WHAT IS MICROLOCAL ANALYSIS?

COLLIN KOFROTH

Abstract. In this note, we discuss the major themes of microlocal analysis. We will discuss two fundamental objects in the field, namely pseudodifferential operators and wavefront sets. We will spend time motivating the former using classical and quantum mechanics. Next, we define these objects explicitly and outline what one can do with them, then we will demonstrate their utility for tackling a variety of problems. This note is expository note and is kept at the level of an advanced undergraduate to early graduate student level. We will occasionally make references to higher-level concepts for the more advanced reader.

This serves as an expansion of a three-part lecture given at the UNC Graduate Mathematics Association (GMA) visions seminar and UNC graduate analysis seminar on the major themes of microlocal analysis. Graduate students at UNC will notice tools from the first-year analysis, geometry/topology, and methods of applied mathematics course sequences arising throughout.

Contents

1. Introduction

Microlocal analysis, loosely, is a subfield of mathematical analysis where one works in phase space $\mathbb{R}^n_x \times \mathbb{R}^n_\xi$ (more broadly the *cotangent bundle* of a smooth manifold) rather than in configuration space the $\mathbb{R}^n_x$. Here, one typically views the x-variable as representing the position of a particle and the ξ-variable representing its momentum (or its frequency). A bit more technically, one studies both the location and *(co)directions* along which a *distribution* is singular (i.e. not represented by a smooth function). This information is stored in an object called the *wavefront set* (to be discussed in more detail in §4). Roughly speaking, a point (x_0, ξ_0) in phase space is not in the wavefront set provided there is a conic neighborhood around ξ_0 for which the Fourier transform of the distribution, localized near x_0, is rapidly decreasing. This tie-in with the Fourier transform leads to

Date: August 31, 2021.

another fundamental object of importance, namely *pseudodifferential operators*. Such objects are generalizations of Fourier multipliers, where the symbol can now depend on both phase space variables. Pseudodifferential operators can be viewed, in a sense to be defined later, as an encapsulation of the classical-quantum correspondence (sometimes referred to as wave-particle duality).

A particular flavor of microlocal analysis, called *semiclassical analysis*, is more directly related to the classical-quantum correspondence. It develops the same framework as microlocal analysis, but it introduces a small parameter h (motivated by the Planck constant $\hbar$, for a very direct relation to quantum mechanics) and one considers the asymptotics as $h \to 0$ (called the *semiclassical* or *high energy* limit). The addition of the parameter h allows one to explicitly track parameter dependence and, thus, is very useful in subjects such as high-energy eigenfunction asymptotics (and many, many others). We remark that one can transition between the two frameworks using variable rescaling.

Our discussion will center around pseudodifferential operators, wavefront sets, and applications thereof. We will largely stick to the microlocal framework, although the semiclassical one is beneficial for motivation. Microlocal analysis is widely applicable to both linear and non-linear PDE theory, such as general relativity, scattering theory for quantum particles, inverse problems, and fluid dynamics. We will be focused on linear problems. Much of what we do will be on $\mathbb{R}^n$, but many of the objects make perfectly good sense on any smooth manifold M if we replace $\mathbb{R}^n_x \times \mathbb{R}^n_\xi = T^*\mathbb{R}^n$ by T^*M and work in coordinate charts, loosely speaking (if M is non-compact, then extra care is needed in certain places). If one does not feel comfortable with manifolds, one may just think of $\mathbb{R}^n$ (unless we say compact manifold without boundary, in which case it is best to think of the unit sphere $\mathbb{S}^{n-1}$ or the flat torus $\mathbb{T}^{n-1}$).

For more in-depth accounts on microlocal analysis, we refer the reader to [3], [6],[5], [7]. The first three references have the most depth (especially the four-volume treatise [5]), with the three-book series [6] frequently applying the tools to PDE theory (particularly in volumes two and three). The final reference ([7]) is a set of concise, readable, and motivated lectures notes. For a semiclassical approach, see [8] and [2], the former being the standard reference on semiclassical analysis and the latter discussing semiclassical analysis in great generality in Appendix E and applying such tools to the theory of scattering resonances. We will heavily borrow from these resources, especially [6],[7], and [8].

1.1. How to Read.

These notes are primarily written for students who have seen ODE theory and know real analysis. Functional analysis and distribution theory would be useful, but it is okay to frame things in simpler terms for until learning such subjects.

We will be a tad imprecise at times for the sake of exposition, especially for reading by non-analysts. For the same reason, this survey will be less focused on proofs and more focused on general ideas. With that being said, we will utilize certain terminology from analysis and PDE (distribution, self-adjoint, etc.) without definition. The first three sections should be largely understandable without any knowledge of distribution theory.

2. Motivation

Before delving into the key objects, we will start with background on classical and quantum mechanics. For a more in-depth perspective, see [4]. Our goal will be to provide a list of analogues between the two subjects.

2.1. Classical Mechanics.

We begin by recalling Newton's second law of motion on $\mathbb{R}^n$.

Newton's Second Law: Let $x(t) = (x_1(t), \cdots, x_n(t))$ be the position of a classical particle of mass m at time t which is being acted on by a force F. Since $x(t) \in \mathbb{R}^n$, we call $\mathbb{R}^n$ the *configuration space.* Then, Newton's second law of motion tells us that

$$ma = F \iff m\frac{d^2x}{dt^2} = F(x(t)).$$

We denote *momentum* of the particle by

$$\xi := mv = m\frac{dx}{dt}.$$

We can convert Newton's second law in a first-order system of ODEs

$$\begin{cases} \dot{x} = \dfrac{1}{m}\xi \\ \dot{\xi} = F(x(t)) \end{cases}$$

Solutions live in $T^*\mathbb{R}^n := \mathbb{R}^n_x \times \mathbb{R}^n_\xi$, which we call the *phase space.*

Suppose that F is a smooth, conservative force with potential V (that is, $\nabla V = -F$). The expression for the total energy of the particle is given by kinetic+potential energy, i.e.

$$H(x, \xi) := \frac{1}{2}m|v|^2 + V(x) = \frac{|\xi|^2}{2m} + V(x).$$

We will call this the (mechanical) *Hamiltonian.* Observe that we can use H to generate a system of ODEs

$$\begin{cases} \dot{x} = \partial_\xi H = \dfrac{1}{m}\xi \\ \dot{\xi} = -\partial_x H = -\nabla V(x) = F(x), \end{cases}$$

which are the same equations as before! They are referred to as *Hamilton's equations* or the *Hamiltonian system* generated by H.

