

G25.2651: Statistical Mechanics

Notes for Lecture 6

I. THE FUNDAMENTAL POSTULATES OF QUANTUM MECHANICS

The fundamental postulates of quantum mechanics concern the following questions:

1. How is the physical state of a system described?
2. How are physical observables represented?
3. What are the results of measurements on quantum mechanical systems?
4. How does the physical state of a system evolve in time?
5. The uncertainty principle.

A. The physical state of a quantum system

The physical state of a quantum system is represented by a vector denoted

$$|\Psi(t)\rangle$$

which is a column vector, whose components are *probability amplitudes* for different states in which the system might be found if a measurement were made on it.

A probability amplitude α is a complex number, the square modulus of which gives the corresponding probability P_α

$$P_\alpha = |\alpha|^2$$

The number of components of $|\Psi(t)\rangle$ is equal to the number of possible states in which the system might be observed. The space that contains $|\Psi(t)\rangle$ is called a Hilbert space \mathcal{H} . The dimension of \mathcal{H} is also equal to the number of states in which the system might be observed. It could be finite or infinite (countable or not).

$|\Psi(t)\rangle$ must be a unit vector. This means that the inner product:

$$\langle\Psi(t)|\Psi(t)\rangle = 1$$

In the above, if the vector $|\Psi(t)\rangle$, known as a Dirac “ket” vector, is given by the column

$$|\Psi(t)\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \vdots \end{pmatrix}$$

then the vector $\langle\Psi(t)|$, known as a Dirac “bra” vector, is given by

$$\langle\Psi(t)| = (\psi_1^* \quad \psi_2^* \quad \cdots)$$

so that the inner product becomes

$$\langle\Psi(t)|\Psi(t)\rangle = \sum_i |\psi_i|^2 = 1$$

We can understand the meaning of this by noting that ψ_i , the components of the state vector, are probability amplitudes, and $|\psi_i|^2$ are the corresponding probabilities. The above condition then implies that the sum of all the probabilities of being in the various possible states is 1, which we know must be true for probabilities.

B. Physical Observables

Physical observables are represented by linear, hermitian operators that act on the vectors of the Hilbert space. If A is such an operator, and $|\phi\rangle$ is an arbitrary vector in the Hilbert space, then A might act on $|\phi\rangle$ to produce a vector $|\phi'\rangle$, which we express as

$$A|\phi\rangle = |\phi'\rangle$$

Since $|\phi\rangle$ is representable as a column vector, A is representable as a matrix with components

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ \cdot & \cdot & \cdot & \cdots \end{pmatrix}$$

The condition that A must be hermitian means that

$$A^\dagger = A$$

or

$$A_{ij} = A_{ji}^*$$

C. Measurement

The result of a measurement of the observable A must yield one of the eigenvalues of A . Thus, we see why A is required to be a hermitian operator: Hermitian operators have *real* eigenvalues. If we denote the set of eigenvalues of A by $\{a_i\}$, then each of the eigenvalues a_i satisfies an eigenvalue equation

$$A|a_i\rangle = a_i|a_i\rangle$$

where $|a_i\rangle$ is the corresponding eigenvector. Since the operator A is hermitian and a_i is therefore real, we have also the left eigenvalue equation

$$\langle a_i|A = \langle a_i|a_i$$

The probability amplitude that a measurement of A will yield the eigenvalue a_i is obtained by taking the inner product of the corresponding eigenvector $|a_i\rangle$ with the state vector $|\Psi(t)\rangle$, $\langle a_i|\Psi(t)\rangle$. Thus, the probability that the value a_i is obtained is given by

$$P_{a_i} = |\langle a_i|\Psi(t)\rangle|^2$$

Another useful and important property of hermitian operators is that their eigenvectors form a complete orthonormal basis of the Hilbert space, when the eigenvalue spectrum is non-degenerate. That is, they are linearly independent, span the space, satisfy the orthonormality condition

$$\langle a_i|a_j\rangle = \delta_{ij}$$

and thus any arbitrary vector $|\phi\rangle$ can be expanded as a linear combination of these vectors:

$$|\phi\rangle = \sum_i c_i |a_i\rangle$$

By multiplying both sides of this equation by $\langle a_j|$ and using the orthonormality condition, it can be seen that the expansion coefficients are

$$c_i = \langle a_i|\phi\rangle$$

The eigenvectors also satisfy a closure relation:

$$I = \sum_i |a_i\rangle\langle a_i|$$

where I is the identity operator.

