Recitation 1 (Sep. 15, 2017)

1.1 Simultaneous Diagonalization

In the last lecture, we discussed the situations in which two operators can be simultaneously diagonalized. As we stated, two diagonalizable operators A and B are simultaneously diagonalizable if and only if they commute, $[A, B] = 0$. Note that this is equivalent to the statement that there exists a basis consisting of simultaneous eigenkets of both operators.

Let's go through the proof once again. Consider a basis $\{|a_i\rangle\}$ for which the diagonalizable operators A and B have

$$A|a_i\rangle = a_i|a_i\rangle, \quad B|a_i\rangle = b_i|a_i\rangle. \tag{1.1}$$

This is a basis in which A and B are simultaneously diagonal; the $|a_i\rangle$ are eigenkets of both operators. Because the eigenvalues a_i and b_i are regular numbers in $\mathbb{C}$, they commute with one another and with the operators A and B, so we find that

$$AB|a_i\rangle = a_i b_i|a_i\rangle = BA|a_i\rangle. \tag{1.2}$$

The kets $\{|a_i\rangle\}$ form a basis, and so this shows that $AB = BA$, i.e., $[A, B, =]0$. This proves the forward direction: if A and B are simultaneously diagonalizable, then they commute.

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

$$A|a_i\rangle = a_i|a_i\rangle. \tag{1.3}$$

This means that A is diagonal in this basis. We can order the basis kets $|a_i\rangle$ so that eigenkets with the same eigenvalue are grouped together; in this case, A corresponds to a matrix

$$A = \begin{pmatrix} a_1 & & & & & & \\ & \ddots & & & & & \\ & & a_1 & & & & \\ & & & a_2 & & & \\ & & & & \ddots & & \\ & & & & & a_2 & \\ & & & & & & \ddots \end{pmatrix}. \tag{1.4}$$

Now, we see that in this basis

$$\begin{aligned} A(B|a_i\rangle) &= AB|a_i\rangle \\ &= BA|a_i\rangle \\ &= Ba_i|a_i\rangle \\ &= a_i(B|a_i\rangle), \end{aligned} \tag{1.5}$$

so $B|a_i\rangle$ is an eigenket of A with eigenvalue a_i. This means that $B|a_i\rangle$ must be a superposition of the eigenkets of A with eigenvalue a_i. Thus, B is block diagonal in the basis $\{|a_i\rangle\}$. To see this, note that multiplying a matrix by a vector can be viewed as taking a linear combination of the matrix columns:

$$\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = c_1 \begin{pmatrix} | \\ v_1 \\ | \end{pmatrix} + c_2 \begin{pmatrix} | \\ v_2 \\ | \end{pmatrix} + \dots + c_n \begin{pmatrix} | \\ v_n \\ | \end{pmatrix} \tag{1.6}$$

This is telling us that the matrix maps the first basis vector to the vector v_1, the second basis vector to the vector v_2, and so on. So, if $B|a_i\rangle$ is a superposition of kets with eigenvalue a_i, then the ith column of B must only have nonzero entries in rows corresponding to kets with eigenvalues of a_i under the action of A. Since we grouped the eigenvalues together, as seen in Eq. (1.4), this means that B must consist of blocks,

$$B = \begin{pmatrix} B_1 & & \\ & B_2 & \\ & & \ddots \end{pmatrix}, \tag{1.7}$$

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. The block diagonal form of B in this basis tells us that the eigenspaces of A (subspaces spanned by the eigenkets with a particular eigenvalue) are invariant subspaces of the operator B, meaning that each eigenspace is mapped into itself under the action of B.

We can then diagonalize B by diagonalizing each of the blocks B_i. Diagonalizing the block B_i only mixes eigenkets of A with the same eigenvalue, so this process leaves A diagonal. Thus, A and B are simultaneously diagonalizable. This proves the backward direction: if A and B commute, then they are simultaneously diagonalizable.

1.2   Hilbert Spaces

Thus far in the course, we have been talking about bras and kets living in Hilbert spaces, but we have not carefully defined what a Hilbert space is. A *Hilbert space* is a complete inner product space; more specifically, it is a real or complex inner product space that is complete with respect to the metric induced by the inner product.

Let's start piecing apart this definition. We have already discussed vector spaces and inner products. An *inner product space* is simply a vector space V over a field $\mathbb{F}$ (either $\mathbb{R}$ or $\mathbb{C}$) endowed with an inner product; that is, a map $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{F}$ satisfying the following properties for all $a, b \in \mathbb{F}$ and $u, v, w \in V$:

(a) Linearity in the second argument:

$$\langle u, av + bw \rangle = a\langle u, v \rangle + b\langle u, w \rangle. \tag{1.8}$$

(b) Conjugate symmetry:

$$\langle u, v \rangle = \langle v, u \rangle^*. \tag{1.9}$$

(c) Positive-definiteness:

$$\langle v, v \rangle \geq 0, \tag{1.10}$$

and $\langle v, v \rangle = 0$ if and only if $v = \mathbf{0}$.

