C/CS/Phys 191 Operators, Physical Postulates, Hamiltonians
9/11/03
Fall 2003 Lecture 6

Operators

In lecture 3 we defined the operator $P = |v\rangle\langle v|$ which projects an arbitrary state onto the state $|v\rangle$. Now for an orthonormal basis $\{|j\rangle\}$ we can define the set of projection operators $P_j = |j\rangle\langle j|$ which obey the so-called "completeness relation" $\sum_{j=1}^{k} P_j = \sum_j |j\rangle\langle j| = 1$.

A linear operator maps states (kets) onto linear combinations of other states (kets). Suppose a ket $|b\rangle$ is mapped to a ket $|a\rangle$: the operator for this is denoted by the outer product $|a\rangle\langle b|$. So the action of linear operators can easily be written in our bra-ket language, e.g.,

$$\begin{aligned} X|\psi\rangle &= |a\rangle\langle b|\psi\rangle \\ Y|\psi\rangle &= |c\rangle\langle d|\psi\rangle \\ XY|\psi\rangle &= |a\rangle\langle b|c\rangle\langle d|\psi\rangle. \end{aligned}$$

If one these kets is a superposition of states, e.g., $|a\rangle = \alpha|0\rangle + \beta|1\rangle$, then the resulting state is also a superposition, i.e.,

$$X|\psi\rangle = (\langle b|\psi\rangle\alpha)|0\rangle + (\langle b|\psi\rangle\beta)|1\rangle.$$

So the bra-ket notation is naturally suited to the linear nature of quantum mechanical operators.

The inner product in the center of the last equation is a number, so clearly the "product" XY is also an operator. We often denote operators by the notation $\hat{X}$. Note that the order of these operators matters: applying $\hat{X}\hat{Y}$ to $|\psi\rangle$ results in a state proportional to $|a\rangle$, while applying $\hat{Y}\hat{X}$ results in a state proportional to $|c\rangle$.

Now lets consider how to express an operator that acts on states in a Hilbert space spanned by an orthonormal set $|j\rangle$. We can write the operator in terms of its action on these basis states, by making use of the completeness relation:

$$\begin{aligned} \hat{X}|j\rangle &= \hat{I}\hat{X}|j\rangle \\ &= \sum_{j'} |j'\rangle\langle j'|\hat{X}|j\rangle \\ &= \sum_{j'} X_{j'j}|j'\rangle, \end{aligned}$$

where $X_{j'j}$ is the $j'j$th element of the matrix representing the linear action of $\hat{X}$ on the basis. Furthermore,

$$\begin{aligned} \hat{X} &= \hat{I}\hat{X}\hat{I} \\ &= \sum_{j,j'} |j\rangle\langle j|\hat{X}|j'\rangle\langle `j'| \\ &= \sum_{j,j'} X_{jj'}|j\rangle\langle j'|. \end{aligned}$$

The diagonal matrix element X_{jj} is often referred to as the "expectation value" of $\hat{X}$ on state $|j\rangle$.

An important global characteristic of operators is their trace:

$$Tr\hat{X} = \sum_j X_{jj}.$$

For finite dimensional spaces the trace is easy to evaluate and is easily seen to be independent of basis (hint: insert the unit operator in above equation).

From now on we will drop the "$\hat{X}$" notation, unless essential to avoid misunderstanding, and simply refer to the operator as X.

A general operator A has a number of related operators that have their analogs in matrix algebra. The operator transpose A^T is defined by

$$A^T = \sum_{jj'} |j\rangle\langle j'|A|j\rangle\langle j'|$$

and the operator complex conjugate A^* by

$$A^* = \sum_{jj'} |j\rangle\langle j|A|j'\rangle^*\langle j'|$$

If $A = A^T$, then A is a symmetric operator, while if $A = -A^T$ it is skew-symmetric. A very important related operator is the Hermitian adjoint

$$A^\dagger = (A^*)^T = \sum jj'|j\rangle\langle j'|Aj\rangle^*\langle j'|$$

If $A = A^\dagger$, then A is Hermitian.

Hermitian operators are essential to quantum mechanics. A basic postulate of quantum mechanics is that physically meaningful entities of classical mechanics, such as momentum, energy, position, etc., are represented by Hermitian operators. Dirac called these entities "observables". Hermitian operators have some useful properties that again have their analog in matrix algebra. Thus, starting from the basic definition of Hermitian adjoint

$$\langle k|A|k'\rangle^* = \langle k'|A^\dagger|k\rangle$$

which means that if

$$A|\psi\rangle = |\psi'\rangle$$

that

$$\langle\psi'| = \langle\psi|A^\dagger,$$

one can easily show that

$$(BA)^\dagger = A^\dagger B^\dagger.$$

Now if both A and B are Hermitian operators, $A^\dagger B^\dagger = BA$, whence

$$(BA)^\dagger = AB.$$

For this product operator to be also Hermitian, we require $AB = BA$ and this is only true if A and B commute. This commutation property is so important in quantum mechanics that we define a special notation for it. The commutator of two operators is defined as the operator

$$C = AB - BA = [A, B]$$

and the operators A and B commute if $C = [A, B] = 0$. Note that this result implies that if the commutator $[A, B] \neq 0$ and A, B are both observables, then the product AB is not an observable. We say that "A and B are incompatible observables".

