Mathematical physics notes

John Kerl

February 9, 2008

Abstract

These are, in main part, class notes from Jan Wehr's splendid fall 2007 mathematical physics (Math 541) course at the University of Arizona. Some additional reference material drawn from my other graduate coursework is added as appendices.

This is very much a work under construction. I expect general completeness by end of term, with probably some final clean-up over the winter break.

Contents

1   Introduction

1.1   Reason for the course

Math 541 is a course in mathematical physics — offered through a mathematics department, although it is cross-listed with Physics 541.

Why offer such a course?

- It is an essential course for students going on in mathematical physics.

- It is important for students of mathematics to have at least an appreciation for where certain mathematical structures came from.

- It is important for students of physics to have at least an appreciation for a rigorous approach to mathematical physics.

- For the sake of the forward progress of mathematics and physics, students of each discipline must learn to speak one another's languages.

Certain mathematical structures and theories come from physics. For example:

- Symplectic geometry originates with the Hamiltonian approach to classical mechanics.

- Probability theory originates in part from statistical mechanics. (It also owes a huge debt to gambling!) In particular, we mention central limit theorems and the theory of large deviations, which is an active area of current research.

- The spectral theory of operators in Hilbert spaces, along with von Neumann and C^* algebras, originates in quantum theory.

1.2   About these notes

Dr. Wehr is taking a best-of approach from several texts. It seems worthwhile to me to record his systematic treatment in writing, in a way (searchable and indexed, as handwritten notes can never be) that will be useful for future reference.

A disclaimer: lectures and written material are not the same. The former is, as it should be, bracketed by class sessions; one reviews material from the previous class, and summarizes at the end. One omits certain details due to time constraints (for example, "Now check that the axioms for a metric are satisfied"). Writing notes is then more than a mere exercise in transcription — it includes removal of segués and insertion of details. This is a work in progress — as of the current writing, these notes retain their conversational tone (which I would like to retain) with some details still not fully fleshed out.

The course is being taught with the point of view that the students have seen elementary quantum mechanics; in fact this is true for most of the students in the class. In the main body of the paper we move from spectral theory directly to many-body theory. For my own reference, however, I want to create a handy, short, example-driven list of basic prerequisite facts: the essentials of what I need to know to be conversant with quantum theory, assuming first-year graduate coursework as a given. These constitute the appendices. In fact, the appendices represent my attempt to document, if not fill in, the gaps in my own knowledge.

My longer-term goal is to create a handy reference for myself containing what I need to know about quantum mechanics for my graduate study — given the pure-math core-course sequence in algebra, analysis, and geometry-topology as prerequisite material. These notes will be more and less than a text — more examples, less filler and formality. Also, I will summarize, in this one location, key concepts from various disciplines on which quantum mechanics depends: linear and multilinear algebra, probability, statistics, measure theory, and functional analysis.

My attitude toward rigor in reference papers I write for myself (such as this one) is the following: If a proof is short, and if I can remember it after reading it, and if the proof contributes something to my understanding, I will write it out. Else I will give a reference. In fact, papers such as this one are primarily *not* mathematical; they contain far more intuition than rigor. So be it.

1.3   Overview of the course

The plan is as follows:

- Give some motivation for quantum theory.

- Go through spectral theory in as much rigor as possible, while leaving time for further topics. (We may need to abbreviate or skip some proofs.) It is possible to delve so deeply into spectral theory that at the end of the semester there is time only for study of the infinite well and the simple harmonic oscillator. In fact, the course has been taught this way before. This semester, we will make time for some topics of current interest in physics.

- Extend/specialize the spectral theory to systems of many (usually identical) particles. We will use the tools of second quantization and Fock spaces.

- Discuss Bose-Einstein condensates and the BCS theory of superconductivity. Last, mention superfluids if time permits.

The main text for the course is [**MR**], with two important deviations: (1) we will take some time at the beginning to develop quantum-theoretical concepts from scratch, and (2) we will omit chapter 4 of the text (on the electron gas) entirely.

1.4   Notation

The integers, reals, and complexes are denoted by $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ respectively.

For $z = x + iy \in \mathbb{C}$, the complex conjugate $x - iy$ of z is written either $\overline{z}$ or z^*. The modulus or absolute value of z, which is $\sqrt{x^2 + y^2} = \sqrt{z\overline{z}}$, is written $|z|$. (Notice in particular that $z\overline{z} = |z|^2$.) The real part x and the imaginary part y of z are written $\mathrm{Im}(z)$ and $\mathrm{Re}(z)$, respectively.

For a map A from an object V to an object W, the kernel of A is written $\ker(A)$; the image of A is written either $\mathrm{im}(A)$ or $A(V)$. Injective and surjective mappings are written $A \hookrightarrow B$ and $A \twoheadrightarrow B$, respectively. Actions of mappings on elements are written $a \mapsto b$.

The indicator function of a subset A of a set X is written 1_A. It has values

$$1_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A. \end{cases}$$

We write
$$A := B$$
when we are defining what A is. We say "A is defined to be B." For example,
$$\mathbb{C}^+ := \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}.$$

Proofs are ended by $\Box$; examples are bracketed between $\triangleright$ and $\triangleleft$. Definitions, on the other hand, terminate with the end of a paragraph.

2   Main ideas of quantum theory

In this section, we sketch some ideas in order to fix intuition, before diving headlong into technical details. You may think of it as a cartoon version.

2.1   Classical vs. quantum mechanics

Consider a one-particle system. This may conjure up a frightening view of an all-but-empty universe, but consider that the particle may be acted on by forces originating with fields or other particles. For now, we are simply focusing our spotlight on the behavior of a single particle.

The classical-mechanical view of the state of a single-particle system at time t is

$$(\mathbf{q}(t), \mathbf{p}(t)) \in \mathbb{R}^3 \times \mathbb{R}^3,$$

where $\mathbf{q}$ is the *position* of the particle, and $\mathbf{p} = m\dot{\mathbf{q}} = m\mathbf{v}$ is the *momentum* of the particle, where m is its mass. The *time evolution* of the particle is given by the *initial conditions* $(\mathbf{q}(0), \mathbf{p}(0))$ and a *law of evolution* describing $\dot{\mathbf{q}}$ and $\dot{\mathbf{p}}$. This is a system of *ordinary differential equations*, and they maybe (wildly) non-linear. Symplectic geometry addresses this subject further.

By contrast, the quantum-mechanical view of the state of a single-particle system is a function $\psi(\mathbf{x})$, for $\mathbf{x} \in \mathbb{R}^3$, from which position and momentum are derived:

$$\mathbf{x} \mapsto (\mathbf{q}, \mathbf{p}).$$

We again want to know the time evolution of the system. (We will intentionally mix between writing $\psi(\mathbf{x})$ and $\psi(\mathbf{x}, t)$. The reason for this will be given in section [write and xref].) The *initial conditions* are $\psi(\mathbf{x}, 0)$; the *law of evolution* is the *Schrödinger equation*, which is a PDE (since ψ is a function of both $\mathbf{x}$ and t) but is (happily) a linear PDE[1].

Why do we care about only $\mathbf{q}$ and $\mathbf{p}$? It turns out that many physical quantities are functions of position and momentum. For example, the classical kinetic energy is

$$T = \frac{m\|\mathbf{v}\|^2}{2} = \frac{\|\mathbf{p}\|^2}{2m}.$$

More generally, we can look at

$$F(\mathbf{q}(t), \mathbf{p}(t))$$

evolving in time. These functions of the state variables $\mathbf{q}$ and $\mathbf{p}$ are called *observables*. They are not, in fact, arbitrary functions of state; rather, we will view them as *linear operators* acting on the state. In particular, we will view $\mathbf{q}$ and $\mathbf{p}$ themselves as linear operators acting on ψ. (There are all sorts of linear operators; the ones in this course will all be multiplication or differentiation operators, or some combination thereof.)

The possible functions ψ will form a *Hilbert space* $\mathcal{H}$; observables will be *unbounded linear operators* in $\mathcal{H}$. We say "in" rather than "on" because these operators may not be defined on all of $\mathcal{H}$. This is not merely an annoyance — rather, it is key to the power of the theory.

---

[1]This is not to say that nonlinear equations don't appear in quantum theory; they do. The basic underlying PDEs *are* linear, but may (especially with multi-particle systems) become quite complicated. There exist approximation techniques which (perhaps drastically) reduce the number of variables, but which do so at the price of introducing nonlinearity in the approximation equations.

2.2 Hilbert space and linear operators

The particle state is encapsulated in a non-zero complex-valued function $\psi \in L^2(\mathbb{R}^3, \mathbb{C})$.

Definition 2.1. The L^2 condition, or *square-integrability condition*, means[2] that

$$\int_{\mathbb{R}^3} |\psi(\mathbf{x})|^2 \, d\mathbf{x} = \int_{\mathbb{R}^3} \psi^*(\mathbf{x})\psi(\mathbf{x}) \, d\mathbf{x} < +\infty.$$

(When time dependence is included, i.e. $\psi : \mathbb{R}^4 \to \mathbb{C}$, we require square-integrability only in the spatial coordinates.)

We will treat any two such functions differing only a complex constant, i.e. $\psi_1(\mathbf{x}) = c\psi_2(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^3$, as being the same. We *normalize* so that we assume

$$\|\psi\|_{L^2}^2 = 1$$

where

$$\|\psi\|_{L^2}^2 = \int_{\mathbb{R}^3} |\psi(\mathbf{x})|^2 \, d\mathbf{x}.$$

The physical interpretation of ψ is that $|\psi|^2$ — which is real-valued — plays the role of a probability-density function as discussed in appendix F. [xxx make sure to mention, here and/or in the appendix, griffith's caveat: probability in terms of repeated identical experiments. Also foreshadow (with xref) the collapse of the wave function and make it clear we can't immediately repeat the experiment on the *same* system.]