We will call any smooth function $H : T^*\mathbb{R}^n \to \mathbb{R}$ a Hamiltonian, or a *classical observable.* By observable, we mean that it can be "observed" by taking measurements of our system. For example, one can perform an experiment and calculate the position and the momentum of a particle, after which we can compute the value of the Hamiltonian at the time when we performed the experiment. Any classical observable generates a corresponding Hamilton system, from which we can get the position and momentum of the particle. Conservation of energy means that H is constant along solutions of Hamilton's equations, and one can readily verify that

$$\partial_t H(x(t), \xi(t)) = 0$$

by the chain rule. In the special case of

$$H(x, \xi) = \frac{|\xi|^2}{2m} + V(x),$$

we get the standard law of the conservation of mechanical energy. In particular, trajectories for are represented by the level sets of H, which are called *energy levels.*

Each Hamiltonian H induces a flow

$$\Phi^t_H : T^*\mathbb{R}^n \to T^*\mathbb{R}^n$$

called the *Hamiltonian flow*, which is the solution $(x(t), \xi(t))$ to Hamilton's equations with initial conditions $(x(0), \xi(0)) = (x, \xi)$. That is,

$$\Phi^t_H(x, \xi) = (x(t), \xi(t)).$$

It is also the flow of the *Hamiltonian vector field*

$$X_H = (\partial_\xi H, -\partial_x H),$$

on $T^*\mathbb{R}^n$. In particular, $\Phi_H^t = e^{tX_H}$, with the latter notation being more popular (and suggestive).

The Hamiltonian vector field adds an interesting structure to phrase space. Let X_H be the Hamiltonian vector field corresponding to $H \in C^\infty(T^*\mathbb{R}^n)$, and let f be another function in $C^\infty(T^*\mathbb{R}^n)$. One can readily compute that

$$X_H f = \nabla_x f \cdot \nabla_\xi H - \nabla_\xi f \cdot \nabla_x H.$$

This is called the *Poisson bracket* of f and H, and it is denoted via $\{f, H\}$.

Remark. We are adopting the standard physics notation since we are using physics as a means of motivation. In math, people often denote it by $\{H, f\}$, which is the negative of the above (due to the anti-commutativity of the Poisson bracket which is stated immediately after this remark). If one wishes to use this notation, then one will need to flip signs in some of the equations featuring the Poisson bracket.

The Poisson bracket has many important properties. Here are a few:

(1) Anti-commutativity: $\{f, g\} = -\{g, f\}$
(2) Jacobi identity:

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$$

(3) Commutator relation: $X_{\{f,g\}} = [X_f, X_g]$, where $[\cdot, \cdot]$ denotes the Lie bracket of vector fields

The first two properties, along with obvious bilinearity, form $\{\cdot, \cdot\}$ turns $(C^\infty(T^*\mathbb{R}^n), \{\cdot, \cdot\})$ into a Lie algebra.

It can be readily computed that

$$\partial_t(f \circ \Phi_H^t) = \{f, H\}(\Phi_H^t),$$

which demonstrates that the evolution of an observable along the flow governed by another observable is given by taking their Poisson bracket.

There are other interesting and relevant classical objects that one can talk about such as the natural symplectic structure on $T^*\mathbb{R}^n$ and the diffeomorphisms that preserve this structure (symplectomorphisms/canonical transformations), but we will soldier on to avoid making our list of objects too large. See Chapter 2 in [8] for more on elementary symplectic geometry.

2.2. **Quantum Mechanics.**

Quantum mechanics takes place on Hilbert spaces, particularly the space

$$L^2(\mathbb{R}^n) := \left\{ f : \mathbb{R}^n \to \mathbb{C} : \int_{\mathbb{R}^n} |f(x)|^2 \, dx < \infty \right\},$$

with the Hilbert space structure given by the inner product

$$\langle f, g \rangle = \int_{\mathbb{R}^n} f(x) \overline{g(x)} \, dx.$$

The evolution of a non-relativistic quantum particle of mass m is governed by the solution $\psi_t(x) := \psi(t,x)$ of the Schrödinger equation

$$\begin{cases} i\hbar\partial_t\psi_t = \hat{H}\psi_t, \qquad \hat{H} = -\dfrac{\hbar^2}{2m}\Delta + V(x) \\ \|\psi_0\|_{L^2(\mathbb{R}^n)} = 1, \end{cases}$$

where $\hbar = h/2\pi$ is the reduced Planck's constant, V is the potential, and

$$\Delta = \sum_{j=1}^{n} \partial_j^2$$

is the Laplacian on $\mathbb{R}^n$. We always assume that $\|\psi_0\|_{L^2(\mathbb{R}^n)} = 1$ (focusing on what are called *pure states*). The solution, at least formally, will also be L^2-normalized.

We remark that $\hat{H}$ is a densely-defined, unbounded, self-adjoint linear operator on $L^2(\mathbb{R}^n)$ (we will often not clarify that the operator is linear because they always will be). If one is not comfortable with unbounded operators, let us just say that it is symmetric:

$$\left\langle \hat{H}u, v \right\rangle = \left\langle u, \hat{H}v \right\rangle$$

for all u, v in a nice, dense subspace of $L^2(\mathbb{R}^n)$ such as the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ of rapidly decreasing smooth functions (i.e. smooth functions whose derivatives of every order decay faster than any polynomial).

The operator $\hat{H}$ is called the *quantum Hamiltonian*, and the solution ψ_t to the PDE is called the *wave function* (or quantum state at time t). Importantly, $|\psi_t(x)|^2\,dx$ represents the probability density for the location of the particle. That is, for any Borel measurable set $U \subset \mathbb{R}^n$, the probability that the particle is in B at time t is

$$\text{Prob(particle is in } U \text{ at time } t) = \int_U |\psi_t(x)|^2\,dx.$$

So, we no longer get the exact location of a particle (unlike in classical mechanics), but rather a way to find the probability that the particle is within a given set. This tie-in with probability is why we prefer to work with states of norm one.

The analogues of position and momentum are similarly more complicated. They manifest as two densely-defined, unbounded, self-adjoint linear operators on $L^2(\mathbb{R}^n)$. In particular, we define the *position operator* $Q_j : \mathcal{D} \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ by

$$Q_j(f) = x_j f(x),$$

where

$$\mathcal{D} = \{f \in L^2(\mathbb{R}^n) : x_j f \in L^2(\mathbb{R}^n)\},$$

and the *momentum operator* $P_j : \mathcal{D} \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ by

$$P_j(f) = \hbar D_j f, \qquad D_j = \frac{1}{i}\frac{\partial}{\partial x_j},$$

where

$$\mathcal{D} = \left\{ f \in L^2(\mathbb{R}^n) : \frac{\hbar}{i}\partial_j f \in L^2(\mathbb{R}^n) \right\}.$$

In practice, we often write x_j for Q_j and $\hbar D_j$ for P_j since they are simply multiplication and differentiation operators, respectively. Notice that

$$\hat{H} = \frac{1}{2m}\sum_{j=1}^{n} P_j^2 + V(Q),$$

where $V(Q)f(x) = V(x)f(x)$ acts as a multiplication operator. One can readily see the similar structure to the mechanical Hamiltonian

$$H = \frac{1}{2m}|\xi|^2 + V(x) = \frac{1}{2m}\sum_{j=1}^{n} \xi_j^2 + V(x)$$

from earlier.