Averaging over many individual measurements of A gives rise to an average value or expectation value for the observable A , which we denote $\langle A \rangle$ and is given by

$$\langle A \rangle = \langle \Psi(t) | A | \Psi(t) \rangle$$

That this is true can be seen by expanding the state vector $|\Psi(t)\rangle$ in the eigenvectors of A :

$$|\Psi(t)\rangle = \sum_i \alpha_i(t) |a_i\rangle$$

where α_i are the amplitudes for obtaining the eigenvalue a_i upon measuring A , i.e., $\alpha_i = \langle a_i | \Psi(t) \rangle$. Introducing this expansion into the expectation value expression gives

$$\begin{aligned} \langle A \rangle(t) &= \sum_{i,j} \alpha_i^*(t) \alpha_j(t) \langle a_i | A | a_j \rangle \\ &= \sum_{i,j} \alpha_i^*(t) \alpha_j(t) a_i \delta_{ij} \\ &= \sum_i a_i |\alpha_i(t)|^2 \end{aligned}$$

The interpretation of the above result is that the expectation value of A is the sum over possible outcomes of a measurement of A weighted by the probability that each result is obtained. Since $|\alpha_i|^2 = |\langle a_i | \Psi(t) \rangle|^2$ is this probability, the equivalence of the expressions can be seen.

Two observables are said to be compatible if $AB = BA$. If this is true, then the observables can be diagonalized simultaneously to yield the same set of eigenvectors. To see this, consider the action of BA on an eigenvector $|a_i\rangle$ of A . $BA|a_i\rangle = a_i B|a_i\rangle$. But if this must equal $AB|a_i\rangle$, then the only way this can be true is if $B|a_i\rangle$ yields a vector proportional to $|a_i\rangle$ which means it must also be an eigenvector of B . The condition $AB = BA$ can be expressed as

$$\begin{aligned} AB - BA &= 0 \\ [A, B] &= 0 \end{aligned}$$

where, in the second line, the quantity $[A, B] \equiv AB - BA$ is known as the commutator between A and B . If $[A, B] = 0$, then A and B are said to commute with each other. That they can be simultaneously diagonalized implies that one can simultaneously predict the observables A and B with the same measurement.

As we have seen, classical observables are functions of position x and momentum p (for a one-particle system). Quantum analogs of classical observables are, therefore, functions of the operators X and P corresponding to position and momentum. Like other observables X and P are linear hermitian operators. The corresponding eigenvalues x and p and eigenvectors $|x\rangle$ and $|p\rangle$ satisfy the equations

$$\begin{aligned} X|x\rangle &= x|x\rangle \\ P|p\rangle &= p|p\rangle \end{aligned}$$

which, in general, could constitute a *continuous* spectrum of eigenvalues and eigenvectors. The operators X and P are not compatible. In accordance with the Heisenberg uncertainty principle (to be discussed below), the commutator between X and P is given by

$$[X, P] = i\hbar I$$

and that the inner product between eigenvectors of X and P is

$$\langle x | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}$$

Since, in general, the eigenvalues and eigenvectors of X and P form a continuous spectrum, we write the orthonormality and closure relations for the eigenvectors as:

$$\begin{aligned}
\langle x|x'\rangle &= \delta(x-x') & \langle p|p'\rangle &= \delta(p-p') \\
|\phi\rangle &= \int dx|x\rangle\langle x|\phi\rangle & |\phi\rangle &= \int dp|p\rangle\langle p|\phi\rangle \\
I &= \int dx|x\rangle\langle x| & I &= \int dp|p\rangle\langle p|
\end{aligned}$$

The probability that a measurement of the operator X will yield an eigenvalue x in a region dx about some point is

$$P(x, t)dx = |\langle x|\Psi(t)\rangle|^2 dx$$

The object $\langle x|\Psi(t)\rangle$ is best represented by a continuous function $\Psi(x, t)$ often referred to as the *wave function*. It is a representation of the inner product between eigenvectors of X with the state vector. To determine the action of the operator X on the state vector in the basis set of the operator X , we compute

$$\langle x|X|\Psi(t)\rangle = x\Psi(x, t)$$

The action of P on the state vector in the basis of the X operator is consequential of the incompatibility of X and P and is given by

$$\langle x|P|\Psi(t)\rangle = \frac{\hbar}{i} \frac{\partial}{\partial x} \Psi(x, t)$$

Thus, in general, for any observable $A(X, P)$, its action on the state vector represented in the basis of X is

$$\langle x|A(X, P)|\Psi(t)\rangle = A\left(x, \frac{\hbar}{i} \frac{\partial}{\partial x}\right) \Psi(x, t)$$

