Introduction to Semiclassical Analysis

Stepan Malkov

Last Updated: January 2026

This is an expository document with some background in symplectic geometry and quantum mechanics, as well as notes on quantization and semiclassical analysis taken during a reading course. These notes largely follow Zworski's "Semiclassical Analysis."

Contents

1 Symplectic Geometry Primer

1.1 Primer on Manifolds

We briefly review some manifold theory.

Definition 1.1.1. A **derivation** D is a function on an algebra that satisfies the product rule, i.e. $D(fg) = (Df)g + f(Dg)$. An **antiderivation** on a graded algebra satisfies the graded product rule $D(fg) = (Df)g + (-1)^{\deg f} fDg$. The **degree** of a derivation is m if $\deg Df = \deg f - m$.

A **vector bundle of rank** n is an assignment $\pi : E \to M$ on a (C^∞) manifold M of an n-dimensional vector space to each point of the manifold such that $\phi : \pi^{-1}(U) \cong U \times k^n$ locally and restricts to a vector space isomorphism on the **fibers** $\pi^{-1}(\{m\}), m \in M$. ϕ is called a **local trivialization** of π, and E, M are called the **total and base space** of the bundle, respectively. Note that E can be given a manifold structure. Given two vector bundles E, F, one can naturally construct the **direct sum** bundle $E \oplus F$, the **tensor product bundle** $E \otimes F$, the **Hom bundle** $\mathrm{Hom}(E, F)$, and the **dual vector bundle** E^* over M. The **trivial bundle** is globally isomorphic to $M \times k^n$.

If M is a (C^∞) manifold, then for $p \in M$, then **tangent space** T_pM is the vector space of **derivations** on $C^\infty(M)$, i.e. linear functionals satisfying the product rule. The **tangent bundle of** M is the corresponding vector bundle, defined as

$$TM := \{(p, \eta) : p \in M, \eta \in T_pM\}.$$

A **vector field** on M is a **section** of the tangent bundle, i.e. a continuous inverse $X : M \to TM$ to the natural projection $\pi : TM \to M$, with $X_p = \eta$. One writes $X \in \Gamma(TM)$. Alternatively, a vector field is a derivation $X : C^\infty(M) \to C^\infty(M)$. The **cotangent bundle** T^*M is the corresponding dual bundle, defined as

$$T^*M := \{(p, \eta^*) : p \in M, \eta^* \in T_p^*M\},$$

and a **differential** k**-form** ω is a section of the k-th exterior power $\Lambda^k(T^*M)$, that is, an alternating multilinear map on TM. In particular, 0-forms are smooth functions on M, $T^*M = \Lambda^1(M)$, and there is a wedge product operation $\Lambda : \Lambda^k(T^*M) \times \Lambda^l(T^*M) \to \Lambda^{k+l}(T^*M)$.

Definition 1.1.2. We define the **exterior derivative** $d : \Lambda^k(T^*M) \to \Lambda^{k+1}(T^*M)$ as the unique linear map such that:

(a) $(df)(X) = X(f)$ for all vector fields X on M.

(b) $d \circ d = 0$.

(c) d is an antiderivation of degree 1, i.e. $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^{\deg \alpha} \alpha \wedge d\beta$.

Given a map $F : M \to N$ between manifolds and a k-form ω on N, one may define the **pushforward** $F_*\eta$ as an element of $T_{F(p)}N$ such that

$$(F_*\eta)(f) = \eta(f \circ F).$$

pullback $F^*\omega$ as a k-form on M such that

$$(F^*\omega)(Y) = \omega(F_*Y).$$

Note that the pushforward of a vector field might not necessarily be a vector field. A vector field X locally induces a **smooth-one parameter family of diffeomorphisms** $\Phi : \mathbb{R} \to \mathbf{Diff}(M)$ such that

$$\Phi'(0) = X$$

A vector field that induced a global flow, that is, a flow defined for all $t \in \mathbb{R}$, is called a **complete vector field.** Given two vector fields X, Y, their **Lie bracket** is another vector field defined as

$$[X, Y](f) = X(Y(f)) - Y(X(f)).$$

The Lie bracket measures the lack of commutativity between the two flows. Now, if one wants to define a directional derivative on a manifold, one needs a vector field to define a direction. Thus, we define the **Lie derivative** of a function $f \in C^\infty(M)$ with respect to a vector field X to be

$$\mathcal{L}_X(f)(x) := \lim_{h\to 0} \frac{f(\Phi^h(x)) - f(x)}{h} = X(f)(x),$$

where Φ is the flow generated by X. Similarly, one may define the Lie derivative of a vector field by

$$\mathcal{L}_X(Y) = [X, Y].$$

Even more importantly, we can define the Lie derivative of a differential form. For that, we first define the **contraction** of a form $i_X : \Lambda^k(M) \to \Lambda^{k-1}(M)$ as an antiderivation such that

$$i_X(\omega)(\cdot) = \omega(X, \cdot),$$

with $i_X f = 0$ by convention. Then, the Lie derivative treats the contraction using the Leibniz rule, i.e.

$$\mathcal{L}_X(i_Y\omega)(\cdot) = i_{\mathcal{L}_X Y}\omega(\cdot) + i_Y(L_X\omega)(\cdot).$$

For instance, for a 1-form ω,

$$\mathcal{L}_X(\omega) = X(\omega(Y)) - \omega([X, Y]).$$

Moreover, notice that

$$\mathcal{L}_X f = i_X df = df(X) = X(f).$$

Formally, the Lie derivative of a differential form is defined as

$$\mathcal{L}_X\omega = \lim_{h\to 0} \frac{\Phi^{h*}\omega - \omega}{h},$$

which is nothing but the "derivative" of the form. Note that since the exterior derivative commutes with evaluation and pullbacks, it commutes with the Lie derivative. Then, we have the following "magical" fact.

Theorem 1.1.1 (Cartan's Magic Formula)**.** $\mathcal{L}_X\omega = di_X\omega + i_X d\omega$.

Proof. Both sides of the formula are derivations that commute with d, so it suffices to prove the identity for 0 forms, which follows immediately. ☐

The power of differential forms is that one can integrate them. Given a smooth manifold M of dimension n covered by a partition of unity (U_i, ϕ^*), and an n-form ω on M we define the integral of ω over M to be

$$\int_U \omega := \sum_i \int_{U\cap U_i} \phi_i^*\omega,$$

where if $U \subset \mathbb{R}^n$ is open, $\int_U f dx_1 \wedge ... \wedge dx_n := \int_U f dx_1...dx_n$. In other words, cover the manifold by charts, pullback to Euclidean space, and integrate regularly. We define a manifold to be **orientable** if there exists a form on it that never vanishes. In particular, orientation is determined by the sign of a form. A **manifold with boundary** is a manifold M with a boundary ∂M which is locally diffeomorphic to the half space $\mathbb{R}^n \cap \{x \leqslant 0\}$. The importance of these concepts is arguably the most fundamental theorem of differential topology:

Theorem 1.1.2 (Generalized Stokes Theorem)**.** *If M is a orientable manifold of dimension n with boundary (with the induced orientation) and ω is a compactly supported $(n-1)$-form on M, then $\int_M d\omega = \int_{\partial M} \omega$.*

A differential form ω is **closed** if $d\omega = 0$ and exact if $\omega = d\alpha$ for some form α. Clearly, every exact form is closed. The form

$$\frac{ydx}{x^2+y^2} - \frac{xdy}{x^2+y^2}.$$

on $\mathbb{R}^2\backslash\{0\}$ can be checked to be closed. However, by Stokes's theorem, if it were exact, it would integrate to 0 over S^1, which it does not. However, one does have the following nice fact:

Theorem 1.1.3 (Poincare Lemma)**.** *Every closed form on an open ball in $\mathbb{R}^n$ is exact.*

Proof. WLOG, take the unit ball. We construct a linear operator $\alpha : \Lambda^k \to \Lambda^{k-1}$ such that $d\alpha + \alpha d = 1$. For closed ω, define

$$(\alpha\omega)(X) = \int_0^1 t^{k-1}(i_X)\omega(tX)dx$$

□

1.2 Riemannian and Pseudo-Riemannian Manifolds

Definition 1.2.1. A **(pseudo-)Riemannian manifold** is a manifold M equipped with a smooth positive (semi)definite symmetric bilinear form $g_p : T_pM \times T_pM \to \mathbb{R}$. The bilinear form $g = (g_{ij})$ on a Riemannian manifold induces a norm $T_pM \to \mathbb{R}$ by $\|v\| = \sqrt{g(v,v)}$.

Example 1.2.1. If $U \subset \mathbb{R}^n$, $(U, \cdot)$ with the dot product is a Riemmanian manifold.

Example 1.2.2. Let $i : S \hookrightarrow (M, g)$ be an embedded submanifold of a (pseudo-)Riemannian manifold. Then, the restriction (S, i^*g) is a (pseudo-)Riemannian manifold.

Definition 1.2.2. A diffeomorphism $f : (M, g_1) \to (N, g_2)$ is called an **isometry** if $f^*g_2 = g_1$.

Definition 1.2.3. The bilinear form g on a Riemannian manifold M induces a length on C^1 curves $\gamma : [0,1] \to M$ by

$$L(\gamma) = \int_0^1 \sqrt{g(\gamma'(t), \gamma'(t)}dt,$$

and a corresponding metric

$$d(x,y) = \inf_{\gamma(0)=x,\gamma(1)=y} L(\gamma).$$

A curve γ from x to y is called a **geodesic** if $d(x,y) = L(\gamma)$.

1.3 Primer on Lagrangian and Hamiltonian Mechanics

One of the main reasons for studying symplectic geometry comes from Lagrangian and Hamiltonian mechanics. In elementary classical physics, one describes the behavior of a physical system S using Newton's second law

$$F = ma = m\ddot{x},$$

where $F = F(x, \dot{x}, t)$ is determined solely based on the positions and velocities x of the particles. In principle, this implies that the state of S is uniquely determined by its position and velocity. We now recall that from the forces acting on S one is able to extract two quantities - the **kinetic energy** $T = T(\dot{x}, t)$ and the **potential energy** $V = V(x, t)$. We now restrict ourselves to considering two particular combinations of these values, the **Lagrangian**

$$L(x, \dot{x}, t) = T - V = \text{kinetic energy - potential energy}$$

and **Hamiltonian**

$$H(x, \dot{x}, t) = T + V = \text{kinetic energy - potential energy}.$$

In the classical case of a particle in a potential, recall that $T = \frac{1}{2}m\|\dot{x}\|^2$ and $V = V(x)$. Now, it turns out (by miracle of some sort) that the Euler-Lagrange equations for the critical points of the **action functional**

$$\mathcal{L}(s) = \int_0^s L(x, \dot{x}, t)dt,$$

which are

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} = 0,$$

in the classical case become

$$-\partial_{x_i} V - m\ddot{x}_i = 0,$$

which is identical to the motion predicted by Newton's second law. Motivated by this example, for a system with $n = N - C$ degrees of freedom (where C is the number of constraints), we introduce the **generalized positions/coordinates** $q = (q_i)_{i=1}^n$ and the corresponding Lagrangian $L = L(q, \dot{q}, t)$. Note that under this formulation, the Euler-Lagrange equations are a typically a system of n second-order ODE in q_i. Moreover, by defining the **generalized momenta** $p_i := \frac{\partial L}{\partial \dot{q}_i}$, we may write these as a system of $2n$ first-order ODE for (p_i, q_i) in terms of the Hamiltonian H, namely,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

Proof. Given a Lagrangian L, the Hamiltonian may be defined as the Legendre transform of L, so it satisfies

$$L(q, \dot{q}, t) + H(p, q, t) := \langle p, \dot{q} \rangle.$$

Then,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \dot{p}_i = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} = -\frac{\partial H}{\partial q_i}.$$

□

Remark 1.3.1. Note that by chain rule,

$$\frac{\partial L}{\partial q}\dot{q} + \frac{\partial L}{\partial \dot{q}}\ddot{q} + \frac{\partial L}{\partial t} + \frac{\partial H}{\partial p}\dot{p} + \frac{\partial H}{\partial q}\dot{q} + \frac{\partial H}{\partial t} = \langle \dot{p}, \dot{q} \rangle + \langle p, \ddot{q} \rangle,$$

which implies

$$\frac{\partial H}{\partial t} + \frac{\partial L}{\partial t} = 0.$$

In particular, H is constant w.r.t. time iff L is.

Example 1.3.1. Note that in the classical example, $H = \langle m\dot{x}, \dot{x} \rangle - (\frac{1}{2}m\|\dot{x}\|^2 - V(x)) = \frac{1}{2}m\|\dot{x}\|^2 + V(x)$ is precisely the total energy of the system, which should be the typical interpretation of the Hamiltonian.

2 Quantum Mechanics Primer

2.1 Wave Primer

Let ψ be a fixed sinusoidal wave with wavelength λ. We then define its **wave number** $k := \frac{2\pi}{\lambda}$, so that $e^{ik\cdot x}$ has wavelength λ. If ψ oscillates sinusoidally in time with period T, define the **frequency and angular frequency** to be $f := \frac{1}{T}, \omega := 2\pi f$, so that the general formula for a wave with wavelength λ and period T is

$$\psi(x,t) = e^{ik\cdot x - \omega t}.$$

If $\omega^2 = c|k|^2$, then ψ solves the classical **wave equation** $\partial_{tt}\psi = c^2\Delta\psi$. We refer to different solution $\omega = \omega(k)$ of this relation as the different **modes** of the equation. Then, $\psi(x,t) = \psi(x - vt, 0)$ for any vector v such that $v \cdot k = \omega$, so the normal **phase velocity** of a wave with wavevector k is $v_p := \frac{\omega}{|k|}\widehat{k}$. This is the "speed" at which the crests of a wave travel. Now, for a general wavepacket ψ decomposed into its constituent waves, if wavenumber k travels at an angular frequency of $\omega = \omega(k)$, the **group velocity** of the packet can be computed as

$$\int A(k)e^{i(k\cdot x - \omega t)}dk \approx e^{i(k_0\cdot x - \omega_0 t)} \int A(k)e^{i(k-k_0)\cdot(x - \nabla\omega_0 t)}dk,$$

and is defined to be $v_g = \nabla\omega$. The group velocity is the velocity of the envelope, i.e. the term on the right in the expansion above, and holds asymptotically as the wave becomes mostly monochromatic. For example, if $\omega(k) = c|k|$, the group velocity is $c\widehat{k}$. Notice that if ω changes nonlinearly with k, then different waves move with different speeds, a phenomenon known as **dispersion,** and the function $\omega(k)$ gives one the **dispersion relation** for the waves at hand. To that end, we define a **dispersive equation** to be one where the dispersion relation satisfies

$$\det D^2\omega(k) \not\equiv 0.$$

For example, the wave equation is not dispersive, while the heat and Schrodinger equation are dispersive.

2.2 Mathematical Formalism

We now consider the following abstract picture: Let $(M, \mathcal{L})$ be a mechanical system with Lagrangian $\mathcal{L}$, where $M = M(q)$ is a manifold of even dimension describing the phase space of a system of n particles with positions q and momenta $\dot{q}$. The tangent bundle $TM = TM(q, \dot{q})$ contains the position and velocity information for the system.

Definition 2.2.1. If M is a smooth manifold,we define a **Lagrangian** as an element $L \in C^\infty(TM)$. For $m \in M$, we can then define a function

$$L_m : T_m M \to \mathbb{R}, v_m \to L(v_m),$$

whose Frechet derivative is

$$L'_m : T_m M \to T^*_m M, v \to L'_m(v) \in T^*_m M.$$

This induces a map of vector bundles

$$F_L : TM(q, \dot{q}) \to T^*M(q, p), \dot{q} \to p$$

where p are the **generalized momenta** of the system. Compare this to the regular definition $\dot{q} \to p = \frac{\partial L}{\partial \dot{q}}$.

Theorem 2.2.1. *If L is convex and bounded below by a positive quadratic on each fiber, then $F : TM \to T^*M$ is a diffeomorphism. Moreover, if one defines $H \in C^\infty(T^*M)$ according to $H(q,p) = pF_L^{-1}(p) - L$ and uses the standard double dual identification, then $F_H = F_L^{-1}$. H is known as the* ***Hamiltonian.***

Example 2.2.1. Let $L(v,x) = \frac{1}{2}mv^2 - V(x)$ on $M = \mathbb{R}, TM = TM^* = \mathbb{R} \times \mathbb{R}$. Then, the Lagrangian is convex, and so

$$F : TM \to T^*M, (v, x) \to (mv, x)$$

is a diffeomorphism (where the action of mv on TM is to be interpreted as multiplication by mv.) Then, H is precisely the standard Hamiltonian $H(p,x) = \frac{p^2}{2m} + V(x)$.

For a given set of coordinates q on M, note that TM^* has a canonical symplectic structure $\omega = dq \wedge dp$ as described above. Then, the Hamiltonian vector field X of the Hamiltonian H precisely generates Hamilton's equations of motion. More generally, one interested in the space of **observables** $C^\infty(M)$. Now, the one-parameter group of diffeomorphisms Φ_t generated by X induces a flow g_t of some observable g, given by the Lie derivative

$$\mathcal{L}_X(g_t) = X(g_t) = \omega(H_{g_t}, X) := \{H_{g_t}, X\} = \{g_t, H\},$$

where we define the **Poisson bracket** of two vector fields as above. The conclusions from this are as follows:

(a) An observable is time-independent iff it commutes with the Hamiltonian under the Poisson bracket. In particular, one has that the evolution of observables is given by **Liouville's equation**

$$g_t = e^{iL_g t} g_0$$

for the operator $L = -i\{g, H\}$. In other words, the total derivative

$$\frac{dg}{dt} = \frac{\partial g}{\partial t} + \frac{\partial g}{\partial q}\dot{q} + \frac{\partial g}{\partial p}\dot{p} = 0.$$

(b) If ω^n is the volume form on T^*M, the Hamiltonian flow is a symplectomorphism and therefore preserves the form, which is known as **Liouville's theorem.** In other words, volumes in phase space are preserved under Hamilton's equations.