Definition 2.2. We define[3] a *scalar product* or *inner product* on $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C})$

$$(\phi, \psi) := \int_{\mathbb{R}^3} \phi^*(\mathbf{x})\psi(\mathbf{x}) d\mathbf{x}.$$

(For the physical interpretation, see [write and xref].) With this inner product,

$$\mathcal{H} = (L^2(\mathbb{R}^3, \mathbb{C}), (\cdot, \cdot))$$

is a *Hilbert space*.

Notation 2.3. The standard mathematical notation for the inner product is (ϕ, ψ) as written above. The physics notation (called *Dirac notation* or *bra-ket notation*) for the inner product (ϕ, ψ) is

$$\langle \phi \mid \psi \rangle.$$

We will later [write and xref] define what the symbols $\langle \phi \mid$ and $\mid \psi \rangle$ mean individually.

So far, this is a poor man's quantum theory: we only have a single particle, and we have no spin — without which (as we will see below [write and xref]) there is no superconductivity. Later, we will modify and extend the theory to include multiple particles as well as spin.

An *observable quantity* will be represented by a *self-adjoint operator* in $\mathcal{H}$. We say "in" rather than "on" since there will be domain issues: these operators may not be defined on all of the functions in $\mathcal{H}$. (We will see why in a few paragraphs.)

---

[2]The use of a superscript asterisk for complex conjugation is a physics convention which we will follow.
[3]Conjugating the first function in the integral rather than the second is another physics convention which we will follow.

Definition 2.4. Observable position $\mathbf{q}$ has an operator representation. It is defined by the vector-times-scalar product

$$\psi \mapsto \mathbf{x}\psi(\mathbf{x}).$$

I.e. if $\mathbf{x} = (x^1, x^2, x^3)$ then we have

$$\psi \xmapsto{q^\alpha} x^\alpha \psi$$

being an operator $q^\alpha : \mathcal{H} \to \mathcal{H}$.

The domain issue is that even if ψ is square integrable, $\mathbf{x}\psi$ might not be: multiplication by $\mathbf{x}$ is an *unbounded linear operator*. We will leave this issue at present; for now, just be aware of difficulties to come.

Definition 2.5. Similarly, the observable momentum $\mathbf{p}$ has an operator representation. This is $p^\alpha : \mathcal{H} \to \mathcal{H}$ given by

$$\psi \xmapsto{p^\alpha} -i\hbar \frac{\partial \psi}{\partial x^\alpha}.$$

The value $\hbar$ is called *Planck's constant*; its precise value and units are unimportant at present. The key thing to know for now is that it is small. This is why quantum effects are typically visible only microscopically. Again, there are domain issues: functions ψ which are square-integrable certainly do not need to be differentiable.

Definition 2.6. The *commutation relation* is

$$[\mathbf{q}, \mathbf{p}] = \mathbf{q}\mathbf{p} - \mathbf{p}\mathbf{q}$$

where the multiplication is composition of operators, which we recall is non-commutative in general.

This is also called the *canonical commutation relation* or *Heisenberg's commutation relation*. This is a misnomer; it describes *lack* of commutativity.

Proposition 2.7. *We have*

$$[\mathbf{q}, \mathbf{p}] = i\hbar$$

i.e.

$$[q^\alpha, p^\beta] := q^\alpha p^\beta - p^\beta q^\alpha = i\hbar \delta_{\alpha\beta}.$$

Proof. Using ψ as a test function, we have (using the product rule)

$$\begin{aligned}
[q^\alpha, p^\beta]\psi(\mathbf{x}) &= q^\alpha p^\beta \psi - p^\beta q^\alpha \psi \\
&= -i\hbar x^\alpha \frac{\partial \psi}{\partial x^\beta} + i\hbar \frac{\partial}{\partial x^\beta}(x^\alpha \psi) \\
&= -i\hbar x^\alpha \frac{\partial \psi}{\partial x^\beta} + i\hbar x^\alpha \frac{\partial \psi}{\partial x^\beta} + i\hbar \frac{\partial x^\alpha}{\partial x^\beta}\psi \\
&= i\hbar \delta_{\alpha\beta}\psi.
\end{aligned}$$

Removing the test function and collecting the coordinate functions into column vectors, we have

$$[\mathbf{q}, \mathbf{p}] = i\hbar.$$

$\square$

2.3 Fourier transforms

Definition 2.8. The *Fourier transform* of $\psi(\mathbf{x})$, written

$$\tilde{\psi}(\mathbf{k}) = \mathcal{F}(\psi(\mathbf{x})),$$

is

$$\tilde{\psi}(\mathbf{k}) := (2\pi)^{-3/2} \int_{\mathbb{R}^3} e^{-i\mathbf{k}\cdot\mathbf{x}} \psi(\mathbf{x})\, d\mathbf{x}.$$

Remark 2.9. The 3 in $(2\pi)^{-3/2}$ comes from the dimension of $\mathbb{R}^3$; in d dimensions it would be $(2\pi)^{-d/2}$.

Remark 2.10. There are numerous definitions in the literature involving the scale factor out front and the $-i$ in the integrand. We have carefully chosen the scale factor to make the Fourier transform an *isometry*, so that

$$\int_{\mathbb{R}^3} |\tilde{\psi}(\mathbf{k})|^2\, d\mathbf{k}$$

if ψ is normalized to 1. More generally,

$$\int_{\mathbb{R}^3} |\tilde{\psi}(\mathbf{k})|^2\, d\mathbf{k} = \int_{\mathbb{R}^3} |\psi(\mathbf{x})|^2\, d\mathbf{x}.$$

Remark 2.11. For the Fourier transform to be easily defined, we actually require $\psi \in L^1(\mathbb{R}^3, \mathbb{C}) \cap L^2(\mathbb{R}^3, \mathbb{C})$. We gloss this detail for now.

Definition 2.12. We will soon need the inverse transform, so we'll define it now:

$$\psi(\mathbf{x}) = \mathcal{F}^{-1}(\tilde{\psi}(\mathbf{k})) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} e^{+i\mathbf{k}\cdot\mathbf{x}} \tilde{\psi}(\mathbf{k})\, d\mathbf{k}.$$

The interpretation is that $|\tilde{\psi}|^2$ is the density of distribution of momenta. [xxx similar to caveats for distribution of $\mathbf{q}$: write and xref.] That is, the probability that $\mathbf{p}$ is in a (Borel) subset B of $\mathbb{R}^3$ is

$$\int_B |\tilde{\psi}(\mathbf{k})|^2\, d\mathbf{k}.$$

A crucial fact is how $\mathbf{p}$ acts in $\mathbf{k}$-space. In $\mathbf{x}$-space it was a derivative operator — what about here? To find out, let's carefully define it.

Definition 2.13. The momentum operator in $\mathbf{k}$-space is really shorthand for the compostion of three operators:

$$\mathbf{p}\tilde{\psi}(\mathbf{k}) := \mathcal{F} \circ \mathbf{p} \circ \mathcal{F}^{-1} \tilde{\psi}.$$

Writing out the steps (which are easy, but require justification) we obtain

$$\mathbf{p}\tilde{\psi}(\mathbf{k}) = \hbar\mathbf{k}\tilde{\psi}(\mathbf{k}).$$

That is, $\mathbf{p}$ is a multiplication operator in $\mathbf{k}$-space, just as $\mathbf{q}$ is a multiplication operator in $\mathbf{x}$-space. We call $\tilde{\psi}(\mathbf{k})$ the *wave function* of the particle in the momentum representation; $\psi(\mathbf{x})$ is called the wave function of the particle in the position representation. (When we discuss Bose-Einstein condensation in section [write and xref], we will see that the condensation appears explicitly only in the momentum representation.)

We emphasize that the Fourier transform turned a differentiation operator in the position representation into a multiplication operator in the momentum representation. Also, we will see in section 2.5, the potential-energy operator will be a multiplication operator and the kinetic-energy operator will be a differentiation

operator. It is (generally speaking) easier to solve problems involving position or potential energy in the position representation, and to solve problems involving momentum or kinetic energy in the momentum representation: in each case, the operators of interest become multiplication operators. One of the challenges of quantum mechanics is that there is no obvious representation which simplifies position *and* momentum, or potential *and* kinetic energy, simultaneously.

2.4   The measurement theorem

Suppose that A is a linear operator corresponding to an observable quantity. Assume self-adjointness of A. The linear-algebra point of view of self-adjointness (glossing details which occur in infinite-dimensional vector spaces, and ignoring domain issues whereby A is not defined on all of $\mathcal{H}$) is that for all $\phi, \psi \in \mathcal{H}$,

$$(A\phi, \psi) = (\phi, A\psi).$$

It is known that this implies that the spectrum of A is real. In the finite-dimensional case, the spectrum consists only of the (real) eigenvalues a_n, and there exists a spanning set of orthonormal eigenvectors ϕ_n such that

$$A\phi_n = a_n\phi_n.$$

Interpretation 2.14 (First postulate of quantum mechanics)**.** The interpretation is as follows. Suppose that ψ is the wave function of our one-particle system. Then measurement of the quantity to which A corresponds has as its possible outcomes the numbers a_n with probabilities $|(\phi_n, \psi)|^2$. If this is to work, we would want the probabilities to be non-negative (which we have by positive-definiteness of the inner product), and we would want them to sum to 1. In fact, they do.

