16. More about Operators

According to the second "principle" of quantum mechanics (see Lesson 12), every observable quantity corresponds to a linear *operator* that acts in the appropriate vector space. So, for example, the most important operators for a structureless particle in one dimension are

$$x, \quad -i\hbar\frac{\partial}{\partial x}, \quad \text{and} \quad -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x), \tag{1}$$

which correspond to position, momentum, and energy, respectively. The eigen*values* of each such operator are the values that you can obtain when you measure the corresponding quantity, and the corresponding eigen*vectors* are the quantum states for which the measured quantity has each of those well-defined values. If you measure the quantity for some other quantum state, the probability of getting any particular value is the square modulus of the *component* of the state vector along the direction of the corresponding eigenvector.

In order for this whole scheme to make sense, these eigenvalues must always be *real* numbers. In addition, the corresponding eigenvectors must form a complete, orthonormal basis (called an *eigenbasis*) for the vector space, so that any other normalized vector in the space can be uniquely resolved into components whose square moduli (the probabilities) add up to 1. It's now time to investigate some of the mathematics of the operators that have these essential properties.

Mathematicians say that an operator A is *Hermitian* if, for all vectors ψ_1 and ψ_2,

$$\langle A\psi_1, \psi_2\rangle = \langle \psi_1, A\psi_2\rangle. \tag{2}$$

That is, A can operate on *either* vector in an inner product with the same effect.

From this definition, it is easy to prove the following two theorems about Hermitian operators:

1. The eigenvalues of a Hermitian operator are real.
2. Eigenvectors of a Hermitian operator that have distinct eigenvalues are orthogonal to each other.

I suggest that you take a moment right now to try to prove each of these results, starting from the definition (2). If you get stuck, look in Griffiths (or some other textbook) for a hint. You do need to assume that the eigenvectors of your operator are normalizable, which is not the case when the eigenvalues are continuous (as they are for x and $-i\hbar\,\partial/\partial x$); Griffiths has a nice discussion of how to handle continuous eigenvalues.

Moreover, we can also prove a theorem that is almost the converse of the two theorems above:

3. If a linear operator has a complete, orthonormal set of eigenvectors and a corresponding set of real eigenvalues, then it is Hermitian.

This theorem means that any operator that corresponds to an observable quantity in quantum mechanics must be Hermitian. Here is the proof.

Let's call the operator A, its normalized eigenvectors α_n, and its eigenvalues a_n. Now consider the two arbitrary vectors ψ_1 and ψ_2 that appear in the inner product in equation 2. Because the set $\{\alpha_n\}$ is complete, we can expand each of these vectors in the α_n basis:

$$\psi_1 = \sum_n c_n \alpha_n, \qquad \psi_2 = \sum_n d_n \alpha_n, \tag{3}$$

for some sets of complex coefficients $\{c_n\}$ and $\{d_n\}$. So let's insert these expansions into the left-hand side of equation 2:

$$\langle A\psi_1, \psi_2 \rangle = \Big\langle A \sum_n c_n \alpha_n, \sum_m d_m \alpha_m \Big\rangle. \tag{4}$$

(Notice how I've used a different index, m, in the second sum, so I won't confuse it with the index in the first sum.) Because A is linear, we can move it inside the first sum. And because the inner product obeys the algebraic rules that one would expect of a "complex dot-product," we can move both of the sums, as well as the coefficients, outside of the inner product, picking up a $*$ on c_n, to obtain

$$\langle A\psi_1, \psi_2 \rangle = \sum_n \sum_m c_n^* d_m \langle A\alpha_n, \alpha_m \rangle. \tag{5}$$

But α_n is an eigenvector of A with real eigenvalue a_n, and the eigenvectors are orthonormal, so this becomes

$$\langle A\psi_1, \psi_2 \rangle = \sum_n \sum_m c_n^* d_m a_n \langle \alpha_n, \alpha_m \rangle = \sum_n \sum_m c_n^* d_m a_n \delta_{mn} = \sum_n c_n^* d_n a_n. \tag{6}$$

Through a completely analogous set of manipulations, you can show that the *right*-hand side of equation 2 reduces to exactly the same expression, and this completes the proof that A is Hermitian.

