been discussing, really provides the best and most comprehensive picture of the physical world. In any case, it seems safe to say that the problem of "determinism in Nature" is no longer considered to be a settled matter, as it was before the invention of quantum mechanics, and indeed, it will probably remain in an unresolved state for some time to come.

4-5 MOTION OF A PARTICLE IN ONE DIMENSION

The foregoing development of the theory of quantum mechanics has been carried out in terms of general observables associated with a general one-dimensional system. We wish now to apply these results to the specific system of a mass m moving along the x-axis in a potential field $V(x)$—a system which we discussed from the standpoint of classical mechanics in Chapter 3. Experiments tell us that our classical treatment of this system is entirely adequate for a "tangible" particle moving over "visible" distances. However, experiments also tell us that our classical description is *not* universally valid; it fails, for example, to correctly describe the behavior of an electron (mass $\simeq 10^{-27}$ gram) on a scale of the order of an atomic diameter (distance $\simeq 10^{-8}$ centimeter). Now, we can expect that the *quantum* treatment of such a system will be valid in both cases; thus we expect that, on the one hand, the quantum description will reduce to the classical description in the *macroscopic* limit, and on the other hand that it will account for such nonclassical phenomena as quantized observables and the wave-particle duality in the *microscopic* limit.

In Sec. 4-5a we shall define and discuss the relevant observable operators for a mass m moving on the x-axis in a potential field $V(x)$. In Sec. 4-5b we shall indicate how these operators lead to a dualistic "wave-particle" behavior. In Sec. 4-5c we shall discuss the way in which the classical description appears as a limiting case of the quantum description. Finally, in Sec. 4-5d we shall work out a simple "quantum mechanics problem" which is typical of those considered in virtually all texts and courses on elementary quantum mechanics.

4-5a Formation of the Observable Operators The Schrödinger Equations and the Position Probability

In classical mechanics, the system consisting of a particle moving along the x-axis has two basic observables—namely, the "position" x and the "momentum" p. Many other observables can be expressed

as *functions* of position and momentum; for example, the observables "velocity" and "energy" are given respectively by the functions $v = p/m$ and $E = p^2/2m + V(x)$. Now in quantum mechanics the situation is very much the same: position and momentum are still valid observables, as are also most well-behaved functions of position and momentum. Our last postulate stipulates how we are to form the appropriate *operators* to represent these observables.

Postulate 6. For a particle confined to the x-axis, the observables "position" and "momentum" are represented respectively by the operators

$$\hat{X} = x \tag{4-46}$$

and

$$\hat{P} = -i\hbar \frac{d}{dx} \tag{4-47}$$

Moreover, any observable which in classical mechanics is some well-behaved function of position and momentum, $f(x,p)$, is represented in quantum mechanics by the operator $f(\hat{X},\hat{P})$:

$$\mathcal{A} = f(x,p) \quad \text{implies} \quad \hat{A} = f(\hat{X},\hat{P}) = f\left(x, -i\hbar \frac{d}{dx}\right) \tag{4-48}$$

According to Eqs. (4-46) and (4-47), the position operator $\hat{X}$ and the momentum operator $\hat{P}$ are the operators which transform any given function $\phi(x)$ into the respective functions

$$[\hat{X}\phi(x)] \equiv x \cdot \phi(x)$$

and

$$[\hat{P}\phi(x)] \equiv -ih \cdot \frac{d\phi(x)}{dx}$$

Probably the most important application of the rule (4-48) is the formation of the "energy" or Hamiltonian operator $\hat{H}$. Since the energy in classical mechanics is given by Eq. (3-6b) [see also Eq. (3-8)], then according to Postulate 6 we have

$$\hat{H} = \frac{1}{2m} \hat{P}^2 + V(\hat{X}) \tag{4-49a}$$

or

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \tag{4-49b}$$

That is, the energy operator $\hat{H}$ transforms any given function $\phi(x)$ into the function

$$[\hat{H}\phi(x)] \equiv -\frac{\hbar^2}{2m}\frac{d^2\phi(x)}{dx^2} + V(x)\phi(x)$$

The position, momentum and energy are not the *only* observables which the system under consideration might have in quantum mechanics. There may be several other observables, some of them having *no* analogues in classical mechanics, which characterize certain other attributes of the system. However, we shall confine our discussion here to a description of the system solely in terms of the three observables defined above.

Before examining some of the immediate consequences of the foregoing expressions for $\hat{X}$, $\hat{P}$ and $\hat{H}$, it is well to make certain that these three operators are Hermitian, as required by Postulate 2. We recall that $\hat{A}$ is said to be an Hermitian operator if and only if $(\phi_1, \hat{A}\phi_2) = (\hat{A}\phi_1, \phi_2)$ for any two $\mathcal{H}$-vectors $\phi_1(x)$ and $\phi_2(x)$. It will be left as an exercise for the reader to show that the position operator in Eq. (4-46) satisfies this requirement [see Exercise 49]. To show that the momentum operator in Eq. (4-47) is Hermitian, we proceed as follows: Using the definition of the inner product in Eq. (2-32), we have for any two $\mathcal{H}$-vectors $\phi_1(x)$ and $\phi_2(x)$

$$(\phi_1, \hat{P}\phi_2) = \int_{-\infty}^{\infty} \phi_1^*(x)\,[\hat{P}\phi_2(x)]\,dx = \int_{-\infty}^{\infty} \phi_1^*(x)\left[-i\hbar\frac{d\phi_2(x)}{dx}\right]dx$$

$$= -i\hbar\int_{-\infty}^{\infty} \phi_1^*(x)\,d\,[\phi_2(x)]$$

Integrating by parts, we obtain

$$(\phi_1, \hat{P}\phi_2) = -i\hbar\left\{\phi_1^*(x)\phi_2(x)\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{\infty} \phi_2(x)\,d\,[\phi_1^*(x)]\right\}$$

