Lecture 13:
Postulates of Quantum Mechanics Revisited

Date Given: 2008/10/29
Date Revised: 2008/10/29
Recall the Postulates we briefly discussed in Section 1.2:

1. The state of a particle is represented by a vector in a physical Hilbert space.
2. The fundamental state variables \( x \) and \( p \) of classical mechanics are replaced by Hermitian operators \( X \) and \( P \) whose matrix elements are well specified in a physical Hilbert space basis consisting of position eigenstates (states with perfectly defined position \( x \)). Any derived dynamical variables \( \omega(x, p) \) are replaced by operators \( \Omega \) defined by the above correspondence.
3. Measurement of any classical variable \( \omega(x, p) \) for a quantum state yields only the eigenvalues of the corresponding operator \( \Omega \), with the probability of obtaining the eigenvalue \( \omega \) given by the squared norm of the projection of the state onto the eigenstate corresponding to \( \omega \).
4. The state vector evolves according to the Schrödinger equation.

We now have the language to interpret what is meant by these postulates. We do that in this section.
Postulate 1: Representation of Particle States

The state of a particle is represented by a vector $|\psi(t)\rangle$ in a physical Hilbert space.

We now know what is meant by this statement mathematically, in a generic sense: the state $|\psi(t)\rangle$ is an element in an inner product space; members of such spaces have the following important properties:

- They can be added together linearly, with coefficients that are just numbers.
- An inner product is defined that provides for definitions of orthogonality and normalization.
- There exist orthonormal sets of basis states and all states can be written as linear combinations of them.
- In terms of any particular basis, the vector corresponding to the state has a column matrix representation, and the corresponding dual vector in the dual vector space has a row matrix representation that is the conjugate transpose of the column matrix representation. When the basis is orthonormal, the inner product is equivalent to matrix multiplication of the column and row matrices.
Postulate 1: Representation of Particle States (cont.)

- Operators can act on the states and return new states, and any operator has a matrix representation for any particular choice of orthonormal basis.
- There are Hermitian operators that have real eigenvalues and a set of eigenvectors yields as an orthonormal basis.
- There are unitary operators that can be used to rotate from one orthonormal basis to another and which themselves have unit-modulus eigenvalues and orthonormal sets of eigenvectors.
- The space has been restricted so that the $K$ operator is Hermitian.
Normalization Considerations

One implication of this postulate is that, when we take a linear combination of states, $|\chi\rangle = \alpha |\psi\rangle + \beta |\phi\rangle$, we will in general want to normalize the result; that is, we should define

$$|\chi\rangle = \frac{\alpha |\psi\rangle + \beta |\phi\rangle}{\sqrt{\alpha^2 + \beta^2}}$$

so that $|\chi|^2 = \langle \chi | \chi \rangle = 1$ if $|\psi|^2 = \langle \psi | \psi \rangle = 1$ and $|\phi|^2 = \langle \phi | \phi \rangle = 1$ (you can check this by writing out $\langle \chi | \chi \rangle$). As we will see, this convention ensures measurement probabilities will be automatically normalized for $|\chi\rangle$. 

Section 4.2 Postulates Revisited: Postulate 1: Representation of Particle States
The independent variables $x$ and $p$ that describe completely the state of a particle in classical mechanics are represented by Hermitian operators $X$ and $P$ in the physical Hilbert space of states, with $X$ and $P$ having the following matrix elements when using the position basis for the physical Hilbert space:

$$
\langle x | X | x' \rangle = x \delta (x - x') \\
\langle x | P | x' \rangle = -i \hbar \frac{d}{dx} \delta (x - x')
$$

(4.1)

Any arbitrary classical dynamical variable $\omega(x, p)$ has a corresponding Hermitian operator

$$
\Omega(X, P) = \omega(x \rightarrow X, p \rightarrow P)
$$

(4.2)

where we simply replace $x$ and $p$ in $\omega$ with $X$ and $P$ to obtain $\Omega(X, P)$.

