M393C NOTES: TOPICS IN MATHEMATICAL PHYSICS

ARUN DEBRAY
SEPTEMBER 28, 2017

These notes were taken in UT Austin's M393C (Topics in Mathematical Physics) class in Fall 2017, taught by Thomas Chen. I live-TEXed them using vim, so there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu. Thanks to Yanlin Cheng for fixing some typos.

Contents

Lecture 1.

The Lagrangian formalism for classical mechanics: 8/31/17

The audience in this class has a very mixed background, so this course cannot and will not assume any physics background. We'll first discuss classical and Lagrangian mechanics. Quantum mechanics is, of course, more fundamental, and though historically people obtained quantum mechanical mechanics from classical mechanics, it should be possible to go in the other direction.

We'll start, though, with classical and Lagrangian mechanics. This involves understanding symplectic and Poisson structures, and the principle of least action, the beautiful insight that classical mechanics can be formulated variationally; there is a Lagrangian L and an action functional

$$S = \int_{t_0}^{t_1} L \, \mathrm{d}t,$$

and the system evolves through paths that extremize the action functional.

The history of the transition from classical mechanics to quantum mechanics to quantum field theory happened extremely quickly in the historical sense, all fitting into one lifetime. JJ Thompson discovered the electron in 1897, and in 1925, GP Thompson, CJ Dawson, and LH Germer discovered that it had mass. This led people to discover some inconsistencies with classical physics on small scales, ushering in quantum mechanics, with all of the famous names: Einstein, Schrödinger, Heisenberg, and more. The basic equations of quantum mechanics fall in linear dispersive PDE for functions living in the Hilbert space, typically L^2 or the Sobolev space H^1 (since energy involves a derivative).

One of the key new constants in quantum mechanics is *Planck's constant* $\hbar := h/2\pi$. It has the same units as the classical action S, and therefore they are comparable. There is a sense in which quantum mechanics is the regime in which $S/\hbar \approx 1$, and classical mechanics is the regime in which $S/\hbar \gg 1$. In this sense, quantum mechanics is the physics of very small scales. Sometimes people take a "semiclassical limit," and say they're letting $\hbar \to 0$, but this makes no sense: $\hbar$ is a physical quantity. Instead, it's more accurate to say taking a semiclassical limit lets $(S/\hbar)^{-1} \to 0$.

If you want to analyze a fixed number of electrons, life is good. They will always be there, and so on. But this is a problem for photons, as there are physical processes which create photons, and processes which destroy photons. Thus imposing a fixed number of quantum particles is a constraint — and the theory which describes the quantum physics of arbitrary numbers of quantum particles, quantum field theory, was worked out a little later. In this case, the Hilbert space is a direct sum over the Hilbert subspace of 1-particle states, 2-particle states, etc., and is called *Fock space*. The symplectic and Poisson structures of classical mechanics, transformed into commutation relations of operators in quantum mechanics, is again interpreted as commutation relations of creation and annihilation operators.

The mathematics of quantum field theory is rich and diverse, drawing in more PDE as well as large amounts of geometry and topology. But there's a problem — many important integrals and power series don't converge. And this is not a formal series problem: it's too central. Physicists have used renormalization as a formal trick to solve these divergences; it feels like a dirty trick that produces incredibly accurate results agreeing with experiment. But again there are problems: renormalization expresses Fock space and the commutation relations in terms of the noninteracting case, and the results you get don't necessarily agree with what you did *a priori*.

For example, quantum field theory contains a Hamiltonian H whose spectrum is of interest. One can imagine starting with the noninteracting Hamiltonian H_0 and perturbing it by some small operator W: $H := H_0 + W$. You're often interested in the resolvent

$$\begin{aligned} R(z) &= (H - z)^{-1} \\ &= (H_0 - z)^{-1} \sum_{\ell=0}^{\infty} \Big(W(H_0 - z)^{-1} \Big)^{\ell}. \end{aligned}$$

The issue is that adding W does not do nice things to the spectrum, and this is part of the complexity of quantum field theory.

Let λ denote the interaction, and N denote the number of particles, and suppose $\lambda \sim 1/N$ as we let $N \to \infty$. Then, the equations describing the mean field theory for this system are complicated, typically nonlinear PDEs. Typical examples include the nonlinear Schrödinger equation, the nonlinear Hartree equation, the Vlasov equation, or the Boltzmann equation. We'll hopefully see some of these equations in this class.

This is a lot of stuff that's tied together in complicated and potentially confusing ways, and hopefully in this class we'll learn how to make sense of it.

Classical mechanics and symplectic geometry In classical mechanics, we think of objects in idealized ways, e.g. thinking of a stone as a point mass at its center of mass. Thus, we're studying the motion of idealized point masses (or particles, in the strictly classical sense). We do this by letting time be $t \in \mathbb{R}$; at a time t, the particles $x_1, \dots, x_N$ have positions $\mathbf{q}(t) := (q_1(t), \dots, q_N(t))$, with $q_i(t) \in \mathbb{R}^d$; these are called "generalized coordinates."

Classical mechanics says that the kinematics of particles can be completely described by their position and velocity. Thus the motion of a system is completely determined by $\mathbf{q}(t)$ and $\dot{\mathbf{q}}(t) := \frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t}$.

The next question: what determines the motion? The answer is the Newtonian equations of motion: $\ddot{\mathbf{q}}$ is expressed as a function of $\dot{\mathbf{q}}$ and $\mathbf{q}$ using *Hamilton's principle*, also known as the *principle of least action*.

(1) Let $\mathbf{q} \in C^2([t_0, t_1], \mathbb{R}^{Nd})$ be a curve in $\mathbb{R}^{Nd}$. We associate to $\mathbf{q}$ a weight function $L(\mathbf{q}, \dot{\mathbf{q}})$ called the *Lagrangian*.

(2) Given $\mathbf{q}$ as above, define the *action functional*

$$S[\mathbf{q}] := \int_{t_0}^{t_1} L(\mathbf{q}(t), \dot{\mathbf{q}}(t))\, \mathrm{d}t.$$

(3) Then, among all C^2 curves with $\mathbf{q}(t_0)$ and $\mathbf{q}(t_1)$ fixed, the curve that minimizes S is the one that satisfies the equations of motion.

Now let $\mathbf{q}_\bullet(t)$ be a C^2 family of curves $[t_0, t_1] \times \mathbb{R} \to \mathbb{R}^{Nd}$ and that $\mathbf{q}_0$ minimizes S. Then,

$$\partial_s|_{s=0}\, S[\mathbf{q}_s] = 0.$$

We can apply this to the Lagrangian to derive the equations of motion.

$$\begin{aligned}\partial_s|_{s=0}\, S[\mathbf{q}_s] &= \int_{t_0}^{t_1} \left((\nabla_{\mathbf{q}_s} L)\cdot \partial_s \mathbf{q}_s(t) + (\nabla_{\dot{\mathbf{q}}_s} L)\cdot \partial_s \dot{\mathbf{q}}_s(t) \right) \mathrm{d}t \bigg|_{s=0} \\ &= \int_{t_0}^{t_1} \left(\nabla_{\mathbf{q}_s} L - (\nabla_{\dot{\mathbf{q}}_s} L)^\bullet \right)\big|_{s=0} \cdot \underbrace{\partial_s|_{s=0}\, \mathbf{q}_s(t)}_{\delta \mathbf{q}(t)} \,\mathrm{d}t + (\nabla_{\dot{\mathbf{q}}_0} L)\cdot \underbrace{(\partial_s|_{s=0}\mathbf{q}(t))}_{=0} \Bigg|_{t_0}^{t},\end{aligned}$$

where $\delta \mathbf{q}(t)$ is the variation. For all variations, this is nonzero. Thus, minimizers of S satisfy the *Euler-Lagrange equations*

$$\nabla_{\mathbf{q}} L - (\nabla_{\dot{\mathbf{q}}} L)^\bullet = 0. \tag{1.1}$$

We'll now impose some conditions on L that come from reasonable physical principles.

Additivity: if we analyze a system $A \cup B$ which is a union of two subsystems A and B that don't interact, then

$$L_{A\cup B} = L_A + L_B.$$

Uniqueness: Assume L_1 and L_2 differ only by a total time derivative of a function $f(\mathbf{q}(t), t)$; then, they should give rise to the same equations of motion:

$$\begin{aligned} S_2 &= S_1 + \int_{t_0}^{t_1} \partial_t f(\mathbf{q}(t), t)\,\mathrm{d}t \\ &= S_1 + f(\mathbf{q}(t_1), t_1) - f(\mathbf{q}(t_0), t_0), \end{aligned}$$

so the minimizers for S_1 and S_2 are the same.

Galilei relativity principle: The physical laws of a closed system are invariant under the symmetries of the *Galilei group* parameterized by $a, v \in \mathbb{R}^d$, $t \in \mathbb{R}$, and $R \in \mathrm{SO}(d)$, the group element $g_{a,v,R,b}$ acts by

$$\begin{aligned} \mathbf{q} &\longmapsto a + vt + Rq \\ t &\longmapsto t + b. \end{aligned}$$

That is, in each component j, $q_j \mapsto a + vt + Rq_j$.

This actually determines L for a system consisting of a single particle. By homogeneity of space (by the Galilei group contains translations), L can only depend on $V = \dot{q}$. Since space is isotropic (because the Galilei group contains rotations), L should depend on v^2. Next, the Euler-Lagrange equations imply

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v} - \frac{\partial L}{\partial q} = 0,$$

and since L does not depend on q, $\frac{\partial L}{\partial q} = 0$, so $\frac{\partial L}{\partial v}$ must be a constant.

Now we consider Galilei invariance of v If $v \mapsto v + \varepsilon$, the equations of motion must be invariant, so

$$L[(v')^2] = L[(v+\varepsilon)^2] = L(v^2) + \frac{\partial L}{\partial v^2} 2v \cdot e + O(\varepsilon),$$

and this should only differ by a total time derivative $\dot{q}$:

$$F(\dot{q}) \cdot \dot{q} = \partial_t G,$$

where $F(\dot{q})$ is a constant, and $\frac{\partial L}{\partial v^2}$ is also constant. This latter constant is denoted m, and called the *mass*, and the Lagrangian expresses its kinetic energy:

$$L(v) = \frac{1}{2} m v^2.$$

Now imagine adding N particles, which we assume don't interact. Then additivity tells us they have masses $m_1, \dots, m_N$, and the Lagrangian is

$$L = \frac{1}{2} \sum_{j=1}^{N} m_j v_j^2.$$

If the particles are interacting, there's some potential function $U(q_1, \dots, q_N)$, and the Lagrangian is instead

$$L = \frac{1}{2}\sum_{j=1}^{N} m_j v_j^2 - U(q_1, \dots, q_N).$$

Now, by (1.1),

$$m_j \ddot{q}_j = -\partial_{q_j} U = F,$$

and this is called the *force*. This is Newton's second law $F = ma$.

Symmetries and conservation laws There's a general result called Noether's theorem which shows that any symmetry of a physical system leads to a conserved quantity. We'll see the presence of symmetry in classical mechanics and then how it changes in quantum mechanics.

For example, the systems we saw above had symmetries under time translation invariance $t \mapsto t + b$, so the Lagrangian doesn't depend on t, just on $\mathbf{q}$ and $\dot{\mathbf{q}}$. Therefore

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} L &= \sum_j \left(\frac{\partial L}{\partial q_j} \dot{q}_j + \frac{\partial L}{\partial \dot{q}_j} \ddot{q}_j \right) \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \sum_{j=1}^{N} \left(\frac{\partial L}{\partial \dot{q}_j} \right) \cdot \dot{q}_j,\end{aligned}$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}t} \underbrace{\left(\sum_{j=1}^{N} \frac{\partial L}{\partial \dot{q}_j} \cdot \dot{q}_j - L \right)}_{E} = 0.$$

The quantity E is the *energy* of the system, and time translation invariance tells is that energy is conserved. The component $p_j := \frac{\partial L}{\partial \dot{q}_j}$ is called the j^{th} *canonical momentum*.

The homogeneity of space, told to us by invariance under the Galilei translations $q_j \mapsto q_j + \varepsilon$, tells us that

$$\begin{aligned}\delta L &= \sum_i \frac{\partial L}{\partial \dot{q}_j} \cdot \varepsilon \\ &= \varepsilon \frac{\mathrm{d}}{\mathrm{d}t} \sum \frac{\partial L}{\partial \dot{q}_j} = 0.\end{aligned}$$

Thus, the quantity

$$\mathbf{p} := \sum_{j=1}^{N} \frac{\partial L}{\partial \dot{q}_j}$$

is conserved, and is constant. This is called the *total momentum*, so translation-invariance gives you conservation of momentum. In the same way, rotation-invariance around any center gives you conservation of angular momentum around any center.

Hamiltonian dynamics The Euler-Lagrange equations express $\ddot{\mathbf{q}}$ as a second-order ODE. One might want to reformulate this into a first-order ODE; there are many ways to do this. There's one that's particularly important. Since

$$p_j = \frac{\partial L}{\partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}),$$

then it looks like one could solve for $\dot{\mathbf{q}}$ in terms of $\mathbf{p}$ and $\mathbf{q}$.

