Notes for 8.321

Lectures by Senthil Todadri; notes taken by Sanath Devalapurkar

Abstract. These are my notes from 8.321, graduate quantum mechanics, offered in Fall 2018 at MIT, and taught by Senthil Todadri. The usual disclaimers apply: any mistakes are mine, and I've failed to faithfully reproduce the jokes and side discussions that took place during class. These notes are in no way official.

1. Introduction

We have a midterm on October 22, in the same room (4-265). Homework is assigned every Monday, and due the following Tuesday. It's worth 70% of your grade. Don't take this class if you've already taken 8.04–8.06. Let's get started. Today, we'll talk about the postulates of quantum mechanics, and talk about the mathematical preliminaries. Nothing about history, though.

1.1. Weirdness in quantum mechanics. Here's an example of a "weird" phenomenon in quantum mechanics. Imagine an atom with spin 1/2. Imagine we measure the z-component of spin for this atom. What quantum mechanics tells us is that you always find one of two values: $S^z = \pm\hbar/2$. If anyone ever says "this is true", you should always ask how they measured it. So: how do we measure S^z?

There are many ways to do this, but there's an old setup by Stern and Gerlach (the "Stern-Gerlach apparatus"). You have a beam of atoms, and you make it go through an inhomogeneous magnetic field pointing along the z-axis. Then, the beam of atoms will deflect. This'll tell you what its spin is. The reason is that $E = -\vec{\mu} \cdot \vec{B}$. This will be $-\mu_z B$, where $\vec{\mu}$ is the magnetic moment of the atom, which is proportional to the spin $\vec{S}$. Therefore, $F = -\partial E/\partial z = \mu_z \partial B/\partial z$. It follows that the sign of F depends precisely on the spin of μ_z. In other words, different μ_Z states have different vertical displacements.

We can also arrange this so that the inhomogeneous magnetic field is pointing along the x and y axes. Suppose you have Stern-Gerlach apparati in all direction. This'll allow you to measure any component of spin. For a spin 1/2 atom, we have $S^z = \pm\hbar/2$. Suppose we come in with our beam, and put it through a Stern-Gerlach apparatus which separates the z component. Then, let us restrict to the positive S^z component. Then, let us measure the z component again. Common sense would say that all of these atoms lie in the resulting positive S^z channel, and nothing in the negative S^z channel.

Alernatively, we could do the following. As before, run our beam through a device filtering the z direction. Block the negative S^z beam. Now, measure in the x direction. What quantum mechanics says that you'll get two beams, with both positive and negative S^x. We can repeat what we did earlier: we block the negative S^x beam, and measure S^z again. Now the prediction

from quantum mechanics is that there are both positive and negative S^z beams! This is very weird from a classical perspective.

The upshot is that not all things that you measure simultaneously get the saame answers. This issue of uncertainty in measurement is what makes quantum mechanics diverge from classical mechanics. Let's now move on to the postulates; we'll spend the next several weeks trying to understand them and their consequences.

1.2. Postulates of quantum mechanics.

(1) The state of a quantum mechanics at time t is given by a vector (ray) $|\psi\rangle$ in a complex Hilbert space $\mathcal{H}$.
(2) The observables are Hermitian operators in $\mathcal{H}$ whose eigenvectors form a complete set.
(3) A measurement of $\widehat{A}$ returns one of its eigenvalues.
   (a) If $\widehat{A}$ is measured in a state $|\psi\rangle$, then the probability that you get a is $\langle\psi|M_a|\psi\rangle$, where $M_a = \sum_{j|a_j=a} |a_j\rangle\langle a_j|$ is the "measurement operator".
   (b) After measurement, the system is in a state $|\widetilde{\psi}_a\rangle \propto M_a|\psi\rangle$. (In other words, the proportionality constant ensures that the state is normalized.) This is called the "collapse of the wavefunction".
(4) Time evolution is given by the map $|\psi(t)\rangle \mapsto |\psi(t')\rangle = U(t',t)|\psi(t)\rangle$, where $U(t',t)$ (known as the *time evolution operator*) is a unitary operator, i.e., $U^\dagger U = 1$. If $t' = t+\epsilon$, then you find the Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle,$$

where H is the Hermitian operator known as the Hamiltonian.

How did we end up at these postulates? Nobody's derived them, they are postulates! There are two different reasons for liking this formulation: it's internally consistent (i.e., they are not mutually incompatible), so it's a sensible theory. It also agrees beautifully with experiment. Problems in either front implies that you have a bad theory, but we do not know of any logical inconsistency in this formulation. As far as experiments are concerned, quantum mechanics passes with flying colors everyday. Every once in a while, people try to change these things.

1.3. The mathematical preliminaries. This might be a review for some of you. But if not, this is a chance for you to internalize it. Let's begin by defining the concept of a Hilbert space. I'm mostly only going to talk about vector spaces. Later on, we'll hint at some of the more sophisticated thinking that goes into the special kind of structure that goes into a Hilbert space.

We begin by recalling the definition of a vector space V. This is a collection of objects $|\alpha\rangle$ with the following properties. There is an an addition law "+" such that:

(1) $|\alpha\rangle + |\beta\rangle = |\gamma\rangle \in V$ and $|\gamma\rangle$ is unique.
(2) $|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle$, i.e., the addition is commutative.
(3) $(|\alpha\rangle + |\beta\rangle) + |\gamma\rangle = |\alpha\rangle + (|\beta\rangle + |\gamma\rangle)$, i.e., the addition is associative.
(4) There is a vector $|0\rangle$ such that $|0\rangle + |\alpha\rangle = |\alpha\rangle$ for all $|\alpha\rangle \in V$.
(5) For all $|\alpha\rangle \in V$, there is $-|\alpha\rangle \in V$ such that $|\alpha\rangle + (-|\alpha\rangle) = |0\rangle$.
(6) For some field F (we'll only be interested in $F = \mathbf{R}, \mathbf{C}$), we can multiply $c \in F$ with $|\alpha\rangle \in V$ to get $c|\alpha\rangle \in V$. Moreover, the following properties are satisfied:
   (a) $c(d|\alpha\rangle) = cd|\alpha\rangle$;
   (b) $1|\alpha\rangle = |\alpha\rangle$;
   (c) $c(|\alpha\rangle + |\beta\rangle) = c|\alpha\rangle + c|\beta\rangle$;
   (d) $(c+d)|\alpha\rangle = c|\alpha\rangle + d|\alpha\rangle$.

If $F = \mathbf{R}$, then the vector space is called a real vector space, and similarly for $F = \mathbf{C}$. Let us look at some examples.

(1) Vectors in $\mathbf{R}^n$ form a real vector space.
(2) The space of states of a spin 1/2 system forms a vector space, in the following way: all vectors are of the form $c_+|+\rangle + c_-|-\rangle$, where $|+\rangle, |-\rangle$ are basis vectors, and $c_\pm \in \mathbf{C}$. This is a complex vector space.
(3) Yet another example, more entertaining perhaps, is the space of all functions on $[0, 1]$. Clearly we can add and multiply two functions, and scale by a scalar. It's a more complicated vector space than the two examples above.

Having defined a vector space, let's define the concept of a subspace. A subset $V \subset W$ of a vector space W is called a subspace if V itself is a vector space. A ray is a one-dimensional subspace of the form $\{c|\alpha\rangle\}$ for some $|\alpha\rangle \in V$.

The next concept I want to recall is that of linear independence. Say that $|\alpha_1\rangle, \cdots, |\alpha_n\rangle$ are linearly independent if and only if $\sum_{j=1}^n c_j|\alpha_j\rangle = 0$ implies that $c_j = 0$. Let's define a maximal linearly independent set to be a linearly independent set $\{|\alpha_j\rangle\}$ such that there is no $|\beta\rangle$ for which $\{|\alpha_j\rangle, |\beta\rangle\}$ is a linearly independent set. The set $\{|\alpha_j\rangle\}$ is called a basis for V. The number of elements such a basis is the dimension of V. The dimension of a vector space can be finite, countable infinite (sometimes called "denumerably infinite"), or uncountably infinite.

Let's again look at some examples.

(1) Of course, $\mathbf{R}^n$ has finite dimension (n itself).
(2) Consider the space of square integrable functions in $[0, 1]$ with $f(0) = f(1) = 0$. You can check that this forms a vector space. What's its dimension? The theory of Fourier series gives you a basis: you can expand f as

$$f(x) = \sum_{k=1}^{\infty} c_k \sin(\pi k x),$$

so you can take the functions $\sin(\pi k x)$ as a collection of basis vectors. It follows that the space is countably infinite.
(3) Consider the space of square integrable functions on $[-\infty, \infty]$. What's its dimension? Well, it's clearly not finite; is it countably infinite-dimensional? It's actually countably infinite-dimensional. From a physics point of view, the way to see this is as follows. Such a space can be thought of as wavefunctions in a quantum-mechanical system (where the square integrability corresponds to asking that the wavefunction be normalized). We know that every such wavefunction is a linear combination of the eigenstates of the simple harmonic oscillator; but there are only countably many such eigenstates (indexed by the integers).

    Let me plant a small seed of confusion. Since you know some quantum mechanics, you might be tempted to take a different basis: the collection of position eigenstates. The wavefunction is the Dirac delta function. But it's not clear that you get a countable basis — the delta function can be positioned anywhere. The upshot is that the position eigenstates are not square integrable, so those functions do not lie in the vector space itself! We still get away with it in quantum mechanics, and it's one of the sins that physicists commit.

If $|\alpha_1\rangle, \cdots, |\alpha_n\rangle$ form a basis for V, then any vector $|\beta\rangle$ can be expanded as a sum $|\beta\rangle = \sum_i c_i|\alpha_i\rangle$ for some scalars c_i.

So far, we've just defined a vector space and some properties. I want to further enrich this definition by adding more structure. The thing I want to define is the notion of an inner

product. It's supposed to generalize the notion of the scalar product. An inner product is a map $(-,-): V \times V \to \mathbf{C}$ (or whatever your field is), with the following properties.

(1) $(|\alpha\rangle, c|\beta\rangle) = c(|\alpha\rangle, |\beta\rangle)$ and $(c|\alpha\rangle, |\beta\rangle) = c^\star(|\alpha\rangle, |\beta\rangle)$;
(2) $(|\alpha\rangle, |\beta\rangle)^\star = (|\beta\rangle, |\alpha\rangle)$;
(3) $(|\alpha\rangle, |\beta\rangle + |\beta'\rangle) = (|\alpha\rangle, |\beta\rangle) + (|\alpha\rangle, |\beta'\rangle)$;
(4) $(|\alpha\rangle, |\alpha\rangle) \geq 0$;
(5) $(|\alpha\rangle, |\alpha\rangle) = 0$ implies that $|\alpha\rangle = |0\rangle$.

Here are some examples.

(1) If $V = \mathbf{C}^n$, then $|z\rangle = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ and $z_i \in \mathbf{C}$, with $(|z\rangle, |w\rangle) = z_1^\star w_1 + \cdots + z_n^\star w_n$.
(2) Let V be the space of complex functions on $[0,1]$. Define the inner product via

$$(f,g) = \int_0^1 f^*(x) g(x) dx.$$

We'll use the word "kets" to refer to vectors. Kets are said to be "orthogonal" when their inner product is zero. The norm of $N = \sqrt{(|\alpha\rangle, |\alpha\rangle)} = ||\alpha||$.

Recitation 1

Generalities. I'm pretty sure that you've seen this equation:

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = \widehat{H}|\psi(t)\rangle. \tag{1}$$

This is called the time-dependent Schrödinger equation. If you think about it from the classical perspective, the state vector is a point in phase space. It contains almost all of the information about the quantum space. The operator $\widehat{H}$ is called the Hamiltonian operator.

This equation is hard to solve in general — the Hamiltonian is very complicated in general. The space of all possible states (a "Hilbert space") can be very large (infinite-dimensional, in many cases). The Hamiltonian can be a function of time, and this makes the equation a lot harder to solve. For now, but not forever, we will assume that $\widehat{H}$ is not a function of time.

When $\widehat{H}$ doesn't depend on time (i.e., $\partial_t \widehat{H} = 0$), I can reduce this to the time-*independent* Schrödinger equation:

$$\widehat{H}|\psi\rangle = E|\psi\rangle.$$

This is version you're probably more familiar with. Why is this a useful equation to solve? How is this going to help us understand the evolution of the system? The reason is that quantum mechanics is a linear theory, i.e., adding two solutions gives another solution.

This is an important type of equation, which you've probably seen in linear algebra: it says that E is an eigenvalue for $\widehat{H}$, and $|\psi\rangle$ is an eigenvector for $\widehat{H}$. This implies that the equation can be solved (usually) only for vertain values of E. What we're going to do is assume that we can solve the time-independent Schrödinger equation if $E \in \{E_n\}$, where $\{E_n\}$ is some fixed set (the set of eigenvalues). We will assume that there is a *unique* eigenvector/eigenstate $|\psi_n\rangle$ for each eigenvalue E_n. (If this eigenvector was not unique, then the other states would be called degenerate states; in this case, a few things we would like to be true would not be true without more work.)

For any $|\psi\rangle \in \mathcal{H}$ (the Hilbert space), there is a vector $\vec{c} \in \mathbf{C}^N$ (where $N = \dim \mathcal{H}$, which in principle, could be (countable) infinity) such that

$$|\chi\rangle = \sum_{j=1}^{N} c_j |\psi_j\rangle.$$

This is a completeness relation, which says that the $|\psi_j\rangle$ span the Hilbert space. If $E_i \neq E_j$, then the $|\psi_j\rangle$'s are also orthogonal, as we'll see later, so that $\{|\psi_j\rangle\}$ forms a basis of the Hilbert space.

We'll now return to Equation (1). In this case, we have:

$$i\hbar\frac{d}{dt}|\psi_j\rangle = E_j|\psi_j\rangle.$$

Hopefully, this should be a very familiar differential equation: you find that

$$|\psi_j(t)\rangle = e^{-\frac{iE_jt}{\hbar}}|\psi_j(0)\rangle.$$

Since the $|\psi_j\rangle$ form a basis, we have solved the Schrödinger equation in general. Namely, if we write $|\chi(0)\rangle = \sum_{j=1}^{N} p_j|\psi_j(0)\rangle$ for some complex numbers p_j (forget about normalization, etc. for now), we find that

$$|\chi(t)\rangle = \sum_{j=1}^{N} p_j e^{-\frac{iE_jt}{\hbar}}|\psi_j(0)\rangle.$$

What was the point of all of that? This is the standard formula for quantum mechanics: diagonalize the Hamiltonian (i.e., solve the time-independent Schrödinger equation), then use completeness to solve the time-dependent Schrödinger equation.

The infinite square well. Now, we will move on to an example which (hopefully) everyone has seen before. In nonrelativistic quantum mechanics, if I have a system of M particles, the standard form of the Hamiltonian looks like

$$\widehat{H} = \sum_{j=1}^{M} \frac{\widehat{p_j}^2}{2m_j} + \widehat{V}.$$

We're going to let $M = 1$, and work with the special case of the infinite square well of length L. Namely, we set

$$V(x) = \begin{cases} \infty & x_0 \text{ or } x > L \\ 0 & 0 \leq x \leq L \end{cases}$$

Next class, you'll learn about bras and kets, but essentially, you have been working with $\langle x|\psi\rangle = \psi(x)$. Moreover, the momentum operator acts as

$$\widehat{p_j} = -i\hbar\frac{d}{dx_j}.$$

We will multiply the time-independent Schrödinger equation by the bra $|x\rangle$ to get:

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi(x) = \widehat{H}\langle x|\psi\rangle = E\langle x|\psi\rangle.$$

This is perhaps a more familiar version of the Schrödinger equation.

The potential for the infinite square well is sorta funny, because of these infinities. All this is saying, though, is that we want to solve the Schrödinger equation *inside* the well: if $\psi(x)$ was zero outside the well, then the probability that the particle is near x is (assuming normalization of $\psi(x)$) given by $p(x)dx = |\psi(x)|^2dx$. Then, the expectation of the potential energy is

$$\langle V\rangle = \int dx\ p(x)V(x),$$

which would be infinite outside L. That's not allowed. We're therefore going to impose that $\psi(x) = 0$ if $x \notin (0, L)$, and $\psi(0) = \psi(L) = 0$ (the boundary conditions).

We will also demand that the integral

$$Q = \int_0^L |\psi(x)|^2 dx$$

is a positive real number. In other words, we want the wavefunction to not blow up, and to not vanish. By rescaling $\psi(x)$ by $1/\sqrt{Q}$, we get a new function, $\widetilde{\psi}(x)$ — then $|\widetilde{\psi}(x)|^2 dx$ is a probability distribution. This is just saying that $\psi(x)$ is square integrable. Are there any other constraints that we need to impose? The reason this is important is that we are now trying to identify the Hilbert space. For the moment, we will not impose anything else. However, one thing you might think to demand is that the second derivative of $\psi(x)$ actually exists, because if the energy is finite (as it must be), you'll end up with a problem. The reason we won't demand this is that we can obtain better mathematical properties this way, even though it may introduce other non-physical states.

The Hilbert space, therefore, is

$$\mathcal{H}_{\text{isw}} = \{\text{continuous } \psi : [0, L] \to \mathbf{C} | 0 < \int_0^\infty |\psi(x)|^2 dx < \infty,\ \psi(0) = \psi(L) = 0\}.$$

To really make it into a Hilbert space, we need to endow this with an inner product. The inner product is: if we have states described by functions f and g, then

$$\langle f|g\rangle = \int_0^L f^*(x) g(x) dx.$$

We now just have to worry about the Schrödinger equation inside the well, in which case the time-independent Schrödinger equation becomes

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

This has a simple solution, as you have probably seen already:

$$\psi(x) = a\sin(kx) + b\cos(kx),$$

where

$$k = \frac{\sqrt{2mE}}{\hbar}.$$

What do the extra conditions on our wave function translate into? For example, we find that $\psi(0) = b = 0$. What about the other boundary condition? Well:

$$\psi(L) = a\sin(kL) = 0.$$

I *could* take $a = 0$, but the resulting $\psi = 0$ does not lie in the Hilbert space (it is not normalizable). Therefore, $\sin(kL) = 0$. Then, if we set $k_n = n\pi/L$ for all positive integers n — in which case $\sin(kL)$ vanishes. Therefore, we find that the energy is

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 n^2 \pi^2}{2mL^2}.$$

So energy is quantized, hence *quant*um mechanics. Normalizing, we find that

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

Fourier theory proves that the $\{\psi_n\}$ indeed form a complete basis. It's actually amazing how far this simple example goes.

The Hamiltonian being a Hermitian operator has a bunch of important consequences. In the above example, if we restrict to states such that $\psi(x)^* = \psi(x)$, then

$$\langle \widehat{p} \rangle = \int \psi^*(x) \left(-i\hbar \frac{d}{dx} \right) \psi dx = 0,$$

so such a particle can't be moving! You can define a probability current as:

$$J(x,t) = \frac{i\hbar}{2m} \left(\frac{d\psi^*}{dx}\psi - \psi^* \frac{d\psi}{dx} \right).$$

Using the Schrödinger equation, you can prove that

$$\frac{d}{dt}|\psi(x,t)|^2 + \frac{d}{dx}J(x,t) = 0.$$

This is a continuity equation: probability is converved. If I evaluate the first term at x, then I'm just looking at the difference of $J(x,t)$ at the endpoints, which, unless I'm doing something ridiculous, would be zero. Clearly $J(x,t) = 0$ for ψ real, so you conclude that the particle cannot move if ψ is real-valued.

2. More math

We finished last class by defining the notion of an inner product. The first thing I want to talk about is what's known as the dual space. Assume we have a vector space V, called the "ket" space. We will introduce the "bra" space, which is a vector space that is dual (in some sense) to the "ket" space. Namely, for every ket $|\alpha\rangle \in V$, define a bra $\langle\alpha| \in V^*$ (the bra space) such that $\langle\alpha|\beta\rangle = (|\alpha\rangle, |\beta\rangle)$. Namely, consider the space of linear functions $\gamma : V \to \mathbf{C}$. Every such map will be associated with $\langle\gamma|$ such that for any basis $\{|\alpha_j\rangle\}$,

$$\gamma : |\alpha_j\rangle \to c_j \text{ implies that } \langle\gamma|\alpha_j\rangle = c_j.$$

It is easy to check that the set of such $\langle\gamma|$ form a vector space V^*, which we'll call the ket space.

The next concept I want to define is that of an orthonormal basis. This generalizes the notion of orthonormal vectors in $\mathbf{R}^3$. This is a basis $|\psi_i\rangle$ such $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. Since any $|\alpha\rangle$ can be expanded as $|\alpha\rangle = \sum_i c_i|\psi_i\rangle$, we can easily compute the coefficients c_i by $c_i = \langle\psi_i|\alpha\rangle$. We therefore find that

$$|\alpha\rangle = \sum_i |\psi_i\rangle\langle\psi_i|\alpha\rangle.$$

This is known as a completeness relation.

The reason for introducing all of this stuff was that the state of a quantum mechanical system would be described by a ray in a Hilbert space. For the most part, you don't really need to know everything about what exactly a Hilbert space is, as long as you follow your physics nose. We will still go over the definition of a Hilbert space later, though.

Let's now move on to the second postulate: observables are Hermitian operators in $\mathcal{H}$. The word observable really just means what it usually means in English. We therefore need to understand what operators and Hermitian operators on. An operator is something that acts on kets $|\alpha\rangle$ from the left to product other kets: $X \cdot |\alpha\rangle = X|\alpha\rangle$. Two operators A and B are equal if and only if $A|\alpha\rangle = B|\alpha\rangle$ for all $|\alpha\rangle$.