Having been through the exercise of constructing the infinite-dimensional generalization of inner product spaces, we now understand what is meant by the operators $X$ and $P$ and their matrix elements. Postulate 3 tells us how the above matrix elements are related to measurements. The extension to arbitrary classical variables $\omega(x, p)$ is also clear, modulo the issue of having to deal with ambiguous combinations of $x$ and $p$ (i.e., if one has the classical quantity $x p$, should one use $X P$, $P X$, their sum, or their difference?).
One thing that will not be clear yet, and cannot be discussed until you have seen Hamiltonian mechanics in Ph106a, is why we make the above choice for $P$. This choice was a clever guess by the creators of quantum mechanics based on analogy to classical mechanics. An alternative version of this postulate makes this more clear: it takes $X$ as defined above but then makes the requirement $[X, P] = i \hbar$. This latter relation is the quantum mechanical analogue of the classical Poisson bracket $\{x, p\} = 1$. Written this way, the path from classical mechanics to quantum mechanics is explicit. But that does not amount to a proof. By definition, a postulate can never be explicitly proven, but it can be motivated and then checked that it gives physically reasonable and correct results in particular situations.
Postulate 3: Results of Measurements of Classical Variables

Let \{\ket{\omega}\} denote the set of eigenstates of the Hermitian operator with eigenvalues \(\omega\). If a particle is in an arbitrary state \(\ket{\psi}\), then measurement of the variable corresponding to the operator \(\Omega\) will yield only the eigenvalues \{\omega\} of \(\Omega\). The measurement will yield the particular value \(\omega\) for that variable with relative probability \(P(\omega) = |\langle \omega | \psi \rangle|^2\) and the system will change from state \(\ket{\psi}\) to state \(\ket{\omega}\) as a result of the measurement being made.

Let’s break the statement down carefully:

1. The eigenvalues of \(\Omega\) are the only values the measured quantity may take on.

2. The measurement outcome is fundamentally probabilistic, and the relative probability of a particular allowed outcome \(\omega\) is given by finding the projection of \(\ket{\psi}\) onto the corresponding eigenstate \(\ket{\omega}\). By relative probability, we simply mean that the ratio of the probabilities of two outcomes is given by \(P(\omega_1)/P(\omega_2) = |\langle \omega_1 | \psi \rangle|^2 / |\langle \omega_2 | \psi \rangle|^2\). The absolute probability of a particular outcome requires a normalizing factor that sums over all possible measurement outcomes, to be discussed later. This implies that, if \(\ket{\psi}\) is an eigenstate of \(\Omega\), then the measurement will always yield the corresponding eigenvalue.

3. The measurement process itself changes the state of the particle to the eigenstate \(\ket{\omega}\) corresponding to the measurement outcome \(\omega\). This is the equivalent of applying the projection operator \(P_\omega\) (but one only knows which \(P_\omega\) to use after the measurement has been done!) and then renormalizing the state.
Where the math ends and the physics starts

As we noted in Section 1.2, we could have made a more classical interpretation of the expansion of $|\psi\rangle$ in terms of the eigenvectors $\{|\omega\rangle\}$: that the result of the measurement would be the weighted sum of the eigenvalues, weighted by the norms of the expansion coefficients $|\langle \omega | \psi \rangle|^2$ rather than $\langle \omega | \psi \rangle$ because the former is guaranteed to be real while the latter is not. But we do not do that. It is a physical assumption that the expansion coefficients are to be interpreted as probabilities of the allowed outcomes, not as weighting factors.

Also, we could have assumed that measurement is not an operation that changes $|\psi\rangle$: we could have said that $|\psi\rangle$ evolves in some way independent of any measurements that take place. Even if we had said that the action of a measurement on $|\psi\rangle$ is to act with the corresponding operator $\Omega$ on $|\psi\rangle$, we would not arrive at this postulate. It is entirely outside of the mathematical structure to assume that a measurement to which the operator $\Omega$ corresponds results in $|\psi\rangle$ collapsing to one of the $\{|\omega\rangle\}$ via application of the appropriate $P_\omega$. It is an assumption that yields correct predictions for experimental results.

Note that there is no way to write the collapse process explicitly as an operator. The problem is that you only know which $P_\omega$ to apply after you have obtained the measured value $\omega$; the act of measurement does not, for example, correspond to the operator $M = \sum_j P(\omega_j)P_{\omega_j} = \sum_j |\langle \omega_j \psi(t) \rangle|^2 |\omega_j \rangle \langle \omega_j |$. 
Degeneracy

In the case of degenerate eigenvalues, the obvious generalization of the above postulate is to replace $P(\omega) = |\langle \omega | \psi \rangle|^2$ with

$$P(\omega) = |P_\omega |\psi \rangle|^2 = \langle \psi | P_\omega P_\omega | \psi \rangle = \langle \psi | P_\omega | \psi \rangle$$ (4.3)

where $P_\omega$ is the projection operator for the $\omega$ subspace,

$$P_\omega = \sum_{\omega_j=\omega} |\omega_j \rangle \langle \omega_j |$$ (4.4)

and where we have written out three equivalent expressions using the fact that the projection operator $P_\omega$ is Hermitian and satisfies $P^2_\omega = P_\omega$. This expression results in

$$P(\omega) = \sum_{\omega_j=\omega} |\langle \omega_j | \psi \rangle|^2$$ (4.5)

