MICROLOCAL ANALYSIS AND EVOLUTION EQUATIONS: LECTURE NOTES FROM 2008 CMI/ETH SUMMER SCHOOL
November 7, 2012

JARED WUNSCH

Contents

1. Introduction

The point of these notes, and the lectures from which they came, is not to provide a rigorous and complete introduction to microlocal analysis—many good ones now exist—but rather to give a quick and impressionistic feel for how the subject is used in practice. In particular, the philosophy is to crudely axiomatize the machinery of pseudodifferential and Fourier integral operators, and then to see what problems this enables us to solve. The primary emphasis is on application of commutator methods to yield microlocal energy estimates, and on simple parametrix constructions in the framework of the calculus of Fourier integral operators; the rigorous justification of the computations is kept as much as possible inside a black box. By contrast, the author has found that lecture courses focusing on a careful development of the inner workings of this black box can (at least when he is the lecturer) too easily bog down in technicality, leaving the students with no notion of why one might suffer through such agonies. The ideal education, of course, includes both approaches. . .

A wide range of more comprehensive and careful treatments of this subject are now available. Among those that the reader might want to consult for supplementary reading are [17], [7], [22], [24], [26], [2], [28], [16] (with the last three focusing on the "semi-classical" point of view, which is not covered here). Hörmander's treatise [11], [12], [13], [14] remains the definitive reference on many aspects of the subject.

Some familiarity with the theory of distributions (or a willingness to pick it up) is a prerequisite for reading these notes, and fine treatments of this material include [11] and [6]. (Additionally, an appendix sets out the notation and most basic concepts in Fourier analysis and distribution theory.)

Much of the hard technical work in what follows has been shifted onto the reader, in the form of exercises. Doing at least some of them is essential to following the exposition. The exercises that are marked with a "star" are in general harder or longer than those without, in some cases requiring ideas not developed here.

The author has many debts to acknowledge in the preparation of these notes. The students at the CMI/ETH summer school were the ideal audience, and provided helpful suggestions on the exposition, as well as turning up numerous errors and inconsistencies in the notes (although many more surely remain). Discussions with Andrew Hassell, Michael Taylor, András Vasy, and Maciej Zworski were very valuable in the preparation of these lectures and notes. Rohan Kadakia kindly corrected a number of errrors in the final version of the manuscript.

Finally, the author wishes to gratefully acknowledge Richard Melrose, who taught him most of what he knows of this subject: a strong influence of Melrose's own excellent lecture notes [17] can surely be detected here.

2. Prequel: energy methods and commutators

This section is supposed to be like the part of an action movie before the opening credits: a few explosions and a car chase to get you in the right frame of mind, to be followed by a more careful exposition of plot.

2.1. **The Schrödinger equation on $\mathbb{R}^n$.** Let us consider a solution ψ to the *Schrödinger equation* on $\mathbb{R} \times \mathbb{R}^n$:

$$(2.1) \qquad\qquad i^{-1}\partial_t \psi - \nabla^2 \psi = 0.$$

The complex-valued "wavefunction" ψ is supposed to describe the time-evolution of a free quantum particle (in rather unphysical units). We'll use the notation $\Delta = -\nabla^2$ (note the sign: it makes the operator positive, but is a bit non-standard).

Consider, for any self-adjoint operator A, the quantity

$$\langle A\psi, \psi \rangle$$

where $\langle \cdot, \cdot \rangle$ is the sesquilinear L^2-inner product on $\mathbb{R}^n$. In the usual interpretation of QM, this is the expectation value of the "observable" A. Since $\partial_t \psi = i\nabla^2 \psi = -i\Delta\psi$, we can easily find the time-evolution of the expectation of A :

$$\partial_t \langle A\psi, \psi \rangle = \langle \partial_t(A)\psi, \psi \rangle + \langle A(-i\Delta)\psi, \psi \rangle + \langle A\psi, (-i\Delta)\psi \rangle.$$

Now, using the self-adjointness of Δ and the sesquilinearity, we may rewrite this as

$$(2.2) \qquad \partial_t \langle A\psi, \psi \rangle = \langle \partial_t(A)\psi, \psi \rangle + i\langle [\Delta, A]\psi, \psi \rangle$$

where $[S, T]$ denotes the *commutator $ST - TS$* of two operators (and $\partial_t(A)$ represents the derivative of the operator itself, which may have time-dependence). Note that this computation is a bit bogus in that it's a formal manipulation that we've done without regard to whether the quantities involved make sense, or whether the formal integration by parts (i.e. the use of the self-adjointness of Δ) was justified. For now, let's just keep in mind that this makes sense for sufficiently "nice" solutions, and postpone the technicalities.

If you want to learn things about $\psi(t, x)$, you might try to use (2.2) with a judicious choice of A. For instance, setting $A = \text{Id}$ shows that the L^2-norm of $\psi(t, \cdot)$ is conserved. Additionally, choosing $A = \Delta^k$ shows that the H^k norm is conserved (see the appendix for a definition of this norm). In both these examples, we are using the fact that $[\Delta, A] = 0$.

A more interesting example might be the following: set $A = \partial_r$, the radial derivative. We may write the Laplace operator on $\mathbb{R}^n$ in polar coordinates as
$$\Delta = -\partial_r^2 - \frac{n-1}{r}\partial_r + \frac{\Delta_\theta}{r^2}$$
where Δ_θ is the Laplacian on S^{n-1}; thus we compute
$$[\Delta, \partial_r] = 2\frac{\Delta_\theta}{r^3} - \frac{(n-1)}{r^2}\partial_r.$$

Exercise 2.1. Do this computation! (Be aware that ∂_r is not a differential operator with smooth coefficients.)

This is kind of a funny looking operator. Note that Δ is self-adjoint, and ∂_r wants to be anti-self-adjoint, but isn't quite. In fact, it makes more sense to replace ∂_r by

$$A = (1/2)(\partial_r - \partial_r^*) = \partial_r + \frac{n-1}{2r},$$

which corrects ∂_r by a lower-order term to be anti-self-adjoint.

Exercise 2.2. Show that

$$\partial_r^* = -\partial_r - \frac{n-1}{r}.$$

Trying again, we get by dint of a little work:

$$(2.3) \qquad [\Delta, \partial_r + \frac{n-1}{2r}] = \frac{2\Delta_\theta}{r^3} + \frac{(n-1)(n-3)}{2r^3},$$

provided n, the dimension, is at least 4.

Exercise 2.3. Derive (2.3), where you should think of both sides as operators from Schwartz functions to tempered distributions (see the appendix for definitions). What happens if $n = 3$? If $n = 2$? Be very careful about differentiating negative powers of r in the context of distribution theory...

Why do we like (2.3)? Well, it has the very lovely feature that both summands on the RHS are *positive operators*. Let's plug this into (2.2)

and integrate on a finite time interval:

$$
i^{-1}\langle A\psi, \psi\rangle\big|_0^T = \int_0^T \left\langle \frac{2\Delta_\theta}{r^3}\psi, \psi\right\rangle + \left\langle \frac{(n-1)(n-3)}{2r^3}\psi, \psi\right\rangle dt
$$
$$
= \int_0^T 2\left\|r^{-1/2}\slashed{\nabla}\psi\right\|^2 dt + \frac{(n-1)(n-3)}{2}\left\|r^{-3/2}\psi\right\|^2 dt,
$$

where $\slashed{\nabla}$ represents the (correctly scaled) angular gradient: $\slashed{\nabla} = r^{-1}\nabla_\theta$, where ∇_θ denotes the gradient on S^{n-1}.

Now, we're going to turn the way we use this estimate on its head, relative to what we did with conservation of L^2 and H^k norms: the *left*-hand-side can be estimated by a constant times the $H^{1/2}$ norm of the initial data. This should be at least plausible for the derivative term, since morally, half a derivative can be dumped on each copy of u, but is complicated by the fact that ∂_r is not a differential operator on $\mathbb{R}^n$ with smooth coefficients. The following (somewhat lengthy) pair of exercises goes somewhat far afield from the main thrust of these notes, but is necessary to justify our $H^{1/2}$ estimate.

In the sequel, we employ the useful notation $f \lesssim g$ to indicate that $f \le Cg$ for some $C \in \mathbb{R}^+$; when f and g are Banach norms of some function, C is always supposed to be independent of the function.

Exercise* 2.4.
   (1) Verify that for $u \in \mathcal{S}(\mathbb{R}^n)$ with $n \ge 3$, $|\langle \partial_r u, u\rangle| \lesssim \|u\|_{H^{1/2}}^2$.
       HINT: Use the fact that
       $$
       \partial_r = \sum |x|^{-1}x^j \partial_{x^j}.
       $$

       Check that $x/|x|$ is a bounded multiplier on both L^2 and H^1, and hence, by interpolation and duality, on $H^{-1/2}$. An efficient treatment of the interpolation methods you will need can be found in [25]. You will probably also need to use *Hardy's inequality* (see Exercise 2.5).
   (2) Likewise, show that the $\langle r^{-1}u, u\rangle$ term is bounded by a multiple of $\|u\|_{H^{1/2}}^2$ (again, use Exercise 2.5).

Exercise 2.5. Prove *Hardy's inequality*: if $u \in H^1(\mathbb{R}^n)$ with $n \ge 3$, then

$$
\frac{(n-2)^2}{4} \int \frac{|u|^2}{r^2}\, dx \le \int |\nabla u|^2\, dx.
$$

HINT: In polar coordinates, we have for $u \in \mathcal{S}(\mathbb{R}^n)$

$$
\int \frac{|u|^2}{r^2}\, dx = \int_{S^{n-1}} \int_0^\infty |u|^2 r^{n-3}\, dr\, d\theta.
$$

Integrate by parts in the r integral, and apply Cauchy-Schwarz.

So we obtain, finally, the *Morawetz inequality*: if $\psi_0 \in H^{1/2}(\mathbb{R}^n)$, with $n \geq 4$ then

(2.4)
$$2 \int_0^T \left\| r^{-1/2} \overline{\nabla} \psi \right\|^2 dt + \frac{(n-1)(n-3)}{2} \int_0^T \left\| r^{-3/2} \psi \right\|^2 dt \lesssim \|\psi_0\|_{H^{1/2}}^2.$$

Now remember that we've been working rather formally, and there's no guarantee that either of the terms on the LHS is finite a priori. But the RHS is finite, so *since both terms on the LHS are positive, both must be finite, provided $\psi_0 \in H^{1/2}$*. (This is a dangerously sloppy way of reasoning—see the exercises below.) So we get, at one stroke two nice pieces of information: if $\psi_0 \in H^{1/2}$, we obtain the finiteness of both terms on the left.

Let's try and understand these. The term

$$\int_0^T \left\| r^{-3/2} \psi \right\|^2 dt$$

gives us a weighted estimate, which we can write as

(2.5) $\psi \in r^{3/2} L^2([0, T]; L^2(\mathbb{R}^n))$

for any T, or, more briefly, as

(2.6) $\psi \in r^{3/2} L^2_{\text{loc}} L^2.$

(The right side of (2.5) denotes the Hilbert space of functions that are of the form $r^{3/2}$ times an element of the space of L^2 functions on $[0, T]$ with values in the Hilbert space $L^2(\mathbb{R}^n)$; note that whenever we use the condensed notation (2.6), the Hilbert space for the time variables will precede that for the spatial variables.) So ψ can't "bunch up" too much at the origin. Incidentally, our whole setup was translation invariant, so in fact we can conclude

$$\psi \in |x - x_0|^{3/2} L^2_{\text{loc}} L^2$$

for *any* $x_0 \in \mathbb{R}^n$, and ψ can't bunch up too much anywhere at all.

How about the other term? One interesting thing we can do is the following: Choose x_0, x_1 in $\mathbb{R}^n$, and let X be a smooth vector field with support disjoint from the line $\overline{x_0 x_1}$. Then we may write X in the form

$$\mathsf{X} = \mathsf{X}_0 + \mathsf{X}_1$$

with X_i smooth, and $\mathsf{X}_i \perp (x - x_i)$ for $i = 0, 1$; in other words, we split X into angular vector fields with respect to the origin of coordinates placed at x_0 and x_1 respectively. Moreover, we can arrange that the coefficients of X_i be bounded in terms of the coefficients of X (provided we

bound the support uniformly away from $\overline{x_0 x_1}$). Thus, we can estimate for any such vector field X and any $u \in \mathcal{C}_c^\infty(\mathbb{R}^n)$

$$\int |\mathsf{X}u|^2 \, dx \lesssim \int \left||x - x_0|^{-1/2} \not\nabla_0 u\right|^2 dx + \int \left||x - x_1|^{-1/2} \not\nabla_1 u\right|^2 dx$$

where $\not\nabla_i$ is the angular gradient with respect to the origin of coordinates at x_i. Since for a solution of the Schrödinger equation, (2.4) tells us that the time integral of each of these latter terms is bounded by the squared $H^{1/2}$ norm of the initial data, we can assemble these estimates with the choices $\mathsf{X} = \chi \partial_{x^j}$ for any $\chi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ to obtain

$$\int_0^T \|\chi \nabla \psi\|^2 \, dt \lesssim \|\psi_0\|_{H^{1/2}}^2.$$

In more compact notation, we have shown that

$$\psi_0 \in H^{1/2} \implies \psi \in L^2_{\text{loc}} H^1_{\text{loc}}.$$

This is called the *local smoothing estimate.* It says that on average in time, the solution is locally half a derivative smoother than the initial data was; one consequence is that in fact, with initial data in $H^{1/2}$, the solution is in H^1 in space at *almost every time.*

Exercise 2.6. Work out the Morawetz estimate in dimension 3. (This is in many ways the nicest case.) Note that our techniques yield no estimate in dimension 2, however.

In fact, if all we care about is the local smoothing estimate (and this is frequently the case) there is an easier commutator argument that we can employ to get just that estimate. Let $f(r)$ be a function on $\mathbb{R}^+$ that equals 0 for $r < 1$, is increasing, and equals 1 for $r \geq 2$. Set $A = f(r)\partial_r$ and employ (2.2) just as we did before. The commutant $f(r)\partial_r$ (as opposed to just ∂_r) has the virtue of actually being a smooth vector field on $\mathbb{R}^n$. So we can write

$$[\Delta, f(r)\partial_r] = -2f'(r)\partial_r^2 + 2r^{-3}f(r)\Delta_\theta + R$$

where R is a first order operator with coefficients in $\mathcal{C}_c^\infty(\mathbb{R}^n)$. As we didn't bother to make our commutant anti-self-adjoint, we might like to fix things up now by rewriting

$$[\Delta, f(r)\partial_r] = -2\partial_r^* f'(r)\partial_r + 2r^{-3}f(r)\Delta_\theta + R'$$

where R' is of the same type as R. Note that both main terms on the right are now nonnegative operators, and also that the term containing

∂_r^* is not, appearances to the contrary, singular at the origin, owing to
the vanishing of f' there. Thus we obtain, by another use of (2.2),

$$(2.7) \quad \int_0^T \left\| \sqrt{f'(r)}\partial_r\psi \right\|^2 dt + \int_0^T \left\| \sqrt{f(r)}r^{-1/2}\nabla\!\!\!/\,\psi \right\|^2 dt$$

$$\lesssim \int_0^T |\langle R'\psi, \psi\rangle|\, dt + |\langle f(r)\partial_r\psi, \psi\rangle|\big|_0^T.$$

Now the first term on the RHS is bounded by a multiple of $\|\psi_0\|_{H^{1/2}}^2$ (as
R' is first order with coefficients in $\mathcal{C}_c^\infty(\mathbb{R}^n)$); the second term is likewise
(since f is bounded with compactly supported derivative, and zero near
the origin). This gives us an estimate of the desired form, valid on any
compact subset of supp $f \cap$ supp f', which can be translated to contain
any point.

Exercise 2.7. This exercise is on giving some rigorous underpinnings
to some of the formal estimates above. It also gets you thinking about
the alternative, Fourier-theoretic, picture of how might think about
solutions to the Schrödinger equation.[1]

  (1) Using the Fourier transform,[2] show that if $\psi_0 \in L^2(\mathbb{R}^n)$, there
      exists a unique solution $\psi(t,x)$ to (2.1) with $\psi(0,x) = \psi_0$.
  (2) As long as you're at it, use the Fourier transform to derive the
      explicit form of the solution: show that

$$\psi(t,x) = \psi_0 * K_t$$

      where K_t is the "Schrödinger kernel;" give an explicit formula
      for K_t.
  (3) Use your explicit formula for K_t to show that if $\psi_0 \in L^1$ then
      $\psi(T,x) \in L^\infty(\mathbb{R}^n)$ for any $T \neq 0$.
  (4) Show using the first part, i.e. by thinking about the solution
      operator as a Fourier multiplier, that if $\psi_0 \in H^s$ then $\psi(t,x) \in$
      $L^\infty(\mathbb{R}_t; H^s)$, hence give another proof that H^s regularity is con-
      served.
  (5) Likewise, show that the Schrödinger evolution in $\mathbb{R}^n$ takes Schwartz
      functions to Schwartz functions.
  (6) *Rigorously* justify the Morawetz inequality if $\psi_0 \in \mathcal{S}(\mathbb{R}^n)$. Then
      use a density argument to rigorously justify it for $\psi_0 \in H^{1/2}(\mathbb{R}^n)$.

---

[1]If you want to work hard, you might try to derive the local smoothing estimate
from the explicit form of the Schrödinger kernel derived below. It's not so easy!

[2]See the appendix for a very brief review of the Fourier transform acting on
tempered distributions and L^2-based Sobolev spaces.

2.2. **The Schrödinger equation with a metric.** Now let us change our problem a bit. Say we are on an n-dimensional manifold, or even just on $\mathbb{R}^n$ endowed with a complete non-Euclidean Riemannian metric g. There is a canonical choice for the Laplace operator in this setting:

$$\Delta = d^* d$$

where d takes functions to one-forms, and the adjoint is with respect to L^2 inner products on both (which of course also involve the volume form associated to the Riemannian metric). This yields, in coordinates,

$$(2.8) \qquad \Delta = -\frac{1}{\sqrt{g}} \partial_{x^i} g^{ij} \sqrt{g} \partial_{x^j},$$

where $\sum_{i,j=1}^n g^{ij} \partial_{x^i} \otimes \partial_{x^j}$ is the dual metric on forms (hence g^{ij} is the inverse matrix to g_{ij}) and g denotes $\det(g_{ij})$.

Exercise 2.8. Check this computation!

Exercise 2.9. Write the Euclidean metric on $\mathbb{R}^3$ in spherical coordinates, and use (2.8) to compute the Laplacian in spherical coordinates.

We can now consider the Schrödinger equation with the Euclidean Laplacian replaced by this new "Laplace-Beltrami" operator. By standard results in the spectral theory of self-adjoint operators,[3] there is still a solution in $L^\infty(\mathbb{R}; L^2)$ given any L^2 initial data—this generalizes our Fourier transform computation in Exercise 2.7—but its form and its properties are much harder to read off.

Computing commutators with this operator is a little trickier than in the Euclidean case, but certainly feasible; you might certainly try computing $[\Delta, \partial_r + (n-1)/(2r)]$ where r is the distance from some fixed point.

Exercise 2.10. Write out the Laplace operator in Riemannian polar coordinates, and compute $[\Delta, \partial_r + (n-1)/(2r)]$ near $r = 0$.

But what happens when we get beyond the injectivity radius? Of course, the r variable doesn't make any sense any more. Moreover, if we try to think of ∂_r as the operator of differentiating "along geodesics emanating from the origin" then at a conjugate point to 0, we have the problem that we're somehow supposed to be be simultaneously differentiating in two different directions. One fix for this problem is to employ the calculus of pseudodifferential operators, which permits us to construct operators that behave differently depending on what

---

[3]The operator Δ is manifestly formally self-adjoint, but in fact turns out to be essentially self-adjoint on $\mathcal{C}_c^\infty(X)$ for X any complete manifold.

direction we're looking in: we can make operators that separate out the different geodesics passing through the conjugate point, and do different things along them.

2.3. The wave equation. Let

$$\Box u \equiv (\partial_t^2 + \Delta)u = 0$$

denote the wave equation on $\mathbb{R} \times \mathbb{R}^n$ (recall that $\Delta = -\sum \partial_{x^i}^2$). For simplicity of notation, *let us consider only real-valued solutions in this section.*

The usual route to thinking about the energy of a solution to the wave equation is as follows. We consider the integral

$$(2.9) \qquad\qquad 0 = \int_0^T \langle \Box u, \partial_t u \rangle \, dt$$

where $\langle \cdot, \cdot \rangle$ is the inner product on $L^2(\mathbb{R}^n)$. Then integrating by parts in t and in x gives the conservation of

$$\|\partial_t u\|^2 + \|\nabla u\|^2.$$

We can recast this formally as a commutator argument, if we like, by considering the commutator with the indicator function of an interval:

$$0 = \int_{\mathbb{R}} \langle [\Box, 1_{[0,T]}(t)\partial_t]u, u \rangle \, dt.$$

The integral vanishes, at least formally, by self-adjointness of $\Box$—it is in fact a better idea to think of this whole thing as an inner product on $\mathbb{R}^{n+1}$:

$$\langle [\Box, 1_{[0,T]}(t)]\partial_t u, u \rangle_{\mathbb{R}^{n+1}}.$$

Having gone this far, we might like to replace the indicator function with something smooth, to give a better justification for this formal integration by parts; let $\chi(t)$ be a smooth approximator to the indicator function with $\chi' = \phi_1 - \phi_2$ with ϕ_1 and ϕ_2 nonnegative bump functions supported respectively in $(-\epsilon, \epsilon)$ and $(T-\epsilon, T+\epsilon)$, with $\phi_2(\cdot) = \phi_1(\cdot - T)$ Let $A = \chi(t)\partial_t + \partial_t \chi(t)$. Then we have

$$[\Box, A] = 2\partial_t \chi' \partial_t + \partial_t^2 \chi' + \chi' \partial_t^2,$$

and by (formal) anti-self-adjointness of ∂_t (and the fact that u is assumed real),

$$
\begin{aligned}
0 = \langle [\Box, A]u, u \rangle_{\mathbb{R}^{n+1}} &= -2\langle \chi' \partial_t u, \partial_t u \rangle_{\mathbb{R}^{n+1}} + 2\langle \chi' u, \partial_t^2 u \rangle_{\mathbb{R}^{n+1}} \\
&= -2\langle \chi' \partial_t u, \partial_t u \rangle_{\mathbb{R}^{n+1}} + 2\langle \chi' u, \nabla^2 u \rangle_{\mathbb{R}^{n+1}} \\
&= -2\langle \chi' \partial_t u, \partial_t u \rangle_{\mathbb{R}^{n+1}} - 2\langle \chi' \nabla, \nabla u \rangle_{\mathbb{R}^{n+1}} \\
&= -2 \int_{\mathbb{R}^{n+1}} \phi_1(t)(|u_t|^2 + |\nabla u|^2)\, dt\, dx \\
&\quad + 2 \int_{\mathbb{R}^{n+1}} \phi_2(t)(|u_t|^2 + |\nabla u|^2)\, dt\, dx.
\end{aligned}
$$

Thus, the energy on the time interval $[T - \epsilon, T + \epsilon]$ (modulated by the cutoff ϕ_2) is the same as that in the time interval $[-\epsilon, \epsilon]$ (modulated by ϕ_1).

We can get fancier, of course. Finite propagation speed is usually proved by considering the variant of (2.9)

$$
\int_{-T_2}^{-T_1} \int_{|x|^2 \le t^2} \Box u\, \partial_t u\, dx\, dt,
$$

with $0 < T_1 < T_2$. Integrating by parts gives negative boundary terms, and we find that the energy in

$$
\{t = -T_1, |x|^2 \le T_1^2\}
$$

is bounded by that in

$$
\{t = -T_2, |x|^2 \le T_2^2\}.
$$

Hence if the solution has zero Cauchy data (i.e. value, time-derivative) on the latter surface, it also has zero Cauchy data on the former.

Exercise 2.11. Go through this argument to show finite propagation speed.

Making this argument into a commutator argument is messier, but still possible:

Exercise* 2.12. Write a positive commutator version of the proof of finite propagation speed, using smooth cutoffs instead of integrations by parts. (An account of energy estimates with smooth temporal cutoffs, in the general setting of Lorentzian manifolds, can be found in [27, Section 3].)

There is of course also a Morawetz estimate for the wave equation! (Indeed, this was what Morawetz originally proved.)

Exercise* 2.13. Derive (part of) the Morawetz estimate: Let u solve

$$\Box u = 0, (u, \partial_t u)|_{t=0} = (f, g)$$

on $\mathbb{R}^n$, with $n \geq 4$. Show that

$$\left\| r^{-3/2} u \right\|_{L^2_{\text{loc}}(\mathbb{R}^{n+1})} \lesssim \|f\|^2_{H^1} + \|g\|^2_{L^2};$$

this is analogous to the weight part of the Morawetz estimate we derived for the Schrödinger equation. There is in fact no need for the local L^2 norm—the global spacetime estimate works too: prove this estimate, and use it to draw a conclusion about the long-time decay of a solution to the wave equation with Cauchy data in $\mathcal{C}^\infty_c(\mathbb{R}^n) \oplus \mathcal{C}^\infty_c(\mathbb{R}^n)$.

HINT: consider $\langle [\Box, \chi(t)(\partial_r + (n-1)/(2r))] u, u \rangle_{\mathbb{R}^{n+1}}$.

3. THE PSEUDODIFFERENTIAL CALCULUS

Recall that we hoped to describe a class of operators enriching the differential operators that would, among other things, enable us to deal properly with the local smoothing estimate on manifolds, where conjugate points caused our commutator arguments with ordinary differential operators to break down. One solution to this problem turns out to lie in the calculus of pseudodifferential operators.

3.1. **Differential operators.** What kind of a creature is a pseudodifferential operator? Well, first let's think more seriously about *differential* operators. A linear differential operator of order m is something of the form

$$(3.1) \qquad\qquad P = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha$$

where $D_j = i^{-1}(\partial/\partial x^j)$ and we employ "multiindex notation:"

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

$$|\alpha| = \sum \alpha_j.$$

We will always take our coefficients to be smooth:

$$a_\alpha \in \mathcal{C}^\infty(\mathbb{R}^n).$$

We let

$$\text{Diff}^m(\mathbb{R}^n)$$

denote the collection of all differential operators of order m on $\mathbb{R}^n$ (and will later employ the analogous notation on a manifold).

If $P \in \text{Diff}^m(\mathbb{R}^n)$ is given by (3.1), we can associate with P a function by formally turning differentiation in x^j into a formal variable ξ_j with $(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$:

$$p(x, \xi) = \sum a_\alpha(x)\xi^\alpha.$$

This is called the "total (left-) symbol" of P; of course, knowing p is equivalent to knowing P. Note that $p(x, \xi)$ is a rather special kind of a function on $\mathbb{R}^{2n}$: it is actually polynomial in the ξ variables with smooth coefficients. Let us write $p = \sigma_{\text{tot}}(P)$.

Note that

$$\sigma_{\text{tot}} : P \mapsto p$$

is *not* a ring homomorpism: we have

$$PQ = \sum_{\alpha, \beta} p_\alpha(x)D^\alpha q_\beta(x)D^\beta,$$

and if we expand out this product to be of the form

$$\sum_\gamma c_\gamma(x)D^\gamma,$$

then the coefficients c_γ will involve all kinds of derivatives of the q_β's. This is a pain, but on the other hand life would be pretty boring if the ring of differential operators were commutative.

If we make do with less, though, composition of operators doesn't look so bad. We let $\sigma_m(P)$, the *principal symbol* of P, just be the symbol of the top-order parts of P :

$$\sigma_m(P) = \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha.$$

Note that $\sigma_m(P)$ is a *homogeneous* degree-m polynomial in ξ, i.e., a polynomial such that $\sigma_m(P)(x, \lambda\xi) = \lambda^m \sigma_m(P)(x, \xi)$ for $\lambda \in \mathbb{R}$. As a result, we can reconstruct it from its value at $|\xi| = 1$, and it makes sense for many purposes to just consider it as a (rather special) smooth function on $\mathbb{R}^n \times S^{n-1}$. It turns out to make more invariant sense to regard the principal symbol as a homogeneous polynomial on $T^*\mathbb{R}^n$, so that once we have scaled away the action of $\mathbb{R}^+$, we may regard it as a function on $S^*\mathbb{R}^n$, the *unit* cotangent bundle of $\mathbb{R}^n$, which is simply defined as $T^*\mathbb{R}^n/\mathbb{R}^+$ (or identified with the bundle of unit covectors in, say, the Euclidean metric). To clarify when we are talking about the symbol on $S^*\mathbb{R}^n$, we define[4]

$$\hat{\sigma}_m(P) = \sigma_m(P)|_{|\xi|=1} \in \mathcal{C}^\infty(S^*\mathbb{R}^n).$$

[4]The reader is warned that this notation is not a standard one.

Now it *is* the case that the principal symbol is a homomorphism:

Proposition 3.1. *For P, Q differential operators of order m resp. m',*

$$\sigma_{m+m'}(PQ) = \sigma_m(P)\sigma_{m'}(Q).$$

(and likewise with $\hat{\sigma}$).

Exercise 3.1. Verify this!