Let us define the addition of operators as follows: $(X+Y)|\alpha\rangle = X|\alpha\rangle + Y|\alpha\rangle$. Then,

$$X + Y = Y + X,\ X + (Y + Z) = (X + Y) + Z.$$

A linear operator is one which satisfies

$$X(c_\alpha|\alpha\rangle + c_\beta|\beta\rangle) = c_\alpha X|\alpha\rangle + c_\beta X|\beta\rangle.$$

Most of the time, we'll encounter linear operators, but later on, we'll need something known as an anti-linear operator. This is an operator for which

$$X(c_\alpha|\alpha\rangle + c_\beta|\beta\rangle) = c_\alpha^\star X|\alpha\rangle + c_\beta^\star X|\beta\rangle.$$

Time reversal is an anti-linear operator, and that's the only thing in the course which'll be anti-linear.

Let's ask about the operation of operators on bras. These act on bras from the *right*: $\langle\beta|X$ is defined by

$$(\langle\beta|X)|\alpha\rangle = \langle\beta|(X|\alpha\rangle).$$

We now want to define a multiplication operation on operators: $(XY)|\alpha\rangle = X(Y(|\alpha\rangle))$. Note, however, that $XY \neq YX$, but $X(YZ) = (XY)Z =: XYZ$. Let's develop a feeling for operators.

An operator can be described by a matrix after choosing a basis. I'm going to start by writing an important equation (probably the most important in the math part of the course). The identity operator 1, which we define $1|\alpha\rangle = |\alpha\rangle$ for all $|\alpha\rangle$, satisfies

$$\boxed{1 = \sum_{a'} |a'\rangle\langle a'|,}$$

where a' is an orthonormal basis. Let's prove this statement. Indeed:

$$1|\alpha\rangle = |\alpha\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha\rangle.$$

Any operator X can be written as

$$X = 1 \cdot X \cdot 1 = \sum_{a',a''} |a'\rangle\langle a'|X|a''\rangle\langle a'| = \sum_{a',a''} (|a'\rangle\langle a''|)\langle a'|X|a''\rangle.$$

The data of $\langle a'|X|a''\rangle$ completely specifies the operator X — so the operator X can be specified by a matrix after picking a basis. The intuition behind operators just comes from thinking about matrices. Suppose we have two matrix representations for X and Y; let us ask about the matrix representation for XY. Indeed:

$$\begin{aligned} XY &= (1 \cdot X \cdot 1)(1 \cdot Y \cdot 1) \\ &= \sum_{a',a''} \sum_{b',b''} |a'\rangle\langle a'|X|a''\rangle\langle a''|b'\rangle\langle b'|Y|b''\rangle\langle b''| \\ &= \sum_{a',a'',b''} |a'\rangle\langle a'|X|a''\rangle\langle a''|Y|b'\rangle|b''\rangle \\ &= \sum_{a',b''} (|a'\rangle\langle b''|) \sum_{a''} \langle a'|X|a''\rangle\langle a''|Y|b'\rangle. \end{aligned}$$

The entries of the matrix given by $\sum_{a''}\langle a'|X|a''\rangle\langle a''|Y|b'\rangle$ precisely correspond to the product of the matrices associated to X and Y.

The adjoint of an operator X, denoted $X^\dagger$ (also called the "Hermitian conjugate"), is defined by $X^\dagger|\alpha\rangle$ is dual to $\langle\alpha|X$. If $X^\dagger = X$, then the operator is "Hermitian", or "self-adjoint". Note that $(XY)^\dagger = Y^\dagger X^\dagger$; indeed, for an arbitrary $|\alpha\rangle$, consider $|\beta\rangle = Y|\alpha\rangle$ and $|\gamma\rangle = X|\beta\rangle$. Then $\langle\beta| = \langle\alpha|Y^\dagger$ and $\langle\gamma| = \langle\beta|X^\dagger$, so that $\langle\gamma| = \langle\alpha|Y^\dagger X^\dagger$, so since $\langle\gamma| = \langle\alpha|(XY)^\dagger$. That implies the desired result.

In a particular basis, we can represent A as a matrix $\langle a'|A|a''\rangle$; then, $\langle a'|A^\dagger|a''\rangle = (\langle a''|A|a'\rangle)^\star$. In other words, the matrix representation of $A^\dagger$ is the complex conjugate of the transpose of the matrix representation of A.

Let us look at some examples.

(1) This is called the "outer product". Given a ket $|\alpha\rangle$ and a bra $|\beta\rangle$, define an operator $O = |\alpha\rangle\langle\beta|$, so that
$$O|\gamma\rangle = |\alpha\rangle\langle\beta|\gamma\rangle.$$
It follows that
$$|\gamma\rangle O^\dagger = \langle\beta|\gamma\rangle^\dagger\langle\alpha| = \langle\gamma|\beta\rangle\langle\alpha|,$$
so that $O^\dagger = |\beta\rangle\langle\alpha|$.

(2) Let $\mathcal{H}$ be the (Hilbert) space of square integrable functions in $[-\infty, \infty]$. Let $A = d/dx$. Define the inner product by
$$\langle f_1|f_2\rangle = \int_{-\infty}^{\infty} f_1^\star(x) f_2(x) dx.$$
What is the adjoint of A? We have
$$\langle f_1|A|f_2\rangle = \int_{-\infty}^{\infty} f_1^\star(x)\frac{d}{dx} f_2(x) dx = \int_{-\infty}^{\infty}\left(-\frac{df_1^\star}{dx}\right) f_2 dx,$$
so that
$$\langle f_1|A|f_2\rangle = -\langle f_2|A|f_1\rangle^*,$$
so that $A^\dagger = -A$. Such an operator is called "anti-Hermitian". What can we do to product a Hermitian operator that is related to the derivative? The answer is to consider the operator $-i\frac{d}{dx}$. Well:
$$\left(-i\frac{d}{dx}\right)^\dagger = i\left(\frac{d}{dx}\right)^\dagger = -i\frac{d}{dx},$$
so that $-i\frac{d}{dx}$ is Hermitian — this is (up to a factor of $\hbar$) how momentum is represented in quantum mechanics. Momentum is something we can observe, so this is compatible with the second postulate of quantum mechanics.

(3) The identity operator 1.

(4) The inverse operator (which doesn't necessarily exist) is such that $A^{-1}A = AA^{-1} = 1$.

(5) Unitary operators are another example: these are operators for which $U^{-1} = U^\dagger$. We'll be using these all the time. In this case:
$$(\langle\beta|U^\dagger)(U|\alpha\rangle) = \langle\beta|U^{-1}U|\alpha\rangle = \langle\beta|\alpha\rangle.$$

(6) The projection operator: $A^2 = A$. For example, $A = |\alpha\rangle\langle\alpha|$ with $\langle\alpha|\alpha\rangle = 1$.

If $A|a'\rangle = a|a'\rangle$, then $|a'\rangle$ is said to be an eigenket of A, and a is an eigenvalue of A. The "spectrum" of A is the set of eigenvalues of A.

Theorem 1. *If $A = A^\dagger$, then all eigenvalues of A are real.*

Proof. Let a', a'' be eigenvalues of A, so that $A|a'\rangle = a'|a'\rangle$ and $A|a''\rangle = a''|a''\rangle$, so that $\langle a''|A = \langle a''|(a'')^\star$ since $A = A^\dagger$. Then:
$$\langle a''|A|a'\rangle = a'\langle a''|a'\rangle$$
and
$$\langle a''|A|a'\rangle = (a'')^\star\langle a''|a'\rangle.$$
It follows that
$$(a' - (a'')^\star)\langle a''|a'\rangle = 0.$$
If $|a'\rangle = |a''\rangle$, since $\langle a'|a''\rangle \neq 0$, we conclude that
$$(a')^\star = (a'')^\star = a'.$$
This means that a' is real. □

Moreover, if $|a'\rangle \neq |a''\rangle$, then $a' \neq a''$, so that $\langle a'|a''\rangle = 0$. In other words, eigenkets corresponding to distinct eigenvalues are orthogonal. Normalize $|a'\rangle$ such that $\langle a'|a'\rangle = 1$; then $\{|a'\rangle\}$ is such that $\langle a'|a''\rangle = \delta_{a',a''}$, i.e., the eigenkets are orthonormal. Since the eigenkets span the space of all states, we get an orthonormal basis. The point is that the eigenstates of a Hermitian operator can be used to provide a basis for the vector space. This is very useful in quantum mechanics. We can write $A = \sum_{a'} a'|a'\rangle\langle a'|$; then

$$A|b'\rangle = \sum_{a'} |a'\rangle\langle a'|b'\rangle = b'|b'\rangle.$$

Let us define a few more operations. The trace of an operator A is

$$\mathrm{Tr}(A) = \sum_i \langle a_i|A|a_i\rangle.$$

If you choose $\{a_i\}$ to be the basis given by the eigenkets of the operator A, then $\mathrm{Tr}(A) = \sum_i a_i$.

If $|a_i\rangle$ and $|b_i\rangle$ are two orthonormal bases, we can define an operator U for which $U|a_i\rangle = |b_i\rangle$. (We choose some order of the basis.) Then, $\langle b_i| = \langle a_i|U^\dagger$, so we can clearly write

$$U = U\sum_i |a_i\rangle\langle a_i| = \sum_i |b_i\rangle\langle a_i|,$$

so that

$$U^\dagger = \sum_i |a_i\rangle\langle b_i|.$$

Then, we see that

$$UU^\dagger = \sum_{ij} |b_i\rangle\langle a_i|a_j\rangle\langle b_j| = \sum_i |b_i\rangle\langle b_i| = 1 = U^\dagger U.$$

Therefore, U is unitary. We've proved that if you have two distinct orthonormal bases for the same Hilbert space, and U is any operator which takes one of these bases to the other, then U is unitary. (The analogue of a unitary operator for $\mathbf{R}^3$ is a rotation.)

Consider a vector $|\alpha\rangle$, and two distinct bases $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$. Then we can write

$$|\alpha\rangle = \sum_i c_i|a_i\rangle = \sum_i d_i|b_i\rangle.$$

Then

$$|\alpha\rangle = \sum_j d_j|b_j\rangle = \sum_j d_j U|a_j\rangle = \sum_{i,j} d_j|a_i\rangle\langle a_i|U|a_j\rangle,$$

so that

$$c_i = \sum_j \langle a_i|U|a_j\rangle d_j = \sum_j U_{ij}d_j,$$

with $U_{ij} = \langle a_i|U|a_j\rangle$; this is the matrix representing the operator U in the $\{|a_i\rangle\}$ basis. Similarly, for any operator X, if we have

$$X = \sum_{ij} |a_i\rangle X_{ij}\langle a_j| = \sum_{k\ell} |b_k\rangle Y_{k\ell}\langle b_\ell|,$$

then

$$X_{ij} = \sum_{k\ell} U_{ik}Y_{k\ell}U^\dagger_{\ell j}.$$

3. More math, and measurement

We begin with a theorem.

Theorem 2. *A Hermitian matrix $H = \langle\phi_i|H|\phi_j\rangle$ can always be diagonalized by a unitary transformation.*

Proof. Let $\{|\phi_i\rangle\}$ be the orthonormal basis in the problem statement, and let $\{|h_i\rangle\}$ be an orthonormal basis of eigenkets of the Hermitian operator H. Then, there is a uniitary operator U such that $|h_i\rangle = U|\phi_i\rangle$. It follows that

$$\langle h_i|H|h_j\rangle = \delta_{ij}h_i = |\phi_i\rangle U^\dagger HU|\phi_j\rangle,$$

which means that $U^\dagger_{ik}H_{k\ell}U_{\ell j}$ is a diagonal matrix (where we sum over repeated indices); this is easy. □

We now turn to simultaneous diagonalization. The claim is that two operators A and B are simultaneously diagonalizable if and only if $[A, B] = 0$. Here is the proof. Let $|\alpha_i\rangle$ be a basis of eigenkets for A, so that $A|\alpha_i\rangle = a_i|\alpha_i\rangle$. If B and A are simultaneously diagonalizable, then $B|\alpha_i\rangle = b_i|\alpha_i\rangle$. Then

$$AB|\alpha_i\rangle = a_ib_i|\alpha_i\rangle = b_ia_i|\alpha_i\rangle = BA|\alpha_i\rangle,$$

which implies that $[A, B] = 0$ since $|\alpha_i\rangle$ form a basis.

For the converse, if $[A, B] = 0$, and $A|\alpha_i\rangle = a_i|\alpha_i\rangle$, then $A(B|\alpha_i\rangle) = a_iB|\alpha_i\rangle$, so $B|\alpha_i\rangle$ is another eigenket of A with the same eigenvalue a_i. In general, this implies that B is block diagonal in the $\{|\alpha_i\rangle\}$ basis, with each block being a subspace of diagonal eigenkets of A. More precisely, suppose our matrix A looks like

$$A = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_1 & & & \\ & & & a_2 & & \\ & & & & \ddots & \\ & & & & & a_2 \end{pmatrix},$$

where there is a $m_1 \times m_1$ block of a_1's on the diagonal, and a $m_2 \times m_2$ block of a_2's on the diagonal. Then, you can diagonalize B in each block without affecting A. It follows that A and B can be simultaneously diagonalizable.

We're now going to move on to the third postulate, which is about measurement. Consider a quantum system in a normalized state $|\psi\rangle \in \mathcal{H}$, and we have an observable which is described by a Hermitian operator A. Recall that the measurement postulate states:

(3) A measurement of A returns one of its eigenvalues.
   (a) If A is measured in a state $|\psi\rangle$, then the probability that you get a is $\langle\psi|M_a|\psi\rangle$, where $M_a = \sum_{j|a_j=a} |a_j\rangle\langle a_j|$ is the "measurement operator". In other words, the probability is $\sum_{j|a_j=a} |\langle a_j|\psi\rangle|^2$.
   (b) After measurement, the system is in a state $|\widetilde{\psi_a}\rangle \propto M_a|\psi\rangle$. (In other words, the proportionality constant ensures that the state is normalized.) This is called the "collapse of the wavefunction".

The operator M_a is projection onto the subspace with eigenvalue a. When you make a measurement, the wavefunction will collapse, and if you measure immediately, then you will get the same answer. This seems reasonable, right? It's somewhat subtle, because you could have destroyed the state in the process of measurement. A good example is if you're detecting a photon. What will usually happen is that the photon gets absorbed into your device, so you

can't remeasure — the photon is gone! One has to refine what the notion of a measurement is: one, where you destroy the thing you're measuring; and two, where you don't destroy the thing you're measuring. Experimentalists have invented a name for these classes, where the former is called a demolition measurement, and the latter is called a nondemolition measuremeant.

Here are a few comments.

(1) The probability

$$\text{Prob}(A = a_i) = \sum_{j|a_j=a_i|} |\langle a_j|\psi\rangle|^2.$$

We therefore need

$$\sum_{a_i} \text{Prob}(A = a_i) = \sum_j |\langle a_j|\psi\rangle|^2 = \sum_j \langle\psi|a_j\rangle\langle a_j|\psi\rangle = \langle\psi|\psi\rangle = 1,$$

since the system is normalized.

(2) For any observable A and state $|\psi\rangle$, the "expectation value" of A is

$$\langle A\rangle = \sum_{a_i} \text{Prob}(A = a_i)a_i = \sum_{a_i} a_j \sum_{j|a_j=a_i} \langle\psi|a_j\rangle\langle a_j|\psi\rangle = \langle\psi|A|\psi\rangle,$$

where we use $A = \sum_i a_i|a_i\rangle\langle a_i|$.

Let's start doing an example, and move towards physics. This is a spin-1/2 system. Everything we've talked about this class has built up stuff in the abstract. Let's see how it all works out in this simple example, which is a good example to keep in mind. These days, such a system is usually called a "qubit". The state space will be spanned by the eigenstates of S^z. We know that there are two states, $|+\rangle$, corresponding to $S^z = \hbar/2$, and $|-\rangle$, corresponding to $S^z = -\hbar/2$. The Hilbert space is $\mathcal{H} = \{|\psi\rangle = c_+|+\rangle + c_-|-\rangle|c_+, c_- \in \mathbf{C}\}$, so it is $\mathbf{C}^2$. Two vectors $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ are physically equivalent, because the state of a system is defined by a ray in the Hilbert space.

This means that only the relative phase of (c_+, c_-) is physical. We also want to demand that the states are normalized, i.e., $\langle\psi|\psi\rangle = 1$, which translates to $|c_+|^2 + |c_-|^2 = 1$. We are free to choose c_+ to be real; then c_- will be a complex numbert. Let us parametrize $c_+ = \cos\theta/2$, so that $c_- = e^{i\phi}\sin\theta/2$, with $0 \le \theta \le \pi$, and $0 \le \phi < 2\pi$. The data we need to specify a spin-1/2 system is therefore precisely θ and ϕ. This is the same data that you need to specify a point on a two-dimensional sphere, i.e., that each state of a spin-1/2 system can be represented as a point on the surface of the unit sphere S^2. This has a name; it's called the "Bloch sphere". It's sometimes useful to think about the state geometrically.

Let's understand the Bloch sphere representation. Let's ask what different point on the sphere correspond to. The north pole is $c_+ = 1$ and $c_- = 0$, so $|\psi\rangle = |+\rangle$, and it has eigenvalue $\hbar/2$; similarly, the south pole is $c_+ = 0$ and $c_- = 1$, so $|\psi\rangle = |-\rangle$, and it has eigenvalue $-\hbar/2$. You'll explore this more in the homework.

Let's now move on to operators. The identity operator is obvious:

$$1 = |+\rangle\langle+| + |-\rangle\langle-|,$$

so that its representation in this basis is $1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. The matrix S^z is therefore

$$S^z = \frac{\hbar}{2}(|+\rangle\langle+| - |-\rangle\langle-|),$$

whose matrix representation is $S^z = \frac{\hbar}{2}\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. You'll need a lot of familiarity with Pauli matrices as we move forward. Here are the Pauli matrices:

$$\sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

so that $S^z = \frac{\hbar}{2}\sigma^z$. Let's define two other operators, S^x and S^y:

$$S^x = \frac{\hbar}{2}(|+\rangle\langle -| + |-\rangle\langle +|),$$

whose matrix representation is $S^x = \frac{\hbar}{2}\sigma^x$, where

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Also:

$$S^y = \frac{\hbar}{2}(-i)(|+\rangle\langle -| - |-\rangle\langle +|),$$

so that its matrix representation is $S^y = \frac{\hbar}{2}\sigma^y$, with

$$\sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

You can, and should, check explicitly that

(1) $[S^a, S^b] = i\hbar\epsilon^{abc}S^c$, where ϵ^{abc} is the fully antisymmetric tensor in a, b, c, with $\epsilon^{123} = 1$.
(2) $\{S^a, S^b\} = \frac{\hbar^2}{2}\delta^{ab}$, where $\{A, B\} = AB + BA$ is the anticommutator. A corollary of this is that
$$S^2 = S^a S^a = (S^x)^2 + (S^y)^2 + (S^z)^2 = \frac{3}{4}\hbar^2 1.$$
(3) $[S^2, S^a] = 0$.

The simplest way to check all of these things is to just do the simple computation with the 2×2-matrices; later, we'll find deep reasons for why these results are true.

We've described states and operators, so let us now move on to measurements. Take a state $|\psi\rangle = c_+|+\rangle + c_-|-\rangle$, which is normalized. Then

$$\text{Prob}(S^z = \hbar/2) = |\langle +|\psi\rangle|^2 = |c_+|^2 = \cos^2\theta/2,$$

and

$$\text{Prob}(S^z = -\hbar/2) = |\langle -|\psi\rangle|^2 = |c_-|^2.$$

Suppose instead of S^z, I measure spin in some direction. I can look at the associated operator, diagonalize, and look at the eigenstates, from which I can deduce the probability of observing each eigenstate. Namely, you can get these from $\text{Prob}(S^x = \pm\hbar/2)$, and in general $\text{Prob}(\vec{S}\cdot\hat{n} = \pm\hbar/2)$, where $\hat{n}$ is any unit vector. Implicit in these is the claim that the eigenvalues of these operators are $\pm\hbar/2$. For example, the eigenkets of S^x are $\frac{1}{\sqrt{2}}(|+\rangle \pm |-\rangle)$, as you should check. The sign in the superposition corresponds to the sign of the eigenvalue. Therefore

$$\text{Prob}(S^x = \hbar/2) = \left|\frac{1}{\sqrt{2}}(\langle +| + \langle -|)(c_+|+\rangle + c_-|-\rangle)\right|^2 = \frac{1}{2}|c_+ + c_-|^2.$$

This stuff is very concrete. What I'll urge you to do is to go back to lecture 1, and make sure that you understand where all that (Stern-Gerloch, etc.) comes from (since you now understand the quantum mechanics of a spin-1/2 system). I asserted in the first lecture that it's not possible to simultaneously measure S^z and S^x. Let me talk about what observables can be simultaneously measured.

I'll return now to the general development of the theory, but keep the spin-1/2 system in mind. The general question we want to ask is when two things can be measured simultaneously precisely. (Position and momentum are non-examples, as are S^z and S^x.) This translates into a question about the operators. Let A and B be Hermitian operators, and measure A so that the wavefunction collapses to an eigenket $|a\rangle$ of A. Next, we will measure B. If $|a\rangle$ is also an eigenket of B, then the measurement of B will give a sharp value. Therefore, A and B can be

simultaneously measured if all eigenkets of A are also eigenkets of B. (They are "compatible".) This implies that $[A, B] = 0$. Conversely, if $[A, B] = 0$, then we proved earlier that A and B are simultaneously diagonalizable, so their eigenkets can be chosen to be the same. This is a very crisp condition for compatibility of observables — you can measure one without disturbing the other.