Proposition 2.15. *We have*

$$\sum_n |(\phi_n, \psi)|^2 = 1.$$

Proof. First, we know we've normalized ψ so that

$$\|\psi\|^2 = (\psi, \psi) = 1.$$

Since the eigenfunctions ϕ_n are a spanning set, we have

$$\psi = \sum_n c_n \phi_n$$

with

$$c_n = (\phi_n, \psi).$$

Then

$$1 = (\psi, \psi) = \left(\sum_m c_m \phi_m, \sum_n c_n \phi_n \right) = \sum_m \sum_n (c_m \phi_m, c_n \phi_n)$$

$$= \sum_m \sum_n c_m^* c_n \delta_{mn}$$

$$= \sum_n |c_n|^2$$

by the orthonormality of the ϕ_n's. But the c_n's were my shorthand for (ϕ_n, ψ) so we have

$$\sum_n |(\phi_n, \psi)|^2 = 1$$

which is what we wanted to show. □

Next we can calculate the *mean* of A. [write and xref to appendix.] This is the sum over each possible measurement times the probability of that measurement. We have

$$(\psi, A\psi) = \left(\sum_m (\phi_m, \psi)\, \phi_m, \sum_n (\phi_n, \psi)\, A\, \phi_n \right)$$

$$= \left(\sum_m (\phi_m, \psi)\, \phi_m, \sum_n (\phi_n, \psi)\, a_n\, \phi_n \right)$$

$$= \sum_n a_n |(\phi_n, \psi)|^2$$

which is the mean of the value measured.

Interpretation 2.16 (Second postulate of quantum mechanics)**.** Immediately after measurement, the state of the system is described by ϕ_n *only*. This is called the *collapse of the wave function*. The evolution of the single-particle system thereafter is described by the Schrödinger equation, with new initial conditions.

2.5   The Schrödinger equation

The wave function $\psi(\mathbf{x}, t)$ satisfies (as verified by experiment)

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi$$

with H being a self-adjoint operator in $\mathcal{H}$. This H is called the *energy operator* or *Hamiltonian operator* for the system. It corresponds to classical energy just as the $\mathbf{q}$ and $\mathbf{p}$ operators correspond to the classical notions of position and momentum. A common set-up for a single particle is

$$H = T + V$$

where T is the *kinetic energy* and V is the *potential energy*. Classically, potential energy comes from such things as gravity, spring tension, electrical fields, etc. This is often position-dependent, although it could also be time-dependent (e.g. an electrical field which varies sinusoidally). The classical kinetic energy is

$$H = \frac{m\|\mathbf{v}\|^2}{2} = \frac{\|\mathbf{p}\|^2}{2m}$$

where m is the mass of the particle.

We want to mimic this as much as possible. We have

$$H = T + V = \frac{\mathbf{p}^2}{2m} + V(\mathbf{q}).$$

Classically and quantum-mechanically, we write

$$\mathbf{p}^2 = \sum_{\alpha=1}^{3} p^\alpha \cdot p^\alpha.$$

Classically, the p^α's are numbers; quantum-mechanically they are differentiation operators. It is key to note that $\mathbf{p}^2$ is still a linear operator, even though it looks quadratic. Writing out the second derivatives, we have (making note of various notations in common use)

$$T = -\frac{\hbar^2}{2m} \frac{\partial}{\partial x_j^2} = -\frac{\hbar^2}{2m}\Delta = -\frac{\hbar^2}{2m}\nabla^2,$$

i.e. $-\hbar^2/2m$ times the Laplacian operator. We also set V to be the operator which multiplies by $V(\mathbf{x})$.

With these definitions, the Schrödinger equation becomes

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi$$
$$= T\psi + V\psi$$
$$= -\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{x})\psi.$$

The next task is to attempt to write down some solutions. We will soon [write and xref] in some detail the **harmonic oscillator**, where

$$V(\mathbf{q}) = \frac{1}{2}\gamma\|\mathbf{q}\|^2.$$

Here, in this very special case, the Schrödinger equation can be solved in explicit detail.

2.6 The ODE perspective and the evolution operator

We can think of the Schrödinger equation as follows. Consider the mapping

$$t \mapsto \psi_t \in \mathcal{H}.$$

Then the Schrödinger equation is an $\mathcal{H}$-valued ODE

$$i\hbar\frac{d\psi_t}{dt} = H\psi_t.$$

(This is merely formal manipulation of symbols; think of it as a device for guessing solutions.) If $\mathcal{H}$ were a finite-dimensional vector space, with H represented by a matrix, then we would write

$$\psi_t = e^{-\frac{i}{\hbar}Ht}\psi_0 = U(t)\psi_0$$

where we define the **evolution operator** to be

$$U(t) = e^{-\frac{i}{\hbar}Ht}.$$

In an ODE course, we would define the **operator exponential** to be

$$e^{-\frac{i}{\hbar}Ht} = \sum_{n=0}^{\infty}\frac{\left(-\frac{i}{\hbar}Ht\right)^n}{n!}.$$

However, for an unbounded operator H, the series does not converge.

We may use a different procedure to define the operator exponential. First, diagonalize: find a family ψ_n of eigenfunctions, which we hope spans $\mathcal{H}$, with corresponding eigenvalues ε_n such that

$$H\psi_n = \varepsilon_n\psi_n.$$

Define an action on eigenfunctions

$$e^{-\frac{i}{\hbar}Ht}\psi_n e^{-\frac{i}{\hbar}\varepsilon_n t}\psi_n$$

which we hope extends by linearity. One reason for concern is the following. If

$$\psi = \sum_n c_n\psi_n$$

with

$$\sum_n |c_n|^2 = 1$$

and we try to write

$$U(t)\psi = \sum_n d_n \psi_n$$

then there is no guarantee that

$$\sum_n |d_n|^2 = 1.$$

A concrete example [write and xref: SHO] is

$$H\psi_n = \hbar(n + 1/2)\psi_n.$$

The coefficients here are not square-summable.

We do want to be able to define $U(t)$ on all of $\mathcal{H}$, but this is not automatic in the infinite-dimensional setting. This is called the **extension problem**. We will see in [write and xref] that there is in fact a one-to-one correspondence between groups of evolution operators and [missed it — xxx type up].

What properties do we expect of an evolution operator? Here is a list:

(1) Observables are real, so operators are self-adjoint. (There exists complex observables, but their real and imaginary parts correspond to self-adjoint operators.) We sometimes write $H^* = H$. Using mathematicians' notation, this means that for all $\phi, \psi \in \mathcal{H}$,

$$(H\phi, \psi) = (\phi, H\psi).$$

The Dirac notation reads simply $\langle \phi \mid H \mid \psi \rangle$ in either — this notation was designed for self-adjoint operators and so it does not permit us to communicate about non-self-adjoint operators.

(A remark: the class of self-adjoint operators is not closed under linear combinations. In particular, if A and B are self-adjoint, $A + B$ may not be. We will see below [write and xref] that a and $a^\dagger$ are non-self-adjoint linear combinations of self-adjoint operators.)

(2) From this it follows that

$$U(t)^* = U(t)^{-1}.$$

Here is an informal proof, using formal manipulation of symbols:

$$U^* = (e^{-\frac{i}{\hbar}Ht})^* = e^{+\frac{i}{\hbar}Ht} = e^{-\frac{i}{\hbar}H\cdot(-t)} = (e^{-\frac{i}{\hbar}H\cdot(t)})^{-1}$$

where we think of backward evolution undoing forward evolution.

(3) Preservation of normalization of ψ under U: i.e. if $\int |\psi(\mathbf{x})|^2 \, d\mathbf{x} = 1$, then $\int |U\psi(\mathbf{x})|^2 \, d\mathbf{x} = 1$. If $|\psi|^2$ is a probability density function, time evolution shouldn't change that. This is easy to show using unitarity of U, since **unitary operators** preserve norms in Hilbert spaces. (A consequence is that eigenvalues of U will lie on the unit circle in the complex plane.)

(4) $U(t_1 + t_2) = U(t_1) \circ U(t_2)$: A wave function should be able to evolve forward by t_1 time units, starting from initial conditions $\psi(\mathbf{x}, 0)$, then evolve further with new initial conditions $\psi(\mathbf{x}, t_1)$. That is, the $U(t)$'s form a **unitary group of operators**[4]

One in fact may make the following grand summary: *Quantum mechanics says that the energy operator generates the group of time-evolution operators.*

This is based mathematically on the theory of passing back and forth between H and $U(t)$, which we will see [write and xref] in **Stone's theorem** and the **spectral theorem**.

---

[4]One would rather say *group of unitary operators*, but the name has stuck.

2.7 A few words about many-body theory

For a while now we will be either restricting our attention to single-particle systems, or talking about abstract Hilbert spaces. We will not handle multi-particle systems until [write and xref]. However, it is important to at least sketch at this point what some of these other Hilbert spaces are — lest we make the mistake early on of thinking of $\mathcal{H}$ as $L^2(\mathbb{R}^3, \mathbb{C})$ only.