Lemma 1.2. *Let $f \in C^2(\mathbb{R}^n, \mathbb{R})$ be such that its Hessian $D^2 f$ is uniformly positive definite, i.e. there's an $\alpha > 0$ such that*

$$D^2 f(x)(h,h) = \sum_{i,j} \frac{\partial^2 f}{\partial x_j \partial x_\ell} h_j h_\ell \geq \alpha \|h\|^2$$

uniformly in $x \in \mathbb{R}^n$, then there is a unique solution to

$$Df(x) = y$$

for every $y \in \mathbb{R}^n$.

Proof. Let $g(x,y) := f(x) - \langle x, y\rangle$. Then, $\nabla_x g(x,y) = \nabla f - y$, and $D^2 g = D^2 f$. Hence it suffices to check for $y = 0$.

The positive definite assumption on $D^2 f$ means f is strictly convex, and hence has at most a single critical point, at which $\nabla f = 0$. Thus it remains to check that there's at least one solution.

If you Taylor-expand, you get that

$$f(x) = f(0) + \langle Df(0), x\rangle + \frac{1}{2} D^2 f(sx)(x,x) + \cdots,$$

so for all x,

$$f(x) \geq f(0) - |\nabla f(0)||x| + \frac{\alpha}{2}|x|^2.$$

Thus, there's an $R > 0$ such that if $|x| \geq R$, then $f(x) \geq f(0)$, so f has at most one minimum in the ball $\overline{B_R(0)}$, so by compactness, it has a minimum x_0, which must be the global minimum, so $Df(x_0) = 0$. ⊠

Definition 1.3. Suppose f is continuous on $\mathbb{R}^n$. Then, its *Legendre transform* or *Legendre-Fenchel transform* is

$$f^*(y) := \sup_{x \in \mathbb{R}^n} (\langle y, x\rangle - f(x)).$$

You can think of this as measuring the distance from the graph of f to the line cut out by $\langle y, x\rangle$ (i.e. between the two points with minimum distance).

Lecture 2.

The Hamiltonian formalism for classical mechanics: 9/5/17

Last time, we discussed Lemma 1.2, that if $f\colon \mathbb{R}^n \to \mathbb{R}$ is C^2 and its Hessian is uniformly positive definite, then there's a unique solution to $\nabla f(x) = y$ for all $y \in \mathbb{R}^n$. We then defined the Legrendre-Fenchel transform of f: $f^*(y)$ geometrically means the minimal distance from $f(x)$ to the hyperplane $\langle y, x\rangle = 0$. It has the following key properties:

Theorem 2.1. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a C^2 function with uniformly positive definite Hessian. Then,*

(1)
$$f^*(y) = \langle y, x(y)\rangle - f(x(y)),$$
where $x(y)$ is the unique solution to $\nabla f(x) = y$ guaranteed by Lemma 1.2, and

(2) *$f^*(y)$ is C^2 and strictly convex.*

(3) *If $n = 1$, $\nabla(f^*) = (\nabla f)^{-1}$.*

(4) *For all $x, y \in \mathbb{R}^n$,*
$$f(x) + f^*(y) \geq \langle y, x\rangle,$$
with equality iff $x = x(y)$ is the unique solution to $\nabla f(x) = y$.

(5) *The Legendre-Fenchel transform is involutive, i.e. $(f^*)^* = f$.*

We'll use this in the Hamiltonian formalism of classical mechanics. One motivation for the Hamiltonian formalism is that the Lagrangian formalism produces second-order ODEs, and it would be nice to have an approach that gives first-order equations. There are many ways to do that, but this one has particularly nice properties.

Suppose we have generalized coordinates $\mathbf{q}$ and $\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$. You might ask whether we can solve for $\dot{q}_i = \dot{q}_i(\mathbf{q}, \mathbf{p})$. If we assume $D^2_{\mathbf{v}} L(\mathbf{q}, \mathbf{v})$ is uniformly positive definite, then $\mathbf{p} = \nabla_{\dot{\mathbf{q}}} L(\mathbf{q}, \dot{\mathbf{q}})$ has a unique solution.

Definition 2.2. The *Hamiltonian* H is the Legendre-Fenchel transform of L for $\mathbf{q}$ fixed, i.e.

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}) &:= \sup_{\mathbf{v} \in \mathbb{R}^n} (\langle \mathbf{p}, \mathbf{v}\rangle - L(\mathbf{q}, \mathbf{v})) \\ &= \langle \mathbf{p}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p})\rangle - L(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p})). \end{aligned}$$

Theorem 2.3. *Assume the matrix*

$$\left[\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}\right] \tag{2.4}$$

is uniformly positive definite. Then, the Euler-Lagrange equations

$$\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)^{\bullet} - \frac{\partial L}{\partial \mathbf{q}} = 0$$

are equivalent to

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}, \qquad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}. \tag{2.5}$$

(2.4) is called the *mass matrix* of the system, and (2.5) is called the *Hamiltonian equations of motion*.

Proof. Since $p_j = \frac{\partial L}{\partial \dot{q}_j}$,

$$\begin{aligned} \frac{\partial H}{\partial p_i} &= \dot{q}_i + \sum_{j=1}^{n} \left(p_j \frac{\partial \dot{q}_j}{\partial p_i} - \frac{\partial L}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial p_i} \right) \\ &= \dot{q}_i. \end{aligned}$$

Similarly, since $\frac{\partial q_j}{\partial q_i} = \delta_{ij}$ and $\frac{\partial L}{\partial \dot{q}_j} = p_j$, then

$$\begin{aligned} \frac{\partial H}{\partial q_i} &= \sum_{j=1}^{n} \left(p_j \frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial L}{\partial q_j} \frac{\partial q_j}{\partial q_i} - \frac{\partial L}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial q_i} \right) \\ &= -\left(\frac{\partial L}{\partial \dot{q}_i}\right)^{\bullet} = \dot{p}_i. \end{aligned}$$

⊠

This leads to the Hamiltonian formalism, which starts with the Hamiltonian and works towards the physics from there. We begin on a phase space $\mathbb{R}^{2n}$ with coordinates $(\mathbf{q}, \mathbf{p})$, and a Hamiltonian $H\colon \mathbb{R}^{2n} \to \mathbb{R}$. Let

$$J := \begin{bmatrix} 0 & \mathbf{1}_n \\ -\mathbf{1}_n & 0 \end{bmatrix}$$

denote the *symplectic normal matrix*.[1]

The *Hamiltonian vector field* for this system is

$$X_H := J \nabla H = \begin{bmatrix} \nabla_{\mathbf{p}} H \\ -\nabla_{\mathbf{q}} H \end{bmatrix}.$$

Then, the Hamiltonian equations of motion (2.5) may be expressed in terms of the flow for X_H.

This "Hamiltonian structure" on $\mathbb{R}^{2n}$ is closely related to a complex structure: $J^2 = -1$ is closely reminiscent of $i^2 = -1$. Indeed, if

$$\mathbf{z} := (\mathbf{q} + i\mathbf{p}),$$

then

$$\begin{aligned} i\dot{\mathbf{z}} &= i(\dot{\mathbf{q}} + i\dot{\mathbf{p}}) \\ &= i(\nabla_{\mathbf{p}} H - i \nabla_{\mathbf{q}} H) \\ &= (\nabla_{\mathbf{q}} + i \nabla_{\mathbf{p}}) H. \end{aligned}$$

This is an example of a *Wirtinger derivative*:

$$\begin{aligned} \partial_z &= \frac{1}{2}(\partial x - i \partial y) \\ \partial_{\overline{z}} &= \frac{1}{2}(\partial x + i \partial y) \end{aligned}$$

[1] More generally, one can formulate this system on any symplectic manifold, in which case J is the symplectic form in Darboux coordinates. But we won't worry about this right now.

Example 2.6 (Harmonic oscillator)**.** Let

$$H(q,p) = \frac{1}{2}q^2 + \frac{1}{2}p^2,$$

so

$$H(z,\overline{z}) = \frac{1}{2}z\overline{z}.$$

In this case, the Hamiltonian equations of motion are

$$\begin{aligned} i\dot{z} &= 2\partial_{\overline{z}}H = z \\ z(0) &= z_0, \end{aligned}$$

so we recover

$$z(t) = z_0 e^{it},$$

as usual for a harmonic oscillator. ◀

We can also study Hamiltonian PDEs, which include several interesting systems of equations. But they got erased before I could write them down. :(One of them includes the *nonlinear Schrödinger equation*: for $x \in \mathbb{R}^d$, the system

$$\mathcal{H}[u,\overline{u}] = \int \left(\frac{1}{2}|\nabla u|^2 + \frac{1}{2p}|u|^{2p} \right) \mathrm{d}x,$$

which leads to the equations of motion (the Schrödinger equation)

$$i\dot{u} = -\Delta u + |u|^{2p-2}u.$$

The solutions of these equations tend to be interesting: Hamiltonian flow (the flow generated by X_H) isn't a gradient flow, but rather gradient flow twisted by J. We call this flow $\Phi_t\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, with $x(t) = \Phi_t(x_0)$ and $x(t) = \Phi_{t,s}(x(s))$.

Theorem 2.7. *H is conserved by Φ_t.*

Proof.

$$\frac{\mathrm{d}}{\mathrm{d}t}H(x(t)) = \nabla_{\mathbf{x}}H \cdot \dot{\mathbf{x}} = \nabla_{\mathbf{x}}H \cdot J\nabla_{\mathbf{x}}H = 0,$$

because J is skew-symmetric. ⊠

Definition 2.8. In this situation, the *symplectic form* is the skew-symmetric form $\omega \in \Lambda^2((\mathbb{R}^{2n})^*)$ defined by

$$\omega(X,Y) := \langle Y, JX \rangle.$$

The pair $(\mathbb{R}^{2n},\omega)$ is a symplectic vector space; the space of invertible matrices preserving this form is called the *symplectic group*

$$\mathrm{Sp}(2n,\mathbb{R}) := \{M \in \mathrm{GL}_{2n}(\mathbb{R}) \mid M^{\mathrm{T}}JM = J\}.$$

Now we can prove some properties of the Hamiltonian flow.

Theorem 2.9. *Let Φ_t be the Hamiltonian flow generated by X_H. Then,*

(1) $x(t) = \Phi_{t,s}(x(s))$,
(2) $\Phi_{s,s} = \mathrm{id}$, *and*
(3) $D\Phi_{t,s}(x) \in \mathrm{Sp}(2n,\mathbb{R})$.

Conversely, if $\Phi_{t,s}$ is the local flow generated by a vector field X such that locally (in x) (3) holds, then X is locally Hamiltonian, in that there's a G such that $X = X_G$.

Definition 2.10. A diffeomorphism $\phi\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ with $D\phi \in \mathrm{Sp}(2n,\mathbb{R})$ is called a *symplectomorphism*.

Proof sketch of Theorem 2.9. Since

$$\partial_t D\Phi_{t,s}(x) = DX_H(\Phi_{t,s}(x)) \cdot D\Phi_{t,s}(x),$$

then it suffices to check that if

$$\Gamma(t,s,x) := D\Phi_{t,s}^{\mathrm{T}}(x)JD\Phi_{t,s}(x),$$

then

$$\frac{\mathrm{d}}{\mathrm{d}t}\Gamma = 0.$$

Definition 2.11. The *Liouville measure* μ_L on $\mathbb{R}^{2n}$ is the measure induced by $\omega^{\wedge n}$, i.e.

$$\int_{\mathbb{R}^{2n}} f \,\mathrm{d}\mu_L := \int_{\mathbb{R}^{2n}} f\omega^{\wedge n}.$$

Theorem 2.12 (Liouville). *Let $\Phi_{t,s}$ be the Hamiltonian flow. Then, for every Borel set B, $|\Phi_{t,s}(B)| = |B|$. Hence $\Phi_{t,s}$ preserves the Lesbegue measure and the Liouville measure.*

Proof. If $\varphi\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is a diffeomorphism, then

$$\int_B f(x)\,\mathrm{d}x = \int_{\varphi^{-1}(B)} (f\circ\varphi)|\det D\varphi(x)|\,\mathrm{d}x,$$

and $\det D\Phi_t = 1$. ⊠

The next theorem is a conservation property.

Theorem 2.13. *Let $\Phi_{t,s}$ be the flow generated by an arbitrary vector field X, $D \subset \mathbb{R}^{2n}$ be a bounded region, and $D_{t,s} := \Phi_{t,s}(D)$. Then, for every $f \in C^1(\mathbb{R}^n)$,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{D_{t,s}} f\,\mathrm{d}x = \int_{D_{t,s}} (\partial_t f + \operatorname{div}(fX))\,\mathrm{d}x.$$

Proof. By the group property ($\Phi_{t,s} = \Phi_{t,s_1}\circ\Phi_{s_1,s}$) it suffices to prove it for $s = 0$ and at $t = 0$. In this case

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}\int_{D_t} f\,\mathrm{d}x = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}\int_D (f\circ\Phi_t)\det D\Phi_t\,\mathrm{d}x$$

Since $D\Phi_t = \mathbf{1} + tDX + O(t^2)$, then $\det(D\Phi_t) = 1 + t\operatorname{tr}(DX) + O(t^2)$ and hence

$$\begin{aligned} &= \int_D ((\partial f + \nabla f\cdot X) + f\operatorname{div} X)\,\mathrm{d}x \\ &= \int_D (\partial_t f + \operatorname{div} fX)\,\mathrm{d}x. \end{aligned}$$ ⊠

Corollary 2.14. *Any function $f(t,x)$ for which the* matter content

$$MC(f)(t) := \int_{\Phi_{t,s}(D)} f(x,t)\,\mathrm{d}x$$

remains constant (equivalently, $\frac{\mathrm{d}}{\mathrm{d}t}MC(f)(t) = 0$), must satisfy the continuity equation

$$\partial_t f + \operatorname{div}(fX) = 0. \tag{2.15}$$

In physically interesting cases, the matter content actually represents how much mass is in the system. In the Hamiltonian case, $\operatorname{div} X_H = 0$, so

$$\partial_t f + \nabla f\cdot X_H = 0$$

is equivalent to

$$\partial_t f + \nabla f\cdot J\nabla H = 0.$$

We can rewrite this in terms of the *Poisson bracket*

$$\{f,H\} := \langle\nabla f, J\nabla H\rangle,$$

producing the equation

$$\partial_t f + \{f,H\} = 0.$$

The Poisson bracket can also be defined as

$$\begin{aligned} \{f,H\} &= \omega(X_f, X_H) \\ &= \sum_{j=1}^n\left(\frac{\partial f}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial H}{\partial q_j}\frac{\partial f}{\partial p_j}\right). \end{aligned}$$

We'll see related phenomena in the quantum-mechanical case. What we talk about next, though, will not reappear in quantum mechanics, but it's too beautiful to ignore completely.

