PHYSICS 215A: QUANTUM MECHANICS

IAN LIM
LAST UPDATED NOVEMBER 4, 2019

These notes were taken for Physics 215A, *Quantum Mechanics*, as taught by Mukund Rangamani at the University of California, Davis in fall quarter 2019. I live-TEXed them using Overleaf, and as such there may be typos; please send questions, comments, complaints, and corrections to itlim@ucdavis.edu.
Many thanks to Arun Debray for the LATEX template for these lecture notes: as of the time of writing, you can find him at https://web.ma.utexas.edu/users/a.debray/.

Contents

Lecture 1.

Wednesday, September 25, 2019

The aim of this course is to revisit quantum mechanics in more depth and more mathematical rigor. The official course text is Shankar's Principles of Quantum Mechanics, but the lecturer also recommends Weinberg's QM, Sakurai, and Leslie Valentine. To really understand the foundations of QM, Dirac's book is also useful.

Origins of QM One of the original motivations for QM was explaining the blackbody spectrum. The Rayleigh-Jones curve described the small-temperature limit, and Planck later added a correction, exponential damping at high T. Underlying Planck's model was some suggestion of quantized, discrete energies, and this theory was further developed with Schrödinger's wave mechanics.

Today, there have been many precision tests of quantum mechanics confirming that QM is a good model of small-scale phenomena. We'll spend the rest of today discussing the mathematical formalism which makes QM possible.

Mathematical background Quantum mechanics is basically infinite-dimensional[1] linear algebra. Why care about linear algebra? The key idea is this– QM obeys a superposition principle. Our theory is linear.

To build this theory, we need objects living in a vector space and some operators acting on those vectors. More precisely, we will deal with

- wavefunctions, i.e. states in a Hilbert space
- *calculables*, representing operator expectation values.[2]

Definition 1.1. A *linear vector space* is a collection of elements called *vectors*, denoted by kets $\{|v_i\rangle\}$, on which two operations are defined:

- Addition, $+ \to |v\rangle + |w\rangle$,
- and scalar multiplication, $\alpha|v\rangle, \alpha \in \mathbb{C}$.

These operations satisfy the following properties. For all $|v_i\rangle, |v_j\rangle \in V$, $\alpha \in \mathbb{C}$,

- $|v_i\rangle + |v_j\rangle \in V$
- $\alpha|v_i\rangle \in V$.

From these axioms, it follows that all linear combinations (superpositions) of vectors are allowed. That is, for $|v_i\rangle, |v_j\rangle \in V, \alpha_i, \alpha_j \in \mathbb{C}$,

$$\alpha_i|v_i\rangle + \alpha_j|v_j\rangle \in V. \tag{1.2}$$

From the axioms, we can also prove some useful properties:

- $\exists|0\rangle \in V$ (additive identity) such that $|v\rangle + |0\rangle = |v\rangle$

[1] Well, finite in spin systems and so on.

[2] We won't get into the details of measurement or observation in this course.

- For all $|v\rangle \in V$, $\exists |-v\rangle \in V$ (additive inverse) such that $|v\rangle + |-v\rangle = |0\rangle$.

We will often be sloppy with our notation and denote $|0\rangle \sim 0$, so that $|v\rangle - |v\rangle = 0$, which is secretly the zero vector.

It would be frustrating if our set of vectors was simply impossible to manage, i.e. if the vectors had no nontrivial relationships between each other.[3] Therefore, we will introduce the following definition.

Definition 1.3. A set of vectors $\{|w_i\rangle\}$ comprises a *linearly independent set* if no nontrivial linear combination of them sums to zero, i.e. if

$$\sum_i \alpha_i |w_i\rangle = 0 \implies \alpha_i = 0. \tag{1.4}$$

Definition 1.5. A set of vectors which is not linearly independent (there exists some combination α_i not all zero such that $\sum \alpha_i |w_i\rangle = 0$) is called linearly dependent.

Linear independence allows us to pick a special *basis set*, which we denote $\{|e_n\rangle\}$ such that for *any* $|v\rangle \in V$, there exists a decomposition

$$|v\rangle = \sum \alpha_n |e_n\rangle. \tag{1.6}$$

Let's illustrate this with some examples.

Example 1.7. The space $\mathbb{R}^3$ is a (real) vector space, where vectors can be denoted

$$\mathbf{v} = x\hat{e}_x + y\hat{e}_y + z\hat{e}_z. \tag{1.8}$$

Equivalently in ket notation we could write

$$|v\rangle = x|e_x\rangle + y|e_y\rangle + z|e_z\rangle. \tag{1.9}$$

This generalizes in the obvious way to $\mathbb{R}^n$.

Example 1.10. Consider a 1-qubit system, a quantum spin system with two states. A general state is written[4]

$$\alpha|0\rangle + \beta|1\rangle, \tag{1.11}$$

with $\alpha, \beta \in \mathbb{C}$. This is a complex vector space.

Example 1.12. We can define a *discretuum vector space* with a basis set $|n\rangle, n = 0, 1, 2, \ldots$. A general element is written

$$\sum_{n=0}^{\infty} \alpha_n |n\rangle, \tag{1.13}$$

with $\alpha_n \in \mathbb{C}$. This is of course the space of states for the harmonic oscillator, or more generally any confining potential.

Example 1.14. Consider the space spanned by 2×2 matrices, defined over $\mathbb{C}$. In particular, take

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.15}$$

Example 1.16. We could have a continuum vector space which is the space of functions over some domain.[5] In particular, consider the Hermite polynomials $H_n(x)$, $x \in \mathbb{R}$, the solutions to the harmonic oscillator. This is just a function representation of the solutions. Similarly the Bessel functions from electromagnetism and spherical harmonics from the spherical Laplacian form vector spaces.

The rest of today's discussion centers on how to choose a useful basis.

Definition 1.17. If a vector can be written as

$$|v\rangle = \sum_n \alpha_n |e_n\rangle \tag{1.18}$$

with respect to a basis $\{|e_n\rangle\}$, we say that α_n are *components* of $|v\rangle$ in the basis $\{|e_n\rangle\}$.

[3] Free fields are kind of like this.

[4] Some people prefer $|\uparrow\rangle|\downarrow\rangle$ or $|+\rangle, |-\rangle$.

[5] Strictly this is a countable vector space with continuum elements. For a vector space with uncountably infinite basis elements, consider plane waves and a Fourier decomposition.

In a different basis, the components will generally change, but the *vector does not*.[6]

Normed vector space It's possible to do linear algebra without an inner product. But we're physicists, so we shall define one.

Definition 1.19. An *inner product* is a function of two vectors, denoted in bra-ket notation as

$$(|v\rangle, |w\rangle) = \langle v|w\rangle, \tag{1.20}$$

and obeys the following properties:

- $\langle v|w\rangle \in \mathbb{C}$
- $\langle v|w\rangle = \langle w|v\rangle^*$
- $\langle v|\alpha_1 w_1 + \alpha_2 w_2\rangle = \alpha\langle v|w_1\rangle + \alpha_2\langle v|w_2\rangle$
- $\langle v|v\rangle \geq 0$ with $\langle v|v\rangle = 0 \iff |v\rangle = 0$.

Definition 1.21. We can then define the *norm* of a vector as

$$||v|| = \sqrt{\langle v|v\rangle}. \tag{1.22}$$

Exercise 1.23. Show that the inner product is antilinear in the first argument, i.e.

$$\langle \alpha_1 v_1 + \alpha_2 v_2|w\rangle = \alpha_1^*\langle v_1|w\rangle + \alpha_2^*\langle v_2|w\rangle. \tag{1.24}$$

We can derive some useful properties of an inner product. The inner product must satisfy the *Schwarz inequality*,

$$|\langle w|v\rangle|^2 \leq \langle w|w\rangle\langle v|v\rangle. \tag{1.25}$$

Recall that in $\mathbb{R}^3$, this is just the statement that $\langle w|v\rangle = \mathbf{v}\cdot\mathbf{w} = |\mathbf{v}||\mathbf{w}|\cos\theta$. We can prove it by considering the vector

$$|z\rangle = |v\rangle - \frac{\langle w|v\rangle}{||w||^2}|w\rangle \tag{1.26}$$

and using the positivity of the norm of $|z\rangle$.

It also satisfies the *triangle inequality*,

$$||v + w|| \leq ||v|| + ||w||. \tag{1.27}$$

A basis $\{|e_n\rangle\}$ can be made orthonormal, i.e. such that

$$\langle e_n|e_m\rangle = \delta_{nm}. \tag{1.28}$$

This can always be done by the Gram-Schmidt algorithm.

In an orthonormal basis, the inner product between two vectors is very easy to calculate. With $|v\rangle = \sum \alpha_n|e_n\rangle, |w\rangle = \sum \beta_n|e_n\rangle$, the inner product is given by

$$\langle v|w\rangle = \sum_n \alpha_n^*\beta_n. \tag{1.29}$$

Dual vector space We wish to use the norm to define a *dual vector space*. Strictly we don't need a norm to do this, but it provides a natural way to do so.

Definition 1.30. Given a vector space V, we define the dual vector space V^*, whose elements are *linear functionals* on V. That is, if $F_w \in V^*$, then

$$F_w : V \to \mathbb{C} \tag{1.31}$$

such that

$$F_w(\alpha_1|v_1\rangle + \alpha_2|v_2\rangle = \alpha_1 F_w(|v_1\rangle) + \alpha_2 F_(|v_2\rangle). \tag{1.32}$$

Moreover, we claim that the dual vector space deserves its name; it is an honest vector space. This is pretty easy to check, i.e. that linear combinations of linear functionals are themselves linear functionals.

For a ket vector $|v\rangle \in V$, we can associate a bra vector $\langle v| \in V^*$. Given a basis $\{|e_n\rangle\}$ for V, we may define a dual basis $\{\langle e_m|\}$ for V^*, defined such that

$$\langle e_m|e_n\rangle = \delta_{mn}. \tag{1.33}$$

[6] This is a critical fact in general relativity, where vectors are defined in a tangent bundle.

Note that this immediately implies that any ket vector has a corresponding bra vector in the natural way. That is, we may define an *adjoint operation* such that for

$$|v\rangle = \sum_n \alpha_n |e_n\rangle, \tag{1.34}$$

there exists a corresponding bra vector

$$\langle v| = \sum_n \alpha_n^* \langle e_n|. \tag{1.35}$$

In this sense, the vector space V and its dual vector space V^* are isomorphic.

We can also write the vectors in a funny way:

$$|v\rangle = \sum_n \langle e_n|v\rangle |e_n\rangle \tag{1.36}$$

and similarly

$$\langle v| = \sum_m \langle v|e_m\rangle \langle e_m|. \tag{1.37}$$

That is, the components of a vector are simply given by its projections onto the basis vectors.

Lecture 2.

Tuesday, October 1, 2019

Now that we have defined vectors and vector spaces, let us define linear operators.

Definition 2.1. An *operator* is a map $\mathcal{O} : V \to V$ such that

$$|v'\rangle = \mathcal{O}|v\rangle. \tag{2.2}$$

Linear operators obey

$$\mathcal{O}(\alpha_1|v_1\rangle + \alpha_2|v_2\rangle) = \alpha_1 \mathcal{O}|v_1\rangle + \alpha_2 \mathcal{O}|v_2\rangle. \tag{2.3}$$

Operators form an algebra, i.e. addition and multiplication are defined on operators such that

$$(A + B)|v\rangle = A|v\rangle + B|v\rangle \tag{2.4}$$

$$(AB)|v\rangle = A(B|V\rangle). \tag{2.5}$$

Addition is commutative, but multiplication is not– generically,

$$AB \neq BA. \tag{2.6}$$

Multiplication is however asssociative,

$$A(BC) = (AB)C. \tag{2.7}$$

Linear operators can be thought of as generalizations of matrices. In particular they have *matrix elements* given by $\langle v|\mathcal{O}|w\rangle$ or in a basis,

$$\langle e_n|\mathcal{O}|e_m\rangle = \mathcal{O}_{nm}. \tag{2.8}$$

If $\mathcal{O}|v\rangle = |w\rangle$ and $|v\rangle = \sum v_n|e_n\rangle, |w\rangle = \sum w_m|e_m\rangle$, then it follows

$$w_m = \sum_n \mathcal{O}_{mn} v_n. \tag{2.9}$$

Exercise 2.10. Check Eq. **??** from the definition of the matrix element and inner product.

Sometimes we conflate operators with their matrix elements, but the matrix elements are basis-dependent; the operator is not.

Adjoint of an operator Let us now define the adjoint of an operator. Normally we have

$$\langle v|\mathcal{O}|w\rangle = \langle v|(\mathcal{O}|w\rangle), \tag{2.11}$$

but can we construct some operator that allows us to evaluate this as

$$(\langle v|\mathcal{O})|w\rangle? \tag{2.12}$$

Definition 2.13. If $\mathcal{O}|v\rangle = |w\rangle$, let us define the adjoint $\mathcal{O}^\dagger$ by

$$\langle v|\mathcal{O}^\dagger = \langle w|. \tag{2.14}$$

From this definition, it follows that since $\langle z|v\rangle = \langle v|z\rangle^*$, we have

$$\langle v|\mathcal{O}^\dagger|z\rangle^* = \langle z|\mathcal{O}|v\rangle. \tag{2.15}$$

In terms of matrix elements, we know this is just the conjugate transpose,

$$(\mathcal{O}^\dagger)_{mn} = O^*_{nm}. \tag{2.16}$$

As a matrix, $O^\dagger = (O^*)^T$.

Notice that the adjoint has the following properties:

- $(\alpha\mathcal{O})^\dagger = \alpha^*\mathcal{O}^\dagger$ for $\alpha \in \mathbb{C}$
- $(A+B)^\dagger = \alpha^\dagger + B^\dagger$
- $(AB)^\dagger = B^\dagger A^\dagger$.

Exercise 2.17. Check these properties from the definition of the adjoint.

Some examples of operators include the following:

- $\mathbb{I}$, the identity operator acting as $\mathbb{I}|v\rangle = |v\rangle$.
- Over $\mathbb{C}^2$ spanned by $\{|1\rangle, |0\rangle\}$, the Pauli matrices and the identity element are linear operators. In particular they are complete, and obey the commutators

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \tag{2.18}$$
$$[\sigma_i, \mathbb{I}] = 0. \tag{2.19}$$

In finite dimension, operators are basically matrices. How do we generalize to the infinite-dimensional cases we often see in quantum mechanics?

Example 2.20. Consider first the vector space of smooth functions $\{f(x)\}$ on $\mathbb{R}$. We can define operators $x^m, m \in \mathbb{Z}_{\geq 0}$ which simply multiply these smooth functions,

$$x^m : f(x) \to x^n f(x). \tag{2.21}$$

Example 2.22. The derivative also defines an operator on this vector space. We might define

$$x\frac{\partial}{\partial x} : f(x) \to xf'(x) \tag{2.23}$$

or similarly

$$\frac{\partial}{\partial x}x : f \to \frac{\partial}{\partial x}(xf(x)) = xf' + f \tag{2.24}$$

Sometimes we will write

$$\frac{\partial}{\partial x}x = x\frac{\partial}{\partial x} + 1, \tag{2.25}$$

omitting the function f.

Self-adjoint operators and computables In a sentence, we can describe quantum mechanics as complex matrix linear algebra in infinite dimensions. But let us note that while QM in general is complex, the things we measure must be real. This leads us to introduce the notion of self-adjoint operators.

