

Contents

Preface	2
Prelude: Fourier Transform and Schwartz Space	2
Functions of Moderate Decrease	2
Schwartz Space	3
The Fourier Transform on $\mathcal{S}(\mathbb{R})$	3
Good Kernels for the Fourier Transform	4
Fourier Inversion	5
Plancherel's Theorem	6
Prelude: Sesquilinear and Quadratic Forms on Hilbert Spaces	7
Classical Mechanics	8
Motion in \mathbb{R}^1	8
Conservation of Energy	9
Damping	10
Motion in \mathbb{R}^n	10
Systems of Particles	11
Momentum of a System of Particles	11
Center of Mass	13
Motion of Relative Position under Translation Invariant Potential	13
Angular Momentum	13
Hamiltonian Mechanics	15
Poisson Bracket	15
Conserved Quantities	16
Flow and Liouville's Theorem	16
Hamiltonian Flow and Hamiltonian Generators	17
Introduction to Quantum Mechanics	18
A Taste of Operator Theory	18
Position and Momentum Operators	20
Kinematic Axioms of Quantum Mechanics	23
Hydrogen Atom Measurements and Uncertainty	24
The Schrödinger Equation	25
Solving the Schrödinger Equation	26
The Schrödinger Equation in \mathbb{R}	27
Expected Position and Momentum for Solutions to the Schrödinger Equation	28
The Heisenberg Approach	28
Particle in a Box	30
n -Dimensional Single-Particle Quantum Mechanics	31
Systems of Multiple Particles	32
Dirac Notation	33
The Schrödinger Equation for a Free Particle	34
Using the Fourier Transform to solve the Free Schrödinger Equation	34
Solving via Convolutions	36
Solving by Propagation of the Wave Function: First Approach	37

Introduction to Spectral Theory	37
Goals of Spectral Theory	37
Spectral Theorem on the Position Operator	38
Generalization: Multiplication Operators	39
Spectral Theorem on the Momentum Operator	39
The Spectral Theorem for Bounded Self-Adjoint Operators	40
Preliminaries	40
Spectral Subspaces	44
Projection-Valued Measure	45

Preface

Hello, you have stumbled across my notes on quantum mechanics from a mathematical perspective. Welcome! These notes are going to be largely based on the book *Quantum Theory for Mathematicians* by Brian C. Hall, a book I got on recommendation of a friend at my REU. This is mostly going to be an experiment in seeing the level of analysis I'll be able to put in my brain over the course of a summer without access to much in the way of resources.

As far as sectioning goes, I'll mostly be following the chapters of the book; if there are certain details that are especially important that I can't see myself learning on the fly, I will probably source another book in order to find a more detailed explanation (the appendices of the main book are not particularly strong as far as explanatory power).

There will be footnotes, either with moderately snarky comments or me explaining to myself some of the results that the book assumes the reader instantly understands.

If I'm being honest, this project will probably be a failure as far as actually understanding quantum mechanics goes, but hopefully by the end of it I'll have a decent understanding of spectral theory, measure theory, and maybe solve the Schrödinger equation or quantum harmonic oscillator.

Prelude: Fourier Transform and Schwartz Space

It's probably pretty useful to understand the basics of the Fourier transform before we jump in.ⁱ To do this, the following are notes of some content from the *Princeton Lectures on Analysis*.

Functions of Moderate Decrease

A function f defined on \mathbb{R} that is integrable over a closed and bounded interval may not be integrable over \mathbb{R} . For instance, the function $f = \frac{1}{1+|x|}$ is such a function.

In order to study functions that *are* integrable over \mathbb{R} , we will find it useful to restrict our view to certain classes of functions — namely, ones that decrease “fast enough.”

Definition (Function of Moderate Decrease). A function $f \in C(\mathbb{R})$ is said to be of moderate decrease if there exists $A > 0$ such that

$$|f(x)| \leq \frac{A}{1+x^{1+\varepsilon}}$$

ⁱI actually did this in the middle of learning but decided to move it up to the front, alongside some other preliminaries I'm going to have to draw on.

for all $x \in \mathbb{R}$ and some $\varepsilon > 0$. We denote the set of all functions of moderate decrease as $\mathcal{M}(\mathbb{R})$; $\mathcal{M}(\mathbb{R})$ is a vector space over \mathbb{C} .

If f is of moderate decrease, we can define

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{N \rightarrow \infty} \int_{-N}^N f(x) dx.$$

The integral over \mathbb{R} of $f, g \in \mathcal{M}(\mathbb{R})$ (with $a, b \in \mathbb{C}$) has the following properties:

- Linearity:

$$\int_{-\infty}^{\infty} (af(x) + bg(x)) dx = a \int_{-\infty}^{\infty} f(x) dx + b \int_{-\infty}^{\infty} g(x) dx.$$

- Translation Invariance:

$$\int_{-\infty}^{\infty} f(x - h) dx = \int_{-\infty}^{\infty} f(x) dx.$$

- Scaling under dilation:

$$\delta \int_{-\infty}^{\infty} f(\delta x) dx = \int_{-\infty}^{\infty} f(x) dx$$

- Continuity:

$$\lim_{h \rightarrow 0} \int_{-\infty}^{\infty} |f(x - h) - f(x)| dx = 0.$$

Schwartz Space

The Schwartz space is a subspace of $f \in C^\infty(\mathbb{R})$ where, for all $k, \ell \geq 0$,

$$\sup_{x \in \mathbb{R}} |x^k f^{(\ell)}(x)| < \infty.$$

The space $\mathcal{S}(\mathbb{R}) \subseteq C^\infty(\mathbb{R})$ is also a vector space over \mathbb{C} . Notice that if $f \in \mathcal{S}(\mathbb{R})$, then so too are $f'(x)$ and $xf(x)$.

Important classes of Schwartz functions include the Gaussians and the bump functions (compactly supported $C^\infty(\mathbb{R})$ functions).ⁱⁱ

The Fourier Transform on $\mathcal{S}(\mathbb{R})$

Definition (Fourier Transform). The Fourier transform of a function $f \in \mathcal{S}(\mathbb{R})$ is defined by

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i k t} dt.$$

I will use \mathcal{F} to denote the Fourier transform; $\mathcal{F}(f(t)) = \hat{f}(k)$.

Proposition (Properties of the Fourier Transform). Let $f \in \mathcal{S}(\mathbb{R})$. Then:

- $\mathcal{F}(f(x + h)) = \hat{f}(k) e^{2\pi i h k}$ for $h \in \mathbb{R}$;

ⁱⁱBump functions are fascinating because even though they're C^∞ , they are not analytic (or equivalent to their Taylor series, denoted C^ω). Compare this to functions defined over the complex numbers, where $C^1(\mathbb{C}) = C^\infty(\mathbb{C}) = C^\omega(\mathbb{C})$.

- $\mathcal{F}(f(x)e^{-2\pi i x h}) = \hat{f}(k + h)$ for $h \in \mathbb{R}$;
- $\mathcal{F}(f(\delta x)) = \delta^{-1} \hat{f}(\delta^{-1} k)$ for $\delta > 0$;
- $\mathcal{F}(f'(x)) = 2\pi i k \hat{f}(k)$;
- $\mathcal{F}(-2\pi i x f(x)) = \frac{d\hat{f}}{dk}$.

One of the important facts of the Fourier transform is that it preserves the Schwartz space.

Theorem (Preservation of Schwartz Space under Fourier Transform). *If $f \in \mathcal{S}(\mathbb{R})$, then so too is \hat{f} .*

Proof. Let $f \in \mathcal{S}(\mathbb{R})$. Then, \hat{f} is bounded,ⁱⁱⁱ so for every nonnegative k, ℓ , it is the case that

$$\xi^k \left(\frac{d}{d\xi} \right)^\ell \hat{f}(\xi)$$

is bounded, as it is the Fourier transform of the (bounded) expression

$$\frac{1}{(2\pi i)^k} \left(\frac{d}{dx} \right)^k ((-2\pi i x)^\ell f(x)).$$

Thus, $\hat{f} \in \mathcal{S}(\mathbb{R})$. □

Good Kernels for the Fourier Transform

Consider the fact that

$$\int_{-\infty}^{\infty} e^{-\pi x^2} dx = 1.$$

We can see that $e^{-\pi x^2} \in \mathcal{S}(\mathbb{R})$, meaning we can examine its Fourier transform.

Theorem (Fourier Transform of $e^{-\pi x^2}$). *Let $f(x) = e^{-\pi x^2}$. Then, $\hat{f}(\xi) = f(\xi)$.*

Proof. Define $F(\xi) = \hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i x \xi} dx$. Observe that $F(0) = 1$.

Using the fact that $\frac{df}{dx} = -2\pi x f(x)$ and that $\mathcal{F}(-2\pi i x f(x)) = \frac{d\hat{f}}{d\xi}$, we find by differentiating under the integral sign that

$$\begin{aligned} F'(\xi) &= \int_{-\infty}^{\infty} f(x)(-2\pi i x) e^{-2\pi i x \xi} dx \\ &= i \int_{-\infty}^{\infty} f'(x) e^{-2\pi i x \xi} dx \end{aligned}$$

meaning

$$\begin{aligned} F'(\xi) &= i \left(\frac{d}{d\xi} \hat{f}(\xi) \right) \\ &= -2\pi \xi F(\xi). \end{aligned}$$

Thus, $F(\xi) = e^{-\pi \xi^2}$.^{iv} □

From this, we can see that if $K_\delta(x) = \delta^{-1/2} e^{-\pi x^2/\delta}$, then $\widehat{K_\delta}(\xi) = e^{-\pi \delta \xi^2}$. We can see that as $\delta \rightarrow 0$, K_δ peaks at the origin, while $\widehat{K_\delta}$ becomes flatter.

Importantly, $K_\delta(x) = \delta^{-1/2} e^{-\pi x^2/\delta}$ also satisfies the following properties:

ⁱⁱⁱI'm not sure exactly why.

^{iv}I skipped a step here that was kind of annoying even though I maybe shouldn't have.

- $\int_{-\infty}^{\infty} K_{\delta}(x) dx = 1;$
- $\int_{-\infty}^{\infty} |K_{\delta}(x)| dx \leq M;$
- for every $\eta > 0$, we have $\int_{|x|>\eta} |K_{\delta}(x)| dx \rightarrow 0$ as $\delta \rightarrow 0$.

This makes the collection $\{K_{\delta}\}$ a family of good kernels for the Fourier transform.

In particular, this means that $(f * K_{\delta})(x) \rightarrow f(x)$ uniformly as $\delta \rightarrow 0$ for any $f \in \mathcal{S}(\mathbb{R})$, where $*$ denotes the convolution.

Fourier Inversion

Proposition (Multiplication Identity). *If $f, g \in \mathcal{S}(\mathbb{R})$, then*

$$\int_{-\infty}^{\infty} f(x) \hat{g}(x) dx = \int_{-\infty}^{\infty} \hat{f}(y) g(y) dy.$$

Proof. Consider a continuous multivariate function of moderate decrease in x and y , $|F(x, y)| \leq \frac{A}{(1+x^2)(1+y^2)}$. Set $F_1(x) = \int_{-\infty}^{\infty} F(x, y) dy$ and $F_2(y) = \int_{-\infty}^{\infty} F(x, y) dx$. Then, by Fubini's Theorem, we have

$$\int_{-\infty}^{\infty} F_1(x) dx = \int_{-\infty}^{\infty} F(y) dy.$$

If $F(x, y) = f(x)g(y)e^{-2\pi ixy}$, then $F_1(x) = f(x)\hat{g}(x)$ and $F_2(y) = \hat{f}(y)g(y)$. □

Theorem (Fourier Inversion). *If $f \in \mathcal{S}(\mathbb{R})$, then^v*

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{2\pi i x \xi} d\xi.$$

Proof. We start by claiming that

$$f(0) = \int_{-\infty}^{\infty} \hat{f}(\xi) d\xi.$$

Let $G_{\delta}(x) = e^{-\pi\delta x^2}$ such that $\widehat{G_{\delta}}(\xi) = K_{\delta}(\xi)$. By the multiplication formula, we get

$$\int_{-\infty}^{\infty} f(x) K_{\delta}(x) dx = \int_{-\infty}^{\infty} \hat{f}(\xi) G_{\delta}(\xi) d\xi.$$

Since K_{δ} is a good kernel, the first integral goes to $f(0)$ as $\delta \rightarrow 0$, and since the second integral converges to $\int_{-\infty}^{\infty} \hat{f}(\xi) d\xi$ as $\delta \rightarrow 0$, we find that the claim is true.

Let $F(y) = f(y+x)$, so

$$\begin{aligned} f(x) &= F(0) \\ &= \int_{-\infty}^{\infty} \hat{F}(\xi) d\xi \\ &= \int_{-\infty}^{\infty} \hat{f}(\xi) e^{2\pi i x \xi} d\xi. \end{aligned}$$

□

^vNote the sign change.

This implies the existence of two maps, $\mathcal{F} : \mathcal{S}(\mathbb{R}) \rightarrow \mathcal{S}(\mathbb{R})$ and $\mathcal{F}^* : \mathcal{S}(\mathbb{R}) \rightarrow \mathcal{S}(\mathbb{R})$, where

$$\begin{aligned}\mathcal{F}(f)(\xi) &= \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi} dx \\ \mathcal{F}^*(g)(x) &= \int_{-\infty}^{\infty} g(\xi) e^{2\pi i x \xi} d\xi.\end{aligned}$$

We say \mathcal{F} is the Fourier transform, and \mathcal{F}^* is the inverse Fourier transform; $\mathcal{F} \circ \mathcal{F}^* = \mathcal{F}^* \circ \mathcal{F} = I$.

Plancherel's Theorem

In order to prove that $\|\mathcal{F}\|_{\text{op}} = 1$, we need to understand the Fourier transform's effects on convolutions.

Proposition (Convolutions under the Fourier Transform). *Let $f, g \in \mathcal{S}(\mathbb{R})$. Then,*

- $f * g \in \mathcal{S}(\mathbb{R})$;
- $f * g = g * f$;
- $\widehat{(f * g)}(\xi) = \hat{f}(\xi) \hat{g}(\xi)$.

Proof. To prove that $f * g$ is rapidly decreasing, we observe that for $\ell \geq 0$,

$$\sup_x |x|^\ell |g(x - y)| \leq A_\ell (1 + |y|)^\ell,$$

since g is rapidly decreasing. Thus, we see that

$$\sup_x |x|^\ell |(f * g)(x)| \leq A_\ell \int_{-\infty}^{\infty} |f(y)| (1 + |y|)^\ell dy,$$

meaning $x^\ell (f * g)(x)$ is bounded for every $\ell \geq 0$. Similarly, since

$$\left(\frac{d}{dx}\right)^k (f * g)(x) = \left(f * \left(\frac{d}{dx}\right)^k g\right)(x),$$

it is the case that $f * g \in \mathcal{S}(\mathbb{R})$.

To show that $f * g = g * f$, simply substitute $x - y = u$ in the formula of $f * g$:

$$\begin{aligned}(f * g)(x) &= \int_{-\infty}^{\infty} f(y) g(x - y) dy \\ &= \int_{-\infty}^{\infty} f(x - u) g(u) du \\ &= (g * f)(x)\end{aligned}$$

Finally, consider $F(x, y) = f(y) g(x - y) e^{-2\pi i x \xi}$. We set $F_1(x) = (f * g)(x) e^{-2\pi i x \xi}$ (by integrating with respect to y), and $F_2(x) = f(y) e^{-2\pi i y \xi} \hat{g}(\xi)$. Thus,

$$\int_{-\infty}^{\infty} F_1(x) dx = \int_{-\infty}^{\infty} F_2(y) dy.$$

□

The Schwartz space is a Hilbert space with inner product

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)} dx.$$

Theorem (Plancherel). If $f \in \mathcal{S}(\mathbb{R})$, then $\|\hat{f}\| = \|f\|$.

Proof. Let $f \in \mathcal{S}(\mathbb{R})$. Define $\tilde{f} = \overline{f(-x)}$. Then, $\hat{\tilde{f}}(\xi) = \overline{\hat{f}(\xi)}$. Let $h = f * \tilde{f}$. Clearly,

$$\begin{aligned}\hat{h}(\xi) &= |\hat{f}(\xi)|^2 \\ h(0) &= \int_{-\infty}^{\infty} |f(x)|^2 dx.\end{aligned}$$

Plancherel's theorem follows from the inversion formula applied at $x = 0$, meaning

$$\int_{-\infty}^{\infty} \hat{h}(\xi) d\xi = h(0).$$

□

Prelude: Sesquilinear and Quadratic Forms on Hilbert Spaces

It will be useful to understand the general theory of sesquilinear forms as we go deeper into theories of operators on Hilbert spaces. To do this, I will draw information from the book *Introduction to Hilbert Spaces with Applications* by Debnath and Mikusinski, as well as a result or two from *Quantum Theory for Mathematicians*.

