PHYSICS 215A: QUANTUM MECHANICS

IAN LIM
LAST UPDATED OCTOBER 3, 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 $|v_i\rangle, |v_j\rangle \in V$, $\alpha \in \mathbb{C}$,

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

[1] Well, finite in spin systems and so on.
[2] We won't get into the details of measurement or observation in this course.

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

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

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

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.

[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.

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

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

[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 guaranted 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

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

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.

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