Let $\psi_t(x)$ be the wave function of a quantum particle that solves the Schrödinger equation. We define the *expected position* and *expected momentum*, respectively, as the expectation values

$$\langle Q_j \rangle_{\psi_t} = \langle Q_j \psi_t, \psi_t \rangle\,,$$

and

$$\langle P_j \rangle_{\psi_t} = \langle P_j \psi_t, \psi_t \rangle\,.$$

These tell us what we would expect the position and momentum of the quantum particle to be if we performed a large number of measurements. More specifically, we would perform these measurements on an ensemble of particles which are all represented by the same wave function (all in the same state), then average. We can analogously define the expectation value for any (self-adjoint) operator, so long as the right-hand side of the expression is defined. From here, we obtain a quantum version of Hamilton's equation for the mechanical Hamiltonian, in the form of the *Ehrenfest equations*

$$\begin{cases} \partial_t \langle Q_j \rangle_{\psi_t} = \dfrac{1}{m}\langle P_j \rangle_{\psi_t} \\ \partial_t \langle P_j \rangle_{\psi_t} = -\langle \partial_j V \rangle_{\psi_t} \end{cases}$$

Once again, our quantum-mechanical picture transitions into probabilistic statements. This provides further reasoning for the position and momentum operators Q_j and P_j as being the natural *quantizations* of position and momentum x_j and ξ_j (respectively), whatever that means (we will see in a bit).

Remark. The Ehrenfest equations are not a *perfect* analogue, since we would want the right-hand side of the second equation to read $-\partial_j V(\langle Q_j \rangle_{\psi_t})$. We get a more analogous picture if we use the Heisenberg approach, whereby we view the states as fixed and the operators as evolving in time. In this case, self-adjoint operators $\hat{H}$ in the Schrödinger picture are represented as

$$\hat{H}(t) = e^{it\hat{H}/\hbar}\hat{H}e^{-it\hat{H}/\hbar}$$

in the Heisenberg picture. The Hamiltonian structure comes from looking at time evolution of Q_j and P_j, as opposed to expectation values. We remark that both pictures give consistent expectation values, which means that they are the same in a physical sense.

We define the *semiclassical Fourier transform* as

$$(\mathcal{F}_h f)(\xi) = (2\pi h)^{-n/2}\int_{\mathbb{R}^n} e^{-\frac{i}{h}x\cdot\xi} f(x)\,dx,$$

defined on an appropriate class of functions (e.g. $\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^n), L^2(\mathbb{R}^n)$). Via the (corresponding) Plancherel theorem,

$$\langle P_j \rangle_{\psi_t} = \frac{\hbar}{i}\int \partial_j \psi_t(x)\overline{\psi_t(x)}\,dx = \int \xi_j |\mathcal{F}_h \psi_t(\xi)|^2\,d\xi.$$

This suggests that $|\mathcal{F}_h\psi_t(\xi)|^2\,d\xi$ is the joint probability density of the momentum operator $P = (P_1, \cdots, P_n)$. That is,

$$\text{Prob(particle's momentum is in } B \text{ at time } t) = \int_B |\mathcal{F}_h\psi_t(\xi)|^2\,d\xi.$$

In general, we call a densely-defined, self-adjoint operator on $L^2(\mathbb{R}^n)$ a *quantum Hamiltonian* or *quantum observable.* That is, they can be *observed* via computing expectation values. Even in quantum mechanics, one is really measuring a classical observable, and we get a probabilistic value for it via the expectation value of the corresponding quantum observable. We will discuss a methodology to transition from classical to quantum observables shortly.

Let us return to discussing the solution of the Schrödinger equation. Since this is an evolution equation, we can informally write the solution to the Schrödinger equation as

$$\psi_t = e^{-\frac{it}{\hbar}\hat{H}}\psi_0.$$

Naively, we can see that this solves the equation by "taking a derivative in t." We call $U(t) := e^{-\frac{it}{\hbar}\hat{H}}$ the *Schrödinger propagator/flow*, and it can be given meaning as a unitary operator on $L^2(\mathbb{R}^n)$ (hence, solutions to the Schrödinger equation genuinely retain the same L^2 norm). Justifying this exponentiation is highly nontrivial, and it is rooted in the Borel functional calculus afforded by spectral theorem for unbounded, self-adjoint operators (or via Stone's theorem on one-parameter unitary groups). If $\hat{H}$ is bounded (which is rarely the case), one can make this rigorous using a power series, which is more simplistic.

Take another quantum observable $\hat{G}$. Then, we get the following formula for the expectation values of the evolution of $\hat{G}$ under the Schrödinger flow:

$$\partial_t \langle \hat{G} \rangle_{\psi_t} = \left\langle \frac{[\hat{G}, \hat{H}]}{i\hbar} \right\rangle_{\psi_t}.$$

Compare the latter expression to how one described evolution of a classical observable along the classical flow - the Poisson bracket $\{\cdot,\cdot\}$ plays the role of $(i\hbar)^{-1}[\cdot,\cdot]$. To see how these operations are related to position and momentum, we compute the *canonical commutation relations*

$$\begin{aligned}
\{x_j, \xi_k\} &= \delta_{jk} \\
\{x_j, x_k\} &= 0 \\
\{\xi_j, \xi_k\} &= 0 \\
(i\hbar)^{-1}[Q_j, P_k] &= \delta_{jk} I \\
[Q_j, Q_k] &= 0 \\
[P_j, P_k] &= 0,
\end{aligned}$$

where δ_{jk} denotes the Kronecker delta. These relations underscores the relationships between the Poisson bracket and commutator.

The operators Q_j, P_j, and iI generate a $2n + 1$ dimensional Lie algebra called the *Heisenberg algebra*. It is the Lie algebra of the Lie group H_n called the *Heisenberg group*. See [3] for more on this and how it relates to quantization (coming shortly).

Here is a run-down of the classical-quantum correspondences introduced so far.

	Hamiltonian Mechanics	**Quantum Mechanics**
State Space	Phase space $T^*\mathbb{R}^n = T^*\mathbb{R}^n$	$\mathbb{P}L^2(\mathbb{R}^n)$ = unit elements in $L^2(\mathbb{R}^n)$
Observables	$H \in C^\infty(T^*\mathbb{R}^n, \mathbb{R})$	Self-adjoint operators on $L^2(\mathbb{R}^n)$
Position and momentum	x, ξ	Q, P
Flows	Hamilton flows $\Phi_H^t = e^{tX_H}$	Schrödinger flows $U(t) = e^{-\frac{it}{\hbar}\hat{H}}$
Lie algebra structure	Poisson bracket $\{\cdot, \cdot\}$	Commutator $[\cdot, \cdot]$
Evolution of observables	$\partial_t(f \circ \Phi_H^t) = \{f, H\} \circ \Phi_H^t$	$\partial_t \langle \hat{G} \rangle_{\psi_t} = \left\langle (i\hbar)^{-1}[\hat{G}, \hat{H}] \right\rangle_{\psi_t}$

We would like a process to transition from one framework to the other, which we will call *quantization.* A quantization procedure would ideally preserve all of the properties that we have listed so far (as well as a few others which we did not). Let us start with a more reasonable wish list and say that we are looking for a map Op that sends f (for suitable f, say in $\mathcal{S}(T^*\mathbb{R}^n)$) to a linear operator $Op(f)$ which has the following properties:

(1) **Preservation of identity**: $Op(1) = I$
(2) **Preservation of position and momentum**: $Op(x_j) = Q_j$ and $Op(\xi_j) = P_j$
(3) **Preservation of observables**: $Op(f)$ is self-adjoint/symmetric
(4) **Preservation of the algebraic structure**: $Op(\{f, g\}) = (i\hbar)^{-1}[Op(f), Op(g)]$

While this list does not seem too demanding, it turns out that it is. Such a map extends uniquely to $\mathcal{P}_2$, the space of quadratic polynomials in I, Q_j and P_j. In fact,

$$Op(x_j^2) = Q_j^2, \quad Op(\xi_j^2) = P_j^2, \quad Op(x_j\xi_k) = \frac{1}{2}(Q_jP_k + P_kQ_j).$$

However, it fails for any larger polynomial space, which is a pretty sizable failure.