Eigenvalues/Eigenvectors

Since linear operators can be represented by matrices (on finite dimensional complex vector spaces), all the relevant properties of such matrices follow also for operators. Thus,

- any single Hermitian operator A can be diagonalized by a unitary transformation

  $$U^{\dagger} A U = a,$$

  where $a_{ij} = a\delta_{ij}$.

- elements of the diagonalized form are real eigenvalues $a_1, a_2, ...a_d$ where d is the dimension of the complex vector space. They may be degenerate, i.e., several having the same value. The set $\{a_i\}$ is called the "spectrum" of $\hat{A}$.

- the eigenvalues are the roots of the secular equation

  $$det(A - aI) = 0,$$

  i.e., the roots of an algebraic equation of degree d.

- the basis vectors $|1\rangle, |2\rangle, ...|d\rangle$ that diagonalize A are the eigenvectors (eigenkets) and satisfy

  $$A|n\rangle = a_n|n\rangle.$$

  Hence we may write A in terms of its eigenvectors/eigenvalues as

  $$A = \sum_n |n\rangle a_n \langle n|.$$

  This is known as the spectral decomposition of A.

- eigenvectors with different eigenvalues are orthogonal.

- If $A_i, i = 1, 2, ...K$ is a set of commuting Hermitian operators, i.e.,

  $$[A_i, A_j] = 0$$

  then one can simultaneously diagonalize the operators with the same unitary transformation. The eigenvalues are $a_n^{(i)}$ and the eigenvectors satisfy

  $$\{A_i - a_n^{(i)}\}|a_n^{(1)} a_n^{(2)} ... a_n^{(K)}\rangle = 0$$

  where the ket is labelled by all of its eigenvalues.

- If A, B are Hermitian and do not commute, they cannot be simultaneously diagonalized.

Hermitian operators and unitary evolution

We saw before (lecture 4) that time evolution of quantum systems is unitary. Now again from matrix algebra we know that unitary matrices are related to Hermitian matrices, as

$$U = e^{i}A,$$

since $U^{\dagger} = exp(-iA^{\dagger}) = exp(-iA)$ and hence $UU^{\dagger} = 1$.

What do we mean by the exponential of a linear operator? Think matrix representation:

$$e^{iAt} = 1 + (iAt) + \frac{(iAt)^2}{2} + \frac{(iAt)^3}{3} + ...$$

with

$$A^n = AA...A$$

the *n*-fold product. This is fine as long as the operator A is not dependent on time itself, in which case we need to be more careful.

The unitary time evolution of quantum systems is determined by the Hermitian operator H which corresponds to the observable of the system energy, according to

$$U(t) = e^{-(i/\hbar)Ht}$$

where t is the time and $\hbar$ a fundamental constant, Planck's constant, which has units of energy-time (Joule-sec). The Hermitian operator H is called the "Hamiltonian" and the above equation is a solution of the time dependent Schrodinger equation. We shall give a heuristic derivation of this in the next lecture by combining some physical reasoning with the abstract framework of quantum states and operators.

Fundamental (physical) postulates and the Schrodinger equation

Why do quantum state evolve in time according to this particular operator, and what is the meaning of this operator? To answer this we have to look at quantum mechanics from a more physical perspective. The physical basis of quantum mechanics rests on three fundamental postulates. These are given below in the wording of K. Gottfried and T. M. Yan (Quantum Mechanics: Fundamental, Springer 2003).

I. States, superposition The most complete description of the *state* of any physical system S at any time is provided by some vector $|v\rangle$ in the Hilbert space H appropriate to the system. Every linear combination of such state vectors $|\Psi\rangle$ represents a possible physical state of S.

This last sentence is the *superposition principle* that we have been using from the very beginning. Note the difference between a quantum and a classical description of a physical system. A classical description is complete with specification of the positions and momenta of all particles, each of which can be precisely measured at any time. In contrast, the quantum description is specified by the wave function $|\Psi\rangle$ that lives in an abstract Hilbert space that has no direct connection to the physical world. Classical mechanics is deterministic - particle positions and momenta can be specified for all times using the classical equations of motion. In contrast, quantum mechanics provides a statistical prediction of the outcomes of all observables on the system as the wave function $|\Psi\rangle$ evolves. Both descriptions are "complete" but they differ in the information that can be obtained. The uncertainty principle fundamentally changes the relation between coordinates and momenta in quantum mechanics.