Moreover, the principal symbol has another lovely property that the total symbol lacks: it behaves well under change of variables. If $y = \phi(x)$ is a change of variables, with ϕ a diffeomorphism, and if P is a differential operator in the x variables, we can of course define a pushforward of P by

$$(\phi_* P)f = P(\phi^* f)$$

Then in particular,

$$\phi_*(D_{x^j}) = \sum_k \frac{\partial y^k}{\partial x^j} D_{y^k},$$

hence

$$\phi_*(D_x^\alpha) = D_{x^1}^{\alpha_1} \dots D_{x^n}^{\alpha_n} = \left(\sum_{k_1=1}^n \frac{\partial y^{k_1}}{\partial x^1} D_{y^{k_1}} \right)^{\alpha_1} \dots \left(\sum_{k_n=1}^n \frac{\partial y^{k_n}}{\partial x^n} D_{y^{k_n}} \right)^{\alpha_n};$$

when we again try to write this in our usual form, as a sum of coefficients times derivatives, we end up with a hideous mess involving high derivatives of the diffeomorphism ϕ. *But,* if we restrict ourselves to dealing with principal symbols alone, the expression simplifies in both form and (especially) interpretation:

Proposition 3.2. *If P is a differential operator given by (3.1), and $y = \phi(x)$, then*

$$\sigma_m(\phi_* P)(y, \eta) = \sum_{|\alpha|=m} a_\alpha(\phi^{-1}(y)) \left(\sum_{k_1=1}^n \frac{\partial y^{k_1}}{\partial x^1} \eta_{k_1} \right)^{\alpha_1} \dots \left(\sum_{k_n=1}^n \frac{\partial y^{k_n}}{\partial x^n} \eta_{k_n} \right)^{\alpha_n}$$

where η are the new variables "dual" to the y variables.

This corresponds exactly to the behavior of a function defined on the cotangent bundle: if ϕ is a diffeomorphism from $\mathbb{R}^n_x$ to $\mathbb{R}^n_y$, then it induces a map $\Phi = \phi^* : T^*\mathbb{R}^n_y \to T^*\mathbb{R}^n_x$, and

$$\sigma_m(\phi_* P) = \Phi^*(\sigma_m(P)).$$

Exercise 3.2. Prove the proposition, and verify this interpretation of it.

Notwithstanding its poor properties, it is nonetheless a useful fact that the map

$$\sigma_{\text{tot}} : P \mapsto p$$

is one-to-one and onto polynomials with smooth coefficients; it therefore has an inverse, which we shall denote

$$\text{Op}_\ell : p \mapsto P,$$

taking functions on $T^*\mathbb{R}^n$ that happen to be polynomial in the fiber variables to differential operators on $\mathbb{R}^n$. Op_ℓ is called a "quantization" map.[5] You may wonder about the ℓ in the subscript: it stands for "left," and has to do with the fact that we chose to write differential operators in the form (3.1) instead of as

$$P = \sum_{|\alpha| \le m} D^\alpha a_\alpha(x),$$

with the coefficients on the right. This would have changed the definition of σ_{tot} and hence of its inverse.

Note that $\text{Op}_\ell(x^j) = x^j$ (i.e. the operation of multiplication by x^j) while $\text{Op}_\ell(\xi_j) = D_j$.

Why not, you might ask, try to extend this quantization map to a more general class of functions on $T^*\mathbb{R}^n$? This is indeed how we obtain the calculus of pseudodifferential operators. The tricky point to keep in mind, however, is that for most purposes, it is asking too much to deal with the quantizations of all possible functions on $T^*\mathbb{R}^n$, so we'll deal only with a class of functions that are somewhat akin to polynomials in the fiber variables.

3.2. **Quantum mechanics.** One reason why you might care about the existence of a quantization map, and give it such a suggestive name, lies in the foundations of quantum mechanics.

It is helpful to think about $T^*\mathbb{R}^n$ as being a classical *phase space,* with the x variables (in the base) being "position" and the ξ variables (the fiber variables) as "momenta" in the various directions. The general notion of *classical mechanics* (in its Hamiltonian formulation) is as follows: The state of a particle is a point in the phase space $T^*\mathbb{R}^n$, and moves along some curve in $T^*\mathbb{R}^n$ as time evolves; an *observable* $p(x, \xi)$ is a function on the phase space that we may evaluate at the state (x, ξ) of our particle to give a number (the observation). By contrast, a *quantum* particle is described by a complex-valued *function* $\psi(x)$ on $\mathbb{R}^n$, and a quantum observable is a self-adjoint *operator* P acting on functions on $\mathbb{R}^n$. Doing the same measurement repeatedly on

---

[5]It is far from unique, as will become readily apparent.

identically prepared quantum states is not guaranteed to produce the
same number each time, but at least we can talk about the *expected
value* of the observation, and it's simply

$$\langle P\psi, \psi \rangle_{L^2(\mathbb{R}^n)}.$$

In the early development of quantum mechanics, physicists sought a
way to transform the classical world into the quantum world, i.e. of
taking functions on $T^*\mathbb{R}^n$ to operators on[6] $L^2(\mathbb{R}^n)$. This is, loosely
speaking, the process of "quantization."

We now turn to the question of describing the dynamics in the quan-
tum and classical worlds. To describe how the point in phase space
corresponding to a classical particle in Hamiltonian mechanics evolves
in time, we use the notion of the "Poisson bracket" of two observables.
In coordinates, we can explicitly define

$$\{f, g\} \equiv \sum \frac{\partial f}{\partial \xi_j} \frac{\partial g}{\partial x^j} - \frac{\partial f}{\partial x^j} \frac{\partial g}{\partial \xi_j}$$

(this in fact makes invariant sense on any symplectic manifold). The
map $g \mapsto \{f, g\}$ defines a vector field[7] (the Hamilton vector field)
associated to f :

$$\mathsf{H}_f = \sum \frac{\partial f}{\partial \xi_j} \frac{\partial}{\partial x^j} - \frac{\partial f}{\partial x^j} \frac{\partial}{\partial \xi_j}$$

The classical time-evolution is along the flow generated by the Hamilton
vector field associated to the *energy function* of our system, i.e. the flow
along H_h for some given $h \in \mathcal{C}^\infty(T^*\mathbb{R}^n)$. By contrast, the wavefunction
for a quantum particle evolves in time according to the Schrödinger
equation (2.1), with $-\nabla^2$ in general replaced by a self-adjoint "Hamil-
tonian operator" H whose principal symbol is the energy function h.[8]
By a mild generalization of (2.2), the time derivative of the expectation
of an observable A is related to the commutator

$$[H, A].$$

One of the essential features of quantum mechanics is that

$$\sigma_{m+m'}([H, A]) = i\{\sigma_m(H), \sigma_{m'}(A)\},$$

---

[6]Well, they are not necessarily going to be defined on all of L^2; the techni-
cal subtleties of unbounded self-adjoint operators will mostly not concern us here,
however.

[7]We use the geometers' convention of identifying a vector and the directional
derivative along it.

[8]For honest physical applications, one really ought to introduce the semi-classical
point of view here, carrying Planck's constant along as a small parameter and using
an associated notion of principal symbol.

so that the time-evolution of the quantum observable A is related to the classical evolution of its symbol along the Hamilton flow; this is the "correspondence principle" between classical and quantum mechanics.[9]

3.3. **Quantization.** How might we construct a quantization map extending the usual quantization on fiber-polynomials?

Let $\mathcal{F}$ denote the Fourier transform (see Appendix for details). Then we may write, on $\mathbb{R}^n$,

$$(D_{x^j}\psi)(x) = \mathcal{F}^{-1}\xi_j \mathcal{F} u = (2\pi)^{-n} \int e^{ix\cdot\xi}\xi_j \int e^{-iy\cdot\xi}\psi(y)\, dy\, d\xi$$

$$= \frac{1}{2\pi}\iint \xi_j e^{i(x-y)\cdot\xi}\psi(y)\, dy\, d\xi$$

Likewise, since $\mathcal{F}^{-1}\mathcal{F} = I$, we of course have

$$(x^j\psi)(x) = (2\pi)^{-n}\iint x^j e^{i(x-y)\cdot\xi}\psi(y)\, dy\, d\xi$$

Going a bit further, we see that at least for a fiber polynomial $a(x,\xi) = \sum a_\alpha(x)\xi^\alpha$ we have
(3.2)
$$(\mathrm{Op}_\ell(a)\psi)(x) = \sum a_\alpha(x)D^\alpha\psi(x) = (2\pi)^{-n}\iint a(x,\xi)e^{i(x-y)\cdot\xi}\psi(y)\, dy\, d\xi;$$

stripping away the function ψ, we can also simply write the Schwartz kernel (see Appendix) of the operator $\mathrm{Op}_\ell(a)$ as

$$\kappa\big(\mathrm{Op}_\ell(a)\big) = (2\pi)^{-n}\int a(x,\xi)e^{i(x-y)\cdot\xi}\, d\xi.$$

(Making sense of the integrals written above is not entirely trivial: Given $\psi \in \mathcal{S}(\mathbb{R}^n)$, we can make sense of the ξ integral in (3.2), which looks (potentially) divergent, by observing that

$$(1+|\xi|^2)^{-k}(1+\Delta_y)^k e^{i(x-y)\cdot\xi} = e^{i(x-y)\cdot\xi}$$

for all $k \in \mathbb{N}$; repeatedly integrating by parts in y then moves the derivatives onto ψ. This method brings down an arbitrary negative power of $(1+|\xi|^2)$ at the cost of differentiating ψ, thus making the ξ integral convergent.[10] Similar arguments yield continuity of $\mathrm{Op}_\ell(a)$

---

[9]In the semi-classical setting, the correspondence principle tells that we can in a sense recover CM from QM in the limit when Planck's constant tends to zero. What we have in this setting is a correspondence principle that works at high energies, i.e. in doing computations with high-frequency waves.

[10]This kind of integration by parts argument is ubiquitous in the subject, and somewhat scanted in these notes, relative to its true importance.

as a map $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$, hence we can extend to let $\mathrm{Op}_\ell(a)$ act on $\psi \in \mathcal{S}'$ by duality. For more details, cf. [17].)

Exercise* 3.3. Verify the vague assertions in the parenthetical remark above. You may wish to consult, for example, the beginning of [10].

This of course suggests that we use (3.2) as the *definition* of $\mathrm{Op}_\ell(a)$ for more general observables ("symbols") a. And we do. In $\mathbb{R}^n$, we set

$$(3.3) \qquad (\mathrm{Op}_\ell(a)\psi)(x) = \frac{1}{(2\pi)^n} \int a(x,\xi) e^{i(x-y)\cdot\xi} \psi(y)\, dy\, d\xi.$$

We can define the pseudodifferential operators on $\mathbb{R}^n$ to be just the range of this quantization map on some reasonable set of symbols a, to be discussed below.

On a Riemannian manifold, we can make similar constructions global by cutting off near the diagonal and using the exponential map and its inverse. The pseudodifferential operators are those whose Schwartz kernels[11] near the diagonal look like (3.3) in local coordinates, and that away from the diagonal are allowed to be arbitrary functions in $\mathcal{C}^\infty(X \times X)$. If the manifold is noncompact, we will often assume further that operators are *properly supported*, i.e. that both left- and right-projection give proper maps from the support of the Schwartz kernel to X.

3.4. The pseudodifferential calculus.

Definition 3.3. A function a on $T^*\mathbb{R}^n$ is a *classical symbol* of order m if

- $a \in \mathcal{C}^\infty(T^*\mathbb{R}^n)$
- On $|\xi| > 1$, we have
$$a(x,\xi) = |\xi|^m \tilde{a}(x, \hat{\xi}, |\xi|^{-1}),$$
  where $\tilde{a}$ is a smooth function on $\mathbb{R}_x^n \times S_{\hat\xi}^{n-1} \times \mathbb{R}^+$, and
$$\hat{\xi} = \frac{\xi}{|\xi|} \in S^{n-1}.$$

We then write $a \in S_{\mathrm{cl}}^m(T^*\mathbb{R}^n)$.

It is convenient to introduce the notation
$$\langle \xi \rangle = (1 + |\xi|^2)^{1/2},$$
so that $\langle \xi \rangle$ behaves like $|\xi|$ near infinity, but is smooth and nonvanishing at 0. A fancy way of saying that a is a classical symbol of order

---

[11]For some remarks on the Schwartz kernel theorem, see the Appendix.

m is thus to simply say that a is equal to $\langle\xi\rangle^m$ times *a smooth function on the fiberwise radial compactification of* $T^*\mathbb{R}^n$, denoted $\overline{T}^*\mathbb{R}^n$. This compactification is defined as follows: We can diffeomorphically identify $\mathbb{R}^n_\xi$ with the interior of the unit ball by first mapping it to the upper hemisphere of $S^n \subset \mathbb{R}^{n+1}$ by mapping

$$(3.4) \qquad\qquad \xi \mapsto \left(\frac{\xi}{\langle\xi\rangle}, \frac{1}{\langle\xi\rangle}\right)$$

and identifying this latter space with the interior of the ball. Then $1/\langle\xi\rangle$ becomes a *boundary defining function,* i.e. one that cuts out the boundary nondegenerately as its zero-set; $1/|\xi|$ is also a valid boundary defining function near the boundary of the ball, i.e. away from its singularity.

A very important consequence is that we can write a Taylor series for a near $|\xi|^{-1} = 0$ (the "sphere at infinity") to obtain

$$a(x,\xi) \sim \sum_{j=0}^{\infty} a_{m-j}(x,\hat{\xi})|\xi|^{m-j}, \quad \text{with } a_{m-j} \in \mathcal{C}^\infty(\mathbb{R}^n \times S^{n-1}),$$

and where the tilde denotes an "asymptotic expansion"—truncating the expansion at the $|\xi|^{m-N}$ term gives an error that is $O(|\xi|^{m-N-1})$.[12]

If X is a Riemannian manifold, we may define $S^m_{\mathrm{cl}}(T^*X)$ in the same fashion, insisting that these conditions hold in local coordinates.[13]

(For later use, we will also want symbols in a more general geometric setting: if E is a vector bundle we define

$$S^m_{\mathrm{cl}}(E)$$

to consist of smooth functions having an asymptotic expansion, as above, in the fiber variables. Often, we will be concerned with trivial examples like $E = \mathbb{R}^n_x \times \mathbb{R}^k_\xi$, where we will usually use Greek letters to distinguish the fiber variables.)

The classical symbols are the functions that we will "quantize" into operators using the definition (3.3). As with fiber-polynomials, the symbol that we quantize to make a given operator will transform in a complicated manner under change of variables, but the *top order* part of the symbol, $a_m(x,\hat{\xi}) \in \mathcal{C}^\infty(S^*\mathbb{R}^n)$, will transform invariantly.

---

[12]This does not, of course, mean that the series has to converge, or, if it converges, that it has to converge to a : we never said a had to be *analytic* in $|\xi|^{-1}$, after all.

[13]One should of course check that the conditions for being a classical symbol are in fact coordinate invariant.

Exercise 3.4. We say that a function $a \in \mathcal{C}^\infty(T^*X)$ is a *Kohn-Nirenberg symbol* of order m on T^*X (and write $a \in S^m_{\mathrm{KN}}(T^*X)$) if for all α, β,

$$(3.5) \qquad \sup \langle \xi \rangle^{|\beta|-m} |\partial_x^\alpha \partial_\xi^\beta a| = C_{\alpha,\beta} < \infty.$$

Check that $S^m_{\mathrm{cl},c}(T^*\mathbb{R}^n) \subset S^m_{\mathrm{KN}}(T^*\mathbb{R}^n)$, where the extra subscript c denotes compact support in the base variables. Find examples of Kohn-Nirenberg symbols compactly supported in x that are not classical symbols.[14]

In the interests of full disclosure, it should be pointed out that it is the Kohn-Nirenberg symbols, rather than the classical ones defined above, that are conventionally used in the definition of the pseudodifferential calculus.

At this point, as discussed in the previous section, we are in a position to "define" the pseudodifferential calculus as sketched at the end of the previous section: it consists of operators whose Schwartz kernels near the diagonal look like the quantizations of classical symbols, and away from the diagonal are smooth. While our quantization procedure so far has been restricted to $\mathbb{R}^n$, the theory is in fact cleanest on compact manifolds, so *we shall state the properties of the calculus only for X a compact n-manifold.*[15] Most of the properties continue to hold on noncompact manifolds provided we are a little more careful either to control the behavior of the symbols at infinity, or if we restrict ourselves to "properly supported" operators, where the projections to each factor of the support of the Schwartz kernels give proper maps. We will therefore not shy away from pseudodifferential operators on $\mathbb{R}^n$, for instance, even though they are technically a bit distinct; indeed we will only use them in situations where we could in fact localize, and work on a large torus instead.

Instead of trying to make a definition of the calculus and read off its properties, we shall simply try to axiomatize these objects:

THE SPACE OF PSEUDODIFFERENTIAL OPERATORS $\Psi^*(X)$ ON A COMPACT MANIFOLD X ENJOYS THE FOLLOWING PROPERTIES. (Note that this enumeration is followed by further commentary.)

(I) (Algebra property) $\Psi^m(X)$ is a vector space for each $m \in \mathbb{R}$. If $A \in \Psi^m(X)$ and $B \in \Psi^{m'}(X)$ then $AB \in \Psi^{m+m'}(X)$. Also, $A^* \in \Psi^m(X)$. Composition of operators is associative and distributive. The identity operator is in $\Psi^0(X)$.

---

[14]Note that most authors use S^m to denote S^m_{KN}.

[15]Some remarks about the noncompact case will be found in the explanatory notes that follow.

(II) (Characterization of smoothing operators) We let

$$\Psi^{-\infty}(X) = \bigcap_m \Psi^m(X);$$

the operators in $\Psi^{-\infty}(X)$ are exactly those whose Schwartz kernels are $\mathcal{C}^\infty$ functions on $X \times X$, and can also be characterized by the property that they map distributions to smooth functions on X.

(III) (Principal symbol homomorphism) There is family of linear "principal symbol maps" $\hat{\sigma}_m : \Psi^m(X) \to \mathcal{C}^\infty(S^*X)$ such that if $A \in \Psi^m(X)$ and $B \in \Psi^{m'}(X)$,

$$\hat{\sigma}_{m+m'}(AB) = \hat{\sigma}_m(A)\hat{\sigma}_{m'}(B)$$

and

$$\hat{\sigma}_m(A^*) = \overline{\hat{\sigma}_m(A)}$$

We think of the principal symbol either as a function on the unit cosphere bundle S^*X or as a homogeneous function of degree m on T^*X, depending on the context, and we let $\sigma_m(A)$ denote the latter.

(IV) (Symbol exact sequence) There is a short exact sequence

$$0 \to \Psi^{m-1}(X) \to \Psi^m(X) \overset{\hat{\sigma}_m}{\to} \mathcal{C}^\infty(S^*X) \to 0,$$

hence the principal symbol of order m is 0 if and only if an operator is of order $m-1$.

(V) There is a linear "quantization map" $\mathrm{Op} : S^m_{\mathrm{cl}}(T^*X) \to \Psi^m(X)$ such that if $a \sim \sum_{j=0}^\infty a_{m-j}(x,\hat{\xi})|\xi|^{m-j} \in S^m_{\mathrm{cl}}(T^*X)$ then

$$\hat{\sigma}_m(\mathrm{Op}(a)) = a_m(x,\hat{\xi}).$$

The map Op is onto, modulo $\Psi^{-\infty}(X)$.

(VI) (Symbol of commutator) If $A \in \Psi^m(X)$, $B \in \Psi^{m'}(X)$ then[16] $[A,B] \in \Psi^{m+m'-1}(X)$, and we have

$$\sigma_{m+m'}([A,B]) = i\{\sigma_m(a), \sigma_{m'}(b)\}.$$

(VII) (L^2-boundedness, compactness) If $A = \mathrm{Op}(a) \in \Psi^0(X)$ then $A : L^2(X) \to L^2(X)$ is bounded, with a bound depending on finitely many constants $C_{\alpha,\beta}$ in (3.5). Moreoever, if $A \in \Psi^m(X)$, then

$$A \in \mathcal{L}(H^s(X), H^{s-m}(X)) \text{ for all } s \in \mathbb{R}.$$

---

[16]That the order is $m + m' - 1$ follows from Properties (III), (IV).

Note in particular that A maps $\mathcal{C}^\infty(X) \to \mathcal{C}^\infty(X)$. As a further consequence, note that operators of negative order are compact operators on $L^2(X)$.

(VIII) (Asymptotic summation) Given $A_j \in \Psi^{m-j}(X)$, with $j \in \mathbb{N}$, there exists $A \in \Psi^m(X)$ such that

$$A \sim \sum_j A_j,$$

which means that

$$A - \sum_{j=0}^{N} A_j \in \Psi^{m-N-1}(X)$$

for each N.

(IX) (Microsupport) Let $A = \mathrm{Op}(a) + R$, $R \in \Psi^{-\infty}(X)$. The set of $(x_0, \hat{\xi}_0) \in S^*X$ such that $a(x,\xi) = O(|\xi|^{-\infty})$ for $x, \hat{\xi}$ in some neighborhood of $(x_0, \hat{\xi}_0)$ is well-defined, independent of our choice of quantization map. Its complement is called the *microsupport* of A, and is denoted $\mathrm{WF}' A$. We moreover have

$$\mathrm{WF}'\, AB \subseteq \mathrm{WF}'\, A \cap \mathrm{WF}'\, B, \quad \mathrm{WF}'(A+B) \subseteq \mathrm{WF}'\, A \cup \mathrm{WF}'\, B,$$

$$\mathrm{WF}'\, A^* = \mathrm{WF}'\, A.$$

The condition $\mathrm{WF}'\, A = \emptyset$ is equivalent to $A \in \Psi^{-\infty}(X)$.

COMMENTARY:

(I) If we begin by defining our operators on $\mathbb{R}^n$ by the formula (3.3), with $a \in S^m_{\mathrm{cl}}(T^*\mathbb{R}^n)$, it is quite nontrivial to verify that the composition of two such operators is of the same type; likewise for adjoints. Much of the work that we are omitting in developing the calculus goes into verifying this property.

(II) On a non-compact manifold, it is only among, say, properly supported operators that elements of $\Psi^{-\infty}(X)$ are characterized by mapping distributions to smooth functions.

(III) Note that there is no sensible, *invariant*, way to associate, to an operator A, a "total symbol" a such that $A = \mathrm{Op}(a)$. As we saw before, a putative "total symbol" even for differential operators would be catastrophically bad under change of variables. Moreover, as we also saw for differential operators, it's a little hard to see what the total symbol of the composition is. This principal symbol map is a compromise that turns out to be extremely useful, especially when coupled with the asymptotic summation property, in making iterative arguments.

(IV) A good way to think of this is that $\hat{\sigma}_m$ is just the obstruction to an operator in $\Psi^m(X)$ being of order $m - 1$.

 (V) The map Op is far from unique. Even on $\mathbb{R}^n$, for instance, we can use Op_ℓ as defined by (3.2) but we could also use the "Weyl" quantization

$$(\text{Op}_W(a)\psi)(x) = (2\pi)^{-n} \iint a((x + y)/2, \xi)e^{i(x-y)\cdot\xi}\psi(y)\,dy\,d\xi$$

or the "right" quantization

$$(\text{Op}_r(a)\psi)(x) = (2\pi)^{-n} \iint a(y, \xi)e^{i(x-y)\cdot\xi}\psi(y)\,dy\,d\xi$$

or any of the obvious interpolating choices. On a manifold the choices to be made are even more striking. One convenient choice that works globally on a manifold is what might be called "Riemann-Weyl" quantization: Fix a Riemannian metric g. Given $a \in S_{\text{cl}}^m(T^*X)$, define the Schwartz kernel of an operator A by

$$\kappa(A)(x, y) = (2\pi)^{-n} \int \chi(x, y)a(m(x, y), \xi)e^{i(\exp_y^{-1}(x), \xi)}\,dg_\xi;$$

here χ is a cutoff localizing near the diagonal and in particular, within the injectivity radius; $m(x, y)$ denotes the midpoint of the shortest geodesic between x, y, exp denotes the exponential map, and the round brackets denote the pairing of vectors and covectors. The "Weyl" in the name refers to the evaluation of a at $m(x, y)$ as opposed to x or y (which give rise to corresponding "left" and "right" quantizations respectively—also acceptable choices). The "Riemann" of course refers to our use of a choice of metric.

We will often only employ a single simple consequence of the existence of a quantization map: given $a_m \in \mathcal{C}^\infty(S^*X)$ and $m \in \mathbb{R}$, there exists $A \in \Psi^m(X)$ with principal symbol a_m and with $\text{WF}'\,A = \text{supp}\,a_m$.

(VI) A priori of course $AB - BA \in \Psi^{m+m'}(X)$; however the principal symbol vanishes, by the commutativity of $\mathcal{C}^\infty(S^*X)$. Hence the need for a lower-order term, which is subtler, and noncommutative. That the Poisson bracket is well-defined independent of coordinates reflects the fact that T^*X is naturally a symplectic manifold, and the Poisson bracket is well-defined on such a manifold (see §4.1 below).

Exercise 3.5. Check (by actually performing a change of coordinates) that if $f, g \in \mathcal{C}^\infty(T^*X)$, then $\{f, g\}$ is well-defined, independent of coordinates.

This property is the one which ties classical dynamics to quantum evolution, as the discussion in §3.2 shows.

(VII) Remarkably, the mapping property is one that can be derived from the other properties of the calculus purely algebraically, with the only analytic input being boundedness of operators in $\Psi^{-\infty}(X)$. This is the famous Hörmander "square-root" argument— see [10], as well as Exercise 3.12 below.

On noncompact manifolds, restricting our attention to properly supported operators gives boundedness $L^2 \to L^2_{\mathrm{loc}}$.

The compactness of negative order operators of course follows from boundedness, together with Rellich's lemma, but is worth emphasizing; we can regard $\hat{\sigma}_0$ as the "obstruction to compactness" in general. On noncompact manifolds, this compactness property fails quite badly, resulting in much interesting mathematics.

(VIII) This follows from our ability to do the corresponding "asymptotic summation" of total symbols, which in turn is precisely "Borel's Lemma," which tells us that any sequence of coefficients are the Taylor coefficients of a $\mathcal{C}^\infty$ function; here we are applying the result to smooth functions on the radial compactification of T^*X, and the Taylor series is in the variable $\sigma = |\xi|^{-1}$, at $\sigma = 0$.

(IX) Since the total symbol is not well-defined, it is not so obvious that the microsupport is well-defined; verifying this requires checking how the total symbol transforms under change of coordinates; likewise, we may verify that the (highly non-invariant) formula for the total symbol of the composition respects microsupports to give information about $\mathrm{WF}' AB$.

3.5. **Some consequences.** If you believe that there exists a calculus of operators with the properties enumerated above, well, then you believe quite a lot! For instance:

Theorem 3.4. *Let $P \in \Psi^m(X)$ with $\hat{\sigma}_m(P)$ nowhere vanishing on S^*X. Then there exists $Q \in \Psi^{-m}(X)$ such that*

$$QP - I, PQ - I \in \Psi^{-\infty}(X).$$

In other words, P has an approximate inverse ("parametrix") which succeeds in inverting it modulo smoothing operators.

An operator P with nonvanishing principal symbol is said to be *elliptic*. Note that this theorem gives us, via the Sobolev estimates of (VII), the usual elliptic regularity estimates. In particular, we can deduce

$$Pu \in \mathcal{C}^\infty(X) \implies u \in \mathcal{C}^\infty(X).$$

Exercise 3.6. Prove this.

Proof. Let $q_{-m} = (1/\hat{\sigma}_m(P))$; let $Q_{-m} \in \Psi^{-m}(X)$ have principal symbol q_{-m}. (Such an operator exists by the exactness of the short exact symbol sequence.) Then by (III),

$$\hat{\sigma}_0(PQ_{-m}) = 1,$$

hence by (IV),[17]

$$PQ_{-m} - I = R_{-1} \in \Psi^{-1}(X).$$

Now we try to correct for this "error term:" pick $Q_{-m-1} \in \Psi^{-m-1}(X)$ with

$$\hat{\sigma}_{-m-1}(Q_{-m-1}) = -\hat{\sigma}_{-1}(R_{-1})/\hat{\sigma}_m(P).$$

Then we have

$$P(Q_{-m} + Q_{-m-1}) - I = R_{-2} \in \Psi^{-2}(X).$$

Continuing iteratively, we get a series of $Q_j \in \Psi^{-m-j}$ such that

$$P(Q_{-m} + \cdots + Q_{-m-N}) - I \in \Psi^{-N-1}(X).$$

Using (VIII), pick

$$Q \sim \sum_{j=-m}^{-\infty} Q_j.$$

This gives the desired parametrix:

Exercise 3.7.
  (1) Check that $PQ - I \in \Psi^{-\infty}(X)$.
  (2) Check that $QP - I \in \Psi^{-\infty}(X)$. (HINT: First check that a left parametrix exists; you may find it helpful to take adjoints. Then check that the left parametrix must agree with the right parametrix.)

$\square$

---

[17]The identity operator has principal symbol equal to 1, since the symbol map is a homomorphism.