The inner product naturally induces a norm on V, given by

$$\|v\| = \sqrt{\langle v, v \rangle} \tag{1.11}$$

for all $v \in V$. The properties of the inner product imply that this map is in fact a norm, i.e., a map $\|\cdot\| \colon V \to \mathbb{F}$ satisfying the following properties for all $a \in \mathbb{F}$ and $u, v \in V$:

(a) Absolute homogeneity:

$$\|av\| = |a| \|v\|. \tag{1.12}$$

(b) Subadditivity (the triangle inequality):

$$\|u + v\| \leq \|u\| + \|v\| . \tag{1.13}$$

(c) Definiteness: if $\|v\| = 0$, then $v = \mathbf{0}$.

This norm, in turn, induces a metric on V, given by

$$d(u, v) = \|u - v\| = \langle u - v, u - v \rangle \tag{1.14}$$

for all $u, v \in V$. Once again, the properties of the inner product ensure that this map is a metric, i.e., a map $d \colon V \times V \to \mathbb{F}$ satisfying the following properties for all $u, v, w \in V$:

(a) Identity of indiscernables: $d(u, v) = 0$ if and only if $u = v$.

(b) Symmetry:
$$d(u, v) = d(v, u) . \tag{1.15}$$

(c) Subadditivity (the triangle inequality):

$$d(u, w) \leq d(u, v) + d(v, w) . \tag{1.16}$$

We said that a Hilbert space is a real or complex inner product space that is complete with respect to this induced metric, and so all that remains is to understand what completeness is. A metric space is *complete* if all Cauchy sequences in the space converge to a point within the space. A *Cauchy sequence* is a sequence $\{x_n \in V \mid n \in \mathbb{N}\}$ whose elements become arbitrarily close to one another: for every real $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all natural numbers $m, n > N$,

$$d(x_m, x_n) < \epsilon . \tag{1.17}$$

1.2.1   Why do we care?

Now we have the full picture of what a Hilbert space is: it is a vector space, endowed with an inner product, such that any sequence of points in the space that becomes arbitrarily close together as we progress through the sequence has a limit within the space. Why is this the mathematical structure we want? Why not simply work in an inner product space that may or may not be complete?

From a practical or experimental standpoint, we don't need the completeness condition. Any physical apparatus will only be able to distinguish a finite number of outcomes, and so this concept of convergence will not play a role in the experiments we can actually perform physically. However, completeness is crucial if we want to use the language we are familiar with in quantum mechanics. For example, the time evolution operator is defined as

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar} = \sum_{n=0}^{\infty} \frac{(-i\hat{H}t/\hbar)^n}{n!} . \tag{1.18}$$

This operator is defined by its action on elements of the Hilbert space: for all $|\psi\rangle \in \mathcal{H}$, we have

$$\hat{U}(t)|\psi\rangle = \sum_{n=0}^{\infty} \frac{(-it/\hbar)^n \hat{H}^n |\psi\rangle}{n!} . \tag{1.19}$$

By definition, this infinite sum is defined to be equal to the limit of the sequence of its partial sums. Thus, we can only make sense of this operator if this sequence of finite sums converges to some state within the Hilbert space!

Another more formal issue is embedded in our use of the bra–ket notation. We use kets $|\psi\rangle$ to denote states in the Hilbert space $\mathcal{H}$ and bras $\langle\phi|$ to denote states in its continuous dual space $\mathcal{H}^*$. We regularly conflate two distinct notions: the inner product within the Hilbert space $\mathcal{H}$, and the action of linear functionals (bras) from the dual space on states in the Hilbert space. The reason we are allowed to do this is because every bra $\langle\psi| \in \mathcal{H}^*$ has a unique corresponding ket $|\psi\rangle \in \mathcal{H}$ such that the action of $\langle\psi|$ on any ket $|\phi\rangle \in \mathcal{H}$ is given by the inner product of $|\psi\rangle$ and $|\phi\rangle$:

$$\langle\psi|\phi\rangle = \langle|\psi\rangle, |\phi\rangle\rangle. \tag{1.20}$$

This duality between bras and kets is guaranteed by the *Riesz representation theorem*, which only holds in Hilbert spaces.

Heuristically, you can think of the reason we use Hilbert spaces in quantum mechanics analogously to why we use real numbers in physics. Any actual apparatus can only make measurements to some finite resolution, and so all physical measurements we will ever make will be rational numbers. However, when discussing the physics underlying these measurements, it is convenient to use the real numbers because they have many nice properties that the rational numbers do not.