Definition 2.16. An *integral of motion* is a C^1 function $g\colon \mathbb{R}^{2n} \to \mathbb{R}$ constant along the orbits of the Hamiltonian. Equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t} g(x(t)) = \{g, H\} = 0.$$

Two integrals of motion g_1 and g_2 are *in involution* if $\{g_1, g_2\} = 0$.

Notice that $\{g, g\} = 0$ always.

Generally, Hamiltonian systems are incredibly difficult to solve. There are some cases where they can be solved by hand, e.g. by quadrature classically. It would be nice to know when such a solution exists. If you can find n integrals of motion that are in involuton with each other, you can heuristically reduce the equations into something tractable; this is the contant of the Arnold-Yost-Liouville theorem.

Theorem 2.17 (Arnold-Yost-Liouville)**.** *On the phase space $(\mathbb{R}^{2n}, \omega)$, assume we have n integrals of motion $G_1, \dots, G_n$ which are in involution; further, assume $G_1 = H$. Let $\mathbf{G} = (G_1, \dots, G_n)\colon \mathbb{R}^{2n} \to \mathbb{R}^n$, and consider its level set*

$$\mathcal{M}_{\mathbf{G}}(\mathbf{c}) := \{x \in \mathbb{R}^{2n} \mid \mathbf{G}(x) = \mathbf{c}\},$$

for some $\mathbf{c} \in \mathbb{R}^n$. Assume that the 1-forms $\{\mathrm{d}G_j\}$ are linearly independent (equivalently, the gradients ∇G_j are linearly independent). Then,

(1) *$\mathcal{M}_{\mathbf{G}}(\mathbf{c})$ is a smooth manifold that's invariant under the flow generated by X_H, and*
(2) *if $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$ is compact and connected, it is diffeomorphic to an n-torus $T^n := S^1 \times \cdots \times S^1$.*
(3) *The Hamiltonian flow of H determines a quasiperiodic motion*

$$\frac{\mathrm{d}\boldsymbol{\varphi}}{\mathrm{d}t} = \boldsymbol{\eta}(\mathbf{c}), \qquad \frac{\mathrm{d}\mathbf{I}}{\mathrm{d}t} = \mathbf{0} \tag{2.18}$$

with initial data $(\boldsymbol{\varphi}_0, \mathbf{I}_0)$.

(4) *The Hamiltonian equations of motion can be integrated by quadrature:*

$$\begin{aligned} \mathbf{I}(t) &= \mathbf{I}_0 \\ \boldsymbol{\varphi}(t) &= \boldsymbol{\varphi}_0 + \boldsymbol{\eta}(\mathbf{c})t. \end{aligned} \tag{2.19}$$

Here $\mathbf{I}$ and $\boldsymbol{\varphi}$ are the new coordinates for phase space in which the system can be solved.

We'll prove this next lecture, then move to quantum mechanics.

Lecture 3.

The Arnold-Yost-Liouville theorem and KAM theory: 9/7/17

Today, we're going to prove the Arnold-Yost-Liouville theorem, Theorem 2.17. We keep the notation from that theorem and the notes before it.

One key takeaway from the theorem is that the Hamiltonian equations can be explicitly solved. That is, going from (2.18) to (2.19) is a particularly simple system of ODEs.

Proof sketch of Theorem 2.17. By assumption, $\{\nabla G_j\}$ is linearly independent on $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$. By the implicit function theorem, $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$ is an n-dimensional submanifold of $\mathbb{R}^{2n}$. The gradients $\{\nabla G_j\}$ span the normal bundle of $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$ because it's a level set for them.

Consider $X_{G_j} := J\nabla G_j$. It's a tangent vector:

$$\begin{aligned} \langle X_{G_j}, \nabla G_\ell \rangle &= \langle J\nabla G_j, \nabla G_\ell \rangle \\ &= -\langle J\nabla G_j, JJ\nabla G_\ell \rangle \\ &= \omega\left(X_{G_j}, X_{G_\ell}\right) \\ &= \{G_j, G_\ell\} = 0 \end{aligned} \tag{3.1}$$

for all j and ℓ. We've produced n linearly independent tangent vectors at each point, so $\{X_{G_j}\}_{j=1}^n$ spans $T\mathcal{M}_{\mathbf{G}}(\mathbf{c})$. In particular, $X_H = X_{G_1}$ is tangent to $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$, so $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$ is invariant under its flow. This proves (1).

For part (2), we assume $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$ is compact and connected. Let $\varphi^j_{t_j}$ denote the flow generated by X_{G_j}, so $t_1, \dots, t_n \in \mathbb{R}$ are separate time variables. Because $\{G_j, G_\ell\} = 0$, then G_ℓ is invariant under $\varphi^j_{t_j}$ for any j and ℓ. Thus $\varphi^j_{t_j}$ and $\varphi^\ell_{t_\ell}$ commute, so we may define

$$\varphi_{\mathbf{t}} := \varphi^1_{t_1} \circ \cdots \circ \varphi^n_{t_n}.$$

Pick an $x_0 \in \mathcal{M}_{\mathbf{G}}(\mathbf{c})$ and define $\varphi\colon \mathbb{R}^n \to \mathcal{M}_{\mathbf{G}}(\mathbf{c})$ to send $\mathbf{t} \mapsto \varphi_{\mathbf{t}}(x_0)$. This is transitive in the sense that for all $x \in \mathcal{M}_{\mathbf{G}}(\mathbf{c})$, there's a $\boldsymbol{\tau} \in \mathbb{R}^n$ such that $\varphi_{\boldsymbol{\tau}}(x_0) = x$.

Since $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$ is compact but $\mathbb{R}^n$ isn't, φ cannot be a bijection. Define

$$\Gamma_{x_0} := \{\mathbf{t} \in \mathbb{R}^n \mid \varphi_{\mathbf{t}}(x_0) = x_0\},$$

the *stationary group* of x_0. This is indeed an abelian group, because if $\boldsymbol{\tau} \in \Gamma_{x_0}$, then $n\boldsymbol{\tau} \in \Gamma_{x_0}$ for all $n \in \mathbb{Z}$: if you iterate a loop again and again, you still end up back where you started with. And clearly $\mathbf{0} \in \Gamma_{x_0}$.

Let ${}_{\varepsilon_1}U$ be an ε_1-neighborhood of $\mathbf{0}$ in $\mathbb{R}^n$ and V_{ε_2} be an ε_2-neighborhood of x_0 in $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$; then, there are $\varepsilon_1, \varepsilon_2 > 0$ such that $\varphi|_{U_{\varepsilon_1}} : U_{\varepsilon_1} \to V_{\varepsilon_2}$ is a diffeomorphism. Thus, for sufficiently small ε_2, there's no other fixed point in V_{ε_2}, which means Γ_{x_0} is a discrete subgroup of $(\mathbb{R}^n, +)$.

This means there are vectors $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{R}^n$ such that

$$\Gamma_{x_0} = \left\{ \sum_{i=1}^n m_i \mathbf{e}_i \mid m_1, \dots, m_n \in \mathbb{Z} \right\},$$

and that φ establishes an isomorphism

$$T^n \cong \mathbb{R}^n / \Gamma_{x_0} \longrightarrow \mathcal{M}_{\mathbf{G}}(\mathbf{c}).$$

This proves (2).

Now we need to make the change-of-variables in (3); these new variables are called *action-angle variables*. First note that $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$ is a *Lagrangian submanifold*, i.e. it's half-dimensional and the restriction of ω to it is 0 (it's *isotropic*; an isotropic submanifold of $\mathbb{R}^{2n}$ can be at most n-dimensional). This is because $T\mathcal{M}_{\mathbf{G}}(\mathbf{c})$ is spanned by $\{X_{G_j}\}$, and in (3.1), we proved $\omega(X_{G_j}, X_{G_\ell}) = \{G_j, G_\ell\} = 0$ for all j, ℓ.

Consider the 1-form

$$\Theta := \sum_j p_j \, \mathrm{d}q_j.$$

Then,

$$\mathrm{d}\Theta = \sum_j \mathrm{d}p_j \wedge \mathrm{d}q_j = \omega,$$

so restricted to $\mathcal{M}_{\mathbf{G}}(\mathbf{c})$, Θ is a closed 1-form.

Let $\{\gamma_j\}_{j=1}^n$ be a set of cycles whose homology classes generate $H_1(\mathcal{M}_{\mathbf{G}}(\mathbf{c})) = H_1(T^n) \cong \mathbb{Z}^n$. Then, the *action variables*

$$I_j(\mathbf{c}) := \frac{1}{2\pi} \oint_{\gamma_j} \Theta$$

is independent of the choice of cycle representative of the homology class of γ_j: if D is a 2-chain with $\partial D = \gamma_j - \widetilde{\gamma}_j$ (a cobordism or homotopy from γ_j to $\widetilde{\gamma}_j$), then by Stokes' theorem.

$$\oint_{\gamma_j} \Theta - \oint_{\widetilde{\gamma}_j} \Theta = \int_D \mathrm{d}\Theta = \int_D 0 = 0.$$

One can show that the assignment $(\mathbf{q}, \mathbf{p}) \mapsto (\boldsymbol{\varphi}, \mathbf{I})$ is symplectic, where φ_j is a variable parameterizing γ_j and is called an *angle variable* (since it's valued in S^1). In these coordinates, H only depends on $\mathbf{I}$, not $\boldsymbol{\varphi}$, so

$$\frac{\mathrm{d}\boldsymbol{\varphi}}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{I}} = \boldsymbol{\eta}(\mathbf{c})$$
$$\frac{\mathrm{d}\mathbf{I}}{\mathrm{d}t} = -\frac{\partial H}{\partial \boldsymbol{\varphi}} = 0. \qquad \boxtimes$$

Sometimes the entires of $\boldsymbol{\eta}(\mathbf{c})$ are irrational relative to each other. In this case you'll get dense orbits in the torus, corresponding to lines with irrational slope in $\mathbb{R}^{2n}$ before quotienting by the lattice Γ_{x_0}, and there will not be n integrals of motion.

Kolmogorov-Arnold-Moser (KAM) theory. More generally, if one doesn't have complete integrability, one can make weaker but still interesting statements. For example, one can envision a problem which is completely integrable in the absence of perturbations, and one can study what happens when the dependence on $\boldsymbol{\varphi}$ is small:

$$H(\boldsymbol{\varphi}, \mathbf{I}) = H_0(\mathbf{I}) + \varepsilon H(\boldsymbol{\varphi}, \mathbf{I}).$$

Some systems will lose integrability, though understanding the precise ways they do so is very hard. Such a system is associated to a *frequency vector* $\boldsymbol{\eta}_0 := \boldsymbol{\eta}(\mathbf{I}(t_0))$ satisfying the *Diophantine condition*

$$|\langle \boldsymbol{\eta}_0, \mathbf{n}\rangle| \geq \frac{1}{\langle \mathbf{n}\rangle^\tau}$$

for all $n \in \mathbb{Z}$ for some $\tau > 0$. Here $\langle \mathbf{x}\rangle := \sqrt{1 + |\mathbf{x}|^2}$ is the *Japanese bracket*. This quantitatively captures the qualitative idea that "$\boldsymbol{\eta}_0$ is poorly approximated by rationals."

In this setup, there exists an invariant torus under the flow of H. The proof involves renormalization group flow, though it was not originally discovered in those terms. It's a kind of recursive proof style, and getting into the details would take a long time. It involves a great result called the *shadowing lemma*, which discusses the dynamics of a pendulum.

The pendulum has two equilibria: the bottom is stable ($\varphi = 0$), and the top is unstable (both with no velocity). The phase space is two-dimensional, in φ and $\dot{\varphi}$, and some trajectories are shown in Figure 1. The curves with singularities are called *separatrices*.