Exercise 3.8. Show that an elliptic pseudodifferential operator on a compact manifold is Fredholm. (HINT: You can show, for instance, that the kernel is finite dimensional by observing that the existence of a parametrix implies that the identity operator on the kernel is equal to a smoothing operator, which is compact.)

Exercise* 3.9.
 (1) Let X be a compact manifold. Show that if $P \in \Psi^m(X)$ is elliptic, and has an actual inverse operator P^{-1} as a map from smooth functions to smooth functions, then $P^{-1} \in \Psi^{-m}(X)$. (HINT: Show that the parametrix differs from the inverse by an operator in $\Psi^{-\infty}(X)$—remember that an operator is in $\Psi^{-\infty}(X)$ if and only if it maps distributions to smooth functions.)
 (2) More generally, show that if $P \in \Psi^m(X)$ is elliptic, then there exists a generalized inverse of P, inverting P on its range, mapping to the orthocomplement of the kernel, and annihilating the orthocomplement of the range, that lies in $\Psi^{-m}(X)$.

Exercise* 3.10. Let X be compact, and P an elliptic operator on X, as above, with positive order. Using the spectral theorem for compact, self-adjoint operators, show that if $P^* = P$, then there is an orthornormal basis for $L^2(X)$ of eigenfunctions of P, with eigenvalues tending to $+\infty$. Show that the eigenfunctions are in $\mathcal{C}^\infty(X)$. (HINT: show that there exists a basis of such eigenfunctions for the generalized inverse Q and then see what you can say about P.)

Exercise 3.11. Let X be compact.
 (1) Show that the principal symbol of Δ, the Laplace-Beltrami operator on a compact Riemannian manifold, is just
$$|\xi|_g^2 \equiv \sum g^{ij}(x)\xi_i\xi_j,$$
    the metric induced on the cotangent bundle.
 (2) Using the previous exercise, conclude that there exists an orthonormal basis for $L^2(X)$ of eigenfunctions of Δ, with eigenvalues tending toward $+\infty$.

Exercise 3.12. Work out the Hörmander "square root trick" on a compact manifold X as follows.
 (1) Show that if $P \in \Psi^0(X)$ is self-adjoint, with positive principal symbol, then P has an approximate square root, i.e. there exists $Q \in \Psi^0(X)$ such that $Q^* = Q$ and $P - Q^2 \in \Psi^{-\infty}(X)$. (HINT: Use an iterative construction, as in the proof of existence of elliptic parametrices.)

(2) Show that operators in $\Psi^{-\infty}(X)$ are L^2-bounded.
(3) Show that an operator $A \in \Psi^0(X)$ is L^2-bounded. (HINT: Take an approximate square root of $\lambda I - A^*A$ for $\lambda \gg 0$.)

As usual, let Δ denote the Laplacian on a compact manifold. By Exercise 3.12, there exists an operator $A \in \Psi^1(X)$ such that $A^2 = \Delta + R$, with $R \in \Psi^{-\infty}(X)$. By abstract methods of spectral theory, we know that $\sqrt{\Delta}$ exists as an unbounded operator on $L^2(X)$. (This is a very simple use of the functional calculus: merely take $\sqrt{\Delta}$ to act by multiplication by λ_j on each ϕ_j, where (ϕ_j, λ_j^2) are the eigenfunctions and eigenvalues of the Laplacian, from Exercise 3.11.) In fact, we can improve this argument to obtain:

Proposition 3.5.
$$\sqrt{\Delta} \in \Psi^1(X).$$

Indeed, it follows from a theorem of Seeley that all complex powers of a self-adjoint, elliptic pseudodifferential operator[18] on a compact manifold are pseudodifferential operators.

All proofs of the proposition seem to introduce an auxiliary parameter in some way, and the following (taken directly from [24, Chapter XII, §1]) seems one of the simplest. An alternative approach, using the theory of elliptic boundary problems, is sketched in [26, pp.32-33, Exercises 4–6].

Proof. Let A be the self-adjoint parametrix constructed in Exercise 3.12, so that
$$A^2 - \Delta = R \in \Psi^{-\infty}(X).$$
By taking a parametrix for the square root of A, in turn, we obtain
$$A = B^2 + R'$$
with $B \in \Psi^{1/2}(X)$ and $R' \in \Psi^{-\infty}$, both self-adjoint; then pairing with a test function ϕ shows that
$$\langle A\phi, \phi \rangle \geq \langle R'\phi, \phi \rangle \geq -C\|u\|^2$$
for some $C \in \mathbb{R}$. Thus, A can only have finitely many nonpositive eigenvalues (since it has a compact generalized inverse) hence its eigenvalues can accumulate only at $+\infty$). So we may alter A by the smoothing operator projecting off of these eigenspaces, and maintain
$$A^2 - \Delta = R \in \Psi^{-\infty}(X)$$
(with a different R, of course) while now ensuring that A is positive.

---
[18]Seeley's theorem is better yet: self-adjointness is unnecessary.

Now we may write, using the spectral theorem,

$$(\Delta')^{-1/2} = \frac{1}{2\pi i} \int_\Gamma z^{-1/2}((\Delta') - z)^{-1} \, dz$$

where Γ is a contour encircling the positive real axis counterclockwise, and given by $\operatorname{Im} z = \operatorname{Re} z$ for z sufficiently large, and Δ' is given by Δ minus the projection onto constants (hence has no zero eigenvalue). (The integral converges in norm, as self-adjointness of Δ' yields

$$\left\|((\Delta') - z)^{-1}\right\|_{L^2 \to L^2} \lesssim |\operatorname{Im} z|^{-1}.)$$

Likewise, since $A^2 = \Delta' + R$ (with R yet another smoothing operator) we may write

$$A^{-1} = \frac{1}{2\pi i} \int_\Gamma z^{-1/2}((\Delta') + R - z)^{-1} \, dz$$

Hence

$$(\Delta')^{-1/2} - A^{-1} = \frac{1}{2\pi i} \int_\Gamma z^{-1/2}\left[((\Delta') - z)^{-1} - ((\Delta') + R - z)^{-1}\right] dz$$

$$= \frac{1}{2\pi i} \int_\Gamma z^{-1/2}((\Delta') - z)^{-1} R((\Delta') + R - z)^{-1} \, dz.$$

Now the integrand, $z^{-1/2}((\Delta') - z)^{-1} R((\Delta') + R - z)^{-1}$, is for each z a smoothing operator, and decays fast enough that when applied to any $u \in \mathcal{D}'(X)$, the integral converges to an element of $\mathcal{C}^\infty(X)$ (in particular, the integral converges in $\mathcal{C}^0(X)$, even after application of Δ^k on the left, for any k). Hence

$$(\Delta')^{-1/2} - A^{-1} = E \in \Psi^{-\infty}(X);$$

thus we also obtain

$$(\Delta')^{1/2} = (A^{-1} + E)^{-1} \in \Psi^1(X);$$

as $(\Delta')^{1/2}$ differs from $\Delta^{1/2}$ by the smoothing operator of projection onto constants, this shows that

$$\Delta^{1/2} \in \Psi^1(X). \quad \square$$

4. Wavefront set

If $P \in \Psi^m(X)$ and $(x_0, \xi_0) \in S^*X$, we say P is elliptic at (x_0, ξ_0) if $\hat{\sigma}_m(P)(x_0, \xi_0) \neq 0$. Of course if P is elliptic at each point in S^*X, it is elliptic in the sense defined above. We let

$$\operatorname{ell}(P) = \{(x, \xi) : P \text{ is elliptic at } (x, \xi)\},$$

and let

$$\Sigma_P = S^*X \backslash \operatorname{ell}(X);$$

Σ_P is known as the *characteristic set* of P.

Exercise 4.1.
  (1) Show that $\operatorname{ell} P \subseteq \operatorname{WF}' P$.
  (2) If P is a differential operator of order m of the form $\sum a_\alpha(x) D^\alpha$ then show that $\operatorname{WF}' P = \pi^*(\bigcup \operatorname{supp} a_\alpha)$, while $\operatorname{ell} P$ may be smaller.

The following "partition of unity" result, and variants on it, will frequently be useful in discussing microsupports. It yields an operator that is microlocally the identity on a compact set, and microsupported close to it.

Lemma 4.1. *Given $K \subset U \subset S^*X$ with K compact, U open, there exists a self-adjoint operator $B \in \Psi^0(X)$ with*

$$\operatorname{WF}'(\operatorname{Id} - B) \cap K = \emptyset, \ \operatorname{WF}' B \subset U.$$

Exercise 4.2. Prove the lemma. (HINT: You might wish to try constructing B in the form

$$\operatorname{Op}(\psi \sigma_{\operatorname{tot}}(\operatorname{Id}))$$

where $\sigma_{\operatorname{tot}}(\operatorname{Id})$ is the total symbol of the identity (which is simply 1 for all the usual quantizations on $\mathbb{R}^n$) and ψ is a cutoff function equal to 1 on K and supported in U. Then make B self-adjoint.)

Theorem 4.2. *If $P \in \Psi^m(X)$ is elliptic at (x_0, ξ_0), there exists a microlocal elliptic parametrix $Q \in \Psi^{-m}(X)$ such that*

$$(x_0, \xi_0) \notin \operatorname{WF}'(PQ - I) \cup \operatorname{WF}'(QP - I).$$

In other words, you should think of Q as inverting P *microlocally near* (x_0, ξ_0).

Exercise 4.3. Prove the theorem. (HINT: If B is a microlocal partition of unity as in Lemma 4.1, microsupported sufficiently close to (x_0, ξ_0) and microlocally the identity in a smaller neighborhood, then show

$$W = BP + \lambda \operatorname{Op}(\langle \xi \rangle^m)(\operatorname{Id} - B)$$

is globally elliptic provided $\lambda \in \mathbb{C}$ is chosen appropriately. Now, using the existence of an elliptic parametrix for W, prove the theorem.)

Let u be a distribution on a manifold X. We define the *wavefront set* of u as follows.

Definition 4.3. The wavefront set of u,

$$\operatorname{WF} u \subseteq S^*X,$$

is given by

$$(x_0, \xi_0) \notin \mathrm{WF}\, u$$

if and only if there exists $P \in \Psi^0(X)$, elliptic at (x_0, ξ_0), such that

$$Pu \in \mathcal{C}^\infty.$$

Exercise 4.4. Show that the choice of $\Psi^0(X)$ in this definition is immaterial, and that we get the same definition of $\mathrm{WF}\, u$ if we require $P \in \Psi^m(X)$ instead.

Note that the wavefront set is, from its definition, a closed set. Instead of viewing $\mathrm{WF}\, u$ as a subset of S^*X, we also, on occasion, think of $\mathrm{WF}\, u$ as a *conic* subset of $T^*X \backslash o$, with o denoting the zero section; a conic set in a vector bundle is just one that is invariant under the $\mathbb{R}^+$ action on the fibers.

An important variant is as follows: we say that

$$(x_0, \xi_0) \notin \mathrm{WF}^m\, u$$

if and only if there exists $P \in \Psi^m(X)$, elliptic at (x_0, ξ_0) such that

$$Pu \in L^2(X).$$

Proposition 4.4. $\mathrm{WF}\, u = \emptyset$ *if and only if* $u \in \mathcal{C}^\infty(X)$; $\mathrm{WF}^m\, u = \emptyset$ *if and only if* $u \in H^m_{loc}(X)$.

The wavefront set serves the purpose of measuring not just where, but also in what (co-)direction, a distribution fails to be in $\mathcal{C}^\infty(X)$ (or H^m in the case of the indexed version). It is instructive to think about testing for such regularity, at least on $\mathbb{R}^n$, by localizing and Fourier transforming. Given $(x_0, \hat{\xi}_0) \in S^*\mathbb{R}^n$, let $\phi \in \mathcal{C}^\infty_c(\mathbb{R}^n)$ be nonzero at x_0; let $\gamma \in \mathcal{C}^\infty(\mathbb{R}^n)$ be given by

$$\gamma(\xi) = \psi\big(\big|\frac{\xi}{|\xi|} - \hat{\xi}_0\big|\big)\chi(|\xi|)$$

where ψ is a cutoff function supported near $x = 0$ and $\chi(t) \in \mathcal{C}^\infty(\mathbb{R})$ is equal to 0 for $t < 1$ and 1 for $t > 2$. Think of γ as a cutoff in a cone of directions near ξ_0, but modified to be smooth at the origin. (We will use such a construction frequently, and refer in future to a function such as γ as a "conic cutoff near direction $\hat{\xi}_0$.".)

Now note that $\phi(x)\gamma(\xi)$ is a symbol of order zero, and

$$(4.1) \quad \mathrm{Op}_\ell(\phi(x)\gamma(\xi))^* = \mathrm{Op}_r(\phi(x)\gamma(\xi))u = (2\pi)^{-n}\mathcal{F}^{-1}\gamma(\xi)\mathcal{F}(\phi u).$$

By definition, if $\mathrm{Op}_\ell(\phi(x)\gamma(\xi))^*u \in \mathcal{C}^\infty$, then $(x_0, \xi_0) \notin \mathrm{WF}\, u$. Note that since ϕu has compact support, we automatically have $\mathcal{F}(\phi u) \in$

$\mathcal{C}^\infty$, hence $\mathcal{F}^{-1}\gamma\mathcal{F}(\phi u)$ is rapidly decreasing. Since $\mathcal{F}$ is an isomorphism from $\mathcal{S}(\mathbb{R}^n)$ to itself, we see that it in fact suffices to have

$$\gamma\mathcal{F}(\phi u) \in \mathcal{S}(\mathbb{R}^n)$$

to be able to conclude that $(x_0, \xi_0) \notin \mathrm{WF}\, u$. Conversely, one can check that the class of operators of the form

$$\mathrm{Op}_\ell(\phi(x)\gamma(\xi))^*$$

is rich enough that this in fact amounts to a *characterization* of wavefront set:

Proposition 4.5. *We have $(x_0, \xi_0) \notin \mathrm{WF}\, u$ if and only if there exist ϕ, γ as above with*

$$\gamma\mathcal{F}(\phi u) \in \mathcal{S}(\mathbb{R}^n).$$

Exercise 4.5. Prove the Proposition. (HINT: If $A \in \Psi^0(\mathbb{R}^n)$ is elliptic at (x_0, ξ_0) and $Au \in \mathcal{C}^\infty(\mathbb{R}^n)$, construct $B = \mathrm{Op}_\ell(\phi(x)\gamma(\xi))^*$ as above so that $\mathrm{WF}'\, B$ is contained in the set where A is elliptic. Hence there is a microlocal parametrix Q such that $B(QA - I) \in \Psi^{-\infty}(X)$.)

Note that if u is smooth near x_0, then we have $\phi u \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ for appropriately chosen ϕ, hence there is no wavefront set in the fiber over x_0.

If, by contrast, u is not smooth in any neighborhood of x_0, then we of course do not have $\mathcal{F}(\phi u) \in \mathcal{S}$, although it is in $\mathcal{C}^\infty$; the wavefront set includes the directions in which it fails to be rapidly decaying.

Thus, we can easily see that in fact the projection to the base variables of $\mathrm{WF}\, u$ is the *singular support* of u, i.e. the points which have no neighborhood in which the distribution u is a $\mathcal{C}^\infty$ function.

Exercise 4.6. Let $\Omega \subset \mathbb{R}^n$ be a domain with smooth boundary. Show that $\mathrm{WF}\, 1_\Omega = SN^*(\partial\Omega)$, the spherical normal bundle of the boundary. (HINT: You may want to use the fact that the definition of $\mathrm{WF}\, u$ is coordinate-invariant.)

We have a result constraining the wavefront set of a solution to a PDE or, more generally, a pseudodifferential equation, directly following from the definition:

Theorem 4.6. *If $Pu \in \mathcal{C}^\infty(X)$, then $\mathrm{WF}\, u \subseteq \Sigma_P$.*

Proof. By definition, $Pu \in \mathcal{C}^\infty(X)$ means that $\mathrm{WF}\, u \cap \mathrm{ell}\, P = \emptyset$. $\quad\square$

Theorem 4.7. *If $P \in \Psi^*(X)$, $\mathrm{WF}\, Pu \subseteq \mathrm{WF}\, u \cap \mathrm{WF}'\, P$.*

Exercise 4.7. Prove this, using microlocal elliptic parametrices for the inclusion in $\mathrm{WF}\, u$.

The property of pseudodifferential operators that $\mathrm{WF}\, Pu \subseteq \mathrm{WF}\, u$ is called "microlocality:" the operators are not "local," in that they do move *supports* of distributions around, but they don't move *singularities*, even in the refined sense of wavefront set.

We shall also need related results on Sobolev based wavefront sets in what follows:

Proposition 4.8. *If $P \in \Psi^m(X)$, $\mathrm{WF}^{k-m}\, Pu \subseteq \mathrm{WF}^k\, u \cap \mathrm{WF}'\, P$ for all $k \in \mathbb{R}$.*

Corollary 4.9. *Let $P \in \Psi^m(X)$. If*

$$\mathrm{WF}'\, P \cap \mathrm{WF}^m\, u = \emptyset$$

then

$$Pu \in L^2(X).$$

Exercise 4.8. Prove the proposition (again using a microlocal elliptic parametrix) and the corollary.

We will have occasion to use the following relationship between ordinary and Sobolev-based wavefront sets:

Proposition 4.10.

$$\mathrm{WF}\, u = \overline{\bigcup_k \mathrm{WF}^k\, u}.$$

Exercise 4.9. Prove the proposition.

Exercise 4.10. Let $\square$ denote the wave operator,

$$\square u = D_t^2 u - \Delta u$$

on $M = \mathbb{R} \times X$ with X a Riemannian manifold. Show that the wavefront set of u is a subset of the "wave cone" $\{\tau^2 = |\xi|_g^2\}$ where τ is the dual variable to t and ξ to x in $T^*(M)$.

Exercise 4.11.

(1) Let $k < n$, and let $\iota : \mathbb{R}^k \to \mathbb{R}^n$ denote the inclusion map.
    Show that there is a continuous *restriction map* on compactly supported distributions with no wavefront set conormal to $\mathbb{R}^k$:

$$\iota^* : \{u \in \mathcal{E}'(\mathbb{R}^n) : \mathrm{WF}\, u \cap SN^*(\mathbb{R}^k) = \emptyset\} \to \mathcal{E}'(\mathbb{R}^k).$$

HINT: Show that it suffices to consider u supported in a small neighborhood of a single point in $\mathbb{R}^k$. Then take the Fourier transform of u and try to integrate in the conormal variables to obtain the Fourier transform of the restriction.

(2) Show that, with the notation of the previous part,
$$\mathrm{WF}\,\iota^* u \subseteq \iota_*(\mathrm{WF}\,u)$$
where $\iota_* : T^*_{\mathbb{R}^k}\mathbb{R}^n \to T^*\mathbb{R}^k$ is the naturally defined projection map.

(3) Show that both the previous parts make sense, and are valid, for restriction to an embedded submanifold Y of a manifold X.

(4) Show that if u is a distribution on $\mathbb{R}^k_x$ and v is a distribution on $\mathbb{R}^l_y$ then $w = u(x)v(y)$ is a distribution on $\mathbb{R}^{k+l}$ and

$$\mathrm{WF}\,w \subseteq \big[(\operatorname{supp} u, 0) \times \mathrm{WF}\,v\big] \cup \big[\mathrm{WF}\,u \times (\operatorname{supp} v, 0)\big] \cup \mathrm{WF}\,u \times \mathrm{WF}\,v.$$

(HINT: Localize and Fourier transform, as in (4.1).)

You might wonder: given P, can the wavefront set of a solution to $Pu = 0$ be any closed subset of Σ? The answer is no, there are, in general, further constraints. To talk about them effectively, we should digress briefly back into geometry.

4.1. **Hamilton flows.** We now amplify the discussion §3.2 of Hamiltonian mechanics and symplectic geometry, generalizing it to a broader geometric context.

Let N be a symplectic manifold, that is to say, one endowed with a closed, nondegenerate[19] two-form. (Our prime example is $N = T^*X$, endowed with the form $\sum d\xi_j \wedge dx^j$; by Darboux's theorem, every symplectic manifold in fact locally looks like this.)

Given a real-valued function $a \in \mathcal{C}^\infty(N)$, we can make a *Hamilton vector field* from a as follows: by nondegeneracy, there is a unique vector field H_a such that $\iota_{\mathsf{H}_a}\omega \equiv \omega(\cdot, \mathsf{H}_a) = da$.

Exercise 4.12. Check that in local coordinates in T^*X,
$$\mathsf{H}_a = \sum_{j=1}^n \frac{\partial a}{\partial \xi_j}\partial_{x^j} - \frac{\partial a}{\partial x^j}\partial_{\xi_j}.$$

Thus, for any smooth function b, we may define the Poisson bracket
$$\{a, b\} = \mathsf{H}_a(b)$$

Exercise 4.13. Check that the Poisson bracket is antisymmetric.

It is easy to verify that the flow along H_a preserves both the symplectic form and the function a : we have from Cartan's formula (and since ω is closed):
$$\mathcal{L}_{\mathsf{H}_a}(\omega) = d\iota_{\mathsf{H}_a}\omega = d(da) = 0;$$

---

[19]Nondegeneracy of ω means that contraction with ω is an isomorphism from T_pN to T^*_pN at each point.

also,

$$\mathsf{H}_a(a) = da(\mathsf{H}_a) = \omega(\mathsf{H}_a, \mathsf{H}_a) = 0.$$

The integral curves of the vector field H_a are called the *bicharacteristics* of a and those lying inside $\Sigma_a = \{a = 0\}$ are called *null bicharacteristics*.

Exercise* 4.14.

   (1) Show that the bicharacteristics of $|\xi|_g = (\sigma_2(\Delta))^{1/2}$ project to X to be geodesics. The flow along the Hamilton vector field of $|\xi|_g$ is known as *geodesic flow*.
   (2) Show that the null bicharacteristics of $\sigma_2(\Box)$ are lifts to $T^*(\mathbb{R} \times X)$ of geodesics of X, traversed both forward and backward at unit speed.

Recall that the setting of symplectic manifolds is exactly that of Hamiltonian mechanics: given such a manifold, we can regard it as the phase space for a particle; specifying a function (the "energy" or "Hamiltonian") gives a vector field, and the flow along this vector field is supposed to describe the time-evolution of our particle in the phase space.

Exercise 4.15. Check that the phase space evolution of the harmonic oscillator Hamiltonian, $(1/2)(\xi^2 + x^2)$ on $T^*\mathbb{R}$, agrees with what you learned in physics class long ago.

4.2. **Propagation of singularities.**

Theorem 4.11 (Hörmander)**.** *Let $Pu \in \mathcal{C}^\infty(X)$, with $P \in \Psi^m(X)$ an operator with real principal symbol. Then $\mathrm{WF}\, u$ is a union of maximally extended null bicharacteristics of $\hat{\sigma}_m(P)$ in S^*X.*

We should slightly clarify the usage here: to make sense of these null bicharacteristics, we should actually take the Hamilton vector field of the homogeneous version of the symbol, $\sigma_m(P)$; this is a homogeneous vector field, and its integral curves thus have well-defined projections onto S^*X. If the Hamilton vector field should be "radial" at some point $q \in T^*X$, i.e. coincide with a multiple of the vector field $\xi \cdot \partial_\xi$ there, then the projection of the integral curve through q is just a single point in S^*X, and the theorem gives no further information about wavefront set at that point.

For $P = \Box$, the theorem says that the wavefront set lies in the "light cone," and propagates forward and backward at unit speed

along geodesics. If we take the fundamental solution to the wave equa-tion[20] $u = \sin(t\sqrt{\Delta}/\sqrt{\Delta})\delta_p$, it is not hard to compute that in fact for small, nonzero time,[21]

$$\text{WF}\, u \subseteq N^*\{d(\cdot, p) = |t|\} \equiv \mathcal{L};$$

This is a generalization of Huygens's Principle, which tells us that in $\mathbb{R} \times \mathbb{R}^n$, the support of the fundamental solution is on this expanding sphere (but which is a highly unstable property). Note that $\mathcal{L}$ is in fact the bicharacteristic flowout of all covectors in Σ projecting to $N^*(\{p\})$ at $t = 0$, and under this interpretation, $\mathcal{L} \subset T^*(\mathbb{R} \times X)$ makes sense for *all* times, not just for short time, regardless of the metric geometry. We shall return to and amplify this point of view in §9.

Exercise 4.16.

  (1) Suppose that $\Box u = 0$ on $\mathbb{R} \times \mathbb{R}^n$ and $u(t, x) \in \mathcal{C}^\infty$ for $(t, x) \in (-\epsilon, \epsilon) \times B(0, 1)$ for some $\epsilon > 0$. Show, using Theorem 4.11, that $u \in \mathcal{C}^\infty$ on $\{|x| < 1 - |t|\}$. Can you show this more directly using the energy methods described in §2.3?
  (2) Suppose that $\Box u = 0$ on $\mathbb{R} \times \mathbb{R}^n$ and $u(t, x) \in \mathcal{C}^\infty$ for $(t, x) \in (-\epsilon, \epsilon) \times (B(0, 1) \backslash B(0, 1/2))$ for some $\epsilon > 0$. Show, using the theorem, that $u \in \mathcal{C}^\infty$ in $\{|x| < 1 - |t|\} \cap \{|t| \in (3/4, 1)\}$

Proof. [22] Note that we already know that $\text{WF}\, u \subseteq \Sigma_P$ by Theorem 4.6, hence what remains to be proved is the flow-invariance.

Let $q \in \Sigma_P \subset S^*X$. By homogeneity of $\sigma_m(P)$, we can write the Hamilton vector field in T^*X in a neighborhood of q as

$$(4.2) \qquad\qquad \mathsf{H}_p = |\xi|^{m-1}(\mathsf{V} + h\mathsf{R}),$$

where R denotes the radial vector field $\xi \cdot \partial_\xi$, h is a function on S^*X, and V is the pullback under quotient of a vector field on S^*X itself, i.e. V is homogeneous of degree zero with no radial component, hence of the form $\sum_j f_j(x, \hat{\xi})\partial_{\hat{\xi}_j} + g_j(x, \hat{\xi})\partial_{x^j}$. Note that if a is homogeneous of degree l then

$$(4.3) \qquad\qquad\qquad \mathsf{R}a = la.$$

(Exercise: Verify these consequences of homogeneity.)

---

[20]This is the spectral-theoretic way of writing the solution with initial value 0 and initial time-derivative δ_p.

[21]Well, I am cheating a bit here, as we don't haven't stated any results allowing us to relate the wavefront set of Cauchy data for the wavefront set of the solution to the equation. To understand how to do this, you should read [17].

[22]This proof is very close to those employed by Melrose in [17] and [18].

By the comments above, we may take $\mathsf{V} \neq 0$ near q; otherwise the theorem is void. Thus, without loss of generality, we may employ a coordinate system $\alpha_1, \ldots, \alpha_{n-1}$ for S^*X in which

(4.4) $$\mathsf{V} = \partial_{\alpha_1},$$

hence using $\alpha, |\xi|$ as coordinates in T^*X,

$$\mathsf{H}_p = \partial_{\alpha_1} + h\mathsf{R};$$

we may shift coordinates so that $\alpha(q) = 0$. We split the α variables into α_1 and $\alpha' = (\alpha_2, \ldots, \alpha_{n-1})$.

Since $\mathrm{WF}\, u$ is closed, it suffices to prove the following: if $q \notin \mathrm{WF}\, u$ then $\Phi_t(q) \notin \mathrm{WF}\, u$ for $t \in [-1, 1]$, where Φ_t denotes the flow generated by V.[23] (This will show that the intersection of $\mathrm{WF}\, u$ with the bicharacteristic through q is both open and closed, hence is the whole thing.)

We can make separate arguments for $t \in [0, 1]$ and $t \in [-1, 0]$, and will do so (in fact, we will leave one case to the reader).

For simplicity, let us take $Pu = 0$; we leave the case of an inhomogeneous equation for the reader (it introduces extra terms, but no serious changes will in fact be necessary in the proof).

Since $\mathrm{WF}\, u$ is closed, our assumption that $q \notin \mathrm{WF}\, u$ tells us that there is in fact a 2δ-neighborhood of 0 in the α coordinates that is disjoint from $\mathrm{WF}\, u$; we are trying to extend this regularity along the rest of the set $(\alpha_1, \alpha') \in [0, 1] \times 0$. We proceed as follows: let

(4.5) $$s_0 = \sup\{s : \mathrm{WF}^s\, u \cap \{(\alpha_1, \alpha') \in [0, 1] \times B(0, \delta)\} = \emptyset\}.$$

Pick any $s < s_0$. We will show that in fact

(4.6) $$\mathrm{WF}^{s+1/2}\, u \cap \{(\alpha_1, \alpha') \in [0, 1] \times B(0, \delta)\} = \emptyset,$$

thus establishing that $s_0 = \infty$, which is the desired result (by Proposition 4.10). One can regard this strategy as iteratively obtaining more and more regularity for u along the bicharacteristic (i.e. the idea is that we start by knowing *some* possibly very bad regularity, and we step by step conclude that we can improve upon this regularity, half a derivative at a time). More colloquially, the idea is that the "energy," as measured by testing the distribution u by pseudodifferential operators, should be comparable at different points along the bicharacteristic curve.