Definition (Sesquilinear Form). Let E be a \mathbb{C} -vector space. A sesquilinear form^{vi} is a map $\varphi : E \times E \rightarrow \mathbb{C}$ with the following properties:

- (a) $\varphi(\alpha x_1 + \beta x_2, y) = \alpha \varphi(x_1, y) + \beta \varphi(x_2, y)$;
- (b) $\varphi(x, \alpha y_1 + \beta y_2) = \bar{\alpha} \varphi(x, y_1) + \bar{\beta} \varphi(x, y_2)$;

for any $\alpha, \beta \in \mathbb{C}$ and $x, x_1, x_2, y, y_1, y_2 \in E$.

A familiar sesquilinear form is the inner product, but we want to develop properties and ideas that are generalized over all sesquilinear forms rather than simply sticking to the inner product.

Definition (Classes of Sesquilinear Forms). Let φ be a sesquilinear form on E .

- (a) We say φ is symmetric if $\varphi(x, y) = \overline{\varphi(y, x)}$ for all $x, y \in E$.^{vii}
- (b) We say φ is positive if $\varphi(x, x) \geq 0$ for every $x \in E$.
- (c) We say φ is *strictly* positive if $\varphi(x, x) = 0$ if and only if $x = 0$.
- (d) If E has a normed space, then φ is bounded if $|\varphi(x, y)| \leq K \|x\| \|y\|$ for some $K > 0$ and all $x, y \in E$.^{viii} The operator norm is

$$\|\varphi\| = \sup_{\|x\|, \|y\|=1} |\varphi(x, y)|.$$

Note that, by the definition of the operator norm, $|\varphi(x, y)| \leq \|\varphi\| \|x\| \|y\|$ for all $x, y \in E$.

Definition (Quadratic Form). For a particular sesquilinear form φ on E , the function $\Phi : E \rightarrow \mathbb{C}$ defined by $\Phi(x) = \varphi(x, x)$ is known as the quadratic form for φ .

If E has a norm, then $\Phi(x)$ is bounded if $|\Phi(x)| \leq K \|x\|^2$ for all $x \in E$ and some $K > 0$. The norm of Φ is defined by

$$\|\Phi\| = \sup_{\|x\|=1} |\Phi(x)|.$$

^{vi}Debnath and Mikusinski call this a bilinear functional.

^{vii}I'd personally prefer to call this "conjugate symmetry."

^{viii}The Cauchy-Schwarz inequality over Hilbert spaces with the standard inner product is a particular case of this.

For a bounded quadratic form Φ on a normed space, $|\Phi(x)| \leq \|\phi\| \|x\|^2$.

The standard inner product in an inner product space is a special case where $\|x\|^2 = \langle x, x \rangle$.

Theorem (Polarization Identity). *For any φ a sesquilinear form on E (and associated quadratic form Φ),*

$$4\varphi(x, y) = \Phi(x + y) - \Phi(x - y) + i\Phi(x + iy) - i\Phi(x - iy)$$

for all $x, y \in E$.

Proof of Polarization Identity. For any $\alpha, \beta \in \mathbb{C}$, we have

$$\Phi(\alpha x + \beta y) = |\alpha|^2 \Phi(x) + \alpha \bar{\beta} \varphi(x, y) + \bar{\alpha} \beta \varphi(y, x) + |\beta|^2 \Phi(y).$$

Keeping $\alpha = 1$, the cases of $\beta = 1, -1, i, -i$ yield

$$\Phi(x + y) = \Phi(x) + \varphi(x, y) + \varphi(y, x) + \Phi(y); -\Phi(x - y) = -\Phi(x) + \varphi(x, y) + \varphi(y, x) - \Phi(y); i\Phi(x + iy) = i\Phi(x) + \varphi(x, y)$$

Adding all these together, we get the polarization identity. \square

Proposition (Passing Quadratic Forms into Operators). *Let Q be a quadratic form over a Hilbert space \mathbf{H} . Then, there is a unique $A \in \mathcal{B}(\mathbf{H})^{\text{ix}}$ such that $Q(\psi) = \langle \psi, A\psi \rangle$ for all $\psi \in \mathbf{H}$.*

If $Q(\psi) \in \mathbb{R}$ for all $\psi \in \mathbf{H}$, then A is self-adjoint.

Proof. Since Q is bounded, so too is the sesquilinear form from which it originates (which we will call L). Therefore, there exists a constant C such that $|L(\phi, \psi)| \leq C \|\phi\| \|\psi\|$, and for any $\phi \in \mathbf{H}$, the linear functional $\psi \mapsto L(\phi, \psi)$ is bounded.

By the Riesz representation theorem, there exists a unique $\chi \in \mathbf{H}$ with $\|\chi\| \leq C \|\phi\|$, and $L(\phi, \psi) = \langle \chi, \psi \rangle$.

Define $B : \mathbf{H} \rightarrow \mathbf{H}$ by $B\phi = \chi$ — thus meaning $L(\phi, \psi) = \langle B\phi, \psi \rangle$. We can see that B is bounded linear. Setting $A = B^*$, we find the desired operator; we can see that A is unique since, if $\langle \phi, A\psi \rangle = 0$ for all $\phi, \psi \in \mathbf{H}$, then A is the zero operator.

If $Q(\psi)$ is real for all $\psi \in \mathbf{H}$, then

$$\begin{aligned} \langle \phi, A\psi \rangle &= L(\phi, \psi) \\ &= \overline{L(\psi, \phi)} \\ &= \overline{\langle \psi, A\phi \rangle} \\ &= \langle A\phi, \psi \rangle, \end{aligned}$$

meaning A is self-adjoint. \square

Classical Mechanics

Motion in \mathbb{R}^1

Let $x(t)$ denote position. Then, $v(t) = \frac{dx}{dt} = \dot{x}(t)$ is velocity (where the \cdot denotes derivative with respect to time), $a(t) = \dot{v}(t) = \ddot{x}(t)$, etc.

^{ix}The set of all bounded linear operators on \mathbf{H} into itself.

Considering Newton's second law, $F(x(t)) = m\ddot{x}(t)$, every exact solution requires initial conditions of $x(t_0)$ and $v(t_0)$. Solutions to Newton's second law are known as trajectories.

Considering a spring of constant k , $F(x) = -kx$ yields the differential equation $m\ddot{x} + kx = 0$. The general solution is

$$x(t) = a \cos(\omega t) + b \sin(\omega t),$$

with $\omega = \sqrt{k/m}$ denoting the frequency. The spring is an example of a simple harmonic oscillator.

Conservation of Energy

For a general force function $F(x)$, the kinetic energy is $\frac{1}{2}mv^2$, and the potential energy is

$$V(x) = - \int F(x) dx,$$

meaning $F(x) = -\frac{dV}{dx}$. The total energy is thus found as

$$E(x, v) = \frac{1}{2}mv^2 + V(x).$$

Proposition (Conservation of Energy). *If a particle with trajectory $x(t)$ satisfies $m\ddot{x} = F(x)$, then the energy E is conserved.*

Proof.

$$\begin{aligned} \frac{d}{dt} E(x(t), \dot{x}(t)) &= \frac{d}{dt} \left(\frac{1}{2}m(\dot{x}(t))^2 + V(x(t)) \right) \\ &= m\dot{x}(t)\ddot{x}(t) + \frac{dV}{dx}\dot{x}(t) \\ &= \dot{x}(t)(m\ddot{x}(t) - F(x(t))). \end{aligned}$$

□

By using the conservation of energy, we can reduce the second order differential equation $F(x) = m\ddot{x}$ to a system of first order differential equations in $x(t)$ and $v(t)$ respectively:

$$\begin{aligned} \frac{dx}{dt} &= v(t) \\ \frac{dv}{dt} &= \frac{1}{m}F(x(t)). \end{aligned}$$

If $(x(t), v(t))$ satisfies this set of equations, then $x(t)$ satisfies Newton's second law. We say the set of all possible (x, v) forms the phase space for the particle in \mathbb{R}^1 .

In phase space, conservation of energy implies that the set of all (x, v) must lie on the level curve of the energy function: $\{(x, v) \mid E(x, v) = E(x_0, v_0)\}$.

Using the conservation of energy, we find that, though Newton's second law is a second order differential equation in time, it is actually a first order differential equation:

$$\begin{aligned} \frac{m}{2}(\dot{x}(t))^2 + V(x(t)) &= E(x(t_0), v(t_0)) \\ \dot{x}(t) &= \sqrt{\frac{2(E_0 - V(x(t)))}{m}} \end{aligned}$$

Damping

Suppose we also introduce a force that depends on velocity — in the case of a damped simple harmonic oscillator, the equation for force changes from $F = -kx$ to $F = -kx - \gamma\dot{x}$, with $\gamma > 0$. The damping force acts in the opposite direction of velocity, meaning the particle slows down.

The equation of motion is then

$$m\ddot{x} + \gamma\dot{x} + kx = 0.$$

For γ small, the solutions are a sum sines and cosines multiplied by some exponential decay factor, but for γ large, the solutions are only the exponential decay.

Proposition (Energy Conservation in Damped Systems). *Suppose a particle moves along $x(t)$ that satisfies $F(x, \dot{x}) = F_1(x) - \gamma\dot{x}$, with $\frac{dV}{dx} = -F_1(x)$ and $\gamma > 0$. Then,*

$$\frac{d}{dt}E(x(t), \dot{x}(t)) = -\gamma\dot{x}(t)^2.$$

Proof.

$$\begin{aligned} \frac{d}{dt}E(x(t), \dot{x}(t)) &= \dot{x}(t) (m\ddot{x}(t) - F_1(x(t))) \\ &= \dot{x}(t) (m\ddot{x}(t) - (m\ddot{x}(t) + \gamma\dot{x}(t))) \\ &= -\gamma\dot{x}(t)^2 \end{aligned}$$

□

Motion in \mathbb{R}^n

The position of a particle $\mathbf{x} = (x_1, \dots, x_n)$ lends itself to velocity $\mathbf{v} = (v_1, \dots, v_n) = (\dot{x}_1, \dots, \dot{x}_n)$, and $\mathbf{a} = (\ddot{x}_1, \dots, \ddot{x}_n)$. Similar to in \mathbb{R}^1 , Newton's second law is denoted

$$m\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x}(t), \dot{\mathbf{x}}(t)).$$

Proposition (Conservation of Energy in n Dimensions). *The energy function*

$$E(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}m\|\dot{\mathbf{x}}\|^2 + V(\mathbf{x})$$

is only satisfied where $\mathbf{F} = -\nabla V$.

Proof.

$$\begin{aligned} \frac{d}{dt} \left(\frac{1}{2}m\|\dot{\mathbf{x}}\|^2 + V(\mathbf{x}) \right) &= m \sum_{j=1}^n \dot{x}_j \ddot{x}_j + \sum_{j=1}^n \frac{\partial V}{\partial x_j} \dot{x}_j(t) \\ &= \dot{\mathbf{x}}(t) (m\ddot{\mathbf{x}}(t) + \nabla V) \\ &= \dot{\mathbf{x}}(t) (\mathbf{F}(\mathbf{x}) + \nabla V(\mathbf{x})), \end{aligned}$$

which is equal to zero only if $-\nabla V = \mathbf{F}$.

□

If \mathbf{F} is a smooth \mathbb{R}^n valued function on $U \subset \mathbb{R}^n$, then \mathbf{F} is conservative if there exists a smooth real-valued function V such that $\mathbf{F} = -\nabla V$.

In other words, \mathbf{F} is conservative if \mathbf{F} is a gradient field, implying that $\nabla \times \mathbf{F} = 0$.

If $\mathbf{F}(\mathbf{x}, \mathbf{y}) = -\nabla V(\mathbf{x}) + \mathbf{F}_2(\mathbf{x}, \mathbf{y})$, with $\mathbf{v} \cdot \mathbf{F}_2 = 0$ for all \mathbf{x} and \mathbf{v} , then energy is conserved along a given trajectory.

Systems of Particles

Let $\mathbf{x}^j = (x_1^j, x_2^j, \dots, x_n^j)$ denote the j th particle of a system of N particles. Newton's second law is thus reformulated as

$$m_j \ddot{\mathbf{x}}^j = \mathbf{F}^j(\mathbf{x}^1, \dots, \mathbf{x}^N, \dot{\mathbf{x}}^1, \dots, \dot{\mathbf{x}}^N).$$

The total energy is determined by

$$E(\mathbf{x}^1, \dots, \mathbf{x}^N, \mathbf{v}^1, \dots, \mathbf{v}^N) = \left(\sum_{j=1}^N \frac{1}{2} m_j \|\mathbf{v}^j\|^2 \right) + V(\mathbf{x}^1, \dots, \mathbf{x}^N).$$

Proposition (Conservation of Energy in a System of Particles). *The energy function is constant along each trajectory if $\nabla^j V = -\mathbf{F}^j$, where ∇^j denotes the gradient with respect to \mathbf{x}^j .*

The force function along a simply connected domain U in \mathbb{R}^{nN} satisfies $\nabla^j V = -\mathbf{F}^j$ if and only if

$$\frac{\partial F_k^j}{\partial x_m^l} = \frac{\partial F_m^l}{\partial x_k^j}$$

for all j, k, l, m .

Proof.

$$\begin{aligned} \frac{dE}{dt} &= \sum_{j=1}^N (m_j \dot{\mathbf{x}}^j \cdot \ddot{\mathbf{x}}^j + \nabla^j V \cdot \dot{\mathbf{x}}^j) \\ &= \sum_{j=1}^N \dot{\mathbf{x}}^j (m_j \ddot{\mathbf{x}}^j + \nabla^j V) \\ &= \sum_{j=1}^N \dot{\mathbf{x}}^j (\mathbf{F}^j + \nabla^j V), \end{aligned}$$

which is equal to zero if $\nabla^j V = -\mathbf{F}^j$.

Applying a higher dimension version of $\nabla \times \mathbf{F}$ to each coordinate pair (a, b) , we find the identity that shows \mathbf{F} is a gradient field. \square

Momentum of a System of Particles

The momentum of a particle \mathbf{p}^j is defined by

$$\mathbf{p}^j = m_j \dot{\mathbf{x}}^j.$$

Observe that $\frac{d\mathbf{p}^j}{dt} = m_j \ddot{\mathbf{x}}^j = \mathbf{F}^j$. The total momentum is then

$$\mathbf{p} = \sum_{j=1}^N \mathbf{p}^j.$$

Newton's third law, which states "for every action there is an equal and opposite reaction" applies if

- $\mathbf{F}^j = \sum_{k \neq j} \mathbf{F}^{j,k}(\mathbf{x}^j, \mathbf{y}^j);$
- $\mathbf{F}^{j,k}(\mathbf{x}_j, \mathbf{x}_k) = -\mathbf{F}^{k,j}(\mathbf{x}^k, \mathbf{x}^j).$

If each \mathbf{F}^j is also a conservative force, then satisfying these conditions yields potential energy in the form of

$$V(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N) = \sum_{j < k} V^{j,k}(\mathbf{x}^j - \mathbf{x}^k).$$

Proposition (Newton's Third Law and Conservation of Momentum). *If the system of particles satisfies the conditions of*

- $\mathbf{F}^j = \sum_{k \neq j} \mathbf{F}^{j,k}(\mathbf{x}^j, \mathbf{x}^k)$
- and $\mathbf{F}^{j,k}(\mathbf{x}^j, \mathbf{x}^k) = -\mathbf{F}^{k,j}(\mathbf{x}^k, \mathbf{x}^j)$,

then total momentum is conserved.

Proof.

$$\begin{aligned} \frac{d\mathbf{p}}{dt} &= \sum_{j=1}^N \mathbf{F}^j \\ &= \sum_{j=1}^N \sum_{k \neq j} \mathbf{F}^{j,k}(\mathbf{x}^j, \mathbf{x}^k), \end{aligned}$$

and since $\mathbf{F}^{j,k}(\mathbf{x}^j, \mathbf{x}^k) + \mathbf{F}^{k,j}(\mathbf{x}^k, \mathbf{x}^j) = 0$, we find $\frac{d\mathbf{p}}{dt} = 0$. □

Proposition (Translation Invariance of Potential). *Let V denote the potential for a conservative force. Then, momentum is conserved if and only if V is translation invariant, meaning that for all $\mathbf{a} \in \mathbb{R}^n$,*

$$V(\mathbf{x}^1 + \mathbf{a}, \mathbf{x}^2 + \mathbf{a}, \dots, \mathbf{x}^N + \mathbf{a}) = V(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N).$$

Proof. Let $\mathbf{a} = t\mathbf{e}_k$. Then, differentiating at $t = 0$ with respect to t , we find

$$\begin{aligned} 0 &= \sum_{j=1}^N \frac{\partial V}{\partial x_k^j} \\ &= - \sum_{j=1}^N F_k^j \\ &= - \sum_{j=1}^N \frac{dp_k^j}{dt} \\ &= - \frac{dp_k}{dt}, \end{aligned}$$

with p_k denoting the k th component of \mathbf{p} . Therefore, \mathbf{p} is constant in time.