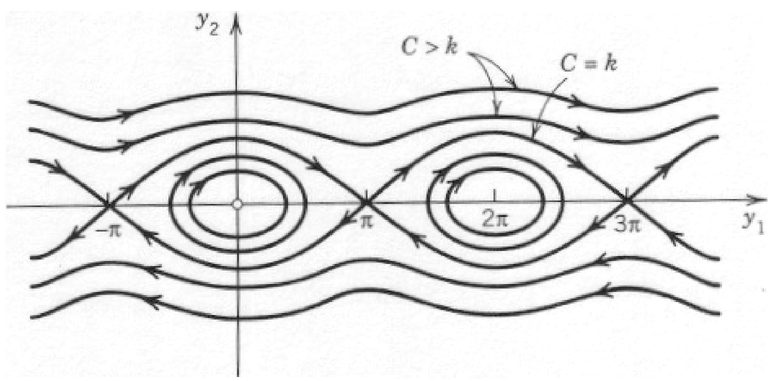

FIGURE 1. The phase diagram of a pendulum. Source: https://physics.stackexchange.com/q/162577.

.

Given a sequence of 0s and 1s, one may construct a parametric perturbation of the pendulum, regularly bumping it a small amount based on whether 0 or 1 is present.[2]. The shadowing lemma states that these trajectories uniformly approximate real trajectories. There's a rich theory here: the proof is a fixed-point argument, and there's interesting geometry of the *homoclinic points*, where two trajectories meet. These tend to be concentrated near the unstable equilibrium.

Quantum mechanics. Though quantum mechanics was discovered later than classical mechanics, it's actually much more fundamental. This suggests that one can derive classical mechanics as some sort of limit of quantum mechanics where Planck's constant is small, and indeed we can do this. We'll do this in three ways.

(1) The first is to use the Weiner transform to derive the Liouville equations from quantum mechanics in a semiclassical limit.

[2] TODO: did I get this right?

(2) The second case is to use a path integral to rediscover the principle of least action.
(3) The third way is to use observables and something called the Ehrenfest theorem.

Schrödinger discovered the *Schrödinger equation*, one of the cornerstones of quantum mechanics:

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

where $\psi(t,x) \in L^2$ and

$$\|\psi\|_{L^2}^2 = \int |\psi(t,x)|^2 \, \mathrm{d}x = 1,$$

Schrödinger arrived at this equation by (somewhat heuristically) studying quantization. Electrons had been observed (by de Broglie) to sometimes behave as particles and sometimes behave as waves. If an electron behaves like a particle, it has momentum $\hbar k$, where k is something called a *wave vector*. If you look at it as a wave, you get something like $i\hbar\nabla e^{-ikx}$, where $P := i\hbar\nabla$ is called the *momentum operator*. The Schrödinger equation (a guess within his PhD thesis) replaced the true momentum in the Hamiltonian

$$H(x,p) = \frac{1}{2m}p^2 + V(x)$$

with the momentum operator $i\hbar\nabla$, giving is $-\hbar^2\Delta$.

Lecture 4.

The Schrödinger equation and the Wigner transform: 9/12/17

Today we're going to begin by asking, how does one derive (well, guess) the Schrödinger equation? This involves an interesting and relevant digression on the Hamilton-Jacobi equation.

From the principle of least action, we know the Euler-Lagrange equations (1.1). Assume $q_0(t)$ is a solution to these equations. Take a one-parameter variation (s, q_s) from (t_0, q_0) to (t, q). The *Hamilton principal function* is

$$S(t,q) = \int_{(t_0,q_0)}^{(t,q)} L(q(s), \dot{q}(s)) \, \mathrm{d}s.$$

The variation with respect to q is

$$\begin{aligned}\delta S &= \int_{t_0}^{t} \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q}\right) \mathrm{d}s \\ &= \int_{t_0}^{t} \partial_s\left(\frac{\partial L}{\partial \dot{q}}\delta q\right) \mathrm{d}s \\ &= \frac{\partial L}{\partial \dot{q}}\delta q\bigg|_{t_0}^{t}.\end{aligned}$$

Since $p = \frac{\partial L}{\partial \dot{q}}$ and $\delta q(t_0) = 0$, this is

$$= (p\delta q)(t).$$

Hence, $p = \frac{\partial S}{\partial q}$ and

$$L = \frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\partial S}{\partial t} + \sum_j \frac{\partial S}{\partial q_j}\dot{q}_j,$$

so

$$\begin{aligned}\frac{\partial S}{\partial t} &= L - \sum_j p_j\dot{q}_j \\ &= -H(q, \nabla_q S).\end{aligned}$$

This is called the *Hamilton-Jacobi equation*.

The link with the Schrödinger equation: let's take for an ansatz that we have a wavefunction

$$\psi(t,x) = a(t,x)e^{-iS(t,x)/\hbar}.$$

This does not come entirely out of left field: if you want to exponentiate the action, you have to make it dimensionless, and that's exactly what dividing by $\hbar$ accomplishes. Then,

$$\begin{aligned} i\hbar\partial_t\psi &= i\hbar\dot{a}e^{-iS/\hbar} + \frac{\hbar}{\hbar}\frac{\partial S}{\partial t}ae^{-iS/\hbar}. \\ &= -H(q, \nabla S)\psi + O(\hbar) \\ &= \left(-\frac{1}{2}(\nabla S)^2 + V(x)\right)\psi + O(\hbar). \end{aligned}$$

Compare with

$$\begin{aligned} -\frac{\hbar^2}{2}\Delta a e^{-iS/\hbar} &= -\frac{\hbar^2}{2}\left(-\frac{i}{\hbar}\Delta S + \left(\frac{i\nabla S}{\hbar}\right)^2\right)ae^{-iS./\hbar} + O(\hbar) \\ &= \frac{1}{2}(\nabla S^2 a e^{-iS/\hbar} + O(\hbar). \end{aligned}$$

Putting these together, we arrive at

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

That is, the Schrödinger equation is an $O(\hbar)$-deformation of the Hamilton-Jacobi equations.

We'd like to solve this equation. Precisely, given a $\psi_0 \in L^2(\mathbb{R}^n)$, we'd like to find ψ such that

$$\begin{aligned} i\partial_t\psi &= -\Delta\psi + V(x)\psi = H\psi \\ \psi(t=0) &= \psi_0. \end{aligned} \tag{4.1}$$

Here H is the Hamiltonian.

We'd like to apply spectral theory to solve this, but $-\Delta$ is unbounded, with the domain

$$\{f \in L^2 \mid \|-\Delta f\|_{L^2} < \infty\},$$

which is dense in L^2. It is self-adjoint, in the formal sense, but because it (and pretty much every operator in quantum mechanics) is unbounded, the analysis is trickier. For the moment, we'll consider a regularized Hamiltonian.

Recall that we have a Fourier transform $\mathcal{F}\colon L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ given by

$$\begin{aligned} \widehat{f}(\xi) &= \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} f(x)e^{-i\xi\cdot x}\,\mathrm{d}x \\ \check{g}(x) &= \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} g(\xi)e^{i\xi\cdot x}\,\mathrm{d}\xi. \end{aligned}$$

Here, $g \mapsto \check{g}$ is the inverse Fourier transform. This was defined on Schwartz-class functions by the formulas above, then using the Plancherel theorem and the density of Schwartz functions in L^2, it extends to L^2. The Laplacian turns into multiplication under the Fourier transform:

$$\mathcal{F}(-\Delta f)(\xi) = \xi^2\widehat{f}(\xi).$$

Now we will regularize the Laplacian: define

$$\mathcal{F}(-\Delta_R f)(\xi) := \xi^2\chi_R(|\xi|)\widehat{f}(\xi),$$

where $R \gg 1$ and χ_R is a smooth bump function equal to 1 on $[0, R]$ and 0 on $[2R, \infty)$. Hence, for any finite R, Plancherel's theorem allows us to calculate that

$$\|-\Delta_R f\| \le (2R)^2,$$

where we use the operator norm. If we assume that $V \in L^\infty(\mathbb{R}^n)$, then

$$\|V(x)\psi\|_{L^2} \le \|V\|_{L^\infty}\|\psi\|_{L^2},$$

so the regularized Hamiltonian

$$H_R := -\Delta_R + V$$

is bounded.

Definition 4.2. Let A be an operator on L^2, possibly unbounded. We define the adjoint operator A^* to satisfy $(\phi, A\psi) = (A^*\phi, \psi)$ for all $\phi, \psi \in L^2$. A is *symmetric* if $(\phi, A\psi) = (A\phi, \psi)$ for all ϕ, ψ in the domain of A; if A and A^* have the same domain, this implies $A = A^*$, and A is called *self-adjoint*.

Theorem 4.3. *If A is bounded, then symmetric implies self-adjoint.*

Theorem 4.4. *If A is a bounded, self-adjoint operator, then there is an L^2 solution to*

$$\begin{aligned} i\partial_t \psi &= -\Delta\psi + V(x)\psi = A\psi \\ \psi(t=0) &= \psi_0, \end{aligned} \tag{4.5}$$

where $\psi_0 \in L^2$, which is given by

$$\psi(t) = e^{-itA}\psi_0.$$

Here,

$$e^A := \sum_{j=0}^{\infty} \frac{A^j}{j!}. \tag{4.6}$$

The particular case e^{-itA} is really nice: it's an isometry, because

$$\|e^{itA}\psi_0\|_{L^2} = \|\psi_0\|_{L^2},$$

and it's unitary:

$$(e^{itA})^* = e^{-itA} = (e^{itA})^{-1}.$$

Exercise 4.7. Check that the infinite sum in (4.6) converges, so that e^A is well-defined, and $\|e^{itA}\| \leq e^{t\|A\|}$ for all t.

Now, what does this all mean physically? Quantum mechanics considers a particle whose position and velocity at time t are probabilistically given by some probability density $\psi(t, x)$, such that

$$\|\psi(t)\|_{L^2} = \|\psi_0\|_{L^2} = 1.$$

Measuring physical facts about this system is expressed through *observables*, self-adjoint operators $A \colon L^2 \to L^2$: the expected value of A with respect to the distribution $\psi(t, x)$ is

$$\langle A \rangle_{\psi(t)} := \int \overline{\psi}(t, x)(A\psi)(t, x)\,\mathrm{d}x = (\psi, A\psi).$$

Because this system satisfies the Schrödinger equation (4.1), there are several conserved quantities. Consider

$$\begin{aligned} \partial_t(\psi, H\psi) &= \left(\frac{1}{i}H\psi, H\psi\right) + \left(\psi, H\left(\frac{1}{i}H\psi\right)\right) \\ &= -\left(H\psi, \frac{1}{i}H\psi\right) + \left(H\psi, \frac{1}{i}H\psi\right). \end{aligned}$$

In our case, we'd use H_R instead of H. The *energy* of the system is

$$E[\psi] := \frac{1}{2}(\psi, H\psi),$$

and by the above, this is a conserved quantity. The L^2 *mass* is also conserved:

$$M[\psi] := \|\psi\|_{L^2}^2.$$

The Wigner transform. We'll now discuss the Wigner transform, a noncommutative version of the Fourier transform. As is customary with the Fourier transform and related phenomena, we will be cavalier about factors of 2π arising from the transform; if you don't like this, it's possible to avoid with the harmonic analysts' convention

$$\widehat{f}(\xi) = \int f(x) e^{-2\pi i \xi \cdot x} \, \mathrm{d}x,$$

where making these factors precise is easier. We'll also ignore some factors of $\hbar$.

Consider the function

$$\widehat{\rho}(t, \xi) := \langle e^{ix\cdot\xi} \rangle_{\psi(t)} = \int \underbrace{|\psi(t,x)|^2}_{\rho(t,x)} e^{-ix\cdot\psi} \, \mathrm{d}\xi,$$

so that $\rho(t, x)$ is a probability distribution in x for a given t. The *momentum operator* $P = i\nabla_x$, on the other hand, satisfies

$$\langle P \rangle_{\psi(t)} = \int |\widehat{\psi}(t,\xi)| \xi \, \mathrm{d}\xi,$$

and hence defines another natural probability density $\mu(t, \xi)$ via

$$\langle e^{-iP\eta} \rangle_{\psi(t)} = \int \underbrace{|\widehat{\psi}(t,\xi)|^2}_{\mu(t,\xi)} e^{-i\xi\cdot\eta} \, \mathrm{d}\xi = \widehat{\mu}(t, \eta).$$

The two probability distributions $\widehat{\rho}$ and μ ought to be related, but they're not Fourier transforms from each other. Maybe in quantum mechanics, it doesn't make sense to separate the densities in x (position) and ξ (momentum), and to instead consider a probability density on the entirety of phase space of a solution ψ to (4.1). In particular, let

$$\widehat{W}(t, \xi, \eta) := \left\langle e^{-i(x\cdot\xi + P\cdot\eta)} \right\rangle_{\psi(t)}.$$

Here x and P do not commute. Accordingly, the *Wigner transform* of ψ is

$$W(t, x, v) := (\widehat{W})^\vee (t, x, v). \tag{4.8}$$

In the semiclassical limit, as $\hbar \to 0$, this will converge to the Liouville equation as in classical mechanics.

Remark. For a general solution ψ of the Schrödinger equation, its Wigner transform is not positive definite, and hence doesn't define a probability density. However, we can make it positive definite: if

$$G(x, v) = e^{-c_1 x^2 - c_2 v^2}$$

is a Gaussian, then the convolution

$$H(t, x, v) := (W * G)(t, x, v)$$

is positive definite, and, suitably normalized, it defines a probability density function. The function H is called a *Husini function*, and is very useful in applied math, specifically in the study of wave equations. ◂

The definition (4.8) is great for telling us what and why the Wigner transform is, but not so much how to calculate anything with it. Fortunately, there's an explicit formula.