---

[23] Of course, we are assuming here that the interval $[-1, 1]$ remains in our coordinate neighborhood; rescale the coordinates if necessary to make this so.

Now we prove the estimates that yield (4.6) via commutator methods. Let $\phi(s)$ be a cutoff function with

(4.7)
$$\phi(t) > 0 \text{ on } (-1, 1),$$
$$\operatorname{supp} \phi = [-1, 1]$$

Let $\phi_\delta(s) = \phi(\delta^{-1}s)$; arrange that $\sqrt{\phi} \in \mathcal{C}^\infty$. Let χ be a cutoff function equal to 1 on $(0, 1)$ and with $\chi' = \psi_1 - \psi_2$, with ψ_1 supported on $(-\delta, \delta)$ and ψ_2 on $(1 - \delta, 1 + \delta)$; we will further assume that $\sqrt{\chi}, \sqrt{\psi_i} \in \mathcal{C}^\infty$.

Exercise 4.17. Verify that cutoffs with these properties exist.

In our coordinate system for S^*X, let
$$\hat{a} = \phi_\delta(|\alpha'|)\chi(\alpha_1)e^{-\lambda\alpha_1} \in \mathcal{C}^\infty(S^*X),$$

with $\lambda \gg 0$ to be chosen presently. Passing to the corresponding function on $a \in \mathcal{C}^\infty(T^*X)$ that is homogeneous of degree $2s - m + 2$, we have

(4.8) $\mathsf{H}_p(a) = |\xi|^{2s+1}\big(-\lambda\phi_\delta(|\alpha'|)\chi(\alpha_1)e^{-\lambda\alpha_1}$
$$+ \phi_\delta(|\alpha'|)(\psi_1 - \psi_2)e^{-\lambda\alpha_1} + h(\alpha)(2s - m + 2)a\big)$$

with h given by (4.2). Since a 2δ coordinate neighborhood of the origin was assumed absent from $\operatorname{WF} u$, we have in particular ensured that $\operatorname{supp} \phi_\delta(|\alpha'|)\psi_1(\alpha_1)$ is contained in $(\operatorname{WF} u)^c$. We also have $\operatorname{supp} \hat{a} \subset (\operatorname{WF}^s u)^c$ by (4.5), since $s < s_0$.
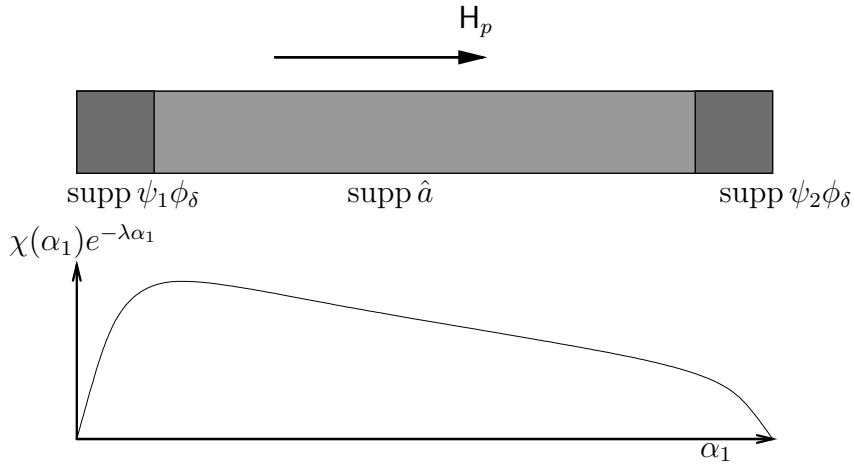

FIGURE 1. The support of the commutant and its value along the line $\alpha' = 0$. The support of the term $\psi_1(\alpha_1)\phi_\delta(|\alpha'|)$ is arranged to be contained in the complement of $\operatorname{WF} u$, while the support of the whole of a is arranged to be in the complement of $\operatorname{WF}^s u$.

Let $A \in \Psi^{2s-m+2}(X)$ be given by the quantization of a.[24] Since $\sigma_m(P)$ is real by assumption, we have $P^* - P \in \Psi^{m-1}(X)$. (Exercise: Check this!) Thus the "commutator" $P^*A - AP$, which is a priori of order $2s + 2$, has vanishing principal symbol of order $2s + 2$, hence it in fact lies in $\Psi^{2s+1}(X)$, and we may write

$$(P^*A - AP) = [P, A] + (P^* - P)A,$$

with

$$(4.9) \quad i\sigma_{2s+1}([P, A] + (P^* - P)A) = \mathsf{H}_p(a) + \sigma_{m-1}(P^* - P)a$$
$$= -\lambda\phi_\delta(|\alpha'|)\chi(\alpha_1)e^{-\lambda\alpha_1}|\xi|^{2s+1} + \phi_\delta(|\alpha'|)(\psi_1 - \psi_2)e^{-\lambda\alpha_1}|\xi|^{2s+1}$$
$$+ (i\sigma_{m-1}(P^* - P) + h(\alpha)(2s - m + 2))a,$$

by (4.2),(4.3), and (4.4). If $\lambda \gg 0$ is chosen sufficiently large, we may absorb the third term into the first, and write the RHS of (4.9) as

$$-f(\alpha)\phi_\delta(|\alpha'|)\chi(\alpha_1) + \phi_\delta(|\alpha'|)(\psi_1 - \psi_2)e^{-\lambda\alpha_1}$$

with $f > 0$ on the support of $\phi_\delta\chi$.

Let $B \in \Psi^{(2s+1)/2}(X)$ be obtained by quantization of

$$|\xi|^{s+1/2}(f(\alpha)\phi_\delta(|\alpha'|)\chi(\alpha_1))^{1/2};$$

and let $C_i \in \Psi^{(2s+1)/2}(X)$ be obtained by quantization of

$$|\xi|^{s+1/2}(\phi_\delta(|\alpha'|)\psi_i(\alpha_1))^{1/2}e^{-\lambda\alpha_1/2}.$$

Then by the symbol calculus, i.e. by Properties III, IV of the calculus of pseudodifferential operators,

$$(4.10) \quad i(P^*A - AP) = i(P^* - P)A + i[P, A] = -B^*B + C_1^*C_1 - C_2^*C_2 + R$$

with $R \in \Psi^{2s}(X)$, hence of lower order than the other terms; moreover we have $\mathrm{WF}' R \subset \operatorname{supp} \hat{a}$.

Now we "pair" both sides of (4.10) with our solution u. We have

$$i\langle(P^*A - AP)u, u\rangle = \langle(-B^*B + C_1^*C_1 - C_2^*C_2 + R)u, u\rangle;$$

as we are taking $Pu = 0$, the LHS vanishes.[25] We thus have, rearranging this equation,

$$(4.11) \qquad \|Bu\|^2 + \|C_2u\|^2 = \|C_1u\|^2 + \langle Ru, u\rangle.$$

---

[24]I.e., really A is given by cutting off a near $\xi = 0$ to give a smooth total symbol and quantizing that.

[25]In the case of an inhomogeneous equation, it is of course here that extra terms arise.

I claim that the RHS is finite: Recall that R lies in $\Psi^{2s}(X)$. Let Λ be an operator of order $-s$, elliptic on $\mathrm{WF}' R$ and with $\mathrm{WF}' \Lambda$ contained in the complement of $\mathrm{WF}^s u$.

Exercise 4.18. Show that such a Λ exists.

Thus, letting Υ be a microlocal parametrix for Λ on $\mathrm{WF}' R$, we have

$$\mathrm{WF}' R \cap \mathrm{WF}'(\mathrm{Id} - \Lambda\Upsilon) = \emptyset,$$

hence

$$R - \Lambda\Upsilon R = E \in \Psi^{-\infty}(X).$$

Thus,

$$|\langle Ru, u \rangle| \leq |\langle \Upsilon Ru, \Lambda^* u \rangle| + |\langle Eu, u \rangle| < \infty$$

by Corollary 4.9 since $\mathrm{WF}' \Upsilon R \cup \mathrm{WF}' \Lambda^* \subset (\mathrm{WF}^s u)^c$ (and since E is smoothing). Returning to (4.11), we also note that the term $\|C_1 u\|^2$ is finite by our assumptions on the location of $\mathrm{WF}^{s+1/2} u$ (and another use of Corollary (4.9)). Thus,

$$\|Bu\| < \infty,$$

and consequently,

$$\mathrm{WF}^{s+1/2} u \cap \mathrm{ell}\, B = \emptyset,$$

which was the desired estimate. $\qquad\qquad\square$

Exercise 4.19. Now see how the argument should be modified to yield absence of $\mathrm{WF}^{s+1/2} u$ on

$$\{\alpha' \in [-1, 0], \alpha' = 0\}.$$

One cheap alternative to going through the whole proof might be to notice that we also have $(-P)u \in \mathcal{C}^\infty$, and that $\mathsf{H}_{-p} = -\mathsf{H}_p$; thus, the "forward propagation" that we have just proved should yield backward propagation along H_p as well.

THE FINE PRINT: Now, having done all that, note that it was a cheat. In particular, we didn't know a priori that we could apply any of the operators that we used to u and obtain an L^2 function, let alone justify the formal integrations by parts used to move adjoints across the pairings. Therefore, to make the above argument rigorous, we need to modify it with an *approximation argument*. This is similar to the situation in Exercise 2.7, except in that case, we had a natural way of obtaining smooth solutions to the equation which approximated the desired one: we could replace our initial data ψ_0 for the Schrödinger equation by, for instance, $e^{-\epsilon\Delta}\psi_0$; the solution at later time is then just $e^{-\epsilon\Delta}\psi$, and we can consider the limit $\epsilon \downarrow 0$. In the general case to which this theorem applies, though, we do not have any convenient families of

smoothing operators commuting with P. So we instead take the tack of smoothing our *operators* rather than the solution u. We should manufacture a family of smoothing operators G_ϵ that strongly approach the identity as $\epsilon \downarrow 0$, and replace A by AG_ϵ everywhere it appears above. If we do this sensibly, then the analogs of the estimates proved above yield the desired estimates in the $\epsilon \downarrow 0$ limit. Of course, we need to know how G_ϵ passes through commutators, etc., so the right thing to do is to take the G_ϵ themselves to be pseudodifferential approximations of the identity, something like

$$G_\epsilon = \mathrm{Op}_\ell(\varphi(\epsilon|\xi|))$$

on $\mathbb{R}^n$, with $\varphi \in \mathcal{C}_c^\infty(\mathbb{R})$ a cutoff equal to 1 near 0. We content ourselves with referring the interested reader to [18] for the analogous development in the "scattering calculus" including details of the approximation argument.

Exercise 4.20.
  (1) Show the following variant of Theorem 4.11: if $P \in \Psi^m(X)$ is an operator with real principal symbol, and $Pu \in \mathcal{C}^\infty(X)$, show that $\mathrm{WF}^k u$ is a union of maximally extended bicharacteristics of P for each $k \in \mathbb{R}$. (Hint: the proof is a subset of the proof of Theorem 4.11.)
  (2) Show the following inhomogeneous variant of Theorem 4.11: if $P \in \Psi^m(X)$ is an operator with real principal symbol, and $Pu = f$, show that $\mathrm{WF}\,u \backslash \mathrm{WF}\,f$ is a union of maximally extended bicharacteristics of P.

Exercise 4.21.
  (1) What does Theorem 4.11 tell us about solutions to the Schrödinger equation? (Hint: not much.)
  (2) Nonetheless: let $\psi(t,x)$ be a solution to the Schrödinger equation on $\mathbb{R} \times X$ with (X, g) a Riemannian manifold; suppose that $\psi(0, x) = \psi_0 \in H^{1/2}(X)$. Define a set $S_1 \subset S^*X$ by

$$q \notin S_1 \iff \text{there exists } A \in \Psi^1(X), \ q \in \mathrm{ell}(A),$$

$$\text{such that } \int_0^1 \|A\psi\|^2 \, dt < \infty.$$

  (In other words, S_1 is a kind of wavefront set measuring where in the phase space S^*X we have $\psi \in L^2([0,1]; H^1(X))$—cf. Exercise 2.7.)
     Show that S_1 is invariant under the geodesic flow on S^*X. (See Exercise 4.14 for the definition of geodesic flow.)

(Hint: use (2.2) with A an appropriately chosen pseudodifferential operator of order zero, constructed much like the ones used in proving Theorem 4.11.)

Reflect on the following interpretation: "propagation of L^2H^1 regularity for the Schrödinger equation occurs at infinite speed along geodesics."

5. TRACES

It turns out to be of considerable interest in spectral geometry to consider the *traces* of operators manufactured from Δ, the Laplace-Beltrami operator on a compact[26] Riemannian manifold. The famous question posed by Kac [15], "Can one hear the shape of a drum," has a natural extension to this context: Recall from Exercise 3.11 that there exists an orthonormal basis ϕ_j of eigenfunctions of Δ with eigenvalues $\lambda_j^2 \to +\infty$; what, one wonders, can one recover of the geometry of a Riemannian manifold from the sequence of frequencies λ_j? Using PDE methods to understand traces of functions of the Laplacian has led to a better understanding of these inverse spectral problems.

Recall from Proposition 3.5 that $\sqrt{\Delta}$ is a first-order pseudodifferential operator on X. It is a slightly inconvenient fact that while $\sqrt{\Delta} \in \Psi^1(X)$, $\sqrt{\Delta} \notin \Psi^1(\mathbb{R} \times X)$: its Schwartz kernel is easily seen to be singular away from the diagonal. But this turns out be be of little practical importance for our considerations here: it is close enough!

Let us now consider the operator

$$(5.1) \qquad\qquad U(t) = e^{-it\sqrt{\Delta}}$$

which can be defined by the functional calculus to act as the scalar operator $e^{-it\lambda_j}$ on each ϕ_j. $U(t)$ is unitary, and indeed is the solution operator to the Cauchy problem for the equation

$$(5.2) \qquad\qquad (\partial_t + i\sqrt{\Delta})u = 0;$$

that is to say, if $u = U(t)f$, we have

$$(\partial_t + i\sqrt{\Delta})u = 0, \quad \text{and } u(0, x) = f(x).$$

Equation (5.2) is easily seen to be very closely related to the wave equation: if u solves (5.2) then applying $\partial_t - i\sqrt{\Delta}$, we see that u also satisfies the wave equation. Of course, (5.2) only requires a single

---

[26]We especially emphasize that X denotes a *compact* manifold throughout this section.

Cauchy datum, unlike the wave equation, so the trade-off is that the Cauchy data of u as a solution to $\Box u = 0$ are constrained: we have

$$u(0, x) = f(x), \quad \partial_t u(0, x) = -i\sqrt{\Delta} f.$$

The real and imaginary parts of the operator $U(t)$ are exactly the solution operators to the (more usual) Cauchy problem for the wave equation with $u(0, x) = f(x), \partial_t u(0, x) = 0$ and with $u(0, x) = 0, \partial_t u(0, x) = -i\sqrt{\Delta} f(x)$ respectively.

Why is the operator $U(t)$ of interest? Well, suppose that we are interested in the sequence of λ_j's. It makes sense to combine these numbers into a generating function, and certainly one option would be to take the exponential sum[27]

$$\sum_j e^{-it\lambda_j}$$

This is, at least formally, nothing but the trace of the operator $U(t)$. One of the principal virtues of this generating function is that if we let $N(\lambda)$ denote the "counting function"

$$N(\lambda) = \#\{\lambda_j \leq \lambda\},$$

then we have

$$N'(\lambda) = \sum_j \delta(\lambda - \lambda_j),$$

hence

$$\sum e^{-it\lambda_j} = (2\pi)^{n/2} \mathcal{F}_{\lambda \to t}(N'(\lambda))(t).$$

This is all a bit optimistic, as $U(t)$ is easily seen to be not of trace class—for example at $t = 0$ it is the identity. So we should try and think of $\operatorname{Tr} U(t)$ as a *distribution.* We do know that for any test function $\varphi(t) \in \mathcal{S}(\mathbb{R})$ and any $f \in L^2(X)$,

$$
\begin{aligned}
\int \varphi(t) U(t) f \, dt &= \int (1 + D_t^2)^{-k} (1 + D_t^2)^k (\varphi(t)) U(t) f \, dt \\
(5.3) \qquad &= \int (1 + D_t^2)^k (\varphi(t)) (1 + D_t^2)^{-k} U(t) f \, dt \\
&= \int (1 + D_t^2)^k (\varphi(t)) (1 + \Delta)^{-k} U(t) f \, dt,
\end{aligned}
$$

---

[27]This choice of generating function, corresponding to taking the wave trace, is of course one choice among many. Some other approaches include taking the trace of the complex powers of the Laplacian or the heat trace. The idea of using (at least some version of) the wave trace originates with Levitan and Avakumovič.

since $D_t^2 U = \Delta U$. Here we can, if we like, consider $(1 + \Delta)^{-k}$ to be defined by the functional calculus; it is in fact pseudodifferential, of order $-2k$. We easily obtain (using either point of view) the estimate:

$$(1 + \Delta)^{-k} U(t) : L^2(X) \to H^{2k}(X);$$

hence, for $k \gg 0$, the operator $(1 + \Delta)^{-k} U(t)$ is of trace class.

Exercise 5.1. Prove that this operator is of trace class for $k \gg 0$. (HINT: One easy route is to think about first choosing k large enough that the Schwartz kernel is continuous, hence the operator is Hilbert-Schmidt; then you can take k even larger to get a trace-class operator, by factoring into a product of two Hilbert-Schmidt operators (see Appendix).)

Equation (5.3) thus establishes that

$$\operatorname{Tr} U(t) : \varphi \mapsto \operatorname{Tr} \int \varphi(t) U(t) \, dt$$

makes sense as a distribution on $\mathbb{R}$. We can thus write

(5.4) $$\operatorname{Tr} U(t) = (2\pi)^{n/2} \mathcal{F}(N')(t).$$

where both sides are defined as distributions. Our next goal is to try to understand the left side of this equality through PDE methods.

Exercise 5.2. Show that if the Schwartz kernel $K(x, y)$ of a bounded, normal operator T on $L^2(X)$ is in $\mathcal{C}^k(X)$ for sufficiently large k, then T is of trace-class and

$$\operatorname{Tr} T = \int K(x, x) \, dg(x).$$

(HINT: Check that K is trace-class as in the previous exercise. Then apply the spectral theorem for compact normal operators, and use the basis of eigenfunctions of K when computing the trace. The crucial thing to check is that if φ_j are the eigenfunctions, then

$$\sum \varphi_j(x) \overline{\varphi_j(y)} = \delta_\Delta,$$

the delta-distribution at the diagonal, since this is nothing but a spectral resolution of the identity operator.)

As a consequence of Exercise 5.2, we can compute the distribution $\operatorname{Tr} U(t)$ in another way if we can compute the Schwartz kernel of $U(t)$. Indeed, knowing even rather crude things about $U(t)$ can give us some useful information here.

Theorem 5.1. *Let Φ_t be the geodesic flow, i.e. the flow generated by the Hamilton vector field of $|\xi|_g \equiv (\sum g^{ij} \xi_i \xi_j)^{1/2}$. Then*

$$\operatorname{WF} U(t) f = \Phi_t(\operatorname{WF} f).$$

We begin with a lemma:

Lemma 5.2. *Let* $(\partial_t + i\sqrt{\Delta})u = 0$. *Then*

$$(x_0, \xi_0) \in \mathrm{WF}\, u|_{t=t_0}$$

if and only if

$$(t = t_0, \tau = -|\xi_0|, x_0, \xi_0) \in \mathrm{WF}\, u.$$

Proof. [28] Suppose $q = (x_0, \xi_0) \in \mathrm{WF}\, u|_{t=t_0}$. Since $\tilde{q} = (t = t_0, \tau = -|\xi_0|, x_0, \xi_0)$ is the only vector in $\Sigma_{\partial_t + i\sqrt{\Delta}}$ that projects to (x_0, ξ_0), it must lie in the wavefront set of u by Exercise 4.11.

The converse is harder. Suppose $q \notin \mathrm{WF}\, u|_{t=t_0}$. Let $v = H(t - t_0)u$, with H denoting the Heaviside function. Then

$$(\partial_t + i\sqrt{\Delta})v = \delta(t - t_0)u(t_0, x) \equiv f.$$

and v vanishes identically for $t < t_0$. By the last part of Exercise 4.11,

$$\tilde{q} \notin \mathrm{WF}\, f,$$

hence (since $\mathrm{WF}\, f$ only lies over $t = t_0$) certainly no points along the bicharacteristic through $\tilde{q}$ lie in $\mathrm{WF}\, f$. Moreover, no points along this bicharacteristic lie in $\mathrm{WF}\, v$ for $t < t_0$ (since v is in fact zero there). Hence by the version of the propagation of singularities in the second part of Exercise 4.20, this bicharacteristic is absent from $\mathrm{WF}\, u$. In particular, $\tilde{q} \notin \mathrm{WF}\, u$. $\qquad\square$

Theorem 5.1 now follows directly[29] from the lemma and Theorem 4.11.

We now require a result on microlocal partitions of unity somewhat generalizing Lemma 4.1:

Exercise 5.3. Let ρ_j, $j = 1, \ldots, N$ be a smooth partition of unity for S^*X. Show that there exists $A_j \in \Psi^0(X)$ with $\mathrm{WF}'\, A_j = \operatorname{supp} \rho_j$, $\hat{\sigma}_0(A_j) = \rho_j$, $A_j^* = A_j$, and

$$\sum_{j=1}^{N} A_j^2 = \mathrm{Id} - R,$$

with $R \in \Psi^{-\infty}(X)$.

---

[28]I am grateful to András Vasy for showing me this proof.

[29]Here is one of the places where we should worry about the fact that $\sqrt{\Delta}$ is not a pseudodifferential operator on $\mathbb{R} \times X$. This problem is seen not to affect the proof of Hörmander's theorem if we note that composing $\sqrt{\Delta}$ with a pseudodifferential operator that is microsupported in a neighborhood of the characteristic set $\{|\tau| = -|\xi|_g\}$ yields an operator that *is* pseudodifferential, and that the symbol calculus extends to such compositions. (The author confesses that this is not entirely a trivial matter.)

For a distribution u, let singsupp u (the "singular support" of u) be the projection of its wavefront set, i.e. the complement of the largest open set on which it is in $\mathcal{C}^\infty$.

Theorem 5.3.

singsupp $\operatorname{Tr} U(t) \subseteq \{0\} \cup \{$*lengths of closed geodesics on X*$\}$.

This theorem is due to Chazarain and to Duistermaat-Guillemin. We begin with the following dynamical result:

Lemma 5.4. *Let L not be the length of any closed geodesic. Then there exists $\epsilon > 0$ and a cover U_i of $S^* X$ by open sets such that for $t \in (L - \epsilon, L + \epsilon)$, there exists no geodesic with start- and endpoints both contained in the same U_i.*

Exercise 5.4.

    (1) Prove the lemma. (HINT: The cosphere bundle is compact.)
    (2) As long as you're at it, show that 0 is an isolated point in the set of lengths of closed geodesics ("length spectrum"), and that the length spectrum is a closed set.

We now prove Theorem 5.3.

Proof. Let L not be the length of any closed geodesic on X. Let U_j be a cover of $S^* X$ as given by Lemma 5.4. Let ρ_j be a partition of unity subordinate to U_j and let A_j be a microlocal partition of unity as in Exercise 5.3. Then, calculating with distributions on $\mathbb{R}^1$, we have

$$\operatorname{Tr} U(t) = \sum_j \operatorname{Tr} A_j^2 U(t) + \operatorname{Tr} RU(t)$$

$$= \sum_j \operatorname{Tr} A_j U(t) A_j + \operatorname{Tr} RU(t)$$

and, more generally,

$$D_t^{2m} \operatorname{Tr} U(t) = \sum_j \operatorname{Tr} A_j \Delta^m U(t) A_j + \operatorname{Tr} R\Delta^m U(t).$$

Let u be a distribution on X; then $\operatorname{WF} A_j u \subseteq \operatorname{WF}' A_j \subset U_j$. Thus Theorem 5.1 gives

$$\operatorname{WF} \Delta^m U(t) A_j u \subseteq \Phi_t(U_j).$$

But by construction, this set is disjoint from U_j and hence from $\operatorname{WF}' A_j$. Hence for any m,[30]

$$A_j \Delta^m U(t) A_j \in L^\infty([L - \epsilon, L + \epsilon]; \Psi^{-\infty}(X));$$

---

[30]We technically have to work just a little to obtain the uniformity in time: observe that $A_j \Delta^m U(t) A_j$ are a continuous (or even smooth) *family* of smoothing

consequently,

$$D_t^{2m} \operatorname{Tr} U(t) \in L^\infty([L - \epsilon, L + \epsilon]). \quad \square$$

Exercise 5.5. Show that in the special case of $X = S^1$, Theorem 5.3 can be deduced from the Poisson summation formula. For this reason it is often referred to as the *Poisson relation*.

One is tempted to conclude from (5.4) and Theorem 5.3 that one can "hear" the lengths of closed geodesics on a manifold, since the right side of (5.4) is determined by the spectrum, and the left side seems to be a distribution from whose singularities we can read off the lengths of closed geodesics. The trouble with this approach is that we do not know with any certainty from Theorem 5.3 that the putative singularities in $\operatorname{Tr} U(t)$ at lengths of closed geodesics are actually there: perhaps the distribution is, after all, miraculously smooth. Thus, proving actual inverse spectral results requires somewhat more care, as we shall see. To this end, we will begin studying the operator $U(t)$ more constructively in the following section.

6. A parametrix for the wave operator

In order to learn more about the wave trace, we will have to bite the bullet and construct an approximation ("parametrix") for the fundamental solution to the wave equation on a manifold. The approach will have a similar iterative flavor to the technique we used to construct an approximate inverse for an elliptic operator, but we have now left the comfortable world of pseudodifferential operators: the parametrix we construct is going to be something rather different. Exactly what, and how to systematize the kinds of calculation we do here, will be discussed later on.

As this construction will be local, we will work in a single coordinate patch, which we identify with $\mathbb{R}^n$; for the sake of exposition, we omit the coordinate maps and partitions of unity necessary to glue this construction into a Riemannian manifold.

Consider once again the "half-wave equation"[31]

$$(6.1) \qquad\qquad (D_t + \sqrt{\Delta})u = 0$$

on $\mathbb{R}^n$, where Δ is the Laplace-Beltrami operator with respect to a metric g. Our goal is to find a distribution u approximately solving

---

operators. We have been avoiding the topological issues necessary to easily dispose of such matters, however.

[31]Remember that $D_t = i^{-1}\partial_t$.

(6.1) with initial data

$$u(0, x, y) = \delta(x - y)$$

for any $y \in \mathbb{R}^n$. Recall that if we let U denote the exact solution to (6.1) with initial data $\delta(x - y)$ then U can also be interpreted as (the Schwartz kernel of) the "solution operator" mapping initial data f to the solution $e^{-it\sqrt{\Delta}}f$ with that initial data, evaluated at time t; this is why we denote it U, as we did above, and why we will often think of our parametrix $u(t, x, y)$ as a family in t of integral kernels of operators on $\mathbb{R}^n$.

We do not expect $U(t, x, y)$ or our parametrix for it to be the Schwartz kernel of a pseudodifferential operator, as it moves wavefront set around, by Theorem 4.11; recall that pseudodifferential operators are *microlocal,* which is to say they don't do that. But we will try and construct our parametrix $u(t, x, y)$ as something of *roughly* the same form, which is to say as an oscillatory integral

$$u(t, x, y) = \int a(t, x, \eta)e^{i\Phi} \, d\eta$$

where the main difference is that the "phase function" $\Phi = \Phi(t, x, y, \eta)$ will be something a good deal more interesting than $(x - y) \cdot \eta$; indeed, this phase function is where all the geometry of the problem turns out to reside.

First, let's write our initial data as an oscillatory integral:

$$\delta(x - y) = (2\pi)^{-n} \int e^{i(x-y)\cdot\eta} \, d\eta.$$

Let us now try, as an Ansatz, modifying the phase as it varies in t, x by setting

(6.2) $$u(t, x, y) = (2\pi)^{-n} \int a(t, x, \eta)e^{i(\phi(t,x,\eta)-y\cdot\eta)} \, d\eta;$$

then if $\phi(0, x, \eta) = x \cdot \eta$ and $a(0, x, \eta) = 1$, we recover our initial data; moreover, if ϕ were to remain unchanged as t varied we would have nothing but a family of pseudodifferential operators. Let us assume that a is a classical symbol of order 0 in η, so that we have an asymptotic expansion

$$a \sim a_0 + |\eta|^{-1}a_{-1} + |\eta|^{-2}a_{-2} + \dots, \quad a_j = a_j(t, x, \hat{\eta}).$$

Let us further assume that ϕ is homogeneous in η of degree 1, hence matches the homogeneity[32] of $x \cdot \eta$.

---

[32]That is is then likely to be singular at $\eta = 0$ will not in fact concern us, as it will turn out that we may as well assume that a vanishes near $\eta = 0$.