Theorem (Groenewold-Van Hove)**.** *There is no ideal quantization on $\mathcal{P}_n$ for $n \geq 3$.*

A compromise that we will need to be willing to make is that the Poisson bracket-commutator relation holds in a weaker sense. Let us loosen (4) to the requirement that

$$Op(\{f, g\}) = \frac{1}{ih}[Op(f), Op(g)] + \mathcal{O}(h),$$

where $h > 0$ is a small parameter and the big $\mathcal{O}$ is taken in some suitable sense. In view of the Bohr correspondence principle, which states that classical mechanics should be recovered from quantum mechanics in the semiclassical limit $h \to 0$, requiring compatibility in an asymptotic sense is reasonable. Again, we underscore that we are now using h to indicate a small *parameter.*

This is achieved using what is called the *Weyl quantization.* What one essentially does is write the classical observable f down using the Fourier inversion formula

$$f(x, \xi) = \int_{\mathbb{R}^n \times \mathbb{R}^n} e^{\frac{i}{h}(x \cdot y + \xi \cdot \eta)} \mathcal{F}_h f(y, \eta)\, dy d\eta,$$

then swap Q_j for x_j and P_j for ξ_j. Using the formula for the Schrödinger representation of the Heisenberg group

$$e^{-\frac{it}{h}(y \cdot Q + \eta \cdot P)}\varphi = e^{-\frac{it}{h}y \cdot x + \frac{it^2}{2h^2}y \cdot \eta}\varphi(x - t\eta),$$

one gets that

$$Op_h^{\mathrm{w}}(f)u(x) = (2\pi h)^{-n} \iint e^{\frac{i}{h}(x-y)\cdot\xi} f\left(\frac{x+y}{2}, \xi\right) u(y)\, dy d\xi.$$

This is called the *semiclassical Weyl quantization* of the *symbol* f, and it is an example of a *semiclassical pseudodifferential operator*, one of the primary objects of interest in semiclassical analysis.

The fact that properties (1)-(3) are met is immediate, but (4) is less obvious. This will be discussed in the subsequent section. We will largely drop the h and move into the microlocal picture ($h = 1$), but we will make note of the achievement of property (4) when it arises.

3. Pseudodifferential Operators

3.1. **What Are They?** For suitable functions $p(x, \xi)$ and $u(x)$, we will define the operator $P(x, D)$ by

$$P(x, D)u(x) = \int e^{ix\cdot\xi} p(x, \xi)\hat{u}(\xi)\, d\xi,$$

where

$$\hat{u}(\xi) = (2\pi)^{-n} \int e^{-ix\cdot\xi} u(x)\, dx.$$

We call P a *pseudodifferential operator* (which we will often abbreviate as ΨDO) and p its *symbol*. This procedure of associating p to P is called the *standard/left/Kohn-Nirenberg* quantization. Notice that pseudodifferential operators are generalizations of Fourier multipliers, which satisfy the above but whose symbols only depend on ξ.

Example (Variable Coefficient Differential Operators). If

$$P(x, D) = \sum_{|\alpha|=0}^{m} a_\alpha(x) D^\alpha, \qquad a_\alpha \in C_b^\infty(\mathbb{R}^n),$$

then

$$p(x, \xi) = \sum_{|\alpha|=0}^{m} a_\alpha(x)\xi^\alpha.$$

We call

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

the *principal symbol* of P. Note that the case of variable coefficient differential operators tells us about

(1) **Multiplication operators**: For example, if $p(x, \xi) = x_j$, then $P(x, D) = x_j$.
(2) **Differential operators**: For example, if $p(x, \xi) = \xi_j$, then $P(x, D) = hD_j$. Another special case is when $P(x, D) = \Delta$, in which case the symbol is simply $p(x, \xi) = -|\xi|^2$. This is worth pointing out due to how ubiquitous the Laplacian is in PDE theory.

In particular, our definition of pseudodifferential operator preserves the classical and quantum definitions of position and momentum, which gives us property (2) on our wish list. For a less trivial example, let $\chi \in C_c^\infty(T^*\mathbb{R}^n)$ be such that $\chi \equiv 1$ for $|(x, \xi)| \leq 1$ and $\chi \equiv 0$ for $|(x, \xi)| > 2$. Generally speaking, good luck explicitly calculating $\chi(x, D)$ like we could for differential operators or multiplication operators! This is actually a very important example of a ΨDO, as it demonstrates one way to use pseudodifferential operators is to *microlocalize* (localize in phase space). We will see another non-trivial example in the form of an *elliptic parametrix* shortly.

We were rather cavalier and imprecise when we defined ΨDO's. In particular, we should ask when the operator formula actually makes sense (the integral converges and gives a sufficiently nice function). If p were a Schwartz function, then we can readily see via integration by parts that Pu is Schwartz, and so

$$P(x,D) : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n), \qquad p \in \mathcal{S}(T^*\mathbb{R}^n).$$

By duality, $P(x,D) : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$. In fact, the same mapping properties are true if $p \in S^m(T^*\mathbb{R}^n)$, where

$$S^m(T^*\mathbb{R}^n) = \left\{ p \in C^\infty(T^*\mathbb{R}^n) : |\partial_x^\alpha \partial_\xi^\beta p(x,\xi)| \leq C_{\alpha\beta} \langle \xi \rangle^{m-|\beta|} \right\}, \qquad \langle \xi \rangle = \left(1 + |\xi|^2\right)^{1/2}$$

(in fact, we can be more general, but let's not). The set $S^m(T^*\mathbb{R}^n)$ is referred to as the *Kohn-Nirenberg* symbol class of order m, and to each $p(x,\xi) \in S^m(T^*\mathbb{R}^n)$, we associate a unique pseudodifferential operator $P(x,D)$ of order m. We denote the space of m'th order pseudodifferential operators as $\Psi^m(\mathbb{R}^n)$. The idea of the Kohn-Nirenberg class is that differentiation in ξ improves decay, and it generates a theory which is highly amenable to working on manifolds due to coordinate invariance.

We remark that if we instead defined the operator

$$P^{\mathrm{w}}(x,D)u(x) = \int e^{ix\cdot\xi} p\left(\frac{x+y}{2}, \xi\right) \hat{u}(\xi)\, d\xi,$$

then this would be the quantization procedure discussed in the prior section with $h = 1$. This quantization (the *Weyl quantization*) yields a symmetric operator, unlike the previously-discussed form, and the calculus that it enjoys is called the *Weyl calculus.* It is extremely similar to the standard calculus with the upshot of possessing self-adjointness and having better remainders (we will see what that means in a moment). The downside is that it is harder to perform computations. However, it turns out that they generate the same calculus at the *principal level* (we will explain this later, but we essentially mean that the highest-order terms match). Since the sharp remainder improvements are often not needed (especially in basic theory), we will work with the standard quantization.