Lemma 4.9.

$$W(t, x, v) = \int \overline{\psi(t, x - y/2)} \psi(t, x + y/2) e^{iy\cdot v} \, \mathrm{d}y.$$

This can be simplified using the density matrix $\Gamma_{xx'} := \overline{\psi}(x)\psi(x')$. So the Wigner transform is the Fourier transform of a density matrix.

Proof. The proof is not fascinating, but will be good practice for a useful technique.

Let A and B be linear operators for which e^A and e^B are well-defined, and assume $[A, B] := AB - BA$ is a scalar multiple of the identity. Then the higher commutators all vanish: $[A, [A, B]] = [[A, B], B] = 0$. Hence, the *Baker-Campbell-Hausdorff* formula for e^{A+B} simplifies greatly to

$$e^{A+B} = e^A e^B e^{-[A,B]/2}. \tag{4.10}$$

We're specifically interested in x_i and P_j, and $[x_i, P_j] = -i\delta_{ij}$, so we may use (4.10):

$$e^{-i(x\cdot\xi + P\cdot\eta)} = e^{-ix\cdot\xi} e^{-iP\cdot\eta} e^{-\xi\cdot\eta/2}.$$

Next, observe that $e^{-iP\cdot\eta}$ acts through a translation by η:

$$\begin{aligned}\left(e^{-iP\cdot\eta}f\right)(x) &= e^{\eta\cdot\nabla}\int \widehat{f}(\xi)e^{i\xi\cdot x}\,\mathrm{d}\xi\\ &= \int \widehat{f}(\xi)e^{i(x+\eta)\cdot\xi}\,\mathrm{d}\xi\\ &= f(x+\eta).\end{aligned}$$

Therefore

$$\widehat{W}(t,\xi,\eta) = \int e^{-ix\xi}e^{-(i/2)\xi\cdot\eta}\overline{\psi(t,x)}\psi(t,x+\eta)\,\mathrm{d}x.$$

If you compute the inverse Fourier transform, which is mechanical, you'll get the desired formula. ⊠

Convergence to the classical Liouville equation. Taking a semiclassical limit means sending $\hbar$ to 0, more or less. Of course, this makes no sense: $\hbar$ is a nonzero physical constant! But it represents the idea that, relative to the scale of $\hbar$, everything is very large. Also, we'll call it ε instead of $\hbar$, which makes it better.

Our Schrödinger equation is, given a potential $V \in C^2(\mathbb{R}^n)$,

$$i\varepsilon\partial_t\psi^\varepsilon = -\frac{\varepsilon^2}{2}\Delta\psi^\varepsilon + V\psi^\varepsilon.$$

Now, the rescaled Wigner transform is

$$W^\varepsilon(t,x,p) = \frac{1}{\varepsilon^n}\int \overline{\psi^\varepsilon(t,x-y/2)}\psi^\varepsilon(t,x+y/2)e^{i(y\cdot P)/\varepsilon}\,\mathrm{d}y.$$

Scaling $y \to \varepsilon y$, this is

$$= \int \overline{\psi^\varepsilon(t,x-\varepsilon y/2)}\psi^\varepsilon(t,x+\varepsilon y/2)e^{iy\cdot P}\,\mathrm{d}y. \tag{4.11}$$

Exercise 4.12. Show that $\partial_t W^\varepsilon(t,x,p)$ is the sum of a *kinetic term* (I) and a *potential term* (II) where

$$(I) = -p\cdot\nabla_x W^\varepsilon(t,x,p) \tag{4.13a}$$

$$(II) = (\text{didn't get this in time}) \tag{4.13b}$$

The Wigner transform has the property that it turns a Schrödinger-like equation into a transport equation, and vice versa.

Lecture 5.

The semiclassical limit of the Schrödinger equation: 9/14/17

"Evaluating an object like (5.11) *looks like it can be damaging to one's health. But it can be done"*

We've been working on the Schrödinger equation

$$\begin{aligned}i\varepsilon\partial_t\psi^\varepsilon(t,x) &= -\frac{\varepsilon^2}{2}\Delta\psi^\varepsilon(t,x) + V(x)\psi^\varepsilon(t,x)\\ \psi^\varepsilon(t=0) &= \psi_0^\varepsilon.\end{aligned}$$

Here, $V \in C^2(\mathbb{R}^n)$, and $\varepsilon = \hbar$, because it seems much more reasonable to say $\varepsilon \to 0$ rather than $\hbar \to 0$ (since $\hbar$ is a physical constant, not a variable!), as we will do when considering its semiclassical limit.[3]

To compute this, we introduced the rescaled Wigner transform Equation (4.11).

Theorem 5.1. *As $\varepsilon \to 0$, $W^\varepsilon \to F$, where*

$$(\partial_t + p\cdot\nabla_X)F(t,x,p) = (\nabla V)(x)\cdot\nabla_p F(t,x,p).$$

[3]For this reason, ε is sometimes known as a *semiclassical parameter*.

Proof. As in Exercise 4.12, we want to write $\partial_t W^\varepsilon(t,x,p)$ as a sum of a kinetic term (I) (4.13a) and a potential term (II) (4.13b). In more detail, if

$$\begin{aligned}(I) &= \frac{i\varepsilon}{2}\int\left(\overline{\psi^\varepsilon\left(t,x-\frac{\varepsilon y}{2}\right)}\Delta\psi^\varepsilon\left(t,x+\frac{\varepsilon y}{2}\right)-\overline{\Delta\psi^\varepsilon\left(t,x-\frac{\varepsilon y}{2}\right)}\psi^\varepsilon\left(t,x+\frac{\varepsilon y}{2}\right)\right)e^{ipy}\,\mathrm{d}y\\ &= i\int\nabla_x\cdot\nabla_y\left(\overline{\psi^\varepsilon\left(t,x-\frac{\varepsilon y}{2}\right)}\psi^\varepsilon\left(t,x+\frac{\varepsilon y}{2}\right)\right)e^{ipy}\,\mathrm{d}y.\end{aligned}$$

Then, the cross terms cancel, which is how you get (4.13a).[4]

The other term is

$$(II) = -\frac{i}{\varepsilon}\int\left(V\left(x+\frac{\varepsilon y}{2}\right)-V\left(x-\frac{\varepsilon y}{2}\right)\right)\overline{\psi^\varepsilon\left(t,x-\frac{\varepsilon y}{2}\right)}\psi^\varepsilon\left(t,x+\frac{\varepsilon y}{2}\right)e^{ipy}\,\mathrm{d}y.$$

For some $s_y\in(-1,1)$, this is

$$= -\frac{i}{\varepsilon}\int\left(\varepsilon\nabla V(x)\cdot y+\frac{1}{2}D^2V\left(x+s_y\frac{\varepsilon y}{2}\right)(\varepsilon y,\varepsilon y)\right)\overline{\psi^\varepsilon(\cdots)}\psi^\varepsilon(\cdots)e^{ipy}\,\mathrm{d}y.$$

Splitting this along the red $+$ sign, call the first part (II_1) and the second (II_2). The first term is what we want, and the second is an error term.

$$\begin{aligned}(II_1) &= -i\int\nabla V(x)\frac{1}{i}\nabla_p\overline{\psi^\varepsilon\left(x-\frac{\varepsilon y}{2}\right)}\psi^\varepsilon\left(x+\frac{\varepsilon y}{2}\right)e^{ipy}\,\mathrm{d}y\\ &= \nabla V(x)\cdot\nabla pW^\varepsilon(t,x,p).\end{aligned}$$

We'd like the error term to go away, but because y is unbounded (this integral is over $\mathbb{R}^n$) we need to make some assumptions. Let's assume $\operatorname{supp}(V)\subset B_R(0)$ is bounded. Then,

$$\left|x+\frac{\varepsilon y}{2}\right|<R,$$

so

$$|y|\le\frac{2}{\varepsilon}(R+|x|).$$

This is not strong enough: it's asymptotic to $1/\varepsilon$, which does not go away (rather the opposite, in fact).

Instead, we'll have to show that (II_2) converges *weakly* to 0, even when V isn't compactly supported. Let $J(x,p)\in\mathcal{S}(\mathbb{R}^n\times\mathbb{R}^n)$ be a test function (Schwartz class), and recall that the Fourier transform sends Schwartz-class functions to Schwartz-class ones. This implies that for all m, n, r, and s,

$$\|x^m\nabla_x^n p^r\nabla_p^s J\|_{L^\infty}<C_{m,n,r,s}.$$

For any such J, its Fourier transform, also Schwartz class, satisfies

$$|\widehat{J}(x,y)|\le\frac{f(x)}{(R^2+y^2)^{m/2}} \tag{5.2}$$

for some R, where $f(x)\to 0$ rapidly as $|x|\to\infty$. Hence, when we integrate,

$$\begin{aligned}&\frac{1}{\varepsilon}\int J(x,p)D^2V(\varepsilon y,\varepsilon y)\overline{\psi^\varepsilon(\cdots)}\psi^\varepsilon(\cdots)e^{ipy}\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}p\\ &\qquad= \frac{1}{\varepsilon}\int\widehat{J}(x,y)D^2V(\varepsilon y,\varepsilon y)\overline{\psi^\varepsilon(\cdots)}\psi^\varepsilon(\cdots)\,\mathrm{d}y\,\mathrm{d}x.\end{aligned}$$

Using (5.2),

$$|(\text{above})|\le\frac{1}{\varepsilon}\int|f(x)|\|D^2V\|_{L^\infty}|\varepsilon y^2|\frac{1}{(R^2+y^2)^m}\left|\overline{\psi^\varepsilon\left(t,x-\frac{\varepsilon y}{2}\right)}\right|\left|\psi^\varepsilon\left(t,x+\frac{\varepsilon y}{2}\right)\right|\,\mathrm{d}x\,\mathrm{d}y.$$

Since V is C^2, $\||f|\cdot\|D^2V\|_{\text{matrix}}\|_{L^\infty}$ is bounded by some constant C. Here we need to assume D^2V grows at most polynomially in $|x|$ as $|x|\to\infty$ and that f is Schwartz. Then,

$$\le\frac{1}{\varepsilon}\int\frac{1}{(R^+y^2)^{n/2+1}}\,\mathrm{d}y\left(\int\left|\psi^\varepsilon\left(t,x-\frac{\varepsilon y}{2}\right)\right|^2\mathrm{d}x\right)^{1/2}\left(\int\left|\psi^\varepsilon\left(t,x+\frac{\varepsilon y}{2}\right)\right|^2\mathrm{d}x\right)^{1/2}.$$

[4] TODO: I think… I didn't get this written down in time.

Each of the L^2 terms is $O(\varepsilon)$, and therefore the whole thing goes as $\varepsilon^2/\varepsilon$, hence $O(\varepsilon)$, which goes to 0 as $\varepsilon \to 0$. ☒

Hence the semiclassical limit of the Schrödinger equation is the Liouville equation, as promised. We're lucky in a sense, because the semiclassical limit came purely by rescaling; in general, one has to be more clever.

Derivation of the principle of least action from the path integral. There's another way to pass from quantum to classical without doing anything so strange as letting $\hbar \to 0$ (so in particular, we can call it $\hbar$ again).

First, let's simplify by removing the $V\psi$ term, obtaining the *free Schrödinger equation*

$$H_0 = -\frac{\hbar^2}{2}\Delta,$$

whose solution with $\psi(t=0) = \psi_0$ is

$$\begin{aligned} \psi(t,x) &= \left(e^{-itH_0/\hbar}\psi_0\right)(x) \\ &= \left(\frac{1}{2\pi i\hbar t}\right)^{d/2} \int e^{-i|x-q_0|^2/(2\hbar t)}\psi_0(q_0)\,\mathrm{d}q_0. \end{aligned} \tag{5.3}$$

If it weren't for the i in the exponent, this would look like a Gaussian. To solve it, we're going to discretize time $[0,t]$ into N intervals of width $\Delta t := t/N$. Let $t_j := j\cdot\Delta t$ and q_j be the variable corresponding to t_j.

$$= \left(\frac{1}{2\pi i\hbar\Delta t}\right)^{dN/2} \int \prod_{j-0}^{N-1} e^{-i|q_{j+1}-q_j|^2/(2\hbar\Delta t)}\bigg|_{q_N=x} \psi_0(q_0)\prod_{j=0}^{N-1}\mathrm{d}q_j. \tag{5.4}$$

Thus, if you let $\mathbf{q}_N := (q_0,\dots,q_N)$ and

$$\mathcal{D}\mathbf{q}_N := \left(\frac{1}{\pi i\hbar\Delta t}\right)^{dN/2}\prod_{j=1}^{N-1}\mathrm{d}q_j,$$

which is a complex-valued measure, then (5.4) simplifies to

$$\int \exp\left(-\frac{i}{\hbar}S_{0,N}(t,\mathbf{q}_N,x)\right)\psi_0(q_0)\,\mathcal{D}\mathbf{q}_N. \tag{5.5}$$

Here $S_{0,N}$ is the discretization of the action:

$$S_{0,N} := \frac{1}{2}\sum_{j=0}^{N-1}\left(\frac{q_{j+1}-q_j}{\Delta t}\right)^2\Delta t. \tag{5.6}$$

As $\Delta t \to 0$, this converges to $(1/2)\int_0^t(\dot{q}(s))^2\,\mathrm{d}s$, and (5.5) resembles more and more the integral of $e^{iS/\hbar}$ over all paths connecting q_0 to x, integrated against $\psi(q_0)$ with respect to q_0. This is an example of a *path integral* (after all, it's an integral over paths).