If \mathbf{p} is conserved, then the sum of all forces is 0 at each point for all t , meaning that for all t ,

$$\begin{aligned} \frac{d}{dt} V(\mathbf{x}^1 + t\mathbf{a}, \mathbf{x}^2 + t\mathbf{a}, \dots, \mathbf{x}^N + t\mathbf{a}) &= \sum_{j=1}^N \nabla^j V(\mathbf{x}^1 + t\mathbf{a}, \mathbf{x}^2 + t\mathbf{a}, \dots, \mathbf{x}^N + t\mathbf{a}) \cdot \mathbf{a} \\ &= - \left(\sum_{j=1}^N \mathbf{F}^j(\mathbf{x}^1 + t\mathbf{a}, \mathbf{x}^2 + t\mathbf{a}, \dots, \mathbf{x}^N + t\mathbf{a}) \right) \cdot \mathbf{a} \\ &= 0 \end{aligned}$$

meaning V is equal at $t = 0$ and $t = 1$. □

Center of Mass

For a system of N particles, the center of mass is denoted

$$\mathbf{c} = \sum_{j=1}^N \frac{m_j}{\sum_{j=1}^N m_j} \mathbf{x}_j.$$

We denote $\sum_{j=1}^N m_j = M$. Differentiating \mathbf{c} , we get

$$\begin{aligned} \frac{d\mathbf{c}}{dt} &= \frac{1}{M} \sum_{j=1}^N m_j \dot{\mathbf{x}}^j \\ &= \frac{\mathbf{p}}{M}. \end{aligned}$$

Notice that if \mathbf{p} is conserved, then $\mathbf{c}(t) = \mathbf{c}(t_0) + (t - t_0) \frac{\mathbf{p}}{M}$.

For a system of two particles, if $V(\mathbf{x}^1, \mathbf{x}^2)$ is invariant under translation, then $V(\mathbf{x}^1, \mathbf{x}^2) = \tilde{V}(\mathbf{x}^1 - \mathbf{x}^2)$, and $\tilde{V}(\mathbf{a}) = V(\mathbf{a}, 0)$.

The positions \mathbf{x}^1 and \mathbf{x}^2 can be recovered from knowledge about \mathbf{c} and the relative position $\mathbf{y} := \mathbf{x}^1 - \mathbf{x}^2$:

$$\begin{aligned} \mathbf{x}^1 &= \frac{\mathbf{c} + m_2 \mathbf{y}}{m_1 + m_2} \\ \mathbf{x}^2 &= \frac{\mathbf{c} - m_1 \mathbf{y}}{m_1 + m_2}. \end{aligned}$$

Thus, we can calculate

$$\begin{aligned} \ddot{\mathbf{y}} &= \ddot{\mathbf{x}}^1 - \ddot{\mathbf{x}}^2 \\ &= -\frac{1}{m_1} \nabla \tilde{V}(\mathbf{x}^1 - \mathbf{x}^2) - \frac{1}{m_2} \nabla \tilde{V}(\mathbf{x}^1 - \mathbf{x}^2). \end{aligned}$$

Motion of Relative Position under Translation Invariant Potential

For a two particle system with translation invariant potential, the relative position $\mathbf{y} = \mathbf{x}^1 - \mathbf{x}^2$ is a solution to the differential equation

$$\mu \ddot{\mathbf{y}} = -\nabla \tilde{V}(\mathbf{y}),$$

where

$$\mu = \frac{m_1 m_2}{m_1 + m_2}.$$

This implies that when momentum is conserved, the relative position of the two particle system evolves as a one-particle system with effective mass μ .

Angular Momentum

A particle moving in \mathbb{R}^2 with position \mathbf{x} , velocity \mathbf{v} , and momentum $\mathbf{p} = m\mathbf{v}$ has angular momentum J denoted as

$$J = x_1 p_2 - x_2 p_1,$$

or $J = \|\mathbf{x} \times \mathbf{p}\| = \|\mathbf{x}\| \|\mathbf{p}\| \sin \phi$, with ϕ measured counterclockwise. In polar coordinates, we find

$$\begin{aligned} J &= mr^2 \frac{d\theta}{dt} \\ &= 2M \frac{dA}{dt}, \end{aligned}$$

where $A = 1/2 \int r^2 d\theta$ denotes the area swept out by $\mathbf{x}(t)$.

Proposition (Conservation of Angular Momentum). *Suppose a particle of mass m is moving in \mathbb{R}^2 under the influence of a conservative force with potential $V(\mathbf{x})$. V is invariant under rotation if and only if J is conserved.*

Proof.

$$\begin{aligned} \frac{dJ}{dt} &= \frac{dx_1}{dt} p_2 + x_1 \frac{dp_2}{dt} - \frac{dx_2}{dt} p_1 - x_2 \frac{dp_1}{dt} \\ &= \frac{1}{m} p_1 p_2 - x_1 \frac{\partial V}{\partial x_2} - \frac{1}{m} p_2 p_1 + x_2 \frac{\partial V}{\partial x_1} \\ &= x_2 \frac{\partial V}{\partial x_1} - x_1 \frac{\partial V}{\partial x_2}. \end{aligned}$$

Alternatively, consider $R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$. Differentiating V along R_θ , we get

$$\begin{aligned} \left. \frac{d}{d\theta} V(R_\theta \mathbf{x}) \right|_{\theta=0} &= \frac{\partial V}{\partial x} \frac{dx}{d\theta} + \frac{\partial V}{\partial y} \frac{dy}{d\theta} \\ &= -x_2 \frac{\partial V}{\partial x_1} + x_1 \frac{\partial V}{\partial x_2} \\ &= -\frac{dJ}{dt}(\mathbf{x}) \end{aligned}$$

Thus, $\frac{dJ}{dt} = 0$ if and only if the angular derivative of V is zero. □

As a result of the conservation of angular momentum, we thus get Kepler's Second Law: if $\mathbf{x}(t)$ is the trajectory of a particle under the influence of a force with rotationally invariant potential, then the area swept out by $\mathbf{x}(t)$ between $t = a$ and $t = b$ is $\frac{b-a}{2m} J$.

In \mathbb{R}^3 , \mathbf{J} is a vector given by $\mathbf{x} \times \mathbf{p}$. Meanwhile, in \mathbb{R}^n , the angular momentum is a skew-symmetric matrix defined by

$$J_{jk} = x_j p_k - x_k p_j.$$

The total angular momentum of a system of N particles in \mathbb{R}^n is given by \mathbf{J} with entries

$$J_{jk} = \sum_{l=1}^N (x_j^l p_k^l - x_k^l p_j^l).$$

Similar to the case of linear momentum, angular momentum is constant in the presence of a conservative force if and only if the potential function V is rotationally invariant. That is,

$$V(R\mathbf{x}^1, R\mathbf{x}^2, \dots, R\mathbf{x}^N) = V(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N)$$

for all rotation matrices R .

Hamiltonian Mechanics

The Hamiltonian is the total energy function, but formulated in terms of position and momentum rather than position and velocity. If a particle in \mathbb{R}^n has the usual energy function, we write

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2m} \sum_{j=1}^n p_j^2 + V(\mathbf{x}),$$

where $p_j = m_j \dot{x}_j$. Observe that the equations of motion can be written as

$$\begin{aligned} \frac{dx_j}{dt} &= \frac{\partial H}{\partial p_j} \\ \frac{dp_j}{dt} &= -\frac{\partial H}{\partial x_j}. \end{aligned}$$

In the basic formulation, we can see that the first equation is just $\dot{x}_j = p_j/m$, and $\dot{p}_j = F_j$. The equations of motion written with Hamiltonians are known as Hamilton's equations.

Poisson Bracket

Let f and g be two smooth functions on \mathbb{R}^{2n} , with each element of \mathbb{R}^{2n} being denoted by (\mathbf{x}, \mathbf{p}) . The Poisson bracket of f and g is equal to

$$\{f, g\}(\mathbf{x}, \mathbf{p}) = \sum_{j=1}^n \left(\frac{\partial f}{\partial x_j} \frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial x_j} \right).$$

The Poisson bracket satisfies the following properties:

- Linearity: $\{f, g + ch\} = \{f, g\} + c\{f, h\}$
- Antisymmetry: $\{g, f\} = -\{f, g\}$
- Product Rule: $\{f, gh\} = \{f, g\}h + g\{f, h\}$
- Jacobi Identity: $\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0$.

It can be easily verified that the following Poisson bracket relations hold:

$$\begin{aligned} \{x_j, x_k\} &= 0 \\ \{p_j, p_k\} &= 0 \\ \{x_j, p_k\} &= \delta_{jk}, \end{aligned}$$

where δ_{jk} denotes the Kronecker delta function.

Proposition (Solutions of Hamilton's Equations). *If $(\mathbf{x}(t), \mathbf{p}(t))$ is a solution to Hamilton's Equations, then for any smooth f on \mathbb{R}^{2n} , we have*

$$\frac{df}{dt} = \{f, H\}.$$

Proof.

$$\begin{aligned} \frac{df}{dt} &= \sum_{j=1}^n \left(\frac{\partial f}{\partial x_j} \frac{dx_j}{dt} + \frac{\partial f}{\partial p_j} \frac{dp_j}{dt} \right) \\ &= \sum_{j=1}^n \left(\frac{\partial f}{\partial x_j} \frac{\partial H}{\partial p_j} + \frac{\partial f}{\partial p_j} \left(-\frac{\partial H}{\partial x_j} \right) \right) \\ &= \{f, H\}. \end{aligned}$$

□

Conserved Quantities

Let $f \in C^1(\mathbb{R}^{2n})$ be called conserved if $f(\mathbf{x}(t), \mathbf{p}(t))$ is independent of t for each solution to Hamilton's equation. Then, f is a conserved quantity if and only if

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

Note that H is also a conserved quantity.

Flow and Liouville's Theorem

Solving Hamilton's equations on \mathbb{R}^{2n} yields a flow Φ_t^x with $\Phi_t(\mathbf{x}, \mathbf{p})$ equal to the solution at time t with initial condition (\mathbf{x}, \mathbf{p}) .

The Φ_t aren't necessarily defined on all of \mathbb{R}^{2n} , but if Φ_t is defined on \mathbb{R}^{2n} for all t , then we say Φ_t is complete.

Proposition (Liouville^{xi}). *The Hamiltonian flow preserves the $2n$ -dimensional measure.*

$$dx_1 dx_2 \cdots dx_n dp_1 dp_2 \cdots dp_n.$$

More specifically, if E is a measurable subset of the domain of Φ_t , then $\mu(\Phi_t(E)) = \mu(E)$.

Proof. Hamilton's equations can be written as

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ p_1 \\ \vdots \\ p_n \end{bmatrix} = \begin{bmatrix} \frac{\partial H}{\partial p_1} \\ \vdots \\ \frac{\partial H}{\partial p_n} \\ -\frac{\partial H}{\partial x_1} \\ \vdots \\ -\frac{\partial H}{\partial x_n} \end{bmatrix}.$$

Hamilton's equations describe the flow along the vector field appearing on the right side — by a result in vector calculus,^{xii} the flow preserves the $2n$ -dimensional area measure if and only if the divergence of the vector field is zero.

$$\begin{aligned} \nabla \cdot \begin{bmatrix} \frac{\partial H}{\partial p_1} \\ \vdots \\ \frac{\partial H}{\partial p_n} \\ -\frac{\partial H}{\partial x_1} \\ \vdots \\ -\frac{\partial H}{\partial x_n} \end{bmatrix} &= \sum_{k=1}^n \frac{\partial}{\partial x_k} \frac{\partial H}{\partial p_k} - \frac{\partial}{\partial p_k} \frac{\partial H}{\partial x_k} \\ &= \sum_{k=1}^n \frac{\partial^2 H}{\partial x_k \partial p_k} - \frac{\partial^2 H}{\partial p_k \partial x_k} \\ &= 0 \end{aligned}$$

□

The condition of zero divergence is equivalent to Φ_t preserving a particular symplectic form ω defined by

$$\omega((\mathbf{x}, \mathbf{p}), (\mathbf{x}', \mathbf{p}')) = \mathbf{x} \cdot \mathbf{p}' - \mathbf{p} \cdot \mathbf{x}',$$

^xthe Φ_t are diffeomorphisms, or differentiable isomorphisms with differentiable inverses

^{xi}not from complex analysis

^{xii}Author's Note: I do not know this result yet, but hopefully I will soon!

meaning that for any t and any $(\mathbf{x}, \mathbf{p}) \in \mathbb{R}^{2n}$, the partial derivatives of Φ_t preserves ω .

Alternatively, this is equivalent to Φ_t preserving Poisson brackets:

$$\{f \circ \Phi_t, g \circ \Phi_t\} = \{f, g\} \circ \Phi_t.$$

Thus, Φ_t is an example of a symplectomorphism.

Hamiltonian Flow and Hamiltonian Generators

We say $f \in C^1(\mathbb{R}^{2n})$ is the Hamiltonian generator of the flow that results from solving Hamilton's equations with f substituted for H :

$$\begin{aligned} \frac{dx_j}{dt} &= \frac{\partial f}{\partial p_j} \\ \frac{dp_j}{dt} &= -\frac{\partial f}{\partial x_j}. \end{aligned}$$

It is possible to see that

$$f_a(\mathbf{x}, \mathbf{p}) = \mathbf{a} \cdot \mathbf{p}$$

yields the flow

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{x}_0 + t\mathbf{a} \\ \mathbf{p}(t) &= \mathbf{p}_0, \end{aligned}$$

and

$$g_b(\mathbf{x}, \mathbf{p}) = \mathbf{b} \cdot \mathbf{x}$$

yields the flow

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{x}_0 \\ \mathbf{p}(t) &= \mathbf{p}_0 - t\mathbf{b}. \end{aligned}$$

Thus, the Hamiltonian flow generated by momentum yields translation in position, and the Hamiltonian flow generated by position yields translation in momentum.

In this light, we can think of *the* Hamiltonian as the Hamiltonian generator that yields time evolution. Other Hamiltonian generators represent some other family of symmetries of the system.

Proposition (Hamiltonian Flow generated by Angular Momentum). *For a particle moving in \mathbb{R}^2 , the Hamiltonian flow generated by*

$$J(\mathbf{x}, \mathbf{p}) = x_1 p_2 - x_2 p_1$$

consists of simultaneous rotations of \mathbf{x} and \mathbf{p} .

$$\begin{aligned} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} &= \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} \\ \begin{bmatrix} p_1(t) \\ p_2(t) \end{bmatrix} &= \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} p_1(0) \\ p_2(0) \end{bmatrix}. \end{aligned}$$

Proof. Plugging J Hamilton's equations, we get

$$\begin{aligned}\frac{dx_1}{dt} &= \frac{\partial J}{\partial p_1} = -x_2 \\ \frac{dp_1}{dt} &= -\frac{\partial J}{\partial x_1} = -p_2 \\ \frac{dx_2}{dt} &= \frac{\partial J}{\partial p_2} = x_1 \\ \frac{dp_2}{dt} &= -\frac{\partial J}{\partial x_2} = p_1.\end{aligned}$$

□

It's important to note that the parameter t in the Hamiltonian flow for J is the rotation, not time. That is, J is the Hamiltonian generator of rotations.

If f is any smooth function, it is the case that the time derivative of any other function g along the Hamiltonian flow generated by f is $\frac{dg}{dt} = \{g, f\}$. In particular, the derivative of H along the flow generated by f is $\{H, f\}$, meaning that f is constant along the flow generated by H if and only if $\{f, H\} = 0$, which is true if and only if H is constant along the flow generated by H .

Thus, we find that f is conserved for solutions of Hamilton's equations if and only if H is invariant under the Hamiltonian flow generated by f . Of particular note, we find that J is conserved if and only if H is invariant under rotations of \mathbf{x} and \mathbf{p} .^{xiii}

Introduction to Quantum Mechanics

Observable quantities such as position and momentum in quantum mechanics are represented by operators on a complex-valued Hilbert space (an inner product space that is complete with respect to the induced metric) — specifically, these quantities are *unbounded* linear operators.

In physics, the inner product is linear in the second factor and conjugate linear in the first factor.^{xiv}

$$\begin{aligned}\langle \phi, \lambda \psi \rangle &= \lambda \langle \phi, \psi \rangle \\ \langle \lambda \phi, \psi \rangle &= \overline{\lambda} \langle \phi, \psi \rangle.\end{aligned}$$

A Taste of Operator Theory

A linear operator $A : \mathbf{H} \rightarrow \mathbf{H}$ is bounded if it has finite operator norm:^{xv}

$$\sup_{\|\psi\| \leq 1} \|A\psi\| < \infty.$$

For each bounded operator A , there exists a unique bounded operator A^* such that $\langle \phi, A\psi \rangle = \langle A^*\phi, \psi \rangle$. The existence of A^* follows from the Riesz representation theorem.