That is, when one has degenerate eigenvalues, the relative probability of obtaining a degenerate eigenvalue is the sum of the relative probabilities for all the states corresponding to that eigenvalue.
In the absence of degeneracy, the above generalization simplifies to our original postulate because $P_\omega = |\omega\rangle\langle\omega|:

$$|\langle\omega|\psi\rangle|^2 = \langle\psi|\omega\rangle\langle\omega|\psi\rangle = \langle\psi|P_\omega|\psi\rangle = \langle\psi|P_\omega P_\omega|\psi\rangle = |P_\omega|\psi\rangle|^2 \quad (4.6)$$

An interesting thing about degeneracy is that states may not completely collapse upon measurement. The measurement will apply the projection operator from Equation 4.4. Therefore, if the particle begins in a state that has non-zero expansion coefficients for more than one of the $|\omega_j\rangle$, then it will retain those nonzero expansion coefficients for all $|\omega_j\rangle$ that correspond to the measured eigenvalue $\omega$. That is,

$$\text{if} \quad |\psi\rangle = \sum_{\omega_j=\omega} c_j |\omega_j\rangle + \sum_{\omega_j \neq \omega} c_j |\omega_j\rangle$$

$$\text{then} \quad P_\omega |\psi\rangle = \left[ \sum_{\omega_k=\omega} |\omega_k\rangle\langle\omega_k| \right] \left[ \sum_{\omega_j=\omega} c_j |\omega_j\rangle + \sum_{\omega_j \neq \omega} c_j |\omega_j\rangle \right] = \sum_{\omega_j=\omega} c_j |\omega_j\rangle$$

The state is collapsed to the subspace $\mathbb{V}_\omega$, but not to a single eigenstate $|\omega\rangle$. One then has to renormalize the resulting state.
Normalization of probabilities

Let us consider three cases:

- **finite-dimensional case:**
  For the finite-dimensional case, the assumption that the relative probability of outcome $\omega$ is given by $|\langle \omega | \psi \rangle|^2$, combined with the very reasonable assumption that there must be some outcome, immediately implies that that absolute probability of outcome $\omega$ is

$$P(\omega_j) = \frac{|\langle \omega_j | \psi \rangle|^2}{\sum_{j=1}^{n} |\langle \omega_j | \psi \rangle|^2}$$  \hspace{1cm} (4.7)

In fact, for a properly normalized state, the denominator is trivial:

$$\sum_{j=1}^{n} |\langle \omega_j | \psi \rangle|^2 = \sum_{j=1}^{n} |\langle \psi | \omega_j \rangle|^2 = \langle \psi | \psi \rangle = 1$$

via the completeness relation $I = \sum_{j=1}^{n} |\omega_j \rangle \langle \omega_j |$. 

Section 4.4

Postulates Revisited: Postulate 3: Results of Measurements of Classical Variables (cont.)
Postulate 3: Results of Measurements of Classical Variables (cont.)

- infinite-dimensional, but considering an operator whose eigenvalues are discretized (though possibly infinite in number)
  e.g., the $K$ operator for our example of functions on the interval $[a, b]$: the above rule continues to hold exactly. The denominator is guaranteed to remain finite in spite of the infinite sum because it is the normalization of the state $|\psi\rangle$.

- infinite-dimensional case and considering an operator whose eigenvalues are a continuum (uncountably infinite)
  e.g., the $X$ operator for our example of functions on the interval $[a, b]$: we must reinterpret the expansion coefficients as a probability density. That is, the probability of obtaining from the measurement corresponding to $\Omega$ a value between $\omega$ and $\omega + d\omega$ is

$$P(\omega)\,d\omega = \frac{\langle \omega | \psi \rangle^2 \,d\omega}{\int_{\omega_-}^{\omega_+} d\omega \,|\langle \omega | \psi \rangle|^2}$$  (4.8)

where $\omega_-$ and $\omega_+$ are the minimum and maximum allowed values of $\omega$, which might be $\pm\infty$. Equivalently, the probability of obtaining a value in the interval $[\omega_1, \omega_2]$ is