For the full Schrödinger equation, with $V \neq 0$, the idea is the same, just with more variables per line. Again subdivide

$$[0,t] = \bigcup_{j=0}^{N_1}[t_j,t_{j+1}],$$

and discretize the classical action, like in (5.6) but with a potential.

$$S_N(t,\mathbf{q}_N,x) := \sum_{j=0}^{N-1}\left(\frac{1}{2}\left(\frac{q_{j+1}-q_j}{\Delta t}\right)^2 + V(q_j)\right)\Delta t. \tag{5.7}$$

Then, define

$$\Psi_N(t,x) := \int e^{-iS_N(t,\mathbf{q}_N,x)/\hbar}\underbrace{\left[\frac{\det(\partial_t^2)}{\det(\partial_t^2 + D^2V)}\right]}_{(*)}\psi_0(q_0)\,\mathcal{D}\mathbf{q}_N. \tag{5.8}$$

The quantity in (*) is called the *Van Vleck-Pauli-Morette determinant*, which is the correction to (5.5) dictated by the potential.

Theorem 5.9. *If $\Psi(t,x) := \lim_{N\to\infty}\Psi_N(t,x)$, then ψ is a strong L^2 solution to the Schrödinger equation with $\Psi(t=0)=\Psi_0$.*

Partial proof. Here s-lim denotes a strong limit. If A and B are two matrices which do not necessarily commute, *Trotter's product formula* establishes that

$$e^{A+B} = \lim_{N\to\infty}\left(e^{(1/N)A}e^{(1/N)B}\right)^N. \tag{5.10}$$

In particular, H_0 and V don't necessarily commute, so if $\Delta t := t/N$,

$$\exp\left(-it\frac{H_0+V}{\hbar}\right) = \operatorname*{s-lim}_{N\to\infty}\left(\exp -i\frac{\Delta t H_0}{\hbar}\exp\left(-i\frac{\Delta t V}{\hbar}\right)\right)^N.$$

Implicit in this composition of operators is a kernel transform.[5] Therefore

$$e^{-it\frac{H_0+V}{\hbar}}(x,q_0) = \operatorname*{s-lim}_{N\to\infty}\int\left(e^{-i\Delta t H_0/\hbar}\right)(x,q_{N-1})e^{-i\Delta t V(q_{N-1})/\hbar}\left(e^{-i\Delta t H_0/\hbar}\right)(q_{N-1},q_{N-2})\cdots e^{-i\Delta t V(q_0)/\hbar}\,\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}. \tag{5.11}$$

If you insert

$$\left(e^{-i\Delta t H_0/\hbar}\right)(q_{j+1},q_j) = \left(\frac{1}{2\pi i\hbar\Delta t}\right)^{d/2}e^{-i|q_{j+1}-q_j|^2/(2\hbar\Delta t)},$$

you get the desired expression for Ψ_N in (5.8), except for the VV-P-M determinant. Now we need to actually evaluate (5.11), which is a very oscillatory integral on a high-dimensional space. Fortunately, we can use a trick from harmonic analysis called the stationary phase formula to assist us.[6]

Theorem 5.12. *Assume Φ and f are C^2 functions on $\mathbb{R}^n$, and let y^* denote the unique solution to $\Delta\Phi(y)=0$. Assume $D^2\Phi(y^*)$ is nondegenerate; then*

$$\int e^{-i\lambda\Phi(y)}f(y)\,\mathrm{d}y = \left(\frac{2\pi i}{\lambda}\right)^{d/2}|\det D^2\Phi(y^*)|^{-1/2}e^{-i\pi\,\mathrm{sign}(D^2\Phi(y^*))/4}e^{-i\lambda\Phi(y^*)}f(y^*) + o\left(\left(\frac{1}{\lambda}\right)^{d/2}\right)$$

as $\lambda\to\infty$.

The cool idea is, since

$$e^{i\lambda cy} = \frac{1}{i\lambda c}\partial_y e^{icy},$$

you can use the regularity of f to trade for factors of $1/\lambda$: the more regular f is, the stronger convergence you can obtain.

Lecture 6.

The stationary phase approximation: 9/19/17

To recap, we wanted to solve the Schrödinger equation, and in order to do so took a kind of path integral: we discretized the action (5.7) and integrated over all (piecewise-linear) possible paths (5.8) between q_0 and $q_N = x$, the point where we wanted to evaluate the answer. These discretized paths are approximations $\mathbf{q}_N^*$ to the classical paths which solve the Euler-Lagrange equations, and one has that

$$\Psi_N(t,x) = \left(\frac{1}{2\pi i\hbar t}\right)^{d/2}\int\exp\left(\frac{-iS_N(t,\mathbf{q}_N^*,x)}{\hbar}\right)\left[\frac{\det(\partial_t^2)}{\det(\partial_t^2+D^2V)}\right]\psi_0(q_0)\,\mathrm{d}q_0. \tag{6.1}$$

We then used Trotter's product formula (5.10) to prove that this converges to solutions $\psi(t)$ strongly (Theorem 5.9).

[5] TODO: what is this explicitly referring to?

[6] For those of you who like topology and geometry, there's a geometric reformulation of this which is related to the Duistermaat-Heckman formula in symplectic geometry.

To prove (6.1), we have to use the *stationary phase formula*: that for $\lambda \gg 1$,

$$\int e^{i\lambda\Phi(x)} f(x)\,\mathrm{d}x = \left(\frac{1}{2\pi i\lambda}\right)^{d/2} e^{i\lambda\Phi(x^*)} e^{i\pi\,\mathrm{sign}(D^2\Phi(x^*))/4} \left(\frac{1}{\det D^2\Phi(x^*)}\right)^{1/2} f(x^*) + o\left(\left(\frac{1}{\lambda}\right)^{d/2}\right), \tag{6.2}$$

where x^* is the *stationary point*, i.e. the point where $\nabla\Phi(x^*) = 0$.

To use (6.2), we need to find the stationary point $\mathbf{q}_N^*$ of $S_N(t, \mathbf{q}_N, x)$, which must satisfy

$$\nabla_{\mathbf{q}_N^*}(t, \mathbf{q}_N^*, x) = 0. \tag{6.3}$$

The Hessian is

$$D^2 S_N(t, \mathbf{q}_N, x) = \frac{1}{\Delta t} \underbrace{\begin{pmatrix} 2\mathbf{1}_d & -\mathbf{1}_d & & & \\ -\mathbf{1}_d & 2\mathbf{1}_d & -\mathbf{1}_d & & \\ & -\mathbf{1}_d & \ddots & & \\ & & & \ddots & -\mathbf{1}_d \\ & & & -\mathbf{1}_d & 2\mathbf{1}_d \end{pmatrix}}_{M_{N,d}} + D^2_{\mathbf{q}} V(\mathbf{q}_N^*)\Delta t.$$

Now, (6.3) is equivalent to the equation

$$\frac{-q_{j+1}^* + 2q_j^* - q_{j-1}^*}{(\Delta t)^2} = -(\nabla_{q_j} V)(q_j^*)$$

for $j = 1, \ldots, N-1$, and this is precisely a discretization of the Newton equations

$$\ddot{q} = -\nabla V(q).$$

Remark. $1/(\Delta t)^2 M_{N,d}$ is a discretization of ∂^2. ◀

From the stationary phase equation,

$$\Psi_N(t, x) = \int K_N(t, \mathbf{q}_N^*, x)\psi_0(q_0)\,\mathrm{d}q_0 + \text{lower-order terms},$$

where

$$\begin{aligned} K_N(t, \mathbf{q}_N^*, x) &= \left(\frac{1}{2\pi i\hbar\Delta t}\right)^{Nd/2} (2\pi i\hbar)^{Nd/2} \left|\det\left(D^2_{\mathbf{q}_N} S_N(t, \mathbf{q}_N^*, x)\right)\right|^{-1/2} \exp\left(\frac{iS_N(t, \mathbf{q}_N^*, x)}{\hbar}\right) \\ &= \left|\det\left(\Delta t\, D_q^2 S_N(t, \mathbf{q}_N^*, x)\right)\right|^{-1/2} \exp\left(\frac{iS_N(t, \mathbf{q}_N^*, x)}{\hbar}\right) \\ &= \left|\det\left(M_{N,d} + D_q^2 V(\mathbf{q}_N^*)(\Delta t)^2\right)\right|^{-1/2} \exp\left(\frac{iS_N(t, \mathbf{q}_N^*, x)}{\hbar}\right), \end{aligned} \tag{6.4}$$

and we know what the Hessian is. We'll use a strange-looking trick to simplify this next: observe that

$$\left(\frac{1}{2\pi i\hbar}\right)^{d/2} = \left(\frac{1}{2\pi i\hbar t}\right)^{d/2} \exp\left(-\frac{i|x-x|}{2\hbar t}\right),$$

which can be interpreted as a free propagator of x with itself. This can be expressed as an action

$$= |\det(M_{N,d})|^{-1/2} \exp\left(\frac{i}{\hbar} S_{0,N}(t, x, x, \ldots, x)\right).$$

Plugging the ratio of these terms back into (6.4),

$$K_N(t, \mathbf{q}_N^*, x) = \left(\frac{1}{2\pi i\hbar t}\right)^{d/2} \int \left|\frac{\det((1/(\Delta t)^2)M_{N,d})}{\det((1/(\Delta t)^2)M_{N,d} + D^2 V)}\right| \exp\left(\frac{iS_N(t, \mathbf{q}_N^*, x)}{\hbar}\right).$$

This ratio of determinants is important — it's the discretization of the VV-P-M determinant that we alluded to last time.

Ehrenfest theorem. The Ehrenfest theorem is another link between the quantum and classical worlds.

Theorem 6.5. *Let $A(t)$ be a linear operator on L^2 and assume $\psi(t)$ is an L^2 solution of the Schrödinger equation, i.e.*

$$i\hbar\partial_t\psi = H\psi$$

and $\psi(t=0) = \psi_0$ for some specified ψ_0. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A(t)\rangle_{\psi(t)} = \frac{1}{i\hbar}\langle [H, A]\rangle_{\psi(t)} + \langle \partial_t A\rangle_{\psi(t)}.$$

(Recall that $\langle A\rangle_\psi = \langle \psi, A\psi\rangle$.) One special case of interest: let $A = x$ be a position variable. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x\rangle_{\psi(t)} = \langle [H, x]\rangle_{\psi(t)},$$

and

$$[H, X] = \left[-\frac{\hbar^2}{2}\Delta + V, x\right].$$

If you calculate it out, this commutator is the gradient, so

$$[H, x]f = -\hbar^2\nabla f = -i\hbar P f,$$

where $P := -i\hbar\Delta$ is the momentum operator.

If on the other hand you apply Theorem 6.5 to P, you get that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle P\rangle_{\psi(t)} = \frac{1}{i\hbar}\langle [H, P]\rangle_{\psi(t)},$$

and in a similar manner,

$$[H, P]f = [V, P]f = -i\hbar(\nabla V)\cdot f.$$

Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle P\rangle_{\psi(t)} = -\langle \nabla V\rangle_{\psi(t)}.$$

What this means is that the classical Hamiltonian equations hold, in the operator sense, in expectation, with respect to $\psi(t)$.

Spectral theory. We're going to spend the next several lectures on spectral theory. We've done some before in the prelim classes, but the operators that arise in quantum physics are not always compact, and so we'll need a more advanced theory.

Definition 6.6. Let A be a linear operator (possibly unbounded) on a Hilbert space $\mathcal{H}$. Its *spectrum* $\sigma(A)$ is the set of $\lambda \in \mathbb{C}$ such that $A - \lambda$ is noninvertible.[7] The *resolvent* $\rho(A) := \mathbb{C} \setminus \sigma(A)$.

The spectrum further subdivides into three types.

- The *point spectrum* $\sigma_p(A)$ is the subset of $\sigma(A)$ where $A - \lambda$ is not injective.
- The *continuous spectrum* $\sigma_c(A)$ is the subset of $\sigma(A)$ where $A - \lambda$ is injective, and the range of $A - \lambda$ is dense, but $(A - \lambda)^{-1}$ is not bounded.
- The *residual spectrum* is the subset of $\sigma(A)$ where the range of $A - \lambda$ is not dense.

Theorem 6.7. *These are all the spectral types: $\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$. Moreover, if A is self-adjoint, $\sigma_r(A) = \varnothing$ and $\sigma(A) \subset \mathbb{R}$.*

Definition 6.8. Assume $(A - \lambda)\psi = 0$ has a nonzero solution $\psi \in \mathcal{H}$. Then, λ is called an *eigenvalue* and ψ an *eigenvector*.

There are also cases that are "almost as good."

Definition 6.9. The sequence $\{\psi_n\} \in \mathcal{H}$ is called a *Weyl sequence* for A and λ if

(1) $\|\psi_n\|_{\mathcal{H}} = 1$,
(2) $\|(A - \lambda)\psi_n\|_{\mathcal{H}}$ goes to 0 as $n \to \infty$, and
(3) $\psi_n \rightharpoonup 0$ as $n \to \infty$.

---

[7] That is, if $\lambda \notin \sigma(A)$, $A - \lambda$ has not just an inverse, but a bounded inverse.

The last condition means that ψ_n *converges weakly* to 0, i.e. $(\phi, \psi_n) \to 0$ for all $\phi \in \mathcal{H}$.

Let $\sigma_d(A)$ denote the *discrete spectrum* of A, i.e. the set of isolated eigenvalues of A with finite multiplicity.