3.2. **What Can We Do With Them?** We can do numerous fun and important things with pseudodifferential operators. Many of the following expressions will make use of *asymptotic expansions.* In the semiclassical calculus, the asymptotics are in h, whereas the asymptotics in microlocal analysis are in terms of the order of the symbols. In the former context, one makes frequent use of the stationary phase lemma.

1. We can compose them: One of the most common operator between two operators is composition. Given $p_j \in S^{m_j}(T^*\mathbb{R}^n)$, the hope is to find $p_1 \# p_2 = q \in S^{m_1+m_2}(T^*\mathbb{R}^n)$ such that $P_1(x,D)P_2(x,D) = Q(x,D) \in \Psi^{m_1+m_2}(\mathbb{R}^n)$. Our hope not forsaken, and one can find such a q. In fact, one is able to obtain an asymptotic expansion

$$q(x,\xi) \sim \sum_{|\alpha \geq 0|} \frac{i^{|\alpha|}}{\alpha!} D_\xi^\alpha p_1(x,\xi) D_x^\alpha p_2(x,\xi),$$

where $\sim$ is taken in the sense that

$$q(x,\xi) - \sum_{|\alpha|=0}^{N} \frac{i^{|\alpha|}}{\alpha!} D_\xi^\alpha p_1(x,\xi) D_x^\alpha p_2(x,\xi) \in S^{m_1+m_2-(N+1)}(T^*\mathbb{R}^n).$$

That is, the difference between the symbol and the truncated sum is of the same order as the next term in the sum. Notice that the first term is $p_1(x,\xi)p_2(x,\xi)$, and so

$$q(x,\xi) = p_1(x,\xi)p_2(x,\xi) + r(x,\xi), \qquad r \in S^{m_1+m_2-1}(T^*\mathbb{R}^n).$$

That is, at the *principal level*, the composition of the operators is given by quantizing the product of the symbols. Notice that if we look at the symbol for $P_2(x,D)P_1(x,D)$, then we find that the first term in the asymptotic expansion is also $p_1(x,\xi)p_2(x,\xi)$. This tells us that the commutator of $P_1(x,D)$ and $P_2(x,D)$ will be of lower order than their composition.

In particular, we find that

$$[P_1(x,D), P_2(x,D)] = Q(x,D) \in \Psi^{m_1+m_2-1}(\mathbb{R}^n),$$

where

$$q(x,\xi) = i\{p_1,p_2\}(x,\xi) \bmod S^{m_1+m_2-2}(T^*\mathbb{R}^n).$$

If we were working in the semiclassical setting, then we would have

$$q(x,\xi) = ih\{p_1,p_2\}(x,\xi) + \mathcal{O}(h^2),$$

(the error would actually be $\mathcal{O}(h^3)$ in the Weyl calculus). This shows that property (4) on our wish list is, indeed, met by the semiclassical Weyl quantization. In particular, the semiclassical Weyl quantization fulfills our entire wish list, making it an "ideal" quantization procedure (in fact, it preserves many classical-quantum analogues). This is why we say that it encapsulates the classical-quantum correspondence.

2. We can find adjoints: Given $p \in S^m(T^*\mathbb{R}^n)$, we have the expression $(P(x,D))^* = Q(x,D) \in \Psi^m(\mathbb{R}^n)$, where $q \in S^m(T^*\mathbb{R}^n)$ has the expansion

$$q(x,\xi) \sim \sum_{|\alpha|\geq 0} \frac{i^{|\alpha|}}{\alpha!} D_\xi^\alpha D_x^\alpha \overline{p(x,\xi)}.$$

Here, the adjoint is taken in the L^2 sense:

$$\langle Pu, v\rangle = \langle u, P^*v\rangle$$

for all u, v in (a dense subspace of) $L^2(\mathbb{R}^n)$. Notice that the first term in the expansion is given by $\overline{p}$.

3. We can almost invert non-zero symbols: Let $P(x,D) \in \Psi^m(\mathbb{R}^n)$ be *elliptic*. This means that there exist $C, r > 0$ so that

$$|p(x,\xi)| \geq C\langle\xi\rangle^m, \qquad \forall|\xi| \geq r.$$

Then, we can construct a two-sided *parametrix* (approximate inverse) $Q(x,D) \in \Psi^{-m}(\mathbb{R}^n)$, i.e.

$$P(x,D)Q(x,D) = I \bmod \Psi^{-\infty}(\mathbb{R}^n)$$
$$Q(x,D)P(x,D) = I \bmod \Psi^{-\infty}(\mathbb{R}^n),$$

where

$$\Psi^{-\infty}(\mathbb{R}^n) = \bigcap_{m\in\mathbb{R}} \Psi^m(\mathbb{R}^n)$$

denotes the space of *smoothing operators* (sometimes also called the *residual class*).

After using a smooth cutoff that is supported in $\{|\xi| \geq r\}$ and identically 1 on $\{|\xi| \geq 2r\}$, we can define a parametrix by inverting the symbol (multiplied by the cutoff) then using a Neumann series argument to construct the asymptotic expansion. In particular, the parametrix is very much a genuine pseudodifferential operator.

Using this parametrix, one can show the following *elliptic regularity* statement. Recall that the singular support of a distribution is the complement of the largest open set where the distribution is represented by a smooth function (that is, a point is *not* in the singular support if the distribution can be represented as a smooth function on a neighborhood of the point).

Theorem. *Let $P(x,D) \in \Psi^m(\mathbb{R}^n)$ be elliptic and $u \in \mathcal{S}'(\mathbb{R}^n)$. Then,*

$$\text{sing supp}(Pu) = \text{sing supp}(u).$$

In particular, if Pu is smooth, then so is u. An example of an elliptic operator is the Laplacian Δ, and the elliptic regularity result proves that Δ is hypoelliptic.

4. We have nice operator mapping properties: If $m \leq 0$ and $P \in \Psi^m(\mathbb{R}^n)$, then

$$P : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$$

is bounded. Notice that this implies that the Weyl quantization genuinely outputs self-adjoint operators given a Kohn-Nirenberg symbol, which was property (3) on the wish list!

This can actually be improved to operators generated by symbols who are smooth and bounded in all derivatives (and we will denote this symbol class as $S^0_{0,0}(T^*\mathbb{R}^n)$). This is called the *Calderón-Vaillancourt theorem.*

Theorem. *If $p \in S^0_{0,0}(T^*\mathbb{R}^n)$, then $P(x,D) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is bounded.*

By using the Fourier multiplier

$$\Lambda^s u = \mathcal{F}^{-1}\left(\langle\xi\rangle^s \mathcal{F}u\right),$$

we actually get that if $P \in \Psi^m(\mathbb{R}^n)$, then

$$P(x,D) : H^s(\mathbb{R}^n) \to H^{s-m}(\mathbb{R}^n),$$

where $H^s(\mathbb{R}^n) = \Lambda^{-s}L^2(\mathbb{R}^n)$ denotes the L^2-based Sobolev space of order m. If the symbol lies in $S^0_{0,0}(T^*\mathbb{R}^n)$ and decays to 0 as $(x,\xi) \to \infty$, then the operator is also *compact* on $L^2(\mathbb{R}^n)$. By the spectral theorem for compact, self-adjoint operators, P can be diagonalized using an L^2-orthonormal basis of eigenfunctions. In particular, this holds for symbols in $S^m(T^*\mathbb{R}^n)$ for $m < 0$.