A wave function of a system of N particles is, in the position representation,

$$\psi(\mathbf{x}_1, \ldots, \mathbf{x}_N), \quad \mathbf{x}_j \in \mathbb{R}^3$$

with the normalization

$$\int_{\mathbb{R}^{3N}} |\psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)|^2 \, d\mathbf{x}_1 \cdots d\mathbf{x}_N = 1.$$

The momentum representation for the wave function is

$$\tilde{\psi}(\mathbf{k}_1, \ldots, \mathbf{k}_N) = (2\pi)^{-N/2} \int_{\mathbb{R}^{3N}} \exp\left(-i \sum_{j=1}^{N} \mathbf{k}_j \cdot \mathbf{x}_j\right) \psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) \, d\mathbf{x}_1 \cdots d\mathbf{x}_N.$$

The physical interpretation is as before: $|\psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)|^2$ is the probability density function for the positions of the particles, and $|\psi(\mathbf{k}_1, \ldots, ,_N)|^2$ is the PDF for their momenta. In particular, for $N = 2$ and $A \subseteq \mathbb{R}^3$, we have

$$\int_{A \times A} |\psi(\mathbf{x}_1, \mathbf{x}_2)|^2 \, d\mathbf{x}_1 \, d\mathbf{x}_2 = P(\mathbf{x}_1 \in A, \mathbf{x}_2 \in A).$$

The key question is, can we *label* particles 1 and 2? Sometimes we cannot distinguish them, and these symmetry requirements drastically reduce the number of possible wave functions. For example, if the two particles are *bosons* (e.g. certain atomic nuclei) then

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \psi(\mathbf{x}_2, \mathbf{x}_1)$$

and more generally for any permutation $\pi \in \mathcal{S}_N$,

$$\psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \psi(\mathbf{x}_{\pi(1)}, \ldots, \mathbf{x}_{\pi(N)}).$$

Likewise, for *fermions*,

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = -\psi(\mathbf{x}_2, \mathbf{x}_1)$$

and more generally

$$\psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \text{sgn}(\pi)\psi(\mathbf{x}_{\pi(1)}, \ldots, \mathbf{x}_{\pi(N)}).$$

Keep in mind that a Hilbert space may be $L^2(\mathbb{R}^{3N}, \mathbb{C})$ — or perhaps a (highly) proper subspace, or a quotient space.

3 Hilbert-space theory

Having completed our sketch of quantum theory, we now take a more careful approach to the underlying mathematics.

First, a caveat: In the introductory version above, we've tacitly assumed that operators in $\mathcal{H}$ are not unlike matrices acting on a finite-dimensional vector space. However, this notion cannot be overused. First, the finite-dimensional viewpoint excludes the case of continuous spectra. Second, consider the following commutation relation from section 2.2:

$$[q^\alpha, p^\beta] = i\hbar\delta_{\alpha\beta}I.$$

Can one have matrices P and Q on a finite-dimensional vector space such that

$$[P, Q] = i\hbar I?$$

No, and here is one reason: take traces. Trace is real-linear and moves through the difference, so

$$\mathrm{tr}(QP - PQ) = \mathrm{tr}(PQ) - \mathrm{tr}(QP).$$

Also, trace is central and so, even when $PQ \neq QP$, $\mathrm{tr}(PQ) = \mathrm{tr}(QP)$. On the right-hand side, we have $\mathrm{tr}(i\hbar I) = i\hbar n$ where n is the dimension of the vector space. These facts yield the absurdity $0 = i\hbar n$. In the infinite-dimensional setting, some operators have a trace defined on them (we say they are *of trace class*); some do not.

A note about notation: As mentioned above, there are mathematicians' and physicists' notations for things. In this section we will use the former, eschewing Dirac notation, in case the student consults mathematical reference texts on this material. However we will keep the convention throughout this course that we conjugate the first argument, not the second, in complex inner products.

3.1 Axioms

Definition 3.1. A **Hilbert space** $\mathcal{H}$ is an infinite-dimensional complex inner-product space where the inner product $(\cdot, \cdot)$ satisfies:

- Conjugate symmetry: for all $\phi, \psi \in \mathcal{H}$, $(\phi, \psi) = \overline{(\psi, \phi)}$. (Note that as a consequence, (ψ, ψ) is real.)

- Right linearity: for all $\phi, \psi \in \mathcal{H}$ and all $a, b \in \mathbb{C}$, $(\phi, a\psi_1 + b\psi_2) = a(\phi, \psi_1) + b(\phi, \psi_2)$.

- Left antilinearity: $(a\phi_1 + b\phi_2, \psi_2) = \bar{a}(\phi_1, \psi) + \bar{b}(\phi_2, \psi)$.

- Positive definiteness: $(\psi, \psi) \geq 0$ for all $\psi \in \mathcal{H}$, with $(\psi, \psi) = 0$ iff $\psi = 0$.

Proposition 3.2. *The inner product induces a **norm** on $\mathcal{H}$:*

$$\|\psi\| := (\psi, \psi)^{1/2}.$$

Proof. Exercise. Make use of the Cauchy-Schwartz inequality. $\qquad\square$

Proposition 3.3. *The norm induces a **metric** on $\mathcal{H}$:*

$$d(\phi, \psi) = \|\phi - \psi\|.$$

Proof. Exercise. $\qquad\square$

Proposition 3.4. $\mathcal{H}$ *is **complete** in this metric.*

Proof outline. One needs to show that a Cauchy sequence in $\mathcal{H}$ converges to something in $\mathcal{H}$. That is, if $\{\phi_n\}$ is a sequence in $\mathcal{H}$ such that $\|\phi_m - \phi_n\| \to 0$ as $m, n \to \infty$, then there exists $\phi \in \mathcal{H}$ such that $\|\phi_n - \phi\| \to 0$ as $n \to \infty$. (This type of convergence is called *convergence in norm*.) □

[xxx projection theorem and the need for spectral analysis? This is why we have Lebesgue measure — not just for esoteric constructions such as integrating the indicator function of the rationals on the unit interval.]

[xxx Riesz-Fischer theorem is non-trivial? What is the context of this comment? The thm is that ψ is square-integrable if its Fourier series converges in ℓ^2 ....]

3.2 Operators in Hilbert spaces

There are many definitions here. Where we are headed is the spectral theorem which says that self-adjoint operators have a so-called *spectral resolution property* which makes them manageable.

Definition 3.5. Let $\mathcal{H}_1, \mathcal{H}_2$ be two Hilbert spaces. A **linear operator** from $\mathcal{H}_1$ to $\mathcal{H}_2$ consists of a **domain of definition** $\mathcal{D}_A$, which is a linear subspace of $\mathcal{H}_1$, and a **linear map** $A : \mathcal{D}_A \to \mathcal{H}_2$.

Often, but not necessarily, $\mathcal{D}_A$ is dense in $\mathcal{H}_1$. If it is, we say that A is **densely defined**. The domains of definition are usually described by elements of an L^2 space having "enough" decay at infinity, or "enough" regularity (degree of differentiability).

Example 3.6. ▷ The position operator $\mathbf{q}$ is, in each coordinate, multiplication by the unbounded function x^α. Let $\mathcal{H}_1 = \mathcal{H}_2 = L^2(\mathbb{R}, \mathbb{C})$. Recalling examples F.1 and F.2, and recalling that $\psi\overline{\psi}$ is a probability density function, let

$$\psi(x) = \sqrt{\frac{1}{\pi(x^2+1)}}.$$

Then

$$\int_{\mathbb{R}} \psi(x)^* \psi(x)\, dx = \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{x^2+1}\, dx = 1.$$

However,

$$\int_{\mathbb{R}} \psi(x)^*\, x\, \psi(x)\, dx = \frac{1}{\pi} \int_{\mathbb{R}} \frac{x}{x^2+1}\, dx$$

diverges because the integrals on each side of the origin are infinite (we do not have absolute integrability). On the other hand, let

$$\phi(x) = \sqrt{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}}$$

Then

$$\int_{\mathbb{R}} \phi(x)^*\, x\, \phi(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x\, e^{-x^2/2}\, dx = 0.$$

Thus, the position operator leaves only one of these two wavefunctions in $\mathcal{H}$. Note in particular that as $|x|$ increases, ϕ goes to zero more quickly than ψ does. ◁

Definition 3.7. The **operator norm** (or sometimes **uniform norm**) of a linear operator A is

$$\|A\|_{\mathrm{op}} := \sup_{f \in \mathcal{D}_A, \|f\|_{\mathcal{H}_1}=1} \|Af\|_{\mathcal{H}_2}$$

or equivalently

$$\|A\|_{\mathrm{op}} = \sup_{0 \neq f \in \mathcal{D}_A} \frac{\|Af\|_{\mathcal{H}_2}}{\|f\|_{\mathcal{H}_1}}.$$

Definition 3.8. A linear operator A is **bounded** (this is bad terminology but universally used) if

$$\|A\|_{\mathrm{op}} < \infty.$$

Exercise 3.9. Show that A is bounded iff it is continuous as a mapping $A : \mathcal{H}_1 \to \mathcal{H}_2$.

One can always extend A from $\mathcal{D}_A$ to $\overline{\mathcal{D}}_A$ by continuity. This means that if $\mathcal{D}_A$ is dense in $\mathcal{H}_1$ then a bounded operator can be uniquely extended to all of $\mathcal{H}_1$.

We have the usual concepts of **kernel** and **image**. If $\ker(A) = \{0\}$ then A^{-1} is defined as a linear map from $\mathrm{im}(A)$ to $\mathcal{D}_A$. The sum $A_1 + A_2$ is defined on the domain $\mathcal{D}_{A_1} \cap \mathcal{D}_{A_2}$. Likewise, if $A : \mathcal{H}_1 \to \mathcal{H}_2$ and $B : \mathcal{H}_2 \to \mathcal{H}_3$ then $B \circ A$ has domain of definition $\{f \in \mathcal{D}_A : Af \in \mathcal{D}_B\}$.

The **graph** of A, written $\mathcal{G}_A$, is the subspace of $\mathcal{H}_1 \oplus \mathcal{H}_2$ consisting of pairs $\{(f, Af) : f \in \mathcal{D}_A\}$. (Beware that the ordered-pair notation looks like an inner product!) Then $\mathcal{H}_1 \oplus \mathcal{H}_2$ is a Hilbert space with inner product

$$((f_1, f_2), (g_1, g_2))_{\mathcal{H}_1 \oplus \mathcal{H}_2} := (f_1, f_2)_{\mathcal{H}_1} + (g_1, g_2)_{\mathcal{H}_2}.$$

Exercise 3.10. Prove this claim.