D. Time evolution of the state vector

The time evolution of the state vector is prescribed by the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = H|\Psi(t)\rangle$$

where H is the Hamiltonian operator. This equation can be solved, in principle, yielding

$$|\Psi(t)\rangle = e^{-iHt/\hbar} |\Psi(0)\rangle$$

where $|\Psi(0)\rangle$ is the initial state vector. The operator

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

is the *time evolution operator* or *quantum propagator*. Let us introduce the eigenvalues and eigenvectors of the Hamiltonian H that satisfy

$$H|E_i\rangle = E_i|E_i\rangle$$

The eigenvectors form an orthonormal basis on the Hilbert space and therefore, the state vector can be expanded in them according to

$$|\Psi(t)\rangle = \sum_i c_i(t) |E_i\rangle$$

where, of course, $c_i(t) = \langle E_i|\Psi(t)\rangle$, which is the amplitude for obtaining the value E_i at time t if a measurement of H is performed. Using this expansion, it is straightforward to show that the time evolution of the state vector can be written as an expansion:

$$\begin{aligned}
|\Psi(t)\rangle &= e^{-iHt/\hbar}|\Psi(0)\rangle \\
&= e^{-iHt/\hbar} \sum_i |E_i\rangle \langle E_i|\Psi(0)\rangle \\
&= \sum_i e^{-iE_i t/\hbar} |E_i\rangle \langle E_i|\Psi(0)\rangle
\end{aligned}$$

Thus, we need to compute all the initial amplitudes for obtaining the different eigenvalues E_i of H , apply to each the factor $\exp(-iE_i t/\hbar)|E_i\rangle$ and then sum over all the eigenstates to obtain the state vector at time t .

If the Hamiltonian is obtained from a classical Hamiltonian $H(x, p)$, then, using the formula from the previous section for the action of an arbitrary operator $A(X, P)$ on the state vector in the coordinate basis, we can recast the Schrödinger equation as a partial differential equation. By multiplying both sides of the Schrödinger equation by $\langle x|$, we obtain

$$\begin{aligned}
\langle x|H(X, P)|\Psi(t)\rangle &= i\hbar \frac{\partial}{\partial t} \langle x|\Psi(t)\rangle \\
H\left(x, \frac{\hbar}{i} \frac{\partial}{\partial x}\right) \Psi(x, t) &= i\hbar \frac{\partial}{\partial t} \Psi(x, t)
\end{aligned}$$

If the classical Hamiltonian takes the form

$$H(x, p) = \frac{p^2}{2m} + U(x)$$

then the Schrödinger equation becomes

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + U(x)\right] \Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t)$$

which is known as the Schrödinger *wave equation* or the *time-dependent* Schrödinger equation.

In a similar manner, the eigenvalue equation for H can be expressed as a differential equation by projecting it into the X basis:

$$\begin{aligned}
\langle x|H|E_i\rangle &= E_i \langle x|E_i\rangle \\
H\left(x, \frac{\hbar}{i} \frac{\partial}{\partial x}\right) \psi_i(x) &= E_i \psi_i(x) \\
\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + U(x)\right] \psi_i(x) &= E_i \psi_i(x)
\end{aligned}$$

where $\psi_i(x) = \langle x|E_i\rangle$ is an eigenfunction of the Hamiltonian.

E. The Heisenberg uncertainty principle

Because the operators X and P are not compatible, $[X, P] \neq 0$, there is no measurement that can precisely determine both X and P simultaneously. Hence, there must be an uncertainty relation between them that specifies how uncertain we are about one quantity given a definite precision in the measurement of the other. Presumably, if one can be determined with infinite precision, then there will be an infinite uncertainty in the other. Recall that we had defined the uncertainty in a quantity by

$$\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}$$

Thus, for X and P , we have

$$\begin{aligned}
\Delta x &= \sqrt{\langle X^2 \rangle - \langle X \rangle^2} \\
\Delta p &= \sqrt{\langle P^2 \rangle - \langle P \rangle^2}
\end{aligned}$$

These quantities can be expressed explicitly in terms of the wave function $\Psi(x, t)$ using the fact that

$$\langle X \rangle = \langle \Psi(t) | X | \Psi(t) \rangle = \int dx \langle \Psi(t) | x \rangle \langle x | X | \Psi(t) \rangle = \int dx \Psi^*(x, t) x \Psi(x, t)$$

and

$$\langle X^2 \rangle = \langle \Psi(t) | X^2 | \Psi(t) \rangle = \int \Psi^*(x, t) x^2 \Psi(x, t)$$