2.3 Postulates of Quantum Mechanics

As a physical theory, quantum mechanics has provided some of the most accurate predictions science has ever made. In this section, we briefly provide a rigorous mathematical formalism, with physical justification, for quantum mechanics.

The fundamental idea is that of **wave-particle duality,** which is extremely simple - every particle is simultaneously a **wavefunction/matter wave/wave packet** $\psi \in S^1 \subset L^2(\mathbb{C}^n)$. How could one then describe the physical properties of a particle through this lens? The fundamental postulate and equation that developed this theory was the axiom of Einstein's energy-momentum equation for a particle of rest mass m_0, namely,

$$E^2 = (pc)^2 + (m_0c^2)^2.$$

In particular, for a massless particle, $E = pc$, and in the nonrelativistic limit $v \ll c$,

$$E = \gamma m_0 c^2 = m_0c^2\sqrt{1 + \left(\frac{p}{m_0c}\right)^2} \approx m_0c^2 + \frac{p^2}{2m_0^2c^2}m_0c^2 = m_0c^2 + \frac{1}{2}m_0v^2,$$

where γ is the **Lorentz factor.** How would this energy correspond to the wavefunction of a particle? This is another fundamental postulate of Planck and Einstein, which states that the energy of the wavefunction is directly proportional to its frequency, namely,

$$E = h\nu = \hbar\omega,$$

where $\nu(\omega)$ is the (angular) frequency of the photon and $\hbar = \frac{h}{2\pi}$ is the **reduced Planck's constant.** Experimentally, this certainly was found to be an exact relation for photons, and de Broglie's insight was to extend it to arbitrary wavefunctions.

Can we obtain a formula for the momentum of a wave using this relation? The key relation that must hold here is that the velocity of a particle should correspond to the group velocity of its wavepacket, i.e.

$$v_g = \frac{pc^2}{E} = \frac{d\omega}{dk} = \frac{d\left(m_0c^2\sqrt{1 + \left(\frac{p}{m_0c}\right)^2}\right)}{\hbar dk} = \frac{pc^2}{E\hbar}\frac{dp}{dk},$$

which yields the fundamental momentum relation $p = \hbar k$. In other words, this is the momentum relation that needs to be satisfied if energy is given by the Planck-Einstein equation and the velocity of the particle is the group velocity of its wavepacket.

We may now use these two relations to derive the dispersion relation

$$\omega(k) = \sqrt{|k|^2c^2 + \left(\frac{m_0c^2}{\hbar}\right)^2} \approx \frac{m_0c^2}{\hbar} + \frac{\hbar|k|^2}{2m_0},$$

which (in the nonrelativistic limit) yields the group velocity $v_g = \frac{|k|c^2}{\omega} \approx \frac{\hbar|k|}{m_0}$. What is the conclusion here? Well, assuming that the energy of a wave is proportional to its frequency, we obtain that momentum, and therefore velocity, is proportional to its wavenumber.

Now finally, how does one then determine the expected position x of a particle given its wavefunction? Well, interpreting $|\psi|^2$ as a probability distribution for the particle, one obtains

$$\langle x \rangle = \int x|\psi|^2 = \langle \psi|x|\psi \rangle$$

for $(x\phi)(x) = x\phi(x)$. We thus call x the (unbounded) **position operator.** In fact, $|\psi(x)|^2$ is the probability distribution for the particle position and thus encodes all its information. How can we determine the probability distribution of k from the wavefunction ψ? Well, the scaled Fourier transform $\widehat{\psi}(\frac{k}{\hbar})$ can be easily seen to give the distribution of momenta for each of the waves making up the wavepacket. By Plancherel,

$$\langle p \rangle = \int \widehat{\psi}^* \hbar k \widehat{\psi} = \int \psi^*(-i\hbar\partial_x)\psi,$$

so $p = -i\hbar\partial_x$ is known as the **momentum operator.** Additionally, one can see that since

$$\int_{\mathbb{R}^{n+1}} \widehat{\psi}^* \hbar k \widehat{\psi} dk = \int_{\mathbb{R}^n} |\phi|^2(p) p dp$$

for $\phi(p) = \hbar^{-n}|\widehat{\psi}(\frac{p}{\hbar})|^2$, that ϕ is the probability distribution for the momentum p. Finally, note that the nonrelativistic energy of a wavepacket (in a zero potential) is given by its kinetic energy $E_k = \frac{1}{2}\frac{p^2}{2m}$, so the total energy can be computed as

$$\langle H \rangle = \left\langle \frac{p^2}{2m} + V(x) \right\rangle = \int \widehat{\psi}^* \left(\frac{1}{2m}\hbar^2 k^2 + V(x) \right) \widehat{\psi} = \int \psi^* \left(-\frac{\hbar^2}{2m}\Delta + V(x) \right) \psi,$$

and we define $H := -\frac{\hbar^2}{2m}\Delta + V(x)$ to be the **Hamiltonian operator**. The probability distribution for H is given by the probability distribution for $\frac{p^2}{2m} + V(x)$, which is

$$\Phi(E) = \frac{1}{2\sqrt{E}}(\phi(\sqrt{E - V(x)}) - \phi(-\sqrt{E - V(x)}))$$

for ϕ as above and $E > 0$.
In this formalism, one sees that a if a measurement λ of an operator A on ψ is exact, then ψ is an eigenvector of A with eigenvalue ψ. What can one say in general about the probability distribution of some observable A? Well, it turns out that these operators can be decomposed into their orthonormal eigenstates ϕ (since they have the nice property of being self-adjoint), in which case one gets according to Parseval that the "proportion" of a state ψ that lives in a state ϕ_n with eigenvalue λ_n is $|\langle \psi, \phi_n \rangle|^2$, and so the probability of $\lambda_n \in X$ should equal $\sum_{\lambda_n \in X} |\langle \psi, \phi_n \rangle|^2$. In general, this leads to the formalism of **positive operator-valued measures (POVM),** which we do not discuss here.

With this being said, we are now ready to describe the fundamental postulates of quantum mechanics:

(a) A quantum particle is described by a complex wavefunction $|\psi\rangle \in L^2(M)$ (also called a **pure state**) on a manifold M such that $\||\psi\rangle\|_2 = 1$ and $\||\psi\rangle\|_{L^2(A)}$ gives the probability of finding the particle in the region A. In particular, $e^{i\theta}|\psi\rangle$ represents the same physical system for any $\theta \in \mathbb{R}$, so mathematically, the space of states can be identified with the **projective Hilbert space** $H\mathbb{P}$. From this perspective, the complex phase of a wavepacket is only relevant when considering interactions with other wavepackets. Working with complex functions in L^2 allows us to take Fourier transforms and inner products.

(b) It follows that the natural formalism for describing an ensemble of quantum particles on Hilbert spaces must be linear and respect constants. One thus defines the total state space for an ensemble of particles to be the **tensor product** of the corresponding Hilbert spaces with the corresponding inner product.

(c) The evolution of every state $|\psi\rangle$ is governed by a unique set of **linear** operators $U_t : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ in a way that must necessarily preserve the total probability being 1 (so by the polarization identity it must also preserve inner products), and lose no information, i.e. be **time-reversible** (supported by physical experiments and classical mechanics considerations). It follows that the time evolution operator U_t must be a bounded linear unitary operator. In view of Stone's theorem on strongly continuous one-parameter unitary semigroups, there thus exists a unique self-adjoint (possibly unbounded) operator $H : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ known as the **Hamiltonian**, and the evolution of $|\psi\rangle$ is given by the **time-dependent Schrodinger equation**

$$i\hbar\partial_t|\psi\rangle = H|\psi\rangle.$$

(d) More generally, any one-parameter family of physical transformations T_r of a quantum system is time-reversible and preserves probability, so it is unitary and corresponds to a self-adjoint operator A known as an **observable.** For example, the generator of physical translations of a system $T_r(\psi(x)) = \psi(x + r)$ is the momentum operator p, and the generator of time translations is the **energy operator** $E := i\hbar\partial_t$. The **spectrum** of A is the set of possible values of the observable, where an eigenvector $|\psi\rangle$ with eigenvalue λ is then a state with the value of the observable exactly equal to λ.

(e) Given a physical time-invariant transformation T and the corresponding observable A, if the state $|\psi\rangle$ is T-invariant, one has

$$e^{i\hbar At}|\psi\rangle = e^{i\hbar\lambda(t)}|\psi\rangle$$

for all t. Since A is time-invariant and $A|\psi\rangle = \frac{\lambda(t)}{t}|\psi\rangle$, so $\lambda(t) = \lambda_A t$. λ_A is then known as the value of the observable A on $|\psi\rangle$. This is an example of **Noether's theorem**, which states that invariance under a physical transformation corresponds to a conserved physical quantity. In particular, since A is self-adjoint, one verifies by the spectral theorem that for such a state $|\psi\rangle$, one has

$$\lambda_A = \langle\psi|A|\psi\rangle,$$

so one defines the **expectation value of** A to be $\langle A\rangle := \langle\psi|A|\psi\rangle = \lambda_A$. More generally, if $|\psi\rangle$ is an aribtrary state and $P_{A,\lambda}$ is the projection onto the eigenspace corresponding to measurement λ of A, then the probability of measuring λ as the value of observable A for $|\psi\rangle$ is $\langle\psi|P_{A,\lambda}|\psi\rangle$, a postulate known as **Born's rule.**

(f) Two observables A, B share a eigenbasis iff they commute. If this holds, the observables are said to be **compatible,** and measurement of one does not affect measurement of the other. Moreover, if the variance of an observable X is given by $\sigma_X^2 = \langle(X - \langle X\rangle)^2\rangle$, Cauchy-Schwarz and the polarization identity yield that for two observables A, B,

$$\sigma_A^2\sigma_B^2 \geqslant \langle(A - \langle A\rangle)\psi, (B - \langle B\rangle)\psi\rangle = \left|\frac{1}{2}\langle AB + BA\rangle - \langle A\rangle\langle B\rangle\right|^2 + \left|\frac{1}{2i}\langle AB - BA\rangle\right|^2,$$

which is known as the **Heiseberg uncertainty principle.** The Fourier transform inequality is a concrete example of the uncertainty principle for position and momentum operators.

(g) Given a Hamiltonian H and a decomposition $|\phi\rangle = \sum a_n(t)|n\rangle$ in terms of an arbitrary orthonormal basis $|n\rangle$ for $L^2(\mathbb{C}^n)$ for some particle with corresponding wavefunction $|\phi\rangle$, direct manipulation yields that

$$\frac{\partial\langle\phi|H|\phi\rangle}{\partial a_n^*} = \langle n'|H|\phi\rangle.$$

If we now define $\pi_n = i\hbar a_n^*$, we recognize that the Schrodinger equation is equivalent to

$$\frac{\partial\langle H\rangle}{\partial \pi_n} = \frac{\partial a_n}{\partial t}, \frac{\partial\langle H\rangle}{\partial a_n} = -\frac{\partial \pi_n}{\partial t},$$

which yields a Hamiltonian system with classical Hamiltonian $\langle H\rangle$, explaining why H is called the Hamiltonian operator.

We now make a crucial remark: for any classical particle with wavefunction ψ, the classical Hamiltonian $\langle H\rangle$ as above is the nonrelativistic total energy $\frac{p^2}{2m} + V(x)$ of the particle, which is the expectation of the operator $-\frac{\hbar^2}{2m}\Delta + V(x)$ as shown before. In particular, since two operators whose expectations agree for all states are identical, the fundamental consequence is that *the time-independent Schrodinger equation for a particle in a potential V is*

$$i\hbar\partial_t\psi = \left(-\frac{\hbar^2}{2m}\Delta + V(x)\right)\psi.$$

(h) **Ehrenfest's Theorem** recovers the classical limit of the formalism through the identity

$$\frac{d}{dt}\langle A\rangle = \frac{1}{i\hbar}\langle [A,H]\rangle + \langle \partial_t A\rangle,$$

which may be derived directly from the time-dependent Schrodinger equation. For instance, it implies that the energy E of a quantum state with a time-independent Hamiltonian is constant. Additionally, it implies that conserved observables of time-independent operators are precisely those that commute with the Hamiltonian (i.e. a measurement is well-defined in time iff the corresponding operator commutes with time evolution). In particular, applying Ehrenfest's theorem to the momentum and position operators yields

$$m\frac{d}{dt}\langle x\rangle = \langle p\rangle, \frac{d}{dt}\langle p\rangle = -\langle \nabla V(x)\rangle.$$

For a classical particle,

$$\langle \nabla V(x)\rangle \approx \nabla V(\langle x\rangle) \approx \nabla V(x_0),$$

which recovers classical Newtonian mechanics $p = mv, F = -\nabla V$.

(i) For a pure state $|\psi\rangle$, define the corresponding projection operator $P_\psi := |\psi\rangle\langle\psi|$ onto $|\psi\rangle$. Then, for an ensemble of states $|\psi\rangle$, a **density operator** ρ is a convex combination of P_ψ. It is easy to see using the spectral theorem that ρ is a density operator iff it is a positive semi-definite self-adjoint operator of trace 1. Then, for an observable A and an ONB $|n\rangle$,

$$\langle\psi|A|\psi\rangle = \sum_n \langle\psi|A|n\rangle\langle n||\psi\rangle = \mathrm{Tr}(|\psi\rangle\langle\psi|A),$$

where $\mathrm{Tr}(B) = \sum_n \langle n|B|n\rangle$, which gives the interpretation of ρ_ψ as a probability distribution for A with respect to $|n\rangle$ (motivated by the formula $\langle A\rangle = \int p(A)Adn$.) This formula can be seen to hold by linearity for density operators. **Gleason's Theorem** is a fundamental result which states that for any function p that meaningfully assigns probabilities to projection operators, there exists a positive semi-definite self-adjoint operator of trace 1, i.e. a density operator ρ such that $p(P) = \mathrm{Tr}(\rho P)$.

Remark 2.3.1. **Wigner's Theorem** states that transformations of the projective Hilbert space $H\mathbb{P}$ that satisfy Born's rule are precisely the unitary or antiunitary operators. However, since composition of antiunitary operators is unitary, it follows that any antiunitary transformation must be discrete (and thus not physical). Concrete examples exist, however - for example, time reversal or momentum reversal symmetry.

Remark 2.3.2. In general, $\langle A \rangle \langle B \rangle \neq \langle AB \rangle$. However, in the classical limit,

Remark 2.3.3. In the relativistic regime, the total energy is given by

$$E^2 = (pc)^2 + (mc^2)^2.$$

Interpreting this equation as an operator equation yields the **Klein-Gordon equation** for relativistic quantum particles

$$\left(\frac{1}{c^2}\partial_t^2 - \Delta + \left(\frac{mc}{\hbar}\right)^2\right)\psi = 0$$

This equation turns out to be great at describing the behavior of spin-0 particles. Factoring it yields the **Dirac equation**

$$(E - c\alpha p - \beta mc^2)\psi = 0,$$

where $\alpha_1, \alpha_2, \alpha_3$ are anticommuting Hermitian matrices. The Dirac equation effectively describes the behavior of spin-$\frac{1}{2}$ particles.

Remark 2.3.4. In classical mechanics, measurements are commutative. But in quantum mechanics, a **measurement** consists of an interaction with the system, which changes the wavefunction and thus the state of the system. For instance, a **polarizer** measures the orientation of light. Thus, the order of measurements matters, as demonstrated in the following experiment: a vertical, horizontal, and diagonal polarizer (in that order) let through no light, while a vertical, diagonal, and horizontal polarizer let through some of the light (since the diagonal polarizer changes the polarization of the light to be diagonal). This corresponds to the fact that observables in general do not commute.

Remark 2.3.5. One sees a wide variety of connections in the evolution laws between quantum and classical mechanics. However, QM is framed in the language of operators on Hilbert spaces, and classical mechanics is framed in terms of Lagrangians on phase space. A theory that recovers classical behavior from quantum behavior in the limit of some small parameter is known as **quantization.** If x, p are position and momentum, $\{x, p\} = 1$ classically, and $[x, p] = i\hbar$ in QM. Unfortunately, according to **Groenewold's theorem,** there cannot exist a quantization map Q that satisfies

(a) $Q_x\psi = x\psi, Q_p\psi = -i\hbar\nabla\psi$.

(b) $f \to Q_f$ is linear.

(c) $[Q_f, Q_g] = i\hbar Q_{\{f,g\}}$.

(d) $Q_{g\circ f} = g(Q_f)$.

In fact, any three of these conditions are inconsistent. The best modern approach is to let 1 and 2 be true and let 3 be true asymptotically as $\hbar \to 0$, an approach known as **deformation quantization.** Alternatively, one may restrict the space of observables up to quadratic terms gets a theory known as **geometric quantization.**

In Table 1, we list some classical observables and their classical analogues, as well as the corresponding symmetries.

Observable	Classical Formula	Quantum Operator	Evolution Operator/Symmetry
Position	x	x	Phase Rotation
Momentum	$p = mv$	$p = -i\hbar\nabla$	Spatial Translation
Kinetic Energy	$E_k = \frac{1}{2}mv^2$	$E_k = -\frac{\hbar}{2m}\Delta$	N/A
Total Energy	$H = \frac{1}{2}mv^2 + V(x)$	$H = -\frac{\hbar}{2m}\Delta + V(x)$	Time Evolution
Angular Momentum	$L = r \times p$	$L = -i\hbar(x \times \nabla)$	Spatial Rotation
Lorentz		Lorentz Symmetry	N/A

Table 1: Classical Observables and their Corresponding Symmetries

2.4 Symplectic Geometry

Now, consider a smooth manifold M with a **symplectic form,** i.e. a closed, non-degenerate differential 2-form ω, non-degenerate meaning $\omega_p(x, y) = 0$ for all $y \in T_pM$ implies $x = 0$. Such a pair (M, ω) is called a **symplectic manifold.**

Example 2.4.1. Consider $U \subset \mathbb{R}^{2n}$ open with coordinates (x, ξ), and define a skew-symmetric bilinear form $\omega(x, \xi) = x^T J\xi$, where

$$J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}.$$

Then, U is a symplectic manifold, and in fact

$$\omega = d\xi \wedge dx := \sum_{i=1}^{n} d\xi_i \wedge dx_i.$$

Example 2.4.2. More generally, consider M is a manifold of dimension n, and take the natural projection map $\pi : N = T^*M \to M$, which induces the maps $d_n\pi : T_nN \to T_mM$. For $n \in N$, we define the **tautological one-form** $\theta \in T^*N$ at $n \in T^*M = N$ to be

$$\theta_n = n \circ d_n\pi = d_n\pi^* n.$$

Then, the **canonical symplectic form** on the total space (i.e. $2n$-dimensional space) T^*M is the form $\omega = -d\theta$.