This construction will be very useful below. The above definitions were probably familiar; here is the first new one.

Definition 3.11. A linear operator A is called **closed** if its graph $\mathcal{G}_A$ is closed in $\mathcal{H}_1 \oplus \mathcal{H}_2$.

This is an elegant definition, but akwward to apply; a more useful characterization is coming soon.

Definition 3.12. A linear operator A is **closable** if $\overline{\mathcal{G}}_A$ is a graph of a linear operator $\overline{A}$, called the closure of A (which typically has a larger domain of definition).

The more useful characterization of closedness is the following. It is presented as an exercise, but a very important one.

Exercise 3.13. A linear operator A is closed if whenever there is a sequence $f_n \in \mathcal{D}_A$ converging (in norm) to $f \in \mathcal{H}$ and $Af_n \to g \in \mathcal{H}_2$, we have $f \in \mathcal{D}_A$ and $Af = g$. (Think of this as a substitute for continuity.)

A weaker special case is the following: A is closable iff $f_n \to 0 \in \mathcal{D}_A$ and $Af_n \to g \in \mathcal{H}_2$ implies $g = 0$.

Remark 3.14. The graph $\mathcal{G}_A$ cannot contain pairs $(0, g)$ with $g \neq 0$, since it already contains the ordered pair $(0, 0)$. This is an obvious fact at face value, but is often useful in proofs.

We want a correct definition of self-adjoint operator $A : \mathcal{H} \to \mathcal{H}$ for use in quantum mechanics. We want H to be self-adjoint for

$$H = \frac{-\hbar^2}{2m} \nabla^2 + V(\mathbf{q}),$$

over a broad class of potentials $V(\mathbf{q})$. However, if the definition is too broad, it will include operators which are unmanageable.

A starting point is to require that

$$(Af, g) = (f, Ag)$$

for all $f, g \in \mathcal{H}$. If A satisfies this, it is said to be a **symmetric operator**. This is meaningful but (as von Neumann realized) this is not enough for a good spectral theorem. (Note that in finite-dimensional vector spaces, symmetry is equivalent to self-adjointness.) In fact, there exist operators which are symmetric but which are terrible to work with, and which cannot be the Hamiltonian of any reasonable physical system.

An operator will be self-adjoint if it is equal to its own adjoint, so we need to define adjoint. Return to $A : \mathcal{H}_1 \to \mathcal{H}_2$.

Definition 3.15. Let $A : \mathcal{H}_1 \to \mathcal{H}_2$ be a linear operator. Suppose that $\overline{\mathcal{D}}_A = \mathcal{H}_1$. Then an **adjoint operator** $A^* : \mathcal{H}_2 \to \mathcal{H}_1$ has domain $\mathcal{D}_{A^*}$ consisting of those $g \in \mathcal{H}_2$ for which there is some $g^* \in \mathcal{H}_1$ such that for all $f \in \mathcal{D}_A$,

$$(Af, g) = (f, g^*).$$

Note that we have defined the domain, but not the operator itself! We want to define the action of A^* on g. Hold g fixed. Consider

$$(Af, g) = (f, g^*).$$

Such a g^* is unique by density. If the above holds, we define $A^* g$ to be g^*. We can check that this is a linear operator on the above-specified domain, straight from definitions. Note that $\mathcal{D}_{A^*}$ could be $\{0\}$, but won't be in cases of interest.

Now suppose (this is a strong hypothesis!) that A is bounded and is defined on all of $\mathcal{H}_1$. Then the bounded functional

$$f \mapsto (Af, g)$$

is (by the Riesz representation theorem) always of the form

$$(f, g^*)$$

for a unique g^*. So, we directly define A^* on all of $\mathcal{H}_2$ by

$$(Af, g) = (f, A^* g).$$

This A^* is a bounded linear operator from all of $\mathcal{H}_2$ to $\mathcal{H}_1$: Note that

$$\|(Af, g)\| \leq \|A\|_{\mathrm{op}} \|f\|_{\mathcal{H}_1} \|g\|_{\mathcal{H}_2}$$

by the Cauchy-Schwartz inequality. This is the constant $\|A\|_{\mathrm{op}} \|g\|_{\mathcal{H}_2}$ times $\|f\|_{\mathcal{H}_1}$.

If A is not defined on on all of $\mathcal{H}_1$, then typically it is hard to explicitly describe $\mathcal{D}_{A^*}$.

Definition 3.16. We write $A_1 \subseteq A_2$ if $\mathcal{D}_{A_1} \subseteq \mathcal{D}_{A_2}$ and $A_1 = A_2|_{\mathcal{D}_{A_1}}$.

Exercise 3.17. Show that $A_1 \subseteq A_2$ implies $A_2^* \subseteq A_1^*$.

It can be shown that A^* is a closed operator, whether A is closed or not.

Definition 3.18. Let M be a linear subspace of $\mathcal{H}$. (Really, M can be merely a set.) Then define the **perpendicular space** (or **orthogonal complement**) of M to be

$$M^\perp := \{g \in \mathcal{H} : (f, g) = 0 \text{ for all } f \in M\}.$$

Proposition 3.19. $M^\perp$ *is always closed, whether M was or not.*

Proof. The condition survives the taking of the limit $g_n \to g$. $\qquad\square$

Exercise 3.20. Show that $\ker(A^*) = (\operatorname{im}(A))^\perp$.

Exercise 3.21. Show that A and A^* are both invertible iff $\ker(A) = 0$ and $\operatorname{im}(A)$ is dense in $\mathcal{H}_2$.

Remark 3.22. Show that $M^{\perp\perp} = \overline{M}$. In particular, $M^\perp = \{0\}$ implies $\overline{M} = \mathcal{H}$.

Proposition 3.23. A^* *is densely defined iff A is closable.*

Remark 3.24. In this case,
$$\overline{A} = A^{**}.$$

Proof. The proof uses the geometric construction of A^*, which we have not seen until now. First we restate the definition of A^* geometrically. The pair
$$(g, g^*) \in \mathcal{G}_{A^*},$$
for $g \in \mathcal{H}_2$ and $g^* \in \mathcal{H}_1$, iff
$$(g^*, -g) \perp \mathcal{G}_A.$$
The graph of A^* is $\mathcal{G}_{A^*} = \mathcal{G}_A^\perp$. Check this by writing out the definition of adjoint. For all $f \in \mathcal{D}_A$,
$$(f, g^*) + (Af, -g) = 0$$
i.e.
$$((f, Af), (g^*, -g))_{\mathcal{H}_1 \oplus \mathcal{H}_2} = 0.$$
(Note that we are using parentheses in two different ways here: the inner pairs are for two-tuples; the outer pair denotes an inner product.) This shows that A^* is a closed operator, since the graph is perpendicular to something. Alternatively, vectors of the form $(Ag^*, -g)$ constitute $\mathcal{G}_A^\perp$ and consequently $\overline{\mathcal{G}}_A$ is the orthogonal complement of these vectors.

The proof is now easy with this setup, and it goes similarly in both directions. If $\mathcal{D}_{A^*}$ is not dense in $\mathcal{H}_2$ then there is $0 \neq h \in \mathcal{H}_2$ such that $h \perp \mathcal{D}_{A^*}$, i.e.
$$(0, h) \perp (A^* g, -g),$$
for all $g \in \mathcal{D}_{A^*}$. This means that
$$(0, h) \in \overline{\mathcal{G}}_A$$
so $\overline{\mathcal{G}}_A \neq \mathcal{G}_A$. This means that A is not closable.

Now suppose that A is closable. Then A^* is densely defined, so A^{**} makes sense, and the graph of A^{**} is $\mathcal{G}_A^{\perp\perp} = \overline{\mathcal{G}}_A$. $\qquad\square$

Definition 3.25. A linear operator $A : \mathcal{H} \to \mathcal{H}$ with dense domain of definition $\mathcal{D}_A$ is called **symmetric** if
$$A \subseteq A^*,$$
i.e. if for all $f, g \in \mathcal{D}_A$,
$$(Af, g) = (f, Ag).$$

Note that since $\mathcal{D}_{A^*} \supseteq \mathcal{D}_A$, A^* is also densely defined. So, by the theorem we just proved, A is closable. That is to say, *a symmetric operator is closable.*

Proposition 3.26. *The expectation*

$$\langle A \rangle = (f, Af)$$

is real if A is symmetric. That is, observables are real.

Proof. Obvious. □

Definition 3.27. A densely defined operator $A : \mathcal{H} \to \mathcal{H}$ is **self-adjoint** if

$$A = A^*.$$

Note that adjoints are always closed, so a self-adjoint operator is necessarily closed.

3.3 Self-adjointness of multiplication operators

Here is a preview of the spectral theorem, involving multiplication operators by real-valued functions. Let $\mathcal{H} = L^2(M, \sigma)$ for some not necessarily linear space M and some σ-finite measure σ, with

$$(f, g) = \int_M \overline{f} g \, d\sigma.$$

Then $\mathcal{H}$ is a Hilbert space (this is a restatement of Riesz-Fischer). Let $a : M \to \mathbb{R}$ be a measurable function. Then $T_a : \mathcal{H} \to \mathcal{H}$ with (as usual)

$$\mathcal{D}_{T_a} = \{f \in \mathcal{H} : a \cdot f \in \mathcal{H}\}$$

i.e. f such that $a \cdot f$ is still square-integrable. On this domain of definition, $T_a f := a \cdot f$.

The following result is very important. It is similar in spirit to diagonalization of matrices. The heart of spectral theory is that *all* self-adjoint operators are one of these in disguise (up to unitary transformations). The reason for all the precise definitions in this section is that they capture operators of this form.

