4

The Postulates—a General Discussion

Having acquired the necessary mathematical training and physical motivation, you are now ready to get acquainted with the postulates of quantum mechanics. In this chapter the postulates will be stated and discussed in broad terms to bring out the essential features of quantum theory. The subsequent chapters will simply be applications of these postulates to the solution of a variety of physically interesting problems. Despite your preparation you may still find the postulates somewhat abstract and mystifying on this first encounter. These feelings will, however, disappear after you have worked with the subject for some time.

4.1. The Postulates‡

The following are the postulates of nonrelativistic quantum mechanics. We consider first a system with one degree of freedom, namely, a single particle in one space dimension. The straightforward generalization to more particles and higher dimensions will be discussed towards the end of the chapter. In what follows, the quantum postulates are accompanied by their classical counterparts (in the Hamiltonian formalism) to provide some perspective.

	Classical Mechanics		Quantum Mechanics
I.	The state of a particle at any given time is specified by the two variables $x(t)$ and $p(t)$, i.e., as a point in a two-dimensional phase space.	I.	The state of the particle is represented by a vector $\|\psi(t)\rangle$ in a Hilbert space.
II.	Every dynamical variable ω is a function of x and p: $\omega = \omega(x, p)$.	II.	The independent variables x and p of classical mechanics are represented

‡ Recall the discussion in the Preface regarding the sense in which the word is used here.

by Hermitian operators X and P with the following matrix elements in the eigenbasis of X‡

$$\langle x|X|x'\rangle = x\delta(x-x')$$
$$\langle x|P|x'\rangle = -i\hbar\delta'(x-x')$$

The operators corresponding to dependent variables $\omega(x,p)$ are given Hermitian operators

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

III. If the particle is in a state given by x and p, the measurement‖ of the variable ω will yield a value $\omega(x,p)$. The state will remain unaffected.

III. If the particle is in a state $|\psi\rangle$, measurement‖ of the variable (corresponding to) Ω will yield one of the eigenvalues ω with probability $P(\omega)\propto|\langle\omega|\psi\rangle|^2$. The state of the system will change from $|\psi\rangle$ to $|\omega\rangle$ as a result of the measurement.

IV. The state variables change with time according to Hamilton's equations:

$$\dot{x} = \frac{\partial\mathscr{H}}{\partial p}$$

$$\dot{p} = -\frac{\partial\mathscr{H}}{\partial x}$$

IV. The state vector $|\psi(t)\rangle$ obeys the *Schrödinger equation*

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle$$

where $H(X,P) = \mathscr{H}(x\to X,\ p\to P)$ is the quantum Hamiltonian operator and $\mathscr{H}$ is the Hamiltonian for the corresponding classical problem.

4.2. Discussion of Postulates I–III

The postulates (of classical and quantum mechanics) fall naturally into two sets: the first three, which tell us how the system is depicted at a given time, and the last, which specifies how this picture changes with time. We will confine our attention to the first three postulates in this section, leaving the fourth for the next.

The first postulate states that a particle is described by a ket $|\psi\rangle$ in a Hilbert space which, you will recall, contains *proper vectors* normalizable to unity as well as

‡ Note that the X operator is the same one discussed at length in Section 1.10. Likewise $P=\hbar K$, where K was also discussed therein. You may wish to go over that section now to refresh your memory.

§ By this we mean that Ω is the same function of X and P as ω is of x and p.

‖ That is, in an ideal experiment consistent with the theory. It is assumed you are familiar with the ideal classical measurement which can determine the state of the system without disturbing it in any way. A discussion of ideal quantum measurements follows.

improper vectors, normalizable only to the Dirac delta functions.‡ Now, a ket in such a space has in general an infinite number of components in a given basis. One wonders why a particle, which had only two independent degrees of freedom, x and p, in classical mechanics, now needs to be specified by an infinite number of variables. What do these variables tell us about the particle? To understand this we must go on to the next two postulates, which answer exactly this question. For the present let us note that the double-slit experiment has already hinted to us that a particle such as the electron needs to be described by a wave function $\psi(x)$. We have seen in Section 1.10 that a function $f(x)$ may be viewed as a ket $|f\rangle$ in a Hilbert space. The ket $|\psi\rangle$ of quantum mechanics is none other than the vector representing the probability amplitude $\psi(x)$ introduced in the double-slit experiment.

When we say that $|\psi\rangle$ is an element of a vector space we mean that if $|\psi\rangle$ and $|\psi'\rangle$ represent possible states of a particle so does $\alpha|\psi\rangle+\beta|\psi'\rangle$. This is called the *principle of superposition*. The principle by itself is not so new: we know in classical physics, for example, that if $f(x)$ and $g(x)$ [with $f(0)=f(L)=g(0)=g(L)=0$] are two possible displacements of a string, so is the superposition $\alpha f(x)+\beta g(x)$. What *is* new is the interpretation of the superposed state $\alpha|\psi\rangle+\beta|\psi'\rangle$. In the case of the string, the state $\alpha f+\beta g$ has very different attributes from the states f and g: it will look different, have a different amount of stored elastic energy, and so on. In quantum theory, on the other hand, the state $\alpha|\psi\rangle+\beta|\psi'\rangle$ will, loosely speaking, have attributes that sometimes resemble that of $|\psi\rangle$ and at other times those of $|\psi'\rangle$. There is, however, no need to speak loosely, since we have postulates II and III to tell us exactly how the state vector $|\psi\rangle$ is to be interpreted in quantum theory. Let us find out.

In classical mechanics when a state (x, p) is given, one can say that any dynamical variable ω has a value $\omega(x, p)$, in the sense that if the variable is measured the result $\omega(x, p)$ will obtain. What is the analogous statement one can make in quantum mechanics given that the particle is in a state $|\psi\rangle$? The answer is provided by Postulates II and III, in terms of the following steps:

Step 1. Construct the corresponding quantum operator $\Omega=\omega(x\to X,\ p\to P)$, where X and P are the operators defined in postulate II.

Step 2. Find the orthonormal eigenvectors $|\omega_i\rangle$ and eigenvalues ω_i of Ω.

Step 3. Expand $|\psi\rangle$ in this basis:

$$|\psi\rangle=\sum_i |\omega_i\rangle\langle\omega_i|\psi\rangle$$

Step 4. The probability $P(\omega)$ that the result ω will obtain is proportional to the modulus squared of the projection of $|\psi\rangle$ along the eigenvector $|\omega\rangle$, that is $P(\omega)\propto|\langle\omega|\psi\rangle|^2$. In terms of the projection operator $\mathbb{P}_\omega=|\omega\rangle\langle\omega|$, $P(\omega)\propto|\langle\omega|\psi\rangle|^2=\langle\psi|\omega\rangle\langle\omega|\psi\rangle=\langle\psi|\mathbb{P}_\omega|\psi\rangle=\langle\psi|\mathbb{P}_\omega\mathbb{P}_\omega|\psi\rangle=\langle\mathbb{P}_\omega\psi|\mathbb{P}_\omega\psi\rangle$.

There is a tremendous amount of information contained in these steps. Let us note, for the present, the following salient points.

‡ The status of the two classes will be clarified later in this chapter.

(1) The theory makes only probabilistic predictions for the result of a measurement of Ω. Further, it assigns (relative) probabilities only for obtaining some eigenvalue ω of Ω. *Thus the only possible values of* Ω *are its eigenvalues.* Since postulate II demands that Ω be Hermitian, these eigenvalues are all real.

(2) Since we are told that $P(\omega_i) \propto |\langle \omega_i | \psi \rangle|^2$, the quantity $|\langle \omega_i | \psi \rangle|^2$ is only the relative probability. To get the absolute probability, we divide $|\langle \omega_i | \psi \rangle|^2$ by the sum of all relative probabilities:

$$P(\omega_i) = \frac{|\langle \omega_i | \psi \rangle|^2}{\sum_j |\langle \omega_j | \psi \rangle|^2} = \frac{|\langle \omega_i | \psi \rangle|^2}{\langle \psi | \psi \rangle} \tag{4.2.1}$$

It is clear that if we had started with a normalized state

$$|\psi'\rangle = \frac{|\psi\rangle}{\langle \psi | \psi \rangle^{1/2}}$$

we would have had

$$P(\omega_i) = |\langle \omega_i | \psi' \rangle|^2 \tag{4.2.2}$$

If $|\psi\rangle$ is a proper vector, such a rescaling is possible and will be assumed hereafter. The probability interpretation breaks down if $|\psi\rangle$ happens to be one of the improper vectors in the space, for in this case $\langle \psi | \psi \rangle = \delta(0)$ is the only sensible normalization. The status of such vectors will be explained in Example 4.2.2 below.

Note that the condition $\langle \psi | \psi \rangle = 1$ is a matter of convenience and not a physical restriction on the proper vectors. (In fact the set of all normalized vectors does not even form a vector space. If $|\psi\rangle$ and $|\psi'\rangle$ are normalized, then an arbitrary linear combination, $\alpha|\psi\rangle + \beta|\psi'\rangle$, is not.)

Note that the relative probability distributions corresponding to the states $|\psi\rangle$ and $\alpha|\psi\rangle$, when they are renormalized to unity, reduce to the same absolute probability distribution. Thus, corresponding to each physical state, there exists not one vector, but a *ray* or "direction" in Hilbert space. When we speak of the state of the particle, we usually mean the ket $|\psi\rangle$ with unit norm. Even with the condition $\langle \psi | \psi \rangle = 1$, we have the freedom to multiply the ket by a number of the form $e^{i\theta}$ without changing the physical state. This freedom will be exploited at times to make the components of $|\psi\rangle$ in some basis come out real.

(3) If $|\psi\rangle$ is an eigenstate $|\omega_i\rangle$, the measurement of Ω is guaranteed to yield the result ω_i. A particle in such a state may be said to have a value ω_i for Ω in the classical sense.

(4) When two states $|\omega_1\rangle$ and $|\omega_2\rangle$ are superposed to form a (normalized) state, such as

$$|\psi\rangle = \frac{\alpha|\omega_1\rangle + \beta|\omega_2\rangle}{(|\alpha|^2 + |\beta|^2)^{1/2}}$$

one gets the state, which upon measurement of Ω, can yield either ω_1 or ω_2 with probabilities $|\alpha|^2/(|\alpha|^2 + |\beta|^2)$ and $|\beta|^2/(|\alpha|^2 + |\beta|^2)$, respectively. This is the peculiar

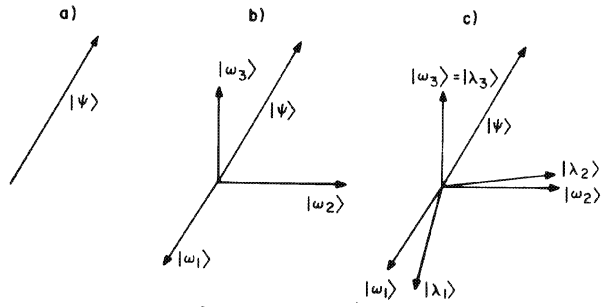

Figure 4.1. (a) The normalized ket in $\mathbb{V}^3(R)$ representing the state of the particle. (b) The Ω basis, $|\omega_1\rangle$, $|\omega_2\rangle$, and $|\omega_3\rangle$. (c) The Ω and the Λ bases. To get the statistical information on a variable, we find the eigenvectors of the corresponding operator and project $|\psi\rangle$ on that basis.

consequence of the superposition principle in quantum theory, referred to earlier. It has no analog in classical mechanics. For example, if a dynamical variable of the string in the state $\alpha f + \beta g$ is measured, one does not expect to get the value corresponding to f some of the time and that corresponding to g the rest of the time; instead, one expects a unique value generally distinct from both. Likewise, the functions f and αf (α real) describe two distinct configurations of the string and are not physically equivalent.