A bounded operator is said to be self-adjoint if $A^* = A$. Self-adjoint operators are nice for a variety of reasons, and as a result we desire for our operators in quantum mechanics to be self-adjoint. However, this

^{xiii}There is another section on Kepler's Laws in the chapter on Classical Mechanics that I didn't really read in depth. I might include it in the future.

^{xiv}Notice that this is different with math, where the inner product is linear in the first factor and conjugate linear in the second factor.

^{xv}I'm using more operator-theoretic language than the book uses because I'm ~~pretentious~~ a mathematician, not a physicist.

brings a significant problem — unbounded self-adjoint operators are not necessarily defined on \mathbf{H} .

We define unbounded operators as linear operators defined on a dense subspace of \mathbf{H} :

$$A : \text{Dom}(A) \subseteq \mathbf{H} \rightarrow \mathbf{H}$$

subject to

$$\overline{\text{Dom}(A)} = \mathbf{H}.$$

In addition to the domain of A not necessarily being equal to \mathbf{H} , the linear functional $\langle \phi, A \cdot \rangle$ is not necessarily bounded (meaning we cannot use the Riesz representation theorem to find $A^* \phi$). The adjoint of A , as a result, will be defined on a subspace of \mathbf{H} .

A vector $\phi \in \mathbf{H}$ is said to belong to the domain $\text{Dom}(A^*)$ if the linear functional $\langle \phi, A \cdot \rangle$ on $\text{Dom}(A)$ is bounded. Then, we define A^* to be the unique vector χ such that $\langle \chi, \psi \rangle = \langle \phi, A\psi \rangle$ for all $\psi \in \text{Dom}(A)$.

Having defined adjoints of an unbounded operator, we can now commit to defining self-adjoint operators. The operator A is symmetric if $\langle \phi, A\psi \rangle = \langle A\phi, \psi \rangle$ — a symmetric operator is self-adjoint if $\text{Dom}(A) = \text{Dom}(A^*)$ and $A^* \phi = A\phi$ for all $\phi \in \text{Dom}(A)$. Finally, A is essentially self-adjoint if the closure of the graph of A in $\mathbf{H} \times \mathbf{H}$ is self-adjoint.

In sum, A is self-adjoint if A and A^* are the same operator with the same domain, more or less.

Definition (Properties of Symmetric Operators). Let A be a symmetric operator on \mathbf{H} . Then, the following hold:

- (1) For all $\psi \in \text{Dom}(A)$, the quantity $\langle \psi, A\psi \rangle$ is real. More generally, if $\psi, A\psi, \dots, A^{m-1}\psi$ belong to $\text{Dom}(A)$, then $\langle \psi, A^m \psi \rangle$ is real.
- (2) Suppose λ is an eigenvector for A . Then, $\lambda \in \mathbb{R}$.

Proof. (1) Since A is symmetric,

$$\begin{aligned} \langle \psi, A\psi \rangle &= \langle A\psi, \psi \rangle \\ &= \overline{\langle \psi, A\psi \rangle}, \end{aligned}$$

for all $\psi \in \text{Dom}(A)$. Similarly, if $\psi, A\psi, \dots, A^{m-1}\psi \in \text{Dom}(A)$, then we use the symmetry of A to show that

$$\begin{aligned} \langle \psi, A^m \psi \rangle &= \langle A^m \psi, \psi \rangle \\ &= \overline{\langle \psi, A^m \psi \rangle}. \end{aligned}$$

- (2) If ψ is an eigenvector for A with eigenvalue λ , then

$$\begin{aligned} \lambda \langle \psi, \psi \rangle &= \langle \psi, A\psi \rangle \\ &= \langle A\psi, \psi \rangle \\ &= \overline{\lambda} \langle \psi, \psi \rangle. \end{aligned}$$

Since ψ is nonzero by definition, it must be the case that $\lambda = \overline{\lambda}$.

□

In physical terms, $\langle \psi, A\psi \rangle$ represents the expected value for measurements of A in the state ψ , with λ representing a possible value of this measurement. This is why we want both numbers to be real.

A self-adjoint A allows us to use the spectral theorem to assign each $\psi \in \mathbf{H}$ a probability measure on the real numbers.

Position and Momentum Operators

Consider a particle moving along the real line with wave function $\psi : \mathbb{R} \rightarrow \mathbb{C}$. Although ψ will evolve over time, let the particle be fixed in time for now.

We want to define ψ to be a unit vector in $L^2(\mathbb{R})$, meaning

$$\int_{\mathbb{R}} |\psi(x)|^2 dx = 1.$$

The probability that the position of the particle belongs to some $E \subseteq \mathbb{R}$ is

$$\int_E |\psi(x)|^2 dx,$$

where E is necessarily a Lebesgue-measurable set.

The expectation value of the position is thus

$$E(x) = \int_{\mathbb{R}} x |\psi(x)|^2 dx,$$

assuming the convergence of the integral, and the m th moment of the position is calculated as

$$E(x^m) = \int_{\mathbb{R}} x^m |\psi(x)|^2 dx,$$

again assuming convergence of the integral.

Definition (Position Operator). The position operator is defined as $X = M_x$, meaning $(X\psi)(x) = x\psi(x)$. With this in mind, we can then see that

$$E(x) = \langle \psi, X\psi \rangle$$

under the standard inner product on $L^2(\mathbb{R})$. The expectation value of X for the state ψ is denoted $\langle X \rangle_{\psi} := \langle \psi, X\psi \rangle$.^{xvi}

The higher moments of position are similarly defined:

$$E(x^m) = \langle \psi, X^m \psi \rangle,$$

where X^m denotes m -degree composition of X .

Since X is an unbounded linear operator, it is not necessarily the case that $X\psi \in L^2(\mathbb{R})$ if $\psi \in L^2(\mathbb{R})$.^{xvii}

Momentum is encoded in the oscillations of the wave function — the de Broglie hypothesis provides a special relationship between the frequency of oscillation (as a function of position at a fixed time) and the momentum.

Proposition (De Broglie Hypothesis). *If the wave function of a particle has spatial frequency k , then the momentum p of the particle is*

$$p = \hbar k,$$

where \hbar denotes Planck's constant.

^{xvi}I don't like this notation either.

^{xvii}Other famous examples of unbounded linear operators on Hilbert spaces include the derivative operator on $A^2(\mathbb{D})$, the space of holomorphic functions on the complex unit disc. I'm doing research on properties of variations of this space.

To be more precise, the de Broglie hypothesis applies to wave functions of the form $\psi(x) = e^{ikx}$, which represents particles that have momentum $p = \hbar k$. Let's develop this a bit further.

Since e^{ikx} is not square integrable over \mathbb{R} , we instead move to the circle, where ψ has period 2π over \mathbb{R} , and

$$\int_0^{2\pi} |\psi(x)|^2 dx = 1.$$

For any integer k , the normalized wave function $\frac{e^{ikx}}{\sqrt{2\pi}}$ will represent the particle with momentum $p = \hbar k$. This momentum value is definite; that is, $p = \hbar k$ with probability 1 for a particle with wave function $\frac{e^{ikx}}{\sqrt{2\pi}}$.

Of note, the functions $\left\{ \frac{e^{ikx}}{\sqrt{2\pi}} \right\}$ for $k \in \mathbb{Z}$ form an orthonormal basis for the Hilbert space of square integrable functions with period 2π .

The wave functions for particles on a circle are thus all of the form

$$\psi(x) = \sum_{k=-\infty}^{\infty} a_k \frac{e^{ikx}}{\sqrt{2\pi}},$$

where the sum is convergent in $L^2([0, 2\pi])$.^{xviii} If ψ is a unit vector, then^{xix}

$$\begin{aligned} \|\psi\|_{L^2([0, 2\pi])}^2 &= \sum_{k=-\infty}^{\infty} |a_k|^2 \\ &= 1. \end{aligned}$$

For a particle with wave function ψ expressed as a Fourier series, the momentum isn't definite. We will have to consider that measurement will yield one of the values of $\hbar k$ with probability $|a_k|^2$.

$$E(p) = \sum_{k=-\infty}^{\infty} \hbar k |a_k|^2,$$

with higher moments defined by

$$E(p^m) = \sum_{k=-\infty}^{\infty} (\hbar k)^m |a_k|^2,$$

assuming absolute convergence.

Definition (Momentum Operator). To encode P , our momentum operator, such that for the wave function $\psi \in L^2([0, 2\pi])$, $E(p^m) = \langle \psi, P^m \psi \rangle$, we must have P satisfying

$$P e^{ikx} = \hbar k e^{ikx}.$$

Thus, we would assume that

$$P = -i\hbar \frac{d}{dx}.$$

Even on the real line, we would still expect $P = -i\hbar \frac{d}{dx}$; though e^{ikx} is not square integrable, we can represent any wave function $\psi \in L^2(\mathbb{R})$ as an integral with the Fourier transform:

$$\begin{aligned} \psi(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \hat{\psi}(k) dk, \\ \hat{\psi}(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \psi(x) dx. \end{aligned}$$

^{xviii}You may recognize these as Fourier series.

^{xix}We use Parseval's identity to relate the L^2 norm of ψ to the ℓ^2 norm of $\{a_k\}_{k \in \mathbb{Z}}$. Try proving it yourself! Hint: use the Pythagorean theorem.

Plancherel's theorem^{xx} states that the Fourier transform is unitary — that is,

$$\int_{-\infty}^{\infty} |\psi(x)|^2 dx = \int_{-\infty}^{\infty} |\hat{\psi}(k)|^2 dk = 1.$$

We can imagine $\hat{\psi}(k)$ as the probability density for the momentum of the particle^{xxi}.

Thus, we have defined the momentum operator as

$$P = -i\hbar \frac{d}{dx},$$

with, for sufficiently nice $\psi \in L^2(\mathbb{R})$

$$\begin{aligned} E(p^m) &= \langle \psi, P^m \psi \rangle \\ &= \int_{-\infty}^{\infty} (\hbar k)^m |\hat{\psi}(k)|^2 dk \end{aligned}$$

for all positive integers m .

Proposition (Commutator of Position and Momentum).

$$[X, P] = i\hbar I,$$

a relation known as the canonical commutation relation.

Proof.

$$\begin{aligned} PX\psi &= -i\hbar \frac{d}{dx} (x\psi(x)) \\ &= -i\hbar \psi(x) - i\hbar x \frac{d\psi}{dx} \\ &= -i\hbar I\psi + XP\psi \end{aligned}$$

□

Remark. Note the parallel between $\{x, p\} = 1$ in the Poisson bracket and $[X, P] = i\hbar I$ in the commutator.

Proposition (Symmetry of Position and Momentum Operators). For all sufficiently nice ϕ and ψ in $L^2(\mathbb{R})$,

$$\begin{aligned} \langle \phi, X\psi \rangle &= \langle X\phi, \psi \rangle \\ \langle \phi, P\psi \rangle &= \langle P\phi, \psi \rangle. \end{aligned}$$

Proof. Let $\phi, \psi \in L^2(\mathbb{R})$ with $x\psi(x), x\phi(x) \in L^2(\mathbb{R})$. Then, since $x \in \mathbb{R}$,

$$\int_{-\infty}^{\infty} \overline{\phi(x)} x\psi(x) dx = \int_{-\infty}^{\infty} \overline{x\phi(x)} \psi(x) dx,$$

with both integrals convergent.

Meanwhile, assume ϕ and ψ are continuously differentiable, vanish at $\pm\infty$, and $\phi, \psi, \frac{d\phi}{dx}, \frac{d\psi}{dx} \in L^2(\mathbb{R})$. Note that $\frac{d\bar{\phi}}{dx} = \overline{\frac{d\phi}{dx}}$. Then, integrating by parts,

$$-i\hbar \int_{-n}^n \overline{\phi(x)} \frac{d\psi}{dx} dx = -i\hbar \overline{\phi(x)} \psi(x) \Big|_{-n}^n + i\hbar \int_{-n}^n \frac{d\bar{\phi}}{dx} \psi(x) dx,$$

^{xx}also known as Parseval's theorem when applied to Fourier series rather than the Fourier transform

^{xxi}well, for p/\hbar , but that's basically the same

meaning

$$\begin{aligned} \int_{-\infty}^{\infty} \overline{\phi(x)} \left(-i\hbar \frac{d\psi}{dx} \right) dx &= i\hbar \int_{-\infty}^{\infty} \overline{\frac{d\phi}{dx}} \psi(x) dx \\ &= \int_{-\infty}^{\infty} \overline{\left(-i\hbar \frac{d\phi}{dx} \right)} \psi(x) dx. \end{aligned}$$

□

Thus, we have shown that X and P are symmetric operators on certain dense subspaces of $L^2(\mathbb{R})$. We will have to wait until later to prove that X and P are essentially self-adjoint.

Kinematic Axioms of Quantum Mechanics

These aren't really axioms^{xxii} (as in, first principles from which quantum mechanics is derived), but they're essential principles of quantum mechanics. For ease of use, I'm labeling them with my personal interpretation of their content (rather than the numbering system used in the book).

Axiom (Principle of Representation). *The state of a quantum mechanical system is represented by $\psi \in \mathbf{H}$ for some Hilbert space \mathbf{H} . If ψ_1 and ψ_2 are two unit vectors in \mathbf{H} with $\psi_1 = c\psi_2$ for some $c \in \mathbb{C}$, ψ_1 and ψ_2 represent the same state.*

Each ψ represents a pure state — there are mixed states, but those will be discussed later.

Axiom (Principle of Correspondence). *To each real-valued f on the classical phase space, there is an unbounded, self-adjoint operator \hat{f} on the Hilbert space.*

For a particle in \mathbb{R}^1 , the phase space in \mathbb{R}^2 is represented by (x, p) for x position and p momentum; the analogue to the classical phase space in quantum mechanics is $L^2(\mathbb{R})$ with the position function $f(x, p) = x$ being M_x (the position operator), and the momentum function $g(x, p) = p$ being $P = -i\hbar \frac{d}{dx}$ (the momentum operator).

Axiom (Principle of Measurement). *If a system is in a state represented by $\psi \in \mathbf{H}$, then the measurement for the m th moment of f satisfies*

$$E(f^m) = \langle \psi, (\hat{f})^m \psi \rangle.$$

In particular, the measurement for f is

$$E(f) = \langle \psi, \hat{f}\psi \rangle.$$

Remark. Note that in the quantum system, we are measuring the classical f . However, rather than a definitive value, we need to find the expectation of the operator \hat{f} using the probabilities derived from ψ .

Since \hat{f} is self-adjoint, $E(f^m)$ is real, we can construct a probability measure $\mu_{A,\psi}$ that is the probability distribution for measurements of A in the state ψ .

Proposition (Eigenvectors). *If a quantum system is in a state described by $\psi \in \mathbf{H}$, and $\hat{f}\psi = \lambda\psi$ for some $\lambda \in \mathbb{R}$ ^{xxiii}, then*

$$E(f^m) = \lambda^m.$$

The aforementioned probability measure consistent with this condition is δ_λ (where f has value λ with probability 1).

^{xxii}Physicists once again caught stealing the valor of mathematicians.

^{xxiii}Recall that the eigenvalues of self-adjoint operators are real.

Essentially, if ψ is an eigenvector for \hat{f} , then measurements for f are determined. Thus, we would need to find the probability measure such that

$$\int_{\mathbb{R}} x^m d\mu = \lambda^m,$$

which only works if $\mu = \delta_\lambda$.

Note that if ψ is a linear combination of eigenvectors for \hat{f} , then the measurements of f are not deterministic.^{xxiv}

Example. Suppose \hat{f} has an orthonormal basis $\{e_j\}$ with distinct real eigenvalues λ_j . Let ψ be a unit vector in \mathbf{H} with the expansion

$$\psi = \sum_{j=1}^{\infty} a_j e_j.$$

Then, the measurement in the state ψ of f will necessarily be some value of λ_j with probability $|a_j|^2$.

Axiom (Principle of Wave Function Collapse). *Suppose a quantum system begins in a state ψ , and a measurement of the observable f is performed. If the measurement is $\lambda \in \mathbb{R}$, then immediately after the measurement, the system will be in a state ψ' such that*

$$\hat{f}\psi' = \lambda\psi',$$

where \hat{f} is the self-adjoint operator representation of f .

Remark. *Since ψ' is an eigenvector of \hat{f} , this means that a second measurement of f occurring immediately after the first measurement will yield the value λ with probability 1 (see the above proposition about eigenvectors).*

The wave function collapse principle only applies to measurements very shortly after the initial measurement; the system will still evolve over time.

Hydrogen Atom Measurements and Uncertainty

The Hamiltonian operator \hat{H} for the hydrogen atom has eigenvalues of the form

$$-\frac{R}{n^2},$$

where R is the Rydberg constant and $n \in \mathbb{N}$ denotes a state. The negative value of these eigenvalues is important; they denote that the electron is bound to the nucleus.