Proposition 3.28. T_a *is self-adjoint.*

Proof. Recalling definition 3.15, we first need to show that the domain $\mathcal{D}_{T_a}$ is dense in $\mathcal{H}$, i.e. that the closure of $\mathcal{D}_{T_a}$ is all of $\mathcal{H}$. Using remark 3.22, we will show that $\mathcal{D}_{T_a}^{\perp} = \{0\}$. Suppose $g \perp \mathcal{D}_{T_a}$; we want to show $g = 0$. We will start by proving a weaker result. Let f be arbitrary in $L^2(M, \sigma)$. For each positive integer N, let

$$M_N = \{m \in \mathcal{H} : |a(m)| \leq N\}.$$

Then

$$\int_{\mathcal{H}} \overline{g} f 1_{M_N} d\sigma = 0$$

since $f 1_{M_N} \in \mathcal{D}_{T_a}$ for all $f \in L^2(M, \sigma)$. We can write this in an equivalent way as

$$\int_{M_N} \overline{g} f d\sigma = 0$$

for all $f \in L^2(M, \sigma)$. This is true in particular for g. So,

$$\int_{M_N} \overline{g} g d\sigma = \int_{M_N} |g|^2 d\sigma = 0$$

which implies $g = 0$ σ-a.e. on M_N. Remember for below that the whole space $L^2(M, \sigma)$ is the countable union of the M_N's.

Second we show the self-adjointness of T_a. Let $g \in \mathcal{D}_{T_a^*}$. This means, by definition 3.15, that there exists a $g^* \in L^2(M, \sigma)$ such that for all $f \in \mathcal{D}_{T_a}$,

$$(g, T_a f) = (g^*, f).$$

(Recall that we then define T_a^* by $T_a^* g := g^*$ as we vary g.) In terms of integrals, this is

$$\int_{\mathcal{H}} \overline{g} T_a f d\sigma = \int_{\mathcal{H}} \overline{g} a f d\sigma = \int_{\mathcal{H}} \overline{g^*} f d\sigma.$$

For uniqueness of g^*, we need $\mathcal{D}_{T_a}$ to be dense in $L^2(M, \sigma)$, which we were partway through proving above. We want to conclude

$$g^* = ag;$$

then $g \in \mathcal{D}_{T_a}$ will follow automatically. The issue (and the reason for the M_N's) is that we don't yet know that ag is square-integrable. However, we do know that

$$\int_{\mathcal{H}} \overline{g} a f d\sigma = \int_{\mathcal{H}} \overline{g^*} f d\sigma$$

for a dense set of f's.

Use

$$\int_{M_N} \overline{g} a f d\sigma = \int_{M_N} \overline{g^*} f d\sigma$$

from which

$$\int_{M_N} (\overline{g} a - \overline{g^*}) f d\sigma = 0.$$

Note $\overline{g} a - \overline{g^*} \in L^2(M_N, \sigma)$; then take $N \to \infty$.

So $g \in \mathcal{D}_{T_a}$, and the action of the operator T_a^* is given by

$$T_a^* g = ag.$$

(Remember a is real-valued.) Since

$$T_a g = ag$$

and g was arbitrary in $\mathcal{D}_{T_a^*}$, T_a is self-adjoint. $\qquad \square$

3.4 Spectrum

Here we generalize the notion of eigenvalue. We define the spectrum of an operator A in a rather backward way: first we define what it means for a complex number *not* to be in the spectrum of A.

For motivation, recall that in finite dimensions, A is always representable by a matrix (which we also call A). Then z is not an eigenvalue of A iff $\ker(A - zI) = 0$, which is the case iff $(A - zI)^{-1}$ exists. We think of the spectrum of A, written $\sigma(A)$, as the eigenvalues of A for this finite-dimensional case.

Definition 3.29. Let A be a densely defined closed operator in $\mathcal{H}$. (Closedness is not necessary, but all our operators are closed.) We say that $z \in \mathbb{C}$ is in the **resolvent set** $\rho(A)$ if there exists a bounded operator

$$R_z := (A - zI)^{-1}$$

defined on all of $\mathcal{H}$. We call R_z the **resolvent operator** of A. (Note that it depends on A as well as z, but this dependence does not appear in the notation.)

The term *resolvent* comes about because of the following. Let $f, g \in \mathcal{H}$. Then $(A - zI)^{-1}g = f$ iff $(A - zi)f = g$ which is to say $Af - zf = g$, and z *resolves* this equation. (If A is a multiplication operator, this is an algebraic equation; if A is a differentiation operator, this is a differential equation.)

Remark 3.30. Physicists call R_z a **Green's function**. One asks: A function of what? To solve for z, one often obtains an integral operator for R_z with a kernel (in the integral sense of the word) which is a Green's function.

Definition 3.31. Let A be a densely defined operator in $\mathcal{H}$. The **spectrum** of A, written $\sigma(A)$, is

$$\sigma(A) := \mathbb{C} \setminus \rho(A).$$

We want to know the kernel and image of $(A - zI)$ in order to find out the domain of its inverse.

Proposition 3.32. *The resolvent $\rho(A)$ is an open set in $\mathbb{C}$, and the operator-valued function $z \mapsto R_z$ is an analytic function.*

Proof. Easy, but deferred until later. The proof is important because it introduces methods of complex analysis.

First we need to define analyticity for operator-valued functions.

Definition 3.33. The operator-valued function $z \mapsto R_z$ is **analytic** if for all z sufficiently close to $z_0 \in \rho(A)$,

$$R_z = \sum_{n=0}^{\infty} C_n(z - z_0)^n$$

with the C_n's all bounded operators defined on all of $\mathcal{H}$, and with convergence in the metric induced by the operator norm.

A sketch of the proof is that one tries to form the inverse $1/(A - zI)$ algebraically. With an eye to using a geometric series we have

$$\frac{1}{A - z_0 I - (z - z_0)I}.$$

Factor out and continue. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

The following corollary is easy.

Corollary 3.34. *Fix $f \in \mathcal{H}$ and take the quadratic form $z \mapsto (f, R_z f)$. This complex-valued function is analytic in $\rho(A)$ for all $f \in \mathcal{H}$.*

This analyticity will be crucial: it will provide a link between classical complex analysis and modern spectral theory.

Proposition 3.35 (Hilbert resolvent identity)**.** *For all $z, z' \in \rho(A)$,*

$$R_z - R_{z'} = (z - z')R_z R_{z'}.$$

Proof. Write $R_z = (1/A - z)$ and expand as a geometric series. Do the algebra, then check it works for operators. Then write $R_z - R_{z'}$ in the following artificial but useful way:

$$R_z - R_{z'} = R_z(A - z'I)R_{z'} - R_{z'}$$

Note that the first term on the right-hand side is the identity on $\mathcal{H}$. This is

$$
\begin{aligned}
R_z - R_{z'} &= R_z(A - zI + (z - z')I)R_{z'} - R_{z'} \\
&= R_{z'} + R_z(z - z')R_{z'} - R_{z'} \\
&= (z - z')R_z R_{z'}.
\end{aligned}
$$

$\square$

Corollary 3.36. *Resolvent operators on the same A commute with one another:*

$$
R_z R_{z'} = R_{z'} R_z.
$$

Of course, this result could have been obtained without making a corollary of the proposition.

[xxx make fwd xrefs to here from the SA section.]

Proposition 3.37. *Suppose $A : \mathcal{H} \to \mathcal{H}$ is self-adjoint. Then:*

(1) $\sigma(A) \subseteq \mathbb{R}$. In particular, R_z is defined on the open half-planes

$$
\mathbb{C}^+ := \{\operatorname{Im}(z) > 0\} \qquad and \qquad \mathbb{C}^- := \{\operatorname{Im}(z) < 0\}.
$$

(2) Recall that R_z is bounded so the adjoint R_z^* exists. We have

$$
R_z^* = R_{\overline{z}}.
$$

Moreover,

$$
\|R_z\|_{\mathrm{op}} \leq |\operatorname{Im}(z)|^{-1}.
$$

Remark. This result is emphatically not true for an arbitrary symmetric operator; self-adjointness is crucial.

Proof. From the definitions, using the antilineary of the inner product, we have

$$
(A - zI)^* = A^* - \overline{z}I.
$$

Since adjoint and inverse commute, it follows that

$$
R_z^* = R_{\overline{z}}.
$$

Now we calculate

$$
\|(A - zI)f\|^2 = (Af - zf, Af - zf).
$$

Let $z = x + iy$. Then

$$
\|(A - zI)f\|^2 = (Af, Af) - x(Af, f) - iy(Af, f) - x(f, Af) + iy(f, Af) + |z|^2(f, f).
$$

Now, (Af, Af) and (f, f) are non-negative and real by the positive-definiteness of the inner product. The terms with x in them are complex conjugates of one another, and x is real, so $x(Af, f) - x(f, Af)$ is real. We have some real number, minus $iy(Af, f)$, plus $iy(f, Af)$, so these latter two terms sum to zero. We claim that

$$
\|(A - zI)f\|^2 = \|(A - xI)f\|^2 + y^2\|f\|^2 = \|(A - zI)f\|^2.
$$

We want to show this is not too small because we are interested in the norm of the inverse operator. The result is

$$
\|(A - zI)f\|^2 \geq y^2\|f\|^2.
$$

We did not assume $z \in \rho(A)$; we only assumed $\text{Im}(z) \neq 0$ and we want to conclude that $z \in \rho(A)$. We know $\ker(A - zI) = 0$ by the inequality. Look at the range of the operator $(A - zI)$. We want R_z defined on all of $\mathcal{H}$: if the kernel of $A - zI$ is zero and if it is onto, then it is invertible on all of $\mathcal{H}$. We have

$$\text{im}(A - zI)^\perp = \ker((A - zI)^*)$$

by an exercise from above [xxx xref]. Using self-adjointness of A, this is

$$\text{im}(A - zI)^\perp = \ker((A - zI)^*) = \ker(A^* - \bar{z}I) = \ker(A - \bar{z}I).$$

Note that $\text{Im}(z) \neq 0$ iff $\text{Im}(\bar{z}) \neq 0$. So $\text{im}(A - zI)$ is dense in $\mathcal{H}$, but this operator is closed so the image is all of $\mathcal{H}$. Thus $(A - zI)^{-1}$ exists. Also we have

$$\|(A - zI)^{-1}\} \leq \frac{1}{y} = \text{Im}(z)^{-1}$$

as desired. $\qquad\square$

Here is another property the resolvent operator satisfies.