In the spin example, we see that S^2 commute with S^a, so you can simultaneously measure S^2 and S^a. However, you cannot do this with S^x and S^y. Once we recognize that some sets of operators can be measured simultaneously, it's natural to ask about the maximal set of such operators. Let us define the concept of a complete set of commuting observables. A "complete set of commuting observables" is a set of observables $\{A, B, C, \cdots\}$ such that $[A, B] = [A, C] = [B, C] = \cdots = 0$, i.e., all operators commute with each other. In order to make it a complete set, we need the further condition that for any $A, B, C, \cdots$, there is at most one solution to the eigenvalue equations $A|\alpha\rangle = a|\alpha\rangle$, $B|\alpha\rangle = b|\alpha\rangle$, $C|\alpha\rangle = c|\alpha\rangle$. This condition asserts that (think about this yourself) that this is a maximal set: there is no other operator you can invent and add to this list which commutes with everything else.

4. Generalized uncertainty principle, and position and momentum operators

4.1. Generalized uncertainty principle. Take two observables A and B, and let $|\psi\rangle$ be a general state. We can ask about the probability distribution of the observable in the state. We define the variance

$$\Delta A^2 = \langle (A - \langle A\rangle)^2\rangle = \langle A^2 - 2A\langle A\rangle + \langle A\rangle^2\rangle = \langle A^2\rangle - \langle A\rangle^2 = \langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2.$$

If $|\psi\rangle$ is an eigenket of A with eigenvalue a, then the variance is $\Delta A^2 = a^2 - a^2 = 0$, which is to be expected (the probability distribution is not interesting). In general, though, this will not be the case: then, there will be a nontrivial variance.

For example, consider a spin 1/2 system in the state $\langle\psi|$. We expect that $\Delta S_z^2 = \langle S_z^2\rangle - \langle S_z\rangle^2 = 0$. However, we can also ask about the variance of other spin operators. For instance, $\Delta S_x^2 = \langle S_x^2\rangle - \langle S_x\rangle^2$. The S_x operator has the property that $S_x^2 = \hbar^2/4 \cdot \mathrm{id}$, so this becomes $\langle +|S_x^2|+\rangle - \langle +|S_x|+\rangle^2$. Since $S_x = \frac{\hbar}{2}\sigma^x$, and $\sigma^x_{++} = 0$, we conclude that $\Delta S_x^2 = \hbar^2/4$.

Define $\Delta A = \sqrt{\Delta A^2}$. Then, the uncertainty relation states:

$$\boxed{\Delta A \Delta B \geq \frac{1}{2}|\langle [A, B]\rangle|.}$$

Note that if A and B commute, then this bound says $\Delta A \Delta B \geq 0$, which makes sense — if they commute, then they are diagonalizable, so you can measure them simultaneously.

Let us prove this result. We will use Schwarz's inequality (on your homework), which says that

$$|\langle a|b\rangle|^2 \leq \langle a|a\rangle\langle b|b\rangle.$$

For a general state $|\psi\rangle$, define the following operators:

$$\delta A = A - \langle A\rangle;\ \delta B = B - \langle B\rangle.$$

These are operators, which depend on the chosen state $|\psi\rangle$. Define new states $|a\rangle = \delta A|\psi\rangle$ and $|b\rangle = \delta B|\psi\rangle$. By the Schwarz inequality, we have

$$|\langle\psi|\delta A\delta B|\psi\rangle|^2 \leq \langle\psi|(\delta A)^2|\psi\rangle\langle\psi|(\delta B)^2|\psi\rangle.$$

We can write

$$\delta A\delta B = \frac{\delta A\delta B + \delta B\delta A}{2} + \frac{\delta A\delta B - \delta B\delta A}{2} = \frac{1}{2}(\{\delta A, \delta B\} + [\delta A, \delta B]),$$

where the first term is Hermitian, and the second is anti-Hermitian. What do we know about the expectation value of a Hermitian operator? It is always real (since the eigenvalues of any

Hermitian operator are all real). Likewise, the expectation value of an anti-Hermitian operator is purely imaginary (since the eigenvalues of an anti-Hermitian operator are all purely imaginary). It follows that the expectation value of $\{\delta A, \delta B\}$ is purely real, and the expectation value of $[\delta A, \delta B]$ is purely imaginary. We conclude that

$$|\langle \delta A \delta B \rangle|^2 = \frac{1}{4}(|\langle [\delta A, \delta B] \rangle|^2 + |\langle \{\delta A, \delta B\} \rangle|^2) \geq \frac{1}{4}|\langle [\delta A, \delta B] \rangle|^2.$$

However,

$$[\delta A, \delta B] = [A - \langle A \rangle, B - \langle B \rangle] = [A, B],$$

which implies that

$$\Delta A^2 \Delta B^2 \geq |\langle \delta A \delta B \rangle|^2 \geq \frac{1}{4}|\langle [A, B] \rangle|^2,$$

as desired.

4.2. Position and momentum. Position and momentum are observables with *continuous* eigenvalues, unlike spin (or most of the operators we've previously considered). In order to discuss operators with continuous eigenvalues, we need to assume that the Hilbert space of states is necessarily infinite-dimensional. I might say a little bit about what a Hilbert space is, although that might've already been talked about in your recitation.

Consider any Hermitian operator ξ with a continuous spectrum, and let $|\xi'\rangle$ and ξ' denote its eigenkets and eigenvalues. What does the inner product of two vectors mean? Namely, how do we think about $\langle \xi' | \xi'' \rangle$? For a discrete spectrum, we know that $\langle a | a' \rangle = \delta_{aa'}$. For a continuous spectrum, we just replace the Kronecker delta by a Dirac delta function:

$$\langle \xi' | \xi'' \rangle = \delta(\xi' - \xi'').$$

We have to what exactly the Dirac delta function means.

I'm assuming you've all seen the Dirac delta function before, but let me recall it anyway. It satisfies the property that

$$\int_{-\infty}^{\infty} \delta(x) dx = 1, \ \delta(x) = 0 \ \forall \ x \neq 0, \ \int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0)$$

for any function $f(x)$. You can think about it as the limit of a series of functions (like a Gaussian distribution).

We can replace the completeness relation $1 = \sum_a \langle a || a \rangle$ with

$$1 = \int d\xi' \langle \xi' || \xi' \rangle.$$

Then, we can write an arbitrary state as

$$|\psi\rangle = \int d\xi' |\xi'\rangle \langle \xi' | \psi \rangle.$$

Armed with all this, we can discuss the inner product of two states:

$$\langle \psi' | \psi \rangle = \int d\xi' \langle \psi' | \xi' \rangle \langle \xi' | \psi \rangle.$$

Note that

$$\langle \xi' | \xi | \xi'' \rangle = \xi' \delta(\xi' - \xi'') = \xi'' \delta(\xi'' - \xi').$$

Let us now specialize to position. Let's start with a particle moving in one spatial dimension, and ask about its position. Define the "position operator x and corresponding eigenstate $|x'\rangle$"

by $x|x'\rangle = x'|x'\rangle$. We postulate that the $|x'\rangle$ form a complete set of states for the Hilbert space of this particle, i.e., that a general state can be written as

$$|\psi\rangle = \int_{-\infty}^{\infty} dx' |x'\rangle\langle x'|\psi\rangle.$$

The object $\langle x'|\psi\rangle$ is a complex number, and that is what is familiar to you as $\psi(x')$.

What about measurement of position? Assume that we are measuring positoin with a detector which registers a click everytime a particle hits the detector. Every detector in real life will have a finite resolution — you'll never have a detector with zero resolution. Let us therefore assume that we have a detector which clicks everytime a particle is between $x' + \Delta/2$ and $x' - \Delta/2$, where Δ is the resolution of your instrument.

When talking about measurement, we have to generalize the collapse postulate: after measurement, the state of the particle is such that the position is definitely between $x' - \Delta/2$ and $x' + \Delta/2$. If Δ is small, then we expect that $\langle x''|\psi\rangle$ will be approximately constant within $x' - \Delta/2 < x'' < x' + \Delta/2$, in which case the probability of particle detection is

$$\text{Prob(particle detection)} = |\langle x'|\psi\rangle|^2 \Delta.$$

To emphasize that Δ is small, let us write it as dx', so this becomes

$$\text{Prob(particle detection)} = |\langle x'|\psi\rangle|^2 dx'.$$

This defines a probability density. We therefore need that

$$\text{Prob(particle is somewhere)} = \int_{-\infty}^{\infty} dx' |\langle x'|\psi\rangle|^2,$$

which is 1 if $|\psi\rangle$ is normalized to be 1. We therefore define $\langle x'|\psi\rangle = \psi(x')$, which is the wavefunction that's familiar to you.

We've only talked about particles in one dimension so far, so let us move on to particles in more dimensions. The generalization is almost trivial, but requires an assumption. In d dimensions, there will be a position vector $\vec{x} = (x_1, \cdots, x_d)$ is the position operator. The assumption that enabled us to proceed was that all these different components are compatible observables (i.e., that they commute with each other). At some level, that's just common sense — but it's an assumption that we must make. In other words, we are assuming that $[x_i, x_j] = 0$. Then, we can define position eigenstates via the equation $\vec{x}|\vec{x'}\rangle = \vec{x'}|\vec{x'}\rangle$, and proceed exactly as before.

Having defined position, let us move on to momentum. All of you know from previous QM classes that momentum can be defined by

$$p = -i\hbar \frac{\partial}{\partial x}.$$

We'll discuss some examples where the definition of momentum via this operator does not correspond to the classical definition as mass times velocity. Let's, for now, define this to be momentum, and run along with it.

In the space of square integrable functions $\psi(x)$, we have

$$p\psi(x) = -i\hbar \frac{d\psi}{dx},$$

and

$$[x, p]\psi(x) = -i\hbar \left(x\frac{d\psi}{dx} - \frac{d}{dx}(x\psi(x)) \right) = i\hbar\psi(x),$$

which can write as an operator equation

$$[x, p] = i\hbar.$$

This is your familiar commutation relation. Using the uncertainty relation, we conclude that

$$\Delta x \Delta p \geq \frac{\hbar}{2},$$

which is the Heisenberg uncertainty relation.

Momentum is a Hermitian operator, so it has eigenstates with real eigenvalues, which should be as legitimate of a basis for the Hilbert space as anything else. Let's understand this momentum basis. Let $|p'\rangle$ denote the eigenstates of p, so that $p|p'\rangle = p'|p'\rangle$. Then $|p'\rangle$ form a basis, so

$$\langle p''|p'\rangle = \delta(p' - p'').$$

We can write the identity operator as

$$1 = \int_{-\infty}^{\infty} dp' |p'\rangle\langle p'|.$$

Once we have this, as usual, we can expand any arbitrary state in this basis:

$$|\psi\rangle = \int_{-\infty}^{\infty} dp' |p'\rangle\langle p'|\psi\rangle,$$

and

$$|\langle p'|\psi\rangle|^2 dp' = \text{Prob(measurement of momentum gives value between } p' \text{ and } p' + dp').$$

Call $\langle p'|\psi\rangle$ the momentum space wavefunction. This is as complete of a description as the position space wavefunction.

It's useful to know how to go back and forth between the position and momentum basis. For this, we need to know how the basis vectors themselves are related to each other. In a usual vector space, we saw that there was a unitary transformation taking one basis to the other, and the same will be true here as well. We need to understand $\langle x'|p'\rangle$. Let us use

$$p'\langle x'|p'\rangle = \langle x'|p|p'\rangle = -i\hbar \frac{\partial}{\partial x}\langle x'|p'\rangle,$$

so that

$$\langle x'|p'\rangle = N e^{ip'x'/\hbar},$$

where N is some constant. We can find N by using the equation $\langle x'|x''\rangle = \delta(x' - x'')$. This implies that

$$\int_{-\infty}^{\infty} dp' \langle x'|p'\rangle\langle p'|x''\rangle = \delta(x' - x'').$$

The left hand side is

$$|N|^2 \int_{-\infty}^{\infty} dp' e^{ip'(x'-x'')/\hbar} = 2\pi\hbar |N|^2 \delta(x' - x''),$$

since the Dirac delta function is the Fourier transform of the identity. It follows that

$$N = \frac{1}{\sqrt{2\pi\hbar}},$$

with the standard choice that N is real. It follows that

$$\boxed{\langle x'|p'\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ip'x'/\hbar}}.$$

This is the familiar fact that the momentum space wavefunctions are plane waves. Once we know this inner product, we can build up a transformation of a general state $|\psi\rangle$ from the position space to the momentum space.

5. More on momentum, and quantum dynamics

For any arbitrary state $|\psi\rangle$, the position state wavefunction is $\langle x'|\psi\rangle = \psi(x')$; this is the familiar wavefunction in quantum mechanics. You should think of it as the ket $|\psi\rangle$ expressed in the position basis. We also defined the momentum space wavefunction $\langle p'|\psi\rangle =: \phi(p')$. How are these two related? It's easy to work this out:

$$\langle x'|\psi\rangle = \int dp' \langle x'|p'\rangle\langle p'|\psi\rangle = \int \frac{dp'}{\sqrt{2\pi\hbar}} e^{ip'x'/\hbar}\phi(p');$$

this is a Fourier transform. Similarly,

$$\langle p'|\psi\rangle = \phi(p') = \int \frac{dx'}{\sqrt{2\pi\hbar}} e^{-ip'x'/\hbar}\psi(x').$$

One thing I won't do is emphasize how to actually solve the Schrodinger equation for various kinds of potentials; I presume you've already seen this — it's just solving differential equations. In particular, I won't do things like solve the Schrodinger equation for the atom. It's important to remember that you can solve it in any basis that you want. A famous example of a system where the position space is very complex, but the momentum basis is simple, is when you want to compute the wavefunction for a particle moving in one dimension in an electric field. Then, the Hamiltonian is

$$H = \frac{p^2}{2m} - e\epsilon x = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} - e\epsilon x.$$

In the p-basis, H goes to $\frac{p^2}{2m} - e\epsilon i\hbar\frac{d}{dp}$, which can be solved easily — much easier than the same equation in the position basis.

Another comment: we built up the Hilbert space, by solving for the position (resp. momentum) eigenstates, and then expanding the wavefunction in terms of this basis. Neither the position or momentum eigenstates live in the Hilbert space: they are not normalizable. But they're useful in building up the set of states that you want. In practice, this is profoundly inconsequential. You get around this by modifying the problem slightly such that everything is normalizable.

The most familiar example is the following. Put the system in a large box of size L, and demand periodic boundary conditions. Then $e^{ip'(x'+L)/\hbar} = e^{ip'x'/\hbar}$. This is satisfied as long as $p'L = 2\pi n\hbar$, for n an integer. The momentum spacing is of order $1/L$. At the end of the day, we'll take $L \to \infty$. For finite L, we require that

$$\int_0^L dx' |\langle x'|p'\rangle|^2 = 1,$$

i.e., it must have one of the momentum eigenstates. It follows that

$$\langle x'|p'\rangle = \frac{1}{\sqrt{L}} e^{ip'x'/\hbar}.$$

Let me make another remark. From the uncertainty principle, we know that the width of the wavefunction $\psi(x')$ (i.e., uncertainty in position) gets related to the width of the wavefunction $\phi(p')$ (i.e., uncertainty in momentum). What are the wavefunctions for which the bound given by the uncertainty principle is saturated? I.e., what are the minimum uncertainty states?

We can pose this question a bit more generally. Since the general uncertainty principle states that for operators A and B,

$$\Delta A \Delta B \geq \frac{1}{2}|\langle [A, B]\rangle|,$$

we can ask when this inequality is saturated. This can be proven to be saturated when

$$(A - \langle B\rangle)|\psi\rangle = \lambda(B - \langle B\rangle)|\psi\rangle,$$

with λ purely imaginary. You will prove this in homework. Let us apply this result to x and p, in a situation where $\langle x\rangle = \langle p\rangle = 0$. Then, the minimum uncertainty states $|\psi\rangle$ satisfy

$$(x - \lambda p)|\psi\rangle = 0.$$

This gives us a first order differential equation. Since λ is purely imaginary, let us write $\lambda = -iw/\hbar$; then, this implies:

$$\left(x' + w\frac{d}{dx'}\right)\psi(x') = \langle x'|x + \frac{iwp}{\hbar}|\psi\rangle = 0.$$

It follows that $\psi(x') = Ae^{-x'^2/2w}$, which is precisely the Gaussian.

Let us now talk about momentum and translation. It is a general result that given any Hermitian operator A, the operator e^{iA} is unitary. From a practical point of view, you can define e^{iA} through the series expansion. Setting $A = p$, we can ask about its exponential. Define an operator $T(a) = e^{-iap/\hbar}$, where $a \in \mathbf{R}$. If you wish, we can write

$$T(a) = 1 - \frac{iap}{\hbar} + \frac{1}{2}\frac{(-iap)^2}{\hbar} + \cdots$$

It follows that

$$T(a)^\dagger = e^{-(-i)ap^\dagger/\hbar} = e^{iap/\hbar},$$

so that

$$T(a)^\dagger T(a) = T(a)T(a)^\dagger = 1,$$

which is just the statement that $T(a)$ is unitary. This operator has some interesting properties.

(1) $T(a)^\dagger = T(-a)$;
(2) $T(a')T(a'') = T(a' + a'')$;
(3) $T^{-1}(a)xT(a) = x + a$. Let me prove this. There's a brute force way of proving this. We will instead prove this by a useful trick. Define

$$F(a) = T^{-1}(a)xT(a) = e^{iap/\hbar}xe^{-iap/\hbar}.$$

Then

$$\begin{aligned}\frac{dF}{da} &= \frac{i}{\hbar}\left(e^{ipa\hbar}pxe^{-ipa/\hbar} + e^{ipa/\hbar}(-xp)e^{-ipa/\hbar}\right)\\ &= \frac{i}{\hbar}e^{ipa/\hbar}(px - xp)e^{-ipa/\hbar}\\ &= \frac{i}{\hbar}e^{ipa/\hbar}(-i\hbar)e^{-ipa/\hbar} = 1.\end{aligned}$$

It follows that $F(a) = F(0) + a = x + a$, as desired.

We also find that

$$T^{-1}(a)xT(a)|x'\rangle = (x + a)|x'\rangle,$$

so that

$$x(T(a)|x'\rangle) = (x + a)(T(a)|x'\rangle) = (x' + a)(T(a)|x'\rangle).$$

We may therefore identify $T(a)|x'\rangle = |x' + a\rangle$. We can also think about $\langle x'|T^{-1}(a)|\psi\rangle = \langle x' + a|\psi\rangle$, and conclude that $T^{-1}(a)|\psi\rangle$ has wavefunction $\psi(x' + a)$ if $|\psi\rangle$ has wavefunction $\psi(x')$. The operator $T(a)$ is called the translation operator. One usually says that "translation is generated by the Hermitian operator p".

There is a conceptual point I want to emphasize. I motivated the discussion of momentum by assuming you know what it is — but I needn't have done it that way. I could have started with the statement that translation is a physical operation, so it must be implemented as a unitary operator. I can do an infinitesimal translation, so let me define translation by a continuous

parameter (a in our case), which allows me to define p. I could have run the whole story backwards, and would've found that p could be represented in the position basis by $-i\hbar d/dx$.

I want to now move on to quantum dynamics. We've yet to discuss the final postulate of quantum mechanics. Recall:

(4) Time evolution is given by the map $|\psi(t)\rangle \mapsto |\psi(t')\rangle = U(t',t)|\psi(t)\rangle$, where $U(t',t)$ (known as the *time evolution operator*) is a unitary operator, i.e., $U^\dagger U = 1$. If $t' = t+\epsilon$, then you find the Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle,$$

where H is the Hermitian operator known as the Hamiltonian.

In general, $|\psi(t')\rangle = U(t',t)|\psi(t)\rangle$ for *some* $U(t',t)$. What can we demand of $U(t',t)$?

(1) Probability conservation; this implies unitarity. Indeed, expand

$$|\psi(t)\rangle = \sum_n c_n(t)|a_n\rangle$$

where $\{|a_n\rangle\}$ is a basis of eigenkets for some observable A. Then, let us write

$$\text{Prob}(A = a_n \text{ at time } t) = p(a_n;t) = |c_n(t)|^2.$$

We must have

$$\sum_n p(a_n,t) = \sum_n |c_n(t)|^2 = 1,$$

i.e., that $\langle\psi(t)|\psi(t)|=\rangle 1$. This must be satisfied for all t. In particular,

$$\langle\psi(t')|\psi(t')\rangle = \langle\psi(t)|U^\dagger(t',t)U(t',t)|\psi(t)\rangle = 1,$$

for all $|\psi(t)\rangle$ with unit norm. We claim that this implies that $U^\dagger(t',t)U(t',t) = 1$ (i.e., U is unitary) if and only if $\langle\psi(t)|U^\dagger(t',t)U(t',t)|\psi(t)\rangle = 1$. Let us write $X(t',t) = U^\dagger(t',t)U(t',t)$; this is a Hermitian operator. We have $\langle\psi|X|\psi\rangle$ for all $|\psi\rangle$. Choose $|\psi\rangle$ to be an eigenket $|n\rangle$ of X with eigenvalue λ_n, so that $\langle\psi|X|\psi\rangle = \lambda_n = 1$, so any eigenvalue of X is just 1. It follows that X is the identity.

It is central to the probabilistic interpretation of quantum mechanics that time evolution is implemented by a unitary operator.

(2) We can demand composition: $U(t_f,t_i) = U(t_f,t)U(t,t_i)$ for all $t_i \le t \le t_F$.

(3) $U(t,t) = 1$.

The combination of these three statements is extraordinarily powerful, as we'll see in the next few lectures.

6. Hamiltonians, the Schrödinger and Heisenberg pictures

We will build up time evolution as as sequence of infinitesimal ones. Consider the operator $U(t+dt,t)$; let us expand it as a power series (the Taylor expansion). Then:

$$U(t+dt,t) = U(t,t) - \left(\frac{i}{\hbar}H\right)dt + O(dt^2),$$

where the operator H is *defined* such that $iH/\hbar$ is the coefficient of dt. Note that $U(t,t) = 1$. Let us now impose the condition that U is unitary; then, we have

$$U^\dagger(t+dt,t) = U^\dagger(t,t) + \left(\frac{i}{\hbar}H^\dagger\right)dt + O(dt^2),$$

so that

$$1 = U^\dagger U = \left(1 + \left(\frac{i}{\hbar}H^\dagger\right)dt + O(dt^2)\right)\left(1 - \left(\frac{i}{\hbar}H\right)dt + O(dt^2)\right).$$

In other words,

$$1 + \frac{i}{\hbar}(H^\dagger - H)dt + O(dt^2) = 1,$$

so $H^\hbar = H$, i.e., H is hermitian. If you work out what happens at higher orders, you'll get some commutation relations between H and higher order operators, but that doesn't give much extra data.