Finally, we have the *sharp Gårding inequality*, which tells us that a positive symbol *nearly* gives a positive pseudodifferential operator.

Theorem. *If $P(x,D) \in \Psi^1(\mathbb{R}^n)$ and $\operatorname{Re} p(x,\xi) \geq 0$, then*

$$\operatorname{Re}\langle P(x,D)u, u\rangle_{L^2} \geq -C\,\|u\|_{L^2}.$$

In fact, the Fefferman-Phong inequality gives the above for $P(x,D) \in \Psi^2(\mathbb{R}^n)$. This is very strong, although it is only known to hold for scalar symbols (the previous theorem holds for matrix-valued symbols).

4. Wavefront Sets

4.1. **What Are They?** To any pseudodifferential operator $P(x,D)$ of order m, we have a *principal symbol* (a bit loosely)

$$\sigma(P) = [p] \in S^m(T^*\mathbb{R}^n)/S^{m-1}(T^*\mathbb{R}^n).$$

This is unique modulo lower-order terms. Using the principal symbol map gives a short exact sequence

$$0 \to \Psi^{m-1}(\mathbb{R}^n) \to \Psi^m(\mathbb{R}^n) \to S^m(T^*\mathbb{R}^n)/S^{m-1}(T^*\mathbb{R}^n) \to 0.$$

This tells us, for example, that a principal symbol of order m equals zero (as an equivalence class) if and only if the operator is in $\Psi^{m-1}(\mathbb{R}^n)$. Notice that all of our prior rules hold at the principal level, e.g. $\sigma(P_1P_2) = \sigma(P_1)\sigma(P_2)$. The principal symbol of the Weyl quantization and standard quantization of a symbol are the same, so they generate the same calculus at the principal level. We also remark that if we are using the semiclassical framework, then the principal symbol is defined modulo $hS^{m-1}(T^*\mathbb{R}^n)$ symbols. In particular, the standard and Weyl calculi match up to $\mathcal{O}(h)$ terms.

Next, define the *characteristic set* (or characteristic variety) as

$$\mathrm{Char}(P) = \{(x,\xi) \in T^*\mathbb{R}^n \setminus o : \sigma(P) = 0\},$$

where $o = \{(x,\xi) \in T^*\mathbb{R}^n : \xi \neq 0\}$ denotes the zero section. Its complement is called the *elliptic set.*

Definition. The wavefront set of a tempered distribution $u \in \mathcal{S}'(\mathbb{R}^n)$ is the set

$$\mathrm{WF}(u) = \bigcap \left\{\mathrm{Char}(P) : P \in \Psi^0(\mathbb{R}^n), Pu \in C^\infty(\mathbb{R}^n)\right\} \subseteq T^*\mathbb{R}^n \setminus o.$$

Equivalently, $(x_0, \xi_0) \notin \mathrm{WF}(u)$ if and only if there exist $\chi \in C_c^\infty(\mathbb{R}^n)$ with $\chi(x_0) \neq 0$ and a conic neighborhood Γ of ξ_0 such that for all N, there exists a C_N so that

$$|\mathcal{F}(\chi u)(\xi)| \leq C_N \langle \xi \rangle^{-N} \qquad \forall \xi \in \Gamma.$$

The latter definition should be compared to the Paley-Wiener theorem, which asserts that the Fourier transform provides an isomorphism between smooth, compactly supported functions on $\mathbb{R}^n$ and entire functions satisfying a growth condition similar to the above.

Note that $\mathrm{WF}(u)$ is a closed, conic subset of $T^*\mathbb{R}^n \setminus o$. The conic nature means that we can regard it as a subset of the *cosphere bundle* $S^*\mathbb{R}^n = \{(x,\xi) \in T^*\mathbb{R}^n : |\xi| = 1\}$. In this perspective, the wavefront set is compact in ξ, which can be a very useful fact to leverage.

Example.

(1) If $u = \delta_{x_0}$, then $\mathrm{WF}(u) = \{(x_0, \xi) : \xi \neq 0\}$.
(2) If u is smooth, then $\mathrm{WF}(u) = \emptyset$.
(3) Let Ω be a bounded domain with smooth boundary. If $u = \mathbb{1}_\Omega$, then $\mathrm{WF}(u) = N^*\partial\Omega \setminus o$, where $N^*\partial\Omega$ denotes the *conormal bundle* of $\partial\Omega$.

4.2. **What Do They Tell Us?** As seen by the definition of the wavefront set, it tells us both where our distribution fails to be represented as a smooth function and in which direction such a singularity propagates (and the examples provide further aid in that intuition). Here, we state some important results pertaining to wavefront sets. The first states that the singular support is the projection of the wavefront set onto the x variable. Since the singular support is the collection of bad points, it makes sense that we get it after projecting away the directions.

Proposition. *If* $\Pi_x : \mathbb{R}_x^n \times \mathbb{R}_\xi^n \to \mathbb{R}_x^n$ *denotes the projection map onto the first coordinate* $\Pi_x(x,\xi) = x$, *then*

$$\text{sing supp}(u) = \Pi_x\left(\mathrm{WF}(u)\right).$$

The next states that smooth functions have empty wavefront set (there are no bad points nor directions) *and* vice-versa.

Proposition. $\mathrm{WF}(u) = \emptyset$ *if and only if* $u \in C^\infty(\mathbb{R}^n)$.

Now, we state an extremely important property of the wavefront set.

Theorem. *Let* $P \in \Psi^m(\mathbb{R}^n)$. *Then,*

$$\mathrm{WF}(u) \subseteq \mathrm{WF}(Pu) \cup \mathrm{Char}(P).$$

As a corollary, we get the statement called *microlocality* (which is an improvement of the analogous statement for singular supports, namely *pseudolocality*). It tells us that all of the interesting information lies within the characteristic set.

Corollary. *Let* $P \in \Psi^m(\mathbb{R}^n)$*. If* $Pu \in C^\infty(\mathbb{R}^n)$*, then*

$$\mathrm{WF}(u) \subseteq \mathrm{Char}(P).$$

If $Pu = 0$, for example, then we know that that bad points and directions of our distribution lie in the set where the principal symbol vanishes.

The last result is a statement called *microlocal regularity* (which is stronger than the previous statement on elliptic regularity).

Corollary. *If* $P \in \Psi^m(\mathbb{R}^n)$ *is elliptic, then*

$$\mathrm{WF}(Pu) = \mathrm{WF}(u).$$

In particular, if $Pu = f \in C^\infty$, then $\mathrm{WF}(u) = \emptyset$, which implies that u is smooth as a consequence of our prior proposition (which is consistent with the elliptic regularity statement).