If an electron is placed into a state with energy $-\frac{R}{n_1^2}$ (where $n_1 > 1$), it will eventually decay into a state with lower energy, $-\frac{R}{n_2^2}$, with $n_2 < n_1$. The process of decay releases a photon with energy

$$E_{\text{photon}} = \frac{R}{n_2^2} - \frac{R}{n_1^2}.$$

The frequency of a photon is proportional to its energy; by observing the frequency of the photon, we can determine the change in energy (and thus the values of n_1 and n_2).

A bound state for the hydrogen atom will be a linear combination of the eigenvectors of \hat{H} with eigenvalues of the aforementioned form.

^{xxiv}The spectral decomposition is very cool.

Just as the standard deviation of a random variable Y is found by $\sigma^2 = E(Y^2) - (E(Y))^2$, we can find the *uncertainty* of a self-adjoint operator A , defined as

$$(\Delta_\psi A)^2 = \langle A^2 \rangle_\psi - (\langle A \rangle_\psi)^2,$$

recalling that $\langle A \rangle_\psi = \langle \psi, A\psi \rangle$.

For a single A , we can find ψ such that $\Delta_\psi A < \varepsilon$ for any $\varepsilon > 0$; however, if A and B do not commute, then $\Delta_\psi A$ and $\Delta_\psi B$ cannot both be made arbitrarily small.

In particular, we know that X and P do not commute; thus yields the famous Heisenberg Uncertainty Principle:

$$(\Delta_\psi X)(\Delta_\psi P) \geq \frac{\hbar}{2}$$

for all ψ such that $\Delta_\psi X$ and $\Delta_\psi P$ are defined.

The Schrödinger Equation

In the previous section, we considered the wave function ψ at a fixed time; however, time is not stationary^{xxv} and neither ought our system be, meaning we need a way to evolve the quantum system over time.

In a classical system, the Hamiltonian is the generator of time evolution^{xxvi}— thus, by the Principle of Representation, there is an operator, \hat{H} , which is the Hamiltonian operator for the system.

We motivated the definition of the momentum operator via the de Broglie hypothesis, which stated that $p = \hbar k$, where k is the spatial frequency of the wave function. Similarly, we can look at the relation between energy and the temporal frequency of the wave function,

$$E = \hbar\omega,$$

to motivate time evolution.

Suppose that ψ_0 has energy E (so ψ_0 is an eigenvector for \hat{H}). Then, the wave function's time dependence should be solely based on the frequency $\omega = E/\hbar$, meaning that if the $t = 0$ state of the system is ψ_0 , the state of the system at any other time t should be

$$\begin{aligned}\psi(t) &= \psi_0 e^{-i\omega t} \\ &= \psi_0 e^{-it \frac{E}{\hbar}}.\end{aligned}$$

This $\psi(t)$ is a function to the differential equation

$$\begin{aligned}\frac{d\psi}{dt} &= -\frac{iE}{\hbar}\psi \\ &= \frac{E}{i\hbar}\psi\end{aligned}$$

with initial value ψ_0 . Finally, rewriting $\hat{H} = E$, we get the Schrödinger equation.

Axiom (Schrödinger Equation). *The time evolution of the wave function ψ in a quantum system is generated by*

$$\frac{d\psi}{dt} = \frac{1}{i\hbar} \hat{H}\psi,$$

where \hat{H} denotes the Hamiltonian operator.

^{xxv}Citation needed

^{xxvi}The Hamiltonian is also equal to the total energy of the system.

Proposition (Schrödinger Equation for Operators). *Let A be a self-adjoint operator on \mathbf{H} . Assuming particular domain conditions hold, then*

$$\frac{d}{dt} \langle A \rangle_{\psi(t)} = \left\langle \frac{1}{i\hbar} [A, \hat{H}] \right\rangle_{\psi(t)},$$

where $\langle A \rangle_{\psi} = \langle \psi, A\psi \rangle$ and $[A, B] = AB - BA$.

Remark. Remember that for a function f acting on the classical phase space, $\frac{df}{dt} = \{f, H\}$ acting along a solution to Hamilton's equations.

Proof. Let $\psi(t)$ be a solution to the Schrödinger equation, with $\psi(t) \in \text{Dom}(A) \cap \text{Dom}(\hat{H})$, $A\psi(t) \in \text{Dom}(\hat{H})$, and $H\psi(t) \in \text{Dom}(A)$. Then,

$$\begin{aligned} \frac{d}{dt} \langle \psi(t), A\psi(t) \rangle &= \left\langle \frac{d\psi}{dt}, A\psi \right\rangle + \left\langle \psi, A \frac{d\psi}{dt} \right\rangle \\ &= \frac{i}{\hbar} \langle \hat{H}\psi, A\psi \rangle - \frac{i}{\hbar} \langle \psi, A\hat{H}\psi \rangle \\ &= \frac{1}{i\hbar} \langle \psi, [A, \hat{H}]\psi \rangle. \end{aligned}$$

□

Remark. If $[A, \hat{H}] = 0$, then nothing interesting happens; thus, for operators to yield time evolution, we need them to not commute with \hat{H} .

For a particle moving in \mathbb{R} , we see that noncommutativity holds for X and P .

If $[A, \hat{H}] = 0$, then we see that $E(A^m)$ is independent of time, effectively meaning that A is conserved.

Proposition (Time Independence of Inner Products of Solutions to the Schrödinger Equation). *If $\phi(t)$ and $\psi(t)$ are solutions to the Schrödinger equation, then $\langle \phi(t), \psi(t) \rangle$ and $\|\psi(t)\|$ are independent of time.*

Proof. Using the product rule for inner products, we find

$$\begin{aligned} \frac{d}{dt} \langle \phi(t), \psi(t) \rangle &= \left\langle \frac{1}{i\hbar} \hat{H}\phi(t), \psi(t) \right\rangle + \left\langle \phi(t), \frac{1}{i\hbar} \hat{H}\psi(t) \right\rangle \\ &= -\frac{1}{i\hbar} \langle \hat{H}\phi(t), \psi(t) \rangle + \frac{1}{i\hbar} \langle \phi(t), \hat{H}\psi(t) \rangle \\ &= 0. \end{aligned}$$

□

Solving the Schrödinger Equation

The Schrödinger equation is an equation of the form

$$\frac{dv}{dt} = Av$$

for some linear operator A over a Hilbert space. Considering the finite-dimensional Hilbert space \mathbb{C}^n , A is an $n \times n$ matrix. The aforementioned equation thus has the solution

$$v(t) = v_0 e^{At},$$

where v_0 is the initial condition. If A is a diagonalizable matrix, we can calculate e^{At} using the eigenvectors.^{xxvii}

The Schrödinger equation simply replaces \mathbb{C}^n with \mathbf{H} and A with $\frac{1}{i\hbar} \hat{H}$.

^{xxvii}In particular, if A is normal, the eigenvectors are the columns of the unitary diagonalization UDU^* , and form an orthonormal basis.

Proposition (Solution to the Schrödinger Equation via Exponentiation). *If \hat{H} is a self-adjoint operator on \mathbf{H} , then*

$$\psi(t) = \psi_0 e^{-it\hat{H}/\hbar}$$

should be a solution to the Schrödinger equation given that $e^{-it\hat{H}/\hbar}$ is reasonably defined.

Remark. *Clearly,^{xxviii} $\psi(t)$ is a solution to the Schrödinger equation.*

If \hat{H} is a bounded operator, then $e^{-it\hat{H}/\hbar}$ can be defined by a convergent power series. However, this is rarely the case.

If \hat{H} is unbounded, then we will have to use the spectral theorem — the full spectral theorem will have to wait for a while, but we can examine the case of a point spectrum.

If $\{e_j\}$ is an orthonormal basis for \mathbf{H} consisting of eigenvectors of \hat{H} with $\hat{H}e_j = \lambda_j e_j$, then we define the exponential by requiring

$$e_j e^{-it\hat{H}/\hbar} = e_j e^{-it\lambda_j/\hbar}.$$

This construction makes $e^{-it\hat{H}/\hbar}$ a unitary operator,^{xxix} and thus bounded.

It is not necessarily true that every self-adjoint operator A has an orthonormal basis (even self-adjoint ones); nevertheless, the spectral theorem tells us that there is a decomposition of \mathbf{H} into generalized eigenspaces for A .

Unfortunately, this doesn't mean we have *solved* the Schrödinger equation. All it tells us is that $e^{-it\hat{H}/\hbar}$ is bounded, meaning it's defined for all $\psi_0 \in \mathbf{H}$; however, ψ_0 must belong to the domain of \hat{H} (which is not \mathbf{H} , but only a dense subset of it) in order to be a solution.

If ψ is an eigenvector of the Hamiltonian, then the equation

$$\hat{H}\psi = E\psi$$

for some $E \in \mathbb{R}$ is known as the time-independent Schrödinger equation. When we solve the time-independent Schrödinger equation, we are trying to find nonzero values of E and their corresponding ψ that satisfy the equation.

If ψ_0 is a solution to the time-independent Schrödinger equation, then $\psi(t) = \psi_0 e^{-itE/\hbar}$ is the solution to the time-dependent Schrödinger equation with initial condition ψ_0 . Since $\psi(t)$ is a constant multiple of ψ_0 , we say that the solution to the time-independent Schrödinger is a stationary state.

The Schrödinger Equation in \mathbb{R}

Consider a particle moving on the real line. The Hamiltonian for such a particle is written in the form of

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

where V denotes the potential. In the quantum case, we may then consider

$$\hat{H} = \frac{P^2}{2m} + V(X)$$

^{xxviii}My professor loves this phrase.

^{xxix}Notice that $|e^{-itE_j/\hbar}| = 1$.

to be the Hamiltonian operator. Here, $V(X)$ denotes multiplication by $V(x)$, where x is an input to ψ . Thus, we find

$$\hat{H}\psi(x) = -\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V(x)\psi(x).$$

An operator of the form above is known as the Schrödinger operator; the Hamiltonian operator refers to the operator that generates time-evolution, regardless of its form. The time-dependent Schrödinger equation thus becomes

$$\frac{\partial\psi(x, t)}{\partial t} = \frac{i\hbar}{2m} \frac{\partial^2\psi(x, t)}{\partial x^2} - \frac{i}{\hbar} V(x)\psi(x, t).$$

The time-dependent Schrödinger equation is thus a linear partial differential equation.^{xxx} For a particle on \mathbb{R} , the time-independent Schrödinger equation is a linear ordinary differential equation with nonconstant coefficients.^{xxxi}

Expected Position and Momentum for Solutions to the Schrödinger Equation

Proposition (Time Evolution of Position and Momentum). *Let $\psi(t)$ be a solution to the time-dependent Schrödinger equation with sufficiently nice V and sufficiently nice $\psi_0 = \psi(0)$. Then,*

$$\begin{aligned} \frac{d}{dt} \langle X \rangle_{\psi(t)} &= \frac{1}{m} \langle P \rangle_{\psi(t)} \\ \frac{d}{dt} \langle P \rangle_{\psi(t)} &= -\langle V'(X) \rangle_{\psi(t)}. \end{aligned}$$

Remark. *We require these assumptions to ensure that \hat{H} is a self-adjoint operator and that the domain conditions hold.*

Note that $-\langle V'(X) \rangle_{\psi(t)} \neq -V'(\langle X \rangle_{\psi(t)})$. This solution to the Schrödinger equation is not the quantum analogue to the solutions of Hamilton's equations.

Although expected position and expected momentum do not exactly follow the classical trajectories, if $\psi(x)$ is very closely concentrated about $x = x_0$, then the particle will approximately follow the classical trajectory.

The Heisenberg Approach

The previous approach (known as the Schrödinger approach) towards analyzing the quantum system comes from understanding how the wave functions evolve over time, with the operators remaining stationary. We will now examine an approach where the operators change over time with the states remaining stationary.

In the Heisenberg approach, every self-adjoint A evolves in time according to the operator differential equation^{xxxi}

$$\frac{dA(t)}{dt} = \frac{1}{i\hbar} [A(t), \hat{H}].$$

Note that since \hat{H} commutes with itself, it remains constant in time — this is akin to the classical Hamiltonian remaining constant along a solution to Hamilton's equations.

^{xxx}Unfortunately I haven't taken PDEs yet, so I can't really explain how to solve this.

^{xxxi}Neither have I taken ODEs, but I will be taking it in Fall 2024!

^{xxxi}I will call this the "modified Schrödinger equation."

The spectral theorem provides us a way to construct a family of unitary operators $e^{-it\hat{H}/\hbar}$ that computes time-evolution of states in the Schrödinger approach; similarly, we can find that the solutions in the operator differential equation are of the form

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

If ψ is the state of the system, then the expectation of $A(t)$ in the state is defined to be $\langle \psi, A(t)\psi \rangle$; then

$$\begin{aligned} \langle \psi, A(t)\psi \rangle &= \langle \psi, e^{it\hat{H}/\hbar} A e^{-it\hat{H}/\hbar} \psi \rangle \\ &= \langle e^{-it\hat{H}/\hbar} \psi, A e^{-it\hat{H}/\hbar} \psi \rangle \\ &= \langle \psi(t), A\psi(t) \rangle, \end{aligned}$$

which is the time-evolved state of the system. Since $\langle \psi(t), A\psi(t) \rangle$ is the expectation of the value of A in the state $\psi(t)$, it must be the case that the Heisenberg approach and the Schrödinger approach are equivalent in their physics.

Proposition (Hamiltonian in the Heisenberg Approach). *Let $\hat{H} = \frac{P^2}{2m} + V(X)$, where V is a bounded below polynomial. Then, for any $t \in \mathbb{R}$,*

$$\hat{H} = \frac{1}{2m} (P(t))^2 + V(X(t)).$$

Remark. Notice that the Hamiltonian is independent of time (since $[\hat{H}, \hat{H}] = 0$), even though P and X depend on time.

Lemma (Moments of Solutions). *Suppose A is a self-adjoint operator on \mathbf{H} and $A(t)$ is a solution to the modified Schrödinger equation with $A(0) = A$. Then, $(A(t))^m$ is also a solution to the modified Schrödinger equation. In other words, the time evolution of the m th power of A is the same as the m th power of the time evolution of A .*

Proof.

$$\begin{aligned} e^{it\hat{H}/\hbar} A^m e^{-it\hat{H}/\hbar} &= \prod_{k=1}^m \left(e^{it\hat{H}/\hbar} A e^{-it\hat{H}/\hbar} \right) \\ &= \left(e^{it\hat{H}/\hbar} A e^{-it\hat{H}/\hbar} \right)^m. \end{aligned}$$

□

Proposition (Analogue to Hamilton's Equations). *Suppose $\hat{H} = \frac{(P(t))^2}{2m} + V(X(t))$. Then,*

$$\begin{aligned} \frac{dX}{dt} &= \frac{1}{m} P(t) \\ \frac{dP}{dt} &= -V'(X(t)). \end{aligned}$$

Remark. Notice that it is the functions $X(t)$ and $P(t)$ that satisfy the analogue to Hamilton's equations; the expectation values satisfy

$$\begin{aligned} \frac{d}{dt} \langle X(t) \rangle_\psi &= \frac{1}{m} \langle P(t) \rangle_\psi \\ \frac{d}{dt} \langle P(t) \rangle_\psi &= -\langle V'(X(t)) \rangle_\psi. \end{aligned}$$

Particle in a Box

We want to solve the time-independent Schrödinger equation for the case of a particle that is confined to move in the interval $x \in [0, L]$. That is, we want to find all the eigenvectors and eigenvalues of $\hat{H}\psi = E\psi$.

The particle has potential 0 for $x \in [0, L]$ and very large C outside $[0, L]$. In the classical case, the particle has to have very high energy to escape the box; in the quantum case, if E is an eigenvalue satisfying $\hat{H}\psi = E\psi$, with $E \ll C$, then ψ rapidly decays outside the box. In general, we expect that the solutions to the time-independent Schrödinger equation vanish as we approach $x = 0$ or $x = L$.

Essentially, we are looking for $\psi \in C^2([0, L])$ such that

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi(x)$$

subject to the boundary conditions $\psi(0) = \psi(L) = 0$.

For $E > 0$, the solution to this differential equation is a linear combination of two complex exponentials (or, since \hat{H} is symmetric, a linear combination of sines and cosines):

$$\psi(x) = a \sin\left(\frac{\sqrt{2mE}}{\hbar}x\right) + b \cos\left(\frac{\sqrt{2mE}}{\hbar}x\right).$$

Implementing the boundary condition $\psi(0) = 0$ yields $b = 0$; imposing $\psi(L) = 0$ *could* give us $a = 0$, but that would be boring. Instead, we want to find solutions such that

$$\sin\left(\frac{\sqrt{2mE}}{\hbar}L\right) = 0$$

that are not identically zero. This forces E to be of the form

$$E_j = \frac{j^2 \pi^2 \hbar^2}{2mL^2},$$

with $j \in \mathbb{N}$. It is easy to verify that for $E \leq 0$, the only solutions are where $\psi = 0$.