(5) When one wants information about another variable Λ, one repeats the whole process, finding the eigenvectors $|\lambda_i\rangle$ and the eigenvalues λ_i. Then

$$P(\lambda) = |\langle \lambda | \psi \rangle|^2$$

The bases of Ω and Λ will of course be different in general. In summary, we have a single ket $|\psi\rangle$ representing the state of the particle in Hilbert space, and it contains the statistical prediction for all observables. To extract this information for any observable, we must determine the eigenbasis of the corresponding operator and find the projection of $|\psi\rangle$ along all its eigenkets.

(6) As our interest switches from one variable Ω, to another, Λ, so does our interest go from the kets $|\omega\rangle$, to the kets $|\lambda\rangle$. There is, however, no need to change the basis each time. Suppose for example we are working in the Ω basis in which

$$|\psi\rangle = \sum_i |\omega_i\rangle\langle\omega_i|\psi\rangle$$

and $P(\omega_i) = |\langle \omega_i | \psi \rangle|^2$. If we want $P(\lambda_i)$ we take the operator Λ (which is some given matrix with elements $\Lambda_{ij} = \langle \omega_i | \Lambda | \omega_j \rangle$); find its eigenvectors $|\lambda_i\rangle$ (which are column vectors with components $\langle \omega_j | \lambda_i \rangle$), and take the inner product $\langle \lambda_i | \psi \rangle$ in this basis:

$$\langle \lambda_i | \psi \rangle = \sum_j \langle \lambda_i | \omega_j \rangle \langle \omega_j | \psi \rangle$$

Example 4.2.1. Consider the following example from a fictitious Hilbert space $\mathbb{V}^3(R)$ (Fig. 4.1). In Fig. 4.1a we have the normalized state $|\psi\rangle$, with no reference

to any basis. To get predictions on Ω, we find its eigenbasis and express the state vector $|\psi\rangle$ in terms of the orthonormal eigenvectors $|\omega_1\rangle$, $|\omega_2\rangle$, and $|\omega_3\rangle$ (Fig. 4.1b). Let us suppose

$$|\omega\rangle=\frac{1}{2}|\omega_1\rangle+\frac{1}{2}|\omega_2\rangle+\frac{1}{2^{1/2}}|\omega_3\rangle$$

This means that the values ω_1, ω_2, and ω_3 are expected with probabilities $\frac{1}{4}$, $\frac{1}{4}$, and $\frac{1}{2}$, respectively, *and other values of ω are impossible.* If instead $|\psi\rangle$ were some eigenvector, say $|\omega_1\rangle$, then the result ω_1 would obtain with unit probability. Only a particle in a state $|\psi\rangle=|\omega_i\rangle$ has a well-defined value of Ω in the classical sense. If we want $P(\lambda_i)$ we construct the basis $|\lambda_1\rangle$, $|\lambda_2\rangle$, and $|\lambda_3\rangle$, which can in general be distinct from the Ω basis. In our example (Fig. 4.1c) there is just one common eigenvector $|\omega_3\rangle=|\lambda_3\rangle$. □

Returning to our main discussion, there are a few complications that could arise as one tries to carry out the steps 1–4. We discuss below the major ones and how they are to be surmounted.

Complication 1: The Recipe $\Omega=\omega(x\to X, p\to P)$ Is Ambiguous. If, for example, $\omega=xp$, we don't know if $\Omega=XP$ or PX since $xp=px$ classically. There is no universal recipe for resolving such ambiguities. In the present case, the rule is to use the symmetric sum: $\Omega=(XP+PX)/2$. Notice incidentally that symmetrization also renders Ω Hermitian. Symmetrization is the answer as long as Ω does not involve products of two or more powers of X with two or more powers of P. If it does, only experiment can decide the correct prescription. We will not encounter such cases in this book.

Complication 2: The Operator Ω Is Degenerate. Let us say $\omega_1=\omega_2=\omega$. What is $P(\omega)$ in this case? We select some orthonormal basis $|\omega, 1\rangle$ and $|\omega, 2\rangle$ in the eigenspace $\mathbb{V}_\omega$ with eigenvalue ω. Then

$$P(\omega)=|\langle\omega, 1|\psi\rangle|^2+|\langle\omega, 2|\psi\rangle|^2$$

which is the modulus squared of the projection of $|\psi\rangle$ in the degenerate eigenspace. This is the result we will get if we assume that ω_1 and ω_2 are infintesimally distinct and ask for $P(\omega_1$ or $\omega_2)$. In terms of the projection operator for the *eigenspace*,

$$\mathbb{P}_\omega=|\omega, 1\rangle\langle\omega, 1|+|\omega, 2\rangle\langle\omega, 2| \tag{4.2.3a}$$

we have

$$P(\omega)=\langle\psi|\mathbb{P}_\omega|\psi\rangle=\langle\mathbb{P}_\omega\psi|\mathbb{P}_\omega\psi\rangle \tag{4.2.3b}$$

In general, one can replace in Postulate III

$$P(\omega)\propto\langle\psi|\mathbb{P}_\omega|\psi\rangle$$

where $\mathbb{P}_\omega$ is the projection operator for the eigenspace with eigenvalue ω. Then postulate III as stated originally would become a special case in which there is no degeneracy and each eigenspace is simply an eigenvector.

In our example from $\mathbb{V}^3(R)$, if $\omega_1 = \omega_2 = \omega$ (Fig. 4.1b) then $P(\omega)$ is the square of the component of $|\psi\rangle$ in the "$x\ y$" plane.

Complication 3: The Eigenvalue Spectrum of Ω Is Continuous. In this case one expands $|\psi\rangle$ as

$$|\psi\rangle = \int |\omega\rangle\langle\omega|\psi\rangle\, d\omega$$

One expects that as ω varies continuously, so will $\langle\omega|\psi\rangle$, that is to say, one expects $\langle\omega|\psi\rangle$ to be a smooth function $\psi(\omega)$. To visualize this function one introduces an auxiliary one-dimensional space, called the ω space, the points in which are labeled by the coordinate ω. In this space $\psi(\omega)$ will be a smooth function of ω and is called the *wave function in the ω space*. We are merely doing the converse of what we did in Section 1.10 wherein we started with a function $f(x)$ and tried to interpret it as the components of an infinite-dimensional ket $|\psi\rangle$ in the $|x\rangle$ basis. As far as the state vector $|\psi\rangle$ is concerned, there is just one space, the Hilbert space, in which it resides. The ω space, the λ space, etc. are auxiliary manifolds introduced for the purpose of visualizing the components of the infinite-dimensional vector $|\psi\rangle$ in the Ω basis, the Λ basis, and so on. The wave function $\psi(\omega)$ is also called the *probability amplitude* for finding the particle with $\Omega = \omega$.

Can we interpret $|\langle\omega|\psi\rangle|^2$ as the probability for finding the particle with a value ω for Ω? No. Since the number of possible values for ω is infinite and the total probability is unity, each single value of ω can be assigned only an infinitesimal probability. One interprets $P(\omega) = |\langle\omega|\psi\rangle|^2$ to be the *probability density* at ω, by which one means that $P(\omega)\, d\omega$ is the probability of obtaining a result between ω and $\omega + d\omega$. This definition meets the requirement that the total probability be unity, since

$$\int P(\omega)\, d\omega = \int |\langle\omega|\psi\rangle|^2\, d\omega = \int \langle\psi|\omega\rangle\langle\omega|\psi\rangle\, d\omega$$
$$= \langle\psi|I|\psi\rangle = \langle\psi|\psi\rangle = 1 \tag{4.2.4}$$

If $\langle\psi|\psi\rangle = \delta(0)$ is the only sensible normalization possible, the state cannot be normalized to unity and $P(\omega)$ must be interpreted as the *relative probability density*. We will discuss such improper states later.

An important example of a continuous spectrum is that of X, the operator corresponding to the position x. The wave function in the X basis (or the x space), $\psi(x)$, is usually referred to as just the wave function, since the X basis is almost always what one uses. In our discussions in the last chapter, $|\psi(x)|^2$ was referred to as the probability for finding the particle *at* a given x, rather than as the probability density, in order to avoid getting into details. Now the time has come to become precise!

Earlier on we were wondering why it was that a classical particle defined by just two numbers x and p now needs to be described by a ket which has an infinite number of components. The answer is now clear. A classical particle has, at any given time, a definite position. One simply has to give this value of x in specifying the state. A quantum particle, on the other hand, can take on any value of x upon measurement and one must give the relative probabilities for *all possible outcomes.* This is part of the information contained in $\psi(x) = \langle x | \psi \rangle$, the components of $|\psi\rangle$ in the X basis. Of course, in the case of the classical particle, one needs also to specify the momentum p as well. In quantum theory one again gives the odds for getting different values of momenta, but one doesn't need a new vector for specifying this; the same ket $|\psi\rangle$ when expanded in terms of the eigenkets $|p\rangle$ of the momentum operator P gives the odds through the wave function in p space, $\psi(p) = \langle p | \psi \rangle$.

Complication 4: The Quantum Variable Ω Has No Classical Counterpart. Even "point" particles such as the electron are now known to carry "spin," which is an internal angular momentum, that is to say, angular momentum unrelated to their motion through space. Since such a degree of freedom is absent in classical mechanics, our postulates do not tell us which operator is to describe this variable in quantum theory. As we will see in Chapter 14, the solution is provided by a combination of intuition and semi-classical reasoning. It is worth bearing in mind that no matter how diligently the postulates are constructed, they must often be supplemented by intuition and classical ideas.

Having discussed the four-step program for extracting statistical information from the state vector, we continue with our study of what else the postulates of quantum theory tell us.

Collapse of the State Vector

We now examine another aspect of postulate III, namely, that the measurement of the variable Ω changes the state vector, which is in general some superposition of the form

$$|\psi\rangle = \sum_{\omega} |\omega\rangle\langle\omega|\psi\rangle$$

into the eigenstate $|\omega\rangle$ corresponding to the eigenvalue ω obtained in the measurement. This phenomenon is called the *collapse or reduction of the state vector.*

Let us first note that any definitive statement about the impact of the measurement process presupposes that the measurement process is of a definite kind. For example, the classical mechanics maxim that any dynamical variable can be measured without changing the state of the particle, assumes that the measurement is an ideal measurement (consistent with the classical scheme). But one *can* think up *nonideal* measurements which *do* change the state; imagine trying to locate a chandelier in a dark room by waving a broom till one makes contact. What makes Postulate III profound is that the measurement process referred to there is an *ideal quantum measurement*, which in a sense is the best one can do. We now illustrate the notion of an ideal quantum measurement and the content of this postulate by an example.

Consider a particle in a momentum eigenstate $|p\rangle$. The postulate tells us that if the momentum in this state is measured we are assured a result p, and that the state will be the same after the measurement (since $|\psi\rangle = |p\rangle$ is already an eigenstate of the operator P in question). One way to measure the momentum of the particle is by *Compton scattering*, in which a photon of definite momentum bounces off the particle.