Now if u solves the half-wave equation, it solves the wave equation, hence we have

$$\Box u = 0;$$

As we seek an approximate solution, we will instead accept

$$\Box u \in \mathcal{C}^\infty((-\epsilon, \epsilon)_t \times \mathbb{R}^n).$$

Our strategy is to plug (6.2) into this equation and see what is forced upon us. To this end, note that if we have an expression

(6.3) $$v = (2\pi)^{-n} \int b(t, x, y, \eta) e^{i(\phi(t,x,\eta) - y \cdot \eta)} \, d\eta;$$

where b is a symbol of order $-\infty$, then v lies in $\mathcal{C}^\infty$, as the integral converges absolutely, together with all its t, x, y derivatives. So terms of this form will be acceptable errors.

Applying $\Box$ to (6.2), we group terms according to their order in η. The "worst case" terms involve factors of η^2, and can only be produced by second-order terms in $\Box$, with all derivatives falling on the exponential term. Since the second-order terms in Δ are just

$$\sum g^{ij}(x) D_i D_j,$$

we can write the term this produces from the phase as $|d_x \phi|_g^2$ or, equivalently, $|\nabla_x \phi|_g^2$. Thus, the equation that we need to solve to make the η^2 terms vanish is just

(6.4) $$(\partial_t \phi)^2 - |\nabla_x \phi|_g^2 = 0.$$

Recall that we further want our phase to agree with the standard pseudodifferential one at time zero, i.e. we want

(6.5) $$\phi(0, x, \eta) = x \cdot \eta.$$

Combining this information with (6.4) we easily see that we in particular have

$$(\partial_t \phi|_{t=0})^2 = |\eta|_g^2,$$

and we need to make an arbitrary choice of sign in solving this to get the initial time-derivative: we will choose[33]

(6.6) $$\partial_t \phi|_{t=0} = -|\eta|_g.$$

If our metric is the Euclidean metric, we can easily solve (6.4), (6.5), and (6.6) by setting

$$\phi(t, x, \eta) = x \cdot \eta - t|\eta|.$$

---

[33]We will use this solution for reasons that will become apparent presently—it is the right one to solve (5.2) and not merely the wave equation.

More generally, the construction of a phase satisfying (6.4),(6.5) and (6.6) is the classic construction of Hamilton-Jacobi theory, and is sketched in the following exercise.

Exercise 6.1.
  (1) Show that equation (6.4) is equivalent to the statement that for each η, the graph of $d_{t,x}\phi(t, x, \eta)$ is contained in the set

$$\Lambda = \{\tau^2 - |\xi|_g^2 = 0\} \subset T^*(\mathbb{R}_t \times \mathbb{R}_x^n)$$

  (where the variables τ and ξ are the canonical dual variables to t and x respectively). The condition (6.5) implies

$$d_x\phi(t, x, \eta)|_{t=0} = \eta \cdot dx.$$

  Equation (6.6) gives further

(6.7) $$d_{t,x}\phi(t, x, \eta)|_{t=0} = -|\eta|\, dt + \eta \cdot dx;$$

  accordingly, for fixed η, let

$$G_0 = \{t = 0, x \in \mathbb{R}^n, \tau = -|\eta|, \xi = \eta\} \subset T^*(\mathbb{R} \times \mathbb{R}^n).$$

  (2) Let H denote the Hamilton vector field of $\tau^2 - |\xi|_g^2$. Show that flow along H preserves Λ and that H is transverse to G_0.
  (3) Show that there is a solution to (6.4),(6.7) for $t \in (-\epsilon, \epsilon)$ where the graph of $d_{t,x}\phi$ is given by flowing out the set G_0 under H. (Among other things, you need to check that the resulting smooth manifold is indeed the graph of the differential of a function.) Show that this solution can be integrated to give a solution to (6.4),(6.5).

Employing the phase ϕ constructed in Exercise 6.1, we have now solved away the homogeneous degree-two (in η) terms in the application of $\square$ to our parametrix. We thus move on to the degree-one terms, which are as follows:

(6.8) $$2D_t\phi D_t a_0 - 2\langle D_x\phi, D_x\rangle_g a_0 + r_1(t, x, y, \eta)$$

where r_1 is a homogeneous function of degree 1 *independent of a_0,* i.e. determined completely by ϕ. Given that ϕ solves the eikonal equation, we can rewrite (6.8) by factoring out $|\nabla_x\phi|$ and noting that our sign choice $\partial_t\phi = -|\nabla_x\phi|$ must persist away from $t = 0$ (for a short time, anyway). In this way we obtain

$$2\partial_t a_0 + 2\left\langle \frac{\nabla_x\phi}{|\nabla_x\phi|_g}, \partial_x \right\rangle_g a_0 - \tilde{r}_1 = 0,$$

with $\tilde{r}_1$ homogeneous of degree 0. This is a *transport equation* that we would like to solve, with the initial condition $a_0(0, x, y, \eta) = 1$ (the

symbol of the identity operator). We can easily see that a solution exists with the desired initial condition $a_0(0, y, \eta) = 1$, as, letting

$$\mathsf{H} = 2\partial_t + 2\Big\langle \frac{\nabla_x\phi}{|\nabla_x\phi|_g}, \partial_x \Big\rangle_g$$

we see that H is a nonvanishing vector field, transverse to $t = 0$, hence we may solve

$$\mathsf{H}a_0 = \tilde{r}_1, \quad a_0|_{t=0} = 1$$

by standard ODE methods.

Now we consider degree-zero terms in η. We find that they are of the form

$$2D_t\phi D_t a_{-1} - 2\langle D_x\phi, D_x\rangle_g a_{-1} + r_0(t, x, y, \eta)$$

where r_0 only depends on a_0 and ϕ (i.e. not on a_{-1}). Thus, we may use the same procedure as above to find a_{-1} with initial value zero, making the degree-zero term vanish. (Note that the vector field H along which we need to flow remains the same as in the previous step.)

We continue in this manner, solving successive transport equations along the flow of H so as to drive down the order in η of the error term. Finally we Borel sum the resulting symbols, obtaining a symbol

$$a(t, x, \eta) \in S_{\mathrm{cl}}^0(\mathbb{R}_{x,y}^{2n} \times \mathbb{R}_\eta^n)$$

such that

$$a(0, x, \eta) = 1,$$

and

$$(6.9) \quad \Box u = \Box\left((2\pi)^{-n}\int a(t, x, \eta)e^{i(\phi(t,x,\eta) - y\cdot\eta)}\,d\eta\right)$$

$$= (2\pi)^{-n}\int b(t, x, y, \eta)e^{i(\phi(t,x,\eta) - y\cdot\eta)}\,d\eta \in \mathcal{C}^\infty((-\epsilon, \epsilon) \times X),$$

since $b \in S^{-\infty}$.

Now we need to check that (6.9) implies that in fact u differs by a smooth term from the actual solution. We will show soon (in the next section) that our choice of the phase implies that[34] $\mathrm{WF}\, u \subset \{\tau < 0\}$. Hence, using this fact, we have

$$(6.10) \qquad (\partial_t - i\sqrt{\Delta})(\partial_t + i\sqrt{\Delta})u = f \in \mathcal{C}^\infty;$$

Now $\partial_t - i\sqrt{\Delta}$ is elliptic on $\tau < 0$, so, letting Q denote a microlocal elliptic parametrix, we have

$$Q(\partial_t - i\sqrt{\Delta}) = I + E$$

---

[34]This can also be verified directly, with localization, Fourier transform, and elbow grease.

with $\mathrm{WF}' E \cap \mathrm{WF}\, u = \emptyset$. Thus, applying Q to both sides of (6.10), we have

$$(\partial_t + i\sqrt{\Delta})u \in \mathcal{C}^\infty.$$

Also, as we have arranged that $a(0, x, \eta) = 1$, we have got our initial data exactly right: $u(0, x, y) = \delta(x - y)$. Letting U denote the actual solution operator to (5.2), we thus find

$$(\partial_t + i\sqrt{\Delta})(u - U) \in \mathcal{C}^\infty, \quad u(0, x, y) - U(0, x, y) = 0;$$

hence by global energy estimates[35] we have

$$u - U \in \mathcal{C}^\infty((-\epsilon, \epsilon) \times \mathbb{R}^n).$$

7. The wave trace

Our treatment of this material (and, in part, that of the previous section) closely follows the treatment in [7], which is in turn based on work of Hörmander [9].

Recall that, if $N(\lambda) = \#\{\lambda_j \leq \lambda\}$ and $U(t)$ is given by (5.1), then

$$(7.1) \qquad \mathrm{Tr}\, U(t) = (2\pi)^{n/2}\mathcal{F}(N'(\lambda)).$$

Thus, the singularities of $\mathrm{Tr}\, U(t)$ are related to the growth of $N(\lambda)$. We think that $\mathrm{Tr}\, U(t)$ should have singularities at zero, together with lengths of closed geodesics; since $U(0)$ is the identity (which has a very divergent trace), the singularity at $t = 0$, at least, seems certain to appear. We will thus spend some time discussing this singularity of the wave trace and its consequences for spectral geometry.

What is the form of the singularity of $\mathrm{Tr}\, U(t)$ at $t = 0$? Our parametrix from the previous section was

$$u(t, x, y) = (2\pi)^{-n} \int a(t, x, \eta)e^{i(\phi(t, x, \eta) - y \cdot \eta)} \, d\eta,$$

where $\phi(t, x, \eta) = x \cdot \eta - t|\eta|_{g(x)} + O(t^2)$, and $a(t, x, \eta) = 1 + O(t)$. Thus,

$$(7.2) \qquad u(t, x, x) = (2\pi)^{-n} \int a(t, x, \eta)e^{i(-t|\eta|_{g(x)} + O(t^2|\eta|))} \, d\eta,$$

where we have used the homogeneity of the phase in writing the error term as $O(t^2|\eta|)$.

Formally, we would now like to conclude that the singularity at $t = 0$ is approximately that of

$$u(t, x, x) = (2\pi)^{-n} \int e^{-it|\eta|_{g(x)}} \, d\eta$$

---

[35]We can either use the estimates developed in §2.3, adapted to this variable coefficient setting, and with a power of the Laplacian applied to the solution (in order to gain derivatives); or we can apply Theorem 4.11, which is overkill.

so that integrating in x would give, if all goes well,

$$\operatorname{Tr} U(t) \sim \int u(t, x, x) \, dx$$

$$\sim (2\pi)^{-n} \iint e^{-it|\eta|_g} \, d\eta \, dx$$

(7.3)
$$= (2\pi)^{-n} \iiint_{\sigma > 0, |\theta| = 1} e^{-it\sigma|\theta|_g} \sigma^{n-1} \, d\sigma \, d\theta \, dx$$

$$= (2\pi)^{-n/2} \iint \mathcal{F}(\sigma^{n-1} H(\sigma))(t|\theta|_g) \, d\theta \, dx,$$

with H denoting the Heaviside function. (Recall that the notation $f \sim g$ means that $(f/g) \to 1$, in this case as $t \to 0$.) If we crudely try to solve (7.1) for $N'(\lambda)$ by applying an inverse Fourier transform to $\operatorname{Tr} U(t)$ and pretending that the singularity of $\operatorname{Tr} U(t)$ at $t = 0$ is all that matters, we find, formally, that (7.3) yields

$$N'(\lambda) \sim (2\pi)^{-n/2} \mathcal{F}^{-1}_{t \to \lambda} \operatorname{Tr} U(t)$$

$$\sim (2\pi)^{-n} \iint_{|\theta| = 1} |\theta|_g^{-1} \left(\frac{\lambda}{|\theta|_g} \right)^{n-1} d\theta \, dx$$

$$= (2\pi)^{-n} \lambda^{n-1} \iint_{|\theta| = 1} |\theta|_g^{-n} \, d\theta \, dx.$$

Integrating would formally yield

$$N(\lambda) \sim (2\pi)^{-n} \frac{\lambda^n}{n} \iint_{|\theta| = 1} |\theta|_g^{-n} \, d\theta \, dx$$

$$= (2\pi)^{-n} \lambda^n \iiint_{|\theta| = 1, \rho \in (0,1)} |\theta|_g^{-n} \rho^{n-1} \, d\rho \, d\theta \, dx$$

$$= (2\pi)^{-n} \lambda^n \iiint_{|\sigma\theta|_g < 1} \sigma^{n-1} \, d\sigma d\theta \, dx,$$

where we have, in the last line, set $\sigma = \rho/|\theta|_g$, with the result that definition of the region of integration now involves the metric. This last quantity can easily be seen to be simply the volume in phase space of the set $|\xi|_g < 1$, otherwise known as the unit ball bundle.[36] Thus, we obtain *formally*

$$N(\lambda) \sim (2\pi)^{-n} \lambda^n \operatorname{Vol}(B^* X) = (2\pi)^{-n} \operatorname{Vol}(\{|\xi|_g < \lambda\}).$$

---

[36]Recall that on a symplectic manifold (N^{2n}, ω) we have a naturally defined volume form ω^n, and it is this volume that we are integrating over the unit ball here.

This is all nonsense, of course, for several different reasons. First, we were very imprecise about dropping higher order terms in t in computing the asymptotics of the trace as $t \to 0$. Furthermore, we formally computed with N' as if it were a smooth function, but of course N' is quite singular (a sum of delta distributions). Moreover, and potentially most seriously, there are in general infinitely many singularities in $\mathrm{Tr}\, U(t)$ that might be contributing to the asymptotic behavior of its Fourier transform: we have been concerning ourselves only with the one near $t = 0$. However: the above argument does give the right leading order asymptotics, the so-called "Weyl Law." What follows is (the outline of) a rigorous version of the above argument.

To begin, we need a cutoff function to localize us near the singularity at $t = 0$, where our parametrix is valid.

Exercise 7.1. Show that there exists $\rho \in \mathcal{S}(\mathbb{R})$ with $\hat{\rho}$ compactly supported, $\hat{\rho}(0) = 1$, $\hat{\rho}(t) = \hat{\rho}(-t)$, $\rho(\lambda) > 0$ for all λ, and $\hat{\rho}$ supported in an arbitrarily small neighborhood of 0. (HINT: Start with a smooth, compactly supported $\hat{\rho}$; convolve with its complex conjugate, and scale.)

We now consider

$$\mathcal{F}^{-1}_{t \to \lambda}\big(\hat{\rho}(t)\, \mathrm{Tr}\, u(t)\big)$$

$$= (2\pi)^{-n-1/2} \iiint \hat{\rho}(t) a(t, x, \eta) e^{i(t(\lambda - |\eta|_g) + O(t^2|\eta|))}\, dx\, d\eta\, dt$$

$$= (2\pi)^{-n-1/2} \iiint \hat{\rho}(t) a(t, x, \lambda\sigma\theta) e^{it\lambda(1 - \sigma + O(t^2\sigma))} (\lambda\sigma)^{n-1}\, dx\, d\sigma\, d\theta\, dt;$$

here we have used the change of variables $\eta = \lambda\sigma\theta$ with $|\theta| = 1$. We now employ the *method of stationary phase* to estimate the asymptotics of the integral in t, σ. If $\hat{\rho}$ is chosen supported sufficiently close to the origin, then the unique stationary point on the support of the amplitude is at $\sigma = 1$, $t = 0$; we thus obtain a complete asymptotic expansion in λ beginning with the terms

$$A\lambda^{n-1} + O(\lambda^{n-2})$$

where

$$A = n(2\pi)^{-n} \mathrm{Vol}(B^*X).$$

Exercise* 7.2. Do this stationary phase computation. If you don't know about the method of stationary phase, this is your chance to learn it, e.g. from [11].

Thus, since $u - U \in \mathcal{C}^\infty((-\epsilon, \epsilon) \times \mathbb{R}^n)$, (7.1) yields

Proposition 7.1.
$$(\rho * N')(\lambda) \sim A\lambda^{n-1} + O(\lambda^{n-2}).$$

We now try to make a "Tauberian" argument to extract the desired asymptotics of $N(\lambda)$ from this estimate.

Lemma 7.2.
$$N(\lambda + 1) - N(\lambda) = O(\lambda^{n-1}).$$

Proof. By Proposition 7.1 and since $N'(\lambda) = \sum \delta(\lambda - \lambda_j)$, we have
$$\sum \rho(\lambda - \lambda_j) \sim A\lambda^{n-1} + O(\lambda^{n-2});$$

thus, by positivity of $\rho(\lambda)$,
$$\big(\inf_{[-1,1]} \rho \big) \,(\#\{\lambda_j : \lambda - 1 < \lambda_j < \lambda + 1\}) \leq \sum \rho(\lambda - \lambda_j) = O(\lambda^{n-1}),$$

and the estimate follows as the infimum is strictly positive. $\qquad\square$

This yields at least a crude estimate:

Corollary 7.3.
$$N(\lambda) = O(\lambda^n).$$

A more technically useful result is:

Corollary 7.4.
$$N(\lambda - \tau) - N(\lambda) \lesssim \langle \tau \rangle^n \langle \lambda \rangle^{n-1}.$$

Exercise 7.3. Prove the corollaries. (For the latter, begin with the intermediate estimate $\langle \tau \rangle \langle |\lambda| + |\tau| \rangle^{n-1}$.)

Now we work harder.

Exercise 7.4. Show that we can antidifferentiate the convolution to get
$$\int_{-\infty}^{\lambda} (\rho * N')(\mu) \, d\mu = (\rho * N)(\lambda).$$

As a result, we of course have
$$(\rho * N)(\lambda) = A\lambda^n/n + O(\lambda^{n-1}) = B\lambda^n + O(\lambda^{n-1})$$
where $B = A/n = (2\pi)^{-n} \operatorname{Vol}(B^*X)$.

Thus, since $\int \rho(\mu) \, d\mu = 1$,
$$N(\lambda) = (N * \rho)(\lambda) - \int (N(\lambda - \mu) - N(\lambda))\rho(\mu) \, d\mu$$
$$= B\lambda^n + O(\lambda^{n-1}) - \int O(\langle \mu \rangle^n \langle \lambda \rangle^{n-1})\rho(\mu) \, d\mu$$
$$= B\lambda^n + O(\lambda^{n-1}),$$

where we have used Corollary 7.4 in the penultimate equality. We record what we have now obtained as a theorem, better known as Weyl's law with remainder term. This form of the remainder term is sharp, and not so easy to obtain by other means.

Theorem 7.5.

$$N(\lambda) = (2\pi)^{-n} \operatorname{Vol}(B^*X)\lambda^n + O(\lambda^{n-1}).$$

As noted above, it is perhaps suggestive to view the main term as the volume of the sublevel set in phase space $\{(x,\xi) : \sigma(\Delta)(x,\xi) \leq \lambda^2\}$. Weyl's law is one of the most beautiful instances of the quantum-classical correspondence, in which we can deduce something about a quantum quantity (the counting function for eigenvalues, also known as energy levels) in terms of a classical quantity, in this case the volume of a region of phase space.

Exercise* 7.5. Show that the error term in Weyl's law is sharp on spheres.

8. Lagrangian distributions

The form of the parametrix that we used for the wave equation turns out to be a special case of a very general and powerful class of distributions, known as *Lagrangian distributions*, introduced by Hörmander. Here we will give a very sketchy introduction to the general theory of Lagrangian distributions, and see both how it systematizes and extends our parametrix construction for the wave equation and how (in principle, at least) it can be made to yield the Duistermaat-Guillemin trace formula, which gives us an explicit description of the singularities of the wave trace.

We begin with a special case of the theory.

8.1. **Conormal distributions.** Let X be a smooth manifold of dimension n and let Y be a submanifold of codimension k. The *conormal distributions* with respect to Y are a special class of distributions having wavefront set[37] in the conormal bundle of Y, N^*Y. Let us suppose that Y is locally cut out by defining functions $\rho_1, \ldots, \rho_k \in \mathcal{C}^\infty(X)$, i.e. that (at least locally), $\{\rho_1 = \cdots = \rho_k = 0\} = Y$, and $d\rho_1, \ldots, d\rho_k$ are linearly independent on Y. Then we may (locally) extend the ρ_j's to a complete coordinate system

$$(x_1, \ldots x_k, y_1, \ldots y_{n-k})$$

---

[37]Recall that we have defined the wavefront set to lie in S^*X but it is often convenient to regard it as a *conic* subset of $T^*X \backslash o$, with o denoting the zero-section.

with

$$x_1 = \rho_1, \ldots, x_k = \rho_k,$$

so that $Y = \{x = 0\}$. In these coordinates, how might we write down some distributions with wavefront set lying only in N^*Y? Well, we can try to make things that are singular in the x variables at $x = 0$, with the y's behaving like smooth parameters. How do we create singularities at $x = 0$? One very nice answer is in the following:

Lemma 8.1. *Let $a(\xi) \in S_{cl}^m(\mathbb{R}_\xi^k)$ for some m. Then* $\operatorname{WF} \mathcal{F}^{-1}(a) \subseteq N^*(\{0\})$.

Proof. Writing

$$\mathcal{F}^{-1}(a)(x) = (2\pi)^{-k/2} \int a(\xi) e^{i\xi \cdot x} \, d\xi,$$

we first note that

$$\mathcal{F}^{-1}(a)(x) \in H^{-m-k/2-\epsilon}(\mathbb{R}^k)$$

for any $a \in S_{cl}^m$ and for all $\epsilon > 0$. Moreover for all j,

$$\begin{aligned}
(x^i D_{x^j}) \mathcal{F}^{-1}(a)(x) &= (2\pi)^{-k/2} \int a(\xi)(x^i D_{x^j}) e^{i\xi \cdot x} \, dx \\
&= (2\pi)^{-k/2} \int x^i \xi_j a(\xi) e^{i\xi \cdot x} \, d\xi \\
&= (2\pi)^{-k/2} \int \xi_j a(\xi) D_{\xi_i} e^{i\xi \cdot x} \, d\xi \\
&= -(2\pi)^{-k/2} \int D_{\xi_i}(\xi_j a(\xi)) e^{i\xi \cdot x} \, d\xi,
\end{aligned}$$

where we have integrated by parts in the final line. Note that if $a \in S_{cl}^m$ then $D_{\xi_i}(\xi_j a(\xi)) \in S_{cl}^m$ too (cf. Exercise 3.4). Thus we also have

$$(x^i D_{x^j}) \mathcal{F}^{-1}(a)(x) \in H^{-m-k/2-\epsilon}(\mathbb{R}^k).$$

Iterating this argument gives

(8.1) $$(x_{i_1} D_{x_{j_1}}) \ldots (x_{i_l} D_{x_{j_l}}) \mathcal{F}^{-1}(a)(x) \in H^{-m-k/2-\epsilon}(\mathbb{R}^k).$$

for all choices of indices and all $l \in \mathbb{N}$. Thus $\mathcal{F}^{-1}a$ is smooth[38] away from $x = 0$. $\qquad\square$

By the same token, we have more generally,

---

[38]We are of course proving more than the lemma states here: (8.1) gives a more precise "conormality" estimate that is valid uniformly across the origin.

Proposition 8.2. *Let $\rho_1, \ldots, \rho_k$ be (local) defining functions for $Y \subset X$ and let*

$$(8.2) \qquad\qquad a \in S_{cl}^{m+(n-2k)/4}(\mathbb{R}_x^n \times \mathbb{R}_\xi^k)$$

be compactly supported in x. Then

$$(8.3) \qquad u(x) = (2\pi)^{-(n+2k)/4} \int_{\mathbb{R}^k} a(x, \theta) e^{i(\rho_1\theta_1 + \cdots + \rho_k\theta_k)} \, d\theta$$

has wavefront set contained in N^*Y. Moreover there exists $s \in \mathbb{R}$ such that if $V_1, \ldots V_l$ are vector fields tangent to Y, then

$$V_1 \ldots V_l u \in H^s.$$

Exercise 8.1. Prove the proposition. You will probably find it helpful to change to a coordinate system $(x_1, \ldots, x_k, y_1, \ldots, y_{n-k})$ in which $x_1, \ldots, x_k = \rho_1, \ldots, \rho_k$. Note that in this coordinate system, any vector field tangent to Y can be written

$$\sum a_{ij}(x, y) x^i \partial_{x^j} + \sum b_j(x, y) \partial_{y^j}.$$

What values of s, the Sobolev exponent in the proposition, are allowable?

Definition 8.3. A distribution $u \in \mathcal{D}'(X)$ is a *conormal distribution* with respect to Y, of order m, if it can (locally) be written in the form (8.3) with symbol as in (8.2).

While it may appear that the definition of conormal distributions depends on the choice of the defining functions ρ_j, this is in fact not the case. The rather peculiar-looking convention on the orders of distributions is not supposed to make much sense just yet.

Note that examples of conormal distributions include $\delta(x) \in \mathbb{R}^n$ (conormal with respect to the origin), and more generally, delta distributions along submanifolds. Also quite pertinent is the example of pseudodifferential operators: if $A = \mathrm{Op}_\ell(a) \in \Psi^m(X)$ then the Schwartz kernel of A is a conormal distribution with respect to the diagonal in $X \times X$, of order m. (This goes at least some of the way to explaining the convention on orders.) Indeed, we could (at some pedagogical cost) simply have introduced conormal distributions and then used the notion to define the Schwartz kernels of pseudodifferential operators in the first place.

8.2. **Lagrangian distributions.** We now introduce a powerful generalization of conormal distributions, the class of *Lagrangian distributions*.[39] We begin by introducing some underlying geometric notions.

---

[39]These were first studied by Hörmander [10].

An important notion from symplectic geometry is that of a *La-grangian submanifold* $\mathcal{L}$ of a symplectic manifold N^{2n}. This is a sub-manifold of dimension n on which the symplectic form vanishes. We can always find local coordinates in which the symplectic form is given by $\omega = \sum dx^i \wedge dy^i$ and $\mathcal{L} = \{y = 0\}$, so there are no interesting local invariants of Lagrangian manifolds.

A *conic Lagrangian manifold* in T^*X is a Lagrangian submanifold of $T^*X \backslash o$ that is invariant under the $\mathbb{R}^+$ action on the fibers. (Here, o denotes the zero-section.)

Among the most important examples of conic Lagrangians are the following: let $Y \subset X$ be any submanifold; then $N^*Y \subset T^*X$ is a conic Lagrangian.

Exercise 8.2. Verify this.

The trick to defining Lagrangian distributions is to figure out how to associate a *phase function* ϕ with a conic Lagrangian $\mathcal{L}$ in T^*X.

Definition 8.4. A nondegenerate phase function is a smooth function $\phi(x, \theta)$, locally defined on a coordinate neighborhood of $X \times \mathbb{R}^k$, such that ϕ is homogeneous of degree 1 in θ and such that the differentials $d(\partial\phi/\partial\theta_j)$ are linearly independent on the set

$$C = \left\{ (x, \theta) : \frac{\partial\phi}{\partial\theta_j} = 0 \text{ for all } j = 1, \ldots, k \right\}.$$

The phase function is said to locally *parametrize* the conic Lagrangian $\mathcal{L}$ if

$$C \ni (x, \theta) \mapsto (x, d_x\phi)$$

is a local diffeomorphism from C to $\mathcal{L}$.

Exercise 8.3.

(1) Show that, in the notation of the definition above, C is au-tomatically a manifold, and the map $C \ni (x, \theta) \mapsto (x, d_x\phi)$ is automatically a local diffeomorphism from C to its image, which is a conic Lagrangian.

(2) Show that if ρ_j are definining functions for $Y \subset X$ then

$$\phi = \sum \rho_j \theta_j$$

is a nondegenerate parametrization of N^*Y.

(3) What Lagrangian is parametrized by the phase function used in our parametrix for the half-wave operator in the Euclidean case, given by

$$\phi(t, x, y, \theta) = (x - y) \cdot \theta - t|\theta|?$$

It turns out that every conic Lagrangian manifold has a local parametrization; the trouble is, in fact, that it has lots of them.

Definition 8.5. A *Lagrangian distribution* of order m with respect to the Lagrangian $\mathcal{L}$ as one that is given, locally near any point in X, by a finite sum of oscillatory integrals of the form

$$(2\pi)^{-(n+2k)/4} \int_{\mathbb{R}^k} a(x,\theta) e^{i\phi(x,\theta)} \, d\theta$$

where

$$a \in S_{\text{cl}}^{m+(n-2k)/4}(\mathbb{R}_x^n \times \mathbb{R}_\theta^k)$$

and where ϕ is a nondegenerate phase function parametrizing $\mathcal{L}$. Let $I^m(X, \mathcal{L})$ denote the space of all Lagrangian distributions on X with respect to $\mathcal{L}$ of order m.

Note that the connection between k, the number of phase variables, and the geometry of $\mathcal{L}$ is not obvious; indeed, it turns out that we have some choice in how many phase variables to use. As there are many different ways to parametrize a given conic Lagrangian manifold, one tricky aspect of the theory of Lagrangian distributions is necessarily the proof that using different parametrizations (possibly involving different numbers of phase variables) gives us the same class of distributions.

The analogue of the *iterated regularity* property of conormal distributions, i.e. our ability to repeatedly differentiate along vector fields tangent to Y, turns out to be as follows:

Proposition 8.6. *Let $u \in I^m(X, \mathcal{L})$. There exists s such that for any $l \in \mathbb{N}$ and for any $A_1, \ldots, A_l \in \Psi^1(X)$ with $\sigma_1(A_j)|_{\mathcal{L}} = 0$, we have*

$$A_1 \ldots A_l u \in H^s(X).$$

Of course, once this holds for one s, it holds for all smaller values; the precise range of possible values of s is related to the order m of the Lagrangian distribution; we will not pursue this relationship here, however. This iterated regularity property of Lagrangian distributions completely characterizes them if we use "Kohn-Nirenberg" symbols (as in Exercise 3.4) instead of "classical" ones (see [14]).