The reason I called this operator H is because it is the Hamiltonian. We therefore learn that

$$|\psi(t+dt)\rangle = \left(1 - \frac{i}{\hbar}H\ dt\right)|\psi(t)\rangle$$

for $dt \to 0$. In other words, we have

$$|\psi(t+dt)\rangle - |\psi(t)\rangle = -\frac{i}{\hbar}H\ dt|\psi(t)\rangle,$$

which gives

$$\boxed{i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle}.$$

This is the Schrödinger equation, in its general, time-evolutionary form. Note that the hermiticity of the Hamiltonian was inferred from the unitarity of time-evolution, which in turn was imposed in order for probability conservation to hold.

Of course, to know anything about the dynamics, we need to precisely what the Hamiltonian is: this'll depend on the system. For those of you who don't know what the Hamiltonian is, here's a brief review. Let's talk about time evolution in classical mechanics. The standard description of dynamics which everyone learns in high school is Newton's law, but you can also use Lagrangians (essentially equivalent to the former) and Hamiltonians. The idea behind Hamiltonians is as follows. Classically, a particle's state can be specified by $q(t)$, the position, and $p(t)$, the momentum. Given $(q(t_0), p(t_0))$, how do we determine $(q(t), p(t))$ for $t > t_0$? This is the problem of evolution. Hamilton wrote down the following equation of motion (the "Hamilton equations of motion"):

$$\frac{dq}{dt} = \frac{\partial H}{\partial p},\ \frac{dp}{dt} = -\frac{\partial H}{\partial q}.$$

Here, $H = H(q,p)$ is the classical Hamiltonian. For instance, a single particle moving in one dimension has

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

Then, the equations become

$$\frac{dq}{dt} = \frac{p}{m},\ \frac{dp}{dt} = -\frac{\partial V}{\partial q};$$

the former just says that $p = mv$, while the latter is Newton's law (the force is the negative of the gradient of the potential).

The space of (q,p) is known as phase space. An important property of classical time evolution is that it preserves the volume of a given region of phase space. This is known as Liouville's theorem, and is an important result in statistical mechanics.

It's useful to formulate classical dynamics in terms of Poisson brackets. You could find the evolution of a function of p, q by first solving for p, q, but you could also try to do it directly. This is why Poisson brackets are useful. Recall:

$$\{f,g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q}.$$

Then, any observable $A(q,p)$ has time evolution determined by

$$\frac{dA}{dt} = \{A, H\}.$$

For instance, let's take $A = q$; then this states that

$$\frac{dq}{dt} = \{q, H\} = \frac{\partial H}{\partial p}.$$

This, of course, is one of Hamilton's equations. You can do the same thing with $A = p$.

Let's go back to the development of dynamics in quantum mechanics. I want to explore a different way to think about time evolution in quantum mechanics (which is equivalent to the Schrödinger equation), which is a completely different point of view. The things we've built up so far are collectively called the Schrödinger picture, and what we're about to develop is called the Heisenberg picture. Historically, Schrödinger and Heisenberg developed their stories independently. Eventually, people realized they're equivalent.

Start with the Schrödinger equation

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle.$$

Let us assume that H is independent of time (which is generally true for a closed system). Write $|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle$; this is the standard solution you would've written down. Let me emphasize again that we are assuming t-independence of H — if this assumption is not satisfied, we *cannot* write the solution in this way; this'll be discussed later. Therefore,

$$U(t,0) = e^{-iHt/\hbar}.$$

The physically significant quantities are matrix elements of operators, i.e.,

$$\langle \psi'(t)|A|\psi(t)\rangle = \langle \psi'(0)|U^\dagger(t,0)AU(t,0)|\psi(0)\rangle.$$

We're implicitly assuming above that A does not explicitly depend on time. The *Schrödinger picture* postulates that the t-dependence of the matrix elements comes from that of the state vectors $|\psi\rangle$ and $|\psi'\rangle$. However, in light of the above equation, we could also have asserted that (and this is the *Heisenberg picture*) the states $|\psi\rangle$ and $|\psi'\rangle$ did not change, but rather the *operator* A changed, according to

$$A_H(t) = U^\dagger(t,0)AU(t,0).$$

We can now move to the infinitesimal world, and ask how the operator A changes under infinitesimal time evolution. We find:

$$\begin{aligned}\frac{dA_H}{dt} &= \frac{i}{\hbar}e^{iHt/\hbar}(HA_H(0) - A_H(0)H)e^{-iHt/\hbar} \\ &= \frac{i}{\hbar}e^{iHt/\hbar}[H, A_H(0)]e^{-iHt/\hbar},\end{aligned}$$

which implies that

$$\boxed{\frac{dA_H}{dt} = \frac{1}{i\hbar}[A_H(t), H].}$$

This is the Heisenberg equation of motion. It's the analogue of the Poisson bracket picture.

There's a similarity between the spatial and time translations. We defined the spatial translation operator to be $T(a) = e^{-iap/\hbar}$, and the time translation operator to be $e^{-iHt/\hbar}$; therefore, momentum describes *spatial* evolution, and energy describes *time* evolution. ("Time translation is generated by the Hermitian operator H.") Soon, we'll talk about further alternate pictures of time evolution in quantum mechanics: the Dirac picture (a hybrid of Schrödinger and Heisenberg), and the Feynman picture.

We will now turn to energy eigenstates. The Hamiltonian H is a Hermitian operator, so it corresponds to some physical observable. Let us call this observable "energy". We can diagonalize H by a unitary transformation. Let $|j\rangle$ be its eigenstates; these form an orthonormal basis. These states are what we'll call *energy eigenstates*. The satisfy the eigenvalue equation $H|j\rangle = E_j|j\rangle$, and the E_j are called the *energy eigenvalues*. An arbitrary state $|\psi\rangle$ can be written as $\sum_j c_j|j\rangle$, and $c_j = \langle j|\psi\rangle$. An arbitrary operator can be written as

$$A = \sum_{j,j'} A_{jj'}|j\rangle\langle j'|.$$

Let's work out time evolution in terms of these states. In the Schrödinger picture, $A_{jj'}$ is t-independent, but

$$|\psi(0)\rangle \to |\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle = \sum_j c_j e^{-iHt/\hbar}|j\rangle = \sum_j c_j e^{-iE_j t/\hbar}|j\rangle.$$

Above, we used the fact that if $|j\rangle$ is an eigenstate of H, then it is also an eigenstate of $e^{-iHt/\hbar}$. Define

$$c_j(t) = c_j e^{-iE_j t/\hbar},$$

so that $\sum_j c_j(t)|j\rangle = |\psi(t)\rangle$. The matrix elements (take the one below, for example)

$$\langle\psi(t)|A|\psi(t)\rangle = \sum_{jj'} c_j^\star(t)c_{j'}(t)A_{jj'} = \sum_{jj'} c_j^\star(0)c_{j'}(0)e^{i(E_j-E_{j'})t/\hbar}A_{jj'}.$$

This is the importance of the energy eigenstates: quantum dynamics is easy in the energy eigenbasis. That's why we spend a lot of time trying to work out the energy eigenstates in introductory quantum mechanics classes.

Let's revisit this story in the Heisenberg picture. The states do not evolve, i.e., the c_j are fixed in time, but the operator A transforms as

$$A \to A(t) = e^{iHt/\hbar}Ae^{-iHt/\hbar} = e^{iHt/\hbar}\left(\sum_{jj'} A_{jj'}|j\rangle\langle j'|\right)e^{-iHt/\hbar} = \sum_{jj'} A_{jj'}e^{i(E_j-E_{j'})t/\hbar}|j\rangle\langle j'|.$$

Therefore, the expectation value is

$$\langle\psi|A(t)|\psi\rangle = \sum_{jj'} A_{jj'}e^{i(E_j-E_{j'})t/\hbar}c_j^\star(0)c_{j'}(0),$$

which agrees with the expectation value coming from the Schrödinger picture.

We will now turn to an example (which we'll return to later). We used this as an example a bit earlier, too. Let's talk about spin precession in a magnetic field. We'll consider the spin-1/2 moment in a B-field with Hamiltonian

$$H = -\frac{ge}{2m}\vec{S}\cdot\vec{B},$$

with $e < 0$ for electrons. The number g is called the "gyromagnetic ratio". Let's assume that $\vec{B} = B\hat{z}$, so that

$$H = -\frac{ge}{2m}S_z B.$$

This is an extraordinary simple system: the Hilbert space is two-dimensional, and the Hamiltonian is particularly simple (it's proportional to S_z, so S_z eigenstates are energy eigenstates), so that

$$H|S_z = +\hbar/2\rangle = \mp\frac{ge}{4m}\hbar B|S_z = \pm\hbar/2\rangle,$$

so the energy eigenvalues are

$$E_\pm = \mp\frac{ge\hbar}{4m}B.$$

Which state has higher or lower energy depends on whether e is positive or negative. Let's assume that $e > 0$. Then the state $|-\rangle = |S_z = -\hbar/2\rangle$ has higher energy than $|+\rangle$, with the energy separation given by $\frac{ge\hbar}{2m}B$. Define the energy separtaion to be

$$\hbar\omega = \frac{ge\hbar}{2m}B,$$

so that $\omega = \frac{geB}{2m}$. This has the dimensions of frequency, and we'll see that this has something to do with oscillations. We can write

$$H = -\omega S_z.$$

We would like to understand the time evolution operator — the Hamiltonian is independent of time, so we can write

$$U(t,0) = e^{-iHt/\hbar} = e^{i\omega S_z t/\hbar}.$$

Let us consider an arbitrary state $|\psi\rangle = c_+|+\rangle + c_-|-\rangle$; what is $|\psi(t)\rangle$? We now have the power to answer such questions. Well,

$$|\psi(t)\rangle = e^{i\omega S_z t/\hbar}(c_+|+\rangle + c_-|-\rangle) = c_+ e^{i\omega t/2}|+\rangle + c_- e^{-i\omega t/2}|-\rangle.$$

If, initially, $c_+ = 1$ and $c_- = 0$, then $|\psi(t)\rangle = e^{i\omega t/2}|+\rangle$, so that

$$\text{Prob}(S_z = \hbar/2) = 1$$

for *all* t. This is true for any energy eigenstate (if the Hamiltonian is t-independent)... For this reason, energy eigenstates are sometimes called stationary states.

If, instead,

$$|\psi\rangle = |S_x = \hbar/2\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle).$$

Then, $c_+ = c_- = 1/\sqrt{2}$. Then

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}(e^{i\omega t/2}|+\rangle + e^{-i\omega t/2}|-\rangle).$$

Let's now compute

$$\text{Prob}(S_x = \hbar/2 \text{ at time } t) = \left|\left(\frac{\langle+| + \langle-|}{\sqrt{2}}\right)|\psi(t)\rangle\right|^2 = \cos^2\left(\frac{\omega t}{2}\right).$$

Similarly,

$$\text{Prob}(S_x = -\hbar/2 \text{ at time } t) = \sin^2\left(\frac{\omega t}{2}\right).$$

Let me say one last thing before concluding. With this solution for the dynamics, we can look at expectation values:

$$\langle S_x\rangle = \frac{\hbar}{2}\left(\cos^2\left(\frac{\omega t}{2}\right) - \sin^2\left(\frac{\omega t}{2}\right)\right) = \frac{\hbar}{2}\cos(\omega t).$$

You see that the average value of S_x oscillates periodically in time, with angular frequency ω. That's what you should expect, from your classical intuition.

7. More on spin precession in a B-field, and general time dependent Hamiltonians

We pick off from last time. Recall that

$$\vec{S}(t) = e^{-i\omega S_z t/\hbar}\vec{S}e^{i\omega S_z t/\hbar},$$

so that

$$\frac{d\vec{S}}{dt} = -\frac{i\omega}{\hbar}\left(e^{-i\omega S_z t/\hbar}[S_x, \vec{S}]e^{i\omega S_z t/\hbar}\right).$$

Since $[S_z, S_z] = 0$, we immediately conclude that $dS_z/dt = 0$. Since $[S_z, S_x] = i\hbar S_y$, we conclude that $dS_x/dt = \omega S_y(t)$, and similarly, $dS_y/dt = -\omega S_x(t)$. Therefore,

$$\frac{d\vec{S}}{dt} = -\omega\vec{S}(t) \times \hat{z};$$

this precisely tells you that the spin precesses. This is the same as for classical spin, but now it is an operator equation. However, after taking the expectation value, we find that

$$\left\langle \frac{d\vec{S}}{dt} \right\rangle = -\omega\langle\vec{S}\rangle \times \hat{z}.$$

In the Schrödinger picture the point on the Bloch sphere is moving around, whereas in the Heisenberg picture, it's the operator that's moving around. This exercise has many interesting applications in physics. Any Hilbert system with a two-dimensional state space has exactly the same mathematical properties. Sakurai discusses an interesting application of this exercise to neutrino oscillations, which is a very cool phenomenon.

Let me talk about other examples of dynamics in quantum mechanics. In this example, we will consider a particle in a potential, and discuss Ehrenfest's theorem. This is a good point to make a general comment about specifying quantum mechanical systems. There are two aspects to specifying a quantum system; the first is specifying what the Hilbert space is, and the second is to specify the Hamiltonian (or some equivalent, like the Lagrangian). Often people only specify the Hamiltonian, but that's not good.

For a particle moving in a potential, the Hilbert space is just the space $L^2(\mathbf{R})$ of square-integrable function on the real line. What is the Hamiltonian? One way to get the quantum Hamiltonian is to just take the classical Hamiltonian (which is a function of q and p), and just replace position and momentum with operators. Of course, this procedure could, in general, be ambiguous; for instance, you can have a product of terms, which, in the quantum mechanical setting, would require an operator ordering (since operator multiplication does not commute). Some care is therefore required, but we will only stick to the naïve way of writing down the Hamiltonian.

It follows from the above discussion that the Hamiltonian is

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

In the Schrödinger picture, we first find the energy eigenstates $|j\rangle$ and the corresponding energy eigenvalues E_j. This allows us to write an arbitrary state as $|\psi\rangle = \sum_j c_j|j\rangle$, with $c_j = \langle j|\psi\rangle$, so that the time evolved state is

$$|\psi(t)\rangle = e^{-iHt/\hbar}|\psi\rangle = \sum_j c_j e^{-iE_j t/\hbar}|j\rangle.$$

We talked about this in the previous class.

Let's try to do this in the Heisenberg picture. Then, we are told to think about the time evolution of operators, so that

$$x_H(t) = e^{iHt/\hbar}x_H(0)e^{-iHt/\hbar},\ p_H(t) = e^{iHt/\hbar}p_H(0)e^{-iHt/\hbar}.$$

The subscript H stands for "Heisenberg"; since that's clear at the moment, I'll just drop it. The equation of motion is

$$\frac{dx_H}{dt} = \frac{1}{i\hbar}[x_H, H].$$

What is the commutator appearing above?

$$\begin{aligned}[x, H] &= \left[x, \frac{p^2}{2m} + V(x)\right] = \left[x, \frac{p^2}{2m}\right] \\ &= \left[x, \frac{p}{2m}\right] p + p \left[x, \frac{p}{2m}\right] \\ &= \frac{i\hbar p}{m}.\end{aligned}$$

Therefore, $\frac{dx}{dt} = \frac{p}{m}$; this the quantum mechanical version of the statement that momentum is mass times velocity. Moreover, we have

$$\begin{aligned}\frac{dp}{dt} &= \frac{1}{i\hbar}\left[p, \frac{p^2}{2m} + V(x)\right] = \frac{1}{i\hbar}[p, V(x)] \\ &= -\frac{i\hbar}{i\hbar}\left(\frac{d}{dx}(V(x)\cdot \bullet) - V(x)\frac{d}{dx}\bullet\right),\end{aligned}$$

so that $\frac{dp}{dt} = -\frac{dV}{dx}$. This is precisely Newton's law. It's important that these are operators — the physics is definitely different, but the formal structure of the mathematics is the same. To make things similar, we can take expectation values, and we find:

$$\frac{d}{dt}\langle x\rangle = \frac{\langle p\rangle}{m}, \ \frac{d}{dt}\langle p\rangle = -\left\langle \frac{dV}{dx}\right\rangle.$$

This is Ehrenfest's theorem. It's important that we take the average of dV/dx, instead of the derivative of the average of V.

We will do a few more examples. Consider ac harged particle in a uniform E-field. Then,

$$H = \frac{p^2}{2m} - qE(t)x,$$

so that there is an operator equation $dp/dt = qE(t)$. It follows that

$$p(t) = p(0) + \int_0^t qE(t')dt'.$$

Similarly, there is an operator equation $dx/dt = p(t)/m$, so that

$$x(t) = x(0) + \frac{tp(0)}{m} + \frac{q}{m}\int_0^t dt' \int_0^t dt'' E(t'').$$

Think about how you would solve this in the Schrödinger formulation — it's really complicated to solve the Schrödinger equation in this case (but you do get the wavefunction as a function of time).

The next example I want to discuss is the simple harmonic oscillator, which I presume you all know how to solve in the Schrödinger picture. Let's examine this in the Heisenberg picture. The Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2.$$

Then, again,

$$\frac{dx}{dt} = \frac{p}{m}, \ \frac{dp}{dt} = -m\omega^2 x.$$

The solution is just

$$x(t) = x(0)\cos\omega t + \frac{p(0)}{m\omega}\sin\omega t,$$
$$p(t) = -\mathfrak{m}\omega x(0)\sin\omega t + p(0)\cos\omega t.$$

It's hard to find out the dynamics using the Schrödinger picture.

Let me now talk about a general time dependent Hamiltonian. (I already gave you an example with the charged particle in a uniform E-field, but nobody complained.) If H is t-independent, then we know that

$$U(t,0) = e^{-iHt/\hbar},$$

but this will not work in the general case. We know that we can always write

$$U(t+dt,t) = 1 - \frac{i}{\hbar}H(t)dt + O(dt^2),$$

with $H(t)$ hermitian — this was the definition of $H(t)$. We therefore find that

$$U(t+dt,t_0) = U(t+dt,t)U(t,t_0) = \left(1 - \frac{i}{\hbar}H(t)dt\right)U(t,t_0)$$

in the limit as $dt \to 0$. In particular,

$$U(t+dt,t_0) - U(t,t_0) = -\frac{i}{\hbar}H(t)U(t,t_0)dt.$$

Dividing by dt, we find that the left hand side is $\frac{d}{dt}U(t,t_0)$ in the limit when $t \to 0$. We therefore find the differential equation

$$i\hbar\frac{d}{dt}U(t,t_0) = H(t)U(t,t_0).$$

If we can solve this, then we find $U(t,t_0)$ in terms of $H(t)$.

If $H(t) = H$ is independent of t, then we obviously get $U(t,0) = e^{-iHt/\hbar}$, as expected. The next simplest case to think about is if $H(t)$ is time-dependent, but $[H(t),H(t')] = 0$ for all t,t'. In this case, the Hamiltonians at different times are different, but mutually compatible, operators, and hence they can be simultaneously diagonalized — therefore, if we diagonalize the Hamiltonian at t, then it will remain diagonalized for all $t' > t$. There is therefore a common eigenbasis for all the operators $H(t)$. Then, we can just solve this differential equation to find that

$$U(t,t_0) = \exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\ H(t')\right).$$

To see this, you can work in the common eigenbasis for the $H(t)$.

Let's move on to the real challenge: the most general case, where $[H(t),H(t')] \neq 0$. For example, you could take the spin in a magnetic field, where the magnetic field changes orientation as time evolves. Then, at different times, the Hamiltonian involves different components of the spin operator, so you're in this setting. Once we know the Hamiltonian, we know the time evolution for an infinitesimal time step, and you can write down the general time evolution as a sum of infinitesimal time steps.

More precisely, let us discretize steps into N steps, say $t_0 < t_1 < \cdots < t_{N-1} < t$, with $\Delta t = t_{i+1} - t_i = \frac{t-t_0}{N}$. Note that the choice of N is up to you. Since N is arbitrary, we can choose to make N go to ∞, and $\Delta t \to 0$. So let us choose $N \gg 0$, so $\Delta t \ll 0$. It follows that

$$U(t_{i+1},t_i) \approx 1 - \frac{i}{\hbar}H(t_i)\Delta t.$$

This equation is only correct to order Δt as $\Delta t \to 0$. We can write this as

$$U(t_{i+1},t_i) \approx e^{-\frac{i}{\hbar}H(t_i)\Delta t},$$

up to errors of order $(\Delta t)^2$. The general time evolution operator is

$$U(t,t_0) = U(t,t_{N-1})\cdots U(t_{i+1},t_i)\cdots U(t_1,t_0) = \prod_{0\le i\le N-1} U(t_{i+1},t_i).$$

The product symbol is frought with danger — what we have to remember is that these operators do not commute with each other, so the operator ordering is absolutely crucial. This equation is meaningless unless we know how to order these operators, but we have specified this above. Therefore

$$U(t,t_0) = \prod_{i=0}^{N-1} e^{-\frac{i}{\hbar}H(t_i)\Delta t}.$$

If the Hamiltonians at different times commuted, we could have just added the exponentials. In the limit of $\Delta t \to 0$, let us formally write this as

$$U(t,t_0) = \mathrm{T}\left[\exp\left(-\frac{i}{\hbar}\int_0^t H(t')dt'\right)\right].$$

The symbol T[−] is called a time-ordered exponential; it is defined to be the product above. It reorders the operators so that they are time ordered (with the later times to the left). Part of the statement here is that the limit exists (it being $U(t,t_0)$), and leads to a well-defined operator on the Hilbert space.