Let us assume the particle is forced to move along the x-axis and that we send in a right-moving photon of energy $\hbar\omega$ that bounces off the particle and returns as a left-moving photon of energy $\hbar\omega'$. (How do we know what the photon energies are? We assume we have atoms that are known to emit and absorb photons of any given energy.) Using momentum and energy conservation:

$$cp' = cp + \hbar(\omega + \omega')$$

$$E' = E + \hbar(\omega - \omega')$$

it is now possible from this data to reconstruct the initial and final momenta of the particle:

$$cp = -\frac{(\hbar\omega + \hbar\omega')}{2} + \sqrt{1 + \frac{m^2c^4}{\hbar^2\omega\omega'}}\,\frac{\hbar\omega - \hbar\omega'}{2}$$

$$cp' = \frac{(\hbar\omega + \hbar\omega')}{2} + \sqrt{1 + \frac{m^2c^4}{\hbar^2\omega\omega'}}\,\frac{\hbar\omega - \hbar\omega'}{2}$$

Solving for ω' and p' in terms of ω and p, one readily sees that for any choice of p, if $\omega \to 0$, then so does ω'. Thus one can always make the change in momentum $p' - p$ arbitrarily small. Hereafter, when we speak of a momentum measurement, this is what we will mean. We will also assume that to each dynamical variable there exists a corresponding ideal measurement. We will discuss, for example, the ideal position measurement, which, when conducted on a particle in state $|x\rangle$, will give the result x with unit probability and leave the state vector unchanged.

Suppose now that we measure the *position* of a particle in a *momentum* eigenstate $|p\rangle$. Since $|p\rangle$ is a sum of position eigenkets $|x\rangle$,

$$|p\rangle = \int |x\rangle\langle x|p\rangle\, dx$$

the measurement will force the system into some state $|x\rangle$. Thus even the *ideal* position measurement will change the state which is not a position eigenstate. Why does a position measurement alter the state $|p\rangle$, while momentum measurement does not? The answer is that an ideal position measurement uses photons of infinitely high momentum (as we will see) while an ideal momentum measurement uses photons of infinitesimally low momentum (as we have seen).

This then is the big difference between classical and quantum mechanics: an ideal measurement of any variable ω in classical mechanics leaves any state invariant,

whereas the ideal measurement of Ω in quantum mechanics leaves only the eigenstates of Ω invariant.

The effect of measurement may be represented schematically as follows:

$$|\psi\rangle \xrightarrow[\Omega \text{ measured, } \omega \text{ obtained}]{} = \frac{\mathbb{P}_\omega|\psi\rangle}{\langle \mathbb{P}_\omega\psi \mid \mathbb{P}_\omega\psi\rangle^{1/2}}$$

where $\mathbb{P}_\omega$ is the projection operator associated with $|\omega\rangle$, and the state after measurement has been normalized. If ω is degenerate,

$$|\psi\rangle \to \frac{\mathbb{P}_\omega|\psi\rangle}{\langle \mathbb{P}_\omega\psi \mid \mathbb{P}_\omega\psi\rangle^{1/2}}$$

where $\mathbb{P}_\omega$ is the projection operator for the eigenspace $\mathbb{V}_\omega$. Special note should be taken of the following point: if the initial state $|\psi\rangle$ were unknown, and the measurement yielded a degenerate eigenvalue ω, we could not say what the state was after the measurement, except that it was some state in the eigenspace with eigenvalue ω. On the other hand, if the initial state $|\psi\rangle$ were known, and the measurement yielded a degenerate value ω, the state after measurement is known to be $\mathbb{P}_\omega|\psi\rangle$ (up to normalization). Consider our example from $\mathbb{V}^3(R)$ (Fig. 4.1b). Say we had $\omega_1 = \omega_2 = \omega$. Let us use an orthonormal basis $|\omega, 1\rangle$, $|\omega, 2\rangle$, $|\omega_3\rangle$, where, as usual, the extra labels 1 and 2 are needed to distinguish the basis vectors in the degenerate eigenspace. If in this basis we know, for example, that

$$|\psi\rangle = \tfrac{1}{2}|\omega, 1\rangle + \tfrac{1}{2}|\omega, 2\rangle + (\tfrac{1}{2})^{1/2}|\omega_3\rangle$$

and the measurement gives a value ω, the normalized state after measurement is known to us to be

$$|\psi\rangle = 2^{-1/2}(|\omega, 1\rangle + |\omega, 2\rangle)$$

If, on the other hand, the initial state were unknown and a measurement gave a result ω, we could only say

$$|\psi\rangle = \frac{\alpha|\omega, 1\rangle + \beta|\omega, 2\rangle}{(\alpha^2 + \beta^2)^{1/2}}$$

where α and β are arbitrary real numbers.

Note that although *we* do not know what α and β are from the measurement, they are not arbitrary. In other words, the system had a well-defined state vector $|\psi\rangle$ before the measurement, though *we* did not know $|\psi\rangle$, and has a well-defined state vector $\mathbb{P}_\omega|\psi\rangle$ after the measurement, although all *we* know is that it lies within a subspace $\mathbb{V}_\omega$.

One of the outstanding features of classical mechanics is that it makes fully deterministic predictions. It may predict for example that a particle leaving $x=x_i$ with momentum p_i in some potential $V(x)$ will arrive 2 seconds later at $x=x_f$ with momentum $p=p_f$. To test the prediction we release the particle at $x=x_i$ with $p=p_i$ at $t=0$ and wait at $x=x_f$ and see if the particle arrives there with $p=p_f$ at $t=2$ seconds.

Quantum theory, on the other hand, makes statistical predictions about a particle in a state $|\psi\rangle$ and claims that this state evolves in time according to Schrödinger's equation. To test these predictions we must be able to

(1) Create particles in a well-defined state $|\psi\rangle$.
(2) Check the probabilistic predictions at any time.

The collapse of the state vector provides us with a good way of preparing definite states: we begin with a particle in an arbitrary state $|\psi\rangle$ and meaure a variable Ω. If we get a nondegenerate eigenvalue ω, we have in our hands the state $|\omega\rangle$. (If ω is degenerate, further measurement is needed. We are not ready to discuss this problem.) Notice how in quantum theory, measurement, instead of telling us what the system was doing *before* the measurement, tells us what it is doing just *after* the measurement. (Of course it does tell us that the original state had some projection on the state $|\omega\rangle$ obtained after measurement. But this information is nothing compared to the complete specifications of the state just *after* measurement.)

Anyway, assume we have prepared a state $|\omega\rangle$. If we measure some variable Λ, immediately thereafter, so that the state could not have changed from $|\omega\rangle$, and if say,

$$|\omega\rangle=\frac{1}{3^{1/2}}|\lambda_1\rangle+\left(\frac{2}{3}\right)^{1/2}|\lambda_2\rangle+0\cdot(\text{others})$$

the theory predicts that λ_1 and λ_2 will obtain with probabilities $1/3$ and $2/3$, respectively. If our measurement gives a λ_i, $i\neq 1, 2$ (or worse still a $\lambda\neq$ any eigenvalue!) that is the end of the theory. So let us assume we get one of the allowed values, say λ_1. This is consistent with the theory but does not fully corroborate it, since the odds for λ_1 could have been $1/30$ instead of $1/3$ and we could still get λ_1. Therefore, we must repeat the experiment many times. But we cannot repeat the experiment with *this* particle, since after the measurement the state of the particle is $|\lambda_1\rangle$. We must start afresh with another particle in $|\omega\rangle$. For this purpose we require a *quantum ensemble*, which consists of a large number N of particles *all in the same state* $|\omega\rangle$. If a measurement of Λ is made on every one of these particles, approximately $N/3$ will yield a value λ_1 and end up in the state $|\lambda_1\rangle$ while approximately $2N/3$ will yield a value λ_2 and end up in a state $|\lambda_2\rangle$. For sufficiently large N, the deviations from the fractions $1/3$ and $2/3$ will be negligible. The chief difference between a classical ensemble, of the type one encounters in, say, classical statistical mechanics, and the quantum ensemble referred to above, is the following. If in a classical ensemble of N particles $N/3$ gave a result λ_1 and $2N/3$ a result λ_2, one can think of the ensemble as having contained $N/3$ particles with $\lambda=\lambda_1$ and the others with $\lambda=\lambda_2$ *before* the

measurement. In a quantum ensemble, on the other hand, every particle is assumed to be in the same state $|\omega\rangle$ prior to measurement (i.e., every particle is potentially capable of yielding *either* result λ_1 or λ_2). Only after that measurement are a third of them forced into the state $|\lambda_1\rangle$ and the rest into $|\lambda_2\rangle$.

Once we have an ensemble, we can measure any other variable and test the expectations of quantum theory. We can also prepare an ensemble, let it evolve in time, and study it at a future time to see if the final state is what the Schrödinger equation tells us it should be.

Example 4.2.2. An example of an ensemble being used to test quantum theory was encountered in the double-slit experiment, say with photons. A given photon of momentum p and energy E was expected to hit the detectors with a probability density given by the oscillating function $|\psi(x)|^2$. One could repeat the experiment N times, sending one such photon at a time to see if the final number distribution indeed was given by $|\psi(x)|^2$. One could equally well send in a macroscopic, monochromatic beam of light of frequency $\omega = E/\hbar$ and wave number $k = p/\hbar$, which consists of a large number of photons of energy E and momentum p. If one makes the assumption (correct to a high degree) that the photons are noninteracting, sending in the beam is equivalent to experimenting with the ensemble. In this case the intensity pattern will take the shape of the probability density $|\psi(x)|^2$, the instant the beam is turned on. □

Example 4.2.3. The following example is provided to illustrate the distinction between the probabilistic descriptions of systems in classical mechanics and in quantum mechanics.

We choose as our classical system a six-faced die for which the probabilities $P(n)$ of obtaining a number n have been empirically determined. As our quantum system we take a particle in a state

$$|\psi\rangle = \sum_{i=1}^{6} C_i |\omega_i\rangle$$

Suppose we close our eyes, toss the die, and cover it with a mug. Its statistical description has many analogies with the quantum description of the state $|\psi\rangle$:

(1) The state of the die is described by a probability function $P(n)$ before the mug is lifted.
(2) The only possible values of n are 1, 2, 3, 4, 5, and 6.
(3) If the mug is lifted, and some value—say $n = 3$—is obtained, the function $P(n)$ collapses to δ_{n3}.
(4) If an ensemble of N such dice are thrown, $NP(n)$ of them will give the result n (as $N \to \infty$).

The corresponding statements for the particle in the state $|\psi\rangle$ are no doubt known to you. Let us now examine some of the key differences between the statistical descriptions in the two cases.

(1) It is possible, *at least in principle*, to predict exactly which face of the die will be on top, given the mass of the die, its position, orientation, velocity, and angular velocity at the time of release, the viscosity of air, the elasticity of the table top, and so on. The statistical description is, however, the only possibility in the quantum case, *even in principle*.

(2) If the result $n=3$ was obtained upon lifting the mug, it is consistent to assume that the die was in such a state *even prior to measurement*. In the quantum case, however, the state after measurement, say $|\omega_3\rangle$, is not the state before measurement, namely $|\psi\rangle$.