8.3. **Fourier integral operators.** Fourier integral operators ("FIO's") quantize classical *maps* from a phase space to itself just as pseudodifferential operators quantize classical *observables* (i.e. functions on the phase space). The maps from phase space to itself that we may quantize in this manner are the *symplectomorphisms*, exactly the class of transformations of phase space that arise in classical mechanics. We

recall that a symplectomorphism between symplectic manifolds is a diffeomorphism that preserves the symplectic form. We further define a *homogeneous symplectomorphism* from T^*X to T^*X to be one that is homogeneous in the fiber variables, i.e. commutes with the $\mathbb{R}^+$ action on the fibers.

An important class of homogeneous symplectomorphisms is those obtained as follows:

Exercise 8.4. Show that the time-1 flowout of the Hamilton vector field of a homogeneous function of degree 1 on T^*X is a homogeneous symplectomorphism.

Given a homogeneous symplectomorphism $\Phi : T^*X \to T^*X$, consider its *graph* $\Gamma_\Phi \subset (T^*X \backslash o) \times (T^*X \backslash o)$. Since Φ is a symplectomorphism, we have
$$\iota^* \pi_L^* \omega = \iota^* \pi_R^* \omega,$$
where ι is inclusion of Γ_Φ in $(T^*X \backslash o) \times (T^*X \backslash o)$, and $\pi_\bullet$ are the left and right projections. If we alter Γ_Φ slightly, forming
$$\Gamma_\Phi' = \{(x_1, \xi_1, x_2, \xi_2) : (x_1, \xi_1, x_2, -\xi_2) \in \Gamma_\Phi\},$$
and let ι' denote the inclusion of this manifold, then we find that a sign is flipped, and
$$(\iota')^* \pi_L^* \omega + (\iota')^* \pi_R^* \omega = 0;$$
since $\Omega = (\pi_L^* \omega + \pi_R^* \omega)$ is just the symplectic form on
$$T^*(X \times X) = T^*X \times T^*X,$$
we thus find that Γ_Φ' is *Lagrangian* in $T^*(X \times X)$. In fact, it is easily to verify that given a diffeomorphism Φ, Γ_Φ' is Lagrangian if and only if Φ is a symplectomorphism.

Exercise 8.5. Check this.

Now we simply define the class of Fourier integral operators (of order m) associated with the symplectomorphism Φ of X to be those operators from smooth functions to distributions whose Schwartz kernels lie in the Lagrangian distributions
$$I^m(X \times X, \Gamma_\Phi').$$
It would be nice if this class of operators turned out to have good properties such as behaving well under composition, as pseudodifferential operators certainly do. We note right off the bat that these operators *include* pseudodifferential operators, as well as a number of other, familiar examples:

(1) $\Psi^m(X) = I^m(X \times X, \Gamma_{\mathrm{Id}}')$.

(2) In $\mathbb{R}^n$, fix α and let $Tf(x) = f(x - \alpha)$ Then T has Schwartz kernel

$$\delta(x - x' - \alpha)$$

which is clearly conormal of order zero at $x - x' - \alpha = 0$. Note that this is certainly *not* a pseudodifferential operator, as it moves wavefront around; indeed, it is associated with the symplectomorphism $\Phi(x, \xi) = (x + \alpha, \xi)$, and it it no coincidence that

$$\mathrm{WF}\, Tf = \Phi(\mathrm{WF}\, f).$$

(3) As a generalization of the previous example, note that if $\phi : X \to X$ is a diffeomorphism, then we may set

$$Tf(x) = f(\phi(x));$$

this is a FIO associated to the homogeneous symplectomorphism

$$\Phi(x, \xi) = (\phi^{-1}(x), \phi^*_{\phi^{-1}(x)}(\xi))$$

induced by ϕ on T^*X.

Exercise 8.6. Work out this last example carefully.

Now it turns out to be helpful to actually consider a broader class of FIO's than we have described so far. Instead of just using Lagrangian submanifolds of $T^*(X \times X)$ given by $\Gamma' = \Gamma'_\Phi$ where Φ is a symplectomorphism, we just require that Γ' be a reasonable Lagrangian (and we allow operators between different manifolds while we are at it):

Definition 8.7. Let X, Y be two manifolds (not necessarily of the same dimension). A *homogeneous canonical relation* from T^*Y to T^*X is a homogeneous submanifold Γ of $(T^*X \backslash o) \times (T^*Y \backslash o)$, closed in $T^*(X \times Y) \backslash o$ such that

$$\Gamma' \equiv \{(x, \xi, y, \eta) : (x, \xi, y, -\eta) \in \Gamma\}$$

is Lagrangian in $T^*(X \times Y)$.

We can view Γ as giving a multivalued generalization of a symplectomorphism, with

$$\Gamma(y, \eta) \equiv \{(x, \xi) : (x, \xi, y, \eta) \in \Gamma\}.$$

and, more generally, if $S \subset T^*Y$ is conic,

(8.4) $\Gamma(S) \equiv \{(x, \xi) : \text{ there exists } (y, \eta) \in S, \text{ with } (x, \xi, y, \eta) \in \Gamma\}.$

Definition 8.8. A Fourier integral operator of order m associated to a homogeneous canonical relation Γ is an operator from $\mathcal{C}_c^\infty(Y)$ to $\mathcal{D}'(X)$ with Schwartz kernel in

$$I^m(X \times Y, \Gamma').$$

Exercise 8.7. Show that a homogeneous canonical relation Γ is associated to a symplectomorphism if and only if its projections onto both factors T^*X and T^*Y are diffeomorphisms.

Exercise 8.8.
  (1) Let $Y \subset X$ be a submanifold. Show that the operation of restriction of a smooth function on X to Y is an FIO.
  (2) Endow X with a metric, and consider the volume form dg_Y on Y arising from the restriction of this metric; show that the map taking a function f on Y to the distribution $\phi \mapsto \int_Y \phi|_Y(y) f(y) dg_Y$ is an FIO. (Think of it as just multiplying f by the delta-distribution along Y, which makes sense if we choose a metric.) What is the relationship between the restriction FIO and this one, which you might think of as an extension map?

In the special case that Γ is a canonical relation that is locally the graph of a symplectomorphism, we say it is a *local canonical graph.*

We now briefly enumerate the properties of the FIO calculus, somewhat in parallel with our discussion of pseudodifferential operators. These theorems are considerably deeper, however. In preparation for our discussion of composition, suppose that

$$\Gamma_1 \subset T^*X \backslash o \times T^*Y \backslash o,$$
$$\Gamma_2 \subset T^*Y \backslash o \times T^*Z \backslash o$$

are homogeneous canonical relations. We say that Γ_1 and Γ_2 are *transverse* if the manifolds

$$\Gamma_1 \times \Gamma_2 \text{ and } T^*X \times \Delta_{T^*Y} \times T^*Z$$

intersect transversely in $T^*X \times T^*Y \times T^*Y \times T^*Z$; here Δ_{T^*Y} denotes the diagonal submanifold.

Exercise 8.9. Show that if either Γ_1 or Γ_2 is the graph of a symplectomorphism, then Γ_1 and Γ_2 are transverse.

In what follows, we will as usual assume for simplicity that all manifolds are compact.[40] In the following list of properties, some are special

---

[40]In the absence of this assumption, we need as usual to add various hypotheses of properness.

to FIO's, that is to say, Lagrangian distributions on product manifolds, viewed as operators; others are more generally properties of Lagrangian distributions per se, hence their statements do not necessarily involve products of manifolds. In the interests of brevity, we focus on the deeper properties, and omit trivialities such as associativity of composition. Note also that for brevity we will systematically confuse operators with their Schwartz kernels.

(I) (Algebra property) If $S \in I^m(X \times Y, \Gamma'_1)$ and $T \in I^{m'}(Y \times Z, \Gamma'_2)$ and Γ_1 and Γ_2 are transverse, then

$$S \circ T \in I^{m+m'}(X \times Z, (\Gamma_1 \circ \Gamma_2)'),$$

where

(8.5) $\Gamma_1 \circ \Gamma_2 = \{(x, \xi, z, \zeta) : (x, \xi, y, \eta) \in \Gamma_1$

and $(y, \eta, z, \zeta) \in \Gamma_2$ for some $(y, \eta)\}$.

Moreover,

$$S^* \in I^m(Y \times X, (\Gamma^{-1})')$$

where Γ^{-1} is obtained from Γ by switching factors.

(II) (Characterization of smoothing operators) The distributions in $I^{-\infty}(X, \mathcal{L})$ are exactly those in $\mathcal{C}^\infty(X)$; composition of an operator $S \in I^m(X \times Y, \Gamma')$ on either side with a smoothing operator (i.e. one with smooth Schwartz kernel) yields a smoothing operator.

(III) (Principal symbol homomorphism) There is family of linear "principal symbol maps"

(8.6) $\sigma_m : I^m(X, \mathcal{L}) \to \dfrac{S_{\mathrm{cl}}^{m+(\dim X)/4}(\mathcal{L}; L)}{S_{\mathrm{cl}}^{m-1+(\dim X)/4}(\mathcal{L}; L)}.$

Here L is a certain canonically defined line bundle on $\mathcal{L}$ (see the commentary below), and $S_{\mathrm{cl}}^m(\mathcal{L}; L)$ denotes L-valued symbols. We may identify the quotient space in (8.6) with

$$\mathcal{C}^\infty(S^*\mathcal{L}; L),$$

and we call the resulting map $\hat{\sigma}_m$ instead. If S, T, are as in (I), with canonical relations Γ_1, Γ_2 intersecting transversely,

$$\sigma_{m+m'}(ST) = \sigma_m(S)\sigma_{m'}(T)$$

and

$$\sigma_m(A^*) = s^* \overline{\sigma_m(A)},$$

where s is the map interchanging the two factors. The product of the symbols, at $(x, \xi, z, \zeta) \in \Gamma_1 \circ \Gamma_2$, is defined as

$$\sigma_m(S)(x, \xi, y, \eta) \cdot \sigma_{m'}(T)(y, \eta, z, \zeta)$$

evaluated at (the unique) (y, η) such that $(x, \xi, y, \eta) \in \Gamma_1$, $(y, \eta, z, \zeta) \in \Gamma_2$.

(IV) (Symbol exact sequence) There is a short exact sequence

$$0 \to I^{m-1}(X, \mathcal{L}) \to I^m(X, \mathcal{L}) \overset{\hat{\sigma}_m}{\to} \mathcal{C}^\infty(S^*\mathcal{L}; L) \to 0.$$

hence the symbol is 0 if and only if an operator is of lower order.

(V) Given $\mathcal{L}$, there is a linear "quantization map"

$$\mathrm{Op} : S_{\mathrm{cl}}^{m+(\dim X)/4}(\mathcal{L}; L) \to I^m(X, \mathcal{L})$$

such that if

$$a \sim \sum_{j=0}^{\infty} a_{m+(\dim X)/4-j}(x, \hat{\xi})|\xi|^{m+(\dim X)/4-j} \in S_{\mathrm{cl}}^{m+(\dim X)/4}(\mathcal{L}; L)$$

then

$$\sigma_m(\mathrm{Op}(a)) = a_{m+(\dim X)/4}(x, \hat{\xi}).$$

The map Op is onto, modulo $\mathcal{C}^\infty(X)$.

(VI) (Product with vanishing principal symbol) If $P \in \mathrm{Diff}^m(X)$ is self-adjoint and $u \in I^{m'}(X, \mathcal{L})$, with $\mathcal{L} \subset \Sigma_P \equiv \{\sigma_m(P) = 0\}$, then

$$Pu \in I^{m+m'-1}(X, \mathcal{L})$$

and

$$\sigma_{m+m'-1}(Pu) = i^{-1}\mathsf{H}_p(\sigma_{m'}(u)),$$

with H_p denoting the Hamilton vector field.

(VII) (L^2-boundedness, compactness) If $T \in I^m(X \times Y, \Gamma)$ is associated to a local canonical graph, then

$$T \in \mathcal{L}(H^s(Y), H^{s-m}(X)) \text{ for all } s \in \mathbb{R}.$$

Negative-order operators of this type acting on $L^2(X)$ are thus compact.

(VIII) (Asymptotic summation) Given $u_j \in I^{m-j}(X, \mathcal{L})$, with $j \in \mathbb{N}$, there exists $u \in I^m(X, \mathcal{L})$ such that

$$u \sim \sum_j u_j,$$

which means that

$$u - \sum_{j=0}^{N} u_j \in I^{m-N-1}(X, \mathcal{L})$$

for each N.

(IX) (Microsupport) The microsupport of $T \in I^m(X \times Y, \Gamma')$ is well defined as the largest conic subset $\tilde{\Gamma} \subset \Gamma$ on which the symbol is $O(|\xi|^{-\infty})$. We have

$$\operatorname{WF} Tu \subseteq \tilde{\Gamma}(\operatorname{WF} u)$$

for any distribution u on Y, where the action of $\tilde{\Gamma}$ on $\operatorname{WF} u$ is given by (8.4). Furthermore,

$$\operatorname{WF}'(S \circ T) \subseteq \operatorname{WF}' S \circ \operatorname{WF}' T.$$

COMMENTARY:

(I) This is a major result. Since FIO's include pseudodifferential operators, this includes the composition property for pseudodifferential operators as a special case. Another special case, when Z a point, yields the statement that an FIO applied to a Lagrangian distribution on the manifold Y with respect to the Lagrangian $\mathcal{L} \subset T^*Y$ is a Lagrangian distribution associated to $\Gamma(\mathcal{L})$, where Γ is the canonical relation of the FIO and $\Gamma(\mathcal{L})$ is defined by (8.4).

One remarkable corollary of this result is as follows: As will be discussed below, what our parametrix construction in §6 really showed was that for t sufficiently small, and fixed, we have

$$e^{-it\sqrt{\Delta}} \in I^0(X \times X, \mathcal{L}_t)$$

where $\mathcal{L}_t$ is the backwards geodesic flowout, for time t, in the left factor of $N^*\Delta$, of the conormal bundle to the diagonal in $T^*(X \times X)$.

Exercise 8.10. Verify this assertion! (Try this now, but fear not: we will discuss this example further in §9 and you can try again then.)

Now $e^{-it\sqrt{\Delta}}$ is a one-parameter group and so the composition property for FIO's allows us to conclude that in fact $e^{-it\sqrt{\Delta}}$ is an FIO for *all* times t, associated to the same flowout described above. The interesting subtlety is that while $\mathcal{L}_t$ is an inward- or outward-pointing conormal bundle for small positive resp. negative time (i.e. in the regime where our parametrix construction worked directly), for t exceeding the injectivity radius, it ceases to be a conormal bundle, while remaining a smooth Lagrangian manifold in $T^*(X \times X)$.

(III) Modulo bundle factors, the principal symbol is defined as follows: if $u \in I^m(X, \mathcal{L})$ is given by

$$u = (2\pi)^{-(n+2k)/4} \int_{\mathbb{R}^k} a(x, \theta) e^{i\phi(x,\theta)} \, d\theta,$$

then $\sigma_m(u)$ is defined by first restricting $a(x, \theta)$ to the manifold

$$C = \{(x, \theta) : d_\theta \phi = 0\};$$

as ϕ is a nondegenerate phase function, this manifold is locally diffeomorphic (via a homogeneous diffeomorphism) to $\mathcal{L}$, hence we may identify $a|_C$ with a function on $\mathcal{L}$; transferring this function to $\mathcal{L}$ via the local diffeomorphism and taking the top-order homogeneous term in the asymptotic expansion gives the principal symbol.

Much has been swept under the rug here—for a proper discussion, see, e.g., [10]. In particular, the line bundle L contains not just the density factors that we have been studiously ignoring—the Schwartz kernel of an operator from functions to functions on X is actually a "right-density" on $X \times X$, i.e. a section of the pullback of the bundle $|\Omega^n(X)|$ in the right factor—but also the celebrated "Keller-Maslov index," which is related to the indeterminacy in choosing the phase function parametrizing the Lagrangian. We will not enter into a serious discussion of these issues here. We have also omitted discussion of the geometry of composing canonical relations, and the fact that transverse canonical relations compose to give a new canonical relation, with a unique point y, η such that $(x, \xi, y, \eta) \in \Gamma_1$, $(y, \eta, z, \zeta) \in \Gamma_2$ whenever $(x, \xi, z, \zeta) \in \Gamma_1 \circ \Gamma_2$.

(VI) There is a more general version of this statement valid for any $P \in \Psi^m(X)$ characteristic on $\mathcal{L}$, but it involves the notion of subprincipal symbol, which requires some explanation; see [5, §5.2–5.3]. Moreover, if we are a little more honest about making this computation work invariantly, so that the symbol has a density factor in it (one factor in the line bundle L,) then we should really write

$$\sigma_{m+m'-1}(Pu) = i^{-1} \mathcal{L}_{\mathsf{H}_p} \sigma_{m'}(u),$$

where $\mathcal{L}_Z$ denotes the Lie derivative along the vector field Z.

(VII) This is fairly easy to prove, as if T of order m is associated to a symplectomorphism from Y to X, it is easy to check from the previous properties that T^*T is an FIO associated with the

canonical relation given by the identity map, and hence

$$T^*T \in \Psi^{2m}(Y),$$

and we may invoke boundedness results for the pseudodifferential calculus. In cases when T is not associated to a local canonical graph, this argument fails badly (i.e. interestingly), and the optimal mapping properties are a subject of ongoing research.

Finally, as with the pseudodifferential calculus, we may define a notion of ellipticity for FIO's, and the above properties imply that (microlocal) parametrices exist for the inverses of elliptic operators associated to symplectomorphisms.

9. The wave trace, redux

Let us briefly revisit our construction of the parametrix for the half-wave equation in the light of the FIO calculus. Here is what we did, in hindsight: we sought a distribution

$$u \in I^m(\mathbb{R} \times X \times X, \mathcal{L})$$

for some Lagrangian $\mathcal{L}$, and some order m, with

$$u(0, x, y) = \delta(x - y)$$

such that

$$(D_t + \sqrt{\Delta_x})u \in I^{-\infty}((-\epsilon, \epsilon) \times X \times X, \mathcal{L}) = \mathcal{C}^\infty((-\epsilon, \epsilon) \times X \times X).$$

We begin by sorting out what m, the order of u, should be. Since

$$u|_{t=0} = \delta(x - y) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i(x-y)\cdot\theta} \, d\theta,$$

we were led us to a solution that for t small was of the form

$$\int_{\mathbb{R}^n} a(t, x, y, \theta) e^{i\Phi(t,x,y,\theta)} \, d\theta$$

with a a symbol of order zero such that $a(0, x, y, \theta) = 1$, and Φ a nondegenerate phase function such that $\Phi(0, x, y, \theta) = (x - y) \cdot \theta$. This was certainly the rough form of our earlier Ansatz; it should now be regarded as a Lagrangian distribution, of course. Since $\dim(\mathbb{R} \times X \times X) = 2n + 1$ and we have n phase variables $\theta_1, \ldots, \theta_n$, the convention on orders of FIO's leads to $m = -1/4$.

Now we address the following question: what Lagrangian $\mathcal{L}$ ought we to choose? Since

$$\Box_{t,x} \in \mathrm{Diff}^2(\mathbb{R} \times X \times X) \subset \Psi^2(\mathbb{R} \times X \times X),$$

we a priori would have

$$\Box u \in I^{7/4}(\mathbb{R} \times X \times X, \mathcal{L});$$

as we would like smoothness of $\Box u$, we ought to start by making the principal symbol of $\Box u$ vanish. The symbol of $\Box$ vanishes only on

$$\Sigma_\Box = \{\tau^2 = |\xi|_g^2\}$$

hence the easiest way to ensure vanishing of the principal symbol is simply to arrange that

(9.1) $\mathcal{L} \subset \Sigma_\Box.$

Now, recall that our initial conditions were to be

$$u(0, x, y) = \delta(x - y),$$

where we may view this as a Lagrangian distribution on $X \times X$ with respect to $N^*\Delta$, the conormal to the diagonal:

$$N^*\Delta = \{(x, y, \xi, \eta) : x = y, \xi = -\eta\}.$$

It is not difficult to check that the requirement that $u|_{t=0}$ gives this lower-dimensional Lagrangian[41] together with the requirement (9.1) that $\mathcal{L}$ should lie in the characteristic set implies that $\mathcal{L} \cap \{t = 0\}$ should just consist of points in $\Sigma_\Box$ *projecting* to points in $N^*\Delta$, i.e. that we should in fact have

$$\mathcal{L} \cap \{t = 0\} = \{(t = 0, \tau = -|\eta|_g, x = y, \xi = -\eta)\} \subset T^*(\mathbb{R} \times X \times X).$$

Here we have chosen the sign $\tau = -|\eta|_g$ in view of our real interest, which is in solving

$$(D_t + \sqrt{\Delta})u = 0$$

rather than the full wave equation;[42] we have thus kept $\mathcal{L}$ inside the characteristic set of $D_t + \sqrt{\Delta}$, which is one of the two components of $\Sigma_\Box$.

Let $\mathcal{L}_0$ now denote $\mathcal{L} \cap \{t = 0\}$. The set $\mathcal{L}_0$ is a manifold on which the symplectic form vanishes (an "isotropic" manifold), of dimension one less than half the dimension of $T^*(\mathbb{R} \times X \times X)$. (Exercise: Check

---

[41]We really ought to think a bit about restriction of Lagrangian distributions here: this is best done by regarding the restriction operator itself as an FIO (cf. Exercise 8.8). We shall omit further discussion of this point, but remark that it should at least seem plausible that the Lagrangian manifold associated to the restriction is the projection (i.e. pullback under inclusion), of the Lagrangian in the ambient space—cf. Exercise 4.11.

[42]We have chosen to emphasize this distinction only at this critical juncture only because as it is in some respects more pleasant to deal with $\Box$ than with the half-wave operator when possible.

this! Most of the work is done already, as $N^*(\Delta)$ is Lagrangian in $T^*(X \times X)$.)

We now proceed as follows to find a Lagrangian (necessarily one dimensional larger) containing $\mathcal{L}_0$: let $\mathsf{H} = \mathsf{H}_\square$ denote the Hamilton vector field of the symbol of the wave operator, in the variables (t, x, τ, ξ). (I.e., take the Hamilton vector field of $\square_{(t,x)}$ on the cotangent bundle of $\mathbb{R} \times X \times X$—nothing interesting happens in y, η.) By construction, $\mathcal{L}_0 \subset \Sigma_\square$; we now *define* $\mathcal{L}$ to be the union of integral curves of H passing through points in $\mathcal{L}_0$. More concretely, these are all backwards unit-speed parametrized geodesics beginning at $(x = y, \xi = -\eta)$, where (x, ξ) evolves along the geodesic flow, and (y, η) are fixed. (Meanwhile, t is evolving at unit speed, and τ is constrained by the requirement that we are in the characteristic set so that $\tau = -|\xi|_g$.) The manifold $\mathcal{L}$ stays inside $\Sigma_\square$ (indeed, inside the component that is $\Sigma_{D_t + \sqrt{\Delta}}$) since H is tangent to this manifold; moreover, $\mathcal{L}$ is automatically Lagrangian since ω vanishes on $\mathcal{L}_0$ and $\sigma_2(\square)$ does as well, so that for $\mathsf{Y} \in T\mathcal{L}_0$, we further have

$$\omega(\mathsf{Y}, \mathsf{H}) = (d(\sigma_2(\square)), \mathsf{Y}) = \mathsf{Y}\sigma_2(\square) = 0.$$

This gives vanishing of ω on the tangent space to $\mathcal{L}$ at points along $t = 0$; to conclude it more generally, just recall that the flow generated by a Hamilton vector field is a family of symplectomorphisms.

Exercise 9.1. Check that $\mathcal{L}$ is in fact the *only* connected conic Lagrangian manifold passing through $\mathcal{L}_0$ and lying in $\Sigma_\square$. (HINT: Observe that H is in fact the *unique* vector at each point along $\mathcal{L}_0$ that has the property $\omega(\mathsf{Y}, \mathsf{H}) = 0$ for all $\mathsf{Y} \in T\mathcal{L}_0$.)

Thus, to recapitulate, if we obtain $\mathcal{L}$ by flowing out $\mathcal{L}_0$ (the lift of the conormal bundle of the diagonal to the characteristic set of $D_t + \sqrt{\Delta}$) along H, the Hamilton vector field of $\square$, we produce a Lagrangian on which $\square$ is characteristic.

Exercise 9.2. Show that the phase function $\phi(t, x, \eta) - y \cdot \eta$ that we constructed explicitly in §6 does indeed parametrize

$$\mathcal{L} = \{(t, \tau, x, \xi, y, -\eta) : \tau = -|\xi|_g, \ (x, \xi) = \Phi_t(y, \eta)$$

(with Φ_t denoting geodesic flow, i.e. the flow generated by the Hamilton vector field of $|\xi|_g$) over $|t| \ll 1$.

Compare our solution to the eikonal equation using Hamilton-Jacobi theory in Exercise 6.1 to what we have done here.

We now remark that while our parametrization of the Lagrangian in §6 worked only for small t, the definition given here of $\mathcal{L} \subset T^*(\mathbb{R} \times$

$X \times X$) makes sense *globally* in t, not merely for short time. When t is small and positive and y fixed, the projection of $\mathcal{L}$ to (x, ξ) is just the inward-pointing conormal bundle to an expanding geodesic sphere centered at y; when t exceeds the injectivity radius of X, $\mathcal{L}$ ceases to be a conormal bundle, but remains a well-behaved smooth Lagrangian.

Let us now return from our lengthy digression on the construction of $\mathcal{L}$ to recall what it gets us. Solving the eikonal equation, i.e. choosing $\mathcal{L}$, has reduced our error term by one order, and we have achieved

$$\Box u \in I^{3/4}(\mathbb{R} \times X \times X, \mathcal{L});$$

to proceed further, we invoke Property (VI) of FIO's, to compute

$$\sigma_{3/4}(\Box u) = i^{-1}\mathsf{H}\sigma_{-1/4}(u);$$

setting this equal to zero yields our first transport equation, and it is solved by simply insisting that $\sigma_{-1/4}(u)$ be constant along the flow, hence equal to 1, its value at $t = 0$ (which was dictated by our δ-function initial data).

Now we have achieved $\Box u = r_{-1/4} \in I^{-1/4}$ Adding an element $u_{-5/4}$ of $I^{-5/4}(\mathbb{R} \times X \times X, \mathcal{L})$ to solve this error away and again applying (VI) yields the transport equation

$$i^{-1}\mathsf{H}(\sigma_{-5/4}(u_{-5/4})) = -\sigma_{-1/4}(r_{-1/4}),$$

which we may solve as before. Continuing in this manner and asymptotically summing the resulting terms, we have our parametrix $u \in I^{-1/4}(\mathbb{R} \times X \times X, \mathcal{L})$.

Now we describe, *very roughly*, how to use the FIO calculus to compute the singularities of $\operatorname{Tr} U(t)$ at lengths of closed geodesics.

Let T denote the operator $\mathcal{C}^\infty(\mathbb{R} \times X \times X) \to \mathcal{C}^\infty(\mathbb{R})$ given by[43]

$$T : f(t, x, y) \mapsto \int_X f(t, x, x) \, dx.$$

Thus, $\operatorname{Tr} U = T(U)$, and we seek to identify this composition as a Lagrangian distribution on $\mathbb{R}^1$; such a distribution is thus conormal to some set of points; as we saw above (and will see again below) these points may only be the lengths of closed geodesics, together with 0.

---

[43]It is here that our omission of density factors becomes most serious: T should really act on *densities* defined along the diagonal, so that the integral over X is well-defined. Fortunately, U itself should be a *right*-density (i.e. a section of the density bundle lifted from the right factor); restricted to the diagonal, this yields a density of the desired type.

The Schwartz kernel of T is the distribution

$$\delta(t - t')\delta(x - y)$$

on $\mathbb{R} \times \mathbb{R} \times X \times X$; it is thus conormal to $t = t', x = y$, i.e. is a Lagrangian distribution with respect to the Lagrangian

$$\{t = t', x = y, \tau = -\tau', \xi = -\eta\}$$

Noting that if we reshuffle the factors into $(\mathbb{R} \times X) \times (\mathbb{R} \times X)$, the distribution $\delta(t-t')\delta(x-y)$ becomes the kernel of the identity operator, we can easily see that the order of this Lagrangian distribution is 0. Thus,

$$T \in I^0(\mathbb{R} \times \mathbb{R} \times X \times X, \Gamma')$$

where the relation $\Gamma : T^*(\mathbb{R} \times X \times X) \to T^*\mathbb{R}$ maps as follows:

$$\Gamma(t, \tau, x, \xi, y, \eta) = \begin{cases} \emptyset, & \text{if } (x, \xi) \neq (y, -\eta) \\ (t, \tau), & \text{if } (x, \xi) = (y, -\eta). \end{cases}$$

Let $\mathcal{L}$ be the Lagrangian for our parametrix u constructed above. If an interval about $L \in \mathbb{R}$ contains no lengths of closed geodesics, then we see that no points in $\mathcal{L}$ lie over $\{(x, \xi) = (y, -\eta)\}$ for t near L, hence $\Gamma(\mathcal{L})$ has no points over this interval, i.e. the composition Tu is smooth in this interval. *This gives another proof of the Poisson relation,* Theorem 5.3.