Recitation: two state systems

Two state systems are quantum systems with a two-dimensional Hilbert space $\mathcal{H} = \{a|0\rangle + b|1\rangle; a,b \in \mathbf{C}\}$, with $|0\rangle$ and $|1\rangle$ forming an orthonormal basis. Let me warn you that not all of the states here are not physically realizable (like the zero vector). What are the physically realizable states? We need the state to be normalized, so that $|a|^2 + |b|^2 = 1$. Moreover, the overall phase on the wavefunction does not affect the probability distributation (and, at least for now, means that it is physically unimportant). Sometimes this is known as the "gauge" degree of freedom.

If we write $a = |a|e^{i\theta}$ and $b = |b|e^{i\theta'}$, then after multiplying by $e^{-i\theta}$, we get $a = |a|$ and $b = |b|e^{i(\theta'-\theta)}$. Let $\theta' - \theta = \phi$. We can now set a to be real, and b to be what it is above. The physically realizable states in a two-dimensional Hilbert space therefore looks like

$$\mathbf{P}(\mathcal{H}) = \left\{\cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\theta}\sin\left(\frac{\theta}{2}\right)|1\rangle \mid 0 \le \theta \le \pi,\ 0 \le \phi < 2\pi\right\}.$$

The notation $\mathbf{P}(\mathcal{H})$ is called the *projective Hilbert space*. It's also denoted $\mathbf{C}P^1$; this is Bloch sphere, which should be thought of as the unit sphere in $\mathbf{R}^3$. An arbitrary unit vector $\hat{n}$ would have coordinates given by $(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. This is the eigenvector corresponding to the eigenvalue $\hbar/2$ of the $S_{\hat{n}}$ operator.

Let us now talk about neutrino oscillations. Hopefully everyone is familiar with the particle content of the standard model. Neutrinos (ν_e, ν_μ, ν_τ — these are called the *flavor eigenstates*) are extremely light particles, up until too long ago, we thought to be massless. They are leptons (don't interact with the strong force) and have no electric charge. This also means that they are really hard to detect, because they only weakly interact via the weak force and gravity. The weak force was originally used to describe β-decay (neutrons $\to$ protons, electrons, and $\overline{\nu_e}$).

For simplicity, let us only consider the states $|\nu_e\rangle$ and $|\nu_\mu\rangle$. An electron neutrino reacts with an electron, a mu neutrino reacts with a muon, and a tau neutrino reacts with a tau particle; these are eigenstates the weak force Hamiltonian. When the neutrinos are propagating and not interacting, they are governed by the free Hamiltonian H_0. It's conceivable that the weak

Hamiltonian and the free Hamiltonian don't have a shared eigenbasis. This is a weird idea, but turns out to be true in nature.

The "mass eigenstates" (eigenstates of H_0) are going to be denoted $|\nu_1\rangle$ and $|\nu_2\rangle$. Let's say that have corresponding eigenvalues E_1 and E_2. The most general thing we can write down is that $|\nu_1\rangle = \cos\theta|\nu_e\rangle + e^{i\theta}|\nu_\mu\rangle$ and $|\nu_2\rangle = -\sin\theta|\nu_e\rangle + e^{i\theta}\cos\theta|\nu_\mu\rangle$; we've dropped the factor of $1/2$, so $0 \le \theta \le 2\pi$. Redefine the phase of $|\nu_\mu\rangle$ to $e^{-i\theta}|\nu_\mu\rangle$; then, we can remove the phase $e^{i\theta}$ in the equations for $|\nu_1\rangle$ and $|\nu_2\rangle$ above.

The only physical parameter we have here is θ, called the "mixing angle". As an aside, in the three neutrino case, you get three different mixing angles, and one phase denoted δ_{cp}. The latter phase is interesting is because, if it is nonzero, it violates CP symmetry. That explains why there's more matter than antimatter in the universe.

In any case, we find that

$$|\nu_e\rangle = \cos\theta|\nu_1\rangle - \sin\theta|\nu_2\rangle$$
$$|\nu_\mu\rangle = \sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle.$$

Let's consider an actual physical example. In the sun, the first nuclear reaction that starts nuclear fusion is the reaction

$$p^+ + p^+ \to p^+ + n + e^+\nu_e;$$

the $p^+ + n$ is usually called a "deutrino". The sun therefore produces tons of electron neutrinos ν_e. Well:

$$|\nu_e(t)\rangle = \cos\theta e^{-iE_1t/\hbar}|\nu_1\rangle - \sin\theta e^{-iE_2t/\hbar}|\nu_2\rangle.$$

If $E_1 \neq E_2$, there's a relative phase — therefore, we're no longer just an electron neutrino; there's a mixing of the muon neutrino. Therefore,

$$\langle\nu_e|\nu_e(t)\rangle = \cos^2 e^{-iE_1t/\hbar} + \sin^2\theta e^{-iE_2t/\hbar}$$

is the probability amplitude that the electron neutrino stays an electron neutrino after propagating for a time t. The obvious thing to do, in order to compute the probability, is take the absolute value of the probability squared. You can find that the electron neutrino is observed as an electron neutrino after propagating for a time t is

$$P(\nu_e \to \nu_e) = 1 - \sin^2(2\theta)\sin^2\left(\frac{(E_1 - E_2)t}{2\hbar}\right).$$

Let $\Delta E = E_1 - E_2$. Neutrinos are very light, and they are almost relativistic. Therefore,

$$E_i = \sqrt{p^2c^2 + m_ic^4} \approx pc\left(1 + \frac{m_i^2c^2}{2p^2}\right),$$

where m_i is the mass of $|\nu_i\rangle$. Suppose m_1 and m_2 are different; then, you'll find that

$$\Delta E = \frac{\Delta m^2c^2}{2p},$$

where $\Delta m^2 = m_1^2 - m_2^2$. It follows that

$$P(\nu_e \to \nu_e) = 1 - \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2c^3t}{4p\hbar}\right).$$

People have observed this probability to be *less* than one. It follows that $\theta \neq 0$, and $\Delta m^2 \neq 0$. For the muon and the electron neutrino, we have $\sin^2(2\theta) = 0.846 \pm 0.021$, and $\Delta m^2 = (7.53 \pm 0.18) \times 10^{-5}$ eV.

8. Dyson series, the Dirac picture, and spin $1/2$ in an ac-field

Last time, we observed that

$$U(t,t_0) = \lim_{N\to\infty,\Delta t\to 0, N\Delta t = t-t_0} [U(t,t_{N-1})\cdots U(t_{N-1},t_{N-2})\cdots U(t_1,t_0)]$$

with $U(t_{i+1},t_i) = e^{-iH(t_i)\Delta t/\hbar}$. We wrote this as

$$U(t,t_0) = \mathrm{T}\left[\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} dt' H(t')\right)\right].$$

Alternatively, we can start from the differential equation

$$i\hbar\frac{d}{dt}U(t,t_0) = H(t)U(t,t_0)$$

and try to solve for $U(t,t_0)$ in terms of a formal power series in H. We solve this iteratively. First, since

$$U\int_{t_0}^{t} dt' \frac{d}{dt'}U(t',t_0) = -\frac{i}{\hbar}\int_{t_0}^{t} dt' H(t')U(t',t_0),$$

we have

$$U(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt' H(t')U(t',t_0)$$

for $t' \le t$. Using the same equation to write $U(t',t_0)$ as an integral involving times $t'' \le t'$, we have

$$U(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt' H(t') + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H(t')H(t'')U(t'',t_0).$$

Now, keep going; you'll get

$$U(t,t_0) = \sum_{n\ge 0}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_1 \cdots \int_{t_0}^{t_{n-1}} dt_n H(t_1)H(t_2)\cdots H(t_n).$$

The ordering of the Hamiltonians is important. We can rewrite this as follows:

$$\boxed{U(t,t_0) = \sum_{n\ge 0}\left(-\frac{i}{\hbar}\right)^n \frac{1}{n!}\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t} dt_n \mathrm{T}\left[H(t_1)H(t_2)\cdots H(t_n)\right]}.$$

This is known as the "Dyson series".

We'll now move on to the interaction picture, which is due to Dirac. This is a mixture of the Schrödinger and Heisenberg pictures. Suppose the Hamiltonian has the form $H(t) = H_0(t) + V(t)$, where the time evolution due to H_0 is known, and $V(t)$ is a perturbation. The idea is to remove the evolution due to H_0 from the state.

Define the "interaction ket" $|\psi(t)\rangle_I = U_0^{-1}(t)|\psi(t)\rangle_S$, where $|\psi(t)\rangle_S$ is the "Schrödinger" ket. This should be contrasted with the Heisenberg picture, where we removed *all* time dependence from the ket. Then, the evolution equation is

$$i\hbar\frac{d}{dt}|\psi(t)\rangle_I = V_I(t)|\psi(t)\rangle_I,$$

where $V_I(t) = U_0^{-1}(t)V(t)U_0(t)$. The operators evolve according to $A_I(t) = U_0^{-1}(t)AU_0(t)$. Roughly speaking, we're changing our frame of reference to just riding the wave given by $H_0(t)$.

Let us derive the equation of motion for the evolution operator directly in the interaction picture. Define $U_I(t)$ by

$$U(t) = U_0(t)U_I(t).$$

We have

$$i\hbar\frac{d}{dt}U(t) = (H_0 + V)U(t),$$
$$i\hbar\frac{d}{dt}U_0(t) = H_0U_0(t).$$

This gives

$$i\hbar\frac{d}{dt}U(t) = i\hbar\frac{dU_0}{dt}U_I + i\hbar U_0\frac{dU_I}{dt} = H_0U_0U_I + i\hbar U_0\frac{dU_I}{dt} = H_0U + i\hbar U_0\frac{dU_I}{dt}.$$

It follows that

$$\boxed{i\hbar\frac{dU_I}{dt} = U_0^{-1}VU = U_0^{-1}VU_0U_I = V_IU_I}.$$

Once we know how to transform the original perturbation $V(t)$ into this frame of reference, we can determine the remaining time evolution U_I.

The example we're about to talk about is incredible relevant to, e.g., nuclear magnetic resonance. This is the example of a spin $1/2$ particle in an ac field. Assume that we have a spin $1/2$ particle in an S^z direction, and a smaller field. So the B-field has

$$\vec{B}(t) = B_0\hat{z} + B_1(\cos\omega t\ \hat{x} + \sin\omega t\ \hat{y}).$$

The Hamiltonian is

$$H = -\frac{ge}{2m}\vec{S}\cdot\vec{B}(t).$$

In the case when $B_1 = 0$, we have two energy eigenstates (spin up and spin down along the z-axis), which have an energy splitting of

$$\Delta E = \left|\frac{geB_0\hbar}{2m}\right|.$$

This is the spectrum in the absence of the time-dependent perturbation. We'll use the interaction picture to study the case of the time-dependent perturbation, i.e., when $B_1 \neq 0$.

We'll see that there's a subtlety in this interaction picture. What we are given is the Hamiltonian H — nobody tells us what H_0 and V are, so we have to decide this for ourselves. Here, it seems like there is an obvious choice. Let's run with this choice; later, we'll see that we'll have to modify our choice in other situations.

Namely, we will try

$$H_0 = -\frac{ge}{2m}S_zB_0,\ V(t) = -\frac{ge}{2m}\vec{S}_\perp\cdot\vec{B}_\perp(t),$$

where $\vec{B}_\perp = B_1(\cos\omega t\ \hat{x} + \sin\omega t\ \hat{y})$. Let's write $H_0 = \omega_0S_z$, where $\omega_0 = \left|\frac{geB_0}{2m}\right|$. We know from earlier that

$$U_0(t) = e^{-\frac{i}{\hbar}H_0t} = e^{-i\omega_0tS_z/\hbar}.$$

We know that the evolution is given by

$$i\hbar\frac{d}{dt}|\psi(t)\rangle_I = V_I(t)|\psi(t)\rangle_I.$$

We would therefore like to understand V_I. By definition:

$$V_I(t) = -\frac{ge}{2m}B_1e^{i\omega_0S_zt/\hbar}(\cos\omega t\ S_x + \sin\omega t\ S_y)e^{-i\omega_0S_zt/\hbar}.$$

Each one of these operators can be expressed as a product of 2×2-matrices (the Pauli matrices), so you can do it in the brute-force way. Instead, we can be slicker. It is convenient to write this

in terms of $S^{\pm} = S_x + iS_y$. Clearly S_x and S_y are linear combinations of $S^{\pm}$. If you know how $S^{\pm}$ is transformed, then you know the above expression. In general, we are considering

$$e^{i\phi\sigma^z/2} S^+ e^{-i\phi\sigma^z/2} = \left(\cos\frac{\phi}{2} + \sigma^z i \sin\frac{\phi}{2}\right) S^+ \left(\cos\frac{\phi}{2} - \sigma^z i\sin\frac{\phi}{2}\right).$$

This can be rewritten as

$$S^+ \left(\cos\frac{\phi}{2} - i\sigma^z \sin\frac{\phi}{2}\right)^2 = \frac{\hbar}{2}(\sigma^x + i\sigma^y)(\cos\phi - i\sigma^z\sin\phi) = S^+ e^{i\phi}.$$

It follows that

$$e^{i\phi S^z/\hbar} S^+ e^{-i\phi S^z/\hbar} = S^+ e^{i\phi}.$$

This also implies that (by taking †):

$$e^{i\phi S^z/\hbar} S^- e^{-i\phi S^z/\hbar} = S^- e^{-i\phi}.$$

Returning back to the problem: we have

$$\begin{aligned} V_I(t) &= -\frac{ge}{2m} B_1 e^{i\omega_0 S_z t/\hbar} (\cos\omega t\ S_x + \sin\omega t\ S_y) e^{-i\omega_0 S_z t/\hbar} \\ &= -\frac{ge}{2m} B_1 e^{i\omega_0 S_z t/\hbar} \left(\frac{S^+ e^{i\omega t} + S^- e^{-i\omega t}}{2}\right) e^{-i\omega_0 t S_z/\hbar} \\ &= -\frac{1}{2}\frac{ge}{2m} B_1 \left(S^+ e^{i(\omega_0+\omega)t} + S^- e^{-i(\omega_0+\omega)t}\right). \end{aligned}$$

Let us analyze the simple case $\omega = -\omega_0$ (the "resonant drive"). Then,

$$V_I(t) = -\frac{ge}{4m} B_1 (S^+ + S^-) = -\frac{ge}{2m} B_1 S_x.$$

We therefore have

$$U_I(t) = e^{igeB_1 S_x t/2m\hbar}.$$

This rotates about the x-axis by the angle $\frac{ge}{2m\hbar} B_1 t$. The frequency is therefore $\omega_R = \left|\frac{geB_1}{2m}\right|$, and this is called the "Rabi frequency". So

$$U_I\left(t + \frac{4\pi}{\omega_t}\right) = U_I(t),$$

which implies that the wavefunction is periodic in time with period $\frac{4\pi}{\omega_t}$. These are known as "Rabi oscillations".

Let's look at an example. If $|\psi(0)\rangle = \binom{1}{0}$, i.e., a spin up state. Then

$$|\psi(t)\rangle_I = \exp(i\omega_R t\sigma^x/2)|\psi(0)\rangle = \begin{pmatrix} \cos\omega_R t/2 \\ i\sin\omega_R t/2 \end{pmatrix}.$$

After time π/ω_R, the state gets flipped to a spin down state.

9. Path integrals

We will first discuss the off-resonant drive, i.e., what happens if $\omega \neq \omega_R = -\omega_0$. If you choose

$$V_I(t) = \frac{B_1}{2}(S^+ e^{i(\omega_0+\omega)t} + S^- e^{-i(\omega_0+\omega)t}),$$

which still depends on t, so the interaction picture will still be hard. We wouldn't have gained very much.

Let's exploit the freedom we have in the interaction picture. We will make a different choice of H_0 and V; namely, define $H = \widetilde{H}_0 + \widetilde{V}$, with

$$\widetilde{H}_0 = \omega S_z,\ \widetilde{V} = -\gamma B_{eff} S_z - \gamma B_1 (S_x \cos\omega t + S_y \sin\omega t),$$

where $\gamma = ge/2m$ and $B_{eff} = B_0 + \frac{\omega}{\gamma}$. We'll see that this is a sensible choice.

With this choice, let us go to the rotating frame, and calculate:

$$\widetilde{V}_I(t) = \widetilde{U}_0^\dagger(t)\widetilde{V}\widetilde{U}_0(t) = e^{i\omega S_z t/\hbar}\left[-\gamma B_{eff}S_z - \frac{\gamma B_1}{2}\left(S^+e^{i\omega t} + S^-e^{-i\omega t}\right)\right]e^{-i\omega S_z t/\hbar}.$$

All operators in the first term in the square brackets above commute with S_z, so it remains unchanged. It remains to understand the term in parentheses. Last time, we computed that

$$e^{i\phi S^z/\hbar}S^+e^{-i\phi S^z/\hbar} = S^+e^{i\phi},\ e^{i\phi S^z/\hbar}S^-e^{-i\phi S^z/\hbar} = S^-e^{-i\phi}.$$

It follows that

$$\widetilde{V}_I(t) = -\gamma(B_{eff}S_z + B_1S_x),$$

which is t-independent. By choosing the right rotating frame, we're able to reduce the interaction Hamiltonian to something that is t-independent. The spin therefore precesses about the net field $\vec{\widetilde{B}} = (B_1\hat{x}, 0, (B_0 + \frac{\omega}{\gamma})\hat{z})$, which leads to modified Rabi oscillations at a frequency $\omega_R = \gamma\sqrt{B_{eff}^2 + B_1^2}$.

We will now move on to path integrals. This is due to Feynman. Nobody will ever solve the Schrödinger equation for the harmonic oscillator or the Hydrogen atom using the path integral formalism, but it is useful for qualitatively understanding a system. Let's start with the Schrödinger picture. Write

$$|\psi(t)\rangle = \sum_{a'} c_{a'}(t)|a'\rangle,$$

where a' is an energy eigenvalue. Then $c_{a'}(t) = e^{-iE_{a'}(t-t_0)/\hbar}c_{a'}(t_0)$. For particles in position space in any dimension, define $u_{a'}(x) = \langle x|a'$. This is the energy wavefunction. Then

$$\psi(x,t) = \sum_{a'} e^{-iE_{a'}(t-t_0)/\hbar}c_{a'}(t_0)u_{a'}(x).$$

Let's rewrite this as

$$\psi(x,t) = \int dx' K(x,t;x',t_0)\psi(x',t_0),$$

with

$$K(x,t;x',t_0) = \sum_{a'}\langle x|a'\rangle e^{-iE_{a'}(t-t_0)/\hbar}\langle a'|x'\rangle = \langle x|U(t,t_0)|x'\rangle.$$

This K is known as the *propagator*. This is independent of ψ — it depends only on the energy eigenstates. All information in U is available in K, since it is just the matrix representation of U in the position basis. It follows that K satisfies

$$\left(i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2}{2m}\nabla^2 - V(x)\right)K(x,t;x',t_0) = 0.$$

We also need

$$\lim_{t\to t_0} K(x,t;x',t_0) = \delta^{(d)}(x-x').$$

A related quantity is the "retarded" propagator

$$K_{\text{ret}}(x,t;x',t_0) = \theta(t-t_0)K(x,t;x',t_0),$$

where

$$\theta(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}.$$

We only use this because you should be able to propagate from t_0 to t only if $t > t_0$. Note that θ has a discontinuity, so its derivative is the Dirac delta function. In other words, $\theta'(t) = \delta(t)$. The retarded propagator satisfies a modified equation:

$$\left(i\hbar\partial_t + \frac{\hbar^2}{2m}\nabla^2 - V(x)\right) K_{\text{ret}}(x,t;x',t_0) = i\hbar\delta^{(d)}(x-x')\delta(t-t_0).$$

This means that K_{ret} is what's called a Green's function for the Schrödinger equation. Another way to think about K is that

$$K(x,t;x',t_0) = \langle x,t|x',t_0\rangle,$$

because $\langle x|U(t,t_0) = \langle x,t|$ (by definition, if you want).

Let's calculate the function K in certain cases. We will compute K for a free particle in one dimension. Then

$$K(x,t;x',t') = \int dp \langle x|p\rangle e^{-\frac{i}{\hbar}\frac{p^2}{2m}(t-t')}\langle p|x'\rangle = \int \frac{dp}{2\pi\hbar} e^{ip(x-x')-\frac{p^2}{2m\hbar}(t-t')}.$$

Now, what is $\int_0^\infty dx\ e^{-ax^2}$? Well, its square is

$$\int_0^\infty dy \int_0^\infty dx\ e^{-ax^2-ay^2} = \int_0^{2\pi} d\theta \int_0^\infty r dr e^{-ar^2} = \frac{2\pi}{2}\int_0^\infty d(r^2)e^{-ar^2} = \frac{\pi}{a},$$

so the integral evaluates to $\sqrt{\frac{\pi}{a}}$.

In any case, we compute that

$$K(x,t;x',t') = \sqrt{\frac{m}{2\pi i\hbar(t'-t)}}\exp\left(i\frac{m(x-x')^2}{2\hbar(t-t')}\right).$$

Let us now talk about properties of K. Why should we spend our efforts in computing K? There are many reasons.

(1) There is a deep relationship with quantum statistical mechanics. Define

$$G(t) = \int d^dx\ K(x,t;x,0).$$

Then

$$G(t) = \int d^dx\ \sum_{a'}\langle x|a'\rangle e^{-iE_{a'}t/\hbar}\langle a'|x'\rangle = \sum_{a'} e^{-iE_{a'}t/\hbar}.$$

Let $t = -i\hbar\beta$; then

$$G(-i\hbar\beta) = \sum_{a'} e^{-\beta E_{a'}};$$

this is the partiion function in statistical mechanics. What may happen is that your quantum mechanical system might be at equilibrium, and then your partition function would describe all thermodynamics quantities.