(3) If N such dice are tossed and covered with N mugs, there will be $NP(1)$ dice with $n=1$, $NP(2)$ dice with $n=2$, etc. in the ensemble *before and after the measurement*. In contrast, the quantum ensemble corresponding to $|\psi\rangle$ will contain N particles all of which are in the same state $|\psi\rangle$ (that is, each can yield any of the values $\omega_1, \ldots, \omega_6$) before the measurement, and $NP(\omega_i)$ particles in $|\omega_i\rangle$ *after* the measurement. Only the ensemble before the measurement represents the state $|\psi\rangle$. The ensemble after measurement is a mixture of six ensembles representing the states $|\omega_1\rangle, \ldots, |\omega_6\rangle$.‡ □

Having seen the utility of the ensemble concept in quantum theory, we now define and discuss the two statistical variables that characterize an ensemble.

Expectation Value

Given a large ensemble of N particles in a state $|\psi\rangle$, quantum theory allows us to predict what fraction will yield a value ω if the variable Ω is measured. This prediction, however, involves solving the eigenvalue problem of the operator Ω. If one is not interested in such detailed information on the state (or the corresponding ensemble) one can calculate instead an average over the ensemble, called the *expectation value*, $\langle\Omega\rangle$. The expectation value is just the mean value defined in statistics:

$$\begin{aligned}\langle\Omega\rangle &= \sum_i P(\omega_i)\omega_i = \sum_i |\langle\omega_i|\,\psi\rangle|^2\omega_i \\ &= \sum_i \langle\psi\,|\,\omega_i\rangle\langle\omega_i|\,\psi\rangle\omega_i \qquad (4.2.5)\end{aligned}$$

But for the factors ω_i multiplying each projection operator $|\omega_i\rangle\langle\omega_i|$, we could have used $\sum_i |\omega_i\rangle\langle\omega_i| = I$. To get around this, note that $\omega_i|\omega_i\rangle = \Omega|\omega_i\rangle$. Feeding this in and continuing, we get

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

Now we can use $\sum_i |\omega_i\rangle\langle\omega_i| = I$ to get

$$\langle\Omega\rangle = \langle\psi|\Omega|\psi\rangle \qquad (4.2.6)$$

‡ This is an example of a *mixed* ensemble. These will be discussed in the digression on density matrices, which follows in a while.

There are a few points to note in connection with this formula.

(1) To calculate $\langle\Omega\rangle$, one need only be given the state vector and the operator Ω (say as a column vector and a matrix, respectively, in some basis). There is no need to find the eigenvectors or eigenvalues of Ω.
(2) If the particle is in an eigenstate of Ω, that is $\Omega|\psi\rangle=\omega|\psi\rangle$, then $\langle\Omega\rangle=\omega$.
(3) By the average value of Ω we mean the average over the ensemble. A given particle will of course yield only one of the eigenvalues upon measurement. The mean value will generally be an inaccessible value for a single measurement unless it accidentally equals an eigenvalue. [A familiar example of this phenomenon is that of the mean number of children per couple, which may be 2.12, although the number in a given family is restricted to be an integer.]

The Uncertainty

In any situation described probabilistically, another useful quantity to specify besides the mean is the *standard deviation*, which measures the average fluctuation around the mean. It is defined as

$$\Delta\Omega=\langle(\Omega-\langle\Omega\rangle)^2\rangle^{1/2} \tag{4.2.7}$$

and often called the root-mean-squared deviation. In quantum mechanics, it is referred to as the *uncertainty in* Ω. If Ω has a discrete spectrum

$$(\Delta\Omega)^2=\sum_i P(\omega_i)(\omega_i-\langle\Omega\rangle)^2 \tag{4.2.8}$$

and if it has a continuous spectrum,

$$(\Delta\Omega)^2=\int P(\omega)(\omega-\langle\Omega\rangle)^2\,d\omega \tag{4.2.9}$$

Notice that $\Delta\Omega$, just like $\langle\Omega\rangle$, is also calculable given just the state and the operator, for Eq. (4.2.7) means just

$$\Delta\Omega=[\langle\psi|(\Omega-\langle\Omega\rangle)^2|\psi\rangle]^{1/2} \tag{4.2.10}$$

Usually the expectation value and the uncertainty provide us with a fairly good description of the state. For example, if we are given that a particle has $\langle X\rangle=a$ and $\Delta X=\Delta$, we know that the particle is likely to be spotted near $x=a$, with deviations of order Δ.

So far, we have concentrated on the measurement of a single variable at a time. We now turn our attention to the measurement of more than one variable at a time. (Since no two independent measurements can really be performed at the same time, we really mean the measurement of two or more dynamical variables in rapid succession.)

Exercise 4.2.1 (Very Important). Consider the following operators on a Hilbert space $\mathbb{V}^3(C)$:

$$L_x=\frac{1}{2^{1/2}}\begin{bmatrix}0&1&0\\1&0&1\\0&1&0\end{bmatrix},\qquad L_y=\frac{1}{2^{1/2}}\begin{bmatrix}0&-i&0\\i&0&-i\\0&i&0\end{bmatrix},\qquad L_z=\begin{bmatrix}1&0&0\\0&0&0\\0&0&-1\end{bmatrix}$$

(1) What are the possible values one can obtain if L_z is measured?

(2) Take the state in which $L_z=1$. In this state what are $\langle L_x\rangle$, $\langle L_x^2\rangle$, and ΔL_x?

(3) Find the normalized eigenstates and the eigenvalues of L_x in the L_z basis.

(4) If the particle is in the state with $L_z=-1$, and L_x is measured, what are the possible outcomes and their probabilities?

(5) Consider the state

$$|\psi\rangle=\begin{bmatrix}1/2\\1/2\\1/2^{1/2}\end{bmatrix}$$

in the L_z basis. If L_z^2 is measured in this state and a result $+1$ is obtained, what is the state after the measurement? How probable was this result? If L_z is measured immediately afterwards, what are the outcomes and respective probabilities?

(6) A particle is in a state for which the probabilities are $P(L_z=1)=1/4$, $P(L_z=0)=1/2$, and $P(L_z=-1)=1/4$. Convince yourself that the most general, normalized state with this property is

$$|\psi\rangle=\frac{e^{i\delta_1}}{2}|L_z=1\rangle+\frac{e^{i\delta_2}}{2^{1/2}}|L_z=0\rangle+\frac{e^{i\delta_3}}{2}|L_z=-1\rangle$$

It was stated earlier on that if $|\psi\rangle$ is a normalized state then the state $e^{i\theta}|\psi\rangle$ is a physically equivalent normalized state. Does this mean that the factors $e^{i\delta_i}$ multiplying the L_z eigenstates are irrelevant? [Calculate for example $P(L_x=0)$.]

Compatible and Incompatible Variables

A striking feature of quantum theory is that given a particle in a state $|\psi\rangle$, one cannot say in general that the particle has a definite value for a given dynamical variable Ω: a measurement can yield any eigenvalue ω for which $\langle\omega|\psi\rangle$ is not zero. The exceptions are the states $|\omega\rangle$. A particle in one of these states can be said, as in classical mechanics, to have a value ω for Ω, since a measurement is assured to give this result. To produce such states we need only take an arbitrary state $|\psi\rangle$ and measure Ω. The measurement process acts as a filter that lets through just one component of $|\psi\rangle$, along some $|\omega\rangle$. The probability that this will happen is $P(\omega)=|\langle\omega|\psi\rangle|^2$.

We now wish to extend these ideas to more than one variable. We consider first the question of two operators. The extension to more than two will be

straightforward. We ask:

Question 1. Is there some multiple filtering process by which we can take an ensemble of particles in some state $|\psi\rangle$ and produce a state with well-defined values ω and λ for two variables Ω and Λ?

Question 2. What is the probability that the filtering will give such a state if we start with the state $|\psi\rangle$?

To answer these questions, let us try to devise a multiple filtering scheme. Let us first measure Ω on the ensemble described by $|\psi\rangle$ and take the particles that yield a result ω. These are in a state that has a well-defined value for Ω. We *immediately* measure Λ and pick those particles that give a result λ. Do we have now an ensemble that is in a state with $\Omega=\omega$ and $\Lambda=\lambda$? Not generally. The reason is clear. After the first measurement, we had the system in the state $|\omega\rangle$, which assured a result ω for Ω, but nothing definite for Λ (since $|\omega\rangle$ need not be an eigenstate of Λ). Upon performing the second measurement, the state was converted to

$$|\psi'\rangle=|\lambda\rangle$$

and we are now assured a result λ for Λ, but nothing definite for Ω (since $|\lambda\rangle$ need not be an eigenstate of Ω).

In other words, the second filtering generally alters the state produced by the first. This change is just the collapse of the state vector $|\omega\rangle=\sum|\lambda\rangle\langle\lambda|\omega\rangle$ into the eigenstate $|\lambda\rangle$.

An exception occurs when the state produced after the first measurement is unaffected by the second. This in turn requires that $|\omega\rangle$ also be an eigenstate of Λ. The answer to the first question above is then in the affirmative only for the simultaneous eigenstates $|\omega\lambda\rangle$. The means for producing them are just as described above. These kets satisfy the equations

$$\Omega|\omega\lambda\rangle=\omega|\omega\lambda\rangle \tag{4.2.11}$$

$$\Lambda|\omega\lambda\rangle=\lambda|\omega\lambda\rangle \tag{4.2.12}$$

The question that arises naturally is: When will two operators admit simultaneous eigenkets? A necessary (but not sufficient) condition is obtained by operating Eq. (4.2.12) with Ω, Eq. (4.2.11) with Λ, and taking the difference:

$$(\Omega\Lambda-\Lambda\Omega)|\omega\lambda\rangle=0 \tag{4.2.13}$$

Thus $[\Omega, \Lambda]$ must have eigenkets with zero eigenvalue if simultaneous eigenkets are to exist. A pair of operators Ω and Λ will fall into one of the three classes:

A. Compatible: $[\Omega, \Lambda]=0$
B. Incompatible: $[\Omega, \Lambda]=$ something that obviously has no zero eigenvalue
C. Others

Class A. If two operators commute, we know a complete basis of simultaneous eigenkets can be found. Each element $|\omega\lambda\rangle$ of this basis has well-defined values for Ω and Λ.

Class B. The most famous example of this class is provided by the position and momentum operators X and P, which obey the *canonical commutation rule*

$$[X, P] = i\hbar \tag{4.2.14}$$

Evidently we cannot ever have $i\hbar|\psi\rangle = 0|\psi\rangle$ for any nontrivial $|\psi\rangle$. *This means there doesn't exist even a single ket for which both X and P are well defined.* Any attempt to filter X is ruined by a subsequent filtering for P and vice vesa. This is the origin of the famous *Heisenberg uncertainty principle*, which will be developed as we go along.

Class C. In this case there are *some* states that are simultaneous eigenkets. There is nothing very interesting we can say about this case except to emphasize that even if two operators don't commute, one can still find a few common eigenkets, though not a full basis. (Why?)

Let us now turn to the second question of the probability of obtaining a state $|\omega\lambda\rangle$ upon measurement of Ω and Λ in a state $|\psi\rangle$. We will consider just case A; the question doesn't arise for case B, and case C is not very interesting. (You should be able to tackle case C yourself after seeing the other two cases.)