We won't prove these theorems, but a proof will be posted (either on Canvas or the course website).

Theorem 6.10 (Weyl criterion). *$\sigma_c(A)$ is the set of $\lambda \in \mathbb{C}$ for which there exists a Weyl sequence.*

Theorem 6.11. *If $U\colon \mathcal{H} \to \mathcal{H}$ is unitary, then $\sigma(U^*AU) = \sigma(A)$.*[8]

Proof. This follows because $U^*AU - \lambda = U^*(A-\lambda)U$, which is true because $U^*U = \mathbf{1}$, and the fact that U is an isometric isomorphism, hence preserves injectivity, surjectivity, and density. ⊠

Different authors use different conventions/definitions for these things, so be careful.

Lecture 7.

Spectral theory: 9/21/17

"Let me write this down in the hope that the errors today are new errors, not old ones."

We started with a correction of a derivation from the last lecture. I don't know where it fits in the notes, unfortunately.

$$\begin{aligned}\left(\frac{1}{2\pi i\hbar t}\right)^{d/2} &= \left(\frac{1}{2\pi i\hbar t}\right)^{d/2} \exp\left(-\frac{i|x-x|^2}{2\hbar t}\right)\\ &= \left(\frac{1}{2\pi i\hbar\Delta t}\right)^{dN/2} \int \exp\left(-\frac{i}{\hbar}S_{0,N}(t,\mathbf{q}_N,x)\right)\mathrm{d}\mathbf{q}_N.\end{aligned}$$

$\mathrm{d}\mathbf{q}_N$ is a product of $N-1$ integrands $\mathrm{d}q_j$, rather than N integrands.

$$= \left(\frac{1}{2\pi i\hbar\Delta t}\right)^{dN/2} (2\pi i\hbar)^{\frac{d(N-1)}{2}} \left|\det\left(\frac{1}{\Delta t}\right)M_{d,N}\right|^{-1/2},$$

and again, $M_{d,N}$ is an $(N-1)\times(N-1)$-matrix.

$$= \left(\frac{N}{2\pi i\hbar t}\right)^{d/2} \left|\det M_{d,N}\right|^{-1/2}.$$

As the determinant of the $(N-1)\times(N-1)$-matrix

$$A_{N-1} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & & & \\ & & \ddots & & \\ & & & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}$$

is N, the determinant of $M_{d,N} = A_{N-1} \otimes \mathbf{1}_d$ is N^d. The good news is, in the limit the answer is the same. (There was also a correction in the spectral theory part of the notes, which has already been incorporated.)

Example 7.1.

(1) Let $\mathcal{H} = L^2(\mathbb{R}^d)$ and for a monotone continuous function $g\colon \mathbb{R} \to \mathbb{R}$ with range $[-M, M]$, let A_g be the operator sending f to $A_g f(x) = g(x)f(x)$. Then, $\sigma(A_g) = [-m, M]$.

If $\lambda \notin \mathrm{Im}(g)$, then $(A_g - \lambda)^{-1} = 1/(g(x)-\lambda)$ is bounded.

Now assume $\lambda \in \mathrm{Im}(g)$; since g is monotone, $g^{-1}(\lambda)$ is either a point or an interval.

- Suppose $|g^{-1}(\lambda)| = 0$. Then, $A_g - \lambda$ is injective, because $((A_g-\lambda)f)(x) = 0$ iff $(g(x)-\lambda)f(x) = 0$ implies $f = 0$ almost everywhere, so $\lambda \in \sigma_c(A_g)$.
- Suppose $|g^{-1}(\lambda)| > 0$. Then, there functions $f \in L^2(\mathbb{R})$ not almost everywhere zero with $((A_g-\lambda)f)(x) = 0$, so $\lambda \in \sigma_p(A_g)$.

(2) Let A be multiplication by x acting on $L^2(\mathbb{R})$. Then, $\sigma(A) = \sigma_c(A) = \mathbb{R}$, essentially by the previous example. This is the spectrum of a position operator in quantum mechanics.

[8] TODO: this is also true for σ_p, σ_c, and σ_r, right?

(3) Let $P = -i\nabla$, which is the momentum operator, again in $d = 1$. If U denotes the Fourier transform, then $U^*PU = \xi$, so once again $\sigma(P) = \sigma_c(P) = \mathbb{R}$.
(4) Suppose $A\psi = -\Delta\psi$. Then, $U^*(-\Delta)U = \xi^2$, so we can just look at the spectrum of that, which is entirely the continuous spectrum, which is $\mathbb{R}_{\geq 0}$, so $\sigma(-\Delta) = \sigma_c(-\Delta) = \mathbb{R}_{\geq 0}$.
(5) If $A = -\Delta$ and $d > 1$, then

$$U^*(-\Delta)U = \sum_{j=1}^{d} \xi_j^2,$$

which has the same range, and therefore it's still true that $\sigma(-\Delta) = \sigma_d(-\Delta) = \mathbb{R}_{\geq 0}$. ◀

Theorem 7.2. *Suppose $H = -\Delta + V$ on $L^2(\mathbb{R}^d)$, where V is a continuous function such that $V(x) \to 0$ as $|x| \to \infty$. Then, $\sigma_{\mathrm{ess}}(H) = \mathbb{R}_{\geq 0}$.*

This is tricky, because these operators do not commute. The proof is a nice application of a variant of Weyl sequences.

Definition 7.3. Let A be a linear operator on $L^2(\mathbb{R}^d)$. A *spreading sequence* for A and λ is a sequence $\{\psi_n\}$ such that

- $\|\psi_n\| = 1$,
- for any bounded $B \subset \mathbb{R}^d$, there's an N_B such that $\operatorname{supp}(\psi_n) \cap B = \varnothing$ for $n > N_B$, and
- $\|(A - \lambda)\psi_n\| \to 0$ as $n \to \infty$.

Proposition 7.4. *A spreading sequence for A, λ is also a Weyl sequence.*

Proof of Theorem 7.2. Consider the sequence

$$\psi_n(x) := e^{i\sqrt{\lambda}x} \frac{1}{n^{d/2}} \phi\left(\frac{|x - 2n\operatorname{sign}(x)|}{n}\right),$$

where ϕ is a bump function with total integral 1 and supported on $(-1, 1)$. We're going to show this is a spreading sequence for H when $\lambda \geq 0$.

We have

$$|x - 2n\operatorname{sign}(x)| \leq n \iff n \leq |x| \leq 3n,$$

and therefore the support of ψ_n is unbounded as $n \to \infty$. It's also quick to check tht $\|\psi_n\|^2 = 1$. Finally, let's compute

$$\|(H - \lambda)\psi_n\| \leq \underbrace{\|(-\Delta - \lambda)\psi_n\|_{L^2}}_{(I)} + \underbrace{\|V\psi_n\|_{L^2}}_{(II)}. \tag{7.5}$$

Since the Fourier transform is norm-preserving,

$$(I) = \|(\xi^2 - \lambda)\widehat{\psi}_n\|,$$

and

$$\widehat{\psi}_n(\xi) = e^{2in|\xi|} \underbrace{n^{d/2}\widehat{\phi}\Big(n\Big(|\xi| - \lambda^{1/2}\Big)\Big)}_{\chi_n(\xi)}.$$

$|\chi_n|^2$ is concentrated around $|\xi| - \lambda^{1/2}$, and in fact converges weakly to a δ-function supported there, which means that for any test function g, as $n \to \infty$,

$$\int g(|\xi|)\chi_n^2(|\xi|)\,\mathrm{d}|\xi| \longrightarrow g(\lambda^{1/2}).$$

Therefore, assuming λ is in the image of ξ^2,

$$\int (\xi^2 - \lambda^2)|\chi_n(\xi)|^2\,\mathrm{d}\xi \longrightarrow 0.$$

The other piece of (7.5) also goes to 0:

$$
\begin{aligned}
(II)^2 &= \|V\psi_n\| \\
&= \int V(x)^2 |\psi_n(x)|^2 \, \mathrm{d}x \\
&\leq \sup_{|x|\in[n,3n]} |V(x)|^2 \int |\psi_n|^2, \qquad (7.6)
\end{aligned}
$$

and we know $\|\psi_n\| = 1$ and $V(x) \to 0$ as $|x| \to \infty$, so (7.6) does go to 0 as $n \to \infty$, and therefore for $\lambda \geq 0$, $\{\psi_n\}$ is a spreading sequence, hence a Weyl sequence by Proposition 7.4, and by Theorem 6.10, we're done. ⊠

Lecture 8.

The spectral theory of Schrödinger operators: 9/26/17

Note: I came in 20 minutes late and may have missed some material.
References: Reed-Simon, Hislop-Sigal, Yoshida, Kato.

Definition 8.1. T is *essentially self-adjoint* on $\mathcal{H}$ if its closure is self-adjoint.

Theorem 8.2. *Let T be a symmetric operator on $\mathcal{H}$. Then, the following are equivalent:*

(1) *T is essentially self-adjoint on $\mathcal{H}$.*
(2) $\ker(T^* \pm i) = \{0\}$.
(3) $\mathrm{Im}(T \pm i)$ *is dense in $\mathcal{H}$.*

Now we'll talk about the spectral theorem.

For motivation, consider an $n \times n$ matrix with complex entries. It has n eigenvalues $\sigma(A) = \{\lambda_j\}$, and assume that it has n linearly independent eigenvectors v_j, so that it may be diagonalized: let $T = (v_1 \ \ldots \ v_n)$ and

$$
\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.
$$

Then,

$$
A = T^{-1}\Lambda T = \sum \lambda_j P_j,
$$

where if E_j is the matrix with a 1 in position (j, j) and 0 elsewhere, $P_j := T^{-1}E_jT$.

The *operator norm* $\|A\|_{\mathrm{op}}$ is defined to be the supremum of the set of eigenvalues of A. If $A = A^*$ (i.e. it's *Hermitian*), then $T^{-1} = T^*$, i.e. it's *unitary*.

Now suppose f is a function with a convergent power series expansion $f = \sum a_n x^n$. For a matrix A we can write

$$
\begin{aligned}
f(A) &= \sum a_n A^n \\
&= T^{-1}\left(\sum a_n \Lambda^n\right) T \\
&= T^{-1} \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} T \\
&= \sum_j f(\lambda_j) P_j.
\end{aligned}
$$

If Γ is a contour enclosing $\sigma(A)$, we can also write this as

$$
f(A) = \frac{1}{2\pi i} \sum_j \oint_\Gamma \mathrm{d}z \, \frac{f(z)}{\lambda_j - z} P_j,
$$

or, if f is analytic,

$$= \frac{1}{2\pi i} \oint_\Gamma \mathrm{d}z \, \frac{f(z)}{A - z}.$$

Now we generalize to infinite-dimensional Hilbert spaces.

Theorem 8.3 (Spectral theorem for bounded Hermitian operators)**.** *Let A be a bounded Hermitian operator on a Hilbert space $\mathcal{H}$, Ω be a complex domain containing $\sigma(A)$, and $f\colon \mathbb{C} \to \mathbb{C}$ be analytic in Ω. If $\Gamma$is a contour in Ω encircling $\sigma(A)$, then*

$$f(A) = \frac{1}{2\pi i} \oint_\Gamma f(z)(A - z)^{-1} \, \mathrm{d}z.$$

Since $A = A^*$, $\sigma(A) \subset \mathbb{R}$; since A is bounded, so is its spectrum, and therefore the picture makes sense.

This integral may be understood in the following way: the numbers $(\psi, f(A)\varphi)$ over all $\psi, \varphi \in \mathcal{H}$ determine $f(A)$ uniquely, and

$$(\psi, f(A)\varphi) = \frac{1}{2\pi i} \oint_\Gamma f(z)(\psi, (A - z)^{-1}\varphi),$$

and the inner product is analytic in z in a neighborhood of Γ.

In quantum mechanics, we need to also understand unbounded operators. In this case, the spectrum is real, but may be unbounded, and the idea is to consider the contour that's the boundary of the rectangle $[-1/\varepsilon, 1/\varepsilon] \times [\varepsilon, \varepsilon]$, and let $\varepsilon \searrow 0$.

Theorem 8.4 (Spectral theorem for unbounded, Hermitian operators)**.** *Let A be an* unbounded *Hermitian operator on $\mathcal{H}$ and $f\colon \mathbb{R} \to \mathbb{R}$ be a Borel function.*[9] *Then,*

$$f(A) = \frac{1}{2\pi i} \lim_{\varepsilon \searrow 0} \operatorname{Im} \int_{-\infty}^{\infty} f(\lambda)(A - \lambda - i\varepsilon)^{-1} \, \mathrm{d}\lambda.$$

The proof is long and not terribly instructive, so we won't go into it. Instead, we'll focus specifically on Schrödinger operators.

Theorem 8.5. *Let $H = -\Delta + V$, where $V\colon \mathbb{R}^d \to \mathbb{R}$ is continuous, $V \geq 0$, and $V(x) \to \infty$ as $|x| \to \infty$. Then,*

(1) *H is self-adjoint on $L^2(\mathbb{R}^d)$,*
(2) *$\sigma(H) = \sigma_d(H)$ is the set $\{\lambda_j\}$ of eigenvalues, and*
(3) *$\lambda_j \to \infty$ as $j \to \infty$.*

Partial proof. For self-adjointness, see Hislop-Segal. We next show there does not exist a spreading sequence for any λ: assume $\{\psi_n\}$ is such a sequence; then, for any λ in the essential spectrum, $(\psi_n, (H - \lambda)\psi_n) \to 0$. And this is

$$\begin{aligned}
(\psi_n, (H - \lambda)\psi_n) &= (\psi_n, (-\Delta)\psi_n) + (\psi_n, V\psi_n) - \lambda \\
&= \int |\nabla \psi_n|^2 + \int V|\psi_n|^2 - \lambda \\
&\geq \inf_{y \in \operatorname{supp}(\psi_n)} (V(y)) - \lambda,
\end{aligned}$$

and this goes to ∞, since $\{\psi_n\}$ is a spreading sequence, providing a contradiction.