One can also use wavefront sets to determine conditions under which one can multiply distributions (see [5]). The condition, called the *Hörmander criterion*, stipulates that the wavefront sets of the two distributions cannot contain opposite directions (if v is a direction in one wavefront set, then $-v$ cannot be in the other).

5. Applications

Finally, we provide a few applications of ΨDO's and wavefront sets. There are many more, but three applications seems reasonable.

1. Index Theory: Non-analysts might find it interesting that ΨDO's are used in popular proofs of the Atiyah-Singer index theorem. This theorem asserts that, for an elliptic (pseudo)differential operator on a smooth, compact manifold without boundary, the analytical index (Fredholm index) of the operator is equal to its topological index (more complicated definition). Even if one does not understand the prior jargon exactly, one can still appreciate the implications of the result; the theorem establishes a very general correspondence between analytical and topological data, and we can compute one by computing the other. It might feel reminiscent of the (Chern-) Gauss-Bonnet theorem, and in fact, the Atiyah-Singer index theorem implies this result. It turns out to be more beneficial to work with ΨDO's in the proof due to the fact that there far more ΨDO's than differential operators, making them less restrictive to work with (e.g. elliptic parametrices are pseudodifferential, even for elliptic *differential* operators). That is fairly reductive, but index theory in out of my personal depth. See the second volume of [6] for a proof in the case where the elliptic operator is a twisted Dirac operator.

2. Spectral Theory: An important result in spectral theory is the *Weyl law*, which provides asymptotics on the number of eigenvalues less than a fixed number (energy level) λ. The asymptotics as $\lambda \to \infty$ are equivalent (via scaling) to the high-energy limit $h \to 0$. There are many versions of the Weyl law, but we will only state one (perhaps, we should call our version *a* Weyl law). It is proved using microlocal methods. To start, let (M, g) be a compact Riemannian manifold without boundary, and $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \nearrow \infty$ denote the eigenvalues (with multiplicity) of the negative Laplacian

$$-\Delta_g = -\frac{1}{\sqrt{\det g}} \sum_{i,j=1}^{n} \frac{\partial}{\partial x^i} \left(g^{ij} \sqrt{\det g} \frac{\partial}{\partial x^j} \right),$$

where (g^{ij}) denotes the ij-th component of the inverse metric tensor. Then, the Weyl law is stated as follows.

Theorem (Weyl)**.** *Let (M, g) be a compact Riemannian manifold without boundary, and define the spectral counting function $N(\lambda)$ via $N(\lambda) = \#\{\lambda_j \le \lambda\}$. Then,*

$$\begin{aligned} N(\lambda) &= (2\pi)^{-n} \operatorname{Vol}(B^*M)\lambda^{n/2} + \mathcal{O}(\lambda^{(n-1)/2}) \\ &= (2\pi)^{-n} \operatorname{Vol}(\{(x,\xi) \in T^*M : |\xi|_g^2 \le \lambda\}) + \mathcal{O}(\lambda^{(n-1)/2}). \end{aligned}$$

The volume given to the unit co-ball B^*M is the Liouville measure inherited from T^*M.

For a microlocal proof of this result, see [7]; this proof is rather complicated, as it uses the wave trace. For a proof of the semiclassical variant, see [8], which uses the Helffer-Sjöstrand functional calculus (one can also see this set of lecture notes on the aforementioned subject which is based off of [8] but is somewhat shorter and is a bit simpler due to the lack of potential). A nice resource on eigenfunctions of the Laplacian can be found in [1]. It also features an elementary proof on a rectangle.

Example. Let $M = \mathbb{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2 = \mathbb{S}^1 \times \mathbb{S}^1$. Then, the eigenvalues and eigenfunctions of $-\Delta$ are given by

$$\lambda_{jk} = j^2 + k^2, \qquad j,k \in \mathbb{Z}$$

and

$$\varphi_{jk}(x,y) = e^{i(jx+ky)}, \qquad j,k \in \mathbb{Z}$$

respectively. Suppose that we want to see the accuracy of the Weyl law for $N(100)$. We can compute (say using a computer) that $N(100) = 317$. If we compute the leading term in the Weyl law, then we get

$$(2\pi)^{-2} \operatorname{Vol}(\{(x,\xi) \in T^*\mathbb{R}^2 : x \in \mathbb{T}^2,\ |\xi|^2 \le 5\}) = (2\pi)^{-2}((2\pi)^2(100\pi) = 100\pi \approx 314.$$

In particular, the relative error in estimating $N(100)$ using the leading order term in the Weyl law is

$$\frac{|317 - 100\pi|}{317} \approx 9 \cdot 10^{-3}.$$

In general, computing $N(\lambda)$ is an open problem which is related to the *Gauss circle problem.*

3. PDE theory: Previously, we described the wavefront set as encoding the points and directions along which a distribution is singular. This can be seen very plainly in the context of PDEs, where we can show that singularities propagate along the Hamiltonian flow generated by the principal symbol of the operator with initial data in the characteristic set.

Consider the pseudodifferential equation

$$Pu = f, \qquad u \in \mathcal{D}'(M), \qquad P \in \Psi^m(M),$$

where M is a smooth manifold. Call $p = \sigma(P)$, which we will assume is real-valued and positively homogeneous of degree m in ξ (such as a homogeneous polynomial of degree m in ξ). We call an integral curve of the Hamiltonian vector field X_p in $\operatorname{Char}(P)$ a *null bicharacteristic*. Every null bicharacteristic has initial data in $\operatorname{Char}(P)$, and the flow leaves $\operatorname{Char}(P)$ invariant (since $X_p p = \{p,p\} = 0$, it is constant along integral curves).

Theorem (Propagation of Singularities)**.** *Suppose that $P \in \Psi^m(M)$ has real-valued principal symbol, $u \in \mathcal{D}'$, and $Pu = f$. Then,* $\operatorname{WF}(u) \setminus \operatorname{WF}(f) \subset \operatorname{Char}(P)$ *is a union of (maximally extended) null bicharacteristics of P. Said another way,* $\operatorname{WF}(u) \setminus \operatorname{WF}(f)$ *is invariant under the flow generated by X_p within* $\operatorname{Char}(P)$.