Example 2.4.3. If $(M_1, \omega_1), (M_2, \omega_2)$ are symplectic manifolds and $\lambda_1, \lambda_2 \in \mathbb{R}\backslash\{0\}$, then $(M_1 \times M_2, \lambda_1\pi_1^*\omega_1 + \lambda_2\pi_2^*\omega_2)$ is a symplectic manifold, where π_1, π_2 are the natural projection maps.

Remark 2.4.1. By a classical linear algebra result, every odd-dimension skew-symmetric matrix is singular, implying that every symplectic manifold has even dimension.

If $\kappa : (M, \omega) \to (N, \omega')$ is a diffeomorphism between symplectic manifolds that preserves the symplectic form, i.e. $\kappa^*\omega' = \omega$, one calls κ a **symplectomorphism.** A vector field X is called symplectic if $\mathcal{L}_X\omega = 0$.

Proposition 2.4.1. *TFAE:*

(a) $X \in \Gamma(TM)$ *is symplectic.*

(b) $\Phi_t : M \to M$ *is a symplectomorphism for all* t*.*

(c) $i_X\omega$ *is closed.*

Proof. If Φ_t is a symplectomorphism for all t, then

$$\mathcal{L}_X\omega = \lim_{t\to 0}\frac{\Phi_t^*\omega - \omega}{t} = 0$$

since $\Phi_t^*\omega = \omega$ for all t. Conversely, if X is symplectic, it suffices to show $\frac{d}{dt}\Phi_t^*\omega = 0$. But indeed,

$$\frac{d}{dt}\Phi_t^*\omega = \lim_{h\to 0}\frac{\Phi_h^*\Phi_t^*\omega - \Phi_t^*\omega}{h} = \lim_{h\to 0}\frac{\Phi_t^*(\Phi_h^*\omega - \omega)}{h} = \Phi_t^*\mathcal{L}_X\omega = 0.$$

Now, from Cartan's magic formula and the fact ω is closed, $di_X\omega + i_X d\omega = 0$ is equivalent to $d(i_X\omega) = 0$. □

Note that a nondegenerate 2-form ω induces an isomorphism of bundles

$$\widetilde{\omega} : TM \to T^*M : X_p \to \omega(X_p, \cdot).$$

2.5 Hamiltonian Vector Fields

The isomorphism above has an inverse

$$\omega^{-1} : T^*M \to TM,$$

which induces a map $\Gamma(T^*M) = \Lambda^1(M) \to \Gamma(TM)$. In particular, every 1-form yields a vector field X.

Definition 2.5.1. For a 0-form $f \in C^\infty(M)$, the **Hamiltonian vector field of** f is uniquely given by $H_f = \omega^{-1}(df)$. Concretely, H_f is a vector field such that

$$\omega(H_f, Y) = df(Y) = Y(f),$$

or equivalently, $i_{H_f}\omega = df$.

Example 2.5.1. With the standard symplectic form $d\xi \wedge dx$ on $U \subset \mathbb{R}^n$, the Hamiltonian vector field of f is

$$\left(\sum_{i=1}^n d\xi_i \wedge dx_i\right)(X, \sum a_i\partial_{x_i} + b_i\partial_{\xi_i}) = \sum_{i=1}^n \partial_{x_i} f dx_i(X) + \partial_{\xi_i} f d\xi_i(X),$$

so $a_i = \partial_{\xi_i} f, b_i = -\partial_{x_i} f$, i.e. $H_f = f_\xi\partial_x - f_x\partial_\xi$.

Example 2.5.2. If one takes the function $f(x,\xi) = \frac{1}{2}\|\xi\|^2 + V(x)$ (motivated by energy conservation from physics), then the Hamiltonian flow generated by f is

$$\dot{x} = \xi, \dot{\xi} = -\nabla V(x),$$

which is equivalent to Newton's laws

$$\dot{x} = v, \dot{v} = -\nabla V.$$

Proposition 2.5.1. *The flow generated by a Hamiltonian vector field H_f is a symplectomorphism.*

Proof. It suffices to show that $i_{H_f}\omega$ is closed. Indeed, $d(i_{H_f}\omega) = d^2 f = 0$. □

Theorem 2.5.1 (Jacobi's Theorem)**.** *If κ is a symplectomorphism, then $H_f = \kappa_* H_{\kappa^* f}$.*

Proof. Indeed, unpacking definitions,

$$\omega(H_{\kappa^* f}, Y) = Y(\kappa^* f) = Y(f \circ \kappa),$$

so

$$\kappa^* \omega(H_{\kappa^* f}, Y) = \omega(\kappa_* H_{\kappa^* f}, \kappa_* Y) = \omega(H_{\kappa^* f}, Y) = Y(f \circ \kappa) = \kappa_* Y(f) = \omega(H_f, \kappa_* Y),$$

and by the nondegeneracy of ω, we are done. □

Theorem 2.5.2 (Darboux's Theorem)**.** *If η is a closed nondegenerate 2-form, then locally, there exists a symplectomorphism κ such that $\kappa^* \eta = \omega$.*

2.6 Lagrangian Manifolds

Definition 2.6.1. An **isotropic submanifold** Λ of a symplectic manifold is a manifold such that $\omega|_\Lambda = 0$. An isotropic submanifold of a manfiold M of dimension $\frac{1}{2} \dim M$ is called a **Lagrangian submanifold.** Equivalently, Λ is Lagrangian iff for $i : \Lambda \to M$ being the inclusion map, $i^* \omega = 0$.

Example 2.6.1. If M is a manifold and T^*M is given the canonical symplectic structure, the zero section $\Lambda = \{(x, \xi) : \xi = 0\}$ is a Lagragian submanifold.

Proposition 2.6.1. *Lagrangian submanifolds of the form $M_{\xi_0} := \{(x, \xi) \in T^*M : \xi = \xi_0\}$ are in bijection with closed 1-forms on M.*

Proof. Consider the pair of maps $s, \pi : M \tilde{\to} T^*M$ mapping $x \to (x, \xi_0)$, and let θ be the tautological one-form on T^*M. Then, for $m \in M$

$$(s^* \theta)_m(v) = \theta_{(m, \xi_0)}(ds_* v) = d\pi^* \xi_0 ds_* v = d(s \circ \pi)^* \xi_0(v) = \xi_0(v),$$

i.e. $s^* \theta = \xi_0$. Now, note that $M_{\xi_0} \cong M \overset{s}{\hookrightarrow} T^*M$, so

$$M_{\xi_0} \text{ Lagrangian } \iff i^* d\theta = 0 \iff s^* d\theta = d(s^* \theta) = d\xi_0 = 0.$$

□

Proposition 2.6.2. *Let $\Lambda \subset M$ be a Lagrangian submanifold. Then, $\Omega := \xi dx$, the form satisfying $d\Omega = \omega$, can be locally represented on Λ as $\Omega = d\phi$ for some smooth ϕ.*

Proof. Let $\gamma : U \subset M \to B(0, 1)$ be a smooth diffeomorphism. Then,

$$d((\gamma^{-1})^* \Omega) = (\gamma^{-1})^* \omega = 0,$$

so by Poincare's theorem, $(\gamma^{-1})^* \Omega = d\phi$ for some smooth ϕ. Pulling back to U concludes the proof. □

We now show that a Lagrangian submanifold is locally the graph of a generating function.

Theorem 2.6.1. *If $\Lambda \subset M$ is a Lagrangian submanifold, then on some open U,*

$$\Lambda \cap U = \{(x', -\partial_{\xi''} \phi; \partial_{x'} \phi, \xi''\} \cap U$$

for a smooth function $\phi = \phi(x', \xi'') : \mathbb{R}^n \to \mathbb{R}$.

Proof. Consider a coordinate chart $\rho : \mathbb{R}^n \to \Lambda \subset \mathbb{R}^{2n}$. Then, the $2n \times n$ Jacobian $\partial\rho(0)$ has full rank, so we pick n linearly independent rows and call the corresponding coordinates x', ξ''. We then define the map

$$p(x, \xi) = (x', \xi'').$$

Thus, $p \circ \rho : V \to \mathbb{R}^k \times \mathbb{R}^{n-k}$ has an invertible Jacobian, and so is a local diffeomorphism around 0, which means we can use the local coordinates (x', ξ''), where

$$\Lambda \cap U = \{(x', f, g, \xi''\} \cap U$$

for $f = f(x', \xi''), g = g(x', \xi'')$. Thus, by the previous proposition,

$$\omega|_\Lambda = \langle g, dx' \rangle + \langle \xi'', \partial_{x'} f dx' + \partial_{\xi''} f d\xi'' \rangle = \langle \partial_{x'}\psi, dx' \rangle + \langle \partial_{\xi''}\psi, d\xi'' \rangle$$

for some function ψ. Finally, setting $\phi = \psi - \langle f, \xi'' \rangle$ yields

$$f = -\partial_{\xi''}\phi, g = \partial_{x'}\phi.$$

□

Example 2.6.2. When $k = n$, $\Lambda \cap U = \{(x, \partial\phi(x)\}$, so that $\Omega|_\Lambda = d\phi = \partial\phi dx$.

Example 2.6.3. Given a symplectomorphism κ, consider the **twisted graph**

$$\Lambda_\kappa := \{(x, y, \xi, -\eta) : (x, \xi) = \kappa(y, \eta)\}.$$

This can be easily checked to be a Lagrangian submanifold with symplectic form $\omega = d\eta \wedge dy + d\xi \wedge dx$. Then, if $(x, y, \xi, \eta) \to (y, x)$ has a full rank Jacobian, the theorem states that the Lagrangian submanifold may be locally described in (x, y) as before.

2.7 Review of Schwartz Functions

Definition 2.7.1. The **Schwartz space** $\mathcal{S}(\mathbb{R}^n)$ is the vector subspace

$$\{f \in C^\infty(\mathbb{R}^n) : \|f\|_k = \sup_{|\alpha|+|\beta| \leqslant k} |D^\alpha x^\beta f| < \infty \ \forall k \geqslant 1\},$$

i.e. the space of all smooth functions decaying faster than any polynomial. The countable collection of seminorms $\|\cdot\|_k$ define a locally convex topology on §, which is in fact metrizable, with

$$d(f, g) = \sum_n 2^{-n} \frac{\|f\|_k}{1 + \|f\|_k}.$$

In fact, $\mathcal{S}$ is complete under this metric, making $\mathcal{S}$ a **Frechet space.** In particular, $f_n \to f$ iff $\|f_n\|_k \to \|f\|_k$ for all k.

Definition 2.7.2. The space $\mathcal{S}'(\mathbb{R}^n)$ of **tempered distributions** is the dual space of $\mathcal{S}(\mathbb{R}^n)$, that is, the space of continuous linear functionals on the Schwartz space. One typically considers the weak-* topology on $\mathcal{S}'$, with $\phi_n \to \phi$ in $\mathcal{S}'$ iff $\phi_n(f) \to \phi(f)$ for all $f \in \mathcal{S}$.

Remark 2.7.1. Note that $\mathcal{S} \hookrightarrow \mathcal{S}'$ by $u(v) = \int_{\mathbb{R}^n} uv dx$

We now state some nice of properties of the Schwartz space.

Proposition 2.7.1. *(a) $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is an isomorphism, and if one defines the Fourier transform of a distribution by $\langle \widehat{f}, \phi \rangle = \langle f, \widehat{\phi} \rangle$, it extends to an isomorphism $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$.*

One main technique used often in semiclassical analysis is that of integral operators.

Theorem 2.7.1 (Schwartz Kernel Theorem). *If $k \in \mathcal{S}'(\mathbb{R}^{2n})$, one may define an integral operator on $K : \mathcal{S} \to \mathcal{S}'$ by*

$$\langle u, Kv \rangle := \langle u \otimes v, k \rangle, \tag{1}$$

where $u \otimes v := u(x)v(y)$. That is, Kv acts on u the same way that k acts on $u \otimes v$. Informally, one then writes

$$Kv = \int k(\cdot, y)v(y)dy,$$

so that

$$\langle u, Kv \rangle = \iint k(x, y)v(y)u(x)dydx.$$

*One calls $k \in \mathcal{S}'(\mathbb{R}^{2n})$ the **kernel** of the map K. Conversely, every map $K : \mathcal{S} \to \mathcal{S}'$ can be characterized by (1), i.e. there exists a unique distribution $k \in S'(\mathbb{R}^{2n})$ satisfying (1).*

2.8 Notation

Throughout, we use the following notation:

$$\partial_x := (\partial_{x_1}, ..., \partial_{x_n}), \quad D_{x_j} := \frac{1}{i}\partial_{x_j}, D = \frac{1}{i}\nabla.$$

We use the **scaled Fourier transform**

$$\mathcal{F}_h(u)(\xi) := \int_{\mathbb{R}^n} e^{-\frac{i}{h}\langle y,\xi \rangle} u(x)dx, \quad \mathcal{F}_h^{-1}(v)(x) := \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n} e^{\frac{i}{h}\langle y,\xi \rangle} v(\xi)d\xi.$$

We use M_j to denote the multiplication operator by x_j.

2.9 Sobolev Spaces

We are interested in defining spaces with derivatives of functions being in some suitable L^p space.

Definition 2.9.1. For $k \in \mathbb{R}, 1 < p < \infty$. Define the **Sobolev space**

$$W^{k,p} = \{u \in L^p : \mathcal{F}^{-1} \langle \xi \rangle^k \mathcal{F}u \in L^p\}$$

with $\|u\|_{W^{k,p}} = \|\mathcal{F}^{-1} \langle \xi \rangle^k \mathcal{F}u\|_p$. $W^{k,p}$ is a Banach space, and the special case

$$H^k := W^{k,2}$$

is a Hilbert space.

Remark 2.9.1. For natural k, an equivalent norm on $W^{k,p}$ is $\|u\|_{W^{k,p}} = \left(\sum_{|\alpha| \leqslant k} \|D^\alpha u\|_p^p\right)^{\frac{1}{p}}$. Thus, intuitively, one should consider the space $W^{k,p}$ as the space of L^p functions with k derivatives in L^p.

One of the main tools of Sobolev spaces is that they come with certain nice embeddings into more familiar spaces. We will now build up to the most general Sobolev embedding theorems.

For simplicity, suppose $u \in W^{1,p}(\mathbb{R}^n)$. We first consider the case

$$1 \leqslant p < n.$$

Can we establish an estimate of the form

$$\|u\|_{p^*} \lesssim \|\nabla u\|_p?$$

From scaling invariance, we obtain that p, p^* must satisfy the relation

$$1 - \frac{n}{p} + \frac{n}{p^*} = 0 \Longleftrightarrow p^* := \frac{np}{n-p},$$

which is known as the **Sobolev conjugate** of p and satisfies $p^* > p$. Then, in fact, we have:

Theorem 2.9.1 (Gagliardo-Nirenberg-Sobolev Inequality)**.** *For $1 \leqslant p < n$, we have*

$$\|u\|_{p^*} \lesssim \|\nabla u\|_p,$$

i.e. $W^{1,p} \hookrightarrow L^{p^*}$ is continuous.

Proof. Note that

$$|u| = \left| \int u_{y_i} dy_i \right| \leqslant \int |\nabla u| dy_i.$$

Thus, by generalized Hölder,

$$\int |u|^{\frac{n}{n-1}} dx_1 \leqslant \int \prod_i \left(\int |\nabla u| dy_i \right)^{\frac{1}{n-1}} dx_1 \leqslant \left(\int |\nabla u| dy_1 \right)^{\frac{1}{n-1}} \left(\prod_{i=2}^n \iint |\nabla u| dx_1 dy_i \right)^{\frac{1}{n-1}}.$$

Next, we pull out the y_2 factor, integrate with respect to x_2, and use generalized Hölder again to get

$$\int |u|^{\frac{n}{n-1}} dx_1 dx_2 \leqslant \iint |\nabla u| dx_1 dy_2 \prod_{i \geqslant 3} \left(\iiint |\nabla u| dx_1 dx_2 dy_i \right)^{\frac{1}{n-1}}.$$

Continuing in this fashion, we get

$$\int |u|^{\frac{n}{n-1}} dx \leqslant \left(\int |Du| dx \right)^{\frac{1}{n-1}},$$

which yields the estimate for $p = 1$. For $p > 1$, apply the estimate to $v = |u|^\gamma$ and use Hölder to get that

$$\frac{\gamma n}{n-1} = \frac{np}{n-p} = p^*.$$

This yields the general Gagliardo-Nirenberg inequality. □

Theorem 2.9.2 (Sobolev Embedding Theorem)**.** *Let $u \in W^{k,p}(\mathbb{R}^n)$ for $k \in \mathbb{N}, 1 \leqslant p < \infty$. Then, if $n, q > p$ and $l > k$, satisfy*

$$\frac{1}{p} - \frac{k}{n} = \frac{1}{q} - \frac{l}{n},$$

then one has a continuous embedding

$$W^{k,p}(\mathbb{R}^n) \subseteq W^{l,q}(\mathbb{R}^n).$$

If the equality is replaced by $<$, *and* $\mathbb{R}^n$ *is replaced by a bounded open set* U, *then the embedding is compact. Moreover, if* $pk > n, r \in \mathbb{N}$, *and*

$$r = \lfloor \frac{pk-n}{p} \rfloor, \alpha = \left\{ \frac{pk-n}{p} \right\},$$

then one has a continuous embedding

$$W^{k,p}(\mathbb{R}^n) \subset C^{r,\alpha}(\mathbb{R}^n).$$

Proof. For the Hölder space inclusion, we rely on **Morrey's inequality,** which states that □

Corollary 2.9.1. *If* $pk > n$, *then* $W^{k,p}(\mathbb{R}^n) \subset C(\mathbb{R}^n)$, *and thus consists of continuous functions.*

3 Symbol Calculus

We now introduce the main technique of semiclassical analysis - pseudodifferential operators.