Proposition (Eigenvectors of the Solution). *The following eigenvectors are a solution to the Schrödinger of the particle in a box satisfying $\psi(0) = \psi(L) = 0$:*

$$\psi_j = \sqrt{\frac{2}{L}} \sin\left(\frac{j\pi x}{L}\right).$$

The ψ_j correspond to each value of E_j , and form an orthonormal basis for $L^2([0, L])$.

Proof. It has already been verified that the ψ_j are eigenvectors with eigenvalue E_j . To verify that the ψ_j are orthonormal, we see that

$$\begin{aligned} \frac{2}{L} \int_0^L \sin^2\left(\frac{j\pi x}{L}\right) dx &= \frac{2}{L} \left(\frac{L}{2}\right) \\ &= 1, \end{aligned}$$

and

$$\frac{2}{L} \int_0^L \sin\left(\frac{j_1\pi x}{L}\right) \sin\left(\frac{j_2\pi x}{L}\right) dx = 0.$$

To verify that it is a basis, we use the fact that the ψ_j are a Fourier sine series for $L^2([0, L])$. □

The Hamiltonian operator for the particle in a box with $V = 0$ is

$$\hat{H}\psi = -\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2}.$$

This operator isn't defined over $L^2([0, L])$, but only a dense subspace. The domain of \hat{H} should be chosen such that \hat{H} is essentially self-adjoint, meaning

$$\langle \phi, \hat{H}\psi \rangle = \langle \hat{H}\phi, \psi \rangle$$

for $\phi, \psi \in \text{Dom}(\hat{H})$. For this to hold, ϕ and ψ must satisfy the boundary conditions necessary for the boundary terms in integration by parts to be 0.

n -Dimensional Single-Particle Quantum Mechanics

It is relatively straightforward to generalize from a quantum particle moving in \mathbb{R} to one moving in \mathbb{R}^n . The Hilbert space is $L^2(\mathbb{R}^n)$, and instead of one position operator, there are n :

$$X_j\psi(\mathbf{x}) = x_j\psi(\mathbf{x}).$$

Similarly, there are n momentum operators given by

$$P_j\psi(\mathbf{x}) = -i\hbar \frac{\partial \psi}{\partial x_j}.$$

As with \mathbb{R} , $[X_j, P_j] = i\hbar I$, but it is also the case that $[X_j, X_k] = 0$ and $[P_j, P_k] = 0$.

Proposition (Canonical Commutation Relations in n Dimensions). *The position and momentum operators satisfy*

$$\begin{aligned} \frac{1}{i\hbar} [X_j, X_k] &= 0 \\ \frac{1}{i\hbar} [P_j, P_k] &= 0 \\ \frac{1}{i\hbar} [X_j, P_k] &= \delta_{jk} I \end{aligned}$$

for all $1 \leq j, k \leq n$.

These are the counterparts to the classical Poisson bracket, where $\frac{1}{i\hbar} [\cdot, \cdot]$ plays the same role as the Poisson bracket.

The Hamiltonian operator in n dimensions is defined analogously to the classical Hamiltonian:

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

where $V(\mathbf{x})$ results from applying V to the family $\mathbf{x} = (x_1, \dots, x_n)$; we may also identify $V(\mathbf{x})$ with the operator of multiplication by $V(\mathbf{x})$, in which case we can write the Hamiltonian operator as

$$\hat{H}\psi(\mathbf{x}) = -\frac{\hbar^2}{2m} \Delta \psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}),$$

where $\Delta = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$.

Now that we are in multiple dimensions, we can now introduce the angular momentum operator, \hat{J}_{jk} :

$$\hat{J}_{jk} = X_j P_k - X_k P_j.$$

As in the classical case, $\hat{J}_{jk} = 0$ when $j = k$; when $j \neq k$, X_j and P_k commute, meaning the order of the factors in \hat{J}_{jk} is not important. In particular,

$$\hat{J}_{jk} + -i\hbar \left(x_j \frac{\partial}{\partial x_k} - x_k \frac{\partial}{\partial x_j} \right).$$

The operator $\left(x_j \frac{\partial}{\partial x_k} - x_k \frac{\partial}{\partial x_j} \right)$ as the angular derivative, $\frac{\partial}{\partial \theta}$ in the (x_j, x_k) plane.

When $n = 3$, we use the quantum version of the angular momentum vector;

$$\begin{aligned}\hat{J}_1 &= X_2 P_3 - X_3 P_2 \\ \hat{J}_2 &= X_3 P_1 - X_1 P_3 \\ \hat{J}_3 &= X_1 P_2 - X_2 P_1.\end{aligned}$$

When $n = 3$, every \hat{J}_{jk} is either $\hat{J}_1, \hat{J}_2, \hat{J}_3$ or their negative.

Systems of Multiple Particles

Suppose we have a system of N particles moving in \mathbb{R}^n . If the particles are of different types, then the Hilbert space is $L^2(\mathbb{R}^{nN})$. The wave function ψ is a function of $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N$, with \mathbf{x}^j a vector in \mathbb{R}^n .

If ψ is a unit vector in $L^2(\mathbb{R}^{nN})$, then $|\psi(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N)|^2$ is the joint probability distribution of the positions of the N particles.

We introduce position operators X_k^j to denote the k th component of the position of the j th particle, and similarly so with momentum operators P_k^j . The Hamiltonian operator is then

$$\hat{H}\psi(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N) = \sum_{j=1}^N \frac{\hbar}{2m_j} \Delta_j \psi(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N) + V(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^N) \psi(\mathbf{x}),$$

where m_j denotes the mass of the j th particle, and Δ_j denotes the Laplacian with respect to \mathbf{x}^j , fixing all other variables.

The Hilbert space for a composite system is taken to be the tensor product of the individual Hilbert spaces; there is a natural isomorphism^{xxxiii} between $L^2(\mathbb{R}^{nN})$ and the N copies of \mathbb{R}^n .

If the particles are identical, then things get more complicated — to start, we are supposed to believe that identical particles are indistinguishable, meaning if we exchange \mathbf{x}^1 and \mathbf{x}^2 with each other, then they represent the same state. Thus, we get that

$$\psi(\mathbf{x}^2, \mathbf{x}^1, \mathbf{x}^3, \dots, \mathbf{x}^N) = u\psi(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \dots, \mathbf{x}^N)$$

for some complex number u with $|u| = 1$.^{xxxiv} Exchanging again yields $u^2 = 1$, meaning $u = 1$ or $u = -1$.

Particles with $u = 1$ are known as bosons, and particles with $u = -1$ are known as fermions. The classification of a particle is determined by its spin — we also say that particles without spin are bosons.

For a collection of N identical spinless particles in \mathbb{R}^3 , we say that the Hilbert space is the symmetric subspace of $L^2(\mathbb{R}^{3N})$, or the space of all functions in $L^2(\mathbb{R}^{3N})$ that are invariant under variable permutation.

^{xxxiii}Category theory term

^{xxxiv}We used the fact that $\psi_1 = c\psi_2$ means ψ_1 and ψ_2 are equivalent states.

Dirac Notation

Dirac notation is physicists' preferred way to deal with ideas such as inner products and operators acting on vectors in Hilbert space.

Definition (Bras and Kets). A vector $\psi \in \mathbf{H}$ is referred to as a ket, denoted $|\psi\rangle$.

A continuous linear functional in \mathbf{H}^* is a bra; we use $\langle\phi|$ to denote the unique linear functional such that $\langle\phi|(\psi) = \langle\phi, \psi\rangle$. The existence of $\langle\phi|$ follows from the Riesz representation theorem.

We denote inner products as $\langle\phi | \psi\rangle$.

If A is an operator on \mathbf{H} , and ϕ is a vector in \mathbf{H} , we can form the linear functional $\langle\phi| A$ to denote the linear map $\psi \mapsto \langle\phi | A\psi\rangle$. Physicists like to write this as $\langle\phi| A|\psi\rangle$, to denote either the linear functional $\langle\phi| A$ applied to the vector $|\psi\rangle$ or the linear functional ϕ applied to the vector $A|\psi\rangle$.

Definition (Outer Products). For any ϕ and ψ in \mathbf{H} , the outer product, $|\phi\rangle\langle\psi|$ denotes the operator given by

$$(|\phi\rangle\langle\psi|)(\chi) = \langle\psi | \chi\rangle |\phi\rangle.$$

In math terms,

$$\chi \xrightarrow{|\phi\rangle\langle\psi|} \langle\psi, \chi\rangle \phi.$$

Notationally, if a family of vectors is labeled by indices, we write the ket vectors with purely the indices. For instance, if $\{\phi_n\} \subseteq \mathbf{H}$, we write $|n\rangle$ rather than $|\phi_n\rangle$.

If an operator \hat{H} has an orthonormal basis of eigenvectors ψ_n , the decomposition would be in the form

$$I = \sum_n |n\rangle\langle n|,$$

where $|n\rangle$ is a unit vector and $|n\rangle\langle n|$ denotes the orthogonal projection onto the one dimensional subspace spanned by $|n\rangle$.

Physicists also like to denote the complex conjugate of z as z^* and the adjoint of an operator as $A^{\dagger \text{xxxv}}$; self-adjoint operators are known as Hermitian operators.

To express the adjoint of an operator using Dirac notation, we say that if A is a bounded operator on \mathbf{H} , then A^\dagger is the unique bounded operator such that

$$\langle\psi| A = \langle A^\dagger \psi|.$$

Physicists tend to define the irreducible canonical commutation relations between certain operators on a Hilbert space in order to define said Hilbert space — this is because the Stone–von Neumann theorem^{xxxvi} says that if certain conditions hold, the canonical commutation relations uniquely define the Hilbert space up to unitary equivalence.

Given this irreducible representation, and a vector $\psi \in \mathbf{H}$, the position wave function is defined by

$$\psi(x) = \langle x | \psi\rangle.$$

We can analogously define the momentum wave function.

^{xxxv}I don't like this either.

^{xxxvi}We will return to this in the future.

The Schrödinger Equation for a Free Particle

We want to solve the Schrödinger equation using a variety of methods for a “free” particle in \mathbb{R} — that is, a particle with identically zero potential. The free Schrödinger equation is thus

$$\frac{d\psi}{dt} = \frac{i\hbar}{2m} \frac{\partial^2 \psi}{\partial x^2}$$

subject to the initial condition $\psi(x, 0) = \psi_0(x)$. Before fully solving the equation, we begin by observing the time evolution of the expectation of the position and momentum operators:

$$\begin{aligned} \frac{d}{dt} \langle X \rangle_{\psi(t)} &= \frac{1}{m} \langle P \rangle_{\psi(t)} \\ \frac{d}{dt} \langle P \rangle_{\psi(t)} &= 0. \end{aligned}$$

Thus, the expectation of X is linear in time:

$$\begin{aligned} \langle X \rangle_{\psi(t)} &= \langle X \rangle_{\psi_0} + \frac{t}{m} \langle P \rangle_{\psi_0} \\ \langle P \rangle_{\psi(t)} &= \langle P \rangle_{\psi_0}. \end{aligned}$$

The free Schrödinger equation is a special case where the expectation of X and P exactly follows the classical solution.

Using the Fourier Transform to solve the Free Schrödinger Equation

We want to find solutions to the free Schrödinger equation of the form

$$\psi(x, t) = e^{i(kx - \omega(k)t)},$$

where k is the spatial frequency and $\omega(k)$ is the frequency in time for a given spatial frequency.^{xxxvii} Plugging this solution form into the free Schrödinger equation yields

$$\omega(k) = \frac{\hbar k^2}{2m}.$$

Formulae that express the frequency in time as a function of spatial frequency in a solution of some partial differential equation are called dispersion relations.

We can also see that the solution form can be written as

$$\psi(x, t) = e^{ik\left(x - \frac{\omega(k)}{k}t\right)},$$

meaning that time evolution shifts the initial function to the write by $\frac{\omega(k)}{k}t$, so $\psi(x, t)$ is moving to the right over time with speed $\frac{\omega(k)}{k}$.

Definition (Phase Velocity). The phase velocity of a particle with momentum $p = \hbar k$ is

$$\begin{aligned} \frac{\omega(k)}{k} &= \frac{\hbar k}{2m} \\ &= \frac{p}{2m}, \end{aligned}$$

which is half the velocity of a classical particle with momentum p .

^{xxxvii}I think my PDEs knowledge is lacking in order to explain why the solution is of this form.

The “real” velocity, $\frac{p}{m}$ is referred to as the group velocity, whereas the phase velocity originates in the pure exponential solution to the free Schrödinger equation.

Proposition (Solution to the Schrodinger Equation). Suppose ψ_0 is a Schwartz function. Let $\hat{\psi}_0$ denote the Fourier transform of ψ_0 , and define $\psi(x, t)$ by

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{\psi}_0(k) e^{i(kx - \omega(k)t)} dk.$$

Then, $\psi(x, t)$ solves the free Schrödinger equation.^{xxxviii}

Proof. Notice that, by differentiating under the integral sign, and remembering that $\omega(k) = \frac{\hbar k^2}{2m}$, we have^{xxxix}

$$\frac{d\psi}{dt} = \frac{i\hbar}{2m} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} -k^2 \hat{\psi}_0(k) e^{i(kx - \omega(k)t)} dk \right).$$

To show that the integral on the right is equal to $\frac{d^2\psi}{dx^2}$, we will need to show that the integral indeed converges. Taking the derivative of the e^{ikx} , we find that

$$\left| \frac{e^{ik(x+h)} - e^{ikx}}{h} \right| \leq |k|$$

for all $h > 0$. Thus, with $|k\hat{\psi}_0(k)|$ as our dominating function, the dominated convergence theorem states that

$$\lim_{h \rightarrow 0} \int_{-\infty}^{\infty} \hat{\psi}_0(k) \frac{e^{i(k(x+h) - \omega(k)t)} - e^{i(kx - \omega(k)t)}}{h} dk = \int_{-\infty}^{\infty} -ik \hat{\psi}_0(k) e^{i(kx - \omega(k)t)} dk.$$

Following a similar pattern, we find that

$$\begin{aligned} \frac{d\psi}{dt} &= \frac{i\hbar}{2m} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} -k^2 \hat{\psi}_0(k) e^{i(kx - \omega(k)t)} dk \right) \\ &= \frac{i\hbar}{2m} \frac{d^2\psi}{dx^2}. \end{aligned}$$

Finally, we use the Fourier inversion formula to show that $\psi(x, 0) = \psi_0(x)$.^{xl}

□

Proposition (Fourier Transform of $\psi(x, t)$). If $\psi(x, t)$ is defined as

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{\psi}_0(k) e^{i(kx - \omega(k)t)} dk,$$

then the Fourier transform of $\psi(x, t)$ with respect to x , keeping t fixed, is

$$\hat{\psi}(k, t) = \hat{\psi}_0(k) e^{-i\frac{\hbar k^2 t}{2m}}.$$

Proof. Rewriting $\psi(x, t)$ as

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \left(\hat{\psi}_0 e^{-i\omega(k)t} \right) dk,$$

we can see that the inner expression must be the Fourier transform of $\psi(x, t)$ with respect to x (seeing as the function e^{ikx} does not have any values of t inside it, and the Fourier transform is injective since it is unitary). Plugging in $\omega(k) = \frac{\hbar k^2}{2m}$, we get the final expression. □

^{xxxviii}The reason why physicists insert a factor of $\sqrt{2\pi}$ is because, rather than dealing with cycles per second, k is angular frequency (or radians per second). To normalize will require a factor of 2π at the conversion from x to k , but we also want the Fourier transform to remain unitary.

^{xxxix}I did not show the convergence of the integral here, but it follows a similar argument to the convergence of the derivative with respect to x .

^{xl}The book's proof is quite a bit more vague than this proof, and I'll probably have to go back and clarify a lot of this in the future.

From these two results, we can see that for any $\psi_0 \in L^2(\mathbb{R})$, we define $\psi(x, t)$ to be the unique element of $L^2(\mathbb{R})$ whose Fourier transform with respect to x is $\hat{\psi}_0(k)e^{-i\frac{\hbar k^2 t}{2m}}$. Notice here that ψ_0 and ψ need not be differentiable, but they will still satisfy a version of the Schrödinger equation.

Solving via Convolutions

Recall that

$$\mathcal{F}(\phi * \psi) = \hat{\phi}\hat{\psi},$$

where $*$ denotes convolution. In other words, the *inverse* Fourier transform of the function $\hat{\phi}\hat{\psi}$ is $\phi * \psi$ ^{xli}. Thus, by the property of the Fourier transform kernel, we ought to have

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

where

$$K_t = \frac{1}{\sqrt{2\pi}} \mathcal{F}^{-1} \left(e^{-i\frac{\hbar k^2 t}{2m}} \right).$$

However, unfortunately, $e^{-i\frac{\hbar k^2 t}{2m}}$ is not the Fourier transform of some function in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, since it does not tend to zero at ∞ . Fortunately, we can calculate $K_t(x)$ using regular techniques of integration.

$$\begin{aligned} K_t(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} e^{-i\frac{\hbar k^2 t}{2m}} dk \\ &= \sqrt{\frac{m}{i2\pi\hbar t}} e^{i\frac{mx^2}{2\hbar t}}, \end{aligned}$$

where the square root has positive real part. This is known as the fundamental solution of the free Schrödinger equation.

