PHYSICS 215A: QUANTUM MECHANICS

IAN LIM
LAST UPDATED OCTOBER 16, 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, $+ \to |v\rangle + |w\rangle$,

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

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

- 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 $|v\rangle + |0\rangle = |v\rangle$
- 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

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

[3] Free fields are kind of like this.

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

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 $\{|e_n\rangle\}$.

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

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

[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.

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

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

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

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. 2.9 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. 3.3 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. 5.7 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. 5.5 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 6.8 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.