If, by contrast, there is a closed geodesic of length L, then

$$\{(L, \tau) : \tau < 0\} \in \Gamma(\mathcal{L}).$$

Note that in effect we get a contribution from every (x, ξ) lying along the geodesic, and that in particular, the fiber over (L, τ) of the projection on the left factor

$$\left(T^*\mathbb{R} \times \Delta_{T^*(\mathbb{R} \times X \times X) \times T^*(\mathbb{R} \times X \times X)}\right) \cap (\Gamma \times \mathcal{L}) \to T^*\mathbb{R}$$

(giving the composition $\Gamma(\mathcal{L})$) consists of at least a whole geodesic of length L, rather than a single point. Thus, the composition of these canonical relations *is not transverse* and the machinery described thus far does not apply. In [3], Duistermaat-Guillemin remedied this deficiency by constructing a theory of composition of FIO's with canonical relations intersecting *cleanly*.

Definition 9.1. Two manifolds X, Y intersect *cleanly* if $X \cap Y$ is a manifold with $T(X \cap Y) = TX \cap TY$ at points of intersection.

For instance, pairs of coordinate axes intersect cleanly but not transversely in $\mathbb{R}^n$. In general, in the notation of Property (I), if the intersection of the product of canonical relations $\Gamma_1 \times \Gamma_2$ with the partial diagonal $T^*X \times \Delta \times T^*Z$ is clean, we define the *excess*, e, to be the dimension

of the fiber of the projection from this intersection to $T^*X \times T^*Z$; this is zero in the case of transversality. Duistermaat-Guillemin show:

$$S \circ T \in I^{m+m'+e/2}(X \times Z, (\Gamma_1 \circ \Gamma_2)')$$

i.e. composition goes as before, but with a change in order. In addition the symbol of the product is obtained by *integrating* the product of the symbols over the e-dimensional fiber of the projection in what turns out to be an invariant way.

Let us now assume that there are finitely many closed geodesics of length L, and that they are *nondegenerate* in the following sense. For each closed bicharacteristic (i.e. lift to S^*X of a closed geodesic) $\gamma \subset S^*X$, pick a point $p \in \gamma$ and let $Z \subset S^*X$ be a small patch of a hypersurface through p transverse to γ. Shrinking Z as necessary, we can consider the map $P_\gamma : Z \to Z$ taking a point to its first intersection with Z under the bicharacteristic flow on S^*X. This is called a *Poincaré map*. Since $P_\gamma(p) = p$, we can consider $dP_\gamma : T_pZ \to T_pZ$. We say that the closed geodesic is *nondegenerate* if $\mathrm{Id} - dP_\gamma$ is invertible. Note that this condition is independent of our choices of p and Z, as are the eigenvalues of $\mathrm{Id} - dP_\gamma$.

The following is due to Duistermaat-Guillemin [3]:

Theorem 9.2. *Assume that all closed geodesics of length L on X are nondegenerate. Then*

$$\lim_{t \to L}(t - L)\operatorname{Tr} U(t) = \sum_{\gamma \text{ of length } L} \frac{L}{2\pi}i^{\sigma_\gamma}|\mathrm{Id} - dP_\gamma|^{-1/2},$$

where P_γ is the Poincaré map corresponding to the geodesic γ, and σ_γ is the number of conjugate points along the geodesic.

A proof of this theorem requires understanding the symbol of the clean composition Tu (where u is our parametrix for the half-wave equation). This lies beyond the scope of these notes. We merely note that we are in the setting of clean composition with excess 1, hence locally near $t = L$,

$$Tu \in I^{0-1/4+1/2}(\mathbb{R}, \{t = L, \tau < 0\}).$$

This Lagrangian is easily seen to be parametrized, locally near $t = L$, by the phase function with one fiber variable[44]

$$\phi(t,\theta) = \begin{cases} (t-L)\theta, \theta < 0, \\ 0, \theta \geq 0; \end{cases}$$

hence we may write

$$Tu = (2\pi)^{-3/4} \int_0^\infty a(t,\theta) e^{-i(t-L)\theta} \, d\theta,$$

where $a \in S^0(\mathbb{R} \times \mathbb{R})$ has an asymptotic expansion $a \sim a_0 + |\theta|^{-1} a_{-1} + \ldots$. Our task is to find the leading-order behavior of Tu, and this is of course dictated by its principal symbol. To top order, a is given by the constant function $a_0(L, 1)$, hence Tu is (to leading order) a universal constant times $a_0(L, 1)$ times the Fourier transform of the Heaviside function, evaluated at $t - L$. Thus, the limit in the statement of the theorem is, up to a constant factor, just the value of $a_0(L, 1)$. The whole problem, then, is to compute the principal symbol of this clean composition, and we refer the interested reader to [3] for the (rather tricky) computation.[45]

10. A GLOBAL CALCULUS OF PSEUDODIFFERENTIAL OPERATORS

10.1. **The scattering calculus on $\mathbb{R}^n$.** We now return to some of the problems discussed in §2, involving operators on noncompact manifolds. Recall that the Morawetz estimate on $\mathbb{R}^n$, for instance, hinged upon a *global* commutator argument, involving the commutator of the Laplacian with $(1/2)(D_r + D_r^*)$ on $\mathbb{R}^n$. Generalizing this estimate to noncompact manifolds will require some understanding of differential and pseudodifferential operators that is uniform near infinity. Recall that thus far, we have focused on the calculus of pseudodifferential operators on compact manifolds; in discussing operators on $\mathbb{R}^n$, we have avoided as far as possible any discussion of asymptotic behavior at spatial infinity. Thus, our next step is to discuss a calculus of operators—initially just on $\mathbb{R}^n$—that involves sensible bounds near infinity.

Thus, let us consider pseudodifferential symbols defined on all of $T^*\mathbb{R}^n$ with no restrictions on the support in the base variables, with

---

[44]This phase function should of course be modified to make it smooth across $\theta = 0$, but making this modification will only add a term in $\mathcal{C}^\infty(\mathbb{R})$ to the Lagrangian distribution we write down.

[45]We note that the factor i^{σ_γ} is the contribution of the (in)famous Keller-Maslov index, and is in many ways the subtlest part of the answer.

asymptotic expansions in *both* the base and fiber variables, both separately and jointly. To this end, note that changing to variables $|x|^{-1}, \hat{x}$, $|\xi|^{-1}$, and $\hat{\xi}$ amounts to *compactifying* the base and fiber variables of $T^*\mathbb{R}^n$ radially, to make the space $B^n_x \times B^n_\xi$, with B^n denoting the closed unit ball. (Recall that we defined a radial compactification map in (3.4), and that while $\langle\xi\rangle^{-1}$ and $\langle x\rangle^{-1}$ are what we should really use as defining functions for the spheres at infinity, $|\xi|^{-1}$ and $|x|^{-1}$ are acceptable substitutes as long as we stay away from the origin in the corresponding variables.) The space $B^n \times B^n$ is a *manifold with codimension-two corners,* i.e. a manifold locally modelled on $[0,1) \times [0,1) \times \mathbb{R}^{2n-2}$; its boundary is the union of the two smooth hypersurfaces $S^{n-1}_x \times B^n_\xi$ and $B^n_x \times S^{n-1}_\xi$. In our local coordinates, $|x|^{-1}$ and $|\xi|^{-1}$ are the defining functions for the two boundary hypersurfaces, i.e. the variables locally in $[0,1)$, while a choice of $n-1$ of each of the $\hat{x}$ and $\hat{\xi}$ variables gives the remaining $\mathbb{R}^{n-2}$.
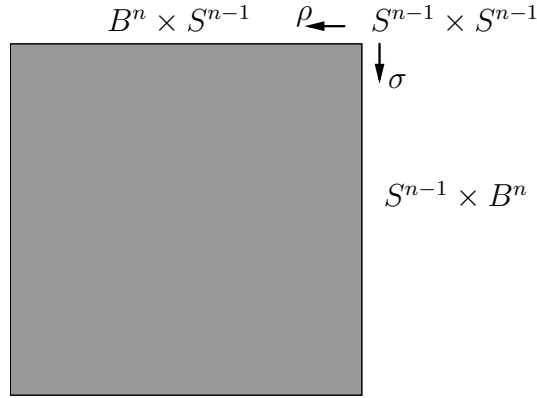

FIGURE 2. The manifold with corners $B^n \times B^n$ in the case $n = 1$. At the top (and bottom) are the boundary faces from $B^n \times S^{n-1}$ arising from the compactification of the second factor—this is "fiber infinity." At left (and right) are the faces from $S^{n-1} \times B^n$, arising from compactification of the first factor—this is "spatial infinity." The corner(s) at which these faces meet is $S^{n-1} \times S^{n-1}$. The functions $\rho = |x|^{-1}$ and $\sigma = |\xi|^{-1}$ can be locally taken as defining functions for the spatial infinity resp. fiber infinity boundary faces. The disconnectedness of $B^n \times S^{n-1}$ and $S^{n-1} \times B^n$ is of course a feature unique to dimension one.

We now let[46]

$$S_{\mathrm{sc}}^{m,l}(T^*\mathbb{R}^n)$$

denote the space of $a \in \mathcal{C}^\infty(T^*\mathbb{R}^n)$ such that[47]

(10.1) $$\langle\xi\rangle^{-m}\langle x\rangle^{-l} a \in \mathcal{C}^\infty(B^n \times B^n).$$

This condition gives asymptotic expansions (i.e., Taylor series) in various regimes:

(10.2)

$$a(x,\xi) \sim \sum |\xi|^{m-j} a_{\bullet,j}(x,\hat{\xi}), \text{ as } \xi \to \infty, \ x \in U \Subset \mathbb{R}^n \cong (B^n)^\circ$$

$$a(x,\xi) \sim \sum |x|^{l-i} a_{i,\bullet}(\hat{x},\xi), \text{ as } x \to \infty, \ \xi \in V \Subset \mathbb{R}^n \cong (B^n)^\circ$$

$$a(x,\xi) \sim \sum |x|^{l-i}|\xi|^{m-j} a_{ij}(\hat{x},\hat{\xi}), \text{ as } x,\xi \to \infty.$$

Finally, let

$$\Psi_{\mathrm{sc}}^{m,l}(\mathbb{R}^n)$$

denote the space consisting of the (left) quantizations of these symbols. The "sc" stands for "scattering."[48]

This is an algebra of pseudodifferential operators, containing all ordinary pseudodifferential operators on $\mathbb{R}^n$ with compactly supported Schwartz kernels. The algebra of scattering pseudodifferential operators enjoys all the good properties of our usual algebra, plus some more that derive from its good behavior at infinity. We can compose operators to get new operators, and if $A \in \Psi_{\mathrm{sc}}^{m,l}(\mathbb{R}^n)$, $B \in \Psi_{\mathrm{sc}}^{m',l'}(\mathbb{R}^n)$, we have $AB \in \Psi_{\mathrm{sc}}^{m+m',l+l'}(\mathbb{R}^n)$. Likewise, adjoints preserve orders. What is novel here, however, is the principal symbol map.

As the symbols defined by (10.1) are those that, up to overall factors, are smooth functions on $B^n \times B^n$, we can define the *principal symbol* of order m,l of the operator $\mathrm{Op}(a)$ as

$$\hat{\sigma}_{m,l}(A) = \langle\xi\rangle^{-m}\langle x\rangle^{-l} a|_{\partial(B^n \times B^n)};$$

this can be further split into pieces corresponding to the restrictions to the two boundary hypersurfaces:

$$\hat{\sigma}_{m,l}(A) = (\hat{\sigma}_{m,l}^\xi(A), \hat{\sigma}_{m,l}^x(A))$$

---

[46]This space should really be called $S_{\mathrm{cl,sc}}^{m,l}$, with the cl once again indicating "classicality" (as opposed to Kohn-Nirenberg type of estimates alone). We omit the cl so as not to clutter up the notation.

[47]We are abusing notation here by ignoring the diffeomorphism of radial compactification, thus identifying $\mathcal{C}^\infty(B^n \times B^n)$ directly with a space of functions on $\mathbb{R}^n \times \mathbb{R}^n$.

[48]This is a space of operators considered by many authors; as we are following roughly the treatment of Melrose [18], we have adopted his notation for the space. Note, however, that we have reversed the sign from his convention for the order l.

where
$$\hat{\sigma}_{m,l}^{\xi}(A)(x,\hat{\xi}) \in \mathcal{C}^{\infty}(B^n \times S^{n-1})$$
is nothing but the ordinary principal symbol, rescaled by a power of $\langle x \rangle$, and
$$\hat{\sigma}_{m,l}^{x}(A)(\hat{x},\xi) \in \mathcal{C}^{\infty}(S^{n-1} \times B^n)$$
is the novel piece of the symbol, measuring the behavior of the operator at spatial infinity. Note that these two pieces of the principal symbol are not independent: they must agree at the *corner,* $S^{n-1} \times S^{n-1}$. We may also choose to think of the principal symbol as
$$\sigma^{m,l}(A) \in S_{\mathrm{sc}}^{m,l}(T^*\mathbb{R}^n)/S_{\mathrm{sc}}^{m-1,l-1}(T^*\mathbb{R}^n),$$
and we will often confuse the symbol with its equivalence class; this is usually less confusing than keeping track of the rescaling factor $\langle \xi \rangle^m \langle x \rangle^l$.

The principal symbol short exact sequence thus reads:
$$0 \to \Psi_{\mathrm{sc}}^{m-1,l-l}(\mathbb{R}^n) \to \Psi^{m,l}(\mathbb{R}^n) \overset{\hat{\sigma}_{m,l}}{\to} \mathcal{C}^{\infty}(\partial(B^n \times B^n)) \to 0.$$

Thus, vanishing of this symbol yields improvement in both orders at once; correspondingly, vanishing of one part of the symbol gives improvement in just one order:

$$0 \to \Psi_{\mathrm{sc}}^{m-1,l}(\mathbb{R}^n) \to \Psi^{m,l}(\mathbb{R}^n) \overset{\hat{\sigma}_{m,l}^{\xi}}{\to} \mathcal{C}^{\infty}(B^n \times S^{n-1}) \to 0,$$

$$0 \to \Psi_{\mathrm{sc}}^{m,l-1}(\mathbb{R}^n) \to \Psi^{m,l}(\mathbb{R}^n) \overset{\hat{\sigma}_{m,l}^{x}}{\to} \mathcal{C}^{\infty}(S^{n-1} \times B^n) \to 0.$$

The symbol of the product of two scattering operators is indeed the product of the symbols,[49] as (equivalence classes of) smooth functions on $\partial(B^n \times B^n)$.

The symbol of the commutator of two scattering operators (which is of lower order than the product in both filtrations) is, as one might suspect, given by i times the Poisson bracket of the symbols.

The residual calculus is particularly nice in this setting: instead of merely consisting of smoothing operators, it consists of operators that are "Schwartzing"—they create decay as well as smoothness:
$$R \in \Psi_{\mathrm{sc}}^{-\infty,-\infty}(\mathbb{R}^n) \iff R : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n).$$

One problem with using the ordinary calculus for global matters is that we can only conclude compactness of operators of negative order

---

[49]It is exactly this innocuous statement, which the reader might think routine, that separates the scattering calculus from many other choices of pseudodifferential calculus on noncompact manifolds: typically the "symbol at infinity" (here $\hat{\sigma}_{m,l}^{x}(\hat{x},\xi)$) will compose under operator composition in a more complex, noncommutative way.

for compactly supported operators. Here, we have a much more precise result:

Proposition 10.1. *An operator in $\Psi_{sc}^{0,0}(\mathbb{R}^n)$ is bounded on $L^2(\mathbb{R}^n)$; an operator of order (m, l) with $m, l < 0$ is compact on $L^2(\mathbb{R}^n)$.*

Associated to the expanded notion of symbol, there is are associated notions of ellipticity (nonvanishing of the principal symbol) and of WF' (lack of infinite order vanishing of the total symbol). We have an associated family of Sobolev spaces:

$$u \in H_{\mathrm{sc}}^{m,l}(\mathbb{R}^n) \iff \forall A \in \Psi_{\mathrm{sc}}^{m,l}(\mathbb{R}^n), \; Au \in L^2(\mathbb{R}^n).$$

Operators in the calculus act on this scale of Sobolev spaces in the obvious way. Since smoothing operators are "Schwartzing," it is not hard to see that

$$H_{\mathrm{sc}}^{-\infty,-\infty}(\mathbb{R}^n) = \mathcal{S}(\mathbb{R}^n).$$

(We will return to an explicit description of these Sobolev spaces shortly.)
There is also an associated wavefront set:

$$\mathrm{WF}_{\mathrm{sc}}\, u \subset \partial(B^n \times B^n)$$

is defined by

$$p \notin \mathrm{WF}_{\mathrm{sc}}\, u \iff \text{ there exists } A \in \Psi_{\mathrm{sc}}^{0,0}(\mathbb{R}^n), \text{ elliptic at } p, \text{ with } Au \in \mathcal{S}.$$

In $(B_x^n)^\circ \times S_\xi^{n-1} \subset \partial(B^n \times B^n)$, (i.e., in the usual cotangent bundle of $\mathbb{R}^n$) this definition just coincides with ordinary wavefront set; but "at infinity," i.e. in $S_x^{n-1} \times B_\xi^n$, it measures something new. To see what, let us consider some examples.

Example 10.2.

(1) *Constant coefficient vector fields on $\mathbb{R}^n$* : If $v \in \mathbb{R}^n$ and $P = i^{-1}v \cdot \nabla$, then, we can write

$$P = \mathrm{Op}_\ell(v \cdot \xi);$$

the principal symbol is thus

$$\sigma_{1,0}(P) = v \cdot \xi$$

(2) Likewise, the symbol of the Euclidean Laplacian Δ is $\sigma_{2,0}(\Delta) = |\xi|^2$. Note that the Laplacian is *not elliptic* in the scattering calculus, as its principal symbol vanishes at $\xi = 0$ on the boundary face $S_x^{n-1} \times B_\xi^n$. This should come as no suprise, as Δ has nullspace in $\mathcal{S}'(\mathbb{R}^n)$ (given by harmonic polynomials) that does not lie in L^2, hence is not consistent with elliptic regularity in

the scattering calculus sense: if Q is elliptic in the scattering calculus,

$$Qu \in \mathcal{S}(\mathbb{R}^n) \implies u \in \mathcal{S}(\mathbb{R}^n).$$

On the other hand, consider $\mathrm{Id} + \Delta$. We have $\mathrm{Id} \in \Psi_{\mathrm{sc}}^{0,0}(\mathbb{R}^n)$, hence adding it certainly does not alter the "ordinary" part of the symbol, living on $(B^n)^\circ \times S^{n-1}$. But it *does* affect the symbol in $S^{n-1} \times B^n$: we have

$$\sigma_{2,0}(\mathrm{Id} + \Delta) = 1 + |\xi|^2;$$

$\mathrm{Id} + \Delta$ *is* an elliptic operator in the scattering calculus, and of course it is the case that $(\mathrm{Id} + \Delta)u \in \mathcal{S}(\mathbb{R}^n)$ implies that u is likewise Schwartz.

(3) If we vary the metric from the Euclidean metric to some other metric g, we may or may not obtain a scattering differential operator; for example, if g were periodic, we certainly would not, as the total symbol of Δ would clearly lack an asymptotic expansion as $|x| \to \infty$. Suppose, however, that we may write in spherical coordinates on $\mathbb{R}^n$

$$g = dr^2 + r^2 \sum h_{ij}(r^{-1}, \theta) d\theta^i d\theta^j \quad \text{for } r > R_0 \gg 0.$$

where h_{ij} is a smooth function of its arguments, and

$$h_{ij}(0, \theta) d\theta^i d\theta^j$$

is the standard metric on the "sphere at infinity." We will call such a metric *asymptotically Euclidean.* Then the corresponding Laplace operator is in the scattering calculus.

Exercise 10.1. Check that this operator does lie in the scattering calculus.

Let Δ denote the Laplacian with respect to an asymptotically Euclidean metric. Then

$$(\mathrm{Id} + \Delta)^{-1} \in \Psi_{\mathrm{sc}}^{-2,0}(\mathbb{R}^n).$$

(4) $\langle x \rangle^2 (\mathrm{Id} + \Delta) \in \Psi_{\mathrm{sc}}^{2,2}(\mathbb{R}^n)$ and has symbol $\langle x \rangle^2 (1 + |\xi|^2)$. This is globally elliptic.

By the last example, we find that

$$u \in H_{\mathrm{sc}}^{2,2}(\mathbb{R}^n) \iff \langle x \rangle^2 (\mathrm{Id} + \Delta)u \in L^2(\mathbb{R}^n);$$

interpolation and duality arguments allow us to conclude more generally that the scattering Sobolev spaces coincide with the usual weighted Sobolev spaces:

$$H_{\mathrm{sc}}^{m,l}(\mathbb{R}^n) = \langle x \rangle^{-l} H^m(\mathbb{R}^n).$$

We now turn to some examples illustrating the scattering wavefront set. Consider the plane wave

$$u(x) = e^{i\alpha \cdot x}.$$

We have

$$(D_{x^j} - \alpha_j)u = 0 \text{ for all } j = 1, \ldots, n.$$

The symbol of the operator $D_{x^j} - \alpha_j$ is $\xi_j - \alpha_j$, hence the intersection of the characteristic sets of these operators is just the points in $S^{n-1} \times B^n$ where $\xi = \alpha$. As a consequence, we have

$$\mathrm{WF}_{\mathrm{sc}}(e^{i\alpha \cdot x}) \subseteq \{(\hat{x}, \xi) \in S^{n-1} \times \mathbb{R}^n : \xi = \alpha\}$$

(here we are as usual identifying $(B^n)^\circ \cong \mathbb{R}^n$). In fact this containment turns out to be equality, as we see by the following characterization of scattering wavefront set.

Proposition 10.3. *Let $p = (\hat{x}_0, \xi_0) \in S^{n-1} \times \mathbb{R}^n$. We have*

$$p \notin \mathrm{WF}_{sc}\, u$$

if and only if there exist cutoff functions $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ nonzero at ξ_0 and $\gamma \in \mathcal{C}^\infty(\mathbb{R}^n)$ nonzero in a conic neighborhood of the direction $\hat{x}_0$ such that

$$\phi \mathcal{F}(\gamma u) \in \mathcal{S}(\mathbb{R}^n).$$

This is of course closely analogous to the characterization of ordinary wavefront set in Proposition 4.5, and is proved in an analogous manner. Note that if u is a Schwartz function in a set of the form

$$\left\{ \left| \frac{x}{|x|} - \hat{x}_0 \right| < \epsilon, |x| > R_0 \right\}$$

for any $\epsilon > 0$, $R_0 \gg 0$, then there is no scattering wavefront set at points of the form $(\hat{x}_0, \xi)$ for any $\xi \in \mathbb{R}^n$. Thus, this new piece of the wavefront set measures the asymptotics of u in different directions toward spatial infinity: $\hat{x}_0$ provides the direction, while the value of ξ_0 records oscillatory behavior of a specific frequency.

There is also, of course, a similar characterization of $\mathrm{WF}_{\mathrm{sc}}\, u$ inside $S^{n-1} \times S^{n-1}$. We leave this as an exercise for the reader.

10.2. **Applications of the scattering calculus.** As an example of how we might use the scattering calculus to obtain global results on manifolds, let us return to the *local smoothing estimate* from §2.1. Recall that if ψ satisfies the Schrödinger equation (2.1) on $\mathbb{R}^n$ with initial data $\psi_0 \in H^{1/2}$, this estimate (or, at least, one version of it) tells us that

$$(10.3) \qquad\qquad \psi \in L^2_{\mathrm{loc}}(\mathbb{R}_t; H^1_{\mathrm{loc}}(\mathbb{R}^n)),$$

hence the solution is (locally) half a derivative smoother than the data, on average. How might we obtain this estimate on a manifold, with Δ replaced by the Laplace-Beltrami operator (which we also denote Δ)? For a start, note that (10.3) fails badly on compact manifolds; in particular, recall that since $[\Delta, \Delta^s] = 0$ for all $s \in \mathbb{R}$, the H^s norms are conserved under the evolution, hence if $\psi_0 \notin H^s$, with $s > 1/2$, then we certainly do not have[50] $\psi \in L^2_{\text{loc}}(\mathbb{R}_t; H^s)$. So if we seek a broader geometric context for this estimate, we had better try noncompact manifolds.

Recall that we initially obtained the estimate by a commutator argument with the Morawetz commutant

$$\partial_r + \frac{n-1}{2r},$$

which actually gave more information; we noted that we could, instead, have used a simpler commutant $f(r)D_r$, with $f(r) = 0$ near $r = 0$, nondecreasing, and equal to 1 for $r \geq 2$ (say): this gives a commutator with a term

$$\chi'(r)D_r^2$$

which, when paired with ψ and integrated in time, tests for H^1 regularity in an annular neighborhood of the origin (which could have been translated to be anywhere); other terms in the commutator are positive also, modulo estimable error terms, and we thus obtain the local smoothing estimate. Generalizing this is tricky, as the positivity of the symbol of the term

$$i[\Delta, D_r]$$

on $\mathbb{R}^n$ is delicate: the symbol of this commutator is given by the Poisson bracket

$$\{|\xi|^2, \xi \cdot \hat{x}\} = 2\xi \cdot \partial_x(\xi \cdot \hat{x}) = \frac{2}{|x|}\left(|\xi|^2 - (\xi \cdot \hat{x})^2\right)$$

which is nonnegative but does actually vanish at $\xi \parallel x$, i.e. in radial directions. If we perturb the Euclidean metric a bit, and replace $|\xi|^2$ with $|\xi|_g^2$, the symbol of the Laplace-Beltrami operator, but leave the inner product $\langle \xi, x \rangle = \sum \xi_j x^j$, then this computation fails to give positivity. So we have to be more careful. We might try to adapt $\sum \xi_j x^j$ to the new metric instead, but this is problematic, as it doesn't really

---

[50]Note that this argument fails on $\mathbb{R}^n$ exactly because of the distinction between local and global Sobolev regularity: there is nothing preventing a solution on $\mathbb{R}^n$ with initial data in $H^{1/2}$ from being locally H^1—or even smooth on arbitrarily large compact sets—in return for having nasty behavior near infinity.

make much invariant sense. Moreover, it seems even more problematic upon interpretation: what positivity of $\{|\xi|_g^2, a\}$ *means* is just that a is increasing along the bicharacteristic flow of $|\xi|_g^2$, i.e. is increasing along (the lifts to the cosphere bundle of) geodesics. This is clearly impossible if there are any closed (i.e., periodic) geodesics, or indeed if there are geodesics that remain in a compact set for all time, hence our difficulty in obtaining an estimate on compact manifolds.

Exercise 10.2. Suppose that a geodesic γ remains in a compact subset of $\mathbb{R}^n$ (equipped with a non-Euclidean metric) for all $t > 0$. Let $p = (\gamma(0), (\gamma'(0))^*) \in T^*\mathbb{R}^n$ (with $*$ denoting dual under the metric). Show that there cannot exist a smooth $a \in \mathcal{C}^\infty(T^*\mathbb{R}^n)$ with $\{|\xi|_g^2, a\} \geq \epsilon > 0$ and $a(p) \neq 0$.

Definition 10.4. Let g be an asymptotically Euclidean metric on $\mathbb{R}^n$, and let γ be a geodesic. We say that γ is *not trapped forward/backward* if
$$\lim_{t \to \pm\infty} |\gamma(t)| = \infty.$$
We say that γ is *trapped* if it is trapped both forward *and* backward. We also use the same notation for the bicharacteristic projecting to γ. Moreover, we say that a point in $S^*\mathbb{R}^n$ along a non-(forward/backward)-trapped geodesic is itself non-(forward/backward)-trapped.

It is a theorem of Doi [4] that the local smoothing estimate (10.3) *cannot hold* near a trapped geodesic. (The total failure of (10.3) on compact manifolds should make this plausible, but it turns out to be considerably more delicate to show that it fails even if the only trapping is, for instance, a single, highly unstable, closed geodesic.) As a result we will require some strong geometric hypotheses in in order to find a general context in which (10.3) holds.

The following is a result of Craig-Kappeler-Strauss [1]:

Theorem 10.5. *Consider ψ a solution to the Schrödinger equation on asymptotically Euclidean space, with $\psi_0 \in H^{1/2}(\mathbb{R}^n)$. The estimate (10.3) holds microlocally at any (x_0, ξ_0) that lies on a nontrapped bicharacteristic, i.e. for any $A \in \Psi^1(\mathbb{R}^n)$ compactly supported and microsupported sufficiently near to (x_0, ξ_0), we have for any $T > 0$,*[51]
$$\int_0^T \|A\psi\|^2 \, dt \lesssim \|\psi_0\|_{H^{1/2}}^2.$$

---

[51]More generally, we can replace the Sobolev exponents $1/2$ and 1 by s and $s + 1/2$ respectively; in particular, L^2 initial data gives an $L^2 H^{1/2}$ estimate.