Theorem (Fundamental Solution to the Schrödinger Equation). *Let $\psi_0 \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$. Then, $\psi(x, t)$ defined by*

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{\psi}_0(k) e^{i(kx - \omega(k)t)} dk$$

is computed for all $t \neq 0$ as

$$\psi(x, t) = \sqrt{\frac{m}{2\pi i \hbar t}} \int_{-\infty}^{\infty} e^{i\frac{m}{2\hbar t}(x-y)^2} \psi_0(y) dy.$$

*This expression is equal to $\frac{1}{\sqrt{2\pi}} (K_t * \psi_0)$, where K_t is the kernel as described above.*

Proof. Let $E \subseteq \mathbb{R}$, and let $\mathbb{1}_E$ denote the indicator function of E . Then, $K_t \mathbb{1}_{[-n, n]}$ belongs to $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ for any $n \in \mathbb{Z}_{\geq 0}$. The Fourier transform thus tells us that

$$\mathcal{F}((K_t \mathbb{1}_{[-n, n]}) * \psi_0) = \sqrt{2\pi} \mathcal{F}(K_t \mathbb{1}_{[-n, n]}) \mathcal{F}(\psi_0).$$

Since $\psi_0 \in L^1(\mathbb{R})$, we can see that $K_t \mathbb{1}_{[-n, n]} * \psi_0$ converges pointwise to $K_t * \psi_0$; similarly, we can see that $K_t \mathbb{1}_{[-n, n]}$ is bounded by a constant function independent of n and

$$\mathcal{F}(K_t \mathbb{1}_{[-n, n]}) \xrightarrow{\text{p.w.}} \frac{1}{\sqrt{2\pi}} e^{-i\frac{\hbar k^2 t}{2m}}.$$

By dominated convergence, it is thus the case that

$$\mathcal{F}((K_t \mathbb{1}_{[-n, n]}) * \psi_0) \rightarrow \hat{\psi} e^{-i\frac{\hbar k^2 t}{2m}}.$$

Thus, $K_t \mathbb{1}_{[-n, n]} * \psi_0$ must converge in $L^2(\mathbb{R})$ (since the Fourier transform is unitary), and the L^2 limit must converge with the pointwise limit. □

^{xli}Barring a factor of $\sqrt{2\pi}$, of course.

A function tends to be the *fundamental* solution to an equation if it converges to the δ function as $t \rightarrow 0$; we can see that $\mathcal{F}(K_t)$ converges to $\frac{1}{\sqrt{2\pi}}$ as $t \rightarrow 0$, meaning $K_t \rightarrow \delta$.^{xlii}

Solving by Propagation of the Wave Function: First Approach

To come soon!^{xliii}

Introduction to Spectral Theory

A matrix A is self-adjoint if $A^* = A$, where A^* denotes the conjugate transpose; by a standard result in linear algebra, there exists an orthonormal basis $\{\mathbf{v}_j\}_{j=1}^n$ for \mathbb{C}^n and real numbers λ_j such that $A\mathbf{v}_j = \lambda_j\mathbf{v}_j$.

When we enter infinite dimensions, things get a little more complicated; there are many self-adjoint operators that have no eigenvalues or eigenvectors. For instance, if A is defined on $\mathbf{H} = L^2([0, 1])$ as

$$(A\psi)(x) = x\psi(x),$$

or $A = M_x$, then A is self-adjoint. However, the only way $x\psi(x) = \lambda\psi(x)$ works is if $\lambda = 0$ or $\psi(x)$ is only supported on $\{\lambda\}$ (which has measure zero). As a result, only the zero element of $L^2([0, 1])$ is an eigenvector for M_x .

It is true that physicists would say that M_x has the eigenvectors $\delta(x - \lambda)$, which do satisfy $x\delta(x - \lambda) = \lambda\delta(x - \lambda)$, but they aren't elements of $L^2([0, 1])$. However, understanding these *generalized* eigenvectors is a problem for the future.

What is valuable, though, is that we can create an orthonormal basis of eigenvectors for *compact* self-adjoint operators (the operators whose image of every bounded set has compact closure); however, in quantum mechanics, operators aren't compact, let alone bounded.

For instance, the position operator X acting on the indicator function $\mathbb{1}_{[n, n+1]}$ has operator norm at least n (and thus, is not bounded on $L^2(\mathbb{R})$).

It can be shown^{xliv} that if $\langle \phi, A\psi \rangle = \langle A\phi, \psi \rangle$ for all $\phi, \psi \in \mathbf{H}$, then A must be bounded; thus, our unbounded operators cannot be defined on \mathbf{H} , only a subspace.

Specifically, unbounded operators in quantum mechanics are densely defined. The adjoint of the unbounded operator A will be a different unbounded operator A^* with its own domain (where $\langle A^*\phi, \psi \rangle = \langle \phi, A\psi \rangle$ for $\phi, \psi \in \text{Dom}(A^*)$); if A^* and A have the same value, and the domains of A^* and A are identical, then we say A is self-adjoint.

Goals of Spectral Theory

In order to fully understand spectral theory, we need to take a step back and ponder what we expect the spectral theorem to do for us. Specifically, we want it to be able to apply functions to operators; for instance, the time-dependent Schrödinger equation is solved by setting

$$\psi(t) = \psi_0 e^{-\frac{it}{\hbar} \hat{H}}.$$

^{xlii}I'm anticipating that I will depend a lot on various tables evaluating Fourier transforms if I take quantum.

^{xliii}I don't really like PDEs and these sections are confusing me, so I'll probably do these much later.

^{xliv}That is, it will be shown in the future.

However, since \hat{H} is unbounded, we can't use the power series representation of e^x to define it; if \hat{H} has an orthonormal basis of eigenvectors $\{e_k\}_k$ with eigenvalues $\{\lambda_k\}_k$, then we can define \hat{H} to be the unique bounded operator with

$$e_k e^{-it\frac{\hat{H}}{\hbar}} = e_k e^{-it\frac{\lambda_k}{\hbar}}$$

for all k .

If \hat{H} doesn't have an orthonormal basis of eigenvectors, we want the spectral theorem to provide a functional calculus that allows us to apply functions to \hat{H} (and specifically, one that has similar properties to the case where \hat{H} has an orthonormal basis).

We would also like the spectral theorem to provide a probability distribution for measurement. In the case where A has an orthonormal basis $\{e_j\}_j$ of eigenvectors with eigenvalues $\{\lambda_j\}_j$, we compute probabilities as follows.

For Borel $E \subseteq \mathbb{R}$, we let $V_E = \overline{\text{span}\{e_j \mid \lambda_j \in E\}}$, and P_E be the orthogonal projection onto V_E . For any unit vector ψ , we define

$$\text{prob}_\psi(A \in E) = \langle \psi, P_E \psi \rangle.$$

In particular, if the eigenvalues are distinct and $\psi = \sum_j c_j e_j$, then the probability of observing λ_j will be $|c_j|^2$.

If A does not have an orthonormal basis of eigenvectors, the spectral theorem should be able to provide a family of projection operators P_E for each Borel $E \subseteq \mathbb{R}$. These will be known as spectral projections, with V_E the spectral subspaces.

In order to achieve these goals, we will create spectral projections using a projection-valued measure, and create the functional calculus via integration.

Meanwhile, to understand generalized eigenvectors, we need to create the *direct integral* to decompose \mathbf{H} with respect to a measure μ into generalized eigenspaces for the self-adjoint operator A .^{xlv} The generalized eigenspace for λ will not be a subspace of \mathbf{H} unless $\mu(\{\lambda\}) > 0$. The direct integral will allow us to formally define "eigenvectors" that are not elements of the Hilbert space.

Spectral Theorem on the Position Operator

Consider the position operator X . We have established that X has no true Hilbert space eigenvectors. If we think that the generalized eigenvectors for X are the distributions $\delta(x - \lambda)$ for $\lambda \in \mathbb{R}$, we can make a guess that V_E should consist of functions whose support is E . The projection P_E can then be calculated by

$$P_E \psi = \mathbb{1}_E \psi.$$

In this case, the probability is

$$\langle \psi, P_E \psi \rangle = \int_E |\psi(x)|^2 dx.$$

Recall from earlier that we claimed that the probability distribution for the position of a particle with state ψ is $|\psi(x)|^2$, meaning this construction holds.

Turning our attention to the functional calculus, if $f(\lambda) = \lambda^m$, then we expect that $f(X)$ should be multiplication by x^m . Similarly, for any function f , $f(X)$ should be multiplication by $f(x)$; in particular, e^{iaX} should be defined as multiplication by e^{iax} , itself a bounded operator on $L^2(\mathbb{R})$.

^{xlv}Think about it as the "continuous" analogue of the direct sum of subspaces for a vector space.

Generalization: Multiplication Operators

Consider the more general setting $\mathbf{H} = L^2(X, \mu)$, with h a real-valued measurable function on X . The multiplication operator on $L^2(X, \mu)$ is then defined

$$M_h \psi = h\psi,$$

with the spectral subspaces

$$V_E = \{\psi \mid \text{supp}(\psi) = h^{-1}(E)\}.$$

The functional calculus can then be defined by

$$f(A) = M_{f \circ h} A.$$

Theorem (The Spectral Theorem: Unitary Equivalence). *A self-adjoint operator A on a separable Hilbert space is unitarily equivalent to a multiplication operator.*

This means there is some σ -finite measure space (X, μ) and a measurable function h on X such that $A = UM_h U^*$ for some unitary operator U .

Spectral Theorem on the Momentum Operator

Consider, now, $P = -i\hbar \frac{d}{dx}$ on $L^2(\mathbb{R})$. The eigenvectors of P are e^{ikx} for $k \in \mathbb{R}$, with eigenvalues $\hbar k$. While $e^{ikx} \notin L^2(\mathbb{R})$, the Fourier transform shows that any function in $L^2(\mathbb{R})$ can be expanded as a superposition of functions of the form e^{ikx} .

As a result, the Fourier coefficients of $\hat{\psi}(k)$ can be expressed in terms of a quasi-inner product of ψ with e^{ikx} :

$$\begin{aligned} \hat{\psi}(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \psi(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \langle e^{ikx}, \psi \rangle_{L^2(\mathbb{R})}. \end{aligned}$$

Physicists say^{xlvi} that $\frac{1}{\sqrt{2\pi}} e^{ikx}$ forms a “continuous orthonormal basis” for $L^2(\mathbb{R})$. The Kronecker delta is replaced by the Dirac δ distribution.

$$\left\langle \frac{1}{\sqrt{2\pi}} e^{ikx}, \frac{1}{\sqrt{2\pi}} e^{ilx} \right\rangle = \delta(k - l),$$

where for continuous f , δ satisfies

$$\int_{-\infty}^{\infty} f(k) \delta(k - l) dk = f(l).$$

The quasi-inner product on e^{ikx} and e^{ilx} can be more rigorously defined by approximation as

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-A}^A e^{-ikx} e^{ilx} dx &= \frac{A \sin(A(k - l))}{\pi A(k - l)} \\ &= f(k). \end{aligned}$$

Taking the limit as A approaches infinity, it is possible to see that

$$\lim_{A \rightarrow \infty} \int_{-A}^A \psi(k) \frac{A \sin(A(k - l))}{\pi A(k - l)} dk = \psi(l),$$

meaning $f(k)$ behaves like the δ distribution.

^{xlvi}They have a tendency to do that.

Definition (Momentum Wave Function). For any $\psi \in L^2(\mathbb{R})$, define $\tilde{\psi}$ by

$$\tilde{\psi}(p) = \frac{1}{\sqrt{\hbar}} \hat{\psi}\left(\frac{p}{\hbar}\right),$$

meaning

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \psi(x) dx.$$

We call $\tilde{\psi}(p)$ the momentum wave function.

Since the Fourier transform is unitary, if ψ is a unit vector then so too is $\hat{\psi}$ and $\tilde{\psi}$; we interpret $|\tilde{\psi}(p)|^2$ to be the probability density of the momentum just as $|\psi(x)|^2$ is the probability distribution of the position.

We can then verify that for nice ψ , $\widehat{P\psi}(p) = p\tilde{\psi}(p)$, meaning the map $\psi \rightarrow \tilde{\psi}$ turns the momentum operator into multiplication by p .

The functional calculus for P can be defined similar to the case for X . For Borel $E \subseteq \mathbb{R}$, define P_E to be the orthogonal projection onto the space of functions ψ for which $\text{supp}(\tilde{\psi}(p)) = E$. If f is any bounded measurable function on \mathbb{R} , define $f(P)\psi$ to be the unique element of $L^2(\mathbb{R})$ where

$$\widehat{f(P)\psi}(p) = f(p)\tilde{\psi}(p).$$

The Spectral Theorem for Bounded Self-Adjoint Operators

Preliminaries

We will let \mathbf{H} denote a separable complex Hilbert space. An operator A on \mathbf{H} is bounded if the operator norm of A is finite:

$$\begin{aligned} \|A\|_{\text{op}} &= \sup_{\|\psi\|=1} \|A\psi\| \\ &< \infty. \end{aligned}$$

For any operator A , I will be denoting the operator norm $\|A\|_{\text{op}}$ as $\|A\|$. The space of bounded operators on \mathbf{H} forms a Banach algebra under the operator norm with multiplication; additionally, $\|AB\|_{\text{op}} \leq \|A\|_{\text{op}} \|B\|_{\text{op}}$. The Banach algebra of bounded operators on \mathbf{H} with respect to the operator norm is denoted $\mathcal{B}(\mathbf{H})$.

For any $A \in \mathcal{B}(\mathbf{H})$, there is a unique operator $A^* \in \mathcal{B}(\mathbf{H})$ such that

$$\langle \phi, A\psi \rangle = \langle A^*\phi, \psi \rangle$$

for all $\phi, \psi \in \mathbf{H}$. An operator $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint if $A^* = A$, and A is nonnegative if $\langle \psi, A\psi \rangle \geq 0$ for all $\psi \in \mathbf{H}$.

Proposition (The C^* Property). For all $A \in \mathcal{B}(\mathbf{H})$, we have

$$\|A^*\| = \|A\|$$

and

$$\|A^*A\| = \|A\|^2.$$

If A is self-adjoint, we have $\|A^2\| = \|A\|^2$.

Proof. For any vector $\phi \in \mathbf{H}$, $\|\phi\| = \sup_{\|\chi\|=1} |\langle \chi, \phi \rangle|$ (by the Cauchy-Schwarz inequality and taking χ a multiple of ϕ). Thus,

$$\|A\| = \sup_{\|\phi\|=\|\psi\|=1} |\langle \phi, A\psi \rangle|.$$

Thus, we get

$$\begin{aligned} \|A^*\| &= \sup_{\|\phi\|=\|\psi\|=1} |\langle \phi, A^*\psi \rangle| \\ &= \sup_{\|\phi\|=\|\psi\|=1} |\langle A\phi, \psi \rangle| \\ &= \sup_{\|\phi\|=\|\psi\|=1} |\langle \psi, A\phi \rangle| \\ &= \|A\|. \end{aligned}$$

Since $\|A^*A\| \leq \|A\|^2$ (by submultiplication of the norm), we have

$$\begin{aligned} \|A^*A\| &= \sup_{\|\phi\|=\|\psi\|=1} |\langle \phi, A^*A\psi \rangle| \\ &= \sup_{\|\phi\|=\|\psi\|=1} |\langle A\phi, A\psi \rangle| \\ &\geq \sup_{\|\psi\|=1} |\langle A\psi, A\psi \rangle| \\ &= \|A\|^2. \end{aligned}$$

□

Proposition (Orthogonal Complements and Adjoints). *For all $A \in \mathcal{B}(\mathbf{H})$, we have*

$$(\text{Range}(A))^\perp = \ker(A^*).$$

Proof. Let $\psi \in (\text{Range}(A))^\perp$. Then, for all $\phi \in \mathbf{H}$, we have

$$\begin{aligned} 0 &= \langle \psi, A\phi \rangle \\ &= \langle A^*\psi, \phi \rangle, \end{aligned}$$

meaning $A^*\psi = 0$, or $\psi \in \ker(A^*)$.

Let $\psi \in \ker(A^*)$. Then, for all $\phi \in \mathbf{H}$, the above relation holds (reading from bottom right upward), implying ψ is orthogonal to every element of the form $A\phi$, so $\psi \in (\text{Range}(A))^\perp$. □

Definition (Spectrum and Resolvent). For $A \in \mathcal{B}(\mathbf{H})$, the resolvent set of A , denoted $\rho(A)$, is the set of all $\lambda \in \mathbb{C}$ such that $(A - \lambda I)$ has a bounded inverse.

The spectrum of A , denoted $\sigma(A)$, is equal to $\mathbb{C} \setminus \rho(A)$.

For $\lambda \in \rho(A)$, the operator $(A - \lambda I)^{-1}$ is called the resolvent of A at λ .