$$P(\omega_1 \leq \omega \leq \omega_2) = \frac{\int_{\omega_1}^{\omega_2} d\omega \,|\langle \omega | \psi \rangle|^2}{\int_{\omega_-}^{\omega_+} d\omega \,|\langle \omega | \psi \rangle|^2}$$  (4.9)
What happens when we have states that are delta-function normalized? For our
example $\Omega = X$, and $|\psi\rangle = |x_0\rangle$ and eigenstate, we have

$$\int_a^b dx \langle x | \psi \rangle^2 = \int_a^b dx \langle x_0 | x \rangle \langle x | x_0 \rangle = \int_a^b dx \delta(x_0 - x) \delta(x - x_0) = \delta(0)$$

which is infinite. Let’s reconsider this expression in the context of Equation 4.9,
moving the normalizing factor to the left side:

$$\left[ \int_a^b dx \langle x | x_0 \rangle^2 \right] P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} dx \langle x | x_0 \rangle^2$$

$$\delta(0)P(x_1 < x < x_2) = \begin{cases} 
\delta(0) & \text{if } x_1 \leq x_0 \leq x_2 \\
0 & \text{if } x_0 < x_1 \text{ or } x_0 > x_2 
\end{cases}$$

$$P(x_1 < x < x_2) = \begin{cases} 
1 & \text{if } x_1 \leq x_0 \leq x_2 \\
0 & \text{if } x_0 < x_1 \text{ or } x_0 > x_2 
\end{cases}$$

where we have taken the somewhat unrigorous step of dividing both sides by
$\delta(0)$ (both sides could be put under an integral sign with an arbitrary function
to be more rigorous). The point is that, while the normalizing factor is formally
infinite, this is only a problem when one considers the differential expression;
one can obtain reasonable results for the probability in any finite interval, which
is what is experimentally accessible.
The above analysis works even if $x_1 \to -\infty$ and $x_2 \to +\infty$.

One runs into a similar problem with the $K$ eigenstates for the infinite interval, but again one can obtain sensible results by only considering probabilities integrated over some finite range.

Moreover, relative probabilities are always well-defined because the infinite normalizing denominator cancels out. For our example, one still gets infinities, but they are sensible infinities: if, given a particle in the eigenstate $|x_0\rangle$, one wants to compare the probability of the particle being in the intervals $[x_1, x_2]$ and $[x_3, x_4]$ that do not overlap, clearly it can only be in one or the other, so the ratio of the two probabilities must either be infinite or zero.

In addition, we can always create states that are reasonably normalized – these simply will not be eigenstates of $X$ or $P$. 
Wavefunctions

Given some continuous eigenvalue $\omega$, the quantity $\langle \omega | \psi \rangle$ can be considered a function of the continuous variable $\omega$. It is conventional to call this quantity the wavefunction and write it as $\psi(\omega)$. The most common use of this nomenclature is for $\psi(x) = \langle x | \psi \rangle$, but it could also be used for $\psi_k(k) = \langle k | \psi \rangle$ when $k$ is continuous. The use of a notation like $\psi(\omega)$ can be confusing because the function $\psi$ is different depending on which operator the eigenvalues correspond to — e.g., above, $\psi(x)$ and $\psi_k(k)$ are in general very different functions — so, the argument of the function, which is normally a dummy variable, means something. To be clear, we will frequently use the labeling $\psi_\omega(\omega) = \langle \omega | \psi \rangle$ so that $\psi_x(x) = \langle x | \psi \rangle$ and $\psi_k(k) = \langle k | \psi \rangle$. This makes it clear what basis we are projecting onto and thus what quantity the “wavefunction” is a function of.

We will refer to $\psi_x(x) = \langle x | \psi \rangle$ as the position-space or coordinate-space wavefunction and $\psi_k(k) = \langle k | \psi \rangle$ as the k-space, or once we have added a $\hbar$ to turn $K$ into the momentum operator $P$, the momentum-space wavefunction. When the eigenvalue is discretized, such as $k$ on a finite interval $[a, b]$, we tend not to use the “wavefunction” language, but this is just semantics and a reluctance to call a quantity that is defined on discrete points a function. There is no truly fundamental difference between a “wavefunction” and the set of expansion coefficients of a state in a particular basis.
Commuting and Non-Commuting Operators