1.3   Hermitian versus Self-Adjoint Operators and the Spectral Theorem

For this section, I will avoid using bras and kets within inner products because of notational clutter, and simply denote elements in the Hilbert space by $u, v, \ldots \in \mathcal{H}$.

In lecture, we talked about adjoint operators, and in particular discussed operators that are equal to their own adjoints, which we called Hermitian. We went on to discuss how Hermitian operators have a complete basis of orthonormal eigenvectors with real eigenvalues. In actuality, we have glossed over some technical details in this discussion.

First, let's define the notion of a bounded linear operator. A linear operator A between normed spaces V and W is *bounded* if there exists some $M > 0$ such that for all $v \in V$,

$$\|Av\|_W \leq M\|v\|_V. \tag{1.21}$$

A linear operator is bounded if and only if it is continuous.

Bounded linear operators work in the way we have discussed in class: if A is a bounded linear operator acting in a Hilbert space $\mathcal{H}$, then there exists a unique bounded linear operator $A^\dagger$, called the *adjoint* of A, that satisfies

$$\left\langle A^\dagger u, v \right\rangle = \langle u, Av \rangle \tag{1.22}$$

for all $u, v \in \mathcal{H}$. The operator A is self-adjoint if $A^\dagger = A$. Note that all linear operators on a finite-dimensional normed space are bounded.

The complications arise when we consider unbounded operators. Note that many of the operators we commonly talk about in quantum mechanics are unbounded: momentum, position, and the Hamiltonians of many systems, to name a few. An unbounded operator (this should be thought of as "not necessarily bounded") $A : X \to Y$ is a linear map from a linear subspace $\mathrm{dom}(A) \subseteq X$ to the space Y. The subset $\mathrm{dom}(A)$ is called the domain of A. Two unbounded operators are equal if they have the same domain and agree on that domain.

Now consider an unbounded operator $A \colon \mathrm{dom}(A) \subseteq \mathcal{H} \to \mathcal{H}$ acting in a Hilbert space $\mathcal{H}$. The operator A is *Hermitian* (or *symmetric*, if you're a mathematician) if it satisfies

$$\langle Au, v \rangle = \langle u, Av \rangle \tag{1.23}$$

for all $u, v \in \mathrm{dom}(A)$. This appears similar to the definition of a Hermitian operator we gave in class, but note that this relation only has to hold for elements in the domain of A (indeed, it wouldn't make sense otherwise). The *Hellinger–Toeplitz theorem* states that any Hermitian operator A with $\mathrm{dom}(A) = \mathcal{H}$ is necessarily bounded. Thus, all of the unbounded operators we discuss in physics, such as the momentum operator, cannot be defined as Hermitian operators on the entire Hilbert space!

Crucially, the spectral theorem only applies to self-adjoint operators, and not to Hermitian ones (contrary to what we stated in the lecture). So next, we must understand what a self-adjoint operator is, and to do so, we must understand adjoint operators of unbounded linear operators. For a densely-defined unbounded linear operator A acting in a Hilbert space $\mathcal{H}$, we define the adjoint $A^\dagger$ of A as follows. The domain of $A^\dagger$ is the set of vectors $u \in \mathcal{H}$ such that

$$v \mapsto \langle u, Av \rangle \tag{1.24}$$

is a continuous linear functional for all $v \in \mathrm{dom}(A)$. For each such u, this linear functional can be extended using linearity and the density of the domain to a unique continuous linear functional on all of $\mathcal{H}$ (see the Hahn–Banach theorem). Then, using the Riesz representation theorem, if $u \in \mathrm{dom}(A^\dagger)$, then there exists a unique $w \in \mathcal{H}$ such that

$$\langle u, Av \rangle = \langle w, v \rangle \tag{1.25}$$

for all $v \in \mathrm{dom}(A)$. We then define $A^\dagger u := w$. Note that A had to be densely-defined in order for its adjoint to be well-defined.

We can now define symmetric and self-adjoint operators in terms of their adjoint operators. An operator A is symmetric if $\mathrm{dom}(A) \subseteq \mathrm{dom}(A^\dagger)$ and $A^\dagger$ agrees with A when restricted to $\mathrm{dom}(A)$. An operator A is self-adjoint if $\mathrm{dom}(A) = \mathrm{dom}(A^\dagger)$ and both operators agree on this domain.