Definition 3.0.1. For a function $a = a(x, \xi) \in \mathcal{S}(\mathbb{R}^{2n})$, a **pseudodifferential operator** is an operator of the form

$$\mathrm{Op}_t(a)u(x) = \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y,\xi \rangle} a(tx + (1-t)y, \xi) u(y) dy d\xi.$$

The function a is called the **symbol** of this operator. $\mathrm{Op}_{\frac{1}{2}}(a) = a^w$ is called the **Weyl (quantum) quantization** of a, and $\mathrm{Op}_1(a) = a$ is called **standard (classical) quantization** of a.

Remark 3.0.1. If $a = a(\xi)$, Fubini implies that $\mathrm{Op}_1(a) = \mathcal{F}_h^{-1} a \mathcal{F}_h u$, the Fourier multiplier of a, where $\mathcal{F}_h$ is the scaled Fourier transform. In fact, rescaling all variables by $x \to h^{-\frac{1}{2}} x$ gives us that $\mathrm{Op}_1(a_h)$ is precisely the Fourier multiplier of a.

Proposition 3.0.1. *For* $a \in \mathcal{S}$, $Op_t(a) : \mathcal{S}' \to \mathcal{S}$ *is well-defined and continuous.*

Proof. Notice that $\mathrm{Op}_t(a)$ is an integral operator with kernel

$$\begin{aligned} K_t(x, y) &= \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y,\xi \rangle} a(tx + (1-t)y, \xi) d\xi \\ &= \mathcal{F}_h^{-1}(a(tx + (1-t)y, \cdot))(x - y). \end{aligned}$$

Sine $a \in \mathcal{S}(\mathbb{R}^{2n})$, $K_t \in \mathcal{S}(\mathbb{R}^{2n})$, so $\mathrm{Op}_t(a)$ maps from $\mathcal{S}'$ according to

$$\mathrm{Op}_t(a)u(x) = \langle K_t(x, \cdot), u \rangle.$$

One may directly show that $\mathrm{Op}_t(a)u \in \mathcal{S}$ by showing it is smooth and rapidly decreasing. Finally, to show the map is continuous, suppose $u_j \to 0$ in $\mathcal{S}'$. Then, by definition, $\mathrm{Op}_t(a)u_j$ go to zero pointwise. Moreover, $|\mathrm{Op}_t(a)u_j(x)| = |\langle K_t(x, \cdot), u \rangle| \leqslant C \sum_{k \leqslant K} \|K_t(x, \cdot)\|_k$, so by Banach-Steinhaus, $\mathrm{Op}_t(a)u_j$ is bounded in $\mathcal{S}$. Since $\mathcal{S}$ is has the Heine-Borel property (i.e. every closed and bounded subset is compact), this implies that the sequence converges to 0 in $\mathcal{S}$, implying continuity. □

Proposition 3.0.2. *The formal adjoint of $Op_t(a)$ is $Op_t(a)^* = Op_{1-t}(\overline{a})$.*

Proof. Since $\mathrm{Op}_t(a)$ is an integral operator with kernel $K_t(x, y)$, the adjoint of the kernel can be easily computed as $K_t(x, y)^* = \overline{K_t(y, x)} = \overline{K_{1-t}(x, y)}$. □

Remark 3.0.2. The formal adjoint is different from a regular adjoint in the sense that it ignores boundary conditions and regularity - in other words, it is an adjoint in some weaker sense.

Proposition 3.0.3. *If $a \in \mathcal{S}'$, then $Op_t(a) : \mathcal{S} \to \mathcal{S}'$ is well-defined and continuous.*

Proof. Notice that the kernel K_t is now a distribution, so the action is well-defined in the sense of the Schwartz kernel theorem by

$$\langle u, \mathrm{Op}_t(a)(v) \rangle = K_t(u \otimes v).$$

□

3.1 Some Examples

Example 3.1.1. If $a = p(\xi)$ is a polynomial, then $\mathrm{Op}_t(a) = p(hD)$.

Example 3.1.2. If $a = p_x(\xi)$ is a polynomial in ξ with coefficients as functions in x, then $\mathrm{Op}_1(a) = p_x(hD)$.

Example 3.1.3. If $a(x, \xi) = \langle x, \xi \rangle$, then

$$\mathrm{Op}_1(a)u = \sum_{j=1}^{n} \mathcal{F}_h^{-1} \xi_j \mathcal{F}_h(y_j u) = hD(xu),$$

so $a(x, \xi) = (1-t)\langle hD, x \rangle + t\langle x, hD \rangle$.

Example 3.1.4. If $a = a(x) \in C^\infty$, then $\mathrm{Op}_t(a) = \mathcal{F}_h^{-1} a \mathcal{F}_h$ is the scaled Fourier multiplier of a, $0 \leqslant t \leqslant 1$.

Proof. Taking the derivative with respect to t yields

$$\begin{aligned} \partial_t \mathrm{Op}_t(a)u(x) &= \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n} \left\langle \int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y, \xi \rangle} \nabla a(tx + (1-t)y)u(y)dy, x - y \right\rangle d\xi \\ &= \frac{h}{i(2\pi h)^n} \int_{\mathbb{R}^n} \nabla_\xi \cdot \left(\int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y, \xi \rangle} \nabla a(tx + (1-t)y)u(y)dy \right) d\xi \\ &= \frac{h}{i(2\pi h)^n} \int_{\mathbb{R}^n} \mathrm{div}_\xi (e^{\frac{i}{h}\langle x, \xi \rangle} \widehat{\alpha}(\xi)) d\xi = 0 \end{aligned}$$

where $\alpha(y) := \nabla a(tx+(1-t)y)u(y)$, and we use the divergence theorem along with the fact $\widehat{\alpha}(\xi) \to 0$ as $|\xi| \to \infty$. □

The above example directly implies the following.

Example 3.1.5. If l is a linear symbol of the form $l(x, \xi) = \langle x^*, x \rangle + \langle \xi^*, \xi \rangle$, then $\mathrm{Op}_t(a)(u) = \langle x^*, x \rangle + \langle \xi^*, hD \rangle$.

Example 3.1.6. If l is a symbol linear in ξ of the form $l(x,\xi) = \langle c, \xi \rangle$, $c = c(x) : \mathbb{R}^n \to \mathbb{R}^n$, then

$$\mathrm{Op}_{\frac{1}{2}}(l) = \frac{1}{2}(\langle hD, c \rangle + \langle c, hD \rangle).$$

(Compare this with Example 2.3.)

Proof. We directly compute

$$\begin{aligned}
\mathrm{Op}_{\frac{1}{2}}(l)u(x) &= \sum_{j=1}^{n} \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} c_j\left(\frac{x+y}{2}\right)\xi_j e^{\frac{i}{h}\langle x-y,\xi\rangle} u(y)dyd\xi \\
&= \sum_{j=1}^{n} \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} c_j\left(\frac{x+y}{2}\right) hD_{x_j}(e^{\frac{i}{h}\langle x-y,\xi\rangle})u(y)dyd\xi \\
&= \sum_{j=1}^{n} \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \left(\frac{h}{2i}\partial_{x_j} c_j\left(\frac{x+y}{2}\right)\right)\xi_j e^{\frac{i}{h}\langle x-y,\xi\rangle}u(y) + c_j\left(\frac{x+y}{2}\right)\xi_j e^{\frac{i}{h}\langle x-y,\xi\rangle} hD_{x_j}u(y)dyd\xi \\
&= \frac{h}{2i}(\nabla c)^w u + hc^w Du = \frac{h}{2i}\nabla c(u) + c(hDu) \\
&= \frac{1}{2}(\langle hD, c\rangle + \langle c, hD\rangle),
\end{aligned}$$

using the fact that $a^w = a$ for $a = a(x)$. □

Remark 3.1.1. This proof can be easily extended to show that

$$\mathrm{Op}_t(l) = t\langle hD, c\rangle + (1-t)\langle c, hD\rangle.$$

Example 3.1.7 (Commutators)**.** $(D_{x_j}a)^w = [D_{x_j}, a^w]$ and $h(D_{\xi_j}a)^w = -[x_j, a^w]$.

Proof. We compute

$$\begin{aligned}
(D_{x_j}a)^w(u) &= \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} D_{x_j}a\left(\frac{x+y}{2},\xi\right)e^{\frac{i}{h}\langle x-y,\xi\rangle}u(y)dyd\xi \\
&= \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left(D_{x_j} + \frac{i}{h}\xi_j + D_{y_j} - \frac{i}{h}\xi_j\right)\left(a\left(\frac{x+y}{2},\xi\right)\right)e^{\frac{i}{h}\langle x-y,\xi\rangle}u(y)dyd\xi \\
&= \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(D_{x_j} + D_{y_j})\left(a\left(\frac{x+y}{2},\xi\right)e^{\frac{i}{h}\langle x-y,\xi\rangle}\right)u(y)dyd\xi \\
&= \frac{1}{(2\pi h)^n}D_{x_j}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} a\left(\frac{x+y}{2},\xi\right)e^{\frac{i}{h}\langle x-y,\xi\rangle}u(y)dyd\xi + \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} a\left(\frac{x+y}{2},\xi\right)e^{\frac{i}{h}\langle x-y,\xi\rangle}D_{y_j}u(y \\
&= (D_{x_j}a^w - a^w D_{x_j})(u) = [D_{x_j}, a^w](u).
\end{aligned}$$

and

$$\begin{aligned}
h(D_{\xi_j}a)^w &= \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} hD_{\xi_j}a\left(\frac{x+y}{2},\xi\right)e^{\frac{i}{h}\langle x-y,\xi\rangle}u(y)dyd\xi \\
&= \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(-hD_{\xi_j}+x_j)(e^{\frac{i}{h}\langle x-y,\xi\rangle})a\left(\frac{x+y}{2},\xi\right)u(y)dyd\xi \\
&+ \frac{1}{(2\pi h)^n}(-x_j)\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y,\xi\rangle}a\left(\frac{x+y}{2},\xi\right)u(y)dyd\xi \\
&= \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} a\left(\frac{x+y}{2},\xi\right)(e^{\frac{i}{h}\langle x-y,\xi\rangle})y_ju(y)dyd\xi \\
&+ \frac{1}{(2\pi h)^n}(-x_j)\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y,\xi\rangle}a\left(\frac{x+y}{2},\xi\right)u(y)dyd\xi \\
&= (a^wx_j - x_ja^w)(u) = -[\xi_j, a^w](u).
\end{aligned}$$

□

Example 3.1.8 (Exponentials of Linear Symbols)**.** If l is a linear symbol,

$$(e^{\frac{i}{h}l})^wu(x) = e^{\frac{i}{h}l(x,hD)}u(x) := e^{\frac{i}{h}\langle x^*,x\rangle+\frac{i}{2h}\langle x^*,\xi^*\rangle}u(x+\xi^*).$$

Moreover,

$$e^{\frac{i}{h}l}e^{\frac{i}{h}m} = e^{\frac{i}{h}\omega(l,m)}e^{\frac{i}{h}(l+m)},$$

where ω is the standard symplectic form.

Proof. By definition of the exponential, $e^{\frac{it}{h}l}u(x)$ is the operator that gives the unique solution of the PDE

$$ih\partial_tv + l(v) = 0, v_0 = u \in \mathcal{S}\,.$$

But, it can be directly shown that

$$v(x,t) = e^{\frac{it}{h}\langle x^*,x\rangle+\frac{it^2}{2h}\langle x^*,\xi^*\rangle}u(x+t\xi^*)$$

solves this PDE. Furthermore, sending $y \to y+\xi^*$,

$$\begin{aligned}
(e^{\frac{i}{h}l})^wu(x) &= \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y,\xi\rangle}e^{\frac{i}{h}(\langle\xi^*,\xi\rangle+\langle x^*,\frac{x+y}{2}\rangle)}u(y)dyd\xi \\
&= \frac{e^{\frac{i}{2h}\langle x^*,x\rangle}}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y,\xi\rangle}\left(e^{\frac{i}{2h}\langle x^*,y+\xi^*\rangle}u(y+\xi^*)\right)dyd\xi \\
&= e^{\frac{i}{h}\langle x^*,x\rangle+\frac{i}{2h}\langle x^*,\xi^*\rangle}u(x+\xi^*),
\end{aligned}$$

where we use the simplification

$$\delta_{y=x} = \int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y,\xi\rangle}d\xi \in \mathcal{S}'\,.$$

Finally, to show the commutation property, one has

$$e^{\frac{i}{h}l}e^{\frac{i}{h}m}u(x) = e^{\frac{i}{h}\langle x_1^*,x\rangle+\frac{i}{2h}\langle x_1^*,\xi_1^*\rangle}e^{\frac{i}{h}\langle x_2^*,x+\xi_1^*\rangle+\frac{i}{2h}\langle x_2^*,\xi_2^*\rangle}u(x+\xi_1^*+\xi_2^*)$$

and

$$e^{\frac{i}{h}(l+m)}u(x) = e^{\frac{i}{h}\langle x_1^*+x_2^*,x\rangle + \frac{i}{2h}\langle x_1^*+x_2^*,\xi_1^*+\xi_2^*\rangle}u(x+\xi_1^*+\xi_2^*),$$

with the difference being

$$e^{\frac{i}{2h}(\langle x_2^*,\xi_1^*\rangle - \langle x_1^*,\xi_2^*\rangle)} = e^{\frac{i}{h}\omega(l,m)}.$$

□

Proposition 3.1.1. *If Q is a real invertible symmetric matrix, then*

$$\mathcal{F}_h(e^{\frac{i}{2h}\langle Qx,x\rangle}) = \frac{(2\pi h)^{\frac{n}{2}} e^{\frac{i\pi}{4}\,sign\;Q}}{|\det Q|^{\frac{1}{2}}} e^{-\frac{i}{2h}\langle Q^{-1}\xi,\xi\rangle}.$$

Example 3.1.9. If Q is an invertible symmetric matrix, then

$$e^{\frac{ih}{2}\langle QD,D\rangle}u(x) = \frac{|\det Q|^{-\frac{1}{2}}}{(2\pi h)^{\frac{n}{2}}} e^{\frac{i\pi}{4}\text{sign } Q} \int_{\mathbb{R}^n} e^{-\frac{i}{2h}\langle Q^{-1}y,y\rangle}u(x+y)dy$$

for $u \in \mathcal{S}$.

Proof. Using the formula for the Fourier transform of a quadratic exponential, we remark that

$$\begin{aligned} e^{\frac{i}{2h}\langle QD,D\rangle}u(x) &= \mathcal{F}_h^{-1} e^{\frac{i}{2h}\langle Q\xi,\xi\rangle}\mathcal{F}_h u(x) \\ &= \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y,\xi\rangle} e^{\frac{i}{2h}\langle Q\xi,\xi\rangle}u(y)dyd\xi \\ &= \frac{1}{(2\pi h)^{\frac{n}{2}}}\frac{e^{\frac{i\pi}{4}\text{sign } Q}}{|\det Q|^{\frac{1}{2}}}\int_{\mathbb{R}^n} e^{-\frac{i}{2h}\langle Q^{-1}(x-y),x-y\rangle}u(y)dy \\ &= \frac{|\det Q|^{-\frac{1}{2}}}{(2\pi h)^{\frac{n}{2}}} e^{\frac{i\pi}{4}\text{sign } Q}\int_{\mathbb{R}^n} e^{-\frac{i}{2h}\langle Q^{-1}y,y\rangle}u(x+y).dy \end{aligned}$$

□

Corollary 3.1.1. *For $u = u(x,y) \in \mathcal{S}(\mathbb{R}^{2n})$, setting $Q = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$ yields*

$$e^{ih\langle D_x,D_y\rangle}u(x,y) = \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^{2n}} e^{-\frac{i}{h}\langle x_1,y_1\rangle}u(x+x_1,y+y_1)dx_1dy_1,$$

and for $u = u(z,w) \in \mathcal{S}(\mathbb{R}^{4n})$, setting $Q = \begin{bmatrix} 0 & J \\ -J & 0 \end{bmatrix}$ yields

$$e^{ih\omega(D_z,D_w)}u(z,w) = \frac{1}{(2\pi h)^{2n}}\int_{\mathbb{R}^{4n}} e^{-\frac{i}{h}\omega(z_1,w_1)}u(z+z_1,w+w_1)dz_1dw_1.$$

Example 3.1.10 (Conjugation by $\mathcal{F}_h$). $\mathcal{F}_h^{-1}a^w\mathcal{F}_h = \omega^*a^w$.

Proof. The Schwartz kernel of $\mathcal{F}_h^{-1}a^w\mathcal{F}_h$ is

$$K_{\mathcal{F}_h^{-1}a^w\mathcal{F}_h} = \frac{1}{(2\pi h)^{2n}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{\frac{i}{h}(\langle x',x\rangle + \langle x'-y',\zeta\rangle - \langle y',y\rangle)} a\left(\frac{x'+y'}{2},\zeta\right) dx'dy'd\zeta.$$

Changing variables by setting $z = \frac{x'+y'}{2}$ gives

$$K_{\mathcal{F}_h^{-1} a^w \mathcal{F}_h} = \frac{1}{(2\pi h)^{2n}} \frac{1}{2^n} \iiint e^{\frac{i}{h}\Phi} a(z,\zeta) dx' dz d\zeta$$

for $\Phi = 2(\langle x', \zeta + \frac{x+y}{2} \rangle - \langle z, y+\zeta \rangle)$. Now, note that

$$\frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n} e^{\frac{2i}{h}\langle x', \zeta + \frac{x+y}{2} \rangle} dx' = 2^n \delta(\zeta + \frac{x+y}{2}),$$

so

$$K_{\mathcal{F}_h^{-1} a^w \mathcal{F}_h} = \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y, z \rangle} a\left(z, -\frac{x+y}{2}\right) dz = K_{\omega^* a^w}.$$

□

Lemma 3.1.1. *For a linear symbol $l = \langle x^*, x \rangle + \langle \xi^*, \xi \rangle$ (which we sometimes identify as $l = (x^*, \xi^*) \in \mathbb{R}^{2n}$) and an arbitrary symbol $a \in \mathcal{S}$, the* ***Fourier transform*** *is given by*

$$\widehat{a}(l) := \int_{\mathbb{R}^{2n}} e^{-\frac{i}{h} l(x,\xi)} a(x,\xi) dx d\xi,$$

so that

$$a^w = \frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{2n}} \widehat{a}(l) (e^{\frac{i}{h} l})^w dl.$$

Moreover, if $a \in \mathcal{S}'$, this formula holds in the sense of distributions.

