Lecture 3 (Sep. 13, 2017)

3.1 Even More Math

3.1.1 More on Matrix Representations

Last time, we described that in a given basis, there is an exact correspondence between $n \times n$ matrices and operators, where n is the dimension of the Hilbert space. Let $\{|a\rangle\}$ form an orthonormal basis, so that any state $|\alpha\rangle$ can be expanded as

$$|\alpha\rangle = \sum_a \langle a|\alpha\rangle |a\rangle := \sum_a c_{\alpha a}|a\rangle \,. \tag{3.1}$$

Similarly, any operator X can be expressed in the form

$$X = \sum_{a_i,a_j} |a_i\rangle\langle a_i|X|a_j\rangle\langle a_j| \,. \tag{3.2}$$

In this basis, all of the information of the operator X is contained in the matrix

$$X_{ij} = \langle a_i|X|a_j\rangle \,. \tag{3.3}$$

The terminology is that the X_{ij} are *matrix elements* of X between states $|a_i\rangle$ and $|a_j\rangle$. As an exercise, you can check that the matrix elements of the product operator XY are given by $\sum_j X_{ij}Y_{jk}$, where X_{ij} and Y_{jk} are the matrix elements of the operators X and Y in some basis, respectively.

We define the *trace* of an operator A as

$$\operatorname{Tr} A = \sum_i \langle a_i|A|a_i\rangle = \sum_i A_{ii} \,. \tag{3.4}$$

If the $|a_i\rangle$ are chosen to be eigenvectors of A, then $A_{ii} = a_i$, so the trace becomes

$$\operatorname{Tr} A = \sum_i a_i \,. \tag{3.5}$$

This is a statement of the familiar fact that the trace of an operator is the sum of its eigenvalues.

3.1.2 Unitary Transformations

Suppose we are given two orthonormal bases $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$. How are these bases related? We can define an operator U by the action

$$U|a_i\rangle = |b_i\rangle \,. \tag{3.6}$$

This implies that

$$\langle b_i| = \langle a_i|U^\dagger \,, \tag{3.7}$$

by definition of the adjoint.

Note that we can write

$$U = U\mathbb{1} = U\sum_i |a_i\rangle\langle a_i| = \sum_i |b_i\rangle\langle a_i| \,, \tag{3.8}$$

so the operator U is simply the sum of the outer products of corresponding vectors from each basis. Then we have

$$U^\dagger = \sum_i |a_i\rangle\langle b_i|\,, \tag{3.9}$$

from which we find that

$$
\begin{aligned}
UU^\dagger &= \sum_{i,j} |b_i\rangle\langle a_i|a_j\rangle\langle b_j| \\
&= \sum_i |b_i\rangle\langle b_i| \\
&= \mathbb{1}\,.
\end{aligned}
\tag{3.10}
$$

Thus, we see that U is unitary. Unitary transformations are precisely those that transform from one orthonormal basis to another.

Consider now a vector $|\alpha\rangle$ and two distinct bases $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$. We can express $|\alpha\rangle$ in two ways as

$$|\alpha\rangle = \sum_i c_i|a_i\rangle = \sum_i d_i|b_i\rangle\,. \tag{3.11}$$

How are these two sets of coefficients $\{c_i\}$ and $\{d_i\}$ related? Using the definition of U, we can write

$$
\begin{aligned}
|\alpha\rangle &= \sum_j d_j|b_j\rangle \\
&= \sum_j d_j U|a_j\rangle \\
&= \sum_{i,j} d_j|a_i\rangle\langle a_i|U|a_j\rangle\,.
\end{aligned}
\tag{3.12}
$$

Thus, we see that

$$c_{ij} = \sum_{i,j}\langle a_i|U|a_j\rangle d_j = U_{ij}d_j\,, \tag{3.13}$$

where we have introduced the shorthand notation in which repeated indices are assumed to be summed over (from now on, we will explicitly state the cases when we are not using this convention). Note that the U_{ij} are simply the matrix elements of the operator U.

A similar approach can be used to show that for any operator that can be expressed in two different bases as

$$X = \sum_i |a_i\rangle X_{ij}\langle a_j| = \sum_{k\ell} |b_k\rangle Y_{k\ell}\langle b_\ell|\,, \tag{3.14}$$

the matrix elements in the two different bases are related by

$$X_{ij} = U_{ik}Y_{k\ell}U^\dagger_{\ell j}\,. \tag{3.15}$$

We leave the proof as an exercise.