(2) Take the Fourier transform, i.e., define

$$\widetilde{G}(E) = -i\int dt G_{\text{ret}}(t)e^{iEt/\hbar},$$

where

$$G_{\text{ret}}(t) = \int d^dx\ K_{\text{ret}}(x,t;x,0).$$

It turns out that $\widetilde{G}(E)$ contains an enormous amount of information. We need to first jump over a hurdle. Explicitly:

$$\widetilde{G}(E) = -i\sum_{a'}\int_0^\infty dt\ e^{i(E-E_{a'})t/\hbar}.$$

Earlier, when we did the computation for K of a free particle, the integral was convergent. But the magnitude of the integrand in this case is always one — and the integral is not always well-defined! We can say that we defined a wrong quantity and give up, or we can invent a fudge.

This fudge was invented by Feynman. He was the master of fudging things in exactly the right way. For convergence, we'll take E to have a small imaginary part $E \to E + i\epsilon$ with $\epsilon > 0$ is small. This is called the "Feynman $i\epsilon$-prescription". Then, the integral is convergent. The sign of ϵ is important.

The integral now becomes

$$\widetilde{G}(E) = -i \sum_{a'} \int_0^\infty dt\ e^{i(E+i\epsilon-E_{a'})t/\hbar} = \sum_{a'} \frac{\hbar}{E - E_{a'} + i\epsilon}.$$

This has poles on the real axis, and each pole describes an energy eigenvalue. It follows that understanding this function $\widetilde{G}$ (in particular, its poles), immediately tell us all of the energy eigenstates.

Let's define a quantity known as the density of states:

$$\rho(E) = \sum_{a'} \delta(E - E_{a'}).$$

This quantity contains all of the information about the spectrum. This function is easily obtained once we know $\widetilde{G}(E)$. Let me explain this. Use

$$\lim_{\epsilon \to 0^+} \mathrm{im} \left(\frac{1}{E - E' + \epsilon} \right) = - \lim_{\epsilon \to 0^+} \frac{\epsilon}{(E - E')^2 + \epsilon^2}.$$

Suppose $E - E' \neq 0$. Then the denominator goes to some nonzero constant, and so the limit goes to zero. If $E = E'$, then the function becomes $1/\epsilon$, so the limit goes to ∞. This is therefore almost a delta function. In any case, you find that the above limit is exactly $-\pi\delta(E - E')$. It follows that

$$\rho(E) = -\frac{1}{\pi\hbar} \lim_{\epsilon \to 0^+} \mathrm{Im}\ \widetilde{G}(E + i\epsilon).$$

Recall that $U(t, t_0) = U(t, t')U(t', t_0)$ for $t > t' > t_0$. An equivalent way to say this in the position basis is

$$K(x, t; x', t_0) = \int d\widetilde{x}\ K(x, t; \widetilde{x}, \widetilde{t}) K(\widetilde{x}, \widetilde{t}; x', t_0).$$

Recitation

In general, if you have a t-dependent operator $B_s(t)$, then

$$B_H(t) = U(t)^\dagger B_s(t) U(t),$$

so that

$$\begin{aligned} \frac{d}{dt} B_H(t) &= \frac{dU^\dagger}{dt} B_s U + U^\dagger \frac{dB_s}{dt} U + U^\dagger B_s \frac{dU}{dt} \\ &= \frac{i}{\hbar} U^\dagger H B_s U + U^\dagger \frac{dB_s}{dt} - \frac{i}{\hbar} U^\dagger B_s H U \\ &= \frac{i}{\hbar} [H, B_s]_H + \left(\frac{dB_s}{dt} \right)_H. \end{aligned}$$

Note that

$$\begin{aligned}[A_H, B_H] &= U^\dagger (A_s U U^\dagger B_s - B_s U U^\dagger A_s) U \\ &= U^\dagger [A_s, B_s] U = [A_s, B_s]_H.\end{aligned}$$

We're now going to study the example of the harmonic oscillator, which is useful everywhere, and in particular, in quantum field theory. The Hamiltonian is

$$H = \frac{\hat{p}^2}{2m} + \frac{1}{2} m\omega^2 \hat{x}^2.$$

One thing that I hope everybody has seen is the following equation

$$H = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right).$$

These are the creation and annihilation operators, also known as the raising and lowering operators. You can write

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + \frac{i}{m\omega} \hat{p} \right).$$

These satisfy the following properties:

$$[\hat{a}, \hat{a}] = 0 = [\hat{a}^\dagger, \hat{a}^\dagger],\ [\hat{a} + \hat{a}^\dagger] = 1.$$

This latter relation is really the defining relation, which makes it extremely useful in quantum field theory. Let us write

$$d = \sqrt{\frac{\hbar}{m\omega}}.$$

It follows that we can write

$$\hat{x}(t) = \frac{d}{\sqrt{2}} (\hat{a}(t) + \hat{a}^\dagger(t)),\ \hat{p}(t) = -\frac{i\hbar}{d\sqrt{2}} (\hat{a}(t) - \hat{a}^\dagger(t)).$$

Now, let us go to the Heisenberg equations of motion. In class, we saw that it was easy to write down the equations of motion for $\hat{x}, \hat{p}$ — but it's even easier to do this for $\hat{a}$:

$$\frac{d\hat{a}_H}{dt} = \frac{i}{\hbar} [H, \hat{a}]_H = i\omega [\hat{a}^\dagger \hat{a}, \hat{a}]_H.$$

The operator $\hat{a}^\dagger \hat{a}$ is called the "number operator" (it tells you what energy level you are at). This becomes

$$i\omega (\hat{a}^\dagger [\hat{a}, \hat{a}] + [\hat{a}^\dagger, \hat{a}] \hat{a})_H = -i\omega \hat{a}_H.$$

This is a differential equation which everyone knows how to solve:

$$\hat{a}_H(t) = e^{-i\omega t} \hat{a}_H(0), \qquad \hat{a}_H^\dagger(t) = e^{i\omega t} \hat{a}_H^\dagger(0).$$

These imply that

$$\begin{aligned}\hat{x}(t) &= \hat{x}(0) \cos(\omega t) + \frac{1}{m\omega} \hat{p}(0) \sin(\omega t) \\ \hat{p}(t) &= \hat{p}(0) \cos(\omega t) - m\omega \hat{x}(0) \sin(\omega t).\end{aligned}$$

One thing that the Heisenberg picture is useful for is talking in quantum mechanics about things that behave classically (take, e.g., Ehrenfest's theorem). Recall that

$$T_{x_0} = e^{-i\hat{p}x_0/\hbar},\ |\widetilde{x}_0\rangle = T_{x_0} |0\rangle.$$

Such states are called "coherent states". We can then compute

$$T_{x_0}^\dagger \hat{x} T_{x_0} = \hat{x} + \left[\hat{x}, -\frac{i}{\hbar} \hat{p} x_0 \right] + \frac{1}{2!} \left[\left[\hat{x}, -\frac{i}{\hbar} \hat{p} x_0 \right], -\frac{i}{\hbar} \hat{p} x_0 \right] + \cdots = \hat{x} + x_0.$$

We've used the *Baker-Campbell-Hausdorff* formula here. Note that

$$T_{x_0}^\dagger \hat{p} T_{x_0} = \hat{p}.$$

It follows that

$$T_{x_0}^\dagger H T_{x_0} = H + m\omega^2 x_0 \hat{x} + \frac{1}{2} m\omega^2 x_0^2.$$

In particular, we have

$$\begin{aligned}\langle \widetilde{x}_0 | H | \widetilde{x}_0 \rangle &= \langle 0|H|0\rangle + m\omega^2 x_0 \langle 0|x|0\rangle + \frac{1}{2} m\omega^2 x_0^2 \langle 0|0\rangle \\ &= \frac{\hbar\omega}{2} + 0 + \frac{1}{2} m\omega^2 x_0^2.\end{aligned}$$

This is just the classical energy of a particle! We don't have a zero point energy ($\langle 0|H|0\rangle = \frac{\hbar\omega}{2}$), though. It follows that

$$\begin{aligned}\langle \widetilde{x}_0 | \hat{x}(t) | \widetilde{x}_0 \rangle &= \langle \widetilde{x}_0 | \hat{x}(0) | \widetilde{x}_0 \rangle \cos(\omega t) + \frac{1}{m\omega} \langle \widetilde{x}_0 | \hat{p}(0) | \widetilde{x}_0 \rangle \sin(\omega t) \\ &= x_0 \cos(\omega t),\end{aligned}$$

and similarly

$$\langle \widetilde{x}_0 | \hat{p}(t) | \widetilde{x}_0 \rangle = -m\omega x_0 \sin(\omega t).$$

Those are exactly the classical energy relations! We have

$$|\widetilde{x}_0\rangle = \exp\left(\frac{x_0}{2\sqrt{d}} (\hat{a}^\dagger - \hat{a}) \right) |0\rangle.$$

The algebra is a little horrible, but you can do it.

A general coherent state of a harmonic oscillator is

$$|\alpha\rangle = \exp(\alpha \hat{a}^\dagger - \alpha^\star \hat{a}) |0\rangle,$$

where α is a complex number. Using the Zassenhaus formula, we have

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} e^{\alpha \hat{a}^\dagger} |0\rangle.$$

An important property of states like this is: if f is any function with a Taylor expansion, then

$$\hat{a} f(\hat{a}^\dagger) |0\rangle = f'(\hat{a}^\dagger) |0\rangle.$$

This is a good exercise. We conclude that

$$\hat{a}|\alpha\rangle = \alpha |\alpha\rangle.$$

These are strange states... They are eigenvectors of some operator which isn't even Hermitian. We therefore have

$$\langle \alpha | \hat{x} | \alpha \rangle = \frac{d}{\sqrt{2}} \langle \alpha | (\hat{a}^\dagger + \hat{a}) | \alpha \rangle = \frac{d}{\sqrt{2}} (\alpha^\star + \alpha) = d\sqrt{2} \mathrm{Re}(\alpha),$$

and similarly

$$\langle \alpha | \hat{p} | \alpha \rangle = -\frac{i\hbar}{d\sqrt{2}} \langle \alpha | (\hat{a} - \hat{a}^\dagger) | \alpha \rangle = \frac{h\sqrt{2}}{d} \mathrm{Im}(\alpha).$$

In other words, we have

$$\boxed{\alpha = \frac{\langle \hat{x} \rangle}{d\sqrt{2}} + i \frac{\langle \hat{p} \rangle d}{\hbar\sqrt{2}}.}$$

It follows that

$$\begin{aligned}|\alpha,t\rangle &= e^{-iHt/\hbar}|\alpha\rangle \\ &= e^{-iHt/\hbar}e^{\alpha\hat{a}^\dagger-\alpha^\dagger\hat{a}}e^{iHt/\hbar}e^{-iHt/\hbar}|0\rangle \\ &= e^{-iHt/\hbar}e^{\alpha\hat{a}^\dagger-\alpha^\dagger\hat{a}}e^{iHt/\hbar}e^{-i\omega t/2}|0\rangle \\ &= e^{-i\omega t/2}e^{\alpha\hat{a}^\dagger(-t)-\alpha^\star\hat{a}(-t)}|0\rangle \\ &= e^{-i\omega t/2}e^{\alpha e^{-i\omega t/2}\hat{a}^\dagger-\alpha^\star e^{i\omega t}\hat{a}}|0\rangle \\ &= e^{-i\omega t/2}|e^{-i\omega t}\alpha\rangle.\end{aligned}$$

10. Path integrals

Recall that $U(t,t_0) = U(t,t')U(t',t_0)$ for $t > t' > t_0$. An equivalent way to say this in the position basis is

$$K(x,t;x',t_0) = \int d\widetilde{x}\ K(x,t;\widetilde{x},\widetilde{t})K(\widetilde{x},\widetilde{t};x',t_0).$$

Break $t-t_0$ into N equal time intervals $t_0 < t_1 < \cdots < t_N = t$ with $\Delta t = \frac{t-t_0}{N}$. Then

$$K(x_N,t_N;x_0,t_0) = \int \prod_{k=1}^{N-1} dx_k\ K(x_N,t_N;x_{N-1},t_{N-1})\cdots K(x_1,t_1;x_0,t_0).$$

This is a sum over all possible N-step trajectories from the point (x_0,t_0) to the point (x_N,t_N). We can let $N\to\infty$, and find that the discrete time steps are close enough together that you can talk about paths in spacetime. Then this becomes a sum over all possible *continuous* trajectories from the point (x_0,t_0) to the point (x_N,t_N).

Feynman proposed that

$$K(x,t;x',t_0) = \int [\mathcal{D}x]\ e^{\frac{i}{\hbar}S[x(t)]}. \tag{2}$$

Most of the lecture will be devoted to understanding this statement. The right hand side is supposed to be a sum over all paths from (x',t_0) to (x,t). The weight (i.e., the integrand) is $e^{\frac{i}{\hbar}S}$, with S being the classical action of the trajectory. Let us remind ourselves of what the classical action refers to.

In classical mechanics, the Lagrangian is $L = \frac{1}{2}m\dot{x}^2 - V$, and the action is $S = \int dt\ L$. The classical trajectory between two spacetime points is extremized by the action: $\frac{\delta S}{\delta x} = 0$. This is the "principle of least action". The action is what is known as a *functional*: it's defined on the space of paths. You take the integral along this path, and this defines the action associated to that path; then, δS denotes the change in S after an infinitesimal variation of the path.

The Lagrangian point of view is different from the Newton and Hamiltonian point of view. In the latter, you specify the initial position and momentum, and in the former, you specify the initial position and velocity. The Lagrangian is different.

In any case, using the equation

$$\frac{\delta S}{\delta x} = 0,$$

you'll find the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}.$$

Using $L = \frac{1}{2}m\dot{x}^2 - V$, you find exactly Newton's law.

Pre-quantum mechanics, you could have taken the point of view that Hamilton and Lagrange were jobless and didn't have anything better to do. But there are incredibly important applications of this point of view on classical mechanics: it provides different ways of understanding quantum mechanics.

I want to make several remarks about Equation (2).

(1) $K = \int [\mathcal{D}x] e^{iS/\hbar}$ clearly obeys the composition law. If you sum over all possible paths from (x', t_0) to (x_1, t_1), and integrate this over all x_1 and sum over all paths from (x_1, t_1) to (x, t), you obtain precisely the sum over all possible paths from (x', t_0) to (x, t).
(2) This gives a very simple connection to classical mechanics. If $\hbar \to 0$, then the term $iS/\hbar$ in the exponent becomes large. This means that the phase varies widely, so you there to be cancellation (because of destructive interference). Thus, you expect the sum to be dominated by a path whose phase has dominant phase, i.e., S is extremized.

Let's now discuss (2). How do you define the measure in the space of paths? The issue of whether the path integral is defined or not is a difficult question which is not really answered in many instances. People go ahead and use it anyway.

Our goal is to make the phrase "sum over paths" precise, and "rigorous". We need a measure on the path space. The strategy is to start with the definition of K, and derive the path integral and the measure, and use this to rederive K for a free particle (as a check).

Recall that

$$\begin{aligned} K(x_N, t_N; x_0, t_0) &= \int \prod_{k=1}^{N-1} dx_k \; K(x_N, t_N; x_{N-1}, t_{N-1}) \cdots K(x_1, t_1; x_0, t_0) \\ &= \int \prod_{k=1}^{N-1} dx_k \; \langle x_N | U(t_N, t_{N-1}) | x_{N-1} \rangle \cdots \langle x_1 | U(t_1, t_0) | x_0 \rangle. \end{aligned}$$

We know that $U(t_{k+1}, t_k) = e^{-i\epsilon H/\hbar}$, where $\epsilon = \Delta t$. For simplicity, let us assume that the Hamiltonian is t-independent. Then, we have

$$\langle x_{j+1} | e^{-i\epsilon H/\hbar} | x_j \rangle = \langle x_{j+1} | e^{-i\epsilon (T+V)/\hbar} | x_j \rangle,$$

with $T = p^2/2m$ and $V = V(x)$ is the potential. When $\epsilon \to 0$, we have that

$$e^{-i\epsilon(T+V)/\hbar} = 1 - \frac{i\epsilon}{\hbar}(T+V) - \frac{\epsilon^2}{2\hbar^2}(T^2 + V^2 + TV + VT) + \cdots .$$

Therefore, as $\epsilon \to 0$, we can write

$$e^{-i\epsilon(T+V)/\hbar} = e^{-i\epsilon T/\hbar} e^{-i\epsilon V/\hbar}$$

up to order ϵ^2. It follows that

$$\begin{aligned} \langle x_{j+1} | e^{-i\epsilon(T+V)/\hbar} | x_j \rangle &\xrightarrow{\epsilon \to 0} \langle x_{j+1} | e^{-i\epsilon p^2/2m\hbar} e^{-i\epsilon V(x)/\hbar} | x_j \rangle \\ &= e^{-i\epsilon V(x_j)/\hbar} \langle x_{j+1} | e^{-i\epsilon p^2/2m\hbar} | x_j \rangle. \end{aligned}$$

Now $\langle x_{j+1}|e^{-i\epsilon p^2/2m\hbar}|x_j\rangle$ is just the propagator for a free particle, which we'd already calculated earlier. We'll compute it again, for posterity.

$$\begin{aligned}\langle x_{j+1}|e^{-i\epsilon p^2/2m\hbar}|x_j\rangle &= \int_{-\infty}^{\infty} dp_j\ \langle x_{j+1}|p_j\rangle\langle p_j|e^{-i\epsilon p^2/2m\hbar}|x_j\rangle \\ &= \int_{-\infty}^{\infty} dp_j\ e^{ip_j(x_{j+1}-x_j)/\hbar}e^{-i\epsilon p_j^2/2m\hbar} \\ &= \sqrt{\frac{m}{2\pi i\hbar\epsilon}}e^{im(x_{j+1}-x_j)^2/2\hbar\epsilon}.\end{aligned}$$

Therefore:

$$\begin{aligned}\langle x_{j+1}|e^{-i\epsilon(T+V)/\hbar}|x_j\rangle &= e^{-i\epsilon V(x_j)/\hbar}\langle x_{j+1}|e^{-i\epsilon p^2/2m\hbar}|x_j\rangle \\ &= \sqrt{\frac{m}{2\pi i\hbar\epsilon}}\exp\left(i\left[\frac{m}{2\hbar\epsilon}(x_{j+1}-x_j)^2 - \frac{\epsilon}{\hbar}V(x_j)\right]\right).\end{aligned}$$

In particular,

$$K(x_n,t_N;x_0,t_0) = \lim_{\epsilon\to 0,N\to\infty,N\epsilon=t_N-t_0}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{\frac{N-1}{2}}\int\prod_{k=1}^{N-1}dx_k\ \exp\left(i\left[\sum_{k=0}^{N-1}\frac{m}{2\epsilon\hbar}(x_{k+1}-x_k)^2 - \epsilon\frac{V(x_k)}{\hbar}\right]\right).$$

As $\epsilon\to 0$, we can view $x_{k+1}-x_k$ as $\epsilon\cdot\frac{dx}{dt}$. Therefore, the integrand becomes

$$\begin{aligned}\exp\left(\frac{i}{\hbar}\left[\epsilon\sum_{k=0}^{N-1}\frac{m}{2}\dot{x}_k^2 - V(x_k)\right]\right) &= \exp\left(\frac{i}{\hbar}\int_{t_0}^{t_N}dt\ \left[\frac{m}{2}\left(\frac{dx}{dt}\right)^2 - V(x)\right]\right) \\ &= \exp\left(\frac{i}{\hbar}\int_{t_0}^{t_N}dt\ L\right),\end{aligned}$$

as $\epsilon\to 0$, and this is what you use to *define* $\int[\mathcal{D}x]\ e^{iS/\hbar}$.

11. More on path integrals and quantum particles in potentials

Last time, we defined the path integral as

$$\begin{aligned}K(x_N,t_N;x_0,t_0) &= \lim_{\epsilon\to 0,N\to\infty,N\epsilon=t_N-t_0}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{\frac{N}{2}}\int\prod_{k=1}^{N-1}dx_k\ \exp\left(i\epsilon\left[\sum_{k=0}^{N-1}\frac{m}{2\hbar}\left(\frac{x_{k+1}-x_k}{\epsilon}\right)^2 - \frac{V(x_k)}{2\hbar}\right]\right) \\ &=: \int[\mathcal{D}x]e^{iS[x(t)]/\hbar}.\end{aligned}$$

Note that we made a typo last time in the prefactor: the exponent of $\frac{m}{2\pi i\hbar\epsilon}$ should be $N/2$, not $(N-1)/2$; this depends on whether you're breaking up $(x_0,t_0)\to(x_N,t_N)$ into $N+1$ steps as opposed to N steps.