Definition 2.26. A *self-adjoint operator* is an operator satisfying

$$\langle w|\mathcal{O}|v\rangle = \langle v|\mathcal{O}|w\rangle^*. \tag{2.27}$$

In other words, $\mathcal{O} = \mathcal{O}^\dagger$, or in terms of matrix elements, $\mathcal{O}_{nm} = \mathcal{O}^*_{mn}$.

For most cases, self-adjoint $\sim$ Hermitian.[7] Self-adjoint operators are nice because their eigenvalues are *real*, meaning that they correspond to observables in our theory.

We'll also use the following definition later, the trace.

Definition 2.28. The *trace* of an operator is defined to be

$$\mathrm{Tr}(\mathcal{O}) = \sum_n \langle e_n|\mathcal{O}|e_n\rangle. \tag{2.29}$$

We claim it is independent of basis, and in fact we will prove it on the first homework.

Eigenspectrum Just as in the finite-dimensional case, we can talk about the eigenvectors and eigenvalues of operators acting on infinite-dimensional spaces.

Definition 2.30. The (nonzero) vector $|w\rangle$ is an eigenvector of the operator $\mathcal{O}$ with eigenvalue α if

$$\mathcal{O}|w\rangle = \alpha|w\rangle, \alpha \in \mathbb{C}. \tag{2.31}$$

Hermitian operators have real eigenvalues, since

$$\mathcal{O}|w\rangle = \alpha|w\rangle \implies \langle w|O^\dagger = \alpha^*|w\rangle. \tag{2.32}$$

Sandwiching with $\langle w|$ or $|w\rangle$ as appropriate, we see that

$$\langle w|\mathcal{O}|w\rangle = \alpha||w||^2 \tag{2.33}$$

and

$$\langle w|\mathcal{O}|w\rangle = \langle w|\mathcal{O}^\dagger|w\rangle = \alpha^*||w||^2 \tag{2.34}$$

since $\mathcal{O} = \mathcal{O}^\dagger$. Hence

$$\alpha||w||^2 = \alpha^*||w||^2 \implies \alpha = \alpha^*, \tag{2.35}$$

provided that $|w\rangle$ is not the zero vector. Note that while the eigenvalues are real, the matrix elements need not be real. For instance, σ_2 has complex entries in the $|0\rangle, |1\rangle$ basis but it is nevertheless Hermitian and has real eigenvalues.

Theorem 2.36. *If A is self-adjoint (Hermitian) then all eigenvalues are real. Eigenvectors corresponding to distinct (non-degenerate) eigenvalues are orthogonal.*

This theorem follows from a simpler lemma:

Lemma 2.37. *If $\langle v|A|v\rangle = \langle v|A|v\rangle^*$ for all $|v\rangle$ then $A = A^\dagger$.*

That is, if every diagonal matrix element is real, then the matrix is Hermitian.

Proof. To prove this lemma, we need to show that the given condition implies

$$\langle v_2|A|v_1\rangle^* = \langle v_1|A|v_2\rangle. \tag{2.38}$$

Define

$$|v\rangle = \alpha_1|v_1\rangle + \alpha_2|v_2\rangle \tag{2.39}$$

for some $\alpha_1, \alpha_2 \in \mathbb{C}$. Then

$$\langle v|A|v\rangle = |\alpha_1|^2\langle v_1|A|v_1\rangle + |\alpha_2|^2\langle v_2|A|v_2\rangle + \alpha_1^*\alpha_2\langle v_1|A|v_2\rangle + \alpha_2^*\alpha_1\langle v_2|A|v_1\rangle. \tag{2.40}$$

We can write $\langle v|A|v\rangle^*$ as well, taking the complex conjugate of the previous expression. We know that

$$\langle v|A|v\rangle = \langle v|A|v\rangle^* \tag{2.41}$$

[7] See Homework 1 for an example of how this can fail in an infinite-dimensional vector space.

for all $|v\rangle$, so setting these equal we find that

$$\alpha_1^*\alpha_2\langle v_1|A|v_2\rangle + \alpha_2^*\alpha_1\langle v_2|A|v_1\rangle = \alpha_1\alpha_2^*\langle v_1|A|v_2\rangle^* + \alpha_2\alpha_1^*\langle v_2|A|v_1\rangle^*. \tag{2.42}$$

Let us evaluate for $\alpha_1 = \alpha_2 = 1$. This yields

$$\langle v_1|A|v_2\rangle + \langle v_2|A|v_1\rangle = \langle v_1|A|v_2\rangle^* + \langle v_2|A|v_1\rangle^*. \tag{2.43}$$

We can also make the choice $\alpha_1 = 1, \alpha_2 = i$ so that

$$i\langle v_1|A|v_2\rangle - i\langle v_2|A|v_1\rangle = -i\langle v_1|A|v_2\rangle^* + i\langle v_2|A|v_1\rangle^*. \tag{2.44}$$

Dividing by i and adding these equations yields

$$\langle v_2|A|v_1\rangle^* = \langle v_1|A|v_2\rangle \implies A = A^\dagger. \tag{2.45}$$

⊠

Some quick definitions:

- An *anti-Hermitian* matrix is one obeying $A^\dagger = -A$.
- A *unitary* matrix is one satisfying $UU^\dagger = \mathbb{I}$. It may be thought of as a complex rotation on vector space.[8]

Lecture 3.

Thursday, October 3, 2019

Spectral theorem The spectral theorem roughly tells us that the eigenvectors of hermitian matrices will be guaranteed to form a good basis. Before we state the theorem formally, let us discuss the following. Given some Hermitian operator $\mathcal{Z}$ with eigenspectrum

$$\mathcal{Z}|z_n\rangle = \zeta_n|z_n\rangle. \tag{3.1}$$

Since we are guaranteed completeness[9] it follows that

$$\mathbb{I}|v\rangle = |v\rangle = \sum|z_n\rangle\langle z_n|v\rangle, \tag{3.2}$$

so in fact reading this as an operator equation, since this is true for any $|v\rangle$,

$$\mathbb{I} = \sum_n |z_n\rangle\langle z_n|. \tag{3.3}$$

We call Eqn. **??** the *resolution of the identity*, and this is true for any complete basis set.

That is, given $|v\rangle \in V$ and $\langle w| \in V^*$, there is another operation, the *outer product* (denoted $\otimes$), which is a map from $(V, V^*) \to V \otimes V^*$. That is, it allows us to take a column vector and a row vector and combine them to form an operator (a matrix). Thus

$$|v\rangle\langle w| \in V \otimes V^* \tag{3.4}$$

is an operator such that

$$(|v\rangle\langle w|)|z\rangle = \langle w|z\rangle|v\rangle. \tag{3.5}$$

In tensor notation, we could write these as contravariant and covariant vectors as V^μ, ω_ν such that the outer product $(V\omega)^\mu{}_\nu$ has the correct indices.

We can also define the adjoint operation on operators in this notation,

$$(|v\rangle\langle w|)^\dagger = |w\rangle\langle v|. \tag{3.6}$$

The adjoint operator may be thought of as a map $V \otimes V^* \to V^* \otimes V$.

Consider now

$$\mathcal{P}_n \equiv |z_n\rangle\langle z_n| \tag{3.7}$$

[8] Strictly we have not excluded reflections but we'll discuss anti-linear operators later.

[9] We haven't shown it yet but suppose there's a vector that cannot be expressed as an eigenvector. We can reason to a contradiction. Probably in Shankar?

the projection operator onto $|z_n\rangle$. Naturally, $\mathcal{P}_n\mathcal{P}_n = \mathcal{P}_n$ since

$$\mathcal{P}_n|v\rangle = |z_n\rangle\langle z_n|v\rangle$$
$$\implies \mathcal{P}_n^2|v\rangle = |z_n\rangle\langle z_n|z_n\rangle\langle z_n|v\rangle = \mathcal{P}_n|v\rangle \text{ if } \langle z_n|z_n\rangle = 1.$$

In this notation, we may equivalently write the identity as

$$\mathbb{I} = \sum_n \mathcal{P}_n. \tag{3.8}$$

That is, the identity is the sum of all the projection operators. Moreover,

$$\mathcal{Z} = \sum_n \zeta_n |z_n\rangle\langle z_n| = \sum_n \zeta_n \mathcal{P}_n(\mathcal{Z}), \tag{3.9}$$

where the $\mathcal{P}_n$ are projecting onto the eigenspectrum of $\mathcal{Z}$.

Since this is given, it becomes easy to define functions of operators $f(\mathcal{Z})$ in terms of their eigenspectrum. Namely,

$$f(\mathcal{Z}) = \sum_n f(\zeta_n)\mathcal{P}_n(\mathcal{Z}) \tag{3.10}$$

Note that an operator cannot have a nontrivial kernel (that is, it cannot have a zero eigenvalue) or else our completeness assumption fails. In addition, some functions have a finite radius of convergence and so the power series is not guaranteed to converge if the eigenvalues are unbounded.[10]

Let us also note that we've been working as though these spaces were finite-dimensional, but there are sometimes complications when we go to infinite dimensions. To every self-adjoint operator we can associate a 1-parameter family of projection operators $\mathcal{P}(\lambda)$, parametrized by some $\lambda \in \mathbb{R}$ such that $\mathcal{P}(\lambda)$ satisfies

(i) $\lambda_1 < \lambda_2 \implies \mathcal{P}(\lambda_1)\mathcal{P}(\lambda_2) = \mathcal{P}(\lambda_2)\mathcal{P}(\lambda_1) = \mathcal{P}(\lambda_1)$. (That is, we always project onto smaller λ.)
(ii) If $\epsilon > 0$ then $\mathcal{P}(\lambda+\epsilon)|v\rangle \to \mathcal{P}(\lambda)|v\rangle$ as $\epsilon \to 0$. (In a sense this family varies continuously.)
(iii) $\mathcal{P}(\lambda)|v\rangle \to 0$ as $\lambda \to -\infty$.
(iv) $\mathcal{P}(\lambda)|v\rangle \to |v\rangle$ as $\lambda \to \infty$.
(v) $\int_{-\infty}^{\infty} \lambda d\mathcal{P}(\lambda) = \mathcal{Z}$.

Discrete spectrum Let's examine these assumptions in the discrete case.

$$\mathcal{P}(\lambda) = \sum_n \Theta(\lambda - \zeta_n)\mathcal{P}_n(\mathcal{Z}), \tag{3.11}$$

where

$$\Theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases} \tag{3.12}$$

is the Heaviside step function. In the degenerate case we can project onto the corresponding subspace.

To put this in physics language, we can project onto energy eigenspaces. That is, λ lets us take sums of the projection operators up to some energy eigenstate of our choice. As $\lambda \to -\infty$ we're looking at energies below the ground state, so there are no states to project on; as $\lambda \to +\infty$ we get all the energies and hence get back the original operator $\mathcal{Z}$.

Continuous spectrum Let's consider a concrete example. For the position operator X, which we define as acting on a function $f(x)$ by

$$Xf(x) = xf(x), \tag{3.13}$$

we could try to solve the eigenvalue equation

$$Xf(x) = \zeta f(x). \tag{3.14}$$

We might be tempted to write the eigenfunctions as delta functions, $\delta(x-\zeta)$. But this would be a serious error, since the delta function is really a distribution in the space of linear functionals, and ought to live under integrals. That is,

$$\int_{-\infty}^{\infty} dx \delta(x) f(x) = f(0) \tag{3.15}$$

[10] See Shankar 1.9.1.

is the statement that the delta function should be thought of as only being meaningful when under an integral and paired with a function; it is more like a dual vector, since it takes a function and gives back a number.

Note that from the delta function we can also write other derivative expressions like

$$\delta'(x) : f(x) \to \mathbb{R}, \tag{3.16}$$

another functional such that

$$\begin{aligned}\delta'[f(x)] &= \int_{-\infty}^{\infty} dx\, \delta'(x) f(x) \\ &= -\int_{-\infty}^{\infty} dx\, \delta(x) f'(x) + (\text{boundary term} \to 0) \\ &= -f'(0).\end{aligned}$$

Given appropriate smoothness of $f(x)$, anyway.

To generalize our one-parameter family of projection operators to the continuous case, we can write the projection operators for the position operator X as

$$\mathcal{P}_X(\lambda) f(x) = \Theta(\lambda - x) f(x). \tag{3.17}$$

That is, we may project onto x up to $x = \lambda$. For notice that

$$\left[\int_{-\infty}^{\infty} \lambda d\mathcal{P}_X(\lambda)\right] f(x) = \int_{-\infty}^{\infty} \lambda \delta(\lambda - x) d\lambda\, f(x) = x f(x). \tag{3.18}$$

Hence by (v) above we see that

$$X = \int_{-\infty}^{\infty} \lambda d\mathcal{P}_X(\lambda). \tag{3.19}$$

Theorem 3.20. *If $\mathcal{A}, \mathcal{B}$ are mutually commuting, self-adjoint operators, each with a complete set of eigenvalues, then $\exists$ a complete orthonormal set of eigenvectors for both $\mathcal{A}$ and $\mathcal{B}$ simultaneously.*

Since we are guaranteed this, the name of the game is to find the maximal set of commuting operators.[11] If we find a complete set of commuting operators, then we can not only simultaneously diagonalize all of them, but any other operator that commutes with the given set also has the same eigenbasis.

In a sense, this is much like how we look for conserved quantities in classical mechanics to avoid solving second-order equations. We'll be particularly interested in operators that commute with the Hamiltonian, since energy eigenstates will coincide with eigenstates of those operators.

Lecture 4.

Tuesday, October 7, 2019

Today we will introduce the postulates of QM. They are as follows:

I. The state of a system is described by a *ray*[12] in Hilbert space.[13]

II. States evolve by unitary evolution,

$$|\psi(t)\rangle = U(t; t_0)|\psi(t_0)\rangle, \tag{4.1}$$

such that $U^\dagger U = \mathbb{I}$.[14]

III. Physical observables[15] correspond to expectation values of linear Hermitian operators which naturally act on states in Hilbert space. That is, for $\mathcal{A} = \mathcal{A}^\dagger$, we have the expectation value $\langle\psi|\mathcal{A}|\psi'\rangle$.

[11] Like we did with angular momentum, for instance, in finding J^2 and J_z.

[12] That is, a vector defined up to normalization. Two vectors that differ only by normalization should be considered as representing the same physical state.

[13] For our purposes, a Hilbert space is a complex vector space.

[14] Intuitively, so long as time evolution is unitary, states just evolve by complex rotations in Hilbert space. Norms and in particular probabilities are therefore left unchanged by unitary evolution.

[15] That is, observables of closed quantum systems where the observer is *not part of the system*. In Mukund's words, we deal with a "meta-observer theory" where the observer can make measurements on a quantum system. There are attempts (cf. Everettian "many-worlds" QM) to describe the observer as part of the system, but this leads to lots of confusion.

IV. *The Born rule.* A measurement of a physical observable $\mathcal{A}$ yields exactly one of its eigenvalues λ with probability $p(\lambda)$ given by

$$\begin{aligned} p(\lambda) &= \langle\psi|\mathcal{P}_\lambda(\mathcal{A})|\psi\rangle \\ &= \langle\psi|a_\lambda\rangle\langle a_\lambda|\psi\rangle \\ &= ||\langle a_\lambda|\psi\rangle||^2 \end{aligned}$$

where $\mathcal{P}_\lambda(\mathcal{A})$ is the projection operator onto $|a_\lambda\rangle$, the eigenstate with eigenvalue λ, and we recall that $\mathcal{A}$ has a spectral decomposition $\mathcal{A} = \sum\lambda|a_\lambda\rangle\langle a_\lambda|$.