3.1.3   Diagonalization of Hermitian Operators

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

Proof. Consider a general orthonormal basis $\{|\phi_i\rangle\}$, and let $\{|h_i\rangle\}$ be the orthonormal basis of eigenstates of the operator H — such a basis exists because H is Hermitian. Because these are both orthonormal bases, there exists a unitary transformation U such that $|h_i\rangle = U|\phi_i\rangle$. Using this operator, we can then write

$$\delta_{ij}h_i = \langle h_i|H|h_j\rangle = \langle \phi_i|U^\dagger H U|\phi_j\rangle\,, \tag{3.16}$$

where the first expression on the left is *not* summed over i. Thus, we see that $U_{ik}^\dagger H_{k\ell} U_{\ell j}$ is a diagonal matrix. This completes the proof. □

3.1.4   Simultaneous Diagonalization

Theorem 3. *Two (diagonalizable) operators A, B are simultaneously diagonalizable if and only if $[A, B] = 0$, where $[\cdot, \cdot]$ is the commutator.*

Proof. Let there be a basis $\{|a_i\rangle\}$ for which the diagonalizable operators A, B have

$$A|a_i\rangle = a_i|a_i\rangle\,, \quad B|a_i\rangle = b_i|a_i\rangle\,. \tag{3.17}$$

This is a basis in which A and B are simultaneously diagonal. In this case, we see that

$$AB|a_i\rangle = a_i b_i|a_i\rangle = BA|a_i\rangle\,, \tag{3.18}$$

and so $AB = BA$. This proves the forward direction.

Now consider two diagonalizable operators A, B that commute, $AB = BA$, and let $\{|a_i\rangle\}$ be a basis of eigenvectors of A,

$$A|a_i\rangle = a_i|a_i\rangle\,. \tag{3.19}$$

Then we have

$$\begin{aligned}
A(B|a_i\rangle) &= AB|a_i\rangle \\
&= BA|a_i\rangle \\
&= Ba_i|a_i\rangle \\
&= a_i(B|a_i\rangle)\,,
\end{aligned} \tag{3.20}$$

so $B|a_i\rangle$ is an eigenket of A with eigenvalue a_i. In general, this means that B is block diagonal in the basis $\{|a_i\rangle\}$ (once we have ordered the basis so as to group eigenkets of A with the same eigenvalue), with each block corresponding to a single eigenvalue of A. We can then diagonalize B within each block, which will leave A diagonal, thereby simultaneously diagonalizing A and B. Visually, we have

$$A = \begin{pmatrix} a_1 & & & & & & \\ & \ddots & & & & & \\ & & a_1 & & & & \\ & & & a_2 & & & \\ & & & & \ddots & & \\ & & & & & a_2 & \\ & & & & & & \ddots \end{pmatrix}\,, \quad B = \begin{pmatrix} B_1 & & \\ & B_2 & \\ & & \ddots \end{pmatrix}\,, \tag{3.21}$$

where the B_i are $k_i \times k_i$ blocks, with k_i the number of occurrences of the eigenvalue a_i on the diagonal of A. Diagonalizing the block B_i only mixes eigenkets of A with the same eigenvalue, so we can diagonalize B while leaving A diagonal. This proves the backward direction. □

3.2   Measurement

Consider a quantum mechanical system. We know that the state of such a system consists of a normalized state $|\psi\rangle \in \mathcal{H}$, and observables are represented by Hermitian operators acting in the Hilbert space $\mathcal{H}$. Recall the third postulate, regarding measurement:

a. The possible results of measuring A are the eigenvalues of a_i of A.

b. Once a measurement is done, and the result is a_i for some i, the system "collapses" to an eigenket $|a_i\rangle$ with a sharp value of A: if we perform a measurement again instantaneously, we are guaranteed to get the same value as the first measurement.

c. The probability that a measurement gives the result $A = a_i$ is given by

$$\begin{aligned}
\text{Prob}(A = a_i) &= \sum_{j:a_j=a_i} |\langle a_j|\psi\rangle|^2 \\
&= \sum_{j:a_j=a_i} \langle\psi|a_j\rangle\langle a_j|\psi\rangle \\
&= \langle\psi|M_{a_i}|\psi\rangle,
\end{aligned} \tag{3.22}$$

where we have defined the *measurement operator*

$$M_{a_i} := \sum_{j:a_j=a_i} |a_j\rangle\langle a_j|. \tag{3.23}$$

This operator is the projector onto the subspace with $A = a_i$.

If we simplify to the case where there is only one eigenket with eigenvalue a_i. Then

$$\text{Prob}(A = a_i) = |\langle a_i|\psi\rangle|^2. \tag{3.24}$$

Historically, there has been a lot of worry about the collapse of the wavefunction. In the modern language, there is a way to understand this collapse in a very palatable way, which we may discuss later in the course.

The concept of measurement is worth pondering. Consider making a measurement on a state, and then immediately making another measurement on it. Can we actually do this? Say our measurement is to see if there is a photon hitting our detector. Prior to measurement, the state consists of a detector and a photon. After measurement, there is a click, and the photon is gone. How can we remeasure? We must separately consider measurements that destroy our state and ones that do not. Measurements that do not destroy the state are called "non-demolition measurements."