Let us check this by using this to recalculate the free particle propagator. We choose $N=2^n$. Define

$$K_N = \left(\frac{mN}{2\pi i\hbar t}\right)^{\frac{N}{2}}\int\prod_{k=1}^{N-1}dx_k\ \exp\left(\frac{imN}{2\hbar t}\left[\sum_{k=0}^{N-1}(x_{k+1}-x_k)^2\right]\right).$$

This is the integral we want to calculate in the limit as $N\to\infty$. The key point is that the exponent is quadratic in the exponent x_j. We'll use some tricks to reduce the amount of work we have to do. The exponetn is

$$\frac{imN}{2\hbar t}\left[x_0^2 + 2x_1^2 + 2x_2^2 + \cdots + 2x_{N-1}^2 + 2x_N^2 - 2x_0x_1 - \cdots - 2x_{N-1}x_N\right].$$

We do all of the odd-labeled integrals first (i.e., the integrals over $x_1, x_3, \cdots$):

$$\begin{aligned}
K_N &= \int \prod_{j \text{ even}} dx_j \prod_{k \text{ odd}} dx_k \left(\frac{m2^n}{2\pi i\hbar t}\right)^{\frac{2^n}{2}} \exp\left(\frac{im2^n}{2\hbar t}\left[\sum_k 2x_k^2 + x_{k-1}^2 + x_{k+1}^2 - 2x_k(x_{k-1}+x_{k+1})\right]\right) \\
&= \int \prod_{j \text{ even}} dx_j \prod_{k \text{ odd}} dx_k \left(\frac{m2^n}{2\pi i\hbar t}\right)^{\frac{2^n-1}{2}} \exp\left(\frac{im2^n}{2\hbar t}\left[\sum_k 2\left(x_k - \frac{x_{k+1}+x_{k-1}}{2}\right)^2 + \frac{x_{k+1}^2+x_{k-1}^2}{2} - x_{k+1}x_{k-1}\right]\right) \\
&= \left(\frac{m2^n}{2\pi i\hbar t}\right)^{\frac{2^n-1}{2}} \int \prod_{\ell=1}^{2^{n-1}-1} dx_{2\ell} \exp\left(\frac{im2^{n-1}}{2\hbar t} \sum_{\ell=0}^{2^{n-1}-1} (x_{2\ell+2} - x_{2\ell})^2\right) \\
&= K_{N/2}.
\end{aligned}$$

It follows that $K_N = K_1$. But you can easily check that

$$K_1 = \sqrt{\frac{m}{2\pi i\hbar t}} \exp\left(\frac{im}{2\hbar t}(x_N - x_0)^2\right).$$

I want to make a few comments. The path integral is an alternate way to study quantum mechanics, which is completely equivalent to the standard formulations. The emphasis in the Feynman formulation is on the action, rather than the Hamiltonian. This has some advantages and some disadvantages. As we saw today, we need to be careful with the integration measure. The propagator is, however, a fairly useless way to calculate stuff, but it is a great intuitive crutch. Once you go to systems with large (e.g., infinitely many) quantum degrees of freedom (e.g., QFT, which is a special case of quantum many-body physics), the path integral becomes an extremely useful way to think about the system.

There are a few technical details which are worth paying attention to. A key step in obtaining the path integral was the calculation of the matrix element

$$\langle x_{j+1}|e^{-i\epsilon H/\hbar}|x_j\rangle = \exp\left(-\frac{i}{\hbar}\frac{\epsilon(x_{j+1}-x_j)^2}{2m} - \frac{i\epsilon V(x_j)}{\hbar}\right).$$

A more symmetric replacement would be to use $V((x_j + x_{j+1})/2)$, which is the same to $O(\epsilon)$.

Second, a particle moving in a magnetic field specified by a vector potential $\vec{A}$ has Lagrangian

$$L = \frac{1}{2}m\dot{x}^2 + e\vec{A}\cdot\dot{x} - V(\vec{x});$$

see your homework. The correct prescription is to evaluate $\vec{A}$ as $\vec{A}((x_j + x_{j+1})/2)$ in the discrete time path integral.

Let's do some physics. We first illustrate the connection to classical mechanics, via the "stationary phase approximation". Consider $\int dx e^{i\lambda f(x)}$ with λ large. Let $f(x)$ have an extremum at $x = x_0$. We expect that the integral is dominated by x near x_0, since $f(x)$ varies very little near x_0, so the phase oscillates slowly as x changes. Near x_0, we have

$$f(x) = f(x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \cdots$$

Therefore,

$$\int dx\ e^{i\lambda f(x)} = e^{i\lambda f(x_0)} \int d(\delta x)\ e^{i\lambda f''(x_0)(\delta x)^2/2}.$$

The tails of this integral contribute very little, so we can replace the limits of this integral by $+\infty$ and $-\infty$. We conclude that

$$\int dx\ e^{i\lambda f(x)} = e^{i\lambda f(x_0)} \sqrt{\frac{2\pi i}{\lambda f''(x_0)}} \left(1 + o(1/\lambda^2)\right).$$

Similarly, $\int [\mathcal{D}x] e^{iS/\hbar}$ is dominated by $x(t) = x_{cl}(t)$ when $\hbar \to 0$ where $\frac{\delta S}{\delta x_{cl}} = 0$. As $\hbar \to 0$, the limit is $K(x,t;x',t')$ is approximately proportional to $e^{iS[x_{cl}(t)]/\hbar}$.

Let us now discuss potentials and electromagnetic fields. In classical mechanics, if you shift $V(x)$ to $V(x) + V_0$, with V_0 indepndent of x. This makes no change to measurable quantities in physics, both in classical and quantum mechanics. In quantum mechanics, $|\psi(t)\rangle \to e^{-iV_0 t/\hbar}|\psi(t)\rangle$, i.e., this shift just corresponds to a phase rotation. The overall phase is not an observable, but if we change V in just one region of space, it leads to observable effects.

Consider a source for particles placed in position A, and a detector in position B. We restrict the path so that it must go through either of two paths such that both of these paths go through metallic cages. Let's assume that the electrostatic potential is V_2 in the top cage, and V_1 in the bottom cage. Within each cage, the potential is a constant, so there is no force on the particles within either cage. The fact that the potential is different between the top and bottom means that there is a phase difference. The phase difference is

$$\phi_2 - \phi_1 = \frac{1}{\hbar} \int_{t_i}^{t_f} dt\ (V_2 - V_1).$$

Let a be the amplitude to go from A to B in time $t_f - t_i$ through the bottom cage. Then, the amplitude to go through the top is $ae^{-i(\phi_2 - \phi_1)}$. In the path integral approach, we're summing over all paths, so the full amplitude is $a(1 + e^{-i(\phi_2 - \phi_1)})$. The probability is therefore $|a|^2(2 + 2\cos(\phi_2 - \phi_1))$. This is an oscillatory function, depending on $\phi_2 - \phi_1$, which in turn depends on $V_2 - V_1$. If you average this probability over a time interval and you set $\hbar \to 0$, this interference vanishes. In other words, the probability in quantum mechanics knows about the potential difference through this interference term.

Next time, we'll talk about the Aharanov-Bohm effect.

12. Quantum mechanics in EM fields and the Aharanov-Bohm effect

Recall

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi$$
$$\vec{B} = \vec{\nabla} \times \vec{A},$$

where $\vec{A}$ is the vector potential. Another way to rewrite this is to use the Maxwell tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, with $A_\mu = (\phi, \vec{A})$, then

$$F_{0i} = F_{-i0} = -E_i,\ F_{ij} = \epsilon_{ijk} B_k,$$

where ϵ_{ijk} is the fully antisymmetric tensor with ϵ_{123}.

In general, a gauge transformation $A_\mu \to A_\mu + \partial_\mu \Lambda(x,t)$ leaves $F_{\mu\nu}$ invariant. The Lagragian of a charged particle in an electromagnetic field is

$$L = \frac{m}{2}\dot{\vec{x}}^2 + \frac{e}{c}\vec{A} \cdot \dot{\vec{x}} - e\phi,$$

and this gives the Lorentz force law, as you've checked on homework.

Given a Lagrangian, the canonical momentum is defined by

$$\vec{p} = \frac{\partial L}{\partial \dot{x}} = m\dot{\vec{x}} + \frac{e}{c}\vec{A}.$$

The Hamiltonian is

$$H = \vec{p}\dot{x} - L = \frac{m}{2}\dot{\vec{x}}^2 + e\phi = \frac{1}{2m}\left(\vec{p} - \frac{e}{c}\vec{A}\right)^2 + e\phi.$$

Note that the canonical momentum is not gauge invariant. The canonical momentum has a special place in the story, since $\vec{p}$ is the object that generated translations. I emphasized earlier

that this is one way to *define* momentum in quantum mechanics. Note that the canonical momentum is different from $m\dot{x}$ when $\vec{A} \neq 0$.

Let us quantize the Hamiltonian. Now, x_i and p_j are operators satisfying $[x_i, p_j] = i\hbar\delta_{ij}$ and

$$H = \frac{p^2}{2m} - \frac{e}{2mc}(\vec{A}\cdot\vec{p} + \vec{p}\cdot\vec{A}) + \frac{e^2}{2mc^2}\vec{A}^2 + e\phi.$$

The Schrödinger equation is therefore

$$i\hbar\frac{\partial\psi}{\partial t} = \left[\frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2 + e\phi\right]\psi.$$

How does the gauge transformation rear its head in the quantum treatment? Let $\vec{A}' = \vec{A} + \vec{\nabla}\Lambda$ and $\phi' = \phi - \frac{1}{c}\partial_t A$. Then the Schrödinger equation becomes

$$\left(i\hbar\frac{\partial}{\partial t} - e\phi' - \frac{e}{c}\frac{\partial\Lambda}{\partial t}\right)\psi = \frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A} + \vec{\nabla}\Lambda\right)^2\psi.$$

These are different equations. Instead, if we let $\psi = \psi' e^{-\frac{ie}{\hbar c}\Lambda(x,t)}$, then ψ' satisfies

$$\left(i\hbar\frac{\partial}{\partial t} - e\phi'\right)\psi' = \frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}'\right)^2\psi.$$

Thus, in quantum mechanics, a gauge transformation looks like $\vec{A}' = \vec{A} + \vec{\nabla}\Lambda$, $\phi' = \phi - \frac{1}{c}\partial_t A$, and $\psi' = \psi e^{\frac{ie}{\hbar c}\Lambda(x,t)}$. In particular, the wavefunction is not gauge invariant. The statement is that $\boxed{\text{all physical observables remain invariant under this transformation}}$.

Let us show how this stuff is manifested in the Aharanov-Bohm effect. Consider a hollow cylindrical shell with a solenoid through the middle, so that there is a magnetic flux $\Phi = \int_{\text{core}} BdA$ through the core. THe charnged particles are constrained to move within the shell, and so the magnetic field is zero where the particles move. There is therefore no Lorentz force on the particle's motion, so classically the magnetic field has no effect whatsoever.

What about in quantum mechanics? Recall that although $\vec{B}$ is zero inside the shell, $\vec{A} \neq 0$. The reason is that, if you take a loop C around the center, then

$$\oint_C \vec{A}\, d\vec{\ell} = \int d\vec{S}\cdot\vec{\nabla}\times\vec{A} = \Phi \neq 0.$$

Choose $\vec{A} = A_\phi e_{\hat{\phi}}$, with A_ϕ independent of ϕ. Assume that C is a circular loop of radius R, so that

$$\oint_C \vec{A}\, d\vec{\ell} = A_\phi(2\pi R) = \Phi,$$

i.e., that $A_\phi = \frac{\Phi}{2\pi R}$. The full vector potential is therefore $\vec{A} = \frac{\Phi}{2\pi R}e_{\hat{\phi}}$. You can check that the curl of this vector potential is zero inside the shell, because the magnetic field is zero inside the shell.

We will address the situation in quantum mechanics using path integrals. The propagator is

$$K(x,t;x',t') = \int [\mathcal{D}x] e^{\frac{i}{\hbar}S[x(t)]}.$$

The action is

$$S[x(t)] = S_0 + \frac{e}{c}\int d\vec{x}\cdot\vec{A},$$

where S_0 is the action without the B-field. Therefore, the amplitude for a path from x to x' is multiplied by $\exp\left(\frac{ie}{\hbar c}\int_x^{x'} d\vec{x}\cdot\vec{A}(x)\right)$. Pick two antipodal points $1, 2$ on the outer boundary

of the shell. There are two possible things a path could do which go around the hollow center, namely go around the left or the right. For all paths P_L going through the left, the amplitude is

$$a(P_L) = a_0 \exp\left(\frac{ie}{\hbar c}\int_{P_L} d\vec{x}\cdot\vec{A}(x)\right).$$

Note that for two different paths P_L and P_L', we the relative phase is

$$\exp\left(\frac{ie}{\hbar c}\left[\int_{P_L} d\vec{x}\cdot\vec{A}(x) - \int_{P_L'} d\vec{x}\cdot\vec{A}(x)\right]\right) = \exp\left(\frac{ie}{\hbar c}\int_{P_L\cup P_L'} d\vec{x}\cdot\vec{A}(x)\right) = 1,$$

since the loop $P_L \cup P_L'$ is a closed loop which does not go through the hollow center — since there is no magnetic field outside the hollow center, the integral must vanish. Similarly, we find that the relative phase is 1 for two paths P_R and P_R'. Therefore

$$a(P_R) = a_0 \exp\left(\frac{ie}{\hbar c}\int_{P_R} d\vec{x}\cdot\vec{A}(x)\right).$$

We have assumed that a_0 is the same for both P_R and P_L, but you can add enough symmetry to guarantee this.

The total amplitude to go from 1 to 2 is therefore

$$\begin{aligned}
&a_0 \exp\left(\frac{ie}{\hbar c}\int_{P_L} d\vec{x}\cdot\vec{A}(x)\right) + a_0 \exp\left(\frac{ie}{\hbar c}\int_{P_R} d\vec{x}\cdot\vec{A}(x)\right)\\
&= a_0 e^{\frac{ie}{\hbar c}\int_{P_R} d\vec{x}\cdot\vec{A}(x)}\left(1 + \exp\left(\frac{ie}{\hbar c}\int_{P_R} d\vec{x}\cdot\vec{A}(x) - \frac{ie}{\hbar c}\int_{P_L} d\vec{x}\cdot\vec{A}(x)\right)\right)\\
&= a_0 e^{\frac{ie}{\hbar c}\int_{P_R} d\vec{x}\cdot\vec{A}(x)}\left(1 + e^{\frac{ie}{\hbar c}\int_C d\vec{x}\cdot\vec{A}(x)}\right),
\end{aligned}$$

where $C = P_R \cup (-P_L)$ encloses the core. The total amplitude is therefore $a_0 e^{\frac{ie}{\hbar c}\int_{P_R} d\vec{x}\cdot\vec{A}(x)}\left(1 + e^{\frac{ie}{\hbar c}\Phi}\right)$, and thus

$$\text{Prob}(1\to 2) = |a_0|^2\left(2 + 2\cos\left(\frac{e\Phi}{\hbar c}\right)\right),$$

which is oscillatory as a function of Φ. These are known as "Aharanov-Bohm oscillations". Next week, you'll deduce this using the Schrödinger equation.

13. Magnetic monopoles

You should review the simple harmonic oscillator this week (raising and lowering operators, etc).

Although we haven't yet seen magnetic monopoles, every theoretical physicist believes that they exist. Whether or not they exist, they are very useful for thinking about physics that does exist. Suppose such monopoles exist. Then, Maxwell's equation $\vec{\nabla}\cdot\vec{B} = 0$ must be modified to $\vec{\nabla}\cdot\vec{B} = 4\pi\rho_m$, where ρ_m is the density of the magnetic charge. Consider a single magnetic charge of strength g. This is a "magnetic monopole". The magnetic field around the monopole is $\vec{B} = \frac{g}{r^2}\vec{e}_r$; this is just Gauss' law but for magnetic fields.

The fun begins when we think about this in quantum mechanics. Consider the quantum mechanics of ac harged particle in this $\vec{B}$-field. We need the vector potential to emulate our previous discussion; but since $\vec{\nabla}\cdot\vec{B} \neq 0$, there is no representation of $\vec{B}$ as $\vec{\nabla}\times\vec{A}$. One could give up at this point, but Dirac didn't. He studied this problem seriously, and reached some amazing conclusions. He studied a bunch of weird things. Must've been smoking something.

We'll take a more modern perspective on Dirac's work. Notice that for $r > 0$, we have $\vec{\nabla}\cdot\vec{B} = 0$ is satisfied, because the monopole is sitting at the origin. This implies that if we

exclude the point at the origin, then locally in each region of space we can find a vector potential $\vec{A}$ such that $\vec{B} = \vec{\nabla} \times \vec{A}$.

Let's think about a sphere of radius r. Then, $\int_{\text{sphere}} \vec{B} \cdot d\vec{S} = 4\pi g$; this is the integral form of Gauss' law. Everywhere on the surface of the sphere, we can write $\vec{B} = \vec{\nabla} \times \vec{A}$ as the curl of $\vec{A}$, so that

$$\int \vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) dV = \int_{\text{sphere}} \vec{B} \cdot d\vec{S} = 4\pi g.$$

The first integral would be zero if we could write $\vec{B}$ as the curl of $\vec{A}$ everywhere.

Let us explicitly demonstrate that $\vec{A}$ is defined locally but not globally, and then deal with it. Let's stick to the sphere of radius r, and consider a circle (with some orientation) at a polar angle of θ. Let's ask about the magnetic flux through the upper cap on the sphere. The magnetic flux is

$$g(2\pi) \int_0^\theta d\theta' \ \sin(\theta') = 2\pi g(1 - \cos\theta).$$

Suppose there is a vector potential everywhere on the circle C. What is the vector potential whose line integral is this flux? We might as well choose $\vec{A}$ to be entirely along the ϕ direction, so $\vec{A} = A_\phi \hat{e}_\phi$. Then

$$\oint_C \vec{A} \cdot d\vec{\ell} = \int_{\text{surface bounded by } C} \vec{B} \cdot d\vec{S} = 2\pi g(1 - \cos\theta).$$

Let us make a guess: take A_ϕ independent of ϕ; then,

$$\oint_C \vec{A} \cdot d\vec{\ell} = 2\pi r A_\phi \sin(\theta).$$

This is because the circumference of the circle is $2\pi r \sin(\theta)$. It follows that

$$A_\phi = \frac{g(1 - \cos\theta)}{r \sin\theta}.$$

You should verify explicitly that $\vec{B} = \vec{\nabla} \times \vec{A}$. This thing clearly blows up as $r \to 0$, but we said we're excluding the point at the origin. We'll stick to $r = 0$. When $\sin\theta \to 0$, this again blows up. This happens at $\theta = 0, \pi$. The case $\theta = 0$ is not problematic, since $\cos\theta \to 1$, so the numerator also goes to zero. The numerator goes to zero faster, though — so $\theta = 0$ is not a problem since $A_\phi \to 0$ in that case.

But if $\theta = \pi$, there is a problem: $\sin\theta = 0$ and $\cos\theta = -1$. The denominator blows up. Since $\theta = \pi$ is the south pole, A_ϕ is only defined in any upper cap of the sphere. To this end, let us define

$$A_\phi^+ = \frac{g(1 - \cos\theta)}{r \sin\theta}.$$

If you look at the flux through a *lower* cap, then you'll find

$$A_\phi^- = -\frac{g(1 + \cos\theta)}{r \sin\theta}.$$

If you examine this function, it's singular at $r = 0$. But at $\theta = \pi$, which was problematic in the upper cap, we no longer have a problem. Similarly, at $\theta = 0$, which was not problematic in the upper cap, we now have a problem. The situation is the opposite of what happened in the case of the upper cap.

If we think about a generic point which is neither the south or the north pole, then we have two vector potentials with the same magnetic field. Therefore the two vector potentials must differ by a gradient:

$$A_\phi^+ - A_\phi^- = \frac{2g}{r \sin\theta},$$

so that

$$\vec{A}^+ - \vec{A}^- = \frac{2g}{r\sin\theta}\hat{e}_\phi = 2g\nabla\phi.$$

As expected, $\vec{A}^+$ and $\vec{A}^-$ are therefore gauge equivalent. Therefore,

$$\vec{A} = \begin{cases} \vec{A}^+ = \frac{g(1-\cos\theta)}{r\sin\theta}\hat{e}_\phi & \text{for } 0 \le \theta < \frac{\pi}{2} + \epsilon \\ \vec{A}^- = -\frac{g(1+\cos\theta)}{r\sin\theta}\hat{e}_\phi & \text{for } \frac{\pi}{2} - \epsilon < \theta \le \pi, \end{cases}$$

with $0 < \epsilon < \pi/2$. This function is well-defined everywhere, up to a gauge transformation.

We can now do quantum mechanics. We can imagine solving the Schrödinger equation. In the region of overlap, the resulting wavefunctions must be related by a gauge transformation. For $0 \le \theta < \pi/2 + \epsilon$, let the wavefunction be $\psi^{(+)}(r, \theta, \phi)$ and for $\pi/2 - \epsilon < \theta \le \pi$, let the wavefunction be $\psi^{(-)}(r, \theta, \phi)$. Last time, we found that

$$\psi^+ = \psi^- e^{2ieg\phi/\hbar c},$$

where e is the charge of the particle. The key point is that when you go all the way around the equator (i.e., $\phi \to \phi + 2\pi$), the wavefunction is single valued: the value cannot change. In order for this to occur, we need

$$\frac{2eg\phi}{\hbar c} = n \in \mathbf{Z}.$$

This implies that

$$g = \frac{n\hbar c}{2e}.$$

One way to think about this is that quantum mechanics requires that magnetic charge g be quantized in units of $\frac{\hbar c}{2e}$. Dirac also used this as an explanation that electric charge is quantized, with the assumption that a magnetic monopole exists.

If you have free time, I would encourage you to do the following computation. Have fun with it! Take a classical magnetic monopole of strength g, and a particle of charge e displaced from one another by a distance d along $\hat{z}$. In space, there's both an electric and a magnetic field. Once there's an $\vec{E}$ and a $\vec{B}$, there's a Poynting vector. In general, there's energy stored in an electromagnetic field. You also have angular momentum. Calculate the angular momentum associated to this configuration of the magnetic monopole and the electric field. This'll be something pointing along $\hat{z}$. You'll find that the angular momentum is independent of d, and only depends on e and g. If we require the total angular momentum we found to be an integer multiple of $\hbar/2$, then we recover the Dirac quantization condition! This us very interesting. Spin can be quantized in half integer multiples of $\hbar$. The bound state of the minimal monopole has internal angular momentum of spin 1/2. This gives us a "mechanical" model for spin. You're putting together two bosons to get a fermion.

I want to now talk about a charged particle in a uniform magnetic field. I'll assume you're completely familiar with the solution of the simple harmonic oscillator. This is a famous problem that was first solved by Landau. The answer goes by the name of "Landau levels". He solved this at an age where he was probably younger than everyone else in this room.