There are different interpretations of measurement in quantum mechanics. Pragmatically speaking, they should all give the same predictions so it doesn't really matter for computations[16] and we will present one of the common interpretations, the *Copenhagen interpretation*. In the Copenhagen interpretation, measurement projects the state vector $|\psi\rangle$ onto one of the eigenstates of the operator being measured.[17] This is what we might call "collapse of the wavefunction." Notice that

$$\mathcal{A}|\psi\rangle = \sum_\lambda \lambda|a_\lambda\rangle\langle a_\lambda|\psi\rangle, \tag{4.2}$$

where the final inner product is the projection of the original state $|\psi\rangle$ onto the eigenstate $\langle a_\lambda|$, and represents a probability amplitude (namely, its modulus squared represents the probability of measuring λ).

Evolution in Hilbert space By our postulate, time evolution is governed by unitary operators.

Theorem 4.3 (Wigner)**.** *Any mapping of a complex vector space onto itself that preserves the inner product is implemented either by*

i) *a unitary operator (linear transformation) or*
ii) *an anti-unitary operator (anti-linear).*

That is, in case i) we have $U(\alpha|\psi\rangle) = \alpha U|\psi\rangle, \alpha \in \mathbb{C}$, whereas for an anti-linear operator in ii) we have $U(\alpha|\psi\rangle) = \alpha^* U|\psi\rangle$. Most operators in physics are unitary apart from one famous example, which is time reversal.[18]

What Wigner's theorem guarantees us is that every symmetry of Hilbert space can be implemented by a unitary operator or an anti-unitary operator. The state evolution is therefore given by the (unitary) time translation operator, which is in turn generated by the Hamiltonian.[19]

Let us now revisit the notion of expectation values, considering the VEV (traditionally "Vacuum Expectation Value") of an operator:

Definition 4.4. The *expectation value* of an operator $\mathcal{A}$ in a state $|\psi\rangle$ is given by

$$\langle\mathcal{A}\rangle = \langle\psi|\mathcal{A}|\psi\rangle \tag{4.5}$$

$$= \sum\lambda\langle\psi|a_\lambda\rangle\langle a_\lambda|\psi\rangle. \tag{4.6}$$

We may also talk about the uncertainty (basically the standard deviation) of an operator.

Definition 4.7. The uncertainty of an operator $\mathcal{A}$ is given in terms of expectation values:

$$\Delta\mathcal{A} = \sqrt{\langle(\mathcal{A} - \langle A\rangle\,\mathbb{I})^2\rangle}. \tag{4.8}$$

This is a quantum generalization of the standard deviation from ordinary statistics– we want to know the root mean squared deviation from the mean value $\langle A\rangle$.

[16] But see J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics* for more on the philosophy of QM.

[17] The counterpoint to this is the Everettian picture. According to Copenhagen, a measurement is a lossy operation, where we lose information upon making a measurement. In many-worlds, the wavefunction instead branches upon making a measurement. If you like, we discover which branch of the wavefunction we were really on, since we are part of the quantum system.

[18] This has to do with the i in the Schrödinger equation. Most anti-unitary operators one encounters in QM are related to time reversal in some way. Moreover the fact that the operator is anti-unitary suggests that in QM and even in QFT, there is a special direction for time.

[19] Here I mean generated in the Lie algebra sense. That is, the Hamiltonian produces infinitesimal (differential) translations in time, and the exponential of an generator gives the full translation. We'll see this soon.

Single-particle QM In single-particle quantum mechanics, we're interested in studying a particle moving in 1 dimension under some external potential V. Classically, we would use an action principle (we'll revisit this in the path integral formulation later). That is, we write down a Lagrangian

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x) \tag{4.9}$$

and we extremize the action $\int dt\mathcal{L}(x,\dot{x})$. However, for this part of the course we will instead use Hamiltonian language:

$$\mathcal{H} = p\dot{x} - \mathcal{L}, \quad p = \frac{\delta\mathcal{L}}{\delta\dot{x}}, \tag{4.10}$$

where the Hamiltonian is the Legendre transform of the Lagrangian, and thus if the potential depends only on x then

$$\mathcal{H} = \frac{1}{2m}p^2 + V(x), \tag{4.11}$$

in terms of the conjugate momentum $p = m\dot{x}$. Then we can write down Hamilton's equations, treating x and p as independent variables to find

$$\dot{x} = \frac{\partial\mathcal{H}}{\partial p}, \quad \dot{p} - \frac{\partial\mathcal{H}}{\partial x}. \tag{4.12}$$

A particle's state is therefore represented by a point in phase space, i.e. the space of positions and momenta. For a particle in 1 dimension, the phase space is $\mathbb{R}^2 = (x,p)$, and in particular the phase space has an additional geometric structure. Namely, it is a *symplectic space*, meaning that that a Poisson bracket is defined on this space.

Definition 4.13. Given two functions $f(x,p), g(x,p)$, the *Poisson bracket* of f,g is simply

$$\{f,g\} = \frac{\partial f}{\partial x}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial x}. \tag{4.14}$$

Hamilton's equations are nice in terms of Poisson brackets– they just become

$$\dot{x} = \{x,\mathcal{H}\}, \quad \dot{p} = \{p,\mathcal{H}\}. \tag{4.15}$$

Note that $\frac{df}{dt} = \{f,\mathcal{H}\} + \frac{\partial f}{\partial t}$, in the case where f may explicitly depend on time. Here, t is time parametrizing trajectories in phase space.

Canonical quantization Now for quantum mechanics. The procedure of going from a classical system to a quantum system is called *canonical quantization*. This is done by taking phase space variables x,p and uplfiting them to Hermitian operators $\hat{x},\hat{p}$. We must also replace classical Poisson bracket relations between position and momentum with operator commutation relations (Lie brackets). Classically, the Poisson bracket between x,p is

$$\{x,p\} = 1. \tag{4.16}$$

In QM, this becomes

$$[\hat{x},\hat{p}] = i\hbar, \tag{4.17}$$

with the commutator defined as

$$[\hat{A},\hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{4.18}$$

The commutator satisfies the following properties:

i) linearity, $[\hat{A},\alpha_1\hat{B} + \alpha_2\hat{B}_2] = \alpha_1[\hat{A},\hat{B}_1] + \alpha_2[\hat{A},\hat{B}_2]$ and similar for the first argument;
ii) The Jacobi identity,
$$[\hat{A},[\hat{B},\hat{C}]] + [\hat{B},[\hat{C},\hat{A}]] + [\hat{C},[\hat{A},\hat{B}]] = 0 \tag{4.19}$$
iii) Leibniz rule, $[\hat{A},\hat{B}\hat{C}] = [\hat{A},\hat{B}]\hat{C} + \hat{B}[\hat{A},\hat{C}]$.

The classical phase space is now replaced by a Hilbert space. All other observables become operators on Hilbert space, and this includes the Hamiltonian,

$$H(x,p) \to \hat{H}(\hat{x},\hat{p}). \tag{4.20}$$

For a single particle in a simple potential, we can write it down immediately:

$$\hat{H}(\hat{x},\hat{p}) = \frac{1}{2m}\hat{p}^2 + V(\hat{x}). \tag{4.21}$$

We should also note that classically, x and p commute. But quantum mechanically, they do not. Consider the classical function xp on phase space. What is its quantum equivalent? Well, it could be $\hat{x}\hat{p}$. Or it could be $\hat{p}\hat{x}$. Or it could be the symmetrized sum $\frac{1}{2}(\hat{x}\hat{p}+\hat{p}+\hat{x})$ (Weyl ordering). A priori, we do not know. This last option is popular, however.[20]

Finally, let us define a wavefunction.

Definition 4.22. A *wavefunction* is a state vector decomposed in the position eigenbasis. That is, let

$$\hat{x}|x\rangle = x|x\rangle \tag{4.23}$$

indicate the position eigenbasis. Then

$$\psi(x) = \langle x|\psi\rangle, \tag{4.24}$$

such that

$$|\psi\rangle = \int_{-\infty}^{\infty} dx\,|x\rangle\langle x|\psi\rangle = \int_{-\infty}^{\infty} dx\,|x\rangle\psi(x). \tag{4.25}$$

Lecture 5.

Thursday, October 10, 2019

Quantum dynamics In a quantum theory, time evolution is given by a unitary operator,

$$|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle. \tag{5.1}$$

Time evolution is determined by the Hamiltonian, a hermitian operator $\hat{H} = \hat{H}^\dagger$, such that

$$e^{i\hat{H}} \tag{5.2}$$

is unitary.[21]

For infinitesimal time evolution, notice that

$$U(t_0+dt,t_0) = \mathbb{I} + \frac{dU}{dt}|_{t=t_0}dt + O(dt^2). \tag{5.3}$$

If we choose U such that

$$\frac{dU}{dt} = -\frac{i\hat{H}}{\hbar}, \tag{5.4}$$

then we can recover

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle, \tag{5.5}$$

the *Schrödinger equation* in vector notation.

For the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}), \tag{5.6}$$

we have

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \left[\frac{\hat{p}^2}{2m} + V(\hat{x})\right]|\psi(t)\rangle. \tag{5.7}$$

We can switch to the position basis, writing

$$\hat{p} = -i\hbar\frac{\partial}{\partial x}, \tag{5.8}$$

[20] For another example of Weyl ordering, we would associate

$$x^3p \to \frac{1}{4}(\hat{x}^3\hat{p} + \hat{x}^2\hat{p}\hat{x} + \hat{x}\hat{p}\hat{x}^2 + \hat{p}\hat{x}^3).$$

[21] This is unitary because $(e^{i\hat{H}})^\dagger = e^{-i\hat{H}^\dagger} = e^{-i\hat{H}} = (e^{i\hat{H}})^{-1}$. More generally it comes from the fact that Hermitian operators form a Lie algebra which generates unitary transformations. We can build finite transformations by exponentiating the appropriate infinitesimal transformation.

from the canonical commutation relation $[\hat{x}, \hat{p}] = i\hbar$. If we then sandwich Eqn. **??** with a $\langle x|$, a position eigenstate, we get

$$i\hbar \frac{\partial}{\partial t}\psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) + V(x)\psi(x,t), \tag{5.9}$$

which is now the Schrödinger equation written in terms of the position space wavefunction.

The Schrödinger equation describes the evolution of $|\psi(t)\rangle$, but we can equivalently use it to understand $U(t, t_0)$. Notice that

$$\frac{\partial}{\partial t}U(t,t_0) = -\frac{i}{\hbar}\hat{H}(\hat{x},\hat{p},t)U(t,t_0) \tag{5.10}$$

with $U(t_0, t_0) = \mathbb{I}$. This follows from Eqn. **??** by substituting in $|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle$ and noting the derivative is with respect to t, leaving $|\psi(t_0)\rangle$ fixed. We can therefore write a formal solution for this equation. It looks almost like an exponential, except for the fact that the operators in the Hamiltonian may not commute at different times.[22]

Hence we get the formal solution for U (sometimes called Dyson's formula), written in terms of *time ordering*:

$$U(t,t_0) = \mathcal{T}\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} dt'\, \hat{H}(\hat{x},\hat{p},t')\right]. \tag{5.11}$$

Here, the curly $\mathcal{T}$ indicates time ordering, i.e. we must write the operators based on the time at which they are evaluated. We won't discuss this too much for now, since a proper discussion of this would basically be tantamount to explaining the path integral.

But the story becomes much simpler if the Hamiltonian is time-independent. In that case, our formal solution has no time ordering ambiguities, and we can immediately write down

$$U(t,t_0) = \exp\left[-\frac{i}{\hbar}(t-t_0)\hat{H}(\hat{x},\hat{p})\right]. \tag{5.12}$$

Moreover this makes it clear that energy eigenstates (i.e. eigenstates of the Hamiltonian) evolve in time by phases, $e^{-iEt/\hbar}$.

Schrödinger and Heisenberg pictures Until now, we have been working in the Schrödinger picture, where operators are left fixed, so that *observables* carry no time dependence, and states are generically time-dependent. Thus expectation values are given by

$$\langle\psi(t)|\hat{A}|\psi(t)\rangle, \tag{5.13}$$

and all the time dependence is in our states' rotation in Hilbert space.

But there is an alternate way to do calculations, the *Heisenberg picture.* In the Heisenberg picture, state vectors carry no time dependence but operators do.[23] For a Schrodinger picture state $|\psi(t)\rangle$, it is certainly true that

$$\hat{A}|\psi(t)\rangle = \hat{A}U(t,t_0)|\psi(t_0)\rangle, \tag{5.14}$$

where t_0 is some fixed earlier time. Since U is unitary we can also write this as

$$\hat{A}|\psi(t)\rangle = U(t,t_0)\Big(U^\dagger(t,t_0)\hat{A}U(t,t_0)\Big)|\psi(t_0)\rangle, \tag{5.15}$$

so that[24]

$$\langle\psi(t)|\hat{A}|\psi(t)\rangle = \langle\psi(t_0)|\Big(U^\dagger(t,t_0)\hat{A}U(t,t_0)\Big)|\psi(t_0)\rangle. \tag{5.16}$$

Call $|\psi(t_0)\rangle$ the *Heisenberg state vector*, defined at some predetermined time t_0, and then the price we pay is that our operators have become time dependent,

$$\hat{A}(t) = U^\dagger(t,t_0)\hat{A}(t_0)U(t,t_0), \tag{5.17}$$

where $\hat{A}(t)$ indicates the Heisenberg picture version of $\hat{A}$.

[22] This leads to equal-time commutation relations in the appropriate picture of QM.

[23] If you like, we regroup all the time dependence into the operators.

[24] Mukund mentioned some mnemonic for why daggers go on the left of operators. I don't know what his is, but this is how I remember it. Write the time-dependent expectation value in Schrödinger picture and regroup the Us into the operator.

Note that for a time-independent Hamiltonian,

$$\hat{A}(t) = e^{\frac{i}{\hbar}(t-t_0)\hat{H}}\hat{A}(t_0)e^{-\frac{i}{\hbar}(t-t_0)\hat{H}}, \tag{5.18}$$

and the only way to evaluate this is to Taylor expand as

$$\hat{A}(t) = \left(\mathbb{I} + \frac{i}{\hbar}(t-t_0)\hat{H} + \left(\frac{i}{\hbar}\right)^2(t-t_0)^2\frac{\hat{H}^2}{2!} + \dots\right)\hat{A}(t_0)\left(\mathbb{I} - \frac{i}{\hbar}(t-t_0)\hat{H} + \left(-\frac{i}{\hbar}\right)^2(t-t_0)^2\frac{\hat{H}^2}{2!} + \dots\right) \tag{5.19}$$

and we claim (exercise) that every other term is of the form of a nested commutator,

$$\hat{A}(t_0) + \frac{i}{\hbar}(t-t_0)[\hat{H}, \hat{A}(t_0)] + (\#)[\hat{H}, [\hat{H}, \hat{A}(t_0)]] + \dots \tag{5.20}$$

where # is some numerical factor we have not worked out yet. This fact will be important on the homework.

Incidentally, this tells us that where states evolve by rotations, operators generically transform in some complicated way. This has become a point of interest in the study of quantum chaos, how the time evolution of operators in the Heisenberg picture may depend very sensitively on the initial conditions.