Case A. Let us first assume there is no degeneracy. Thus, to a given eigenvalue λ, there is just one ket and this must be a simultaneous eigenket $|\omega\lambda\rangle$. Suppose we measured Ω first. We get ω with a probability $P(\omega) = |\langle\omega\lambda|\psi\rangle|^2$. After the measurement, the particle is in a state $|\omega\lambda\rangle$. The measurement of Λ is certain to yield the result λ. The probability for obtaining ω for Ω and λ for Λ is just the product of the two probabilities

$$P(\omega, \lambda) = |\langle\omega\lambda|\psi\rangle|^2 \cdot 1 = |\langle\omega\lambda|\psi\rangle|^2$$

Notice that if Λ were measured first and Ω next, the probability is the same for getting the results λ and ω. Thus if we expand $|\psi\rangle$ in the complete common eigenbasis as

$$|\psi\rangle = \sum |\omega\lambda\rangle\langle\omega\lambda|\psi\rangle \tag{4.2.15a}$$

then

$$P(\omega, \lambda) = |\langle\omega\lambda|\psi\rangle|^2 = P(\lambda, \omega) \tag{4.2.15b}$$

The reason for calling Ω and Λ compatible if $[\Omega, \Lambda] = 0$ is that the measurement of one variable followed by the other doesn't alter the *eigenvalue* obtained in the first measurement and we have in the end a state with a well-defined value for both observables. Note the emphasis on the invariance of the *eigenvalue* under the second measurement. In the non-degenerate case, this implies the invariance of the state vector as well. In the degenerate case, the state vector can change due to the second

measurement, though the eigenvalue will not, as the following example will show. Consider two operators Λ and Ω on $\mathbb{V}^3(R)$. Let $|\omega_3\lambda_3\rangle$ be one common eigenvector. Let $\lambda_1=\lambda_2=\lambda$. Let $\omega_1\neq\omega_2$ be the eigenvalues of Ω in this degenerate space. Let us use as a basis $|\omega_1\lambda\rangle$, $|\omega_2\lambda\rangle$, and $|\omega_3\lambda_3\rangle$. Consider a normalized state

$$|\psi\rangle=\alpha|\omega_3\lambda_3\rangle+\beta|\omega_1\lambda\rangle+\gamma|\omega_2\lambda\rangle \tag{4.2.16}$$

Let us say we measure Ω first and get ω_3. The state becomes $|\omega_3\lambda_3\rangle$ and the subsequent measurement of Λ is assured to give a value λ_3 and to leave the state alone. Thus $P(\omega_3,\lambda_3)=|\langle\omega_3\lambda_3|\psi\rangle|^2=\alpha^2$. Evidently $P(\omega_3,\lambda_3)=P(\lambda_3,\omega_3)$.

Suppose that the measurement of Ω gave a value ω_1. The resulting state is $|\omega_1\lambda\rangle$ and the probability for this outcome is $|\langle\omega_1\lambda|\psi\rangle|^2$. The subsequent measurement of Λ will leave the state alone and yield the result λ with unit probability. Thus $P(\omega_1,\lambda)$ is the product of the probabilities:

$$P(\omega_1,\lambda)=|\langle\omega_1\lambda|\psi\rangle|^2\cdot 1=|\langle\omega_1\lambda|\psi\rangle|^2=\beta^2 \tag{4.2.17}$$

Let us now imagine the measurements carried out in reverse order. Let the result of the measurement be λ. The state $|\psi'\rangle$ after measurement is the projection of $|\psi\rangle$ in the degenerate λ eigenspace:

$$|\psi'\rangle=\frac{\mathbb{P}_\lambda|\psi\rangle}{|\langle\mathbb{P}_\lambda\psi|\mathbb{P}_\lambda\psi\rangle|^{1/2}}=\frac{\beta|\omega_1\lambda\rangle+\gamma|\omega_2\lambda\rangle}{(\beta^2+\gamma^2)^{1/2}} \tag{4.2.18}$$

where, in the expression above, the projected state has been normalized. The probability for this outcome is $P(\lambda)=\beta^2+\gamma^2$, the square of the projection of $|\psi\rangle$ in the eigenspace. If Ω is measured now, both results ω_1 and ω_2 are possible. The probability for obtaining ω_1 is $|\langle\omega_1\lambda|\psi'\rangle|^2=\beta^2/(\beta^2+\gamma^2)$. Thus, the probability for the result $\Lambda=\lambda$, $\Omega=\omega_1$, is the product of the probabilities:

$$P(\lambda,\omega_1)=(\beta^2+\gamma^2)\cdot\frac{\beta^2}{\beta^2+\gamma^2}=\beta^2=P(\omega_1,\lambda) \tag{4.2.19}$$

Thus $P(\omega_1,\lambda)=P(\lambda,\omega_1)$ independent of the degeneracy. *But this time the state suffered a change due to the second measurement* (unless by accident $|\psi'\rangle$ has no component along $|\omega_2\lambda\rangle$). Thus compatibility generally implies the invariance under the second measurement of the *eigenvalue* measured in the first. Therefore, the state can only be said to remain in the same eigenspace after the second measurement. If the first eigenvalue is non-degenerate, the eigenspace is one dimensional and the state vector itself remains invariant.

In our earlier discussion on how to produce well-defined states $|\psi\rangle$ for testing quantum theory, it was observed that the measurement process could itself be used as a preparation mechanism: if the measurement of Ω on an arbitrary, unknown initial state given a result ω, we are sure we have the state $|\psi\rangle=|\omega\rangle$. But this presumes ω is not a degenerate eigenvalue. If it is degenerate, we cannot nail down the state, except to within an eigenspace. It was therefore suggested that we stick to variables with a nondegenerate spectrum. We can now lift that restriction. Let us

say a degenerate eigenvalue ω for the variable Ω was obtained. We have then some vector in the ω eigenspace. We now measure another compatible variable Λ. If we get a result λ, we have a definite state $|\omega\lambda\rangle$, unless the value (ω, λ) itself is degenerate. We must then measure a third variable Γ compatible with Ω and Λ and so on. Ultimately we will get a state that is unique, given all the simultaneous eigenvalues: $|\omega, \lambda, \gamma, \ldots\rangle$. It is presumed that such a set of compatible observables, called a *complete set of commuting observables*, exists. To prepare a state for studying quantum theory then, we take an arbitrary initial state and filter it by a sequence of compatible measurements till it is down to a unique, known vector. Any nondegenerate operator, all by itself, is a "complete set."

Incidentally, even if the operators Ω and Λ are incompatible, we can specify the probability $P(\omega, \lambda)$ that the measurement of Ω followed by that of Λ on a state $|\psi\rangle$ will give the results ω and λ, respectively. However, the following should be noted:

(1) $P(\omega, \lambda) \neq P(\lambda, \omega)$ in general.

(2) The probability $P(\omega, \lambda)$ is not the probability for producing a final state that has well-defined values ω and λ for Ω and Λ. (Such a state doesn't exist by the definition of incompatibility.) The state produced by the two measurements is just the eigenstate of the second operator with the measured eigenvalue.

The Density Matrix—a Digression‡

So far we have considered ensembles of N systems all in the same state $|\psi\rangle$. They are hard to come by in practice. More common are ensembles of N systems, n_i $(i=1, 2, \ldots, k)$ of which are in the state $|i\rangle$. (We restrict ourselves to the case where $|i\rangle$ is an element of an orthonormal basis.) Thus the ensemble is described by k kets $|1\rangle, |2\rangle, \ldots, |k\rangle$, and k *occupancy numbers* $n_1, \ldots, n_k$. A convenient way to assemble all this information is in the form of the *density matrix* (which is really an operator that becomes a matrix in some basis):

$$\rho = \sum_i p_i |i\rangle\langle i| \tag{4.2.20}$$

where $p_i = n_i/N$ is the probability that a system picked randomly out of the ensemble is in the state $|i\rangle$. The ensembles we have dealt with so far are said to be *pure*; they correspond to all $p_i = 0$ except one. A general ensemble is *mixed.*

Consider now the ensemble average of Ω. It is

$$\langle\bar{\Omega}\rangle = \sum_i p_i \langle i|\Omega|i\rangle \tag{4.2.21}$$

The bar on $\langle\bar{\Omega}\rangle$ reminds us that two kinds of averaging have been carried out: a quantum average $\langle i|\Omega|i\rangle$ for each system in $|i\rangle$ and a classical average over the

‡ This digression may be omitted or postponed without loss of continuity.

systems in different states $|i\rangle$. Observe that

$$\begin{aligned}\mathrm{Tr}(\Omega\rho) &= \sum_j \langle j|\Omega\rho|j\rangle \\ &= \sum_j \sum_i \langle j|\Omega|i\rangle\langle i|j\rangle p_i = \sum_i \sum_j \langle i|j\rangle\langle j|\Omega|i\rangle p_i \\ &= \sum_i \langle i|\Omega|i\rangle p_i \\ &= \langle\bar{\Omega}\rangle \end{aligned} \tag{4.2.22}$$

The density matrix contains all the statistical information about the ensemble. Suppose we want, not $\langle\bar{\Omega}\rangle$, but instead $\overline{P(\omega)}$, the probability of obtaining a particular value ω. We first note that, for a pure ensemble,

$$P(\omega) = |\langle\omega|\psi\rangle|^2 = \langle\psi|\omega\rangle\langle\omega|\psi\rangle = \langle\psi|\mathbb{P}_\omega|\psi\rangle = \langle\mathbb{P}_\omega\rangle$$

which combined with Eq. (4.2.22) tells us that

$$\overline{P(\omega)} = \mathrm{Tr}(\mathbb{P}_\omega\rho)$$

The following results may be easily established:

(1)	$\rho^\dagger = \rho$	
(2)	$\mathrm{Tr}\,\rho = 1$	
(3)	$\rho^2 = \rho$	for a pure ensemble
(4)	$\rho = (1/k)I$	for an ensemble uniformly distributed over k states
(5)	$\mathrm{Tr}\,\rho^2 \le 1$	(equality holds for a pure ensemble)

(4.2.23)

You are urged to convince yourself of these relations.

Example 4.2.4. To gain more familiarity with quantum theory let us consider an infinite-dimensional ket $|\psi\rangle$ expanded in the basis $|x\rangle$ of the position operator X:

$$|\psi\rangle = \int_{-\infty}^{\infty} |x\rangle\langle x|\psi\rangle\, dx = \int_{-\infty}^{\infty} |x\rangle\psi(x)\, dx$$

We call $\psi(x)$ the wave function (in the X basis). Let us assume $\psi(x)$ is a Gaussian, that is, $\psi(x) = A\exp[-(x-a)^2/2\Delta^2]$ (Fig. 4.2a). We now try to extract information about this state by using the postulates. Let us begin by normalizing the state:

$$\begin{aligned} 1 = \langle\psi|\psi\rangle &= \int_{-\infty}^{\infty} \langle\psi|x\rangle\langle x|\psi\rangle\, dx = \int_{-\infty}^{\infty} |\psi(x)|^2\, dx \\ &= \int_{-\infty}^{\infty} A^2 e^{-(x-a)^2/\Delta^2}\, dx = A^2(\pi\Delta^2)^{1/2} \quad \text{(see Appendix A.2)} \end{aligned}$$

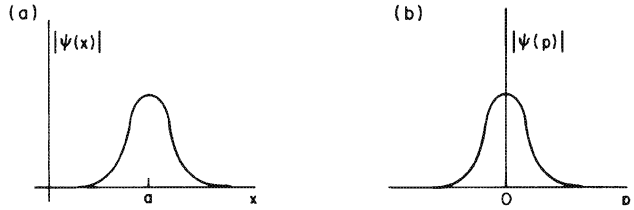

Figure 4.2. (a) The modulus of the wave function, $|\langle x|\psi\rangle| = |\psi(x)|$. (b) The modulus of the wave function, $|\langle p|\psi\rangle| = |\psi(p)|$.