14. Landau levels and composite systems

Consider a particle of chrge e in a magnetic field $\vec{B} = B\hat{z}$ in three dimensions, so that the Hamiltonian is $H = \frac{(\vec{p} - \frac{e}{c}\vec{A})^2}{2m}$. Since only $B_z \neq 0$, we can always choose $A_z = 0$, and A_x and A_y independent of z. The Hamiltonian then becomes

$$H = \frac{p_z^2}{2m} + \frac{\Pi_x^2 + \Pi_y^2}{2m},$$

with $\Pi_{x,y} = p_{x,y} - \frac{e}{c}A_{x,y}$. This is the "kinematic momentum".

Note that $[p_z, H] = 0$, so we can label the energy eigenstates by p_z. Define $H_{2d} = \frac{\Pi_x^2+\Pi_y^2}{2m}$. Now, we notice that

$$[\Pi_x, \Pi_y] = [-i\hbar\partial_x - \frac{e}{c}A_x, -i\hbar\partial_y - \frac{e}{c}A_y] = \frac{i\hbar e}{c}(\partial_x A - \partial_y A_x)\mathbf{1}.$$

Therefore, $[\Pi_x, \Pi_y] = \frac{i\hbar e}{c}B$; in other words, Π_x and Π_y are canonically conjugate. To make this precise, define $X = \frac{c\Pi_x}{eB}$ and $P = \Pi_y$. Then, $[X, P] = i\hbar$, so

$$H_{2d} = \frac{p^2}{2m} + \frac{1}{2m}\left(\frac{eB}{c}\right)^2 X^2.$$

This the Hamiltonian of a simple harmonic oscillator with a frequency $\omega_c = \frac{eB}{mc}$. This is called the "cyclotron frequency".

Recall that classically, with a particle in a perpendicular B-field in two dimensions, we have $\frac{mv^2}{R} = \frac{evB}{c}$, if it is traveling on adefine $X = \frac{c\Pi_x}{eB}$ and $P = \Pi_y$. Then, $[X, P] = i\hbar$, so

$$H_{2d} = \frac{p^2}{2m} + \frac{1}{2m}\left(\frac{eB}{c}\right)^2 X^2.$$

This the Hamiltonian of a simple harmonic oscillator with a frequency $\omega_c = \frac{eB}{mc}$. This is called the "cyclotron frequency".

Recall that classically, with a particle in a perpendicular B-field in two dimensions, we have $\frac{mv^2}{R} = \frac{evB}{c}$, if it is traveling on adefine $X = \frac{c\Pi_x}{eB}$ and $P = \Pi_y$. Then, $[X, P] = i\hbar$, so

$$H_{2d} = \frac{p^2}{2m} + \frac{1}{2m}\left(\frac{eB}{c}\right)^2 X^2.$$

This the Hamiltonian of a simple harmonic oscillator with a frequency $\omega_c = \frac{eB}{mc}$. This is called the "cyclotron frequency".

Recall that classically, with a particle in a perpendicular B-field in two dimensions, the particle will travel in a circular motion with redius R. We have $\frac{mv^2}{R} = \frac{evB}{c}$, so $R = \frac{mvc}{eB}$. Note that v is constant, since $F = \frac{e}{c}\vec{v} \times \vec{B}$ is perpendicular to $\vec{v}$. The time period of the orbit is $\frac{2\pi R}{v} = \frac{2\pi}{\omega_c}$.

Back to the quantum case: using the fact that

$$H_{2d} = \frac{p^2}{2m} + \frac{1}{2m}\left(\frac{eB}{c}\right)^2 X^2,$$

we find that the energy levels are $E_n^{(2d)} = \left(n + \frac{1}{2}\right)\hbar\omega_c$. The full 3d energy states are therefore $E_{n,p_z} = \frac{p_z^2}{2m} + \left(n + \frac{1}{2}\right)\hbar\omega_c$. We are, however, missing the degeneracy of the energy eigenvalues. This is a class of problems where the spectrum of energy levels is easy to determine, but the degeneracies are nontrivial. For example, there are whole quantum theories where the Hamiltonian is zero, but the problem of finding degeneracies is still very interesting.

The spectrum is highly degenerate. For a fixed n, we will see that there are an infinite number of eigenstates. What is the physical origin of this degeneracy? Define new coordinates R_x and R_y through

$$R_x = x + \frac{c}{eB}\Pi_y, \; R_y = y - \frac{c}{eB}\Pi_x.$$

These have dimensions of length. Then, you can check that

$$[R_x, R_y] = -\frac{i\hbar c}{eB}.$$

But then, you also find that

$$[R_x, \Pi_x] = i\hbar - \frac{c}{eB}\left(\frac{i\hbar eB}{c}\right) = 0 = [R_y, \Pi_y].$$

It follows that $[R_i, H] = 0$ but $[R_x, R_y] \neq 0$. This implies that there must be a generacy, as you proved in a homework assignment.

Classically, these are known as "guiding center coordinates". This measures the coordinates on the orbit relative to the center of the orbit, because we are subtracting from x and y the cyclotron radius. The degeneracy of the orbit arises since the cyclotron orbit can be placed anywhere in space: the physics of the orbit should not depend on the location of the orbit.

Pick the following particular gauge: $A_y = Bx$ and $A_x = 0$; this is called the "Landau gauge". Then,

$$H_{2d} = \frac{p_x^2}{2m} + \frac{1}{2m}\left(p_y - \frac{eB}{c}x\right)^2 = \frac{p_x^2}{2m} + \frac{1}{2}m\omega_c^2\left(x - \frac{cp_y}{eB}\right)^2.$$

In this gauge, $[p_y, H_{2d}] = 0$, so we can label energy states by their p_y eigenvalue. We again get a 1d simple harmonic oscillator of frequency ω_c, so that the energy spectrum is $E_n^{(2d)} = \left(n + \frac{1}{2}\right)\hbar\omega_c$.

The wavefunctions, in this gauge, are therefore

$$\psi_{n,p_y}(x, y) = e^{ip_y y/\hbar}\phi_n\left(x - \frac{cp_y}{eB}\right).$$

Here, ϕ_n is the wavefunction of the nth eigenstate of the simple harmonic oscillator. It follows that states with different p_y are degenerate. So, we need to understand how many distinct values of p_y exist for a system with given area.

Consider a sample of size $L_x \times L_y$ with periodic boundary conditions. Then, $y + L_y$ is identified with y, so that $e^{ip_y L_y/\hbar} = 1$, i.e., $p_y = \frac{2\pi n\hbar}{L_y}$, with n an integer. Each ψ_{n,p_y} is therefore centered at $x = \frac{c}{eB}\frac{2\pi n\hbar}{L_y}$. But if my sample has finite size, I cannot have infinitely many such states. For a sample of length L_x, therefore, the number of indepndent states is

$$g = \frac{L_x}{\frac{c}{eB}\frac{2\pi\hbar}{L_y}} = \frac{L_x L_y eB}{2\pi\hbar c} = \frac{eB}{hc}A,$$

where A is the area of the sample. Now, $\frac{hc}{e} = \Phi_0$ is the "elementary quantum flux", so g is the total magnetic flux divided by the elmentary quantum flux. Thus, it is the number of flux quanta piercing the sample.

15. Composite systems; pure versus mixed states, density matrices; entanglement

We will now begin to study composite systems, which are systems made up of composites of two subsystems. Let's say the system is $A + B$, and it's made up of two subsystems, A and B. Let $\mathcal{H}_A$ and $\mathcal{H}_B$ denote the Hilbert spaces of A and B respectively. Then, the Hilbert space of $A + B$ is the tensor product of $\mathcal{H}_A$ and $\mathcal{H}_B$. What does this mean? We will only discuss this for tensor products of vector spaces.

Consider a basis $\{|\phi_i\rangle\}$ of $\mathcal{H}_A$ and $\{|\chi_i\rangle\}$ of $\mathcal{H}_B$. Define new vectors $|\phi_i\rangle \otimes |\chi_j\rangle$; these are elements of the "tensor product space" $\mathcal{H}_A \otimes \mathcal{H}_B$. These vectors are defined to be a basis of $\mathcal{H}_A \otimes \mathcal{H}_B$. In other words, every state of $A + B$, which lives in $\mathcal{H}_A \otimes \mathcal{H}_B$, can be written as a linear combination $\sum_{ij}\psi_{ij}|\phi_i\rangle \otimes |\chi_j\rangle$, where the ψ_{ij} are some complex numbers. Soon enough, we will only write $|\phi_i, \chi_j\rangle$ for $|\phi_i\rangle \otimes |\chi_j\rangle$.

There are operators that act on $\mathcal{H}_{A+B}$. For example, there are operators which act on A alone: $O_A \otimes 1_B$. More generally, we can consider operators of the form $O' = O_A \otimes O_B$.

For example, consider a system of two spin 1/2 particles A and B. Some possible states include:

$$|\psi^{(1)}_{A+B}\rangle = |\uparrow_A\rangle \otimes |\uparrow_B\rangle,\ |\psi^{(2)}_{A+B}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_A\rangle \otimes |\downarrow_B\rangle + |\downarrow_A\rangle \otimes |\uparrow_B\rangle).$$

Note that the Hilbert space is four-dimensional.

We will now talk about the infamous theory of quantum entanglement. Two parts A and B of a quantum system may be "entangled". For example, the spin state $|\uparrow_A\rangle \otimes |\uparrow_B\rangle$ is unentangled. This means that we can specify the quantum state of every subsystem without having to reference the larger system $A + B$, i.e., that each spin by itself is in a well-defined quantum state.

We can contrast this with the state $\frac{1}{\sqrt{2}}(|\uparrow_A\rangle \otimes |\downarrow_B\rangle \pm |\downarrow_A\rangle \otimes |\uparrow_B\rangle)$ we saw above. This is an example of an entangled state: each spin by itself is not in a definite quantum state, although the full system is.

A few more examples: the state $|\uparrow,\downarrow\rangle$ is unentangled. The state $\frac{1}{\sqrt{2}}(|\uparrow,\downarrow\rangle + |\downarrow,\downarrow\rangle)$ is unentangled, since it is $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \otimes |\downarrow\rangle$. Entanglement is the relationship of parts of a subsystem to each other and to the whole system. For example, the state $|\psi_A\rangle \otimes |\psi_B\rangle$ factorizes as a product of parts, so it is unentangled. Entangled parts do not factorize as a product of parts.

How do we describe our part of the world without having to take into account the fact that it's entangled with some other part of the world? Given a composite wavefunction $|\psi_{A+B}\rangle$, define $\rho_A = \mathrm{Tr}_B\, |\psi_{A+B}\rangle\langle\psi_{A+B}|$. This is called the *reduced density matrix*, also sometimes known as the density operator. Why is this called that? Let O_A be an operator that acts only on A. Then, $\langle O_A\rangle = \langle\psi_{A+B}|O_A|\psi_{A+B}\rangle = \mathrm{Tr}_A(O_A\rho_A)$. This therefore contains the information you need in order to describe subsystem A in terms of the system $A + B$.

Let us do an example. Consider a two spin-1/2 system, with $|\psi_{A+B}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_A\downarrow_B\rangle - |\downarrow_A\uparrow_B\rangle)$. Then,

$$\rho_A = \mathrm{Tr}_B\, |\psi_{A+B}\rangle\langle\psi_{A+B}| = \langle\uparrow_B|\psi_{A+B}\rangle\langle\psi_{A+B}|\uparrow_B\rangle + \langle\downarrow_B|\psi_{A+B}\rangle\langle\psi_{A+B}|\downarrow_B\rangle.$$

We will illustrate this for the first term, which is

$$\frac{1}{2}(\langle\uparrow_B|(|\uparrow_A\rangle \otimes |\downarrow_B\rangle - |\downarrow_A\rangle \otimes |\uparrow_B\rangle)(\langle\uparrow_A| \otimes \langle\downarrow_B| - \langle\downarrow_A| \otimes \langle\uparrow_B|)|\psi_B\rangle.$$

It is not hard to see that the total sum is

$$\frac{1}{2}(|\downarrow_A\rangle\langle\downarrow_A| + |\uparrow_A\rangle\langle\uparrow_A|) = \frac{1}{2}1_A.$$

Another example: let $|\psi_{A+B}\rangle = |\uparrow_A\rangle \otimes |\downarrow_B\rangle$, so that

$$\rho_A = \langle\uparrow_B|\uparrow_A\downarrow_B\rangle\langle\uparrow_A\downarrow_B|\uparrow_B\rangle + \langle\downarrow_B|\uparrow_A\downarrow_B\rangle\langle\uparrow_A\downarrow_B|\downarrow_B\rangle = \langle\uparrow_A|\uparrow_A\rangle.$$

In general: choose a basis $|n_A\rangle$ for A; then, $\langle n_A|\rho_A|n_A\rangle = \rho_{nn}$ is the probability of finding A in the state n_A when the full system is in the state ψ_{A+B}. To see this, we note that this probability is exactly

$$\sum_{m_B} |\langle n_A, m_B|\psi_{A+B}\rangle|^2 = \mathrm{Tr}_B\, |\langle n_A|\psi_{A+B}\rangle|^2 = \mathrm{Tr}_B\langle n_A|\psi_{A+B}\rangle\langle\psi_{A+B}|n_A\rangle = \langle n_A|\rho_A|n_A\rangle,$$

where $|m_B\rangle$ is a basis for B.

Clearly the density matrix must satisfy some obvious properties. It follows from the above discussion that the trace of ρ_A must be 1. We also see that $\rho_A = \rho_A^\dagger$ is obviously Hermitian. The diagonal matrix is also positive, i.e., for any $|\phi_A\rangle$, we have $\langle\phi_A|\rho_A|\phi_A\rangle \geq 0$. Since ρ_A is Hermitian, we can diagonalize ρ_A in an orthonormal basis. These imply that the eigenvalues are all real and nonnegative, and the sum of its eigenvalues (which is the trace) is 1.

If $|\psi_{A+B}\rangle = |\phi_A\rangle \otimes |\chi_B\rangle$, then it is easy to calculate that $\rho_A = |\phi_A\rangle\langle\phi_A|$. This is a projection operator onto the subspace generated by $|\phi_A\rangle$. It is easy to check that $\rho_A^2 = \rho_A$.

If on the other hand, A and B are entangled, then $\rho_A^2 \neq \rho_A$. The matrix ρ_A will therefore not be a projection. If we choose a basis for A, then $\rho_A = \sum_a p_a |a\rangle\langle a|$. Then, $\rho_A^2 = \sum_a p_a^2 < \sum_a p_a = 1$ as more than one p_a must be nonzero.

We now introduce some terminology. We've assumed so far that the full system $A+B$ is in a state represented by a ray in Hilbert space. Call this a "pure state". In general, a subsystem A will be in a state described by a density matrix ρ_A, and if ρ_A is such that $\rho_A^2 \neq \rho_A$, then A is said to be in a "mixed state". In this case, $\rho_A = \sum_a p_a |a\rangle\langle a|$, and then we say that A is an "incoherent" mixture of the states $|a\rangle$ (emphatically, this is not called a superposition). Then, p_a is the probability of the state occuring in an ensemble of pure states.

16. Schmidt decomposition and the "no-communication" theorem

We begin by discussing the Schmidt decomposition. Consider a general composite state $|\psi\rangle_{AB}$; this can be written as $\sum_{i,\mu} \psi_{i\mu} |i_A\rangle|\mu_B\rangle$, where $\{|i_A\rangle\}$ and $\{|\mu_B\rangle\}$ are chosen bases for $\mathcal{H}_A$ and $\mathcal{H}_B$. Define a new state $|\widetilde{i}_B\rangle = \sum_\mu \psi_{i\mu}|\mu_B\rangle$; notice that this depends on $|\psi\rangle_{AB}$. Then, our most general state in $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written as

$$|\psi\rangle_{AB} = \sum_i |i_A\rangle|\widetilde{i}_B\rangle.$$

Recall that the density matrix is defined as $\rho_A = \mathrm{Tr}_B(|\psi\rangle_{AB}\langle\psi|_{AB})$. This satisfies $\mathrm{Tr}(\rho_A) = 1$, $\rho_A = \rho_A^\dagger$, and all the eigenvalues of ρ_A are nonnegative. This is equivalent to saying that $\langle\phi_A|\rho_A|\phi_A\rangle \geq 0$ for all $|\phi_A\rangle$.

Let us therefore choose $\{|i_A\rangle\}$ to be the basis in which ρ_A is diagonal, so that $\rho_A = \sum_i p_i |i_A\rangle\langle i_A|$, where $p_i \geq 0$. It follows that

$$\rho_A = \mathrm{Tr}_B(|\psi\rangle_{AB}\langle\psi|_{AB}) = \mathrm{Tr}_B(\sum_{i,j} |i\rangle\langle j| \otimes |\widetilde{i}\rangle\langle\widetilde{j}|) = \sum_{i,j} \langle\widetilde{j}|\widetilde{i}\rangle |i_A\rangle\langle j_A|.$$

When we compare our previous expression for ρ_A to this expression, we learn that $\langle\widetilde{j}|\widetilde{i}\rangle = p_i\delta_{ij}$, where we are not summing over i. Although the $|\widetilde{i}\rangle$ do not form a basis, they are an orthogonal set of vectors. Now, if we define $|i'_B\rangle = \frac{1}{\sqrt{p_i}}|\widetilde{i}_B\rangle$, to make it an orthonormal set of vectors. We conclude that the most general state of $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written as

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i}|i_A\rangle|i'_B\rangle.$$

This is the Schmidt decomposition. It is important to remember that we made a very convenient choice of states.

Suppose that we originally used the bases $|a_A\rangle$ and $|\mu_B\rangle$. Then, $|i_A\rangle = \sum_a |a_A\rangle(U_A)_{ai}$ and $|i'_B\rangle = \sum_\mu |\mu_B\rangle(U_B)_{\mu i'}$, where U_B is not necessarily unitary. Then:

$$\psi_{a\mu} = \sum_{i,j} (U_A)_{ai}\sqrt{p_i\delta_{ij}}(U_B^T)_{j\mu}.$$

This is exactly matrix multiplication. In "matrix notation", this would be written as $\psi = U_A D_{\{\sqrt{p_i}\}} U_B^T$, where $D_{\{\sqrt{p_i}\}}$ is the diagonal matrix whose ith entry is $\sqrt{p_i}$. This is the singular value decomposition from linear algebra.

We now move on to Schmidt numbers. For any (pure) state $|\psi\rangle_{AB} = \sum_{i,\mu} \psi_{i\mu}|i_A\rangle|\mu_B\rangle$, we can associate a positive integer called the Schmidt number as the number of nonzero eigenvalues of ρ_A or ρ_B. (Note that $\rho_B = \sum_i p_i |i'_B\rangle\langle i'_B|$.) This number tells us whether a state is entangled or not. For an unentangled state, the Schmidt number is 1, and for an unentangled state, the Schmidt number is > 1.

Let's talk about quantum correlations. At this point I got distracted and stopped paying attention.

17. Symmetries in QFT

Symmetries are operations that leave the physics unchanged. Consider, for example, a free particle. Then, the Hamiltonian is independent of x, so we can translate freely by x without affecting the ohysics. Inversion is another symmetry: we send $x \mapsto -x, p \mapsto -p$. This preserves the condition $[x,p] = i\hbar$, and leaves the Hamiltonian invariant. Yet another example is time reversal, which is subtle in quantum mechanics. We will address this later. Time translation is another symmetry, as is Galilean symmetry — but this latter symmetry is *very* subtle, so we will not address it here.

Define symmetries as a linear transformation $U : |a\rangle \to |a'\rangle$ such that all measurement outcomes are preserved, i.e.:

$$|\langle b|U^\dagger U|a\rangle|^2 = |\langle b|a\rangle|^2.$$

Theorem 3 (Wigner)**.** *In order for the above equation to be true, U must be either unitary or anti-unitary.*

Recall that U is anti-unitary if $U|a\rangle = (U|a\rangle)^\dagger$. If we assume that U is linear (or anti-linear), then this admits an easy proof. This will be on your homework. The case when U is not necessarily linear is harder, so we will not address it here.

Consider a unitary symmetry defined by a unitary operator U. We will assume that the Hamiltonian is unchanged by this transformation (i.e., $H = U^\dagger H U$) — note that this fails, e.g., for Galilean invariance. The condition that $H = U^\dagger H U$ translates to $[U,H] = 0$. We will now specialize to "continuous symmetries". This means that the symmetry transformation can be continuously built up as a series of infinitesimal transformations starting from the identity. For example, translations and rotations. In contrast to these, there are discrete symmetries which cannot be built up in this way: an example is parity.

Let us focus on continuous symmetries. Consider an infinitesimal transformation $U = 1 - \frac{i\epsilon}{\hbar} G + O(\epsilon^2)$. We proved a while back that if U is unitary, then G must be Hermitian. If $[U,H] = 0$, then $[G,H] = 0$. The Heisenberg equation of motion implies that $dG/dt = 0$. In other words, G is independent of time. We conclude that for every continuous symmetry in quantum emchanics, there is a corresponding conserved quantity. Similarly, if some observable G is conserved, then $U(\theta) = e^{i\theta G}$ is unitary, and $[U,H] = 0$. It is a continuous symmetry. This is called Noether's theorem.

For example, consider translations: $x \mapsto x + a$. The unitary operator that implements this is the translation operator $T(a) = e^{-ipa/\hbar}$. Then, Noether's theorem states that p is Hermitian, which we already knew. In d dimensions, $x_i \mapsto x_i + a_i$ with $1 \le i \le d$. The assumption that the $T_i(a_i)$ all commute imply that the p_i commute for all i. We have assumed this all along.

An interesting point of view is to *define* momentum as the observable that is the conserved quantity corresponding to translation invariance. Then, you can use the fact that $T(a)^\dagger x T(a) = x + a$ to conclude that $[p,x] = -i\hbar$. We can similarly define energy/the Hamiltonian as the observable that is the conserved quantity corresponding to time translations.