Generalizations Here, we have considered the case of a single particle moving in 1D, but more generally we could have a particle moving in d space direction so that the position and momentum operators have d components,

$$\hat{x}_i,\ i = 1, \dots, d, \qquad \hat{p}_j,\ j = 1, \dots, d, \tag{5.21}$$

obeying the commutation relations

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}, \tag{5.22}$$

and in the Schrodinger equation, the $\hat{p}^2$ becomes the Laplacian in d spatial coordinates,

$$\nabla^2 = \sum_{i=1}^{d}\frac{\partial^2}{\partial x_i^2}. \tag{5.23}$$

However, the generalization to multiple particles is more subtle. If our one-particle states live in Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2, \dots$ then the total Hilbert space is not just the direct sum $\oplus$ but the tensor product

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2. \tag{5.24}$$

Hence if $\{|\psi_i\rangle\}_{i=1}^n$ generate $\mathcal{H}_1$ and $\{|\chi_j\rangle\}_{j=1}^m$ generate $\mathcal{H}_2$, then $\mathcal{H}$ is generated by *tensor products*

$$|\psi_i\rangle \otimes |\chi_j\rangle, \tag{5.25}$$

and $\mathcal{H}$ is $m \times n$-dimensional. It is this tensor product structure that is at the heart of the characteristic weirdness of quantum mechanics, namely the phenomenon known as *entanglement*.

The harmonic oscillator (Shankar Ch. 7) The quantum harmonic oscillator is a wonderful thing.[25] To begin, the classical harmonic oscillator is basically just a particle sitting in a parabolic (quadratic) potential. Its Lagrangian is

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2x^2, \tag{5.26}$$

with corresponding Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2. \tag{5.27}$$

Canonical quantization tells us to promote x and p to operators,

$$[\hat{x}, \hat{p}] = i\hbar, \tag{5.28}$$

so that

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2\hat{H} \tag{5.29}$$

$$= -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2\hat{x}^2, \tag{5.30}$$

[25] Not official lecture content. Just my (correct) opinion.

where we have rewritten $\hat{p}$ in the position basis.

Let us try to find the eigenspectrum of the Hamiltonian, i.e. the state vectors such that

$$\hat{H}|\psi(t)\rangle = E|\psi(t)\rangle. \tag{5.31}$$

The Schrödinger equation tells us that energy eigenstates evolve in a special way, namely by

$$|\psi(t)\rangle = e^{-iEt/\hbar}|\psi(0)\rangle. \tag{5.32}$$

Hence the Schrodinger equation written in terms of the position basis wavefunctions becomes

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2\hat{x}^2\right)\psi(x) = E\psi(x). \tag{5.33}$$

This is a second-order ODE for ψ:

$$\frac{d^2\psi}{dx^2} + \left(\frac{2mE}{\hbar^2} - \frac{1}{2}\frac{m^2\omega^2}{\hbar^2}x^2\right)\psi = 0. \tag{5.34}$$

We can simplify by defining

$$\frac{E}{\hbar\omega} = \epsilon, \tag{5.35}$$

$$x = \sqrt{\frac{\hbar}{m\omega}}\xi. \tag{5.36}$$

In these variables, our equation becomes

$$\frac{d^2\psi}{d\xi^2} + (2\epsilon - \xi^2)\psi = 0. \tag{5.37}$$

For such equations, we expect two families of solutions fitting the boundary conditions, one of which will be excluded on physical grounds. Typically we will have a condition on ϵ, telling us that the energy is quantized.

Suppose we didn't have a computer to find the solutions to this equation right away. How do we study this equation? One way is to look at its limiting behavior. As $\xi \to \infty$, our equation reduces to

$$\frac{d^2\psi}{d\xi^2} - \xi^2\psi = 0, \tag{5.38}$$

with solutions

$$\psi = c_1\xi^m e^{\xi^2/2} + c_2\xi^m e^{-\xi^2/2}. \tag{5.39}$$

But this first term grows exponentially as $\xi \to \infty$ (equivalent to $x \to \infty$, so we should exclude that on physical grounds (normalization, if you like).

In the limit $\xi \to 0$, we have instead

$$\frac{d^2\psi}{d\xi^2} + 2\epsilon\psi = 0, \tag{5.40}$$

with solutions

$$\psi = a_1\cos(\sqrt{2\epsilon}\xi) + a_2\sin(\sqrt{2\epsilon}\xi), \tag{5.41}$$

so we haven't learned much from expanding about the origin.

Lecture 6.

Tuesday, October 15, 2019

Last time, we defined the dimensionless variables ξ, ϵ in terms of x and E in order to write the Schrödinger equation in a dimension-free way:

$$\frac{d^2\psi(\xi)}{d\xi^2} + (2\epsilon - \xi^2)\psi(\xi) = 0. \tag{6.1}$$

We suggested that there was some limiting behavior where as $\xi \to \infty$,

$$\psi(\xi) = c_1\xi^k e^{-\xi^2/2} + c_2\xi^k e^{\xi^2/2}, \tag{6.2}$$

setting $c_2 = 0$ for convergence. We also had some expansion in sines and cosines about the origin $\xi \to 0$.

Let us redefine

$$\psi(\xi) = \phi(\xi) e^{-1/2\xi^2} \tag{6.3}$$

so that substituting (exercise), the Schrödinger equation becomes

$$\frac{d^2\phi}{d\xi^2} - 2\xi \frac{d\phi}{d\xi} + (2\epsilon - 1)\phi(\xi) = 0. \tag{6.4}$$

We shall solve this by the Frobenius method, i.e. by expanding ϕ as a power series in ξ. Let

$$\phi(\xi) = \sum_{n=0}^{\infty} a_n \xi^n, \tag{6.5}$$

so that our equation becomes

$$\sum_{n=0}^{\infty} a_n \, n(n-1)\xi^{n-2} - 2\xi \sum_{n=0}^{\infty} a_n \, n\xi^{n-1} + (2\epsilon - 1) \sum_{n=0}^{\infty} a_n \xi^n = 0. \tag{6.6}$$

Note that the first sum really starts from $n = 2$ since the $n = 0, n = 1$ terms are zero, and the second term is similar. Hence we can combine all these sums:

$$\sum_{j=0}^{\infty} \left[(j+1)(j+2)a_{j+2} - (2j + 1 - 2\epsilon)a_j \right] \xi^j = 0. \tag{6.7}$$

That is, we collect terms in powers of ξ, redefining some indices. Note that on the space of smooth functions defined on $\mathbb{R}$, the functions $\{\xi^j\}_{j=0}^{\infty}$ form a complete, linearly independent basis. That means this equation gives us a recursion relation for the coefficients a_j of ϕ, because the only way for the sum of all ξ^j to be zero is if each of the individual coefficients vanish. Hence

$$a_{j+2} = \frac{2j + 1 - 2\epsilon}{(j+1)(j+2)} a_j. \tag{6.8}$$

For a generic second-order equation, we can always find such a recursion relation which will generally involve a_j, a_{j+1}, a_{j+2}.

Let us note that the relation **??** relates even j coefficients to each other, and it relates odd j coefficients, but it does not mix them. This agrees with our intuition that there should be two linearly independent sets of solutions (the power law exponentials in the large ξ limit and the sines and cosines in the small ξ limit).

Hence

$$\phi(\xi) = a_0 \left[1 + \frac{1 - 2\epsilon}{2!}\xi^2 + \frac{(1-2\epsilon)(5-2\epsilon)}{4!}\xi^4 + \dots \right] + a_1 \left[\xi + \frac{(3-2\epsilon)}{3!}\xi^3 + \frac{(3-2\epsilon)(7-2\epsilon)}{5!}\xi^5 + \dots \right]. \tag{6.9}$$

Notice that as $\xi \to 0$, $\phi(\xi)$ is determined by a_0, a_1, which are otherwise unconstrained. To find what happens as $\xi \to \infty$, we have

$$\frac{a_{j+2}}{a_j} = \frac{2j + 1 - 2\epsilon}{(j+1)(j+2)} \approx \frac{2j}{j^2} \approx \frac{2}{j}, \quad j \gg 1. \tag{6.10}$$

That is, in the large ξ limit (equivalently in the limit of j large), for $j \geq j_c \gg 1$, our solution is approximately

$$a_{j_c} \xi^{j_c} \left(1 + \frac{2}{j_c}\xi^2 + \frac{4}{j_c(j_c+2)}\xi^4 + \dots \right). \tag{6.11}$$

That is, pick some j_c large, look at all terms of power ξ^{j_c} or higher, and WLOG consider just the even or odd terms. Then this solution can be explicitly summed; it is

$$a_{j_c} \xi^{j_c} e^{\xi^2}, \tag{6.12}$$

and this tells us that we actually got back this bad exponential behavior in ξ. This is exactly like the $c_2 \xi^k e^{\xi^2/2}$ term we tried to get rid of. We could have expected this– for a generic solution, we have not yet imposed the physical boundary conditions for convergence.

So how do we do this? Independent of a_0, a_1, we have a divergent solution as $\xi \to \infty$. It must be that our infinite sum actually terminates at some finite j. We have three choices to impose the boundary conditions.

i) We can set $a_0 = a_1 = 0$, the trivial solution.
ii) Let a_0 be arbitrary and set $a_1 = 0$.
iii) Let a_1 be arbitrary and set $a_0 = 0$.

In either case ii or iii, we must set

$$\epsilon = \frac{2n+1}{2} \tag{6.13}$$

so that the series terminates at order n, i.e.

$$a_{n+2} = \frac{2n+1-2(2n+1)/2}{(n+1)(n+2)} a_j = 0. \tag{6.14}$$

Since ϵ is related to the (dimensionful) energy E, this constraint tells us the spectrum of the harmonic oscillator.

Let us redefine

$$\psi(\xi) = \phi(\xi) e^{-\frac{1}{2}\xi^2}, \tag{6.15}$$

where the solutions are $\phi_n(\xi)$ with $\epsilon = n + 1/2$. The particular solutions are *Hermite polynomials* $H_n(\xi)$, where the first we are

$$\begin{aligned} H_0(\xi) &= 1 \\ H_1(\xi) &= \xi \\ H_2(\xi) &= 2(2\xi^2 - 1), \end{aligned}$$

where these Hermite polynomials have a generating function

$$e^{-\alpha^2 + 2\alpha\xi} = \sum_{n=0}^{\infty} \frac{H_n(\xi)}{n!} \alpha^n, \tag{6.16}$$

up to normalization. The Hermite polynomials also form a complete basis for smooth functions like the monomials ξ^j.

Restoring units we get some explicit solutions for the harmonic oscillator eigenstates, i.e. the wavefunctions

$$\psi_n(x) = \left(\frac{m\omega}{2^{2n}(n!)^2 \pi\hbar} \right)^{1/4} e^{-\frac{m\omega}{2\hbar}x^2} H_n\left(\sqrt{\frac{m\omega}{\hbar}} x \right) \tag{6.17}$$

satisfy the Schrödinger equation for the harmonic oscillator potential,

$$\hat{H}\psi_n(x) = E_n \psi_n(x), \tag{6.18}$$

with an energy spectrum given by

$$E_n = (n + 1/2)\hbar\omega. \tag{6.19}$$

Looking at the $n = 0$ ground state, we see it is just a (normalized) Gaussian. However, it differs in one important way from the classical solution. For a solution with energy $E_0 = \hbar\omega/2$, a classical particle would never escape the potential well, i.e. it must stay within $|x| \leq \sqrt{\hbar/m\omega}$. But the quantum solution (and more importantly its modulus squared, $|\psi_n(x)|^2$) is nonzero in the classically forbidden region. So there is an (exponentially damped) probability of observing the particle outside the classically forbidden region. Note also that the ground state is node-free, $\psi_0 > 0$ everywhere. More generally, the nth excited state has n zeros corresponding to $H_n(\xi) = 0$.

Here's another takeaway from this problem. The discreteness of energy levels defines the "quantization" of the system, and this is generally true for a confining potential. There are no nodes in the ground state, and the energy levels have some quantization condition. Moreover, there exists an isomorphism of Hilbert spaces between such confining potentials, i.e. the wavefunctions are $\psi_m(x), m = 0, 1, 2, \ldots$ and form some countably infinite set.

Note that confining potentials differ from potentials that vanish at infinity. Potentials of the latter kind will have not just bound states but also scattering states which are not normalizable and will generally form a continuous spectrum.

The algebraic method The Frobenius method requires us to know how to solve differential equations. The series method is very general, but there is another method that will prove useful.[26] Let us define creation and annihilation operators

$$a = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}\hat{x} + \frac{i}{\sqrt{\hbar m\omega}}\hat{p}\right), \tag{6.20}$$

$$a = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}\hat{x} - \frac{i}{\sqrt{\hbar m\omega}}\hat{p}\right). \tag{6.21}$$

Note that $x = x^\dagger, p = p^\dagger$, so

$$[x, p] = i\hbar \implies [a, a^\dagger] = 1. \tag{6.22}$$

Hence we can "factor" the Hamiltonian. Writing

$$\hat{H} = \hbar\omega\left(\frac{p^2}{2m\hbar\omega} + \frac{m\omega}{2\hbar}x^2\right), \tag{6.23}$$

we see that in terms of $a, a^\dagger$ our Hamiltonian becomes

$$\begin{aligned} H &= \frac{\hbar\omega}{2}(a^\dagger a + aa^\dagger) \\ &= \hbar\omega(a^\dagger a + 1/2), \end{aligned} \tag{6.24}$$

where we have used the commutation relation $aa^\dagger - a^\dagger a = 1$ in the second line.

Let us now define the *number operator*

$$N \equiv a^\dagger a. \tag{6.25}$$

It commutes with the Hamiltonian, $[H, N] = 0$, so we may as well find the eigenstates of N since they will be eigenstates of H, with some shifted eigenvalues. Moreover

$$\begin{aligned} [N, a] = [a^\dagger a, a] &= a^\dagger aa - aa^\dagger a \\ &= (a^\dagger a - aa^\dagger)a = -a. \end{aligned} \tag{6.26}$$

Similarly

$$[N, a^\dagger] = a^\dagger. \tag{6.27}$$

Of course, these two relations also hold for the Hamiltonian, since everything commutes with scalars:

$$[H, a] = -\hbar\omega a, \quad [H, a^\dagger] = \hbar\omega a^\dagger. \tag{6.28}$$

That is, acting on some eigenstate $|n\rangle$ of energy E_n, $a^\dagger|n\rangle$ produces a new eigenstate of energy $E_n + \hbar\omega$ and similarly $a|n\rangle$ produces an eigenstate of energy $E_n - \hbar\omega$.

For suppose that N has an eigenvector $|\nu\rangle$ with eigenvalue ν. That is,

$$N|\nu\rangle = \nu|\nu\rangle. \tag{6.29}$$

Then $a|\nu\rangle$ is also an eigenvector of N:

$$N(a|\nu\rangle) = Na|\nu\rangle = a(N-1)|\nu\rangle = (\nu - 1)a|\nu\rangle. \tag{6.30}$$

Moreover

$$(a|\nu\rangle)^\dagger = \langle\nu|a^\dagger, \tag{6.31}$$

so

$$||a|\nu\rangle||^2 = \langle\nu|a^\dagger a|\nu\rangle = \langle\nu|N|\nu\rangle = \nu\langle\nu|\nu\rangle. \tag{6.32}$$

But the norm of this new state depends on ν, so we must be careful. By repeatedly applying a we get a state $a^k|\nu\rangle$ such that

$$N(a^k|\nu\rangle) = (\nu - k)a^k|\nu\rangle. \tag{6.33}$$

But since the norm depends on ν, which can be lowered, there must be some state $a^k|\nu\rangle$ where the norm becomes negative. We must stop, or else $\nu - k$ could become (arbitrarily) negative. All states in our Hilbert space must have non-negative norm.