Proof. The Fourier inversion formula implies that

$$a(x,\xi) = \frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{2n}} e^{\frac{i}{h} l(x,\xi)} \widehat{a}(l) dl$$

(read $dl = dx^* d\xi^*$), so

$$\begin{aligned}
a^w u(x) &= \frac{1}{(2\pi h)^{3n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} e^{\frac{i}{h}\langle x-y, \xi \rangle + \frac{i}{h}\langle x^*, \frac{x+y}{2} \rangle + \frac{i}{h}\langle \xi^*, \xi \rangle} \widehat{a}(l) u(y) dx d\xi dl \\
&= \frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{2n}} \delta(x - y + \xi^*) e^{\frac{i}{h}\langle x^*, \frac{x+y}{2} \rangle} \widehat{a}(l) u(y) dx dl \\
&= \frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{2n}} \widehat{a}(l) e^{\frac{i}{h}\langle x^*, x \rangle + \frac{i}{2h}\langle x^*, \xi^* \rangle} u(x + \xi^*) dl \\
&= \frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{2n}} \widehat{a}(l) (e^{\frac{i}{h} l})^w u(x) dl.
\end{aligned}$$

The verification for distributions is left as an exercise. □

Example 3.1.11 (Composition of Symbols)**.** If $z = (x, \xi)$, the composition of symbols is given by

$$a^w b^w = (a \# b)^w,$$

where the symbol is given by the formula

$$a \# b(z) := e^{\frac{ih}{2}\omega(D_z, D_w)} (a(z) b(z)),$$

i.e.

$$a \# b(z) = \frac{1}{(\pi h)^{2n}} \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} e^{-\frac{2i}{h}\omega(w_1, w_2)} a(z + w_1) b(z + w_2) dw_1 dw_2.$$

Proof. We compute the Fourier transform of the right hand side, by writing out the Fourier expansions of a, b, which yields

$$\begin{aligned}
\widehat{a\#b}(l) &= \int_{\mathbb{R}^{2n}} e^{-\frac{i}{h}l(x,\xi)}(a\#b)(x,\xi)dxd\xi \\
&= \frac{1}{(2\pi h)^{4n}} \int_{\mathbb{R}^{2n}} e^{-\frac{i}{h}l} \int_{\mathbb{R}^{4n}} \widehat{a}(m)\widehat{b}(r)e^{\frac{ih}{2}\omega(D_m,D_r)}e^{\frac{i}{h}(m+r)}dmdrdxd\xi \\
&= \frac{1}{(2\pi h)^{4n}} \int_{\mathbb{R}^{6n}} e^{\frac{i}{h}(m+r-l)(x,\xi)}e^{\frac{ih}{2}\omega}\widehat{a}(m)\widehat{b}(r)dmdrdxd\xi \\
&= \frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{4n}} \delta(r+m-l)\widehat{a}(m)\widehat{b}(r)e^{\frac{ih}{2}\omega}dmdr \\
&= \frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{2n}} \widehat{a}(m)\widehat{b}(l-m)e^{\frac{ih}{2}\omega(m,l-m)}dm,
\end{aligned}$$

so the Fourier inversion formula and the fact that $\omega(m,-m)=0$ imply that

$$(a\#b)^w u(x) = \frac{1}{(2\pi h)^{4n}} \int_{\mathbb{R}^{4n}} \widehat{a}(m)\widehat{b}(l-m)e^{\frac{ih}{2}\omega(m,l)}(e^{\frac{i}{h}l})^w u(x)dmdl.$$

On the other hand, by commutation properties for exponentials of linear symbols and making the substitution $l = r + m$,

$$\begin{aligned}
a^w b^w u(x) &= \frac{1}{(2\pi h)^{4n}} \int_{\mathbb{R}^{4n}} \widehat{a}(m)\widehat{b}(r)(e^{\frac{i}{h}m})^w(e^{\frac{i}{h}r})^w u(x)drdm \\
&= \frac{1}{(2\pi h)^{4n}} \int_{\mathbb{R}^{4n}} \widehat{a}(m)\widehat{b}(l-m)e^{\frac{ih}{2}\omega(m,l)}(e^{\frac{i}{h}l})^w u(x)dmdl.
\end{aligned}$$

The formula for $a\#b$ follows immediately from Corollary 2.0.1. □

3.2 Asymptotic Expansions

Definition 3.2.1. We write $A(D) := \frac{1}{2}\omega(D_x, D_\xi, D_y, D_\eta) = D_\xi D_x - D_y D_\eta$.

Definition 3.2.2. Define the **Poisson bracket** of $f, g \in C^\infty(\mathbb{R}^{2n})$ to be

$$\{f,g\} = \langle \partial_\xi f, \partial_x g \rangle - \langle \partial_x f, \partial_\xi g \rangle.$$

Lemma 3.2.1 (Stationary Phase)**.** *For $a \in C_c^\infty(\mathbb{R}^{4n})$,*

$$\int_{\mathbb{R}^{4n}} e^{\frac{i}{h}\omega(z,w)}a(z,w)dzdw = (2\pi h)^{2n}\left(\sum_{k=0}^{N-1}\frac{h^k}{k!}\left(\frac{\omega(D_x,D_\xi,D_y,D_\eta)}{i}\right)^k a(0,0) + O(h^N)\right)$$

Proposition 3.2.1 (Semiclassical Expansion)**.** *For $a, b \in \mathcal{S}$,*

$$a\#b(x,\xi) = \sum_{k=0}^{N}\frac{i^k h^k}{k!}A(D)^k(a(z)b(z)) + O_{\mathcal{S}}(h^{N+1}),$$

where $\phi \in O_{\mathcal{S}}(h^k)$ means that $\|\phi\|_n \ll h^k$ for all n as $h \to 0$.

Corollary 3.2.1. *In particular,*

$$a\#b = ab + \frac{h}{2i}\{a,b\} + O_{\mathcal{S}}(h^2)$$

and

$$[a^w, b^w] = \frac{h}{i}\{a,b\}^w + O_{\mathcal{S}}(h^3),$$

and if a, b *have disjoint supports,* $a\#b = O_{\mathcal{S}}(h^\infty)$.

Proof. We directly expand the integral formula for $a\#b$ using stationary phase. As in the proof of the stationary phase theorem, we write

$$e^{ihA(D)} = \sum_{k=0}^{N} \frac{(ih)^k}{k!} A(D)^k + \frac{(ih)^{N+1}}{N!} \int_0^1 (1-t)^N e^{ithA(D)} A(D)^{n+1} dt,$$

where the last term is a Fourier multiplier, and so maps $\mathcal{S} \to \mathcal{S}$ uniformly in h and t. Next, we directly compute

$$a\#b = ab + \frac{ih}{2}(\langle D_\xi a, D_y b \rangle - \langle D_x a, D_\eta b \rangle) + O_{\mathcal{S}}(h^2) = ab + \frac{h}{2i}\{a,b\} + O_{\mathcal{S}}(h^2).$$

Similarly,

$$[a^w, b^w] = (a\#b - b\#a)^w = (ab + \frac{h}{2i}\{a,b\} + \frac{1}{2}h^2 A(D)^2(ab) + O_{\mathcal{S}}(h^3))^w - (ba + \ldots)^w = \frac{h}{i}\{a,b\}^w + O_{\mathcal{S}}(h^3)^w,$$

since the second derivatives are identical by Clairaut's. Finally, if the supports of a, b are disjoint, every term in the asymptotic expansion vanishes. ☐

Example 3.2.1. If $a = c_j(x), b = \xi_j$,

$$a^w b^w = (a\#b)^w = (ab)^w + \frac{h}{2i}\{a,b\}^w,$$

since $D^\alpha b = 0$ for $|\alpha| \geqslant 2$, i.e.

$$(c_j \xi_j)^w = c_j h D_{x_j} + \frac{h}{2} D_{x_j} c_j,$$

so summing over all indices yields

$$\langle c, hD \rangle^w = \frac{h}{2}\left(\sum_{j=1}^{n} c_j D_{x_j} + D_{x_j} c_j\right) = \frac{1}{2}(\langle c, hD \rangle + \langle hD, c \rangle),$$

which agrees with our previous calculations.

It turns out that all the quantizations are equivalent.

Theorem 3.2.1 (Changing Quantizations)**.** *If for a fixed operator* A, *there exists a family* $a_t, 0 \leqslant t \leqslant 1$ *such that* $A = Op_t(a_t)$ *for all* t, *then*

$$a_t(x,\xi) = e^{i(t-s)h\langle D_x, D_\xi \rangle} a_s(x,\xi).$$

Proof. Generalizing the decomposition formula for the Fourier transform yields

$$\mathrm{Op}_t(a_t) = \frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{2n}} \widehat{a}_t(l) \mathrm{Op}_t(e^{\frac{i}{h}l}) dl.$$

The theorem then follows from the identity

$$\mathrm{Op}_t(e^{\frac{i}{h}l}) = e^{\frac{i}{h}(s-t)\langle x^*, \xi^* \rangle} \mathrm{Op}_s(e^{\frac{i}{h}l}),$$

which may be checked directly, and the fact that

$$\mathcal{F}_h(e^{i(t-s)h\langle D_x, D_\xi \rangle} a_s(x,\xi))(l) = e^{\frac{i}{h}(t-s)\langle x^*, \xi^* \rangle} \mathcal{F}_h a_s(l).$$

□

It turns out that analogous completely analogous formula hold for the standard quantization $t = 1$. We state the results and omit the proofs.

Theorem 3.2.2. *For $a, b \in \mathcal{S}$, $z = (x, \xi)$, the standard symbol of ab is*

$$c := e^{ih\langle D_x, D_\xi \rangle}(a(z)b(z)),$$

with the integral representation

$$c(x,\xi) = \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^{2n}} e^{-\frac{i}{h}\langle x_1, \xi_1 \rangle} a(x, \xi + \xi_1) b(x + x_1, \xi) dx_1 d\xi_1.$$

Moreover,

$$c(x,\xi) = \sum_{k=0}^{N} \frac{h^k}{k!} (i\langle D_x, D_\xi \rangle^k (a(z)b(z)) + O_{\mathcal{S}}(h^{N+1}),$$

and as a consequence of the changing quantizations formula,

$$a^* = e^{ih\langle D_x, D_\xi \rangle} \overline{a}.$$

4 Symbol Classes

We now extend our definitions of pseudodifferential operators to a much larger class of symbols.

Definition 4.0.1. We call $m : \mathbb{R}^{2n} \to (0, \infty)$ an **order function** if $m(z) \leqslant C\langle z - w \rangle^N m(w)$ for all $z, w \in \mathbb{R}^{2n}$ for some N.

Example 4.0.1. $m(z) = 1, \langle z \rangle^a, \langle x \rangle^a \langle \xi \rangle^b$ are order functions.

Proof.

$$1 + |z|^2 \leqslant 1 + |z - w|^2 + |w|^2 + 2|w||z - w| \leqslant 2(1 + |z - w|^2)(1 + |w|^2),$$

from which one has

$$\langle z \rangle^a \leqslant 2^{\frac{a}{2}} \langle z - w \rangle^a \langle w \rangle^a.$$

□

Remark 4.0.1. Order functions form a unital algebra.

Remark 4.0.2. S is the class of smooth functions with all derivatives in L^∞, while S_δ adds the additional restriction that the derivatives decay in h.

Definition 4.0.2. If m is an order function, define the class

$$S_\delta(m) = \{a \in C^\infty : |\partial^\alpha a| \leqslant C_\alpha h^{-\delta|\alpha|} m\}.$$

We write $S(m) = S_0(m), S = S(1), S_\delta = S_\delta(1)$.

Remark 4.0.3. The standard rescaling $x \to h^{-\frac{1}{2}}x$ yields

$$|\partial^\alpha a_h| = h^{\frac{|\alpha|}{2}}|\partial^\alpha a| \leqslant C_\alpha h^{|\alpha|(\frac{1}{2}-\delta)},$$

so henceforth we assume $0 \leqslant \delta \leqslant \frac{1}{2}$, with $\delta = \frac{1}{2}$ being the **critical scaling.**

Recall the classical version of so-called Borel's Theorem.

Theorem 4.0.1 (Borel's Theorem). *If $\{f_n\} \in C^\infty(U), U \subset \mathbb{R}^n$, then on any open interval I containing 0, there exists $F \in C^\infty(\mathbb{R}^n_x \times \mathbb{R}_t)$ s.t. $\partial_t^k F|_{t=0} = f_k$ on U.*

Definition 4.0.3. We say $a, a_j \in S_\delta(m)$ is **asymptotic** to a sum and write

$$a \sim \sum_{j=0}^{\infty} h^j a_j$$

if

$$a - \sum_{j=0}^{N} h^j a_j = O_{S_\delta(m)}(h^N) \iff \left|\partial^\alpha \left(a - \sum_{j=0}^{N} h^j a_j\right)\right| \leqslant C_{\alpha,N} h^{N-\delta|\alpha|} m.$$

If such an expansion holds, a_0 is called the **principal symbol** of a.

Theorem 4.0.2 (Semiclassical Borel's Theorem). *If $a_j \in S_\delta(m)$, there exists a symbol $a \in S_\delta(m)$ such that*

$$a \sim \sum_{j=0}^{\infty} h^j a_j,$$

with any two such symbols a, a' satisfying $a - a' \in O_{S_\delta(m)}(h^\infty)$.

Proof. Define $\chi \in C_0^\infty([0,\infty))$ that equals 1 on $[0,1]$ and 0 on $[2,\infty)$, and for a rapidly growing sequence $\lambda_j \to \infty$, define

$$a := \sum_{j=0}^{\infty} h^j \chi(\lambda_j h) a_j.$$

Then, one can write

$$a - \sum_{j=0}^{N} h^j a_j = \sum_{j=N+1}^{\infty} (h^j \chi(\lambda_j h)) + \sum_{j=0}^{N-1} h_j(1-\chi(\lambda_j h))a_j + h^N a_N := A + B + C.$$

Now, notice that we have the estimate

$$h^j \chi(\lambda_j h)|\partial^\alpha a_j| \leqslant 2h^{j-1-\delta|\alpha|}\frac{m}{\lambda_j} \leqslant 2^{-j} h^{j-1-\delta|\alpha|} m,$$

since $x\lambda(x) \leqslant 2$. Summing then yields that $A \leqslant h^{N-\delta|\alpha|}m$, B vanishes whenever $\lambda_N h < 1$, and whenever $\lambda_N h \geqslant 1$,

$$B \leqslant \sum_{j=0}^{N} C_{\alpha,j} h^{N-\delta|\alpha|} \lambda_N^N m \leqslant C_{\alpha,N} h^{N-\delta|\alpha|},$$

and the bound is obvious for C, implying the desired estimate, from which the second identity immediately follows. □

Finally, we are able to claim that these symbols define pseudodifferential operators.

Proposition 4.0.1. *For $a \in S_\delta(m), a^w : \mathcal{S} \to \mathcal{S}, \mathcal{S}' \to \mathcal{S}'$ is continuous.*

Proof. Define the operators

$$L_1 := 1 + \langle \xi, ih\partial_y \rangle \langle \xi \rangle^{-2}, L_2 := 1 - 1 + \langle x-y, ih\partial_y \rangle \langle x-y \rangle^{-2}.$$

Then, one can easily check that $e^{\frac{i}{h}\langle x-y,\xi \rangle}$ is unchanged under L_1, L_2, and their conjugates are the same up to flipping the sign of the Japanese bracket. So, integrating by parts yields sufficient decay in x, ξ. Additionally,

$$x_j a^w u(x) = \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^{2n}} (D_{\xi_j} + y_j) e^{\frac{i}{h}\langle x-y,\xi \rangle} a u dy d\xi,$$

so integrating by parts again yields the desired decay. □

We now derive some improved error term estimates for symbols in S_δ. We first recall the following result from stationary phase techniques.

Theorem 4.0.3. *For $\phi(w,z) \in \mathbb{R}^k \times \mathbb{R}^m$ real, quadratic in z, w, and s.t.*

$$|\partial\phi| \geqslant \frac{|w + Kz|}{C},$$

and $a \in C^\infty(\mathbb{R}^k \times \mathbb{R}^m)$ s.t.

$$|\partial^\alpha a| \leqslant C_\alpha \langle w \rangle^M \langle z \rangle^M,$$

the oscillatory integral of a with phase ϕ defines a unique tempered distribution u according to

$$u(\psi) = \lim_{\epsilon \to 0} \int_{\mathbb{R}^{k+m}} e^{i\phi(w,z)} a(w,z) \psi(z) \chi(\epsilon w) dw dz$$

for a cutoff $\chi \in C_c^\infty(\mathbb{R}^m)$.