So the normalized state is

$$\psi(x) = \frac{1}{(\pi\Delta^2)^{1/4}} e^{-(x-a)^2/2\Delta^2}$$

The probability for finding the particle between x and $x+dx$ is

$$P(x)\,dx = |\psi(x)|^2\,dx = \frac{1}{(\pi\Delta^2)^{1/2}} e^{-(x-a)^2/\Delta^2}\,dx$$

which looks very much like Fig. 4.2a. Thus the particle is most likely to be found around $x=a$, and chances of finding it away from this point drop rapidly beyond a distance Δ. We can quantify these statements by calculating the expectation value and uncertainty for X. Let us do so.

Now, the operator X defined in postulate II is the same one we discussed at length in Section 1.10. Its action in the X basis is simply to multiply by x, i.e., if

$$\langle x|\,\psi\rangle = \psi(x)$$

then,

$$\langle x|X|\psi\rangle = \int_{-\infty}^{\infty} \langle x|X|x'\rangle\langle x'|\psi\rangle\,dx' = \int_{-\infty}^{\infty} x\delta(x-x')\psi(x')\,dx'$$
$$= x\psi(x)$$

Using this result, the mean or expectation value of X is

$$\langle X\rangle = \langle\psi|X|\psi\rangle = \int_{-\infty}^{\infty} \langle\psi|x\rangle\langle x|X|\psi\rangle\,dx$$
$$= \int_{-\infty}^{\infty} \psi^*(x)x\psi(x)\,dx$$
$$= \frac{1}{(\pi\Delta^2)^{1/2}} \int_{-\infty}^{\infty} e^{-(x-a)^2/\Delta^2} x\,dx$$

If we define $y = x - a$,

$$\langle X \rangle = \frac{1}{(\pi\Delta^2)^{1/2}} \int_{-\infty}^{\infty} (y + a)\, e^{-y^2/\Delta^2}\, dy$$
$$= a$$

We should have anticipated this result of course, since the probability density is symmetrically distributed around $x = a$.

Next, we calculate the fluctuations around $\langle X \rangle = a$, i.e., the uncertainty

$$\begin{aligned}
\Delta X &= [\langle \psi | (X - \langle X \rangle)^2 | \psi \rangle]^{1/2} \\
&= [\langle \psi | X^2 - 2X\langle X \rangle + \langle X \rangle^2 | \psi \rangle]^{1/2} \\
&= [\langle \psi | X^2 - \langle X \rangle^2 | \psi \rangle]^{1/2} \qquad (\text{since } \langle \psi | X | \psi \rangle = \langle X \rangle) \\
&= [\langle X^2 \rangle - \langle X \rangle^2]^{1/2} \\
&= [\langle X^2 \rangle - a^2]^{1/2}
\end{aligned}$$

Now

$$\langle X^2 \rangle = \frac{1}{(\pi\Delta^2)^{1/2}} \int_{-\infty}^{\infty} e^{-(x-a)^2/2\Delta^2} \cdot x^2 \cdot e^{-(x-a)^2/2\Delta^2}\, dx$$
$$= \frac{1}{(\pi\Delta^2)^{1/2}} \int_{-\infty}^{\infty} e^{-y^2/\Delta^2} (y^2 + 2ya + a^2)\, dy = \frac{\Delta^2}{2} + 0 + a^2$$

So

$$\Delta X = \frac{\Delta}{2^{1/2}}$$

So much for the information on the variable X. Suppose we next want to know the probability distribution for different values of another dynamical variable, say the momentum P.

(1) First we must construct the operator P in this basis.
(2) Then we must find its eigenvalues p, and eigenvectors $|p\rangle$.
(3) Finally, we must take the inner product $\langle p | \psi \rangle$.
(4) If p is discrete, $|\langle p_i | \psi \rangle|^2 = P(p_i)$, and if p is continuous, $|\langle p | \psi \rangle|^2 = P(p)$, the probability density.

Now, the P operator is just the K operator discussed in Section 1.10 multiplied by $\hbar$ and has the action of $-i\hbar\, d/dx$ in the X basis, for if

$$\langle x | \psi \rangle = \psi(x)$$

then

$$\langle x|P|\psi\rangle = \int_{-\infty}^{\infty} \langle x|P|x'\rangle\langle x'|\psi\rangle\, dx'$$

$$= \int_{-\infty}^{\infty} [-i\hbar\delta'(x-x')]\psi(x')\, dx' \qquad \text{(Postulate II)}$$

$$= -i\hbar \frac{d\psi}{dx}$$

Thus, if we project the eigenvalue equation

$$P|p\rangle = p|p\rangle$$

onto the X basis, we get

$$\langle x|P|p\rangle = p\langle x|p\rangle$$

or

$$-i\hbar \frac{d\psi_p(x)}{dx} = p\psi_p(x)$$

where $\psi_p(x) = \langle x|p\rangle$. The solutions, normalized to the Dirac delta function‡ are (from Section 1.10)

$$\psi_p(x) = \frac{1}{(2\pi\hbar)^{1/2}} e^{ipx/\hbar}$$

Now we can compute

$$\langle p|\psi\rangle = \int \langle p|x\rangle\langle x|\psi\rangle\, dx = \int \psi_p^*(x)\psi(x)\, dx$$

$$= \int_{-\infty}^{\infty} \frac{e^{-ipx/\hbar}}{(2\pi\hbar)^{1/2}} \frac{e^{-(x-a)^2/2\Delta^2}}{(\pi\Delta^2)^{1/4}} dx = \left(\frac{\Delta^2}{\pi\hbar^2}\right)^{1/4} e^{-ipa/\hbar}\, e^{-p^2\Delta^2/2\hbar^2}$$

The modulus of $\psi(p)$ is a Gaussian (Fig. 4.2b) of width $\hbar/2^{1/2}\Delta$. It follows that $\langle P\rangle = 0$, and $\Delta P = \hbar/2^{1/2}\Delta$. Since $\Delta X = \Delta/2^{1/2}$; we get the relation

$$\Delta X \cdot \Delta P = \hbar/2$$

‡ Here we want $\langle p|p'\rangle = \delta(p-p') = \delta(k-k')/\hbar$, where $p = \hbar k$. This explains the $(2\pi\hbar)^{-1/2}$ normalization factor.

The Gaussian happens to saturate the lower bound of the uncertainty relation (to be formally derived in chapter 9):

$$\Delta X \cdot \Delta P \geq \hbar/2$$

The uncertainty relation is a consequence of the general fact that anything narrow in one space is wide in the transform space and vice versa. So if you are a 110-lb weakling and are taunted by a 600-lb bully, just ask him to step into momentum space! □

This is a good place to point out that the plane waves $e^{ipx/\hbar}$ (and *all improper vectors*, i.e., vectors that can't be normalized to unity but only to the Dirac delta function) are introduced into the formalism as purely mathematical entities. Our inability to normalize them to unity translates into our inability to associate with them a sensible absolute probability distribution, so essential to the physical interpretation of the wave function. In the present case we have a particle whose relative probability density is uniform in all of space. Thus the absolute probability of finding it in any finite volume, even as big as our solar system, is zero. Since any particle that we are likely to be interested in will definitely be known to exist in some finite volume of such large dimensions, it is clear that no physically interesting state will be given by a plane wave. But, since the plane waves are eigenfunctions of P, does it mean that states of well-defined momentum do not exist? Yes, in the strict sense. However, there do exist states that are both normalizable to unity (i.e., correspond to *proper* vectors) and come arbitrarily close to having a precise momentum. For example, a wave function that behaves as $e^{ip_0x/\hbar}$ over a large region of space and tapers off to zero beyond, will be normalizable to unity and will have a Fourier transform so sharply peaked at $p=p_0$ that momentum measurements will only give results practically indistinguishable from p_0. Thus there is no conflict between the fact that plane waves are unphysical, while states of well-defined momentum exist, for "well defined" never means "mathematically exact," but only "exact to any measurable accuracy." Thus a particle coming out of some accelerator with some advertised momentum, say 500 GeV/c, is in a proper normalizable state (since it is known to be located in our laboratory) and not in a plane wave state corresponding to $|p=500\ \text{GeV}/c\rangle$.

But despite all this, we will continue to use the eigenkets $|p\rangle$ as basis vectors and to speak of a particle being in the state $|p\rangle$, because these vectors are so much more convenient to handle mathematically than the proper vectors. It should, however, be borne in mind that when we say a particle is (coming out of the accelerator) in a state $|p_0\rangle$, it is really in a proper state with a momentum space wave function so sharply peaked at $p=p_0$ that it may be replaced by a delta function $\delta(p-p_0)$.

The other set of improper kets we will use in the same spirit are the position eigenkets $|x\rangle$, which also form a convenient basis. Again, when we speak of a particle being in a state $|x_0\rangle$ we shall mean that its wave function is so sharply peaked at $x=x_0$ that it may be treated as a delta function to a good accuracy.‡

‡ Thus, by the physical Hilbert space, we mean the space of interest to physicists, not one whose elements all correspond to physically realizable states.

Occasionally, the replacement of a proper wave function by its improper counterpart turns out to be a poor approximation. Here is an example from Chapter 19: Consider the probability that a particle coming out of an accelerator with a nearly exact momentum scatters off a target and enters a detector placed far away, and not in the initial direction. Intuition says that the answer must be zero if the target is absent. This reasonable condition is violated if we approximate the initial state of the particle by a plane wave (which is nonzero everywhere). So we proceed as follows. In the vicinity of the target, we use the plane wave to approximate the initial wave function, for the two are indistinguishable over the (finite and small) range of influence of the target. At the detector, however, we go back to the proper wave (which has tapered off) to represent the initial state.

*Exercise 4.2.2.** Show that for a real wave function $\psi(x)$, the expectation value of momentum $\langle P\rangle = 0$. (Hint: Show that the probabilities for the momenta $\pm p$ are equal.) Generalize this result to the case $\psi = c\psi_r$, where ψ_r is real and c an arbitrary (real or complex) constant. (Recall that $|\psi\rangle$ and $\alpha|\psi\rangle$ are physically equivalent.)

*Exercise 4.2.3.** Show that if $\psi(x)$ has mean momentum $\langle P\rangle$, $e^{ip_0x/\hbar}\psi(x)$ has mean momentum $\langle P\rangle + p_0$.

Example 4.2.5. The collapse of the state vector and the uncertainty principle play a vital role in explaining the following extension of the double slit experiment. Suppose I say, "I don't believe that a given particle (let us say an electron) doesn't really go through one slit or the other. So I will set up a light source in between the slits to the right of the screen. Each passing electron will be exposed by the beam and I note which slit it comes out of. Then I note where it arrives on the screen. I make a table of how many electrons arrive at each x and which slit they came from. Now there is no escape from the conclusion that the number arriving at a given x is the sum of the numbers arriving via S_1 and S_2. So much for quantum theory and its interference pattern!"

But the point of course is that quantum theory no longer predicts an interference pattern! The theory says that if an electron of definite momentum p is involved, the corresponding wave function is a wave with a well-defined wave number $k = p/\hbar$, which interferes with itself and produces a nice interference pattern. This prediction is valid only as long as the state of the electron is what we say it is. But this state is necessarily altered by the light source, which upon measuring the position of the electron (as being next to S_1, say) changes its wave function from something that was extended in space to something localized near S_1. Once the state is changed, the old prediction of interference is no longer valid.