Similarly,

$$\langle P \rangle = \langle \Psi(t) | P | \Psi(t) \rangle = \int dx \langle \Psi(t) | x \rangle \langle x | P | \Psi(t) \rangle = \int dx \Psi^*(x, t) \frac{\hbar}{i} \frac{\partial}{\partial x} \Psi(x, t)$$

and

$$\langle P^2 \rangle = \langle \Psi(t) | P^2 | \Psi(t) \rangle = \int dx \Psi^*(x, t) \left(-\hbar^2 \frac{\partial^2}{\partial x^2} \right) \Psi(x, t)$$

Then, the Heisenberg uncertainty principle states that

$$\Delta x \Delta p \gtrsim \hbar$$

which essentially states that the greater certainty with which a measurement of X or P can be made, the greater will be the *uncertainty* in the other.

F. The Heisenberg picture

In all of the above, notice that we have formulated the postulates of quantum mechanics such that the state vector $|\Psi(t)\rangle$ evolves in time but the operators corresponding to observables are taken to be stationary. This formulation of quantum mechanics is known as the *Schrödinger picture*. However, there is another, completely equivalent, picture in which the state vector remains stationary and the operators evolve in time. This picture is known as the *Heisenberg picture*. This particular picture will prove particularly useful to us when we consider quantum time correlation functions.

The Heisenberg picture specifies an evolution equation for any operator A , known as the Heisenberg equation. It states that the time evolution of A is given by

$$\frac{dA}{dt} = \frac{1}{i\hbar} [A, H]$$

While this evolution equation must be regarded as a postulate, it has a very immediate connection to classical mechanics. Recall that any function of the phase space variables $A(x, p)$ evolves according to

$$\frac{dA}{dt} = \{A, H\}$$

where $\{\dots, \dots\}$ is the Poisson bracket. The suggestion is that in the classical limit (\hbar small), the commutator goes over to the Poisson bracket. The Heisenberg equation can be solved in principle giving

$$\begin{aligned} A(t) &= e^{iHt/\hbar} A e^{-iHt/\hbar} \\ &= U^\dagger(t) A U(t) \end{aligned}$$

where A is the corresponding operator in the Schrödinger picture. Thus, the expectation value of A at any time t is computed from

$$\langle A(t) \rangle = \langle \Psi | A(t) | \Psi \rangle$$

where $|\Psi\rangle$ is the stationary state vector.

Let's look at the Heisenberg equations for the operators X and P . If H is given by

$$H = \frac{P^2}{2m} + U(X)$$

then Heisenberg's equations for X and P are

$$\begin{aligned}\frac{dX}{dt} &= \frac{1}{i\hbar}[X, H] = \frac{P}{m} \\ \frac{dP}{dt} &= \frac{1}{i\hbar}[P, H] = -\frac{\partial U}{\partial X}\end{aligned}$$

Thus, Heisenberg's equations for the operators X and P are just Hamilton's equations cast in operator form. Despite their innocent appearance, the solution of such equations, even for a one-particle system, is *highly* nontrivial and has been the subject of a considerable amount of research in physics and mathematics.

Note that any operator that satisfies $[A(t), H] = 0$ will not evolve in time. Such operators are known as constants of the motion. The Heisenberg picture shows explicitly that such operators do not evolve in time. However, there is an analog with the Schrödinger picture: Operators that commute with the Hamiltonian will have associated probabilities for obtaining different eigenvalues that do not evolve in time. For example, consider the Hamiltonian, itself, which it trivially a constant of the motion. According to the evolution equation of the state vector in the Schrödinger picture,

$$|\Psi(t)\rangle = \sum_i e^{-iE_i t/\hbar} |E_i\rangle \langle E_i | \Psi(0)\rangle$$

the amplitude for obtaining an energy eigenvalue E_j at time t upon measuring H will be

$$\begin{aligned}\langle E_j | \Psi(t)\rangle &= \sum_i e^{-iE_i t/\hbar} \langle E_j | E_i\rangle \langle E_i | \Psi(0)\rangle \\ &= \sum_i e^{-iE_i t/\hbar} \delta_{ij} \langle E_i | \Psi(0)\rangle \\ &= e^{-iE_j t/\hbar} \langle E_j | \Psi(0)\rangle\end{aligned}$$

Thus, the squared modulus of both sides yields the probability for obtaining E_j , which is

$$|\langle E_j | \Psi(t)\rangle|^2 = |\langle E_j | \Psi(0)\rangle|^2$$

Thus, the probabilities do not evolve in time. Since any operator that commutes with H can be diagonalized simultaneously with H and will have the same set of eigenvectors, the above arguments will hold for any such operator.