[26] Especially if you ever do quantum field theory.

Hence there exists some state that is annihilated by a, i.e.

$$a|0\rangle = 0, \tag{6.34}$$

and it follows that

$$N|0\rangle = 0. \tag{6.35}$$

Is there a ceiling? By the same manipulations, we find that

$$Na^\dagger|\nu\rangle = (\nu+1)a^\dagger|\nu\rangle, \tag{6.36}$$

so we can construct arbitrarily high number eigenstates with the raising operator. So there is no ceiling in the eigenspectrum.[27] We may write

$$Na^\dagger|0\rangle = 1a^\dagger|0\rangle, \tag{6.37}$$

so we see that $a^\dagger|0\rangle$ is an eigenstate of N with eigenvalue 1. Define

$$|1\rangle = a^\dagger|0\rangle \tag{6.38}$$

and more generally

$$|k\rangle = \frac{1}{\sqrt{k!}}(a^\dagger)^k|0\rangle, \tag{6.39}$$

where the factorial accounts for normalization.

In general,

$$N|k\rangle = k|k\rangle, \quad a|k\rangle = \sqrt{k}|k-1\rangle, \quad a^\dagger|k\rangle = \sqrt{k+1}|k+1\rangle. \tag{6.40}$$

We now see that N has an eigenspectrum in integer steps starting from $|0\rangle$ and going up to arbitrarily high $|k\rangle$.

But now we're done,[28] because H has the same eigenstates as N.

$$H|k\rangle = \hbar\omega(k+1/2)|k\rangle. \tag{6.41}$$

To construct the wavefunctions in position basis, just use

$$a = \frac{1}{\sqrt{2}}(\xi + \frac{d}{d\xi}), \quad a^\dagger = \frac{1}{\sqrt{2}}(\xi - \frac{d}{d\xi}), \tag{6.42}$$

where $\xi = \sqrt{m\omega/\hbar}x$, and then

$$a|0\rangle \implies (\xi + \frac{d}{d\xi})\psi(\xi) = 0 \implies \psi_0(\xi) = a_0 e^{-\xi^2/2}. \tag{6.43}$$

This is a much easier differential equation to solve than the second-order one from earlier.[29] And now we can construct the other states using

$$|k\rangle = \frac{1}{k!}(a^\dagger)^k|0\rangle, \tag{6.44}$$

which gives (explicitly)

$$\psi_k(\xi) = \frac{1}{\sqrt{k!}}\frac{1}{2^{k/2}}\left(\xi - \frac{d}{d\xi}\right)^k e^{-\xi^2/2}, \tag{6.45}$$

and the Hermite polynomials are

$$H_n(\xi) = e^{\xi^2/2}\left(\xi - \frac{d}{d\xi}\right)^n e^{-\xi^2/2}. \tag{6.46}$$

[27] But contrast the angular momentum operator L_z (and its operators L^+, L^-) later. L_z has both minimum and maximum eigenvalues, so the ladder is finite.

[28] We haven't yet shown that the ground state is unique, but we can show that because this space of solutions is isomorphic to the solutions we constructed analytically, the ground state is in fact unique. We could also prove this by contradiction. But see the next footnote for a nice way to do it.

[29] Actually, since this equation is first-order, we can fix a_0 by normalization (up to a complex phase), which completely determines the ground state. This proves uniqueness of the ground state.

Lecture 7.

Thursday, October 17, 2019

Today we'll wrap up our discussion of the harmonic oscillator and then move on to more general topics.

Coherent states Last time, we talked about the creation and annihilation operators. Recall that

$$a = \alpha_1 \hat{x} + \alpha_2 i\hat{p}, \quad a^\dagger = \alpha_1 \hat{x} - \alpha_2 i\hat{p} \tag{7.1}$$

for real constants α_1, α_2. While a and $a^\dagger$ are not hermitian, perhaps eigenstates exist. Let us consider the state

$$|\zeta\rangle = e^{\zeta a^\dagger}|0\rangle \tag{7.2}$$

$$= \left[\sum_{n=0}^{\infty} \frac{\zeta^n}{n!}(a^\dagger)^n\right]|0\rangle \tag{7.3}$$

$$= \sum_{n=0}^{\infty} \frac{\zeta^n}{\sqrt{n!}}|n\rangle. \tag{7.4}$$

We now claim that $|\zeta\rangle$ is an eigenstate of a. We can show this explicitly:

$$a|\zeta\rangle = \sum_{n=0}^{\infty} \frac{\zeta^n}{\sqrt{n!}}|n\rangle = \sum_{n=1}^{\infty} \frac{\zeta^n}{\sqrt{(n-1)!}}|n-1\rangle = \zeta \sum_{n=1}^{\infty} \frac{\zeta^{n-1}}{\sqrt{(n-1)!}}|(n-1)\rangle = \zeta|\zeta\rangle. \tag{7.5}$$

Hence $|\zeta\rangle$ is an eigenstate of a with complex eigenvalue ζ, and this is fine since a is not hermitian. Similarly $\langle 0|e^{\zeta^* a}$ has the interpretation of being an eigenstate of $a^\dagger$ in the sense that

$$\left(\langle 0|e^{\zeta^* a}\right)a^\dagger = \zeta^*\left(\langle 0|e^{\zeta^* a}\right). \tag{7.6}$$

Moreover, the coherent states form a complete basis for the Hilbert space, albeit a very different-looking one than the energy eigenstates. Curiously, it looks as though we have gotten too many states (i.e. these are parametrized by an uncountable infinity of complex parameters ζ, whereas the number states are a countable basis). But we will show that we can write a resolution of the identity in terms of the coherent states, which tells us that we indeed have a basis.

Notice that a has ket eigenvectors, whereas $a^\dagger$ has bra eigenvectors. The converse is not true. The operator a only has *right eigenvectors* but no *left eigenvectors*, i.e. $\nexists \langle\mu| a = f(\mu)\langle\mu|$, and an equivalent statement holds for $a^\dagger$.[30] This is different from the position operator (which is hermitian), such that

$$\hat{x}|x\rangle = x|x\rangle \implies \langle x|\hat{x} = x\langle x|. \tag{7.7}$$

Note this is generally true for any hermitian operator, since $A|v\rangle = \lambda|v\rangle \implies \langle v|A = \langle v|A^\dagger = \lambda^*\langle v| = \lambda\langle v|$.

Let us also remark that

$$\langle\zeta_1|\zeta_2\rangle = \langle 0|e^{\zeta_1^* a}e^{\zeta_2 a^\dagger}|0\rangle. \tag{7.8}$$

How do we evaluate this? We cannot simply add the exponentials since $a, a^\dagger$ do not commute. However, if

$$[A,[A,B]] = [B,[A,B]] = 0 \tag{7.9}$$

then

$$e^A e^B = e^B e^A e^{[A,B]}. \tag{7.10}$$

We haven't proved this formula but maybe later. More generally,

$$e^A e^B = e^C \tag{7.11}$$

where

$$C = A + B + \frac{1}{2}[A,B] + \frac{1}{12}[A,[A,B]] - \frac{1}{12}[B,[A,B]] + \ldots \tag{7.12}$$

[30] Why is this true? Consider writing a general state as a linear combination of the $|n\rangle$ states, and then notice that $a^\dagger$ cannot create the vacuum state. Hence the coefficient of the vacuum state must be zero. And the next coefficient up must therefore be zero. And so on. With a, all the states come down; with $a^\dagger$ we always miss the vacuum $|0\rangle$.

where the ... indicates higher commutators.[31] This is the Baker-Campbell-Hausdorff (BCH) formula.

Let us observe that since $[a, a^\dagger] = 1$, we have $[a, [a, a^\dagger]] = [a^\dagger, [a, a^\dagger] = 0$, so

$$e^{\zeta_1^* a} e^{\zeta_2 a^\dagger} = e^{\zeta_2 a^\dagger} e^{\zeta_1 * a} e^{[\zeta_2 a, \zeta_1 a^\dagger]}, \tag{7.13}$$

where this commutator term is $e^{\zeta_1^* \zeta_2}$, just a c-number. Hence

$$\langle 0|e^{\zeta_1^* a} e^{\zeta_2 a^\dagger}|0\rangle = \langle 0|e^{\zeta_2 a^\dagger} e^{\zeta_1^* a}|0\rangle e^{\zeta_1^* \zeta_2} \tag{7.14}$$

$$= e^{\zeta_1^* \zeta_2} \langle 0| \left(1 + \zeta_2 a^\dagger + \frac{1}{2!}\zeta_2^2 (a^\dagger)^2 + \dots\right)\left(1 + \zeta_1^* a + \frac{1}{2!}(\zeta_1^*)^2 a^2 + \dots\right)|0\rangle \tag{7.15}$$

$$= e^{\zeta_1^* \zeta_2} \langle 0|0\rangle = e^{\zeta_1^* \zeta_2}. \tag{7.16}$$

Rule of thumb working with harmonic oscillators: put all annihilation operators on the right and all the creation operators on the left.[32]

As we've said, the $\{|\zeta\rangle\}$ states for $\zeta \in \mathbb{C}$ are an overcomplete set, but the following identity is true:

$$\mathbb{I} = \frac{1}{2\pi i} \int d^2\zeta \; |\zeta\rangle\langle\zeta| \, e^{-|\zeta|^2}, \tag{7.17}$$

with $d^2\zeta = d\zeta d\zeta^*$.[33] In writing this, we've used the normalization that we just computed, $\langle\zeta|\zeta\rangle = e^{|\zeta|^2}$.

Free particle Now that we have discussed the harmonic oscillator, let us go to a seemingly simpler problem, that of the free particle. The harmonic oscillator is an example of a (regular) confining potential, one where $V(x) \to \infty$ as $|x| \to \infty$. Confining potentials have a discrete spectrum of bound states (those with wavefunction decaying as $|x| \to \infty$).

However, there is another relevant class of states. When particles can escape to infinity, we have *scattering states*. We can understand these by studying the free particle. The free particle has $V = 0$ everywhere, so that the Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m}. \tag{7.18}$$

It's traditional to solve this in the Schrödinger picture, but we're not going to do this just yet. For notice that the Hamiltonian commutes with $\hat{p}$, i.e.

$$[\hat{H}, \hat{p}] = 0. \tag{7.19}$$

Hence momentum eigenstates are also energy eigenstates of H. That is, in the momentum basis,

$$\hat{p}|k\rangle = k|k\rangle, \tag{7.20}$$

where we will require $k \in \mathbb{R}$ in order for certain boundary conditions to be satisfied. Then

$$\hat{H}|K\rangle = \frac{k^2}{2m}|k\rangle, \tag{7.21}$$

which tells us that $E(k) = \frac{k^2}{2m}$. This reminds us of an important lesson– when you have a problem in quantum mechanics, you should think about what basis is best to solve the problem in. This often saves us a lot of time rather than trying to brute-force our way through the Schrödinger equation.

Conversely if we want an energy eigenstate of some energy E, $\hat{H}|E\rangle = E|E\rangle$, then

$$k = \pm\sqrt{2mE}. \tag{7.22}$$

This quadratic relation is forced upon us by the fact we are studying non-relativistic quantum mechanics, where $E = p^2/2m$. Hence a general energy eigenstate can be written as

$$|E\rangle = \alpha_+|k = \sqrt{2mE}\rangle + \alpha_-|k = -\sqrt{2mE}\rangle, \tag{7.23}$$

a linear combination of two momentum eigenstates which correspond to left-moving particles and right-moving particles. Let us now take our eigenvalue expression and sandwich it with $\langle x|$:

$$\langle x|\hat{p}|k\rangle = k\langle x|k\rangle. \tag{7.24}$$

[31] We can also write it more nicely in terms of an adjoint action.

[32] That is, write them in normal ordering. We do this a lot in QFT.

[33] We could probably check this by expressing a general state as a sum of the energy eigenstates.

Equivalently, notice that we can write $\hat{p}$ in the position basis,

$$\langle x|\left(-i\hbar\frac{d}{dx}\right)|k\rangle, \tag{7.25}$$

and as p is self-adjoint we can make it act on the $\langle x|$, namely by rewriting as

$$-i\hbar\frac{d}{dx}\langle x|k\rangle = k\langle x|k\rangle. \tag{7.26}$$

Now $\langle x|k\rangle$ is some function of k and x, with solution

$$\langle x|k\rangle = Ce^{ikx/\hbar}, \tag{7.27}$$

which is simply our plane wave solution. Hence

$$|k\rangle = \int_{-\infty}^{\infty} dx\, |x\rangle\langle x|k\rangle, \tag{7.28}$$

with $\langle x|k\rangle$ as above. We now see that

$$\psi_E(x) = \langle x|E\rangle \tag{7.29}$$

$$= \frac{\alpha_+}{(2\pi\hbar)^{1/2}}e^{\frac{i}{\hbar}\sqrt{2mE}x} + \frac{\alpha_-}{(2\pi\hbar)^{1/2}}e^{-\frac{i}{\hbar}\sqrt{2mE}x}. \tag{7.30}$$

Note that the factor of $1/\sqrt{2\pi\hbar}$ is part of taking the Fourier transform to get back to position space; we can either have factors like this in both the Fourier transform and the inverse transform, or we can keep them all in e.g. the inverse transform. We'll stick with the symmetric normalization for now.

In general, we can now write down a Gaussian wavepacket

$$\psi(x, t=0) = e^{ikx}e^{-\frac{1}{2\delta_x^2}x^2}. \tag{7.31}$$

Notice that $\langle x\rangle = 0$ since this particle is localized to the origin, but it has a nonzero variance,

$$\Delta x = \sqrt{\langle (x-\langle x\rangle)^2\rangle} = \frac{1}{\sqrt{2}}\delta_x. \tag{7.32}$$

Similarly one may compute

$$\langle p\rangle = k\hbar, \quad \Delta p = \frac{\hbar}{\sqrt{2}\delta_x}. \tag{7.33}$$

Hence this looks like a particle initialized at $t=0$ with some momentum $k\hbar$. One may also check readily that

$$\Delta x\Delta p = \frac{\hbar}{2}, \tag{7.34}$$

so Heisenberg uncertainty is saturated.