The spectral theorem, as stated in class, said that every Hermitian operator has an orthonormal basis of eigenkets with real eigenvalues. In actuality, this theorem only applies to self-adjoint operators. The formal way to state the spectral theorem for unbounded operators on infinite-dimensional spaces is to say that a densely-defined self-adjoint operator A is *unitarily equivalent* to a multiplication operator. This means that there is a unitary transformation $U \colon \mathcal{H} \to \mathcal{H}$ and a multiplication operator M such that U is a bijection between $\mathrm{dom}(A)$ and $\mathrm{dom}(M)$ and

$$U^{-1}MUv = Av \tag{1.26}$$

for all $v \in \mathrm{dom}(A)$. Rather formally, a *multiplication operator* M is an operator acting on measurable functions $g(x)$ on a measure space (X, Σ, μ) for which there exists a real-valued measurable function f on X such that

$$(Mg)(x) = f(x)g(x) \tag{1.27}$$

for all functions g such that $f(x)g(x)$ is square-integrable. This more formal definition is useful because it allows us to use the spectral theorem in cases where there is a continuous spectrum.

What if the operators we want to consider are not self-adjoint? A similar concept is that of an *essentially* self-adjoint operator, which is heuristically "good enough" that we can treat it like it's self-adjoint. An operator is *essentially self-adjoint* if its closure is self-adjoint. The closure of an operator is defined as follows: consider the *graph* of the operator A, given by

$$\Gamma(A) = \{(u, v) \in \mathcal{H} \oplus \mathcal{H} \mid v = Au\}. \tag{1.28}$$

If the topological closure of $\Gamma(A)$ in $\mathcal{H} \oplus \mathcal{H}$ also happens to be the graph of some operator, then this operator is called the *closure* of A. If an operator has a closure, it is called *closable* (note that

every Hermitian operator is closable). An operator is essentially self-adjoint if and only if it has a unique self-adjoint extension, which is its closure.

Finally, let's look at an example from physics to see where this actually comes into play. Consider the one-dimensional momentum operator,

$$\hat{p} = -i\hbar \frac{\mathrm{d}}{\mathrm{d}x} \, , \tag{1.29}$$

on the Hilbert space $L^2([0,1])$ of square-integrable functions on the interval $[0,1]$. In order to define this operator on this space of functions, we must specify boundary conditions to specify its domain. Consider the choice

$$\mathrm{dom}(\hat{p}) = \{ f \in C^\infty([0,1]) \mid f(0) = f(1) = 0 \} \, , \tag{1.30}$$

the set of smooth functions on $[0,1]$ that vanish at the endpoints. Using integration by parts, we see that the momentum operator is Hermitian on this domain:

$$
\begin{aligned}
\langle f, \hat{p}g \rangle &= \int_0^1 \mathrm{d}x \; f^*(x) \left(-i\hbar \frac{\mathrm{d}g}{\mathrm{d}x} \right) \\
&= -\int_0^1 \mathrm{d}x \left(-i\hbar \frac{\mathrm{d}f^*(x)}{\mathrm{d}x} \right) g(x) \\
&= \int_0^1 \mathrm{d}x \left(-i\hbar \frac{\mathrm{d}f(x)}{\mathrm{d}x} \right)^* g(x) \\
&= \langle \hat{p}f, g \rangle \, .
\end{aligned}
\tag{1.31}
$$

However, this operator is not essentially self-adjoint on this domain. The closure of $\hat{p}$ has domain

$$\mathrm{dom}\left(\hat{p}^{\mathrm{cl}}\right) = \{\text{twice-differentiable functions } f \in L^2([0,1]) \mid f(0) = f(1) = 0\} \, , \tag{1.32}$$

while the adjoint of $\hat{p}$ has the larger domain

$$\mathrm{dom}\left(\hat{p}^\dagger\right) = \{\text{twice-differentiable functions } f \in L^2([0,1])\} \, . \tag{1.33}$$

Note that the domain of the adjoint of the closure of an operator is equal to the domain of the adjoint, and so this is sufficient to show that $\hat{p}$ is not essentially self-adjoint. The domain of the adjoint has no boundary conditions, because the boundary conditions on the domain of A were sufficient for $\hat{p}$ to be Hermitian with no constraints on the domain of the adjoint.

A better choice of boundary conditions sets the domain of $\hat{p}$ as

$$\mathrm{dom}(\hat{p}) = \{ f \in C^\infty([0,1]) \mid f(0) = f(1) \} \, , \tag{1.34}$$

in which case $\hat{p}$ is essentially self-adjoint, meaning that we can identify a unique self-adjoint extension of $\hat{p}$ and apply the spectral theorem to that extension. You already know the corresponding physics statement: on the second domain, with periodic boundary conditions, there is an orthonormal basis of eigenfunctions of $\hat{p}$, given by

$$f_n(x) = e^{2\pi i n x / \hbar} \, , \tag{1.35}$$

while on the original domain the momentum operator has no eigenfunctions.

8.321 Quantum Theory I
Fall 2017