By “bounded inverse,” it is meant that there exists a bounded operator B such that $(A - \lambda I)B = B(A - \lambda I) = I$. If A is bounded and $A - \lambda I$ is one-to-one and maps \mathbf{H} to \mathbf{H} , then the closed graph theorem states that the inverse map must be bounded. With this in mind, we can also describe the resolvent set as the set of $\lambda \in \mathbb{C}$ such that $A - \lambda I$ is bijective.

Proposition (Properties of the Spectrum and Resolvent). *Let $A \in \mathcal{B}(\mathbf{H})$. Then,*

- $\sigma(A) \subseteq \mathbb{C}$ is closed, bounded, and non-empty.

- If $|\lambda| \geq \|A\|$, then $\lambda \in \rho(A)$.

Lemma (Property of a Contraction Operator). Let $X \in \mathcal{B}(\mathbf{H})$ with $\|X\| < 1$. Then, $I - X$ is invertible and its inverse is given by the series

$$(I - X)^{-1} = \sum_{n=0}^{\infty} X^n.$$

Proof of Lemma. Since the operator norm is submultiplicative, $\|X^m\| \leq \|X\|^m$. The geometric series is, therefore, absolutely convergent, which implies convergence in $\mathcal{B}(\mathbf{H})$.

Multiplying by $(I - X)$, everything on the right cancels except for I ; therefore, the series is the inverse of $(I - X)$. \square

Proof of Properties of Spectrum and Resolvent. Let $\lambda \in \mathbb{C} \setminus \{0\}$. Consider the operator

$$A - \lambda I = -\lambda \left(I - \frac{A}{\lambda} \right).$$

If $|\lambda| > \|A\|$, then $\|A/\lambda\| < 1$, meaning $I - A/\lambda$ is invertible, so $A - \lambda I$ is invertible, meaning $\lambda \in \rho(A)$.

Let $\lambda_0 \in \rho(A)$. For any other value λ , we have

$$\begin{aligned} A - \lambda I &= A - \lambda_0 I - (\lambda - \lambda_0)I \\ &= (A - \lambda_0 I) (A - (\lambda - \lambda_0)(A - \lambda_0 I)^{-1}). \end{aligned}$$

If

$$|\lambda - \lambda_0| < \frac{1}{\|(A - \lambda_0 I)^{-1}\|},$$

both factors on the right will be invertible, so $A - \lambda I$ is also invertible, meaning $\rho(A)$ is open and $\sigma(A)$ is closed.

To show $\sigma(A)$ is nonempty, we note that

$$(A - \lambda I)^{-1} = \left(\sum_{m=0}^{\infty} (\lambda - \lambda_0)^m (A - \lambda_0 I)^{-m} \right) (A - \lambda_0 I)^{-1}.$$

Near any point λ_0 in the resolvent set, the resolvent $(A - \lambda I)^{-1}$ can be computed by the series in powers of $\lambda - \lambda_0$, and the coefficients of the series are elements of $\mathcal{B}(\mathbf{H})$. For any $\psi, \phi \in \mathbf{H}$, the map

$$\lambda \mapsto \langle \phi, (A - \lambda I)^{-1} \psi \rangle$$

will be given by a locally convergent power series with coefficients in \mathbb{C} , meaning the function is holomorphic on $\rho(A)$. We can also see that $\|(A - \lambda I)^{-1}\|$ tends to zero as $|\lambda|$ tends to infinity.

If $\sigma(A)$ were empty, the function would be holomorphic on all of \mathbb{C} and tend to zero at infinity, meaning it would be identical to 0 for all ϕ, ψ , which means $(A - \lambda I)^{-1}$ is the zero operator, which cannot happen. \square

If $A\psi = \lambda\psi$ for some $\lambda \in \mathbb{C}$ and nonzero $\psi \in \mathbf{H}$, then $(A - \lambda I)$ has nonzero kernel, meaning $\lambda \in \sigma(A)$, so any eigenvalue of A is contained in the spectrum of A .

In the infinite-dimensional case, though, points in the spectrum may not be eigenvalues; but, for a bounded self-adjoint operator, the spectrum of A can be defined similar to the finite-dimensional case.

Proposition (Properties of the Spectrum of a Bounded Self-Adjoint Operator). *If $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint, then the following are true.*

- $\sigma(A) \subseteq \mathbb{R}$.
- A number $\lambda \in \mathbb{R}$ belongs to the spectrum of A if and only if there exists a sequence ψ_n of nonzero vectors in \mathbf{H} such that

$$\lim_{n \rightarrow \infty} \frac{\|A\psi_n - \lambda\psi_n\|}{\|\psi_n\|} = 0.$$

Essentially, the spectrum contains all numbers where the error between $A\psi$ and $\lambda\psi$ is small relative to ψ .

Lemma (Lower bound for $A - \lambda I$). *Let $A \in \mathcal{B}(\mathbf{H})$ be self-adjoint. Then, for all $\lambda = a + bi \in \mathbb{C}$, we have*

$$\langle (A - \lambda I)\psi, (A - \lambda I)\psi \rangle \geq b^2 \langle \psi, \psi \rangle.$$

Proof of Lemma. We compute

$$\begin{aligned} \langle (A - \lambda I)\psi, (A - \lambda I)\psi \rangle &= \langle (A - aI)\psi, (A - aI)\psi \rangle + bi \langle \psi, (A - aI)\psi \rangle \\ &\quad - bi \langle (A - aI)\psi, \psi \rangle + b^2 \langle \psi, \psi \rangle, \end{aligned}$$

and since A is self-adjoint,

$$= \langle (A - aI)\psi, (A - aI)\psi \rangle + b^2 \langle \psi, \psi \rangle.$$

□

Proof of Properties of Spectrum of a Bounded Self-Adjoint Operator. To show the first item, we need to show that for any $\lambda \in \mathbb{C}$ with $\lambda = a + bi$, $b \neq 0$, $\lambda \in \rho(A)$.

To start, since $b \neq 0$, we know from the previous lemma that $A - \lambda I$ is injective.^{xlvii} Since $(\text{Range}(A - \lambda I))^\perp = \ker(A - \bar{\lambda}I)$, and $\text{Im}(\bar{\lambda}) \neq 0$, $A - \bar{\lambda}I$ is injective, meaning $\overline{\text{Range}(A - \lambda I)} = \mathbf{H}$.

To show that $\overline{\text{Range}(A - \lambda I)} = \mathbf{H}$, let $\phi \in \mathbf{H}$, and choose a sequence $\phi_n = (A - \lambda I)\psi_n$ in $\text{Range}(A - \lambda I)$ with $\phi_n \rightarrow \phi$. Applying the lemma, replacing ψ with $\psi_n - \psi_m$, we see that $(\psi_n)_n$ is Cauchy, so $\psi_n \rightarrow \psi$ for some $\psi \in \mathbf{H}$. Additionally, since A is bounded,

$$\begin{aligned} (A - \lambda I)\psi &= \lim_{n \rightarrow \infty} (A - \lambda I)\psi_n \\ &= \lim_{n \rightarrow \infty} \phi_n \\ &= \phi. \end{aligned}$$

Thus, $A - \lambda I$ is surjective, so $(A - \lambda I)^{-1}$ is bounded.

To show the second item in the forward direction, let $\lambda \in \sigma(A)$ and ψ_n be a sequence such that

$$\lim_{n \rightarrow \infty} \frac{\|A\psi_n - \lambda\psi_n\|}{\|\psi_n\|} = 0.$$

Suppose toward contradiction that $A - \lambda I$ has an inverse. Let $\phi_n = (A - \lambda I)\psi_n$, meaning $\psi_n = (A - \lambda I)^{-1}\phi_n$; thus, we see

$$\lim_{n \rightarrow \infty} \frac{\|\phi_n\|}{\|(A - \lambda I)^{-1}\phi_n\|} = 0,$$

^{xlvii} $\ker(A - \lambda I) = \{0\}$ since $\|(A - \lambda I)\psi\|^2 \geq b^2 \|\psi\|^2 \neq 0$ for all $\psi \neq 0$.

meaning $A - \lambda I$ is unbounded.

To show the second item in the reverse direction, if for some $\lambda \in \mathbb{R}$, no such sequence exists, then for some $\varepsilon > 0$,

$$\|(A - \lambda I)\psi\| \geq \varepsilon \|\psi\|,$$

meaning $A - \lambda I$ is injective, so the range of A is dense in \mathbf{H} ; by the argument made in the first point, this means $A - \lambda I$ is also bijective, so $\lambda \in \rho(A)$. \square

Example (Spectrum of M_x). Let $\mathbf{H} = L^2([0, 1])$, and let $A = M_x$. Then, A is a bounded, self-adjoint operator whose spectrum is $\sigma(A) = [0, 1]$.

Proof. It is readily apparent that $\|A\psi\| \leq \|\psi\|$ (since $x \leq 1$ for all $x \in [0, 1]$), and $\langle \phi, A\psi \rangle = \|A\phi\| \psi$ for all $\phi, \psi \in \mathbf{H}$, meaning A is bounded and self-adjoint.

Let $\lambda \in (0, 1)$, consider $\psi_n = \mathbb{1}_{[\lambda, \lambda + 1/n]}$, where $\|\psi_n\|^2 = \frac{1}{n}$. Additionally, since $|x - \lambda| \leq \frac{1}{n}$ on $[\lambda, \lambda + 1/n]$, we have

$$\|(A - \lambda I)\psi_n\|^2 \leq \frac{1}{n^3}.$$

Thus, $\lambda \in \sigma(A)$; since this is true for all $\lambda \in (0, 1)$ and $\sigma(A)$ is closed, $[0, 1] \subseteq \sigma(A)$.

Meanwhile, if $\lambda \notin [0, 1]$, the function $\frac{1}{x - \lambda}$ is bounded on $[0, 1]$, meaning $A - \lambda I$ has a bounded inverse consisting of multiplication by $\frac{1}{x - \lambda}$. Therefore, $\sigma(A) = [0, 1]$. \square

Spectral Subspaces

Given a bounded self-adjoint operator A , we hope to associate each Borel $E \subseteq \sigma(A)$ with a closed subspace $V_E \subseteq \mathbf{H}$, where V_E is effectively the closed span of the generalized eigenvectors of A with eigenvalues in E . The subspaces should have the following properties:

- (1) $V_{\sigma(A)} = \mathbf{H}$ and $V_\emptyset = \{0\}$.
- (2) $E \cap F = \emptyset$ implies $V_E \perp V_F$.
- (3) For any E, F , $V_{E \cap F} = V_E \cap V_F$.
- (4) If E_1, E_2, \dots are disjoint, $E = \bigsqcup_j E_j$, then

$$V_E = \bigoplus_j V_{E_j}.$$

- (5) For any E , V_E is invariant under A .
- (6) If $E \subseteq [\lambda_0 - \varepsilon, \lambda_0 + \varepsilon]$, then for $\psi \in V_E$,

$$\|(A - \lambda_0 I)\psi\| \leq \varepsilon \|\psi\|.$$

Combined, these properties essentially allow for an infinite-dimensional version of the finite-dimensional properties that the spectrum holds for self-adjoint matrices.

Projection-Valued Measure

In order to describe closed subspaces of \mathbf{H} , we would like to use the projection operator to do so. Given a closed subspace $V \subseteq \mathbf{H}$, there exists a unique bounded, self-adjoint, idempotent operator P that equals I on V and equals 0 on V^\perp .

We can express the first four properties of spectral subspaces by defining measure that has properties similar to that of a measure.

Definition (Projection-Valued Measure). Let X be a set, with Ω a σ -algebra on X .^{xlviii} A map $\mu : \Omega \rightarrow \mathcal{B}(\mathbf{H})$ is called a projection-valued measure if the following hold:

- (1) For each $E \in \Omega$, $\mu(E)$ is an orthogonal projection.
- (2) $\mu(\emptyset) = 0$ and $\mu(X) = I$.
- (3) If E_1, E_2, \dots are disjoint, then for all $v \in \mathbf{H}$, we have

$$\mu\left(\bigsqcup_{j=1}^{\infty} E_j\right)v = \sum_{j=1}^{\infty} \mu(E_j)v.$$

- (4) For all $E_1, E_2 \in \Omega$, we have $\mu(E_1 \cap E_2) = \mu(E_1)\mu(E_2)$.

If E_1 and E_2 are disjoint, then properties (2) and (4) tell us that $\mu(E_1)\mu(E_2) = 0$, from which it follows that the range of $\mu(E_1)$ and $\mu(E_2)$ are perpendicular; we can verify that $\mu(E_1)\mu(E_2)$ is the projection onto the intersection of the ranges of $\mu(E_1)$ and $\mu(E_2)$. For each $E \in \Omega$, we define a closed subspace $V_E = \text{Range}(\mu(E))$; the collection of V_E then satisfy the first four properties we expect for spectral subspaces.

We will associate a projection-valued measure μ^A for each bounded self-adjoint operator A ; then, $\mu^A(E)$ is a projection onto the spectral subspace corresponding to E . Then, we will introduce operator-valued integration with respect to a projection-valued measure μ^A for A , which will be known as a functional calculus for A .

Observe that for any projection-valued μ and $\psi \in \mathbf{H}$, we can define a positive, real-valued μ_ψ by

$$\mu_\psi(E) = \langle \psi, \mu(E)\psi \rangle$$

for $E \in \Omega$. This allows us to link between projection-valued measure and integration with respect to an ordinary measure.

Definition (Operator-Valued Integration). Let Ω be a σ -algebra in X and $\mu : \Omega \rightarrow \mathcal{B}(\mathbf{H})$ a projection-valued measure. Then, there exists a unique linear map, $f \mapsto \int_\Omega f d\mu$ from the space of bounded, measurable, complex-valued functions on Ω into $\mathcal{B}(\mathbf{H})$, where

$$\left\langle \psi, \left(\int_X f d\mu \right) \psi \right\rangle = \int_X f d\mu_\psi$$

for all f and all $\psi \in \mathbf{H}$, where μ_ψ is given as above. The integral has the following properties:

- (1) For all $E \in \Omega$, we have

$$\int_X \mathbb{1}_E d\mu = \mu(E).$$

In particular, the integral of $\mathbb{1}_X$ is I .

^{xlviii}For more information on σ -algebras, measures, etc., consult my Real Analysis II notes.

(2) For all f , we have

$$\left\| \int_X f d\mu \right\| \leq \sup_{\lambda \in X} |f(\lambda)|.$$

(3) For all f and g , we have

$$\int_X fg d\mu = \left(\int_X f d\mu \right) \left(\int_X g d\mu \right).$$

(4) For all f , we have

$$\int_X \bar{f} d\mu = \left(\int_X f d\mu \right)^*.$$

In particular, if f is real-valued, then $\int_X f d\mu$ is self-adjoint.

We can see from linearity that property (1) applies to all simple functions; as a result, we can integrate an arbitrary bounded measurable function by taking a limit of simple functions converging uniformly to f — the integral of f is the limit in the operator norm topology of the integral of the simple functions.

In this context, multiplication is straightforward:

$$\begin{aligned} \left(\int_X \mathbb{1}_{E_1} d\mu \right) \left(\int_X \mathbb{1}_{E_2} d\mu \right) &= \mu(E_1)\mu(E_2) \\ &= \mu(E_1 \cap E_2) \\ &= \int_X \mathbb{1}_{E_1} \mathbb{1}_{E_2} d\mu. \end{aligned}$$

One of the major reasons this can be the case is that there are many more idempotent elements in $\mathcal{B}(\mathbf{H})$ than in \mathbb{R} (attempting a similar construction of a multiplicative integral in \mathbb{R} yields effectively no benefit).

Proof of Properties of Operator-Valued Integration. Given a projection-valued measure μ and a bounded measurable function f , define $Q_f : \mathbf{H} \rightarrow \mathbb{C}$ by

$$Q_f(\psi) = \int_X f d\mu_\psi,$$

where $\mu_\psi(E) = \langle \psi, \mu(E)\psi \rangle$. If f is an indicator function, then $Q_f(\psi)$ is a bounded quadratic form — as a result, by the boundedness of the form,

$$|Q_f(\psi)| \leq \left(\sup_{\lambda \in X} |f(\lambda)| \right) \|\psi\|^2.$$

Thus, there is a unique bounded operator A_f such that $Q_f(\psi) = \langle \psi, A_f \psi \rangle$ for all $\psi \in \mathbf{H}$.

If $f = \mathbb{1}_E$, then $Q_f(\psi) = \mu_\psi(E) = \langle \psi, \mu(E)\psi \rangle$, meaning $A_f = \mu(E)$. Property (2) follows from the above inequality.

To show property (3), we see that multiplicativity of the integral for indicator functions is built into the definition of the projection-valued measure; we use linearity to extend the multiplicativity to simple functions, then limits for bounded measurable functions.

Finally, for f real valued, $Q_f(\psi)$ is real for all $\psi \in \mathbf{H}$, meaning A_f is self-adjoint. This yields property (4) via linearity. \square