structure of these algebras, which depends on the dimension of V and the signature of Q.

5.3.2. The spin representation. To be written