How does this wavepacket evolve in time? By a unitary operator, as usual.

$$\psi(x,t) = U(t,0)\psi(x,t=0) \tag{7.35}$$

$$= e^{-iHt/\hbar}\psi(x,t=0) \tag{7.36}$$

$$= e^{-\frac{ip^2t}{2m\hbar}}\psi(x,t=0) \tag{7.37}$$

$$= e^{\frac{it}{\hbar}\left(\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\right)}\psi(x,0). \tag{7.38}$$

And we can now study the effect on our Gaussian wavepacket. The solution is

$$\psi(x,t) = \frac{1}{\sqrt{\sqrt{\pi}\delta_x\left(1+\frac{i\hbar t}{m\delta_x^2}\right)}}\exp\left[ik\left(x-\frac{k\hbar t}{2m}\right) - \frac{1}{2\delta_x^2\left(1+\frac{i\hbar t}{m\delta_x^2}\right)}\left(x-\frac{k\hbar t}{m}\right)^2\right]. \tag{7.39}$$

It looks horrible, but it has a nice physical interpretation. What we're seeing is that the peak has moved– it picks up an offset of $\frac{k\hbar t}{m}$, so that

$$\langle x(t)\rangle = \frac{k\hbar t}{m}. \tag{7.40}$$

It also spreads out to have a new width of

$$\Delta x(t) = \frac{1}{\sqrt{2}}\delta_x\sqrt{1+\frac{\hbar^2 t^2}{m^2\delta_x^4}}, \tag{7.41}$$

and in the limit as t grows large, this expression becomes

$$\frac{1}{\sqrt{2}}\frac{\hbar t}{m\delta x}. \tag{7.42}$$

It follows that

$$\langle x\rangle = \frac{\hbar k t}{m} = \frac{\langle p\rangle}{m}t, \tag{7.43}$$

so we see that the expectation values (not operators!) follow classical relations.

As time goes on, we see that the localization of the particle in position fades away, while the total probability to find the particle anywhere remains 1 (as it must under unitary time evolution).

Lecture 8.

Tuesday, October 22, 2019

Last time we discussed the free particle wavefunctions, which are scattering states in $\mathbb{R}^n$. That is, for $\mathcal{H} = \frac{\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}}{2m}$, our solutions are

$$\psi_{\mathbf{k}}(\mathbf{x}) = \frac{1}{(2\pi\hbar)^{d/2}}\exp(i\mathbf{k}\cdot\mathbf{x}), \tag{8.1}$$

with energy given by

$$E(k) = \frac{\mathbf{k}\cdot\mathbf{k}}{2m}\hbar^2. \tag{8.2}$$

The unitary of time evolution guarantees conservation of probability. We say that given a wavefunction $\psi(\mathbf{x},t)$, the quantity $|\psi(\mathbf{x},t)|^2$ gives the probability density to find the particle in the interval (region) $[\mathbf{x},\mathbf{x}+d\mathbf{x}]$,

$$\rho(\mathbf{x},t) = |\psi(\mathbf{x},t)|^2. \tag{8.3}$$

Unitarity says that

$$\psi(\mathbf{x},t) = U(\mathbf{x},t;\mathbf{x}',t')\psi(\mathbf{x}',t'). \tag{8.4}$$

Usually we just indicate the time dependence of U. Since $UU^\dagger = \mathbb{I}$, our probability integrated over some region D is

$$\int_D \rho(\mathbf{x},t)d^dx = \int_D |\psi(\mathbf{x},t)|^2 d^dx. \tag{8.5}$$

Notice now that

$$\frac{d}{dt}\int_D d^dx\,|\psi(\mathbf{x},t)|^2 = \int_D d^dx\left(\psi^*\frac{\partial\psi}{\partial t}+\psi\frac{\partial\psi^*}{\partial t}\right) \tag{8.6}$$

$$= \int_D d^dx\left[\psi^*\left(\frac{i\hbar}{2m}\nabla^2 - i\hbar V(\mathbf{x})\right)\psi + \psi\left(-\frac{i\hbar}{2m}\nabla^2 + i\hbar V(\mathbf{x})\right)\psi^*\right], \tag{8.7}$$

where in the third line we have just used Schrödinger's equation (or its complex conjugate) to replace the time derivatives. Now the potential terms cancel, since they are just multiplication by some numbers. What remains is

$$\frac{d}{dt}\int_D d^dx = \frac{i\hbar}{2m}\int_D d^dx\left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right) \tag{8.8}$$

$$= \frac{i\hbar}{2m}\int_D d^dx\,\boldsymbol{\nabla}\cdot(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*). \tag{8.9}$$

Defining a current, or *probability flux density*, by

$$\mathbf{J}(\mathbf{x},t) = -\frac{i\hbar}{2m}(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*), \tag{8.10}$$

we see that

$$\int_D d^d x \left(\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{J} \right) = 0 \implies \frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{J} = 0. \tag{8.11}$$

This gives us a *local statement of probability conservation*. The probability to find the particle at some location generally changes in time, but the change in its probability is directly related to the divergence of the probability current. In integral form, the probability to find the particle in some region decreases by the flux of the probability current through the boundary of the region. For notice that

$$\int_D d^d x \, \boldsymbol{\nabla} \cdot \mathbf{J} = \int_{\partial D} \mathbf{J} \cdot d\mathbf{S}. \tag{8.12}$$

Global conservation laws are true, but they are hard to impose because they require us to "sample" all of space. Local conservation laws are much more useful in deriving boundary conditions because we only need to check them point by point.

Now consider plane wave solutions,

$$\psi(\mathbf{x}, t) = A e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{8.13}$$

Then the probability current is

$$\mathbf{J} = |A|^2 \frac{\hbar}{m} \mathbf{k}, \tag{8.14}$$

and we see that the probability flux for a momentum eigenstate is given by the momentum eigenvalue. That is, probability flows constantly in the $\mathbf{k}$ direction and its divergence is zero.

Example 8.15. Consider now

$$\psi(\mathbf{x}, t) = A_1 e^{i\mathbf{k}_1\cdot\mathbf{x}} + A_2 e^{i\mathbf{k}_2\cdot x}. \tag{8.16}$$

What do we think the probability current will be? We expect there will be pure $\mathbf{k}_1, \mathbf{k}_2$ terms. But since this is a superposition, there will in general be interference terms:

$$\mathbf{J}(\mathbf{x}, t) = \frac{\hbar}{m} \Big[|A_1|^2 \mathbf{k}_1 + |A_2|^2 \mathbf{k}_2 + (\mathbf{k}_1 + \mathbf{k}_2) \{ \mathrm{Re}(A_1 A_2^*) \cos((\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{x}) - \mathrm{Im}(A_1 A_2^*) \sin(\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{x} \} \Big]. \tag{8.17}$$

The cross-terms represent interference.

Recall our 1D free particle time-dependent solution (wavepacket),

$$\psi(x, t) \propto \exp ik \left(x - \frac{k\hbar t}{2m} \right) - \frac{1}{2\delta_x^2 \left(1 + \frac{i\hbar t}{m\delta_x^2} \right)} \left(x - \frac{k\hbar t}{m} \right)^2. \tag{8.18}$$

Its expected location is

$$\langle x(t) \rangle = \frac{1}{m} k\hbar t, \tag{8.19}$$

and its width is

$$\Delta x = \frac{1}{\sqrt{2}} \delta_x \sqrt{1 + \frac{\hbar^2 t^2}{m^2 \delta_x^4}}. \tag{8.20}$$

This drift in $\langle x(t) \rangle$ comes from the probability flux. Where we are most likely to find the particle drifts and flows, but what is conserved is the area under the (square of the) wavefunction, which is normalized to 1.

Tunneling under a barrier Suppose we have a potential which looks like a step,

$$V(x) = \begin{cases} 0 & x < 0 \text{ or } x > a \\ V_0 & 0 \le x \le a. \end{cases} \tag{8.21}$$

In the left region $x < 0$, we have plane wave scattering states

$$\psi(x) = A_L e^{ikx} + B_L e^{-ikx}, \tag{8.22}$$

and similarly on the right, $x > a$,

$$\psi(x) = A_R e^{ikx} + B_R e^{-ikx}. \tag{8.23}$$

By convention, we will take k to be positive so that the right-movers are e^{ikx} and the left-movers are e^{-ikx}. (Recall that the time dependence is e^{-iEt}, so the signs work out as we expect.) The energy of such a state is

$$E(k) = \frac{\hbar^2 k^2}{2m} < V_0. \tag{8.24}$$

That is, let us restrict ourselves to considering states which would classically just bounce off the potential barrier and could never make it through. We define

$$\alpha^2 = \frac{2m}{\hbar^2}(V_0 - E). \tag{8.25}$$

Then we claim that in the classically forbidden region ($V_0 > E$), our solutions take the form

$$\psi(x) = Ce^{\alpha x} + De^{-\alpha x}. \tag{8.26}$$

For notice that in the barrier region, the Schrödinger equation takes the form

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V_0\right]\psi(x) = E\psi(x) \implies \frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) = (V_0 - E)\psi(x), \tag{8.27}$$

with $V_0 - E > 0$. This suggests that ψ will look like a sum of exponentials in this region, i.e.

$$\frac{d^2}{dx^2}\psi(x) = \alpha^2\psi(x). \tag{8.28}$$

We can make some arguments now on physical grounds. Suppose we send in some wave coming from the left (i.e. right-moving) with energy $E < V_0$. We expect there will be some reflected (left-moving) wave in this region as well. Notice that for a 1D wavefunction

$$\psi(x,t) = Ae^{ikx} + Be^{-ikx}, \tag{8.29}$$

the probability current (cf. Eqn **??**) simplifies to

$$J(x,t) = \frac{\hbar}{m}(|A|^2 - |B|^2)k. \tag{8.30}$$

Now we cannot quite impose the Schrödinger equation at the boundary of the classically forbidden region, but we can demand that the probability current J is continuous along boundaries, and since

$$\mathbf{J}(\mathbf{x},t) = -\frac{i\hbar}{2m}(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*), \tag{8.31}$$

this tells us that ψ, $\boldsymbol{\nabla}\psi$ are therefore continuous across boundaries.[34] That is, we may write

$$\psi(x = 0^-) = \psi(x = 0^+) \quad \psi(x = a^-) = \psi(x = a^+)$$
$$\psi'(x = 0^-) = \psi'(x = 0^+) \quad \psi'(x = a^-) = \psi'(x = a^+).$$

We'll write these boundary conditions in a funny way, as a matrix equation:

$$\underbrace{\begin{pmatrix} 1 & 1 \\ ik & -ik \end{pmatrix}}_{M_1}\begin{pmatrix} A_L \\ B_L \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ \alpha & -\alpha \end{pmatrix}}_{M_2}\begin{pmatrix} C \\ D \end{pmatrix}, \quad \underbrace{\begin{pmatrix} e^{\alpha a} & e^{-\alpha a} \\ \alpha e^{+\alpha a} & -\alpha e^{-\alpha a} \end{pmatrix}}_{M_3}\begin{pmatrix} C \\ D \end{pmatrix} = \underbrace{\begin{pmatrix} e^{ika} & e^{-ika} \\ ike^{ika} & -ike^{-ika} \end{pmatrix}}_{M_4}\begin{pmatrix} A_R \\ B_R \end{pmatrix}. \tag{8.32}$$

Written more compactly,

$$M_1\begin{pmatrix} A_L \\ B_L \end{pmatrix} = M_2\begin{pmatrix} C \\ D \end{pmatrix}, \quad M_3\begin{pmatrix} C \\ D \end{pmatrix} = M_4\begin{pmatrix} A_R \\ B_R \end{pmatrix}, \tag{8.33}$$

so that with a bit of matrix manipulation to solve for $\begin{pmatrix} A_L \\ B_L \end{pmatrix}$, we find that

$$\begin{pmatrix} A_L \\ B_L \end{pmatrix} = M_1^{-1}M_2M_3^{-1}M_4\begin{pmatrix} A_R \\ B_R \end{pmatrix} = \mathcal{T}\begin{pmatrix} A_R \\ B_R \end{pmatrix}, \tag{8.34}$$

[34] This second claim is not quite true for e.g. delta function potentials. We'll come back to this loophole later.

where $\mathcal{T}$ is called the *transfer matrix*. Unfortunately, it is rather complicated:

$$\mathcal{T} = \begin{pmatrix} e^{ika}\left[\cosh(\alpha a) + \frac{i}{2}\sin(\alpha a)\left(\frac{\alpha}{k} - \frac{k}{\alpha}\right)\right] & \frac{i}{2}e^{-ika}\sinh(\alpha a)\left(\frac{\alpha}{k} + \frac{k}{\alpha}\right) \\ -\frac{i}{2}e^{-ika}\sinh(\alpha a)\left(\frac{\alpha}{k} + \frac{k}{\alpha}\right) & e^{-ika}\left[\cosh(\alpha a) - \frac{i}{2}\sin(\alpha a)\left(\frac{\alpha}{k} - \frac{k}{\alpha}\right)\right]. \end{pmatrix} \tag{8.35}$$

Now suppose that we only send in a right-moving wave from $x = -\infty$, such that $B_R = 0$. That is, there is no incoming wave from $x = +\infty$. We would like to know the following:

i) What is the reflected wave?
ii) Is there transmission?

In the right region, the transmitted wave would have flux

$$|A_R|^2 \frac{\hbar k}{m} \tag{8.36}$$

since there is just the right-moving solution. In the left region, the incoming plus reflected wave give flux

$$\frac{\hbar k}{m}(|A_L|^2 - |B_L|^2). \tag{8.37}$$

If we define the transmission and reflection coefficients

$$T = \left|\frac{A_R}{A_L}\right|^2, \quad R = \left|\frac{B_L}{A_L}\right|^2, \tag{8.38}$$

then (total) flux conservation guarantees that

$$R + T = 1. \tag{8.39}$$

In fact, this innocuous-looking statement is the statement of unitarity. If we send in one particle, a "piece" of that particle comes back to you and a "piece" of that particle is transmitted through the barrier, at least at the level of probabilities.

In terms of the transfer matrix elements (which relate A_L, B_L, A_R), we have

$$R = \left|\frac{\mathcal{T}_{21}}{\mathcal{T}_{11}}\right|^2, \quad T = \left|\frac{1}{\mathcal{T}_{11}}\right|^2, \tag{8.40}$$

and by plugging in we see that our transmission coefficient is

$$T = \left(1 + \frac{V_0^2}{4E(V_0 - E)}\sinh^2(\alpha a)\right)^{-1} \approx \frac{4E(V_0 - E)}{V_0} e^{-2\frac{\sqrt{2m(V_0 - E)}}{\hbar}a} \tag{8.41}$$

for $\alpha a \gg 1$. We see that for very wide barriers (large a) or very high barriers (large $\alpha^2 \sim V_0 - E$), the probability of tunneling is exponentially suppressed.

Lecture 9.

Thursday, October 24, 2019

Last time, we discussed tunneling through a step potential. There were only scattering states and no bound states (normalizable states with vanishing probability at ∞). The transmission is roughly proportional to the area under the barrier, as

$$T \propto e^{-2\alpha a}, \quad \alpha = \frac{\sqrt{2m(V_0 - E)}}{\hbar}. \tag{9.1}$$

So tall and/or wide barriers will suppress the transmission probability. Note also that for incoming waves with energy higher than the "step," $E - V_0 > 0$, there is still a reflection probability as well as a transmission probability.

So far, we have discussed the bound states of the harmonic oscillator (a confining potential) as well as scattering states (plane wave states) for the free particle. Let us be more careful about what we mean here.

Bound states For potentials which are unbounded at spatial infinity, if

$$V(x) \to |x| \to \infty, \tag{9.2}$$

then the system has bound states such that

$$\lim_{x\to\pm\infty} \psi(x) = 0. \tag{9.3}$$

That is, a bound state is exponentially small outside of some finite region. On the other hand, a general system wil generically have a mix of bound and scattering states. The classic example is an attractive δ-function potential, where

$$V(x) = -V_0\delta(x). \tag{9.4}$$

For this potential, we have a Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} - V_0\delta(x)\psi(x) = E\psi(x). \tag{9.5}$$

We'll come back to the scattering states; they are e^{ikx} as before. But there is exactly one bound state in this problem. And a scattering state has some probability to be trapped and become a bound state. This tells us there's a nontrivial reflection coefficient.