Now, once in a while some electrons will get to the detectors without being detected by the light source. We note where these arrive, but cannot classify them as coming via S_1 or S_2. When the distribution of just these electrons is plotted; sure enough we get the interference pattern. We had better, for quantum theory predicts it, the state not having been tampered with in these cases.

The above experiment can also be used to demystify to some extent the collapse of the wave function under measurement. Why is it that even the ideal measurement produces unavoidable changes in the state? The answer, as we shall see, has to do with the fact that $\hbar$ is not zero.

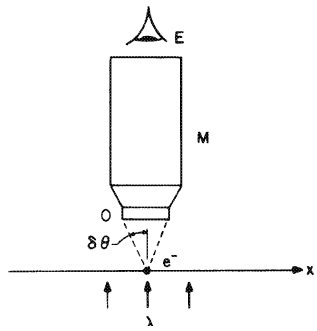

Figure 4.3. Light of frequency λ bounces off the electron, enters the objective O of the microscope, and enters the eye E of the observer.

Consider the schematic set up in Fig. 4.3. Light of wavelength λ illuminates an electron (e^-), enters the objective (O) of a microscope (M) and reaches our eye (E). If $\delta\theta$ is the opening angle of the cone of light entering the objective after interacting with the electron, classical optics limits the accuracy of the position measurement by an uncertainty

$$\Delta X \cong \lambda / \sin \delta\theta$$

Both classically and quantum mechanically, we can reduce ΔX to 0 by reducing λ to zero.‡ In the latter description however, the improved accuracy in the position measurement is at the expense of producing an increased uncertainty in the x component (p_x) of the electron momentum. The reason is that light of wavelength λ is not a continuous wave whose impact on the electron momentum may be arbitrarily reduced by a reduction of its amplitude, but rather a flux of photons of momentum $p = 2\pi\hbar/\lambda$. As λ decreases, the collisions between the electron and the photons become increasingly violent. This in itself would not lead to an uncertainty in the electron momentum, were it not for the fact that the x component of the photons entering the objective can range from 0 to $p \sin \delta\theta = 2\pi\hbar \sin \delta\theta/\lambda$. Since at least one photon must reach our eyes after bouncing off the electron for us to see it, there is a minimum uncertainty in the recoil momentum of the electron given by

$$\Delta P_x \simeq \frac{2\pi\hbar}{\lambda} \sin \delta\theta$$

Consequently, we have at the end of our measurement an electron whose position and momenta are uncertain by ΔX and ΔP_x such that

$$\Delta X \cdot \Delta P_x \simeq 2\pi\hbar \simeq \hbar$$

[The symbols ΔX and ΔP_x are not precisely the quantities defined in Eq. (4.2.7) but are of the same order of magnitude.] This is the famous *uncertainty principle*. There is no way around it. If we soften the blow of each photon by increasing λ or narrowing the objective to better constrain the final photon momentum, we lose in resolution.

‡ This would be the ideal position measurement.

More elaborate schemes, which determine the recoil of the microscope, are equally futile. Note that if $\hbar$ were 0, we could have ΔX and ΔP_x simultaneously 0. Physically, it means that we can increase our position resolution without increasing the punch carried by the photons. Of course $\hbar$ is not zero and we can't make it zero in any experiment. But what we can do is to use bigger and bigger objects for our experiment so that in the scale of these objects $\hbar$ appears to be negligible. We then regain classical mechanics. The position of a billiard ball can be determined very well by shining light on it, but this light hardly affects its momentum. This is why one imagines in classical mechanics that momentum and position can be well defined simultaneously. □

Generalization to More Degrees of Freedom

Our discussion so far has been restricted to a system with one degree of freedom—namely, a single particle in one dimension. We now extend our domain to a system with N degrees of freedom. The only modification is in postulate II, which now reads as follows.

Postulate II. Corresponding to the N Cartesian coordinates $x_1, \ldots, x_N$ describing the classical system, there exist in quantum theory N mutually commuting operators $X_1, \ldots, X_N$. In the simultaneous eigenbasis $|x_1, x_2, \ldots, x_N\rangle$ of these operators, called the *coordinate basis* and normalized as

$$\langle x_1, x_2, \ldots, x_N | x_1', x_2', \ldots, x_N' \rangle = \delta(x_1 - x_1') \ldots \delta(x_N - x_N')$$

(the product of delta functions vanishes unless all the arguments vanish) we have the following correspondence:

$$|\psi\rangle \rightarrow \langle x_1, \ldots, x_N | \psi \rangle = \psi(x_1, \ldots, x_N)$$

$$X_i|\psi\rangle \rightarrow \langle x_1, \ldots, x_N | X_i | \psi \rangle = x_i \psi(x_1, \ldots, x_N)$$

$$P_i|\psi\rangle \rightarrow \langle x_1, \ldots, x_N | P_i | \psi \rangle = -i\hbar \frac{\partial}{\partial x_i} \psi(x_1, \ldots, x_N)$$

P_i being the momentum operator corresponding to the classical momentum p_i. Dependent dynamical variables $\omega(x_i, p_j)$ are represented by operators $\Omega = \omega(x_i \rightarrow X_i, p_j \rightarrow P_j)$.

The other postulates remain the same. For example $|\psi(x_1, \ldots, x_N)|^2 \times dx_1 \ldots dx_N$ is the probability that the particle coordinates lie between $x_1, x_2, \ldots, x_N$ and $x_1 + dx_1, x_2 + dx_2, \ldots, x_N + dx_N$.

This postulate is stated in terms of Cartesian coordinates since only in terms of these can one express the operator assignments in the simple form $X_i \rightarrow x_i$, $P_i \rightarrow -i\hbar\, \partial/\partial x_i$. Once the substitutions have been made and the desired equations obtained in the coordinate basis, one can perform any desired change of variable before solving them. Suppose, for example, that we want to find the eigenvalues and

eigenvectors of the operator Ω, corresponding to the classical variable

$$\omega = \frac{p_1^2 + p_2^2 + p_3^2}{2m} + x_1^2 + x_2^2 + x_3^2 \tag{4.2.24}$$

where x_1, x_2, and x_3 are the three Cartesian coordinates and p_i the corresponding momenta of a particle of mass m in three dimensions. Since the coordinates are usually called x, y, and z, let us follow this popular notation and rewrite Eq. (4.2.24) as

$$\omega = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + x^2 + y^2 + z^2 \tag{4.2.25}$$

To solve the equation

$$\Omega|\omega\rangle = \omega|\omega\rangle$$

with

$$\Omega = \frac{P_x^2 + P_y^2 + P_z^2}{2m} + X^2 + Y^2 + Z^2$$

we make the substitution

$$|\omega\rangle \rightarrow \psi_\omega(x, y, z)$$

$$X \rightarrow x, \qquad P_x \rightarrow -i\hbar \frac{\partial}{\partial x}$$

etc. and get

$$\left[\frac{-\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + x^2 + y^2 + z^2\right]\psi_\omega(x, y, z) = \omega\psi_\omega(x, y, z) \tag{4.2.26}$$

Once we have obtained this differential equation, we can switch to any other set of coordinates. In the present case the spherical coordinates r, θ, and ϕ recommend themselves. Since

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

$$\equiv \nabla^2 \equiv \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]$$

Eq. (4.2.26) becomes

$$\frac{-\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi_\omega}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi_\omega}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi_\omega}{\partial\phi^2}\right]$$
$$+r^2\psi_\omega=\omega\psi_\omega \tag{4.2.27}$$

What if we wanted to go directly from ω in spherical coordinates

$$\omega=\frac{1}{2m}\left(p_r^2+\frac{p_\theta^2}{r^2}+\frac{p_\phi^2}{r^2\sin^2\theta}\right)+r^2$$

to Eq. (4.2.27)? It is clear upon inspection that there exists no simple rule [such as $p_r\rightarrow(-i\hbar\,\partial/\partial r)$] for replacing the classical momenta by differential operators in r, θ, and ϕ which generates Eq. (4.2.27) starting from the ω above. There does exist a complicated procedure for quantizing in non-Cartesian coordinates, but we will not discuss it, since the recipe eventually reproduces what the Cartesian recipe (which seems to work‡) yields so readily.

There are further generalizations, namely, to relativistic quantum mechanics and to quantum mechanics of systems in which particles are created and destroyed (so that the number of degrees of freedom changes!). Except for a brief discussion of these toward the end of the program, we will not address these matters.

4.3. The Schrödinger Equation (Dotting Your *i*'s and Crossing Your $\hbar$'s)

Having discussed in some detail the state at a given time, we now turn our attention to postulate IV, which specifies the change of this state with time. According to this postulate, the state obeys the Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle=H|\psi(t)\rangle \tag{4.3.1}$$

Our discussion of this equation is divided into three sections:

(1) Setting up the equation
(2) General approach to its solution
(3) Choosing a basis for solving the equation

Setting Up the Schrödinger Equation

To set up the Schrödinger equation one must simply make the substitution $\mathscr{H}(x\rightarrow X, p\rightarrow P)$, where $\mathscr{H}$ is the classical Hamiltonian for the same problem. Thus,

‡ In the sense that in cases where comparison with experiment is possible, as in say the hydrogen spectrum, there is agreement.

if we are describing a harmonic oscillator, which is classically described by the Hamiltonian

$$\mathscr{H} = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2 \tag{4.3.2}$$

the Hamiltonian operator in quantum mechanics is

$$H = \frac{P^2}{2m} + \frac{1}{2} m\omega^2 X^2 \tag{4.3.3}$$

In three dimensions, the Hamiltonian operator for the quantum oscillator is likewise

$$H = \frac{P_x^2 + P_y^2 + P_z^2}{2m} + \frac{1}{2} m\omega^2 (X^2 + Y^2 + Z^2) \tag{4.3.4}$$

assuming the force constant is the same in all directions.

If the particle in one dimension is subject to a constant force f, then

$$\mathscr{H} = \frac{p^2}{2m} - fx$$

and

$$H = \frac{P^2}{2m} - fX \tag{4.3.5}$$

For a particle of charge q in an electromagnetic field in three dimensions,

$$\mathscr{H} = \frac{|\mathbf{p} - (q/c)\mathbf{A}(\mathbf{r}, t)|^2}{2m} + q\phi(\mathbf{r}, t) \tag{4.3.6}$$

In constructing the corresponding quantum Hamiltonian operator, we must use the symmetrized form

$$H = \frac{1}{2m}\left(\mathbf{P}\cdot\mathbf{P} - \frac{q}{c}\mathbf{P}\cdot\mathbf{A} - \frac{q}{c}\mathbf{A}\cdot\mathbf{P} + \frac{q^2}{c^2}\mathbf{A}\cdot\mathbf{A}\right) + q\phi \tag{4.3.7}$$

since $\mathbf{P}$ does not commute with $\mathbf{A}$, which is a function of X, Y, and Z.

In this manner one can construct the Hamiltonian H for any problem with a classical counterpart. Problems involving spin have no classical counterparts and some improvisation is called for. We will discuss this question when we study spin in some detail in Chapter 14.

Let us first assume that H has no explicit t dependence. In this case the equation

$$i\hbar|\dot{\psi}\rangle = H|\psi\rangle$$

is analogous to equations discussed in Chapter 1

$$|\ddot{x}\rangle = \Omega|x\rangle$$

and

$$|\ddot{\psi}\rangle = -K^2|\psi\rangle$$

describing the coupled masses and the vibrating string, respectively. Our approach will once again be to find the eigenvectors and eigenvalues of H and to construct the propagator $U(t)$ in terms of these. Once we have $U(t)$, we can write

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle$$

There is no need to make assumptions about $|\dot{\psi}(0)\rangle$ here, since it is determined by Eq. (4.3.1):

$$|\dot{\psi}(0)\rangle = \frac{-i}{\hbar} H|\psi(0)\rangle$$

In other words, Schrödinger's equation is first order in time, and the specification of $|\psi\rangle$ at $t=0$ is a sufficient initial-value datum.