3.2.1   Comments

1. We have said that

$$\text{Prob}(A = a_i) = \sum_{j:a_j=a_i} |\langle a_j|\psi\rangle|^2. \tag{3.25}$$

This can only make sense if the value on the right-hand side is non-negative, and the sum of these probabilities over all eigenvalues is 1. It is clear that the right-hand side is non-negative,

as is the sum of non-negative numbers, and we see that

$$
\begin{aligned}
\sum_i \text{Prob}(A = a_i) &= \sum_j |\langle a_j | \psi \rangle|^2 \\
&= \sum_j \langle \psi | a_j \rangle \langle a_j | \psi \rangle \\
&= \langle \psi | \psi \rangle \\
&= 1 \,,
\end{aligned}
\tag{3.26}
$$

because the state $|\psi\rangle$ is assumed to be normalized.

2. For any observable A and state $|\psi\rangle$, the *expectation value* of A is

$$
\begin{aligned}
\langle A \rangle &:= \sum_{a_i} a_i \, \text{Prob}(A = a_i) \\
&= \sum_{a_i} a_i \sum_{j : a_j = a_i} \langle \psi | a_j \rangle \langle a_j | \psi \rangle \\
&= \langle \psi | A | \psi \rangle \,,
\end{aligned}
\tag{3.27}
$$

where in the final line we have used

$$
A = \sum_{a_i} a_i | a_i \rangle \langle a_i | \,.
\tag{3.28}
$$

3.3   Spin-$\frac{1}{2}$ Systems

Consider a spin-$\frac{1}{2}$ system. The state space is spanned by the eigenstates of, for example, S^z. We denote the state with $S^z = \frac{\hbar}{2}$ by $|+\rangle$ and the state with $S^z = -\frac{\hbar}{2}$ by $|-\rangle$. As a set, the Hilbert space is then

$$
\mathcal{H} = \left\{ |\psi\rangle = c_+ |+\rangle + c_i |-\rangle \mid c_\pm \in \mathbb{C}, |c_+|^2 + |c_-|^2 = 1 \right\} .
\tag{3.29}
$$

This is a subspace of the two-dimensional complex vector space $\mathbb{C}^2$.

The states $|\psi\rangle$ and $e^{i\lambda}|\psi\rangle$, for some $\lambda \in \mathbb{R}$, are physically equivalent by definition (the Hilbert space is actually the space above quotiented by this equivalence relation), so the only physically relevant phase information in a state is the relative phase of the coefficients (c_+, c_-). We can then parameterize these coefficients in the form

$$
c_+ = \cos\frac{\theta}{2}, \quad c_- = e^{i\phi}\sin\frac{\theta}{2},
\tag{3.30}
$$

with $0 \le \theta < \pi$ and $0 \le \phi < 2\pi$. Specifying these two angles specifies the state exactly, and we note that specifying these angles is equivalent to specifying a point on the surface of the unit sphere S^2. This parametrization of the Hilbert space is known as the *Bloch sphere*. For example, the north pole of the Bloch sphere has $c_+ = 1, c_- = 0$, and so represents the state $|\psi\rangle = |+\rangle$, while the south pole has $c_+ = 0, c_- = 1$, and so represents the state $|\psi\rangle = |-\rangle$.

We will now define several important operators in this space. The identity operator is

$$
\mathbb{1} = |+\rangle\langle+| + |-\rangle\langle-| \,,
\tag{3.31}
$$

which can be represented by a matrix

$$
\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
\tag{3.32}
$$

Keep in mind that using an equals sign here is an abuse of notation, as the matrix representation of an operator is formally distinct from the operator itself. The operator S^z is given by

$$S^z = \frac{\hbar}{2}(|+\rangle\langle+| - |-\rangle\langle-|) = \frac{\hbar\sigma^z}{2} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.33}$$

where σ^z is the third Pauli matrix. Similarly, we can write

$$S^x = \frac{\hbar}{2}(|+\rangle\langle-| + |-\rangle\langle+|) = \frac{\hbar\sigma^x}{2} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$S^y = \frac{-i\hbar}{2}(|+\rangle\langle-| - |-\rangle\langle+|) = \frac{\hbar\sigma^y}{2} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{3.34}$$

We can check explicitly (left as an exercise) that these operators satisfy

$$\left[S^a, S^b\right] = i\hbar\epsilon^{abc}S^c$$

$$\left\{S^a, S^b\right\} = \frac{\hbar^2}{2}\delta^{ab}\mathbb{1}$$

$$S^2 := S^a S^a = \frac{3}{4}\hbar^2\mathbb{1} \tag{3.35}$$

$$\left[S^2, S^a\right] = 0,$$

where $[\cdot, \cdot]$ is the commutator, $\{\cdot, \cdot\}$ is the anticommutator, and ϵ^{abc} is the totally antisymmetric three-index tensor (known as the *Levi–Civita symbol*).

MIT OpenCourseWare

8.321 Quantum Theory I
Fall 2017