We now also see the physical relevance of whether two operators corresponding to physical observables commute. Let us first neglect degeneracies. If two Hermitian operators $\Omega$ and $\Lambda$ commute, then, as we proved in Section 3.6, there is a set of common eigenstates $\{|j\rangle\}$ that have eigenvalues $\{\omega_j\}$ and $\{\lambda_j\}$. If $|\psi\rangle$ is an eigenstate $|j\rangle$, then measurements of $\Omega$ and $\Lambda$ will yield the definite values $\omega_j$ and $\lambda_j$. If $|\psi\rangle$ is not an eigenstate, then the measurement outcomes will be correlated: if $\Omega$ yields $\omega_j$, then $\Lambda$ yields $\lambda_j$ because the projection operator $P_{\omega=\omega_j}$ is the same as the projection operator $P_{\lambda=\lambda_j}$. The relative probabilities $P(\omega_j)$ and $P(\lambda_j)$ will of course be equal.

If there are degeneracies, then the correspondence may break down because of incompletely overlapping subspaces. But this is completely consistent with the above statement; what occurs would just be the result of there being multiple eigenstates that contribute to a given $P(\omega)$.

Our archetypal example of two non-commuting operators is $X$ and $P$, which we proved in Section 3.9 (up to a factor of $\hbar$) gives $[X, P] = i\hbar$. These clearly do not commute, implying that there are no states that have definite values of both $X$ and $P$. 
Expectation Values and Uncertainties

Because measurement outcomes are probabilistic, the next most definite quantities to consider are probability-weighted moments of the measurements. The expectation value of an operator $\Omega$ is simply the probability-weighted mean outcome,

$$\langle \Omega \rangle = \sum_j P(\omega_j) \omega_j \quad \text{or} \quad \langle \Omega \rangle = \int_{\omega_-}^{\omega_+} d\omega \, P(\omega) \, \omega$$

(4.10)

We can write this explicitly in terms of $\Omega$ and the state $|\psi\rangle$:

$$\langle \Omega \rangle = \sum_j |\langle \omega_j | \psi \rangle|^2 \omega_j = \sum_j \langle \psi | \omega_j \rangle \langle \omega_j | \psi \rangle \omega_j = \sum_j \langle \psi | \Omega | \omega_j \rangle \langle \omega_j | \psi \rangle$$

(4.11)

$$= \langle \psi | \Omega | \psi \rangle$$

where we used completeness to make the last step. A similar derivation holds for the continuous $\omega$ version so that the same result $\langle \Omega \rangle = \langle \psi | \Omega | \psi \rangle$ holds.
The next moment to consider is the variance of $\omega$, which is conventionally defined as

$$\langle (\Delta \Omega)^2 \rangle = \sum_j P(\omega_j) (\omega_j - \langle \Omega \rangle)^2 \quad \text{or} \quad \langle (\Delta \Omega)^2 \rangle = \int_{\omega_-}^{\omega_+} d\omega \, P(\omega) (\omega - \langle \Omega \rangle)^2$$

(4.12)

Let’s pursue this in the discretized case. First, the above expression can be simplified:

$$\sum_j P(\omega_j) (\omega_j - \langle \Omega \rangle)^2 = \sum_j P(\omega_j) \omega_j^2 - 2\langle \Omega \rangle \sum_j P(\omega_j) \omega_j + \langle \Omega \rangle^2 \sum_j P(\omega_j)$$

$$= \left[ \sum_j P(\omega_j) \omega_j^2 \right] - \langle \Omega \rangle^2$$

where we used the definition of $\langle \Omega \rangle$ to reduce the second term and the normalization of the probability to reduce the third term.
Let's write out the first term in terms of $\Omega$ and $|\psi\rangle$.

$$\sum_{j} P(\omega_j) \omega_j^2 = \sum_{j} \langle \psi | \omega_j \rangle \langle \omega_j | \psi \rangle \omega_j^2 = \sum_{j} \langle \psi | \Omega | \omega_j \rangle \langle \omega_j | \Omega | \psi \rangle = \langle \psi | \Omega^2 | \psi \rangle$$

where we again used completeness. So we have

$$\langle (\Delta \Omega)^2 \rangle = \langle \psi | \Omega^2 | \psi \rangle - \langle \Omega \rangle^2 = \langle \psi | [\Omega^2 - \langle \Omega \rangle^2] | \psi \rangle = \langle \psi | [\Omega - \langle \Omega \rangle]^2 | \psi \rangle \quad (4.13)$$

where we have written three algebraically equivalent forms by using $\langle \psi | \psi \rangle = 1$ and the kind of conversion between $\Omega^2 - \langle \Omega \rangle^2$ and $[\Omega - \langle \Omega \rangle]^2$ that we used on the previous page.