Let us now construct an explicit expression for $U(t)$ in terms of $|E\rangle$, the normalized eigenkets of H with eigenvalues E which obey

$$H|E\rangle = E|E\rangle \tag{4.3.8}$$

This is called the *time-independent Schrödinger equation.* Assume that we have solved it and found the kets $|E\rangle$. If we expand $|\psi\rangle$ as

$$|\psi(t)\rangle = \sum |E\rangle\langle E|\psi(t)\rangle \equiv \sum a_E(t)|E\rangle \tag{4.3.9}$$

the equation for $a_E(t)$ follows if we act on both sides with $(i\hbar\,\partial/\partial t - H)$:

$$0 = (i\hbar\,\partial/\partial t - H)|\psi(t)\rangle = \sum (i\hbar\dot{a}_E - Ea_E)|E\rangle \Rightarrow i\hbar\dot{a}_E = Ea_E \tag{4.3.10}$$

where we have used the linear independence of the kets $|E\rangle$. The solution to Eq. (4.3.10) is

$$a_E(t) = a_E(0)\, e^{-iEt/\hbar} \tag{4.3.11a}$$

or

$$\langle E|\psi(t)\rangle = \langle E|\psi(0)\rangle\, e^{-iEt/\hbar} \tag{4.3.11b}$$

so that

$$|\psi(t)\rangle = \sum_E |E\rangle\langle E|\psi(0)\rangle\, e^{-iEt/\hbar} \tag{4.3.12}$$

We can now extract $U(t)$:

$$U(t) = \sum_E |E\rangle\langle E|\, e^{-iEt/\hbar} \tag{4.3.13}$$

We have been assuming that the energy spectrum is discrete and nondegenerate. If E is degenerate, one must first introduce an extra label α (usually the eigenvalue of a compatible observable) to specify the states. In this case

$$U(t) = \sum_\alpha \sum_E |E, \alpha\rangle\langle E, \alpha|\, e^{-iEt/\hbar}$$

If E is continuous, the sum must be replaced by an integral. The normal modes

$$|E(t)\rangle = |E\rangle\, e^{-iEt/\hbar}$$

are also called *stationary states* for the following reason: the probability distribution $P(\omega)$ for any variable Ω is time-independent in such a state:

$$\begin{aligned} P(\omega, t) &= |\langle\omega|\psi(t)\rangle|^2 \\ &= |\langle\omega|E(t)\rangle|^2 \\ &= |\langle\omega|E\rangle\, e^{-iEt/\hbar}|^2 \\ &= |\langle\omega|E\rangle|^2 \\ &= P(\omega, 0) \end{aligned}$$

There exists another expression for $U(t)$ besides the sum, Eq. (4.3.13), and that is

$$U(t) = e^{-iHt/\hbar} \tag{4.3.14}$$

If this exponential series converges (and it sometimes does not), this form of $U(t)$ can be very useful. (Convince yourself that $|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle$ satisfies Schrödinger's equation.)

Since H (the energy operator) is Hermitian, it follows that $U(t)$ is unitary. We may therefore think of the time evolution of a ket $|\psi(t)\rangle$ as a "rotation" in Hilbert

space. One immediate consequence is that the norm $\langle\psi(t)|\psi(t)\rangle$ is invariant:

$$\langle\psi(t)|\psi(t)\rangle=\langle\psi(0)|U^{\dagger}(t)U(t)|\psi(0)\rangle=\langle\psi(0)|\psi(0)\rangle \tag{4.3.15}$$

so that a state, once normalized, stays normalized. There are other consequences of the fact that the time evolution may be viewed as a rotation. For example, one can abandon the fixed basis we have been using, and adopt one that also rotates at the same rate as the state vectors. In such a basis the vectors would appear frozen, but the operators, which were constant matrices in the fixed basis, would now appear to be time dependent. Any physical entity, such as a matrix element, would, however, come out the same as before since $\langle\phi|\Omega|\psi\rangle$, which is the dot product of $\langle\phi|$ and $|\Omega\psi\rangle$, is invariant under rotations. This view of quantum mechanics is called the *Heisenberg picture*, while the one we have been using is called the *Schrödinger picture*. Infinitely many pictures are possible, each labeled by how the basis is rotating. So if you think you were born too late to make a contribution to quantum theory fear not, for you can invent your own picture. We will take up the study of various pictures in Chapter 18.

Let us now consider the case $H=H(t)$. We no longer look for normal modes, since the operator in question is changing with time. There exists no fixed strategy for solving such problems. In the course of our study we will encounter a time-dependent problem involving spin which can be solved exactly. We will also study a systematic approximation scheme for solving problems with

$$H(t)=H^{0}+H^{1}(t)$$

where H^{0} is a large time-independent piece and $H^{1}(t)$ is a small time-dependent piece.

What is the propagator $U(t)$ in the time-dependent case? In other words, how is $U(t)$ in $|\psi(t)\rangle=U(t)|\psi(0)\rangle$ related to $H(t)$? To find out, we divide the interval $(0-t)$ into N pieces of width $\Delta=t/N$, where N is very large and Δ is very small. By integrating the Schrödinger equation over the first interval, we can write *to first order in* Δ

$$\begin{aligned}|\psi(\Delta)\rangle&=|\psi(0)\rangle+\Delta\left.\frac{d|\psi\rangle}{dt}\right|_{0}\\&=|\psi(0)\rangle-\frac{i\Delta}{\hbar}H(0)|\psi(0)\rangle\\&=\left[1-\frac{i\Delta}{\hbar}H(0)\right]|\psi(0)\rangle\end{aligned}$$

which, to this order

$$=\exp\left[\frac{-i\Delta}{\hbar}H(0)\right]|\psi(0)\rangle$$

[One may wonder whether in the interval $0-\Delta$, one must use $H(0)$ or $H(\Delta)$ or $H(\Delta/2)$ and so on. The difference between these possibilities is of order Δ and hence irrelevant, since there is already one power of Δ in front of H.] Inching forth in steps of Δ, we get

$$|\psi(t)\rangle = \prod_{n=0}^{N-1} e^{-i\Delta H(n\Delta)/\hbar} |\psi(0)\rangle$$

We cannot simply add the exponents to get, in the $N\to\infty$ limit,

$$U(t) = \exp\left[-(i/\hbar)\int_0^t H(t')\,dt'\right]$$

since

$$[H(t_1), H(t_2)] \neq 0$$

in general. For example, if

$$H(t) = X^2 \cos^2 \omega t + P^2 \sin^2 \omega t$$

then

$$H(0) = X^2$$

and

$$H(\pi/2\omega) = P^2$$

and

$$[H(0), H(\pi/2\omega)] \neq 0$$

It is common to use the symbol, called the *time-ordered integral*

$$T\left\{\exp\left[-(i/\hbar)\int_0^t H(t')\,dt'\right]\right\} = \lim_{N\to\infty} \prod_{n=0}^{N-1} \exp[-(i/\hbar)H(n\Delta)\Delta]$$

in such problems. We will not make much use of this form of $U(t)$. But notice that being a product of unitary operators, $U(t)$ is unitary, and time evolution continues to be a "rotation" whether or not H is time independent.

Whether or not H depends on time, the propagator satisfies the following conditions:

$$\begin{aligned} U(t_3, t_2)U(t_2, t_1) &= U(t_3, t_1) \\ U^{\dagger}(t_2, t_1) &= U^{-1}(t_2, t_1) = U(t_1, t_2) \end{aligned} \tag{4.3.16}$$

It is intuitively clear that these equations are correct. You can easily prove them by applying the U's to some arbitrary state and using the fact that U is unitary and $U(t, t) = I$.

Choosing a Basis for Solving Schrödinger's Equation

Barring a few exceptions, the Schrödinger equation is always solved in a particular basis. Although all bases are equal mathematically, some are more equal than others. First of all, since $H = H(X, P)$ the X and P bases recommend themselves, for in going to one of them the corresponding operator is rendered diagonal. Thus one can go to the X basis in which $X \to x$ and $P \to -i\hbar\, d/dx$ or to the P basis in which $P \to p$ and $X \to i\hbar\, d/dp$. The choice between the two depends on the Hamiltonian. Assuming it is of the form (in one dimension)

$$H = T + V = \frac{P^2}{2m} + V(X) \tag{4.3.17}$$

the choice is dictated by $V(X)$. Since $V(X)$ is usually a more complicated function of X than T is of P, one prefers the X basis. Thus if

$$H = \frac{P^2}{2m} + \frac{1}{\cosh^2 X} \tag{4.3.18}$$

the equation

$$H|E\rangle = E|E\rangle$$

becomes in the X basis the second-order equation

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{\cosh^2 x}\right)\psi_E(x) = E\psi_E(x) \tag{4.3.19}$$

which can be solved. Had one gone to the P basis, one would have ended up with the equation

$$\left[\frac{p^2}{2m} + \frac{1}{\cosh^2 (i\hbar\, d/dp)}\right]\psi_E(p) = E\psi_E(p) \tag{4.3.20}$$

which is quite frightening.

A problem where the P basis is preferred is that of a particle in a constant force field f, for which

$$H=\frac{P^2}{2m}-fX \tag{4.3.21}$$

In the P basis one gets a first-order differential equation

$$\left(\frac{p^2}{2m}-i\hbar f\frac{d}{dp}\right)\psi_E(p)=E\psi_E(p) \tag{4.3.22}$$

whereas in the X basis one gets the second-order equation

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}-fx\right)\psi_E(x)=E\psi_E(x) \tag{4.3.23}$$

The harmonic oscillator can be solved with equal ease in either basis since H is quadratic in X and P. It turns out to be preferable to solve it in a third basis in which neither X nor P is diagonal! You must wait till Chapter 7 before you see how this happens.

There exists a built-in bias in favor of the X basis. This has to do with the fact that the x space is the space we live in. In other words, when we speak of the probability of obtaining a value between x and $x+dx$ if the variable X is measured, we mean simply the probability of finding the particle between x and $x+dx$ *in our space*. One may thus visualize $\psi(x)$ as a function in our space, whose modulus squared gives the probability density for finding a particle near x. Such a picture is useful in thinking about the double-slit experiment or the electronic states in a hydrogen atom.

But like all pictures, it has its limits. First of all it must be borne in mind that even though $\psi(x)$ can be visualized as a wave in our space, it is not a real wave, like the electromagnetic wave, which carries energy, momentum, etc. To understand this point, consider a particle in three dimensions. The function $\psi(x, y, z)$ can be visualized as a wave in our space. But, if we consider next a two-particle system, $\psi(x_1, y_1, z_1, x_2, y_2, z_2)$ is a function in a six-dimensional *configuration space* and cannot be visualized in our space.

Thus the case of the single particle is really an exception: there is only one position operator and the space of its eigenvalues *happens to coincide* with the space in which we live and in which the drama of physics takes place.

This brings us to the end of our general discussion of the postulates. We now turn to the application of quantum theory to various physical problems. For pedagogical reasons, we will restrict ourselves to problems of a single particle in one dimension in the next few chapters.