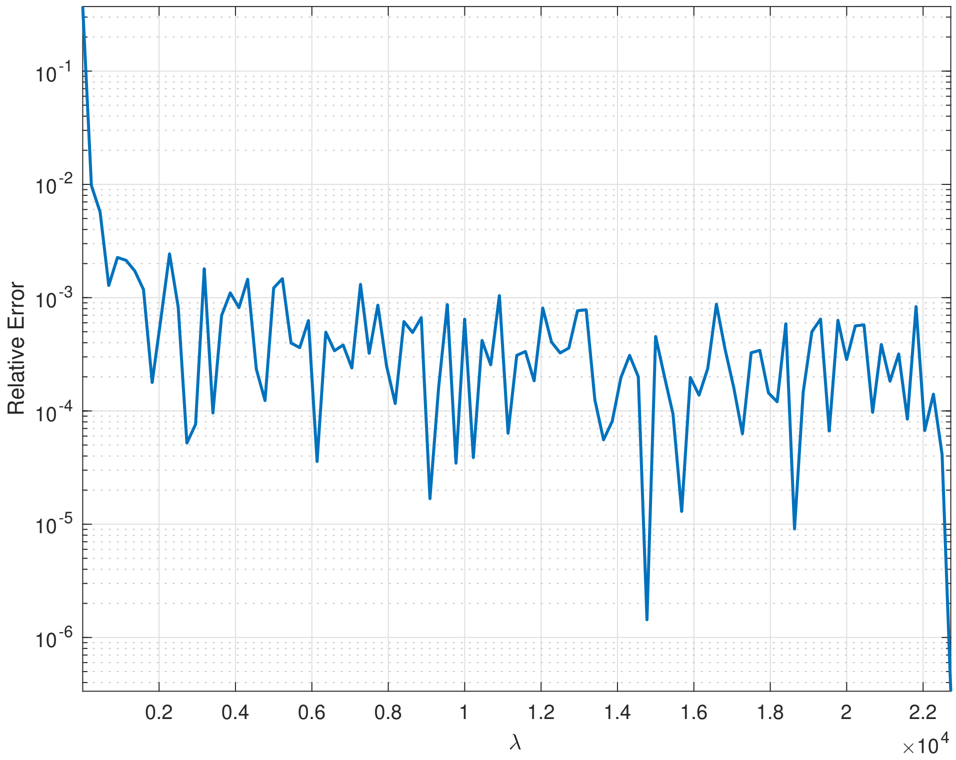

FIGURE 1. Semilog scale plot (in y) for the relative error in estimating $N(\lambda)$ by the leading order term in the Weyl law

Corollary. *If $f \in C^\infty(M)$, then $\mathrm{WF}(u)$ is a union of (maximally extended) null bicharacteristics of P. Said another way, $\mathrm{WF}(u) \subset \mathrm{Char}(P)$ is invariant under the flow generated by X_p within $\mathrm{Char}(P)$.*

One can find a proof of the propagation of singularities in e.g. [5],[6], [7], [8], and [2] (the latter two being semiclassical).

Remark. If $Pu = f$ is an evolution equation with initial data $u(0)$, then an application of the above theorem yields that

$$e^{tX_p}\,\mathrm{WF}(u(0)) = \mathrm{WF}(u(t)),$$

where X_p denotes the Hamiltonian flow projected onto the (x, ξ) variables.

This theorem demonstrates that singularities are propagated along the flow lines generated by X_p. To see this as a statement on *regularity*, notice that if we take $z \in \mathrm{Char}(P) \setminus \mathrm{WF}(u)$, then the corresponding null bicharacteristic through z cannot enter $\mathrm{WF}(u)$ (if it did, then take that as the new initial condition and apply the previous theorem and ODE uniqueness theory to get a contradiction). In particular, this is a *good direction.*

It is proved using a *positive commutator argument*, which is a generalization of the energy method from classical PDE theory to pseudodifferential multipliers. In essence, one looks for a symbol which is positive along X_p in an appropriate sense, then one applies pseudodifferential calculus (e.g. composition rule, ΨDO's mapping properties, the

Gårding inequality). Positive commutator arguments where one utilizes pseudodifferential multipliers is a fantastic example of why ΨDO's are so useful (explicit calculation of the operator from the symbol may well be impossible).

Example. Consider the wave equation

$$\Box u = f, \qquad \Box = \partial_t^2 - c^2\Delta,$$

where $c > 0$. The principal symbol of $\Box$ is $p(t, x, \tau, \xi) = -\tau^2 + c^2|\xi|^2$. This has characteristic set

$$\mathrm{Char}(\Box) = \{(t, x, \tau, \xi) \in T^*\mathbb{R}^n \setminus o : \tau^2 = c^2|\xi|^2\}.$$

Notice that, when projected to the dual variables, this is the *light cone.* We claim that the null bicharacteristics project to straight lines (more broadly, lifts of geodesics to T^*M traversed forward and backward in time when $\mathbb{R}^n$ is replaced by a Riemannian manifold M) at speed c. Indeed, the Hamiltonian system generated by p is

$$\begin{cases} \dot{t}^s = -2\tau^s \\ \dot{\tau}^s = 0 \\ \dot{x}^s = 2c^2\xi^s \\ \dot{\xi}^s = 0 \\ (t^s, \tau^s, x^s, \xi^s)\big|_{s=0} = (t_0, \tau_0, x_0, \xi_0). \end{cases}$$

This has the solution

$$(t^s, \tau^s, x^s, \xi^s)(t_0, \tau_0, x_0, \xi_0) = (t_0 - 2\tau_0 s, \tau_0, x_0 + 2c^2\xi_0 s, \xi_0),$$

In order for this to be a null bicharacteristic, we require that $\tau_0 = \pm c|\xi_0|$ (that is, the initial data lies on the forward/backward light cone). In particular,

$$(t^s, \tau^s, x^s, \xi^s)(t_0, \pm c|\xi_0|, x_0, \xi_0) = (t_0 \mp 2c|\xi_0|s, \pm c|\xi_0|, x_0 + 2c^2\xi_0 s, \xi_0).$$

Let us eliminate the s variable in favor of the time variable t. That is, if we call $t^s = t$, then

$$s = \mp\frac{t - t_0}{2c|\xi_0|}.$$

Then, if we write $\tilde{x}^t = x^s$, we obtain that

$$\tilde{x}^t(t_0, \pm c|\xi_0|, x_0, \xi_0) = x_0 \mp \frac{c\xi_0}{|\xi_0|}(t - t_0).$$

In particular, $\tilde{x}^t$ travels in the direction $\mp\xi_0/|\xi_0|$ at speed c, traversed both forward and backward in time. The statement that singularities/disturbances (and regularity) propagate along the null bicharacteristics indicates that they live on the light cone and propagate forward and backward in time along geodesics - along *light rays.*

References

1. Y. Canzani, *Notes for analysis on manifolds via the laplacian*, `https://canzani.web.unc.edu/wp-content/uploads/sites/12623/2016/08/Laplacian.pdf`.
2. S. Dyatlov and M. Zworski, *Mathematical theory of scattering resonances*, Graduate Studies in Mathematics, vol. 200, American Mathematical Society, Providence, RI, 2019.
3. G.B. Folland, *Harmonic analysis in phase space*, Annals of Mathematics Studies, vol. 122, Princeton University Press, Princeton, NJ, 1989.
4. B. C. Hall, *Quantum theory for mathematicians*, Graduate Texts in Mathematics, no. 267, Springer, New York, 2013.
5. L. Hörmander, *The analysis of linear partial differential operators, I-IV*, Springer-Verlag, New York, 1985.

6. M. E. Taylor, *Partial differential equations, I-III*, second ed., Applied Mathematical Sciences, vol. 115-117, Springer, New York, 2011.
7. J. Wunsch, *Microlocal analysis and evolution equations*, preprint, arXiv:0812.3181, 2008.
8. M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics, vol. 138, Amer. Math. Soc., 2012.

Department of Mathematics, CB 3250, University of North Carolina, Chapel Hill, NC, 27599
Email address: ckofroth@live.unc.edu