Although only Hermitian operators can correspond to observable quantities in quantum mechanics, we do sometimes work with non-Hermitian operators. It's then useful to define the *adjoint*, $A^\dagger$, of an operator A to be the operator that has the same effect on the left side of an inner product that A has on the right:

$$\langle A^\dagger \psi_1, \psi_2 \rangle = \langle \psi_1, A\psi_2 \rangle, \tag{7}$$

for any two vectors ψ_1 and ψ_2. A Hermitian operator, then, is its own adjoint. A *unitary* operator is one whose adjoint is the same as its inverse: $U^\dagger = U^{-1}$, where $U^{-1}U$ is the identity operator.

Often, in quantum mechanics, we will want to discuss two or more operators at the same time. For example, if two Hermitian operators A and B correspond to two observable quantities for a particular system, we might want to know whether these two quantities can be simultaneously well defined. As we shall later see, this question turns out to be closely related to the question of whether A and B *commute*, that is, whether their order matters when they operate, successively, on a vector:

$$AB\psi \equiv A(B\psi) \stackrel{?}{=} B(A\psi) \equiv BA\psi. \tag{8}$$

If $AB\psi = BA\psi$ for all ψ, then we say that operators A and B commute. And whether they commute or not, we define their *commutator*, denoted $[A, B]$, as the difference

$$[A, B] = AB - BA, \tag{9}$$

where it is understood that both sides need to give the same result when acting on an arbitrary vector ψ. So if A and B commute, then their commutator is zero.

Matrix representations

We already saw back in Lesson 7 how to express the Hamiltonian operator in matrix form, to facilitate numerical solutions of the time-independent Schrödinger equation. But the same idea works for any equation involving an operator. Consider the generic equation

$$A\psi_1 = \psi_2, \tag{10}$$

in which an operator A acts on one vector and turns it into another. We can use *any* discrete set of orthonormal basis vectors $\{\alpha_n\}$ (*not* necessarily the eigenvectors of A) to expand both vectors as in equation 3, to obtain

$$\sum_n c_n A\alpha_n = \sum_n d_n \alpha_n. \tag{11}$$

Now use Fourier's trick: take the inner product, from the left, of both sides of this equation with some other basis vector α_m, and use the fact that the basis vectors are orthonormal:

$$\sum_n c_n \langle \alpha_m, A\alpha_n \rangle = \sum_n d_n \langle \alpha_m, \alpha_n \rangle = \sum_n d_n \delta_{mn} = d_m. \tag{12}$$

The inner product inside the sum on the left-hand side is called the *mn matrix element* of A,

$$A_{mn} = \langle \alpha_m, A\alpha_n \rangle, \tag{13}$$

and our original equation can now be written in matrix form:

$$\begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \end{pmatrix}. \tag{14}$$

In summary, we can express the original equation, $A\psi_1 = \psi_2$, in matrix form by simply interpreting ψ_1 and ψ_2 as column vectors of components, and interpreting A as a matrix whose elements are the inner products of equation 13.

Representing operator equations with matrices in this way has the advantage of being concrete and vivid. For spin and other internal degrees of freedom for which the dimension of the vector space is finite and small, we normally express everything in matrix form from the start. But even when the dimension of the vector space is infinite, it's often useful to think of state vectors as columns of components, and to think of operators as matrices. Moreover, as we saw in Lesson 7, you can often get accurate results by keeping only a finite (and manageable) number of rows and columns.

From equation 13 you can prove the following further properties of operator-matrices:

- The adjoint matrix $A^\dagger$ is the conjugate transpose, A^{T*}.
- A Hermitian matrix is its own conjugate transpose, so its real part is symmetric and its imaginary part is antisymmetric.
- The matrix for a Hermitian operator in its *own* eigenbasis is diagonal, with entries equal to its eigenvalues.

I suggest that you now spend a few minutes trying to prove each of these statements.