This means the essential spectrum is empty, so $\sigma(H) = \sigma_d(H)$, which is exactly the isolated eigenvalues.

To get at the limit of the eigenvalues, we'll use a variational characterization of the eigenvalues of an operator.

Theorem 8.6. *Let $\mathcal{H}_h := \operatorname{span}\{\psi_1, \dots, \psi_n\}$, where ψ_i is an eigenvector for the i^{th} lowest eigenvalue λ_i (so $\lambda_1 \leq \lambda_2 \leq \dots$). Then,*

$$\inf_{\{\psi \in \mathcal{H}_n^\perp \cap D(H) \mid \|\psi\| = 1\}} (\psi, H\psi) = \inf\{\sigma(H) \setminus \{\lambda_1, \dots, \lambda_n\}\}.$$

This is true because H is unbounded and its eigenvalues do not accumulate (because there is no essential spectrum). Repeatedly invoking Theorem 8.6, one gets that there's always another eigenvalue λ_{i+1}, and it's at least as big as λ_i, but the eigenvalues cannot accumulate, so they go to ∞. ⊠

[9] This means for every $I \subset \mathbb{R}$ open, $f^{-1}(I)$ is a Borel set.

Not every Schrödinger operator meets the criteria of (8.5), though, including some famous ones.

Example 8.7 (The hydrogen atom)**.** Consider the *Coulomb potential* $V(x) = 1/|x|$ on $\mathbb{R}^3$, which goes to 0 as $|x| \to \infty$, and the Hamiltonian

$$H = -\Delta - \frac{1}{|x|}.$$

Then, the essential spectrum of H is $[0, \infty)$. There are infinitely many eigenvalues below 0, though, and we'll shw this by constructing a sequence $\{u_n\}_{n\geq 1}$ of linearly independent functions with $(u_n, Hu_n) < 0$ for all n.

Pick a $u \in C_0^\infty(\mathbb{R}^3)$ such that $\|u\|_{L^2} = 1$ and $\operatorname{supp} u \subset x \in \mathbb{R}^3 \mid 1 < |x| < 2$. Then, let

$$u_n := 2^{-3n/2} u(2^{-n}x),$$

for $n \in \mathbb{N}$. Since $(u_n, u_m) = \delta_{nm}$, these are linearly independent. Moreover, $(u_n, Hu_m) = 0$ when $n \neq m$:

$$(u_n, Hu_m) = \int \overline{\nabla u_n}(\nabla u_m)\,\mathrm{d}x + \int V(x)\overline{u_n(x)}u_m(x)\,\mathrm{d}x,$$

but u_n and u_m have disjoint domains, so these integrals are both 0. If $m = n$, then we get

$$(u_n, Hu_m) = \int |\nabla u_n|^2 - \int \frac{1}{|x|}|u_n|^2 < 0. \qquad \blacktriangleleft$$

Lecture 9.

The Birman-Schwinger principle: 9/28/17

We've been studying the Schrödinger operator

$$H = -\Delta - \frac{1}{|x|},$$

which corresponds to a hydrogen atom, a single electron bound to a nucleus. We're assuming the nucleus is static and its mass is so large as to make the mass of the hydrogen atom negligible. Last time, we saw that the essential spectrum of H is $\mathbb{R}_+$, the eigenvalues are negative numbers, and there are infinitely many distinct eigenvalues. This implies there's an infinite-dimensional linear subspace on which H is negative.

One might ask, what aspect of this operator leads H to have infinitely many eigenvalues? For which values of $\alpha > 0$ does the operator

$$H = -\Delta - \frac{1}{|x|^\alpha}$$

have infinitely many eigenvalues?

Again we consider the function $u_n(x) := 2^{-3n/2}u(2^{-n}x)$, where $u \in C_0^\infty$ and $\operatorname{supp}(u) \subset \{|x| \mid 1 < |x| < 2\}$, so that $\|u_n\|_{L^2} = 1$.[10] Then,

$$\langle u_n, Hu_m \rangle = C\delta_{n,m},$$

because the support of the derivatives is also disjoint. Thus the kinetic term always vanishes. But the potential term may be nonzero, and is when $m = n$: we get

$$\begin{aligned}\langle u_n, Hu_n \rangle &= \int |\nabla u_n|^2 - \int \frac{1}{|x|^\alpha}|u_n|^2 \\ &= 2^{-2n}\int |\nabla u|^2 - 2^{-\alpha n}\int \frac{|u|^2}{|x|^\alpha}. \end{aligned} \tag{9.1}$$

If $\alpha < 2$, then for all n large enough, this is less than 0, because the second term dominates. This implies there are infinitely many eigenvalues. If $\alpha > 2$, then (9.1) is positive for all n sufficiently large. Does this mean we only have finitely many eigenvalues?

Another potential we might consider is $V(x) = 1/\langle x \rangle^\alpha$ (where $\langle x \rangle$ is the Japanese bracket). Then, $V \in L^{3/2}(\mathbb{R}^3)$. For $\alpha > 2$, do we only have finitely many eigenvalues?

The physical intuition comes from spectroscopic experiments: the hydrogen atom is in a state, and if it absorbs light of a certain energy (color), it can jump to a higher-energy state, and if it emits light of a

[10]One might say that the u_ns are *supported in dyadic shells* and have mutually disjoint supports.

certain energy (color), it falls to a lower-energy state. Every atom has a different potential and hence a different spectrum (of its Hamiltonian and observationally). The essential spectrum represents when the electron has been separated from the atom (ionization).

We'll use something called the *Birman-Schwinger principle* to solve this. Assume $H = -\Delta + V$, where $V < 0$, so $U(x) := -V(x) > 0$. For any $\lambda < 0$, $(-\Delta + V)\phi = \lambda\phi$ iff $(-\Delta - \lambda)\phi = U\phi$, so

$$\phi = (-\Delta - \lambda)^{-1} U\phi.$$

Since $U > 0$ then we can take a square root of it: let $v := U^{1/2}\phi$, so $v = K(\lambda)v$, where

$$K(\lambda) := U^{1/2}(-\Delta - \lambda)^{-1}U^{1/2}.$$

In particular, λ is an eigenvalue of H (for $\lambda < 0$) iff 1 is an eigenvalue of $K(\lambda)$. Therefore the number n_H of $\lambda < 0$ that are eigenvalues of H is the same as the number of $\lambda < 0$ such that 1 is an eigenvalue of $K(\lambda)$.

Proposition 9.2. *n_H is also equal to the number of $\nu < 1$ such that ν is an eigenvalue of $K(0)$.*

We'll prove this in a series of lemmas.

Lemma 9.3. *For all $\lambda < 0$, $\partial_\lambda K(\lambda) > 0$.*

Proof. If $\phi \neq 0$,

$$\begin{aligned}
\partial_\lambda(\phi, K(\lambda)\phi) &= \partial_\lambda\Big(U^{1/2}\phi, (-\Delta - \lambda)^{-1}U^{1/2}\phi\Big) \\
&= \Big(U^{1/2}\phi, (-\Delta - \lambda)^{-2}U^{1/2}\phi\Big) \\
&= \|(-\Delta - \lambda)^{-1}U^{1/2}\phi\|_{L^2}^2 > 0
\end{aligned}$$

⊠

Lemma 9.4. *As $\lambda \to \infty$, $K(\lambda) \to 0$.*

This proof is a nice application of a bunch of tools you learned in your functional analysis course.

Proof. We'll prove this by calculating the integral kernel of $K(\lambda)$, using the integral kernel for $(-\Delta - \lambda)^{-1}$. First,

$$(-\Delta - \lambda)^{-1}(x, y) = \frac{1}{2\pi|x - y|}e^{\sqrt{|\lambda|}|x-y|}. \tag{9.5}$$

The integral kernel is one such that you get a convolution operator after the Fourier and inverse Fourier transforms, and is the infinite-dimensional generalization of matrix multiplication.

$$\begin{aligned}
((-\Delta - \lambda)^{-1}f)(x) &= \Big((\xi^2 - \lambda)^{-1}\widehat{f}\Big)^\vee(x) \\
&= \Big(((|\cdot|^2 - \lambda)^{-1})^\vee * f\Big)(x),
\end{aligned} \tag{9.6}$$

where inside the absolute value is

$$\int \frac{1}{\xi^2 + |\lambda|}e^{i\xi z}\,\mathrm{d}\xi = C\frac{e^{-\sqrt{|\lambda|}}|z|}{2|z|},$$

where we integrated over the ξ such that $|\xi| = \pm i|\lambda|^{1/2}$. Therefore (9.6) is

$$\int G_\lambda(x, y)f(y)\,\mathrm{d}y,$$

where $G_\lambda(x, y)$ is the Green's function, which in this case is either side of (9.5).

Remark. There are two norms one can put on a kernel: the usual operator norm and the *Hilbert-Schmidt norm*

$$\|K\|_{\mathrm{HS}} := \left(\int |K(x, y)|^2\,\mathrm{d}x\,\mathrm{d}y\right)^{1/2}.$$

It turns out this is always at least as big as the operator norm: for any $f \in L^2$,

$$\begin{aligned}\|Kf\|_{L^2}^2 &= (Kf, Kf)_{L^2}\\ &= \int \mathrm{d}x \left| \int K(x,y) f(y)\,\mathrm{d}y \right|^2.\end{aligned}$$

By Cauchy-Schwarz,

$$\begin{aligned}&\leq \int \mathrm{d}x \left(|K(x,y)|^2\,\mathrm{d}y \right) \left(\int |f(y)^2|\,\mathrm{d}y \right)\\ &= \int |K(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y \|f\|_{L^2}^2.\end{aligned}$$

Hence $\|K\|_{\mathrm{op}} \leq \|K\|_{\mathrm{HS}}$.

It'll also be useful to recall the definition of the trace of an integral kernel: if $\{\phi_i\}$ is an orthonormal basis of $L^2(\mathbb{R}^3)$,

$$\operatorname{tr} K := \sum (\phi_j, K\phi_j).$$

Basis-independently, this is also

$$\operatorname{tr} K = \int K(x,x)\,\mathrm{d}x.$$

Putting all this together,

$$K(\lambda) = U^{1/2}(x) \frac{1}{2\pi|x-y|} e^{-\sqrt{|\lambda|}|x-y|} U^{1/2}(y),$$

so

$$\|K(\lambda)\|_{\mathrm{op}} \leq \left(\int \frac{U(x)U(y)}{4\pi^2|x-y|^2} e^{-2\sqrt{|\lambda|}|x-y|}\,\mathrm{d}x\,\mathrm{d}y \right)^{1/2},$$

and this tends to 0 as $\lambda \to -\infty$. ⊠

Proof of Proposition 9.2. Since $\|K(\lambda)\| \to 0$ as $\lambda \to -\infty$, all eigenvalues are less than 1 for λ sufficiently negative. Since $\partial_\lambda K(\lambda) > 0$ for all $\lambda < 0$, then the eigenvalues of $K(\lambda)$ increase monotonically in λ. The idea is that there's this "eigenvalue flow" such that as λ gets more negative, its eigenvalues get closer to 0.

Let $\nu_m(\lambda)$ be the m^{th} eigenvalue of λ. Then, if $\nu_m(\lambda_m) = 1$ for some $\lambda_m < 0$, then $\nu_m(0) > 1$. This means there's a one-to-one correspondence between the eigenvalues $\nu_m(0) > 1$ of $K(0)$ and the points λ_m at which some eigenvalue $\nu_m(\lambda)$ crosses 1, which is, as required, the number of $\lambda < 0$ which have 1 as an eigenvalue of $K(\lambda)$.

But then,

$$\begin{aligned}\{\nu > 1 \mid \nu \text{ is an eigenvalue for } K(0)\} &= \sum_{\substack{\nu_m > 1\\ \text{eigenvalues of } K(0)}} 1\\ &\leq \sum_{\nu_m > 1} \nu_m^2\\ &\leq \sum_{\text{eigenvalues of } K(0)} \nu_m^2\\ &= \operatorname{tr}|K(0)|^2 = \|K(0)\|_{\mathrm{HS}}^2.\end{aligned}$$

This norm is also

$$\frac{1}{(2\pi)^2} \int \frac{U(x)U(y)}{|x-y|^2}\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{(2\pi)^2} \frac{V(x)V(y)}{|x-y|^2}\,\mathrm{d}x\,\mathrm{d}y.$$

The right-hand side is sometimes called the *Rollwick norm* of V. Then, using the Hardy-Littlewood inequality,

$$\leq \frac{1}{(2\pi)^2} C\|V\|_{L^{3/2}}^2. \tag{9.7}$$

The Hardy-Littlewood inequality here depends on the fact that the dimension is 3, and indeed, eigenvalues of Schrödinger operators behave differently in dimension 2.

But the point is, the number of eigenvalues is finite for $V \in L^{3/2}$, and there cannot be any if (9.7) is greater than 1. ⊠