Let's see this explicitly. Consider the bound state(s), since scattering states exist with $E > 0$. How will we solve for the bound state? Let us make the ansatz that our bound state wavefunction must be exponentially damped as the particle goes to $\pm\infty$. That is, for $|x| \to \pm\infty$, ψ takes the form

$$\psi(x) = \begin{cases} A_L e^{\kappa x} + B_L e^{-\kappa x} & x < 0 \\ A_R e^{\kappa x} + B_R e^{-\kappa x} & x > 0, \end{cases} \tag{9.6}$$

with

$$\kappa = \frac{\sqrt{2mE_b}}{\hbar}. \tag{9.7}$$

For notice that away from $x = 0$, the $E < 0$ bound states have some $E = -E_b, E_B > 0$, such that

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = -E_b\psi, \tag{9.8}$$

and therefore the solutions are exponentials. Moreover, $B_L = A_R = 0$ since we want our solutions to decay at infinity (normalizability). But it seems that A_L, B_R are underdetermined, as is E_B. We still need a boundary condition at $x = 0$ (two, actually).

The boundary conditions are as follows:

i) $\psi(x)$ is continuous at the origin for probability flux to be conserved
ii) $\psi'(x)$ is not continuous; its discontinuity will depend on the integral over the delta-function potential.

Continuity immediately gives us that $A_L = B_R$, so

$$\psi(x) = \begin{cases} A_L e^{\kappa x} & x < 0 \\ A_L e^{-\kappa x} & x > 0, \end{cases} \tag{9.9}$$

Suppose now that we integrate the Schrödinger equation between $x = -\epsilon$ and $x = +\epsilon$. Hence

$$\int_{-\epsilon}^{\epsilon} dx \left[-\frac{\hbar^2}{2m}\psi''(x) - V_0\delta(x)\psi(x)\right] = -\int_{-\epsilon}^{\epsilon} dx\, E_b\psi(x). \tag{9.10}$$

The first term gives us

$$-\frac{\hbar^2}{2m}\left[\psi'(\epsilon) - \psi'(-\epsilon)\right] - V_0\psi(0) = -E_b O(\epsilon). \tag{9.11}$$

In the limit as $\epsilon \to 0$, we have

$$-\frac{\hbar^2}{2m}\left[\psi'(0^+) - \psi'\right] - V_0\psi(0) = 0. \tag{9.12}$$

If we plug in our ansatz for $\psi(x)$, we get

$$-\frac{\hbar^2}{2m}[A_L(-\kappa) - A_L(\kappa)] - V_0 A_L = 0, \tag{9.13}$$

which tells us that

$$\kappa = \frac{m}{\hbar^2} V_0. \tag{9.14}$$

This tells us the bound state energy:

$$\kappa = \frac{\sqrt{2mE_b}}{\hbar} \implies E_B = \frac{m}{2\hbar^2} V_0. \tag{9.15}$$

If we fix the normalization of ψ to be $\int_{-\infty}^{\infty} |\psi|^2 dx = 1$, we can uniquely determine A_L:

$$\psi(x) = \begin{cases} \sqrt{\kappa} e^{\kappa x} & x < 0 \\ \sqrt{\kappa} e^{-\kappa x} & x > 0. \end{cases} \tag{9.16}$$

This has interesting physical consequences. For instance, it is responsible for the phenomenon of Anderson localization, where defects in metal disrupt the conductivity of that metal because they force the electron wavefunction to become localized and therefore not free. Note that many-particle systems are different, though. The Anderson localization of many-particle systems violates the "eigenstate thermalization hypothesis," producing unusual interference that destroys conductivity.

For a homework problem, we will consider the transmission and reflection coefficients for a scattering state; the wavefunction is still continuous but the discontinuity in the derivative is defined by the integral of the potential over a vanishingly small region. The scattering will be nontrivial.

What about a finite rectangular well? In such a case, we will have scattering states, but the existence of bound states depends both on the width and height of the well. Bound states have some minimum energy, analogous to the ground state of the infinite square well, so the well must be some minimum depth in order to support bound states.

Note– what if we take the potential to be a derivative of a delta function, $V(x) = -V_0 \delta^{(n)}(x)$? The limit is apparently well-defined but the direct solution is not obvious because the discontinuity in the derivative is also apparently infinite. We can integrate the δ-function term by parts, but it relates the derivative discontinuity to higher derivatives at the origin, which seems ill-defined.[35]

Periodic potentials in 1D Suppose we have a periodic potential with

$$V(x + a) = V(x), \tag{9.17}$$

i.e. it obeys the symmetry of $x \to x + na, n \in \mathbb{Z}$. This is a bit like a 1D lattice.

Claim (Bloch's theorem (Floquet))**.** Wavefunctions solving periodic potentials with period a take the form

$$\psi(x) = e^{iKx} \phi_K(x), \tag{9.18}$$

where

$$\phi_K(x) = \phi_K(x + a), \tag{9.19}$$

where K is called the Bloch momentum and can be taken to lie in the range $-\pi/a \leq K \leq \pi/a$.

That is, the momenta lie in one block of the dual (Fourier-transformed) lattice. This is tantamount to solving the problem on a circle and then arguing that distinct solutions correspond to shifts up to the circle periodicity. Equivalently, an individual solution for the wavefunction is itself periodic up to an overall phase. More generally we will get superpositions of various Bloch momenta K.

We claim that while our problem does not have a continuous symmetry, we can take advantage of the discrete translational symmetry of the lattice. That is, there will be a generator of discrete translations that commutes with the Hamiltonian.

[35] It took Mukund "two minutes" this morning to decide he could not make immediate progress with this problem.

Example 9.20 (Dirac comb)**.** Let us take a potential of this form,

$$V(x) = \frac{\hbar^2}{2m} V_0 \sum_{n=-\infty}^{\infty} \delta(x + na). \tag{9.21}$$

Look at it over the interval $x \in [0, a]$. In such an interval we have free particles,

$$\phi_1 = e^{ikx}, \quad \phi_2 = e^{-ikx}, \tag{9.22}$$

which form linearly independent solutions to the Schrödinger equation. Hence

$$\psi(x) = \alpha_1 e^{ikx} + \alpha_2 e^{-ikx}, \quad 0 \le x \le a. \tag{9.23}$$

Then Bloch's theorem says that in the next interval $a \le x \le 2a$, the wavefunction is

$$\psi(x) = e^{iKa}\left(\alpha_1 e^{ik(x-a)} + \alpha_2 e^{-ik(x-a)}\right). \tag{9.24}$$

We can now impose the boundary conditions. A priori we have four unknowns: α_1, α_2, k, K. At $x = a$, the wavefunction is continuous,

$$\psi(x = a^-) = \psi(x = a^+), \tag{9.25}$$

but its derivative is not,

$$\psi'(x = a^+) = \psi'(x = a^-) + 2V_0\psi(a). \tag{9.26}$$

We get this discontinuity from the same sort of procedure as before, but with a sign flip because the potential is repulsive, not attractive.

Plugging in the wavefunctions, we see that

$$\alpha_1 e^{ika} + \alpha_2 e^{-ika} = e^{iKa}(\alpha_1 + \alpha_2) \tag{9.27}$$

$$ike^{iKa}(\alpha_1 - \alpha_2) = ik(\alpha_1 e^{ika} - \alpha_2 e^{-ika}) + 2V_0(\alpha_1 e^{ika} + \alpha_2 e^{-ika}). \tag{9.28}$$

This can now be written as a matrix equation on the coefficients α_1, α_2 of the form $A\boldsymbol{\alpha} = 0$, and the coefficient matrix A should have vanishing determinant (is non-invertible) in order to get a nontrivial solution for $\boldsymbol{\alpha}$. Using this condition, one finds that

$$\cos Ka = \cos ka + \frac{V_0}{k} \sin ka. \tag{9.29}$$

Notice that the LHS is bounded by ± 1 because of the range of K. This tells us that k is not totally arbitrary. For a wavefunction to exist, the RHS must also be bounded:

$$\left|\cos ka + \frac{V_0}{k} \sin ka\right| \le 1, \tag{9.30}$$

and $E = \frac{\hbar^2 k^2}{2m}$. Solving this for the allowed values of k will give us energy bands with gaps in between, and this is the origin of electronic band structure.

Lecture 10.

Tuesday, October 29, 2019

Logistic note– there is a more challenging homework up this week. The first problem asks us to reason physically about the solutions to a harmonic oscillator with a delta function perturbation. There's also a nice bonus problem to do with the Poschl-Teller potentials.

Let us now derive Bloch's (Floquet's) theorem, which we saw an application of last time. Recall the statement of Bloch's theorem was that the solutions of periodic potentials can be composed of wavefunctions which are also periodic (with the same period) up to a phase.[36]

That is, take $V(x)$ such that

$$V(x + a) = V(x), \quad x \to x + na, n \in \mathbb{Z}. \tag{10.1}$$

[36] Equivalently, we can simultaneously diagonalize the Hamiltonian and the discrete translation operator. This is how it is described in Griffiths.

Say that $\phi_1(x), \phi_2(x)$ are 2 l.i. solutions to

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\phi(x) = E\phi(x). \tag{10.2}$$

Then $\phi_1(x+a), \phi_2(x+a)$ are also solutions– just shift the arguments, and the potential shifts with it. But ϕ_1, ϕ_2 were linearly independent, so the shifted versions must be linear combinations of the original solutions:

$$\phi_1(x+a) = \alpha_{11}\phi_1(x) + \alpha_{12}\phi_2(x) \tag{10.3}$$
$$\phi_2(x+a) = \alpha_{21}\phi_1(x) + \alpha_{22}\phi_2(x). \tag{10.4}$$

If $\psi(x)$ is some such linear combination, then there exists a choice of coefficients such that

$$\psi(x+a) = \lambda\psi(x), \lambda \in \mathbb{C}, |\lambda| = 1. \tag{10.5}$$

In mathematics this is called Floquet's theorem. We can prove this–

$$\psi(x) = A\phi_1(x) + B\phi_2(x) \tag{10.6}$$
$$\begin{aligned}\psi(x+a) &= A\phi_1(x+a) + B\phi_2(x+a) \\ &= A[\alpha_{11}\phi_1(x) + \alpha_{12}\phi_2(x)] + B[\alpha_{21}\phi_1(x) + \alpha_{22}\phi_2(x)].\end{aligned} \tag{10.7}$$

Now if we suppose that

$$\psi(x+a) = \lambda\psi(x) = \lambda[A\phi_1(x) + B\phi_2(x)], \tag{10.8}$$

then by comparison to our expansion **??** we see that

$$A\alpha_{11} + B\alpha_{21} = \lambda A \tag{10.9}$$
$$A\alpha_{12} + B\alpha_{22} = \lambda B. \tag{10.10}$$

The statement this equation has solutions (that the matrix of αs has eigenvalues) is equivalent to the statement that

$$\det\begin{pmatrix}\alpha_{11}-\lambda & \alpha_{21} \\ \alpha_{12} & \alpha_{22}-\lambda\end{pmatrix} = 0. \tag{10.11}$$

The coefficients α_{ij} are fixed by the properties of ϕ_1, ϕ_2. That means we have a quadratic equation for λ which has 2 roots λ_1, λ_2.[37] We therefore have wavefunctions ψ_1, ψ_2 such that

$$\psi_1(x+a) = \lambda_1\psi_1(x), \quad \psi_2(x+a) = \lambda_2\psi_2(x). \tag{10.12}$$

To show these lambdas are of unit norm, $|\lambda_1| = |\lambda_2| = 1$, we need something else. In particular, we need to use normalizability and also prove that $\lambda_1\lambda_2 = 1$. The first part is easier. Notice that

$$\psi_i(x+na) = \lambda_i^n\psi_i(x). \tag{10.13}$$

This suggests that if $|\lambda| \neq 1$, then the wavefunction value grows arbitrarily large as x increases to $\pm\infty$, which breaks normalizability.We conclude that $|\lambda_1| = |\lambda_2| = 1$.

For the second fact, consider the Wronskian[38]

$$W(x) = \psi_1(x)\psi_2'(x) - \psi_1'(x)\psi_2(x). \tag{10.14}$$

It follows from Eqn. **??** that

$$W(x+a) = \lambda_1\lambda_2 W(x). \tag{10.15}$$

But the Wronskian is in fact constant, $W'(x) = 0$, by Schrödinger's equation.[39] Hence

$$W(x+a) = W(x) \implies \lambda_1\lambda_2 = 1. \tag{10.16}$$

[37] Note that the matrix of αs is actually unitary by probability conservation. This is sufficient to let us diagonalize the matrix of αs, i.e. we can indeed solve for the roots λ_1, λ_2.

[38] For a refresher on the Wronskian, see my 204A notes. It tests linear (in)dependence.

[39] This is generally true for two linearly independent solutions to a second-order equation. Taking the derivative, the cross-terms like $\psi_1'\psi_2'$ cancel, and the other terms like $\psi_1''\psi_2$ cancel after we replace second derivatives with their expressions in terms of the original functions.

Combining these two facts we see that

$$\lambda_1 = e^{iKa}, \lambda_2 = e^{-iKa}, \tag{10.17}$$

where K is a real parameter, and WLOG we can restrict

$$-\frac{\pi}{a} \leq K \leq \frac{\pi}{a} \tag{10.18}$$

where a was the period of the potential. Hence we can find solutions to periodic potentials such that

$$\psi(x+a) = e^{\pm iK(x+a)}\phi_K(x+a) = e^{\pm iKa}\psi(x). \tag{10.19}$$

Theorem 10.20 (Bloch). *For $V(x) = V(x+na), n \in \mathbb{Z}$, there exist solutions*

$$\psi(x) = e^{\pm iKx}\phi_K(x), \quad \phi_K(x) = \phi_K(x+a). \tag{10.21}$$

Notice we haven't actually solved the Schrödinger equation yet; we've just said that its solutions can be expressed as linear combinations of wavefunctions which are periodic up to a phase. To find energy eigenvalues, take

$$\psi(x) = A\phi_1(x) + B\phi_2(x), \quad 0 \leq x \leq a. \tag{10.22}$$

That is, it is a linear combination of the basis solutions between 0 and a. We assume nothing about the periodicity of ϕ_1 and ϕ_2. Hence in the next unit cell, it is given by

$$\psi(x) = e^{iKa}[A\phi_1(x-a) + B\phi_2(x-a)], \quad a \leq x \leq 2a, \tag{10.23}$$

where we have just pulled out the overall phase. Continuity of ψ and its derivative at a[40] gives

$$A\phi_1(a) + B\phi_2(a) = e^{iKa}[A\phi_1(0) + B\phi_2(0)], \tag{10.24a}$$

$$A\phi_1'(a) + B\phi_2'(a) = e^{iKa}\left[A\phi_1'(0) + B\phi_2'(0)\right]. \tag{10.24b}$$

For a solution to A and B to exist, the matrix of coefficients must have zero determinant,

$$\det \begin{pmatrix} \phi_1(a) - e^{iKa}\phi_1(0) & \phi_2(a) - e^{iKa}\phi_2(0) \\ \phi_1'(a) - e^{iKa}\phi_1'(0) & \phi_2'(a) - e^{iKa}\phi_2'(0) \end{pmatrix} = 0. \tag{10.25}$$

We can solve this to find

$$\cos Ka = \frac{1}{2W[\phi_1, \phi_2]}\left[(\phi_1(0)\phi_2'(a) + \phi_1(a)\phi_2'(0)) - (\phi_2(0)\phi_1'(a) + \phi_2(a)\phi_1'(0))\right], \tag{10.26}$$

with $W[\phi_1, \phi_2]$ the Wronskian, evaluated at any point in the interval $[0, a]$. Recall it is nonvanishing for any linearly independent functions ϕ_1, ϕ_2. We see that this justifies why we restricted K to lie in $[-\pi/2, \pi/2]$. Moreover, we will generally get certain ranges of K which are allowed and some which are forbidden. This gives rise to a *band structure*, which is a key idea in condensed matter physics, and we saw our first hint of this in the Dirac comb.