Theorem 4.0.4 (Method of Stationary Phase)**.** *For $a \in C_c^\infty$, $\phi'(x_0) = 0, \phi''(x_0) \neq 0$, and ϕ' nonzero everywhere else on the support of a, if I_h is the oscillatory integral of corresponding to a, ϕ, there exist differential operators A_{2k} of order $\leqslant 2k$ s.t.*

$$\left| I_h - \left(\sum_{k=0}^{N-1} A_{2k} a(x_0) h^{k+\frac{1}{2}} \right) e^{\frac{i}{h}\phi(x_0)} \right| \leqslant C_N h^{N+\frac{1}{2}} \sum_{0 \leqslant m \leqslant 2N+2} \|a^{(m)}\|_\infty.$$

Theorem 4.0.5 (Quadratic Phase Asymptotics)**.** *If Q is symmetric invertible, $a \in C_c^\infty$, and $\phi = \frac{1}{2}\langle Qx, x\rangle$ is a quadratic phase, then the oscillatory integral of ϕ and a is*

$$I_h = (2\pi h)^{\frac{n}{2}} \frac{e^{\frac{i\pi}{4} sign\ Q}}{|\det Q|^{\frac{1}{2}}} \left(\sum_{k=0}^{N-1} \frac{h^k}{k!} \left(\frac{\langle Q^{-1}D, D\rangle}{2i} \right)^k a(0) + O(h^N) \right).$$

In particular, if the term on the right is I_N, we have

$$|I_h - I_N| \leqslant C_N h^N \sup_{0 \leqslant m \leqslant 2N+n+1} \|a^{(m)}\|_\infty$$

Example 4.0.2. If $a \in C_c^\infty(\mathbb{R}^{2n})$,

$$\iint_{\mathbb{R}^{2n}} e^{\frac{i}{h}\langle x,y\rangle} a(x,y)dxdy = (2\pi h)^n \sum_{k=0}^{N-1} \frac{h^k}{k!} \left(\frac{\langle D_x, D_y\rangle}{i} \right)^k a(0,0) + O(h^N),$$

and if $a \in C_c^\infty(\mathbb{R}^{4n})$,

$$\iint_{\mathbb{R}^{4n}} e^{\frac{i}{h}\omega(z,w)} a(z,w)dzdw = (2\pi h)^{2n} \sum_{k=0}^{N-1} \frac{h^k}{k!} \left(\frac{\omega(D_x, D_\xi, D_y, D_\eta)}{i} \right)^k a(0,0) + O(h^N),$$

Theorem 4.0.6 (Exponentials of Quadratics in S_δ)**.** *For Q symmetric invertible, $e^{\frac{ih}{2}\langle QD,D\rangle}$: $S_\delta(m) \to S_\delta(m)$ for $0 \leqslant \delta \leqslant \frac{1}{2}$, and if $\delta < \frac{1}{2}$, one has the asymptotic expansion*

$$e^{\frac{ih}{2}\langle QD,D\rangle} a \sim \sum_{k=0}^{\infty} \frac{h^k}{k!} \left(i\frac{\langle QD, D\rangle}{2} \right)^k a$$

for $a \in S_\delta(m)$.

Remark 4.0.4. We do not have an asymptotic expansion for $\delta = \frac{1}{2}$ since the standard rescaling implies that the terms in the expansion at worst satisfy

$$\left| \frac{h^k}{k!} \left(i\frac{\langle QD, D\rangle}{2} \right)^k \partial^\alpha a \right| \leqslant h^{k(1-2\delta)} h^{-\delta|\alpha|} m = h^{-\frac{|\alpha|}{2}} m,$$

since $\langle QD, D\rangle^k a(h^{-\frac{1}{2}}x, h^{-\frac{1}{2}}\xi)$ adds at worst $4k$ ($2k$ from each variable) copies of $h^{-\frac{1}{2}}$ to the derivative.

Proof. Suppose $\delta < \frac{1}{2}$, and proceed as by the formula in Example 3.9. First, note that the quadratic exponential defines an oscillatory integral with $\phi(w) = -\frac{1}{2}\langle Q^{-1}w, w\rangle$, so since Q is invertible, ϕ satisfies the conditions of Theorem 4.2 and thus the quadratic exponential defines a tempered distribution in $\mathcal{S}'$. We define a cutoff $\chi \in C_0^\infty(\mathbb{R}^{2n})$ that equals 1 on the unit ball and vanishes outside the ball of radius 2. Split the integrand in Example 3.9 into the χ term A and $1 - \chi$ term B. For A, since the integrand is now compactly supported, we can apply Theorem 4.5 to obtain the asymptotic expansion

$$e^{\frac{ih}{2}\langle QD,D\rangle} a(z) \sim \sum_{k=0}^{\infty} \frac{h^k}{k!} \left(i\frac{\langle QD, D\rangle}{2} \right)^k a(z).$$

Moreover, since $|w| \leqslant 2$, $m(z+w) \leqslant C \langle w \rangle^N m(z) \leqslant C' m(z)$, and thus in the notation of Theorem 4.5,

$$|\partial^\alpha A| \leqslant |\partial^\alpha A - I_N \partial^\alpha a| + |I_N \partial^\alpha a| \leqslant C_N h^{\frac{1}{2}+N} \sup_{|\beta| \leqslant 2N+n+1} |\partial^{\alpha+\beta} a| + C'_N h^{\frac{n}{2}} |\partial^\alpha a| \leqslant C_N h^{-\delta|\alpha|} m$$

□

Theorem 4.0.7 (Composition in S_δ)**.** *If $a \in S_\delta(m_1), b \in S_\delta(m_2)$, then $a\#b \in S_\delta(m_1 m_2)$. Furthermore,*

$$a\#b = ab + \frac{i}{2h}\{a, b\} + O_{S_\delta(m_1 m_2)}(h^{1-2\delta})$$

and

$$[a^w, b^w] = \frac{h}{i}\{a, b\}^w + O_{S_\delta(m_1 m_2)}(h^{3(1-2\delta)})$$

5 Symbol Calculus on L^2

We now aim to develop the calculus of symbols on L^2, which becomes extremely convenient when dealing with the formalism of quantum mechanics.

Theorem 5.0.1. *If $a \in \mathcal{S}$, then $a^w : L^2 \to L^2$ is bounded uniformly in h.*

Proof. Recall that the a^w is an integral operator with kernel

$$K(x, y) = \frac{1}{(2\pi h)^n} \int e^{\frac{i}{h}\langle x-y, \xi \rangle} a\left(\frac{x+y}{2}, \xi\right) d\xi = \mathcal{F}_h^{-1}\left(a\left(\frac{x+y}{2}, \cdot\right)\right)(x-y).$$

Consequently, since $a \in \mathcal{S}$,

$$\sup_x \int |K(x,y)| dy \ll \sup_x \iint \left| a\left(\frac{x+y}{2}, \xi\right)\right| dy d\xi := C_1 < \infty$$

and likewise,

$$\sup_y \int |K(x,y)| dx := C_2 < \infty.$$

Thus, for $u \in L^2$,

$$\|a^w u\|_2^2 \leqslant \iiint |K(x,y)||K(x,z)||u(y)||u(z)| dx dy dz.$$

Then, either using Schur's test, or Cauchy-Schwarz and the estimate above, one gets

$$\|a^w u\|_2 \leqslant \sqrt{C_1 C_2} \|u\|_2.$$

□

Remark 5.0.1. The norm is independent of h despite the $\frac{1}{(2\pi h)^n}$ factor out front because arbitrary polynomial decay in $\frac{x-y}{h}$ from the fact that $a \in \mathcal{S}$ compenstates for that factor.

Well, the proof is quite straightforward for a Schwartz symbol, but our ultimate goal is to extend these operators to L^2 for symbols in $S_\delta(m)$. For that, we first need some preliminaries, and set $h = 1$ for simplicity.

Definition 5.0.1. Let χ denote a smooth compactly supported cutoff vanishing outside $B(0,2)$ and equal to 1 on the unit ball, with

$$\sum_{\alpha\in\mathbb{Z}^{2n}} \chi_\alpha \equiv 1$$

for $\chi_\alpha = \chi(\cdot - \alpha)$. We write $a_\alpha = \chi_\alpha a$ and define

$$b_{\alpha\beta} := \overline{a_\alpha}\# a_\beta$$

Proposition 5.0.1 (Decay of Mixed Terms). *For $a \in S$ and any $N > 0$, we have the estimates*

$$|\partial^\gamma b_{\alpha\beta}(z)| \leqslant C_{\gamma,N} \langle \alpha - \beta \rangle^{-N} \left\langle z - \frac{\alpha+\beta}{2} \right\rangle^{-N}$$

and

$$\|b^w_{\alpha\beta}\|_{L^2\to L^2} \leqslant C_N \langle \alpha - \beta \rangle^{-N}.$$

Proof. We select a cutoff $\zeta = \zeta(w) \in C_c^\infty(\mathbb{R}^{4n})$ and estimate the cutoff term A and noncutoff term B of $b_{\alpha\beta}$ for the composition of symbols in Example 3.11. The cutoff term vanishes unless ζ intersects χ_α, χ_β, i.e.

$$|z - w_1 - \alpha|, |z - w_2 - \beta| \leqslant 2,$$

so

$$|\alpha - \beta| \leqslant 4 + |w_1| + |w_2| \leqslant 8, \left| z - \frac{\alpha+\beta}{2} \right| \leqslant \frac{1}{2}(4 + |w_1| + |w_2|) \leqslant 8,$$

so since the integrand of A is bounded and compactly supported and since the reciprocals of these two terms can only decay so fast

$$|A| \leqslant C \leqslant C_N \langle \alpha - \beta \rangle^{-N} \left\langle z - \frac{\alpha+\beta}{2} \right\rangle^{-N}.$$

For the first term, the bound is unnecessarily weak and has no meaning in z, since it is compactly supported. The only important part is that as $\alpha - \beta \to \infty$ the term vanishes at some point. and

$$|\partial^\gamma A| \leqslant C \leqslant C_{N,\gamma} \langle \alpha - \beta \rangle^{-N} \left\langle z - \frac{\alpha+\beta}{2} \right\rangle^{-N},$$

and taking derivatives is fine since derivatives of a are in L^∞. For the B term, we use the classic trick of defining an operator

$$L := \frac{\langle \partial\phi, D \rangle}{|\partial\phi|^2}$$

such that $Le^{i\phi} = e^{i\phi}$ for $\phi(z,w) = -2\omega(z,w)$ and doing integrating by parts to extract as many negative powers of w as we want, which certainly makes $|\partial^\gamma B|$ integrable since all the derivatives of a are in L^∞. Now, note that since e^{il} is a unitary operator on L^2, by Fourier inversion

$$\|a^w\|_{L^2\to L^2} \ll \|\widehat{a}(l)\|_1,$$

so using the above estimate,

$$\|b^w_{\alpha\beta}\|_{L^2\to L^2} \ll \|\widehat{b_{\alpha\beta}}\|_1 \leqslant C \max_{|\gamma|\leqslant 2n+1} \|\partial^\gamma b_{\alpha\beta}\|_1 \leqslant \langle \alpha - \beta \rangle^{-N} \int \left\langle z - \frac{\alpha+\beta}{2} \right\rangle^{-N} dz \leqslant C \langle \alpha - \beta \rangle^{-N},$$

where $2n$ is the dimension of the space and $b_{\alpha\beta}$ is smooth, which guarantees that $\langle \xi \rangle^{2n+1} \widehat{b_{\alpha\beta}} \in L^\infty$, and using the Fourier transform property $\|\widehat{u}\|_{L^1(\mathbb{R}^n)} \leqslant C \sup_{\gamma\leqslant n+1} \|\partial^\gamma u\|_{L^1(\mathbb{R}^n)}$. Note that $a_\alpha a_\beta$ are compactly supported, so they are certainly Schwartz and the estimate makes sense. ∎

Definition 5.0.2. A family of bounded operators T_j between Hilbert spaces is called **almost orthogonal** if

$$A = \sup_j \sum_k \|T_j^* T_k\| < \infty, B = \sup_k \sum_j \|T_k^* T_j\| < \infty.$$

Theorem 5.0.2 (Cotlar-Stein Theorem)**.** *If T_j is an almost orthogonal family of bounded operators, then, $\sum T_j \to T$ converges strongly (not in norm) and*

$$\|T\| \leqslant \sqrt{AB}.$$

Theorem 5.0.3 (Symbols in L^2)**.** *If $a \in S$,*

$$\|a^w\|_{L^2 \to L^2} \leqslant C \sum_{|\alpha| \leqslant Mn} \sup |\partial^\alpha a|$$

for some universal constant M. Furthermore, if $a \in S_\delta$,

$$\|a^w\|_{L^2 \to L^2} \leqslant C \sum_{|\alpha| \leqslant Mn} h^{\frac{|\alpha|}{2}} \sup |\partial^\alpha a|.$$

Proof. We have $b_{\alpha\beta} = A_\alpha^* A_\beta$ for $A_\alpha := a_\alpha^w$, and by the previous proposition

$$\|A_\alpha^* A_\beta\|_{L^2 \to L^2} \leqslant C \langle \alpha - \beta \rangle^{-N},$$

so the proof follows by applying the Cotlar-Stein Theorem to $a^w = \sum_\alpha A_\alpha$. The constants in the estimates in the previous proposition depend only on a finite number of derivative of a which grow linearly with dimension, which justifies the finite sum and the universal constant M. The constant M The second estimate follows from a rescaling. □

Remark 5.0.2. This shows that symbols in S classes define strong type $(2, 2)$ operators. More generally, if $m(x, \xi) = \langle \xi \rangle^s$ we can work with Sobolev spaces $S(m) = \{u \in S' : \langle D \rangle^s u \in L^2\} = H^s$, showing that symbols define bounded operators $H^s \to L^2$.

Corollary 5.0.1. *For $a, b \in S_\delta, \delta < \frac{1}{2}$*

$$(ab)^w = a^w b^w + O_{L^2 \to L^2}(h^{1-2\delta})$$

as $h \to 0$.

Proof. Since $a^w b^w - (ab)^w = (a\#b - ab)^w = O(h^{1-2\delta})$, the previous theorem shows that as an operator, $a^w b^w = (ab)^w = O_{L^2 \to L^2}(h^{1-2\delta})$. □

We have the following result for the edge case of $\delta = \frac{1}{2}$.

Corollary 5.0.2. *If $a, b \in S_{\frac{1}{2}}$ have supports of distance γ away from each other, where γ does not depend on h, then*

$$a^w b^w = O_{L^2 \to L^2}(h^\infty).$$

Proof. Since $a, b \in S_{\frac{1}{2}}$, using the regular integration by parts trick with L m times yields

$$(L^*)^M (a(z - w_1) b(z - w_2)) = O(h^{\frac{M}{2}} \langle w \rangle^{-M}),$$

so by setting M large enough we get the h^∞ bound. □

5.1 Compactness

We now consider some compactness results.

Proposition 5.1.1. *If $a \in \mathcal{S}$, $a^w : L^2 \to L^2$ is compact.*

Proof. a^w is an integral operator with Schwartz kernel, so

$$\sup |x^\alpha \partial^\beta a^w u| \leqslant C_{\alpha\beta} \|u\|_2$$

by Cauchy-Schwarz. Then, since

$$\|a^w f_p - a^w f_k\|_2 \leqslant \| \langle x \rangle^{-n} \| \|g_p - g_k\|_\infty$$

for $g_p = \langle x \rangle^N a^w f_p$, it suffices to show a subsequence of g_k converges uniformly. The conclusion thus follows immediately from Arzela-Ascoli. □

We can now also give an improved estimate on the decay of mixed terms:

Proposition 5.1.2 (Decay of Mixed Terms for General Symbols)**.** *For $a \in S(m)$ and $b_{\alpha\beta}$ as defined above,*

$$\|b^w_{\alpha\beta}\|_{L^2 \to L^2} \leqslant C_N m(\alpha) m(\beta) \langle \alpha - \beta \rangle^{-N}.$$

Proof. Follows completely analogously to the proof for S. □

We actually get an improved estimate on compactness for arbitrary decaying order functions.

Theorem 5.1.1 (Compactness for Decaying Functions)**.** *If m is an order function that decays at ∞, $a \in S(m)$, then $a^w : L^2 \to L^2$ is compact.*

Remark 5.1.1. The same is true for all other quantizations, and the converse holds: if $a^w \in S(m)$ is always compact, then m decays.

Proof. Integral operators with L^2 kernels are compact, and compact operators are closed in norm. Thus, setting $A_M = \sum_{|\alpha| \leqslant M} A_\alpha$ as before, we use Cotlar-Stein to show that

$$\|A - A_M\| \leqslant \max \left(\sup_{|\alpha| \geqslant M} \sum_{|\beta| \geqslant M} \|A^*_\alpha A_\beta\|^{\frac{1}{2}}, \sup_{|\alpha| \geqslant M} \sum_{|\beta| \geqslant M} \|A_\alpha A^*_\beta\|^{\frac{1}{2}} \right).$$

Notice that $\sup_{|\alpha| \geqslant M} \sum_{|\beta| \geqslant M} \|A^*_\alpha A_\beta\|^{\frac{1}{2}} \leqslant \sup_{|\alpha| \geqslant M} m(\alpha)$ by the above estimate for the decay of mixed terms. Since m vanishes at ∞, the claim follows. □

5.2 Inverses

We now show that we are able to invert a certain class of elliptic symbols.

Definition 5.2.1. A symbol $a \in S(m)$ is **elliptic** if $|a| \geqslant \gamma m$ for some $\gamma > 0$.

Theorem 5.2.1 (Inverses for Elliptic Symbols)**.** *If $a \in S_\delta(m)$ is elliptic and $m \geqslant 1$, then $a^w : L^2 \to L^2$ is bounded from below for small enough h. If $m = 1$, then a^w is invertible for small enough h.*

Proof. Note that $b = \frac{1}{a} \in S_\delta(\frac{1}{m})$, so

$$a\#b = 1 + r_1, b\#a = 1 + r_2, r_1, r_2 \in h^{1-2\delta} S_\delta.$$

For $m = 1$, $b^w \in S$ and so is bounded in L^2. Thus a^w, b^w are approximate inverses, so they are invertible. Moreover, if $m \geqslant 1$,

$$\|u\|_2 = \|(I + R_2)^{-1} b^w a^w\| \leqslant C \|a^w\|_2$$

since b^w is bounded on L^2. ☐

Finally, we ask what happens when a is real and nonnegative.