Proof. We will prove the theorem by using a commutator argument in the scattering calculus. To begin, we recall from Exercise 4.21 that the set along which microlocal $L^2_{\text{loc}}H^1$ regularity holds is invariant under the geodesic flow. Hence it suffices just to obtain regularity of this form *somewhere* along the geodesic γ. The convenient place to do this is out near infinity.

In order to make a commutator argument, note that it is very useful to have a quantity that behaves monotonically along the flow. We refer to points in $T^*\mathbb{R}^n$ near infinity (i.e. for $|x| \gg 0$) as *incoming* if $\hat{\xi} \cdot \hat{x} < 0$ and *outgoing* if $\hat{\xi} \cdot \hat{x} > 0$ (this corresponds to moving toward or away from the origin, respectively, under asymptotically Euclidean geodesic flow). Heuristically, under the classical evolution, points move from being incoming to being outgoing. More precisely, we observe that the Hamilton vector field of $p \equiv \sigma_{2,0}(\Delta)$ is given by

$$\mathsf{H}_p = -\sum \xi_i \xi_j \frac{\partial g^{ij}(x)}{\partial x^k} \partial_{\xi_k} + 2 \sum \xi_i g^{ij}(x) \partial_{x^j}.$$

Recalling that g^{ij} has an asymptotic expansion with leading term given by the identity metric, we can write this as

(10.4) $$\mathsf{H}_p = 2\xi \cdot \partial_x + O(|x|^{-1}|\xi|)\partial_x + O(|x|^{-1}|\xi|^2)\partial_\xi$$

(where in fact the whole vector field is homogeneous of degree 1 in ξ).

Exercise 10.3. Verify (10.4).

Thus,

$$\mathsf{H}_p(\hat{\xi} \cdot \hat{x}) = \frac{|\xi|}{|x|}\left(1 - (\hat{\xi} \cdot \hat{x})^2\right) + O(|\xi||x|^{-1}).$$

This is thus positive, as long as $\hat{\xi} \cdot \hat{x}$ is away from ± 1, and $|x|$ is large,[52] i.e., as long as we stay away from precisely incoming or outgoing points. Thus, we manufacture a scattering symbol for a commutant that has increase owing to the increase in "outgoingness:" Let $\chi(s)$ denote a smooth function that equals 0 for $s < 1/4$ and 1 for $s > 1/2$, with χ' a square of a smooth function, nonzero in the interior of its support. Let $\chi_\delta(s) = \chi(\delta s)$. We choose

$$a(x, \xi) = |\xi|_g \chi(-\hat{\xi} \cdot \hat{x})\chi_\delta(|x|)\chi(|\xi|_g).$$

Thus a is supported at *incoming* points at which $|x| \geq 1/(4\delta) \gg 0$; the first χ factor localizes near incoming points, and the factor of χ_δ keeps $|x|$ large. (The factor $\chi(|\xi|_g)$ simply cuts off near the origin in ξ to yield a smooth symbol.) Under the flow on the support of a, x

---

[52]Largeness of ξ plays no role because of homogeneity of the Hamilton vector field of the principal symbol of Δ.

tends to decrease and we become more outgoing, so the tendency is the *leave* the support of a along the flow. This is the essential point in the following:

Exercise 10.4. Check that $a \in S_{\mathrm{sc}}^{1,0}(T^*\mathbb{R}^n)$ and that if δ is chosen sufficiently small, we may write

$$\mathsf{H}_p a = -b^2 - c^2$$

where

(1) $b \in S_{\mathrm{sc}}^{1,-1/2}(T^*\mathbb{R}^n)$ is supported in $\operatorname{supp} \chi'(-\hat{\xi} \cdot \hat{x})\chi_\delta(|x|)$
(2) $c \in S_{\mathrm{sc}}^{1,-1/2}(T^*\mathbb{R}^n)$ is supported in $\operatorname{supp} \chi(-\hat{\xi} \cdot \hat{x})\chi_\delta'(|x|)$ and nonzero on the interior of that set.

(Note that $|\xi|_g$ is annihilated by H_p, so the terms containing $|\xi|_g$ simply do not contribute.)

Now let $A \in \Psi_{\mathrm{sc}}^{1,0}(\mathbb{R}^n)$ have principal symbol a. Then we have

$$i[\Delta, A] = -B^*B - C^*C + R$$

with $B = \operatorname{Op}(b)$, $C = \operatorname{Op}(c) \in \Psi_{\mathrm{sc}}^{1,-1/2}(\mathbb{R}^n)$, and $R \in \Psi_{\mathrm{sc}}^{1,-2}(\mathbb{R}^n)$.

Hence,

$$\int_0^T \|C\psi\|^2 \, dt \le \left| \langle A\psi, \psi \rangle \big|_0^T \right| + \left| \int_0^T \langle R\psi, \psi \rangle \, dt \right|.$$

As $\langle A\psi, \psi \rangle$ is bounded by the $L^\infty H^{1/2}$ norm of ψ and hence by $\|\psi_0\|_{H^{1/2}}^2$, and the R term likewise,[53] we obtain

(10.5) $$\int_0^T \|C\psi\|^2 \, dt \lesssim \|\psi_0\|_{H^{1/2}}^2.$$

Exercise* 10.5. Show that for any $R_0 > 0$, there exists $\delta > 0$ sufficiently small that if $(x_0, \xi_0) \in T^*\mathbb{R}^n \cap \{|x| < R_0\}$ lies along a non-backward trapped bicharacteristic, some point on that bicharacteristic with $t \ll 0$ lies in $\operatorname{ell} C$, with $C = \operatorname{Op}(c)$ constructed as above.

Thus, rays starting close to the origin that pass through $|x| \sim \delta^{-1}$ for $t \ll 0$ are incoming when they do so. This is an exercise in ODE. You might begin by showing that if a backward bicharacteristic starting in $\{|x| < R_0\}$ passes through the hypersurface $|x| = R'$ with $R' \gg 0$,

---

[53]In fact, the R term is considerably better than necessary for this step, as it has weight -2 rather than just 0 (which would be all we need to obtain the estimate). The astute reader may thus recognize that we are far from using the full power of the scattering calculus here. A proof of the global estimate in Exercise 10.6 requires a more serious use of the symbol calculus, however, as do the estimates which are the focus of [1], which show that microlocal decay of the initial data yields higher regularity of the solution along bicharacteristics.

then it must have $\hat{\xi} \cdot \hat{x} < 0$ there, and that $\hat{\xi} \cdot \hat{x}$ will keep decreasing thereafter along the backward flow.

Given a non-backward-trapped point $q \in S^*\mathbb{R}^n$, Exercise 10.5 tells us that we may construct a commutant A as above so that the commutator term C is elliptic somewhere along the bicharacteristic through q. Equation 10.5 tells us that we have the desired $L^2 H^1$ estimate on ell C, and the flow-invariance from Exercise 4.21 yields the same conclusion at q. Thus, we have proved the desired result at non-backward-trapped points. It remains to consider non-forward-trapped points.

Suppose, then, that $q = (x_0, \xi_0) \in T^*\mathbb{R}^n$ is non-forward-trapped; then note that $q' = (x_0, -\xi_0)$ is non-backward-trapped. Consider then the function $\overline{\psi}$: if

$$(D_t + \Delta)\psi = 0$$

then

$$(-D_t + \Delta)\overline{\psi} = 0,$$

i.e.

$$\tilde{\psi}(t, x) = \overline{\psi}(T - t, x)$$

again solves the Schrödinger equation. Of course, by unitarity,

$$\left\| \tilde{\psi}(0, x) \right\|_{H^{1/2}} = \|\psi_0\|_{H^{1/2}}.$$

Since q' is non-backward trapped, we thus find that there exists $C \in \Psi_{\mathrm{sc}}^{1,-1/2}(\mathbb{R}^n)$, elliptic at q', with

$$\int_0^T \left\| C\tilde{\psi} \right\|^2 dt \lesssim \left\| \tilde{\psi}(0, x) \right\|_{H^{1/2}}^2 = \|\psi_0\|_{H^{1/2}}^2;$$

on the other hand,

$$\left\| C\tilde{\psi}(t, \cdot) \right\|^2 = \left\| C\overline{\psi}(T - t, \cdot) \right\|^2$$
$$= \left\| \overline{C}\psi(T - t, \cdot) \right\|^2,$$

where

$$C = \mathrm{Op}_\ell(c(x, \xi)), \text{ and } \overline{C} = \mathrm{Op}_\ell(\overline{c}(x, -\xi));$$

thus, $\overline{C}$ tests for regularity at q, and we have obtained the desired estimate at q. $\square$

Corollary 10.6. *On an asymptotically Euclidean space with no trapped geodesics, the local smoothing estimate holds everywhere.*

Exercise* 10.6. (Global (weighted) smoothing.) Show that if there are no trapped geodesics, and $\psi_0 \in L^2$, we have

$$\int_0^T \left\| \langle x \rangle^{-1/2 - \epsilon} \psi \right\|_{H^{1/2}}^2 dt \lesssim \|\psi_0\|_{L^2}^2$$

for every $\epsilon > 0$. (This is a bit involved; a solution can be found, e.g., in Appendix II of [8].)

10.3. **The scattering calculus on manifolds.** We can generalize the description of the scattering calculus to manifolds quite easily, following the prescription of Melrose [18]. Let X be a compact manifold with boundary. We will, in practice, think of the interior, X°, as a non-compact manifold (with a complete metric) that just happens to come pre-equipped with a compactification to X. Our motivating example will be $X = B^n$, where X° is then diffeomorphically identified with $\mathbb{R}^n$ via the radial compactification map. Recall that on $\mathbb{R}^n$, radially compactified to the ball, we used coordinates near S^{n-1}, the "boundary at infinity," given by

$$\rho = \frac{1}{|x|}, \ \theta = \frac{x}{|x|},$$

where in fact ρ together with an appropriate choice of $n - 1$ of the θ's furnish local coordinates near a point. In these coordinates, what do constant coefficient vector fields on $\mathbb{R}^n$ look like? We have

$$\partial_{x^j} = \rho \partial_{\theta^j} - \rho \sum \theta^k \theta^j \partial_{\theta^k} - \rho^2 \theta^j \partial_\rho.$$

Recall moreover that functions in $\mathcal{C}^\infty(B^n)$ correspond exactly, under radial (un)compactification, to symbols of order zero on $\mathbb{R}^n$. So in fact it is easy to check more generally that vector fields on $\mathbb{R}^n$ with zero-symbol coefficients correspond exactly to vector fields on B^n that, near S^{n-1}, take the form

$$a(\rho, \theta)\rho^2 \partial_\rho + \sum b_j(\rho, \theta)\rho \partial_{\theta^j},$$

with $a, b_j \in \mathcal{C}^\infty(B^n)$.

We generalize this notion as follows. Given our manifold X, let $\rho \in \mathcal{C}^\infty(X)$ denote a *boundary defining function*, i.e.

$$\rho \geq 0 \text{ on X}, \ \rho^{-1}(0) = \partial X, \ d\rho \neq 0 \text{ on } \partial X.$$

Let θ^j be local coordinates on ∂X. We define *scattering vector fields on* X to be those that can be written locally, near ∂X, in the form

$$a(\rho, \theta)\rho^2 \partial_\rho + \sum b_j(\rho, \theta)\rho \partial_{\theta^j},$$

with $a, b_j \in \mathcal{C}^\infty(X)$. Let

$$\mathcal{V}_{\mathrm{sc}}(X) = \{\text{scattering vector fields on } X\}$$

Exercise 10.7.

    (1) Show that $\mathcal{V}_{\mathrm{sc}}(X)$ is well-defined, independent of the choices of ρ, θ.

    (2) Let $\mathcal{V}_{\mathrm{b}}(X)$ denote the space of smooth vector fields on X tangent to ∂X. Show that

$$\mathcal{V}_{\mathrm{sc}}(X) = \rho \mathcal{V}_{\mathrm{b}}(X)$$

    (3) Show that both $\mathcal{V}_{\mathrm{sc}}(X)$ and $\mathcal{V}_{\mathrm{b}}(X)$ are Lie algebras.

As we can locally describe the elements of $\mathcal{V}_{\mathrm{sc}}(X)$ as the $\mathcal{C}^\infty$-span of n vector fields, $\mathcal{V}_{\mathrm{sc}}(X)$ is itself the space of sections of a *vector bundle*, denoted

$$^{\mathrm{sc}}TX.$$

There is also of course a dual bundle, denoted

$$^{\mathrm{sc}}T^*X,$$

whose sections are the $\mathcal{C}^\infty$-span of the one-forms

$$\frac{d\rho}{\rho^2}, \quad \frac{d\theta^j}{\rho}.$$

Over X°, we may of course canonically identify $^{\mathrm{sc}}T^*X$ with T^*X, and the canonical one-form on the latter pulls back to give a canonical one-form

$$(10.6) \qquad\qquad \xi\frac{d\rho}{\rho^2} + \eta \cdot \frac{d\theta}{\rho}$$

defining coordinates ξ, η on the fibers of $^{\mathrm{sc}}T^*X$.

The scattering calculus on $\mathbb{R}^n$ is concocted to contain scattering vector fields:

Exercise 10.8. Show that $\Psi^{1,0}_{\mathrm{sc}}(\mathbb{R}^n) \supset \mathcal{V}_{\mathrm{sc}}(B^n)$.

We can, following Melrose, define the scattering calculus more generally as follows. Let $^{\mathrm{sc}}\overline{T}^*X$ denote the *fiber-compactification* of the bundle $^{\mathrm{sc}}T^*X$, i.e. we are radially compactifying each fiber to a ball, just as we did globally in compactifying $T^*\mathbb{R}^n$ to $B^n \times B^n$, only this time, the base is already compact. Now let

$$S^{m,l}_{\mathrm{sc}}(^{\mathrm{sc}}T^*X) = \sigma^{-m}\rho^{-l}\mathcal{C}^\infty(^{\mathrm{sc}}\overline{T}^*X),$$

where σ is a boundary defining function for the fibers. We can (by dint of some work!) quantize these "total" symbols to a space of operators, denoted

$$\Psi_{\mathrm{sc}}^{m,l}(X).$$

(Note that in the case $X = B^n$, we recover what we were previously writing as $\Psi_{\mathrm{sc}}^{m,l}(\mathbb{R}^n)$; the latter usage, with $\mathbb{R}^n$ instead of the more correct B^n, was an abuse of the usual notation.) The principal symbol of a scattering operator is, in this invariant picture, a smooth function on $\partial(^{\mathrm{sc}}\overline{T}^*X)$; or equivalently, an equivalence class of smooth functions on $^{\mathrm{sc}}\overline{T}^*X$; or, in the partially uncompactified picture, an equivalence class of smooth symbols on $^{\mathrm{sc}}T^*X$. (It is this last point of view that we shall mostly adopt.) In the coordinates defined by the canonical one-form (10.6), we have

$$(10.7) \qquad \sigma_{1,0}(\rho^2 D_\rho) = \xi, \ \sigma_{1,0}(\rho D_{\theta^j}) = \eta_j.$$

Recall that the Euclidean metric may be written in polar coordinates as

$$d(\rho^{-1})^2 + (\rho^{-1})^2 h(\theta, d\theta)$$

with h denoting the standard metric on S^{n-1}. We can generalize this to define a *scattering metric* as one on a manifold with boundary X that can be written in the form

$$\frac{d\rho^2}{\rho^4} + \frac{h(\rho, \theta, d\theta)}{\rho^2}$$

locally near ∂X, with ρ a boundary defining function, and h now a smooth *family* in ρ of metrics on ∂X.[54]

Exercise 10.9.

(1) Show that if g is a scattering metric on X, then the Laplace operator with respect to g can be written

$$\Delta = (\rho^2 D_\rho)^2 + O(\rho^3)D_\rho + \rho^2 \Delta_\theta$$

where Δ_θ is the family of Laplacians on ∂X associated to the family of metrics $h(r, \theta, d\theta)$.

(2) Show that for $\lambda \in \mathbb{C}$,

$$\sigma_{2,0}(\Delta - \lambda^2) = \xi^2 + |\eta|_h^2 - \lambda^2.$$

(Note that this entails noticing that you can drop the $O(\rho^3)D_\rho$ terms for different reasons at the the two different boundary faces of $^{\mathrm{sc}}\overline{T}^*X$. The term $-\lambda^2$ is of course only relevant at the

---

[54]The usual definition, as in [18], is a little more general, allowing $d\rho$ terms in h; however, it was shown by Joshi-Sá Barreto that these terms can always be eliminated by appropriate choice of coordinates.

$\rho = 0$ face; it does not contribute to the part of the symbol at fiber infinity, as it is a lower-order term there.)

As a consequence of Exercise 10.9, note as before that for $\lambda \in \mathbb{R}$, the Helmholz operator $\Delta - \lambda^2$ is not elliptic in the scattering sense: there are points in ${}^{\mathrm{sc}}T^*_{\partial X}X$ where $\xi^2 + |\eta|^2_h = \lambda^2$.

We now turn to scattering wavefront set $\mathrm{WF}_{\mathrm{sc}}$, which can, as one might expect, be defined in the usual manner as a subset of

$$\partial({}^{\mathrm{sc}}\overline{T}^*X),$$

hence is a subset of boundary faces at fiber infinity and at spatial infinity (i.e., over ∂X). The scattering wavefront set is the obstruction to a distribution lying in $\dot{\mathcal{C}}^\infty(X)$, where $\dot{\mathcal{C}}^\infty(X)$ denotes the set of smooth functions on X decaying to infinite order at ∂X. This space is the analogue of the space of Schwartz functions in our compactified picture:

Exercise 10.10. Show that pullback under the radial compactification map sends $\dot{\mathcal{C}}^\infty(B^n)$ to $\mathcal{S}(\mathbb{R}^n)$.

By (10.7), it is not hard to see that

$$(\rho^2 D_\rho - \alpha)u = 0 \implies \mathrm{WF}_{\mathrm{sc}}\, u \subset \{\rho = 0, \xi = \alpha\},$$

$$(\rho D_{\theta^j} - \beta)u = 0 \implies \mathrm{WF}_{\mathrm{sc}}\, u \subset \{\rho = 0, \eta_j = \beta\}.$$

The following variant provides a useful family of examples (and can be proved with only a little more thought): if $a(\rho, \theta)$ and $\phi(\rho, \theta) \in \mathcal{C}^\infty(X)$, then[55]

$$\mathrm{WF}_{\mathrm{sc}}\left(a(\rho,\theta)e^{i\phi(\rho,\theta)/\rho}\right) = \{(\rho = 0, \theta, d(\phi(\rho,\theta)/\rho)) : (0,\theta) \in \text{ess-supp}\, a\},$$

where $\text{ess-supp}\, a \subseteq \partial X$ denotes the "essential support" of a, i.e. the points near which a is not $O(\rho^\infty)$.

Of course, if

(10.8) $$(\Delta - \lambda^2)u = f \in \dot{\mathcal{C}}^\infty(X),$$

then we have, by microlocal elliptic regularity,

$$\mathrm{WF}_{\mathrm{sc}}\, u \subset \{\rho = 0,\ \xi^2 + |\eta|^2_h = \lambda^2\}.$$

In fact, there is a propagation of singularities theorem for scattering operators of real principal type that further constrains the scattering wavefront set of a solution to (10.8): it must be invariant under the (appropriately rescaled) Hamilton vector field of the symbol of $\Delta - \lambda^2$.

---

[55]The distribution $ae^{i\phi}$ used here is a simple example of a *Legendrian distribution*. The class of Legendrian distributions on manifolds with boundary, introduced by Melrose-Zworski [19], stands in the same relationship to Lagrangian distributions as scattering wavefront set does to ordinary wavefront set.

Exercise* 10.11. Let $\omega = d(\xi\, d\rho/\rho^2 + \eta \cdot d\theta/\rho)$ and let

$$p = \xi^2 + |\eta|_h^2 - \lambda^2;$$

show that up to an overall scaling factor, the Hamilton vector field of p with respect to the symplectic form ω is, on the face, $\rho = 0$ just

$$\mathsf{H}_p = 2\xi\eta \cdot \partial_\eta - 2|\eta|_{h_0}^2 \partial_\xi + \mathsf{H}_{h_0}$$

where $h_0 = h|_{\rho=0}$, and H_{h_0} is the Hamilton vector field of h_0, i.e. (twice) geodesic flow on ∂X.

Show that maximally extended bicharacteristics of H_p project to the θ variables to be geodesics of length π. (Hint: reparametrize the flow.)

(For a careful treatment of the material in this exercise and indeed in this section, see [18].)

APPENDIX

We give an extremely sketchy account of some background material on Fourier transforms, distribution theory, and Sobolev spaces. For further details, see, for instance, [25] or [11].

Let $\mathcal{S}(\mathbb{R}^n)$, the *Schwartz space,* denote the space

$$\{\phi \in \mathcal{C}^\infty(\mathbb{R}^n) : \sup \left| x^\alpha \partial_x^\beta \phi \right| < \infty \ \forall \alpha, \beta\},$$

topologized by the seminorms given by the suprema. The dual space to $\mathcal{S}(\mathbb{R}^n)$, denoted $\mathcal{S}'(\mathbb{R}^n)$, is the space of tempered distributions.

For $\phi \in \mathcal{S}(\mathbb{R}^n)$, let

$$\mathcal{F}\phi(\xi) = (2\pi)^{-n/2} \int \phi(x) e^{-i\xi \cdot x}\, dx.$$

Then $\mathcal{F}\phi \in \mathcal{S}(\mathbb{R}^n)$, too; indeed, $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is an isomorphism, and its inverse is closely related:

$$\mathcal{F}^{-1}\psi(x) = (2\pi)^{-n/2} \int \psi(\xi) e^{+i\xi \cdot x}\, dx.$$

We can, by duality, then define $\mathcal{F}$ on tempered distributions.

Let $\mathcal{E}'(\mathbb{R}^n)$ denote the space of compactly supported distributions on $\mathbb{R}^n$. When X is a compact manifold without boundary, we let $\mathcal{D}'(X)$ denote the dual space of $\mathcal{C}^\infty(X)$.

We define the (L^2-based) Sobolev spaces by

$$H^s(\mathbb{R}^n) = \{u \in \mathcal{S}'(\mathbb{R}^n) : \langle\xi\rangle^s \mathcal{F}u(\xi) \in L^2(\mathbb{R}^n)\},$$

where $\langle\xi\rangle = (1 + |\xi|^2)^{1/2}$. If s is a positive integer, this definition coincides exactly with the space of L^2 functions having s distributional derivatives also lying in L^2. We note that the operation of multiplication by a Schwartz function is a bounded map on each H^s; this is most

easily proved by interpolation arguments similar to (but easier than) those alluded to in Exercise 2.4—cf. [25].

Throughout these notes we will take for granted the *Schwartz kernel theorem,* not so much as a result to be quoted but as a world-view. Recall that this result says *any* continuous linear operator

$$\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$$

is of the form

$$u \mapsto \int k(x,y)u(y)\,dy$$

for a unique $k \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$; a corresponding result also holds on all the manifolds that we will consider. We thus consistently take the liberty of confusing operators with their Schwartz kernels, although we let $\kappa(A)$ denote the Schwartz kernel of the operator A when we wish to emphasize the difference.

Some results relating Schwartz kernels to traces are important for our discussion of the wave trace. Recall that an operator T on a separable Hilbert space is called *Hilbert-Schmidt* if

$$\sum_j \|Te_j\|^2 < \infty$$

where $\{e_j\}$ is any orthonormal basis. In the special case when our Hilbert space is $L^2(X)$ with X a manifold, the condition to be Hilbert-Schmidt turns out to be easy to verify in terms of the Schwartz kernel: T is Hilbert-Schmidt if and only if $\kappa(T)$, its Schwartz kernel,[56] lies in $L^2(X \times X)$.

A *trace-class* operator is one such that

$$\sum_{i,j} |\langle Te_i, f_j \rangle| < \infty$$

for every pair of orthormal bases $\{e_i\}, \{f_j\}$. It turns out to be the case that an operator T is trace-class if and only if it can be written

$$T = PQ$$

with P, Q Hilbert-Schmidt. The *trace* of a trace-class operator is given by

$$\sum_i \langle Te_i, e_i \rangle$$

---

[56]It is probably best to think of X as a *Riemannian* manifold here, so that the Schwartz kernel is a function, which we can integrate against test functions via the metric density, and likewise integrate the kernel.

over an orthonormal basis: this turns out to be well-defined. We refer the reader to [20] for further discussion of trace-class and Hilbert-Schmidt operators.

References

[1] Craig, W., Kappeler, T., Strauss, W. *Microlocal dispersive smoothing for the Schrödinger equation,* Comm. Pure Appl. Math. 48 (1995), no. 8, 769–860.

[2] Dimassi, Mouez; Sjöstrand, Johannes, *Spectral asymptotics in the semiclassical limit,* London Mathematical Society Lecture Note Series, 268. Cambridge University Press, Cambridge, 1999.

[3] Duistermaat, J. J.; Guillemin, V. W. *The spectrum of positive elliptic operators and periodic bicharacteristics,* Invent. Math. 29 (1975), no. 1, 39–79.

[4] Doi, S.-I. *Smoothing effects of Schrödinger evolution groups on Riemannian manifolds,* Duke Math. J. 82 (1996), no. 3, 679–706.

[5] Duistermaat, J. J.; Hörmander, L. *Fourier integral operators. II,* Acta Math. 128 (1972), no. 3-4, 183–269.

[6] Friedlander, F. G., *Introduction to the theory of distributions* Second edition. With additional material by M. Joshi. Cambridge University Press, Cambridge, 1998.

[7] Grigis, A. and Sjöstrand, J., *Microlocal analysis for differential operators. An introduction.* London Mathematical Society Lecture Note Series, 196. Cambridge University Press, Cambridge, 1994.

[8] *A Strichartz inequality for the Schrödinger equation on non-trapping asymptotically conic manifolds* (with Andrew Hassell and Terence Tao), Comm. PDE., **30** (2005), 157–205.

[9] Hörmander, L., *The spectral function of an elliptic operator,* Acta Math. 121 (1968), 193–218.

[10] L. Hörmander, *Fourier Integral Operators I*, Acta Math. 127 (1971), 79–183.

[11] Hörmander, L. *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Second edition. Grundlehren der Mathematischen Wissenschaften, 256. Springer-Verlag, Berlin, 1990.

[12] Hörmander, L. *The analysis of linear partial differential operators. II. Differential operators with constant coefficients.* Grundlehren der Mathematischen Wissenschaften, 257. Springer-Verlag, Berlin, 1983.

[13] Hörmander, L. *The analysis of linear partial differential operators. III. Pseudodifferential operators.* Grundlehren der Mathematischen Wissenschaften, 274. Springer-Verlag, Berlin, 1985.

[14] Hörmander, L. *The analysis of linear partial differential operators. IV. Fourier integral operators.* Grundlehren der Mathematischen Wissenschaften, 275. Springer-Verlag, Berlin, 1985.

[15] Kac, M. *Can one hear the shape of a drum?* Amer. Math. Monthly 73 1966 no. 4, part II, 1–23.

[16] Martinez, André, *An introduction to semiclassical and microlocal analysis,* Universitext. Springer-Verlag, New York, 2002.

[17] R. Melrose *Lecture notes on microlocal analysis,* available at `www-math.mit.edu/~rbm/Lecture_Notes.html`

[18] R. B. Melrose, *Spectral and scattering theory for the Laplacian on asymptotically Euclidian spaces*, Spectral and scattering theory (Sanda, 1992), Dekker, New York, 1994, pp. 85–130.

[19] R. B. Melrose and M. Zworski, *Scattering metrics and geodesic flow at infinity*, Invent. Math. **124** (1996), no. 1-3, 389–436.

[20] Reed, Michael and Simon, Barry, *Methods of modern mathematical physics I: Functional analysis* Second edition, Academic Press, Inc., New York, 1980.

[21] Seeley, R. T., *Complex powers of an elliptic operator,* 1967 Singular Integrals (Proc. Sympos. Pure Math., Chicago, Ill., 1966) 288–307 Amer. Math. Soc., Providence, R.I.

[22] Shubin, M. A., *Pseudodifferential operators and spectral theory,* Second edition. Springer-Verlag, Berlin, 2001.

[23] Stein, E. M. *Singular integrals and differentiability properties of functions* Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J. 1970.

[24] Taylor, M. E. *Pseudodifferential operators,* Princeton Mathematical Series, 34. Princeton University Press, Princeton, N.J., 1981.

[25] Taylor, M. E., *Partial differential equations. I. Basic theory* Applied Mathematical Sciences, 115. Springer-Verlag, New York, 1996.

[26] Taylor, M. E. *Partial differential equations. II. Qualitative studies of linear equations* Applied Mathematical Sciences, 116. Springer-Verlag, New York, 1996.

[27] A. Vasy, *The wave equation on asymptotically Anti-de Sitter spaces,* Anal. PDE, to appear.

[28] M. Zworski, *Semiclassical analysis*, AMS Graduate Studies in Mathematics, American Mathematical Society, Providence, 2012.

DEPARTMENT OF MATHEMATICS, NORTHWESTERN UNIVERSITY, EVANSTON IL 60208