It is a remarkable fact that k-space can have a nontrivial topology when we consider the bands of allowed values of k in three-dimensional materials. This is closely related to the phenomenon of topological insulators.

To conclude this class, let us discuss some properties of single-particle wavefunctions.

(a) Ground state wavefunctions do not have nodes, i.e. $\psi(x) \neq 0$ anywhere on the open interval. A rough proof is as follows– suppose our wavefunction has a zero. Then we can construct another function (not necessarily a solution to the Schrödinger equation) which has a lower energy (expectation value of the Hamiltonian), which tells us that the original wavefunction must not have been the ground state.[41] This fact is also true in higher dimensions than 1.

(b) The energy spectrum for bound state wavefunctions in 1D (one spatial dimension) is nondegenerate. The proof is by contradiction– suppose two distinct states of the same energy exist, and then show they must in fact be proportional to each other.

[40] The latter is not true for the Dirac comb because of the delta functions.

[41] This sounds like an example of the variational principle. Anyway, the proof is in the notes. This statement will also be important for the homework.

Lecture 11.

Thursday, October 31, 2019

Today we will introduce density matrices, a different formalism for describing quantum states.

To begin with, consider what makes quantum mechanics so quantum. One answer is the uncertainty principle, which reflects the non-commuting nature of various observables in our theory. However, we will argue that the quantum-ness of QM is in fact better-represented by the phenomenon of entanglement.

That is, the uncertainty principle

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{11.1}$$

is simply a consequence of operator non-commutativity. More generally if A, B are non-commuting then

$$(\Delta A)^2 (\Delta B)^2 \geq \left(\frac{i}{2} \langle [A, B] \rangle \right)^2. \tag{11.2}$$

Let us prove this.

Proof. For (non-commuting Hermitian) operators A and B, consider

$$\tilde{A} = A - \langle \psi | A | \psi \rangle, \tag{11.3}$$

$$\tilde{B} = B - \langle \psi | B | \psi \rangle \tag{11.4}$$

for some state $|\psi\rangle$. Hence

$$(\Delta A)^2 = \langle \psi | \tilde{A}^2 | \psi \rangle, \quad (\Delta B)^2 = \langle \psi | \tilde{B}^2 | \psi \rangle. \tag{11.5}$$

Since $\tilde{A}, \tilde{B}$ are just shifted versions of A and B, it follows that

$$[\tilde{A}, \tilde{B}] = [A, B]. \tag{11.6}$$

Let us consider some new states $|\tilde{\psi}_\lambda\rangle$ defined by

$$|\tilde{\psi}_\lambda\rangle \equiv (\tilde{A} + i\lambda \tilde{B}) |\psi\rangle, \quad \lambda \in \mathbb{R}. \tag{11.7}$$

As all such $|\tilde{\psi}_\lambda\rangle$ are in the Hilbert space,

$$\langle \tilde{\psi}_\lambda | \tilde{\psi}_\lambda \rangle \geq 0 \tag{11.8}$$

and expanding, we have

$$0 \leq \langle \psi | (\tilde{A}^\dagger - i\lambda \tilde{B})(\tilde{A} + i\lambda \tilde{B}) | \psi \rangle \tag{11.9}$$

$$\leq \langle \psi | \tilde{A}^2 | \psi \rangle + \lambda^2 \langle \psi | \tilde{B}^2 | \psi \rangle + i\lambda \langle \psi | [\tilde{A}, \tilde{B}] | \psi \rangle \tag{11.10}$$

$$\leq (\Delta A)^2 + \lambda^2 (\Delta B)^2 + i\lambda \langle [A, B] \rangle. \tag{11.11}$$

Note that the commutator of A, B is *antihermitian*, so its eigenvalues are imaginary. This must hold for any $\lambda \in \mathbb{R}$, so in particular it holds for the λ that extremizes the inequality. Taking a derivative with respect to λ gives

$$\frac{d}{d\lambda} \left[(\Delta A)^2 + \lambda^2 (\Delta B)^2 + i\lambda \langle [A, B] \rangle \right] = 0 \implies 2\lambda_* (\Delta B)^2 + i \langle [A, B] \rangle = 0, \tag{11.12}$$

and therefore

$$\lambda_* = -\frac{i}{2} \frac{\langle [A, B] \rangle}{(\Delta B)^2} \tag{11.13}$$

extremizes the inequality. If we plug in this value λ_* back into our inequality we find that

$$\langle \tilde{\psi}_{\lambda_*} | \tilde{\psi}_{\lambda_*} \rangle = (\Delta A)^2 + \frac{1}{4} \frac{(\langle [A, B] \rangle)^2}{(\Delta B)^2} \geq 0, \tag{11.14}$$

which after rearranging is the generalized uncertainty principle. ⊠

Density matrices Having discussed the uncertainty principle, let us consider now what the defining property of quantum mechanics is. What makes it fundamentally different from classical systems? There are two ways to tackle this. One is to show that quantum probability distributions are intrinsically different from classical probability distribtutions. We'll do this later. But another good way to understand this is through the idea of density matrices (sometimes density operators or state operators), a different formalism for quantum ensembles.

Thus far, we have discussed state vectors $|\psi\rangle$. Recall the corresponding projection operators

$$\mathcal{P}_n = |z_n\rangle\langle z_n| \tag{11.15}$$

which project onto the state vector $|z_n\rangle$. That is,

$$\mathcal{P}_n|\psi\rangle = \langle\zeta_n|\psi\rangle|\zeta_n\rangle. \tag{11.16}$$

This suggests that instead of thinking of the vector in Hilbert space, we can think of the corresponding projector $\mathcal{P}_n$, which belongs to the space of operators, i.e. $\mathcal{P}_n \in \mathcal{H}\otimes\mathcal{H}^*$. That is, we identify[42]

$$|\psi\rangle \leftrightarrow |\psi\rangle\langle\psi|\,. \tag{11.17}$$

Now let us define the following.

Definition 11.18. A state/density operator, traditionally denoted as ρ, is an operator associated to a state which has three properties:

(a) Positivity (i.e. $\rho \geq 0$, in the sense that its eigenvalues are non-negative and therefore $\langle\psi|\rho|\psi\rangle \geq 0$)[43]
(b) Hermiticity ($\rho^\dagger = \rho$)
(c) Normalization ($\mathrm{Tr}(\rho) = 1$).

Let us say what we mean by the trace of ρ. In a (pure) state $\rho = |\psi\rangle\langle\psi|$. we have

$$\begin{aligned}\mathrm{Tr}(\rho) &= \mathrm{Tr}(|\psi\rangle\langle\psi|)\\ &= \sum_n \langle e_n|\psi\rangle\langle\psi|e_n\rangle\\ &= \sum_n \langle\psi|e_n\rangle\langle e_n|\psi\rangle\\ &= \langle\psi|\left(\sum_n |e_n\rangle\langle e_n|\right)|\psi\rangle\\ &= \langle\psi|\mathbb{I}|\psi\rangle = \langle\psi|\psi\rangle.\end{aligned} \tag{11.19}$$

Hence if the original state was normalized, then its density matrices ought to be normalized.

Since ρ is hermitian, it has a spectral decomposition,

$$\rho = \sum_i p_i\,|\psi_i\rangle\langle\psi_i|\,, \tag{11.20}$$

where the $|\psi_i\rangle$s are eigenvectors of ρ, i.e.

$$\rho|\psi_i\rangle = p_i|\psi_i\rangle. \tag{11.21}$$

Then positivity implies $p_i \geq 0$ and hermiticity implies $p_i = p_i^*$.[44] The normalization condition implies $\sum_i p_i = 1$.

The last thing we need is a rule to compute expectation values, and it is

$$\langle A\rangle = \mathrm{Tr}(\rho A). \tag{11.22}$$

For notice that

$$\mathrm{Tr}(\rho A) = \mathrm{Tr}(|\psi\rangle\langle\psi|\,A) = \langle\psi|A|\psi\rangle = \langle A\rangle\,. \tag{11.23}$$

So this rule agrees on the simple class of density matrices for state $\rho = |\psi\rangle\langle\psi|$.

Hence we can define the evolution of the density matrix in time:

$$\rho(t) = |\psi(t)\rangle\langle\psi(t)| = U(t;t_0)\,|\psi(t_0)\rangle\langle\psi(t_0)|\,U(t;t_0)^\dagger = (t,t_0)\rho(t_0)U(t,t_0)^\dagger. \tag{11.24}$$

[42] Formally the operator is an endomorphism, i.e. a map $\mathcal{H}\to\mathcal{H}$, but we won't use this language much.

[43] Strictly it is positive semi-definite.

[44] Hermiticity is kind of logically prior here. It doesn't make much sense to talk about complex eigenvalues being positive.

Notice this is different from how other operators transformed in the Heisenberg picture (as $U^\dagger A U$). This tells us that what we've written down should really be interpreted as a Schrödinger picture *state operator*. Although the density matrix is an operator, this describes the time evolution of a state.

Pure states and mixed states Let us now discuss the difference between *pure states* and *mixed states*.

Definition 11.25. Pure states have density operators that are pure projectors. That is, they correspond to a single state we can write down as a vector in the Hilbert space,

$$\rho = |\psi\rangle\langle\psi|\,. \tag{11.26}$$

Equivalently, the spectral decomposition has one nonvanishing eigenvalue $p_1 = 1$, and its value is fixed by normalization. Hence

$$\rho^2 = \rho. \tag{11.27}$$

A weaker condition that is sometimes useful to test for purity is[45]

$$\mathrm{Tr}(\rho^2) = 1. \tag{11.28}$$

In contrast, there are *mixed states*.

Definition 11.29. *Mixed states* are density operators that are convex combinations of pure states.

That is, ρ has a general spectral decomposition

$$\rho = \sum_i p_i \rho_i^{\text{pure}}, \quad 0 \le p_i \le 1. \tag{11.30}$$

This is the general form of a density operator, and it describes an ensemble of pure states ρ_i^{pure} with given probabilities p_i. Density matrices are statistical mixtures of pure states, so that when we compute an expectation value on a general density matrix, we get

$$\langle A\rangle = \mathrm{Tr}(\rho A) \tag{11.31}$$

$$= \sum_i p_i \,\mathrm{Tr}(\rho_i^{\text{pure}} A) \tag{11.32}$$

$$= \sum_i p_i \langle\psi_i|A|\psi_i\rangle. \tag{11.33}$$

That is, we compute the expectation values on the pure states and then we take a classical average weighted by the classical probabilities p_i.

For mixed states,

$$\mathrm{Tr}(\rho^2) \le 1. \tag{11.34}$$

Moreover, one can define an entropy, the *von Neumann entropy*, by

$$S = -\,\mathrm{Tr}(\rho \log \rho). \tag{11.35}$$

The log of an operator is defined by the spectral decomposition of ρ. If ρ is diagonal as a matrix, then

$$S = -\sum p_i \ln p_i, \tag{11.36}$$

the Shannon entropy of the classical probability distribution $\{p_i\}$.[46]

Let us consider a qubit Hilbert space spanned by the state vectors $\{|0\rangle, |1\rangle\}$. It describes a state in $\mathbb{C}^2$. In fact, let us consider a 2-qubit Hilbert space,

$$\mathcal{H}_{\text{qubit}_1} \otimes \mathcal{H}_{\text{qubit}_2}. \tag{11.37}$$

This is a four-dimensional space[47] with basis

$$|0\rangle\otimes|0\rangle, |0\rangle\otimes|1\rangle, |1\rangle\otimes|0\rangle, |1\rangle\otimes|1\rangle. \tag{11.38}$$

We sometimes write this as

$$|00\rangle, |01\rangle, |10\rangle, |11\rangle \tag{11.39}$$

[45] The product of two density matrices is not always a density matrix! In general it is not.

[46] For more on quantum information, check out my Quantum Information notes on GitHub at https://github.com/qmch/partiii-lecture-notes-201819.

[47] Counting complex dimensions.

omitting the tensor products. The general density operator on a 2-qubit Hilbert space is then a 4×4 matrix, such that there is a sum of operators

$$\rho = a_1 |00\rangle\langle 00| + a_2 |00\rangle\langle 01| + a_3 |00\rangle\langle 10| + a_4 |00\rangle\langle 11| \ldots + a_{13} |11\rangle\langle 00| + \ldots + a_{16} |11\rangle\langle 11| . \tag{11.40}$$

Consider something simpler, the EPR state

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = |\text{EPR}\rangle. \tag{11.41}$$

This is a pure state in the 2-qubit Hilbert space. Suppose we want to compute an operator acting only on the first qubit,

$$\langle A_1 \otimes \mathbb{I}_2 \rangle_{\text{EPR}} . \tag{11.42}$$

Then

$$\langle \text{EPR}| A_1 \otimes \mathbb{I}_2 |\text{EPR}\rangle = \text{Tr}(|\text{EPR}\rangle\langle\text{EPR}| \, (A_1 \otimes \mathbb{I}_2) \tag{11.43}$$

$$= \text{Tr}_1(\rho_1 A_1), \tag{11.44}$$

where we claim this is now the expectation value in terms of a *reduced density matrix* ρ_1. We claim that ρ_1 is a mixed state, where

$$\rho_1 = \text{Tr}_2(|\text{EPR}\rangle\langle\text{EPR}|). \tag{11.45}$$

Here, the subscripts now indicate *partial traces*. Notice that the density operator of the EPR state has four terms,

$$\begin{aligned} |\text{EPR}\rangle\langle\text{EPR}| &= \frac{1}{2}(|00\rangle + |11\rangle)(\langle 00| + \langle 11|) \\ &= \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|). \end{aligned}$$

Hence we can rewrite the overall trace as two partial traces,

$$\text{Tr}_{1\otimes 2}(|\text{EPR}\rangle\langle\text{EPR}| \, (A_1 \otimes \mathbb{I}_2)) = \text{Tr}_1[\text{Tr}_2(|\text{EPR}\rangle\langle\text{EPR}| \, (A - 1 \otimes \mathbb{I}_2))] \tag{11.46}$$

$$= \text{Tr}_1[\langle 0_2|(|\text{EPR}\rangle\langle\text{EPR}| \, (A_1 \otimes \mathbb{I}_2))|0_2\rangle + \langle 1_2|(|\text{EPR}\rangle\langle\text{EPR}| \, (A_1 \otimes \mathbb{I}_2))|1_2\rangle] \tag{11.47}$$

$$= \text{Tr}_1\left[\frac{1}{2} |0_1\rangle\langle 0_1| \, A_1 + \frac{1}{2} |1_1\rangle\langle 1_1| \, A_1\right] = \text{Tr}_1(\rho_1 A_1), \tag{11.48}$$

where

$$\rho_1 = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|). \tag{11.49}$$

Notice that the reduced state ρ_1 is now a mixed state.