Proposition 3.38. *Fix $0 \neq f \in \mathcal{H}$. Consider the function*

$$\Phi(z) = (f, R_z f).$$

This is defined on $\rho(A)$ so it is at least defined on $\mathbb{C}^+$ and $\mathbb{C}^-$ (as well as perhaps parts of the real line). This function is analytic and maps $\mathbb{C}^+$ and $\mathbb{C}^-$ into themselves, respectively.

Furthermore, $\Phi(z)$ satisfies

$$\Phi(\bar{z}) = \overline{\Phi(z)}$$

and

$$|\Phi(\bar{z})| \leq (f, f)\text{Im}(z)^{-1}.$$

Remark. We call $\Phi(z)$ a **Herglotz function**.

Proof. The last two points are obvious from the previous theorem; use the Cauchy-Schwartz inequality.

To see that $\Phi(z)$ maps $\mathbb{C}^\pm$ into $\mathbb{C}^\pm$, let

$$R_z f = g.$$

Then

$$f = (A - zI)g.$$

In particular, $g \neq 0$. Also,

$$(f, R_z f) = ((A - zI)g, g) = (Ag, g) - \bar{z}(g, g).$$

Since A is symmetric, (Ag, g) is real. Therefore the imaginary part of the left-hand side is $\text{Im}(\bar{z}(g, g))$. Chasing signs, we find that the right-hand side has imaginary part $\text{Im}(z)(g, g)$. $\qquad\square$

There is a representation theorem for Herglotz functions: for each $\Phi(z)$ there is a $\sigma(x)$ such that we have the following Lebesgue-Stieltjes integral:

$$\Phi(z) = \int_{\mathbb{R}} \frac{d\sigma(x)}{x - z}$$

Here $\sigma(x)$ is right-continuous and non-decreasing, with $\lim_{x \to \infty}(x) = 0$ and and $\lim_{x \to +\infty}(x) < +\infty$. The measure has point masses at discontinuities of σ. As may be seen from the integral representation, the analyticity of $\Phi(z)$ breaks down as z approaches the real line; the function $\sigma(x)$ will encapsulate the continuity and/or discreteness of the spectrum $\sigma(A)$.

3.5 Preview of the spectral theorem

The following analogy is an illustration of the spectral theorem on a finite-dimensional vector space. This makes it easier to follow the proof of the spectral theorem, in particular the choice of M and σ which arise there.

For this section, let $\mathcal{H}$ be an N-dimensional complex Hilbert space. Let $A : \mathcal{H} \to \mathcal{H}$ be self-adjoint (which is equivalent to symmetric in finite dimension). As is well-known, there exists an orthonormal basis of eigenvectors $\{f_i\}$ with eigenvalues a_i such that for $i = 1, 2, \ldots, N$, $Af_i = a_i f_i$. Assume for now that A has a *simple spectrum*, i.e. all the eigenvalues of A are distinct. For each $f \in \mathcal{H}$ there exists a unique representation

$$f = \sum_{i=1}^{N} c_i f_i.$$

Define an operator U trading f for the N-tuple $(c_1, \ldots, c_N) \in \mathbb{C}^N$. Since the f_i's form an orthonormal basis, by writing out the inner product one may show that

$$\|f\|_{\mathcal{H}}^2 = \sum_{i=1}^{N} |c_i|^2.$$

Thus, we may think of $\mathbb{C}^N$ as an L^2 space. However, we do this in a particular way: we write

$$\mathbf{c} = (c_1, \ldots, c_N) \in L^2(\{a_1, \ldots, a_N\}).$$

That is, we put the c_i's as a graph over the a_i's. (This is why we assume the a_i's are distinct.) Then U is a unitary transformation from $\mathcal{H} \to L^2(\{a_1, \ldots, a_N\})$.

This elementary construction achives the same goal as in the spectral theorem. We define $T_a : L^2 \to L^2$ by

$$[T_a(\mathbf{c})]_i = a_i c_i$$

and then we write

$$A = U^{-1} T_a U.$$

Now suppose there are $N/2$ eigenvalues, each of multiplicity 2. Do the above, but with the space

$$L^2(a_1, \ldots, a_{N/2}) \oplus L^2(a_1, \ldots, a_{N/2}).$$

One may generalize to the case of arbitrary multiplicities. In all cases, the L^2 case is the spectral theorem's M.

3.6 Spectral theorem

4 Multi-body systems

Note as of Sep. 27 2007: I had some difficulty following the discussion of spectral theory and will defer typing up any notes until I have some good examples in hand. Also, as of about a week ago, the discussion in class has been closely following the text ([**MR**]), so there I feel reduced urgency for these notes. However I do feel that Martin and Rothen's text could bear some worked examples, which I will include here as I have them.

A Linear algebra

Do as many Hilbert-space constructions as possible in finite dimensions. Symmetrize and antisymmetrize; occupation-number basis. Normalizations.

B   Tensors

Keep it brief and practical. State (with don't-be-scared caveat) the abstract-algebra definition. Then immediately give examples showing that these are just pairs (or n-tuples) with the manipulation rules (which is the practical meaning of the equivalence relations) of scalar-through and multilinearity. These give different ways to write the same thing, and give us some flexibility for computations.

Tensor product of functions in general: given

$$f_1, \ldots, f_n : X \to k$$

where k is a field,

$$(f_1 \otimes \cdots \otimes f_n)(\mathbf{x}_1, \ldots, \mathbf{x}_n) := f_1(\mathbf{x}_1) \cdot \cdots \cdot f_n(\mathbf{x}_n).$$

Maybe talk about vectors and covectors (and how they consume one another by duality) but *only* if necessary for the discussion in this course.

C Analysis results

Cite whom ....

Theorem C.1 (Riesz representation theorem)**.** *Let $\mathcal{H}$ be a Hilbert space. For every $\lambda \in \mathcal{H}^*$ (i.e. a continuous linear operator $\lambda : \mathcal{H} \to \mathbb{R}$) there exists a unique $\phi \in \mathcal{H}$ such that for all $\psi \in \mathcal{H}$,*

$$\lambda(\psi) = (\phi, \psi).$$

Mnemonic: Every row vector comes from a column vector.

Theorem C.2 (Riesz-Fischer theorem)**.** *A function is square integrable (i.e. is in L^2) iff its Fourier coefficients are square-summable (i.e. are in ℓ^2).*

Proof. Sketch; needs words.

Pick (square-summable but otherwise) arbitrary $\psi \in L^2$. [cite what] L^2 has an orthonormal basis $\{\phi_k : k = 1, 2, \ldots\}$. Then

$$(\psi, \psi) = \|\psi\|^2 < \infty$$

$$(\psi, \psi) = \left(\sum_{j=1}^{\infty} c_j \phi_j, \sum_{k=1}^{\infty} c_k \phi_k \right)$$

$$= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} c_j \bar{c}_k \, (\phi_j, \phi_k)$$

$$= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} c_j \bar{c}_k \delta_{jk}$$

$$= \sum_{j=1}^{\infty} |c_j|^2 < \infty.$$

$\square$

D Fourier expansion

D.1 Finite dimensions

Finite dimensional vector space V: column vectors and row vectors. Finite sums; trivial rep theory. Bases for V and V^*:

$$\{\mathbf{e}_1, \ldots, \mathbf{e}_N\} \quad \text{and} \quad \{\mathbf{e}_1^*, \ldots, \mathbf{e}_N^*\}.$$

For $\mathbf{v} \in V$:

$$\mathbf{v} = \sum_{j=1}^{N} a_j \mathbf{e}_j.$$

For $\lambda \in V^*$:

$$\lambda = \sum_{j=1}^{N} b_j \mathbf{e}_j^*.$$

Recall that for an orthonormal basis (namely, $(\mathbf{e}_j, \mathbf{e}_k) = \delta_{jk}$), it's easy to find the a_j's by the sesquilinearity of the inner product:

$$(\mathbf{v}, \mathbf{e}_j) = \left(\sum_{j=1}^{N} a_k \mathbf{e}_k, \mathbf{e}_j \right) = \sum_{j=1}^{N} a_k (\mathbf{e}_k, \mathbf{e}_j) = \sum_{j=1}^{N} a_k \delta_{kj} = a_j.$$

D.2 Infinite dimensions

Separable Hilbert space $\mathcal{H}$; countable basis. (Example: $L^2([0,1])$ has basis functions $\phi_j(x) e^{i2\pi jx}$ for $j \in \mathbb{Z}$.) Cite Riesz rep thm for "row vectors" in Hilbert space. For $\psi \in \mathcal{H}$,

$$\psi = \sum_{j=1}^{\infty} c_j \phi_j.$$

Cite Riesz-Fischer for

$$\sum_{j=1}^{\infty} |c_j|^2 < \infty.$$

When $\mathcal{H} = L^2(\mathbb{R}^3)$, equality as functions means for all $x \in \mathbb{R}^3$,

$$\psi(x) = \sum_{j=1}^{\infty} c_j \phi_j(x).$$

Just as in the finite-dimensional case, it's easy to find the coefficients if the basis is orthonormal:

$$(\psi, \phi_k) = \left(\sum_{j=1}^{N} c_j \phi_j, \phi_k \right) = \sum_{j=1}^{N} c_j (\phi_j, \phi_k) = \sum_{j=1}^{N} c_j \delta_{jk} = c_k.$$

For the case $\mathcal{H} = L^2(\mathbb{R}^3)$, the inner product is represented as an integral and we have

$$(\psi, \phi_k) = \int_{\mathbb{R}^3} \psi(\mathbf{x}) \overline{\phi}_k(\mathbf{x}) \, d\mathbf{x}.$$

D.3   Tensor products of finite-dimensional vector spaces

Do this just as a warm-up and intuition builder for the next subsection.