II. Observables The physically meaningful entities of classical mechanics, such as position (q or x), momentum (p), etc. are represented by Hermitian operators. Following Dirac, we refer to these as "observables". We generalize these today to any physical meaningful entities, i.e., including those observables that have no classical correspondence (e.g., intrinsic spin).

III. Probabilistic interpretation and Measurement A set of N replicas of a quantum system S described by a state $\left|\Psi\right\rangle$ when subjected to measurements for a physical observable A, will yield in each measurement one of the eigenvalues $\{a_1, a_2, ...\}$ of $\hat{A}$ and as $N \to \infty$ this eigenvalue will appear with probability $P_\Psi(a_1), P_\Psi(a_2), ...$ where

$$P_\Psi(a_i) = |\left\langle a_i\right| \left|\Psi\right\rangle|^2$$

and $\left|a_i\right\rangle$ is the eigenvector corresponding to the eigenvalue a_i.

This is precisely the definition of probability in terms of specific outcomes in a sequence of identical tests on copies of S, provided that

$$\sum_i P_\Psi(a_i) = \sum_i |\left\langle a_i\right| \left|\Psi\right\rangle|^2 = \left\langle\Psi\right| \left|\Psi\right\rangle = 1.$$

This is automatically satisfied for states that are normalized to unity.

The expectation value of an observable A in an arbitrary state $\left|\Psi\right\rangle$ also looks like an average over a probability distribution:

$$\begin{aligned} \langle A\rangle_\Psi &= \left\langle\Psi\right|\hat{A}\left|\Psi\right\rangle \\ &= \sum_i \left\langle\Psi\right| \left|a_i\right\rangle a_i \left\langle a_i\right| \left|\Psi\right\rangle \\ &= \sum_i a_i P_\Psi(a_i). \end{aligned}$$

Note that if the state Ψ is an eigenstate of A, then

$$\left\langle\Psi\right|A\left|\Psi\right\rangle = a_j$$

where a_j is the corresponding eigenvalue, i.e., only a single term contributes.

We can generalize this procedure from projection onto eigenstates to projection onto an arbitrary state $\left|\phi\right\rangle$. Thus, the probability to find a quantum system S that is in state $\left|\Psi\right\rangle$ in another state $\left|\phi\right\rangle$ is equal to

$$P_\Psi(\phi) = |\left\langle\phi\right| \left|\Psi\right\rangle|^2.$$

This projection of the ket $\left|\Psi\right\rangle$ onto another state, be it an eigenfunction of some operator $\left|a_i\right\rangle$, a basis function for the Hilbert space $\left|v_i\right\rangle$, or an arbitrary state $\left|\phi\right\rangle$, is referred to as a "probability amplitude", since its square modulus is a probability. Note that the probability amplitude is specified both by $\left|\Psi\right\rangle$ *and* the other state: the latter specifies the "representation" of $\left|\Psi\right\rangle$ which realizes the quantum state in a measurable basis. The probability amplitude is also referred to as the "wave function" in the specified "representation".

A single measurement of th observable A on a state $\left|\Psi\right\rangle$ in the basis (representation) of eigenstates of $\hat{A}$ will yield the value a_i, with probability $P_\Psi(a_i) = |\left\langle a_i\right| \left|\Psi\right\rangle|^2$. This defines the measurement operator

$$\hat{M}_i = \left|a_i\right\rangle \left\langle a_i\right|$$

that acts on the state $\left|\Psi\right\rangle$. The normalized state after measurement is then easily seen to be equal to

$$\frac{\hat{M}_i\left|\Psi\right\rangle}{\sqrt{\left\langle\Psi\right|M_i^\dagger M_i\left|\Psi\right\rangle}}.$$

For a measurement in the $\left|a_i\right\rangle$ basis this is given by

$$\left|i\right\rangle\frac{\left\langle i\right|\left|\Psi\right\rangle}{\sqrt{\left\langle\Psi\right|M_i^\dagger M_i\left|\Psi\right\rangle}},$$

where we have abbreviated

$$\left|a_i\right\rangle \equiv \left|i\right\rangle.$$

For example, suppose we have the linear superposition

$$\left|\Psi\right\rangle = \alpha_1\left|1\right\rangle + \alpha_2\left|2\right\rangle + \alpha_3\left|3\right\rangle + ... + \alpha_k\left|k\right\rangle.$$

Making a single measurement of the observable A on $\left|\Psi\right\rangle$ will result in the outcome a_i with probability

$$P_\Psi(a_i) = |\alpha_i|^2$$

and the resulting state after the measurement is equal to

$$\left|i\right\rangle\left(\frac{\alpha_i}{|\alpha_i|}\right).$$

The measurement of the observable has "collapsed" the state $\left|\Psi\right\rangle$ to a single eigenstate $\left|i\right\rangle \equiv \left|a_i\right\rangle$ of $\hat{A}$ (recall these constitute an orthonormal basis).