Proposition 5.2.1. *If $a \in S$ is elliptic, then for all $\epsilon > 0$ for small enough $h = h(\epsilon)$, $a^w \geqslant \gamma - \epsilon$.*

Proof. We show that $a - \lambda$ is invertible for $\lambda < \gamma - \epsilon$. Indeed, if $b = (a - \lambda)^{-1}$

$$(a - \lambda)\#b = 1 + \frac{h}{2i}\{a - \lambda, b\} + O(h^2) = 1 + O(h^2),$$

where the bracket vanishes since b is a function of $a - \lambda$. Thus, the two are approximate inverses, so $a^w - \lambda$ is invertible with spectrum in $[\gamma - \epsilon, \infty)$. ☐

Lemma 5.2.1. *If $f \in C^2$ is positive and $|D^2 f| \leqslant A$, then $|\nabla f| \leqslant \sqrt{2Af}$.*

Proof. Use Taylor's expansion and set $y = -\frac{1}{A}\nabla f$. ☐

Using this quick lemma, we may obtain a strengthening of the above inequality:

Theorem 5.2.2 (Sharp Garding Inequality)**.** *If $a \in S$ is nonnegative, then $a^w \geqslant -Ch$ for some C and small enough h.*

Remark 5.2.1. The estimate holds true for arbitrary quantizations, and for the Weyl quantization, one actually has $a^w \geqslant -Ch^2$ for h small enough.

Proof. If we fix $\widetilde{h}$ and write $\lambda := \frac{h}{\widetilde{h}}$, our goal is to show that $h(a + \lambda)^{-1} \in S_{\frac{1}{2}}$ independent of $\widetilde{h}$, since then we can argue as in the previous proof. The lemma implies

$$\lambda^{\frac{1}{2}} |\partial a| \leqslant C(\lambda + a),$$

so $|\partial^\beta a|(a + \lambda)^{-1} \leqslant C\lambda^{-1}$ for $|\beta| \geqslant 2$ (since $a \in S$) and $|\partial^\beta a|(a + \lambda)^{-1} \leqslant C\lambda^{-\frac{1}{2}}$ for $|\beta| = 1$. Then, using the estimate

$$\partial^\alpha (a + \lambda)^{-1} = (a + \lambda)^{-1} \sum_k \sum_\alpha C \prod_j (a + \lambda)^{-1} \partial^{\beta_j} a$$

yields

$$|\partial^\alpha (a + \lambda)^{-1}| \leqslant C_\alpha (a + \lambda)^{-1} \lambda^{-\frac{|\alpha|}{2}}.$$

This yields $h(a + \lambda)^{-1} \in \widetilde{h} S_{\frac{1}{2}}$ independent of λ. Setting $b = (a + \lambda)^{-1}$ as usual and expanding using Taylor's formula yields

$$(a + \lambda)\#b = 1 + \int_0^1 (1 - t) e^{ithA(D)} (ihA(D))^2 (a + \lambda) b := 1 + r(z),$$

where the Poisson bracket vanishes as usual and

$$\|r^w\|_{L^2 \to L^2} \leqslant C\widetilde{h} \leqslant \frac{1}{2}$$

for small enough $\widetilde{h}$ since the complex exponential preserves the class $\widetilde{h} S_{\frac{1}{2}}$ and $h(a + \lambda)^{-1} \in \widetilde{h} S_{\frac{1}{2}}$. Thus, $a^w + \gamma + \lambda$ has an approximate inverse, and since $\lambda = \frac{h}{\widetilde{h}}$, this finishes the proof. ☐

Remark 5.2.2. The standard scaling might be generalized by $\widetilde{x} = \left(\frac{\widetilde{h}}{h}\right)^{\frac{1}{2}} x$, performing a change of variables in the small parameter.

6 Applications

6.1 Quantum Harmonic Oscillator

In this chapter, we study the symbol

$$p(x,\xi) = |\xi|^2 + V(x)$$

for a given potential $V : \mathbb{R}^n \to \mathbb{R}$ and the corresponding operator

$$P(x, hD) = -h^2\Delta + V.$$

In particular, setting $V(x) = x^2$ yields the **one-dimensional quantum harmonic oscillator**

$$P = -\Delta^2 + x^2.$$

Definition 6.1.1. Define the **creation** and **annihilation** operators

$$A_\pm := D_x \pm ix.$$

It is easy to check that A_+, A_- are Hermitian conjugates and $P = A_+A_- + 1 = A_-A_+ - 1$.

Theorem 6.1.1. $P \geqslant 1$, *i.e.* $P - I$ *is positive semi-definite. Moreover* $v_0 := e^{-\frac{x^2}{2}}$ *is the smallest eigenfunction of* P *with eigenvalue* 1, *and* $u_n := \frac{A_+^n v_0}{\|A_+^n v_0\|_2} = H_n(x)e^{-\frac{x^2}{2}}$ *are eigenfuctions of* P *with eigenvalues* $2n+1$ *that form an orthonormal eigenbasis of* $L^2(\mathbb{R}^n)$.

Remark 6.1.1. The polynomials $H_n(x)$ have degree n are called the **Hermite polynomials.**

Proof. Since $i[D_x, x] = 1$,

$$\|u\|_2 = \langle\, i[D_x, x]u, u\,\rangle \leqslant 2\|xu\|_2\|D_xu\|_2 \leqslant \|xu\|_2^2 + \|D_xu\|_2^2 = \langle\, Pu, u\,\rangle.$$

Direct calculation shows that $A_-v_0 = 0$ and that v_n is an eigenfunction with eigevalue $2n+1$. Moreover, since $[A_-, A_+] = 2$,

$$\langle\, u_n, u_m\,\rangle = \langle\, A_+^n v_0, A_-^m v_0\,\rangle = \langle\, A_+^{n-1}(A_+A_- + 2)A_-^{n-1}v_0, v_0\,\rangle,$$

so since $A_-v_0 = 0$, this inductively implies that $\{u_n\}$ are orthonormal. It is obvious that $v_n = H_n(x)e^{-\frac{x^2}{2}}$ for some polynomials H_n of degree n and thus form a basis for polynomials. Thus, if $\langle\, u_n, g\,\rangle = 0$ for all n,

$$\int ge^{-\frac{x^2}{2}}p\,dx = 0$$

for all polynomials p and thus $\mathcal{F}(ge^{-\frac{x^2}{2}}) = 0$, so $g = 0$. □

6.2 Higher Dimensions

In the case of the n-dimensional quantum harmonic oscillator, we define

$$u_\alpha(x) = \prod u_{\alpha_j}(x_j) = \prod H_{\alpha_j}(x_j) e^{-\frac{|x|^2}{2}}.$$

Then,

$$Pu_\alpha = (2|\alpha| + n)u_\alpha,$$

so u_α is the eigenfunction with eigenvalue $2|\alpha| + n$. Additionally, we restore the parameter $h > 0$ to obtain

$$u_{\alpha,h}(x) = h^{-\frac{n}{4}} \prod_{j=1}^{n} H_{\alpha_j}(\frac{x_j}{\sqrt{h}}) e^{-\frac{|x|^2}{2h}}$$

with eigenvalues $E(h) = (2|\alpha| + n)h$.

6.3 Weyl's Law

We now introduce the semiclassical formalism to study the asymptotic distribution of eigenvalues of the oscillator.

Lemma 6.3.1 (Weyl's Law, Harmonic Oscillator)**.** *For the QM harmonic oscillator and* $0 \leqslant a < b < \infty$*, one has*

$$|\{a \leqslant E(h) \leqslant b\}| = \frac{1}{(2\pi h)^n}(|\{a \leqslant |\xi|^2 + |x^2| \leqslant b\}| + o(1)).$$

as $h \to 0$.

Proof. If $a = 0$, the above formula for the eigenvalues gives the number of eigenvalues as $|\{\alpha : |\alpha \leqslant R\|\}|$ for $R = \frac{b-nh}{2h}$. Then, using the fact that the volume of $\sum_{i \leqslant n} |x_i| = \frac{1}{n!}$, we approximate

$$|\{0 \leqslant E(h) \leqslant b\}| = \frac{1}{n!}(\frac{b}{2h} - \frac{n}{2})^n + o(R^n) = \frac{1}{n!}\left(\frac{b}{2h}\right)^n + o(h^{-n})$$

as $h \to 0$. Moreover, $|\{|\xi|^2 + |x|^2 \leqslant b\}| = b^n \alpha(2n)$, where $\alpha(k) = \pi^{\frac{k}{2}} \Gamma(1 + \frac{k}{2})^{-1}$ is the k-dimensional ball volume function. Hence

$$|\{0 \leqslant E(h) \leqslant b\}| = \frac{1}{(2\pi h)^n} |\{|\xi|^2 + |x|^2 \leqslant b\}| + o(h^{-n}).$$

□

Remark 6.3.1. This essentially states that the number of eigenvalues that fall within a certain range is proportional to the volume of $a \leqslant |\xi|^2 + |x|^2 \leqslant b$.

Our goal is now to prove the same estimate for an arbitrary quadratic-like potential.

Definition 6.3.1. We say that $V : \mathbb{R}^n \to \mathbb{R}$ is an **admissible potential** if it is smooth and

$$|\partial^\alpha V(x)| \leqslant C_\alpha \langle x \rangle^k, \quad V(x) \geqslant c \langle x \rangle^k \text{ whenever } |x| \geqslant R,$$

for all indices α and appropriate constants $k, c, C_\alpha, R > 0$. In other words, V is a coercive symbol.

Proposition 6.3.1. *Take $P_h := -h^2\Delta + V$. Then, as $h \to 0$, the eigenfunction u_h of P_h with energy E lives on the energy surface $M_E := \{|\xi|^2 + V(x) = E\}$. More formally, if $a \in S$ is a symbol with support disjoint from M_E, then for energies E_h close enough to E, one has*

$$\|a^w u_h\|_2 = O(h^\infty)\|u_h\|_2.$$

Proof. Note that M_E is compact, so there exists a smooth cutoff χ that is 1 on M_E and vanishes on the support of a. Define the symbol

$$b := |\xi|^2 + V(x) - E_h + i\chi$$

and the order function $m := \langle \xi \rangle^2 + \langle x \rangle^k$. Thus, if E_h is close enough to E, b is large outside of K and equals 1 on K, so $|b| \geqslant \gamma m$, i.e. $b \in S(m)$ and $b^{-1} \in S(m^{-1})$. It follows that for some symbol $c \in S(m^{-1})$,

$$b^w c^w = I + r_1^w, c^w b^w = I + r_2^w, r_1^w, r_2^w \in O(h^\infty).$$

Furthermore,

$$a^w c^w b^w = a^w + O(h^\infty), b^w = P_h - E_h + i\chi^w,$$

and

$$a^w c^w \chi^w = O(h^\infty)$$

since a, χ have disjoint support. Thus, if $P_h u_h = E_h u_h$,

$$a^w u_h = a^w c^w (P_h - E_h + i\chi^w)u_h + O(h^\infty) = O(h^\infty).$$

□

We now develop a sharper estimate for the harmonic oscillator:

Proposition 6.3.2. *If $u_h \in L^2$ is an eigenfunction of the harmonic oscillator and $a \in C_c^\infty$, then there exists E_0 depending only on the support of a s.t. for $E_h > E_0$,*

$$\|a^w u_h\|_2 = O\left(\left(\frac{h}{E_h}\right)^\infty\right)\|u_h\|_2$$

Proof. We rescale

$$y = \frac{x}{\sqrt{E}}, \widetilde{h} = \frac{h}{E}, E_{\widetilde{h}} := E_h/E,$$

where E is chosen so that $|E_h - E| \leqslant \frac{E}{4}$. Then,

$$P_h(x) - E_h = E(P_{\widetilde{h}}(y) - E_{\widetilde{h}}).$$

We also introduce the unitary map

$$Uu(y) = E^{\frac{n}{4}} u(\sqrt{E}y)$$

mapping $P_h(x)$ to $P_h(y)$. If we now apply the previous theorem,

$$(P_{\widetilde{h}} - E_{\widetilde{h}})u_{\widetilde{h}} = 0, |E_{\widetilde{h}} - 1| < \delta,$$

and supp $\widetilde{b} \subset \{|y|^2 + |\eta|^2 \leqslant \frac{1}{2}\}$, so

$$\|\widetilde{b}^w u_{\widetilde{h}}\|_2 = O(\widetilde{h}^\infty)\|u_{\widetilde{h}}\|_2.$$

Now, if $\widetilde{b} = 1$ on

$$\{|y|^2 + |\eta|^2 \leqslant \frac{1}{4}\}, \text{supp } a \subset \{|x|^2 + |\xi|^2 \leqslant R^2\},$$

it remains to show that

$$\|\widetilde{a}^w(1 - \widetilde{b}^w)\|_{L^2 \to L^2} = O(\widetilde{h}^\infty).$$

But the support of $\widetilde{a}$ is in $\frac{R^2}{E}$, so for E large enough the supports of $a, 1-b$ are disjoint, and the theorem follows from the h^∞ estimate for symbols with disjoint supports. □

6.4 Projections

We now study the properties of projections:

Proposition 6.4.1. *If* Π *is the projection onto* $E_h \leqslant R$*, where supp* $a \subset \{|\xi|^2 + |x|^2 \leqslant R\}$*, then*

$$a^w(I - \Pi), (I - \Pi)a^w = O(h^\infty).$$

Proof. We see that

$$\|a^w(I - \Pi)\|_{L^2 \to L^2}^2 \leqslant \left(\sum_{R < E_j \leqslant E_1} + \sum_{E_j > E_1} \right) \|a^w u_j\|_2^2 := A + B,$$

where $E_1 = \max(E_0, R)$. For term A, the previous h^∞ estimates and Weyl's law for the harmonic oscillator (which tells us how many terms are in the sum) imply that

$$A \leqslant |\{R < E_j < E_0\}| \max \|a_w u_j\|_2^2 \leqslant C h^{-N} O(h^\infty) = O(h^\infty),$$

where we cover $[R, E_0]$ with with finitely many δ_E to obtain a uniform estimate. Note that Weyl's law yields

$$E_j(h) \geqslant \gamma j^{\frac{1}{n}} h,$$

so using the previous theorem for $E_j(h) > E_0$ yields that for some $\gamma > 0$,

$$\|a^w u_j(h)\|_2^2 \leqslant C_N h^M \left(\frac{h}{E_j(h)} \right)^{N-M} \leqslant C_N h^M j^{-\frac{N-M}{n}}.$$

Consequently, for $N - M \geqslant 2n$ and M large enough, we get that $B = O(h^\infty)$. The proof $(I - \Pi)a^w$ is similar. □

6.5 Resolvent Properties

We begin with the following theorem:

Theorem 6.5.1. *For* $h \leqslant h_0$ *for some* h_0*,* $(P_h - i)^{-1} : L^2 \to L^2$ *is compact. Additionally,*

$$z \to (P_h - z)^{-1}$$

is meromorphic with real simple poles, and $(P_h - i)^{-1}$ *is self-adjoint, so the spectrum of* $(P_h - i)^{-1}$ *is real and discrete, and there exists an orthonormal basis of* L^2 *consisting of eigenfunctions.*

Remark 6.5.1. In fact, you can show that the eigenfunctions of P_h satisfy $u_j(h) \in \mathcal{S}$.

Proof. If

$$m(x,\xi) = 1 + |\xi|^2 + |x|^k,$$

then $p \in S(m), C|p - i| \geqslant m, and p^w = P_h$. For small enough h, m^w has a right inverse

$$(m^w)^{-1} := \left(\frac{1}{m}\right)^w (I + hr^w)^{-1}$$

for $r = h^{-1}((m\#\frac{1}{m}) - 1) \in S$, so if we define the Hilbert space

$$H := \{u \in S' : (I - h^2\Delta + \langle x \rangle^k)u \in L^2\} = (m^w)^{-1}L^2.$$

Since $m^{-1} \to 0$ as $|x| \to \infty$, it defines a compact operator, and hence the inclusion $H \hookrightarrow L^2$ is compact. Since $(P_h - i)^{-1} : L^2 \to H$ is bounddd for small h, it follows that $(P_h - i)^{-1} : L^2 \to L^2$ is a compact operator. Consequently,

$$(P_h - z)^{-1} = (P_h - i)^{-1}(I - (z - i)(P_h - i)^{-1})^{-1}$$

is a meromorphic family of compact operators. Additionally, we easily check that $(P_h - i)^{-1}$ is a self-adjoint compact operator. □

Finally, the last step we need is a theorem regarding estimates on spectra of self-adjoint operators:

Theorem 6.5.2. *Suppose $A \geqslant -c$ is self-adjoint, and the inverse $(A + 2c)^{-1}$ is compact. If for some operator Q of finite rank at most k and $\delta > 0$, one has*

$$\langle Au, u \rangle \geqslant (\lambda + \delta)\|u\|^2 - \langle Qu, u \rangle,$$

then $N_\lambda \leqslant k$. Additionally, if there exists a subspace V with dimension at least k s.t.

$$\langle Au, u \rangle \leqslant (\lambda + \delta)\|u\|^2$$

on V, then $N_\lambda \geqslant k$.

Proof. We appeal to the minimax and maximin principles for characterizing the discrete spectrum. In particular,

$$\lambda_{k+1} = \max_{V \subset H} \min_{w \perp V} \frac{\langle Aw, w \rangle}{\|w\|^2} \geqslant \min_{w \perp Q(H)} \left(\lambda + \delta - \frac{\langle Qw, w \rangle}{\|w\|^2}\right) = \lambda + \delta.$$

Hence $N_\lambda \leqslant k$. Additionally, the minimax formula implies

$$\lambda_k \leqslant \max_V \frac{\langle Av, v \rangle}{\|v\|^2} \leqslant \lambda + \delta,$$

so $N_\lambda \geqslant k$. □

We are now ready to prove the general form of Weyl's law:

Theorem 6.5.3 (Weyl's Law)**.** *For an admissible potential V and $P_h = -h^2\Delta + V(x)$, one has*

$$|\{a \leqslant E_h \leqslant b\}| = \frac{1}{(2\pi h)^n}(|\{a \leqslant |\xi|^2 + V(x) \leqslant b\}| + o(1)).$$

Proof. Let N_λ be the number of energy states $E_h \leqslant \lambda$. Select a smooth cutoff χ that is 1 on $p \leqslant \lambda + \epsilon$ and 0 on $p \geqslant \lambda + 2\epsilon$. Then, for M large,

$$a := p + M\chi - \lambda \geqslant \gamma_\epsilon m$$

for $m = \langle \xi \rangle^2 + \langle x \rangle^m$, hence a is elliptic and therefore invertible for small h.