Oddly and only to ease the transition below, think of

$$V = L^2(\{1, 2, \ldots, N\}).$$

For each index $j = 1, 2, \ldots, N$, we have a complex number at that index. The square-summability condition removes nothing, since all the sums are finite.

Then as is well-known (one hopes!) from abstract algebra, if V has basis $\{\mathbf{e}_j : j = 1, 2, \ldots, N\}$, then $V \otimes V$ has basis

$$\{\mathbf{e}_j \otimes \mathbf{e}_k : j, k = 1, 2, \ldots, N\}.$$

This basis has N^2 elements, and the nth tensor power of V with itself has dimension N^n. I hope this makes it palatable that we can think of

$$V \otimes V = L^2(\{1, 2, \ldots, N\}^2).$$

and in general

$$V^{\otimes n} = L^2(\{1, 2, \ldots, N\}^n).$$

For $F \in V \otimes V$,

$$F = \sum_{j=1}^{N} \sum_{k=1}^{N} c_{jk} \mathbf{e}_j \otimes \mathbf{e}_k;$$

for $F \in V^{\otimes n}$,

$$F = \sum_{r_1=1}^{N} \cdots \sum_{r_n=1}^{N} c_{r_1, \ldots, r_n} \mathbf{e}_{r_1} \otimes \cdots \otimes \mathbf{e}_{r_n}.$$

Something about selection of coordinates ... what is the appropriate notion of inner product? It's done by product of inner products on slots. Write down the formula and maybe an example; prove what is reasonable to prove.

D.4   Tensor products of infinite-dimensional vector spaces

Tensor product of vector spaces: If $\mathcal{H} = L^2(\mathbb{R}^3)$, spell out the identification of $\mathcal{H}^{\otimes n}$ with $L^2(\mathbb{R}^{3n})$. Unlike in the finite-dimensional case, the square-integrability is not for free and we have something to prove.

Certainly $\mathcal{H}^{\otimes n}$ has basis

$$\{\phi_{r_1} \otimes \cdots \otimes \phi_{r_n} : r_1, \ldots, r_n = 1, \ldots, N\}.$$

Define

$$(\phi_j \otimes \phi_k)(\mathbf{x}_1, \mathbf{x}_2) := \phi_j(\mathbf{x}_1)\phi_k(\mathbf{x}_2).$$

As long as $\mathcal{H}^{\otimes n}$ is separable, it has a countable basis $\{\theta_j : j = 1, 2, \ldots\}$. State why (use Fubini's theorem) that $\mathcal{H}^{\otimes n}$ has basis

$$\{\phi_{r_1} \cdot \cdots \cdot \phi_{r_n} : r_1, r_2, \ldots = 1, 2, \ldots\}.$$

Prove these product functions form a complete orthonormal system (w/r/t which inner product?).

E Spectra

E.1 Finite-dimensional case

Cite and follow [**Lax**]. Expand vector $\mathbf{x}$ and $A\mathbf{x}$ in terms of eigenbasis. Sum up.

E.2 Infinite-dimensional case

Continuing to follow Lax and citing the previous section, write out the **spectral resolution** of the operator: replace sum with integral. Define **operator-valued measure**. List its axioms. Cite and follow the Wikipedia article on **projection-valued measure**. Make some mention of **point spectrum** and **continuous spectrum**. How to interpret **singular spectrum**?

F   Probability

See [**Gri**], [**Ker**], [**GS**], and [**FG**], respectively, for treatments of increasing levels of rigor.

F.1   Probability spaces and probability measures

$(\Omega, \mathcal{F}, P)$ from prb MT appx. Meas. space, prob. space; $\mathcal{F} = \mathrm{Bo}(\mathbb{R})$.

F.2   Random variables

fcn from prb spc to meas spc. make sure to note special case $\Omega = \mathbb{R}^d$ and X is identity function. My convention about "random variable" in this document.

$P(X \in [a, b]) = P(X^{-1}([a, b]))$.

$P(X \in A) = P(X^{-1}(A))$.

F.3   Distribution functions, mass functions, and density functions

CDFs $F_X(x)$.

PDFs when they exist: $f_X(x)$. CRVs.

PMFs when they exist: $f_X(x)$. DRVs. (Need both for QM.)

$P(X \in A) = \int_A f_X(x)\, dx$ for all Borel A.

Example F.1. ▷ The **Cauchy random variable** X has PDF

$$f_X(x) = \frac{1}{\pi}\, \frac{1}{x^2 + 1}.$$

You can check using freshman-calculus techniques that this has integral 1.   ◁

Example F.2. ▷ The **standard normal random variable** X has PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}.$$

This also has integral 1.   ◁

What is the relationship between discrete/continuous CDFs and discrete/continuous spectra? Spell it out here or in a separate appendix. Is there any relationship at all? What about empty spectrum? What is "residual spectrum"? Wiki articles on decomposition of spectrum, Borel functional calculus, Lebesgue's decomposition theorem, Stone's theorem on one-parameter unitary groups.

F.4   Expectations

$E[X]$ or $\langle X \rangle$: $\sum x f_X(x)$ or $\int x f_X(x)\, dx$. Absolute-integrability requirement.

$\mathrm{Var}(X) = E[X^2] - E[X]^2$. $\sigma_X(x) = \sqrt{\mathrm{Var}(X)}$.

Example F.3. ▷ The Cauchy random variable from example F.1 has infinite mean: use u-substitution on the parts $x \geq 0$ and $x \leq 0$ and you will find that the mean is undefined because $x\,f_X(x)$ is not absolutely integrable. The problem is that $f_X(x)$ goes to zero at infinity but not fast enough. ◁

Example F.4. ▷ The standard normal random variable from example F.1 has mean 0, as you can verify using u-substitution. Unlike the Cauchy random variable's PDF, this PDF goes to zero at infinity at a higher rate. ◁

F.5   Independence

IID

F.6   The IID paradigm

Let X be a sequence of IID RVs. Form $\overline{X}_n = \sum X_i/n$. (This is called the **sample mean**.) Cite the LLN. Note that $\langle \overline{X}_n \rangle = \langle X \rangle$. Also $\mathrm{Var}(\overline{X}_n) = \sigma^2/n$, so $\sigma(\overline{X}_n) = \sigma/\sqrt{n}$. Emphasize that this is the standard deviation *of the mean itself.*

Example: Experiment is 1000 tosses of a fair coin. (Or, toss 1000 fair coins all at once). Then count heads. 1st time: Maybe 510. 2nd time: 493. 3rd time: 507. Etc. Do *that* many times and the average of your results will tend toward 500. Compute the sample standard deviations of your data; the result will tend toward will be $\sqrt{4000}/\sqrt{N}$. Elaborate.

Caveat about wave-function collapse for QM — do *identical* experiments, not *repeated* experiment.

basic facts about probability functions, with refs

expectations, mean, median ...

variance and stddev

multiple RVs: mention joint PDF.

maybe mention LLN and CLT

PRB, FA/LA, PDEs/ODEs.

Cplx PDF: $\langle Q \rangle = \sum / \int \phi^* Q \phi$

F.7   Multiple random variables

joint CDF and joint PDF.

References

Quantum mechanics:

[**CT**] Cohen-Tannoudji, C. et al. *Quantum Mechanics*. Wiley Interscience, 2006.

[**Gri**] Griffiths, D.J. *Introduction to Quantum Mechanics* (2nd ed.). Pearson Prentice Hall, 2005. This is an undergraduate text. I find it spectacularly clear, but it doesn't use the language of graduate mathematics; this is an introductory text.

[**Lax**] P. Lax. *Functional Analysis*. Wiley Interscience, 2002.

[**Mes**] Messiah, A. *Quantum Mechanics*. Dover, 1999.

[**MR**] Martin, Ph.A. and Rothen, F. *Many-Body problems and Quantum Field Theory* (2nd ed.). Springer, 2004. This is the official textbook for the course.

[**Sak**] Sakurai, J.J. *Modern quantum mechanics* (2nd ed.). Addison Wesley, 1994.

Probability:

[**FG**] Fristedt, B. and Gray, L. *A Modern Approach to Probability Theory*. Birkhauser, 1997.

[**GS**] Grimmett, G. and Stirzaker, D. *Probability and Random Processes*, 3rd ed. Oxford, 2001.

[**Ker**] Kerl, J. *Probability notes*. `http://math.arizona.edu/~kerl/doc/prb.pdf`

Hilbert spaces:

[**AG**] Akhiezer, N.I. and Glazman, I.M. *Theory of linear operators in Hilbert Space*. Dover, 1993.

[**BS**] Berezin, F.A. and Shubin M.A.. *The Schrödinger Equation*. Springer, 1991.

[**RS**] Reed, M. and Simon, B. *Methods of Modern Mathematical Physics* (rev. ed.). Academic Press, 1980.

Index