Lemma 6.5.1. $a^w \geqslant \gamma$ *for* $u \in H$.

Proof. Pick $b \in S(m^{\frac{1}{2}})$ such that $b^2 = a$. Note that b^w is right invertible and

$$(b^w)^{-1} r^w b^w = O(h).$$

Then,

$$a^w = b^w b^w + r^w = b^w(1 + (b^w)^{-1} r_w b^w) b^w \geqslant \|b^w\|_2^2 (1 - O(h)) \geqslant \gamma \|u\|_2^2$$

since b^w is bounded below for small enough h. □

Lemma 6.5.2. *For all* $\delta > 0$,

$$\chi^w = Q + O(h^\infty)$$

and

$$\mathit{rank}(Q) \leqslant \frac{1}{(2\pi h)^n} (|\{p \leqslant \lambda + 2\epsilon\}| + \delta).$$

Proof. Cover $p \leqslant \lambda + 2\epsilon$ with balls B_j, and define the shifted harmonic oscillator

$$P_j(h) := |hD_x - \xi_j|^2 + |x - x_j|^2,$$

and let Π be the projection onto $E_j(h) \leqslant r_j$ for all j. Write $\chi = \sum_j \chi_j$, where χ_j are supported on B_j, and let Π_j be the projection onto the j factor. By the projection theorem, $(I - \Pi_j)\chi_j^w = O(h^\infty)$. Hence, since $\Pi\Pi_j = \Pi_j$

$$(I - \Pi)\chi^w = \sum_j (I - \Pi)(I - \Pi_j)\chi_j^w = O(h^\infty).$$

It thus follows that $\chi^w = Q + O(h^\infty)$ for $Q := \Pi\chi^w$. Moreover, the rank of Q is bounded by the rank of Π, which by Weyl's law for the harmonic oscillator is bounded by the volume of the balls B_j, i.e.

$$\text{rank } Q \leqslant \frac{1}{(2\pi h)^n} \left(|\{p \leqslant \lambda + 2\epsilon\}| + \frac{\delta}{2} + o(1) \right).$$

□

We now use a theorem from the Appendix, which says that under certain conditions, the number of eigenvalues is bounded by the rank of Q. We have

$$\langle P_h u, u \rangle \geqslant (\lambda + \gamma)\|u\|_2^2 - M\langle Qu, u \rangle - O(h^\infty)\|u\|_2^2 \geqslant \lambda \|u\|_2^2 - M\langle Qu, u \rangle.$$

The theorem then implies

$$N_\lambda \leqslant \frac{1}{(2\pi h)^n} (|p \leqslant \lambda + 2\epsilon| + \delta + o(1)),$$

and sending $\epsilon, \delta \to 0$ yields

$$N_\lambda \leqslant \frac{1}{(2\pi h)^n}(|\{p \leqslant \lambda\}|o(1)).$$

We now prove the reverse inequality. We claim that $P_h \leqslant \lambda + \epsilon + O(h^\infty)$ on V_j, which is the subspace on which $E_j(h) \leqslant r_j$. To prove this, find a smooth compactly supported cutoff a that is 1 on $p \leqslant \lambda$ and supported in $p \leqslant \lambda + \frac{\epsilon}{2}$, and let $c := 1 - a$. Then,

$$1 - a^w = c^w = O(h^\infty)$$

since supp $1 - a \cap B_j = \varnothing$. Now, if $b^w = P_h a^w$, since $p \in S(m)$, $a \in S(m^{-1})$, we have $b = pa + O(h) \in S$ so b^w is bounded in L^2. Observe also that $b \leqslant \lambda + \frac{\epsilon}{2}$, so $b^w \leqslant \lambda + \frac{3\epsilon}{4}$. Since $a^w u = u + O(h^\infty)$,

$$\langle P_h u, u \rangle = \langle P_h(a^w - O(h^\infty))u, u \rangle \leqslant (\lambda + \epsilon + O(h^\infty))\|u\|_2^2,$$

which proves the claim. Finally, Let $V = \sum_j V_j$. While the V_i are not orthogonal, they are approximately orthogonal, i.e.

$$\langle u_i, u_j \rangle = O(h^\infty)\|u_i\|\|u_j\|.$$

Since the above estimate holds for each V_i, approximate orthogonality yields

$$\langle Pu, u \rangle \leqslant (\lambda + \delta)\|u\|_2^2.$$

Thus, we conclude that

$$\dim V = \sum_j \dim V_j = \frac{1}{(2\pi h)^n}(|\{p < \lambda\}| - \delta + o(1)),$$

so by the Appendix Theorem,

$$N_\lambda \geqslant \frac{1}{(2\pi h)^n}(|\{p \leqslant \lambda\}| - \delta + o(1)),$$

and we are done. □

7 Agmon Estimates

In the solution of the Schrodinger equation, the intuitive picture is that solutions exponentially decay in classically forbidden regions (that is, regions where $V(x) > E$ where E is the energy of the solution). This section presents a rigorous proof of this fact.

Definition 7.0.1. We define the **semiclassical Sobolev space** $H^k(U)$ with norm

$$\|u\|_{H_h^k} = \left(\sum_{|\alpha| \leqslant k} \|(hD)^\alpha u\|_2^2 \right)^{\frac{1}{2}}.$$

We now prove a standard form of a rescaled elliptic estimate.

Proposition 7.0.1. *For $Q_h = -h^2\Delta + \langle a, hD \rangle + b$, where a, b are smooth complex-valued functions, and $U \subset W$ is precompact, then*

$$\|u\|_{H_h^2(U)} \lesssim \|Q_h u\|_{L^2(W)} + \|u\|_{L^2(W)}.$$

Proof. Let χ be a smooth cutoff that is 1 on U. We multiply $Q_h u$ by $\chi^2 \overline{u}$ and integrate by parts to get

$$\operatorname{Re} \int_W (Q_h u)\chi^2 \overline{u} = \int h^2 \langle \partial(\chi^2 \overline{u}), \partial u \rangle + \operatorname{Re} \langle a, hDu \rangle \chi^2 \overline{u} + \operatorname{Re} b|u|^2 \chi^2 dx$$
$$\geqslant \frac{1}{2} \int_W \chi^2 |hDu|^2 dx - C \int_W |u|^2 dx,$$

where we complete the square to obtain the estimate. Thus, $\|hDu\|_2^2$ is bounded by the term on the RHS of the proposition. A similar integration by parts gives

$$\int_W |\Delta u|^2 = \sum_{i,j} \int \partial_i^2 u \partial_j^2 \overline{u} dx = \sum_{i,j} \int \partial_{ij} u \partial_{ij} \overline{u} = \int_W |D^2 u|^2,$$

so proceeding exactly as above and multiplying by $-\chi^2 h^2 \Delta \overline{u}$ yields that $\|(hD)^2 u\|_2^2$ is also bounded by the RHS. □

Before we continue, we need to introduce an important tool known as conjugation.

Definition 7.0.2. For $\phi \in C^\infty$, the **conjugation** of P_h by ϕ is

$$P_h^\phi := e^{\frac{\phi}{h}} P_h e^{-\frac{\phi}{h}}.$$

Lemma 7.0.1. *For the harmonic oscillator P_h, $P_h^\phi = p^w$ for*

$$p := \langle \xi + i\partial\phi(x), \xi + i\partial\phi(x) \rangle + V(x).$$

Proof. Direct computation and the quantization formula for linear symbols. □

We now are able to prove our first Agmon-type estimate:

Theorem 7.0.1 (Exponential Decay)**.** *Suppose U is an open set that is precompact in the classically forbidden region $\{V > E\}$. Then, for each open set W compactly containing U and λ near E, there exists $\delta > 0$ such that*

$$\|u\|_{L^2(U)} \leqslant Ce^{-\frac{\delta}{h}} \|u\|_{L^2(W)} + C\|(P_h - \lambda)u\|_{L^2(W)}$$

for $u \in C_c^\infty$ and $h < h_0$.

Proof. Pick smooth cutoffs ϕ, ψ, $\psi \equiv 1$ on U, $\phi \equiv 1$ on the support of ψ. Moreover, WLOG suppose $W \subset \{V > E\}$ compactly. Let $A_h := P_h^{\delta\psi}$ be a conjugation of $P_h - \lambda$. By the lemma above, we know the symbol of A_h - in fact, we know it satisfies

$$|\langle \xi + i\delta\partial\psi, \xi + i\delta\partial\psi \rangle + V - \lambda|^2 \gtrsim \langle \xi \rangle^4 > 0$$

for δ small enough, $x \in W$, and λ close enough to E. so the symbol is certainly elliptic in $S(\langle \xi \rangle^4)$.

If ϕ_1 is another cutoff like ϕ such that $\phi_1 \equiv 1$ on the support of ϕ, the above bound implies that

$$B := \langle hD \rangle^{-2} (\phi_1 A_h^* A_h \phi_1 - C^2 \phi_1^2) \langle hD \rangle^{-2} = b^w$$

for $b \in S$ such that $b \geqslant -Ch$. Then, by the sharp Garding inequality, $\langle v, b^w v \rangle \geqslant -Ch\|v\|_2^2$, so setting $v = \langle hD \rangle^2 \phi w$ yields

$$\|A_h \phi w\|_2^2 \geqslant C^2 \|\phi w\|_2^2 - Ch\| \langle hD \rangle^2 \phi w\|_2^2$$

for $w \in H^2_{loc}$. Then, applying the H^2 estimate to bound $\| \langle hD \rangle^2 \phi w\|_2$ from above and making h small enough ultimately yields

$$\|A_h \phi w\|_2 \geqslant \frac{\gamma}{2}\|\phi w\|_2.$$

Putting $w = e^{\frac{\delta\psi}{h}} u$ then yields

$$\|e^{\frac{\delta\psi}{h}} \phi u\|_2 \lesssim \|A_h e^{\frac{\delta\psi}{h}} \phi u\|_2 \lesssim \|e^{\frac{\delta\psi}{h}} \phi(P_h - \lambda)u\|_2 + C\|e^{\frac{\delta\psi}{h}} [P_h, \phi]u\|_2.$$

Now, since $\phi \equiv 1$ on the support of ψ and P_h acts locally, it follows that the supports of ψ and the commutator above are disjoint. Thus, again applying the H^2 bound to the commutator term and using the disjointness of supports yields

$$\|e^{\frac{\delta\psi}{h}} [P_h, \phi]u\|_2 \lesssim \|u\|_2 + \|(P_h - \lambda)u\|_2.$$

Combining these estimates, using the fact that $\psi \equiv 1$ on U, and canceling out the exponential yields the desired inequality

$$e^{\frac{\delta}{h}} \|u\|_2 \leqslant \|e^{\frac{\delta\psi}{h}} \phi u\|_2 \leqslant C\|u\|_2 + C(e^{\frac{\delta}{h}+1})\|(P_h - \lambda)u\|_2.$$

□

Corollary 7.0.1. *If U is contained in the classically forbidden region and u_h is an eigenfunction of P_h with eigenvalue $E_h \to E$ as $h \to 0$, then there exists $\delta > 0$ s.t.*

$$\|u_h\|_{L^2(U)} \leqslant e^{-\frac{\delta}{h}} \|u_h\|_2$$

for small enough h.

7.1 Tunneling

We have shown above the wavefunction decays exponentially in classically forbidden regions. Now we demonstrate that this actually is the decay rate of the wavefunction, i.e. that it is bounded below by an exponential as well. This demonstrates that some mass of the solution will always "tunnel" through classically forbidden barriers.

Definition 7.1.1 (Hormander's Ellipticity Condition)**.** For a symbol p_ϕ, the symbol is said to satisfy Hormander's ellipticity condition if

$$p_\phi = 0 \implies i\{p_\phi, \overline{p_\phi}\} > 0.$$

The last expression is always real since

$$i\{q, \overline{q}\} = i\{\operatorname{Re} q + i \operatorname{Im} q, \operatorname{Re} q - i \operatorname{Im} q\} = 2\{\operatorname{Re} q, \operatorname{Im} q\}.$$

Theorem 7.1.1. *Assuming Hormander's ellipticity condition holds on W, one has*

$$h^{\frac{1}{2}} \|u\|_{L^2(W)} \lesssim \|P_h^\phi u\|_{L^2(W)}$$

for small enough h.

Proof. A straightforward calculation shows that

$$\|P_h^\phi u\|_2^2 = \|P_h^{\phi *}\|_2^2 + \langle [P_h^{\phi *}, P_h^\phi]u, u\rangle .$$

Using the asymptotic expansion for the commutator yields that for fixed M and small enough h,

$$\|P_h^\phi u\|_2^2 \geqslant h\langle (M|p_\phi|^2 + i\{p_\phi, \overline{p_\phi})^w u, u\}\rangle + O(h^2),$$

and Hormander's ellipticity condition implies that the above symbol is elliptic, so that by the sharp Garding inequality,

$$\|P_h^\phi u\|_2^2 \gtrsim Ch\|u\|_2^2 + O(h^2).$$

□

Our goal now is to construct a weight such that p_ϕ satisfies the necessary ellipticity condition.

Lemma 7.1.1. *There exists a positive nonincreasing radial ϕ such that the Hormander ellipticity condition for p_ϕ holds in any annulus around the origin.*

Proof. Suppose ϕ takes the form $\phi = e^{\lambda\psi}$ for $\lambda > 0, \psi$ positive radial to be chosen later. Pick $\psi = \mu - |x|$ for a large enough μ so that ψ is positive. Then $|\nabla u| = 1, |D^2 u| \leqslant C$. Then, after routine calculation, we obtain

$$\frac{i}{2}\{p_\phi, \overline{p_\phi}\} \geqslant 2\lambda^4 e^{3\lambda\psi} - C\lambda^3 e^{3\lambda\psi} - C \geqslant 1$$

on the annulus if λ is chosen to be large enough. □

We also want ψ to be smooth, which we will establish in the proof of the main theorem.

Theorem 7.1.2 (Tunneling Estimate)**.** *If U is a bounded open set and u solves the eigenvalue equation*

$$P_h u = E_h u$$

for $E_h \in [a, b]$ and an admissible potential V, then for some $C > 0$ and small enough h, one has

$$\|u_h\|_{L^2(U)} \geqslant e^{-\frac{C}{h}}\|u_h\|_{L^2(\mathbb{R}^n)}.$$

This is called a ***Carleman estimate.***

Proof. WLOG assume $U = B(0, 3r)$ for $r < \frac{1}{3}$. For R large enough and $a \leqslant \lambda \leqslant b$, one may use the fact that V is bounded from below to conclude that the symbol $p - E_h$ is elliptic on $B(0, R)^c$, i.e.

$$\|(P_h - E_h)v\|_{L^2(B(0,R)^c)} \gtrsim \|v\|_{L^2(B(0,R)^c)}.$$

We now select two smooth radial cutoffs χ_1, χ_2 s.t. $\chi_1 = 1$ on $2r < |x| < R + 2$ and $\chi_2 = 1$ on $|x| > R + 1$ and 0 everywhere else. Applying the above estimate to $\chi_2 u$ yields

$$\|\chi_2 u\|_2 \lesssim \|(P_h - E_h)(\chi_2 u)\|_2 = \|[P_h, \chi_2]u\|_2$$

since $P_h u = E_h u$, and by construction, $[P_h, \chi_2]u$ is supported on $R < |x| < R + 1$. Then, by the H_h^2 estimates of Proposition 7.1, we have

$$\|[P_h, \chi_2]u\|_2 \lesssim h\|u\|_{H_h^1(R<|x|<R+1)} \lesssim h(\|(P_h - E_h)u\|_{L^2(R<|x|<R+1)} + \|u\|_{L^2(R<|x|<R+1)} \lesssim h\|\chi_1 u\|_2,$$

where the first inequality is a direct estimate and the last one follows from the definition of χ_1.

We now apply Theorem 7.3, and the commutator estimate, which gives

$$h^{\frac{1}{2}}\|e^{\frac{\phi}{h}}\chi_1 u\|_2 \lesssim \|e^{\frac{\phi}{h}}[P_h,\chi_1]u\|_2.$$

Since $[P_h,\chi_1]$ is supported on $(r<|x|<2r)\cup(R+2<|x|<R+3)$ and ϕ is nonincreasing, one may estimate

$$\|e^{\frac{\phi}{h}}[P_h,\chi_1]u\|_2 \lesssim he^{\frac{\phi(R+2)}{h}}\|\chi_2 u\|_{H_h^1(R+2<|x|<R+3)} + he^{\frac{\phi(0)}{h}}\|u\|_{H_h^1(|x|<2r)},$$

which again may be estimated by Proposition 7.1 and the previous estimate to give

$$\|e^{\frac{\phi}{h}}\chi_1 u\|_2 \lesssim h^{\frac{1}{2}}e^{\frac{\phi(R+2)}{h}}\|\chi_2 u\|_2 + h^{\frac{1}{2}}e^{\frac{\phi(0)}{h}}\|u\|_{L^2(U)}.$$

Set $A=\phi(R+2)$, and note that

$$e^{\frac{2A}{h}}\chi_1^2 \leqslant 2(e^{\frac{2\phi}{h}}\chi_1^2 + e^{\frac{2A}{h}}\chi_2^2),$$

so plugging this into the $\chi_2-\chi_1$ estimate and adding to the previous estimate yields

$$\|e^{\frac{A}{h}}\chi_2 u\|_2 + \|e^{\frac{\phi}{h}}\chi_1 u\|_2 \lesssim h\|e^{\frac{\phi}{h}}\chi_1 u\|_2 + h^{\frac{1}{2}}\|e^{\frac{A}{h}}\chi_2 u\|_2 + h^{\frac{1}{2}}e^{\frac{\phi(0)}{h}}\|u\|_{L^2(U)}.$$

Taking h sufficiently small and using that $\phi>0$ yields

$$\|\chi_2 u\|_2 + \|\chi_1 u\|_2 \lesssim h^{\frac{1}{2}}e^{\frac{\phi(0)}{h}}\|u\|_{L^2(U)},$$

and since $\chi_1^2+\chi_2^2 \geqslant \frac{1}{2}$ on a set containing U^c, the claim follows. □