Now, the first term on the right vanishes for the following reason: We proved in Sec. 2-3 that the inner product of any two $\mathcal{H}$-vectors $\phi_1(x)$ and $\phi_2(x)$ exists in the sense that

$$|(\phi_1, \phi_2)| \equiv \left|\int_{-\infty}^{\infty} \phi_1^*(x)\phi_2(x)dx\right| < \infty$$

But this can be true only if both the real and imaginary parts of the integrand, $\phi_1^*(x)\phi_2(x)$, approach zero as $x \to \pm\infty$. Thus, the first term on the right in the previous equation vanishes, and we are left with

$$(\phi_1, \hat{P}\phi_2) = i\hbar \int_{-\infty}^{\infty} \phi_2(x) \left[\frac{d\phi_1^*(x)}{dx} dx\right] = i\hbar \int_{-\infty}^{\infty} \left[\frac{d\phi_1(x)}{dx}\right]^* \phi_2(x)\, dx$$

$$= \int_{-\infty}^{\infty} \left[-i\hbar \frac{d\phi_1(x)}{dx}\right]^* \phi_2(x) dx = \int_{-\infty}^{\infty} [\hat{P}\phi_1(x)]^* \phi_2(x)\, dx$$

so

$$(\phi_1, \hat{P}\phi_2) = (\hat{P}\phi_1, \phi_2)$$

which proves that $\hat{P}$ is indeed an Hermitian operator.

Exercise 49.

(a) Prove that the position operator $\hat{X}$, as defined by Eq. (4-46), is an Hermitian operator.

(b) In the same way, show that if $f(x)$ is any well-behaved real function of x, then the operator $f(\hat{X})$ is an Hermitian operator.

Finally, to show that the energy operator H in Eq. (4-49) is Hermitian, we can proceed most simply as follows: We first observe from Eq. (4-49a) that $\hat{H}$ is the *sum* of two operators, namely the kinetic energy operator, $\hat{P}^2/2m$, and the potential energy operator, $V(\hat{X})$. The second operator is Hermitian in consequence of part (b) of the preceding exercise; for the first operator, we note that since $\hat{P}$ is Hermitian, then

$$\left(\phi_1, \frac{1}{2m}\hat{P}^2\phi_2\right) = \frac{1}{2m}(\phi_1, \hat{P}\hat{P}\phi_2) = \frac{1}{2m}(\hat{P}\phi_1, \hat{P}\phi_2)$$

$$= \frac{1}{2m}(\hat{P}\hat{P}\phi_1, \phi_2) = \left(\frac{1}{2m}\hat{P}^2\phi_1, \phi_2\right)$$

Therefore, we see that the kinetic energy operator and the potential energy operator are *each* Hermitian; the Hermiticity of $\hat{H}$ then follows from the theorem proved in Exercise 16 that the sum of two Hermitian operators is itself Hermitian.

We have now postulated the precise forms for the position operator $\hat{X}$, the momentum operator $\hat{P}$, and the energy or Hamiltonian operator $\hat{H}$, and we have demonstrated that these operators are Hermitian as required by Postulate 2. Let us next examine some of the important consequences of so representing these observables by these operators. We consider first the energy operator $\hat{H}$.

It will be recalled that the operator $\hat{H}$ was crucially involved in our statement of Postulate 5. Indeed, with Eq. (4-49b), we find that Eqs. (4-25) and (4-26) take the following respective forms:†

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\eta_n(x) + V(x)\eta_n(x) = E_n\eta_n(x) \tag{4-50}$$

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) + V(x)\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t) \tag{4-51}$$

These two equations are the celebrated Schrödinger equations of quantum mechanics: Eq. (4-50) is called the *time-independent Schrödinger equation*, and Eq. (4-51) is called the *time-dependent Schrödinger equation.* We showed in Sec. 4-4c that if the solutions $\{\eta_n(x)\}$ and $\{E_n\}$ to the first equation can be found, then a general solution $\Psi(x,t)$ to the second equation can be written down immediately, provided $\Psi(x,0)$ is given [see Eq. (4-42)]. However, it is very important, from a logical standpoint, not to confuse these two equations. The time-independent Schrödinger equation (4-50) is the *eigenvalue equation for the energy operator*, whereas the time-dependent Schrödinger equation (4-51) is the *fundamental time-evolution equation for the state vector.* This very basic distinction between the two Schrödinger equations should not become obscured by the similarities in their appearances or by the close relationship between their solutions.

In classical mechanics we know that the motion of a particle in a potential field $V(x)$ is unaltered if we add to $V(x)$ any constant C [see part (a) of Exercise 21]. It is not difficult to show that this is also true in quantum mechanics: If $V(x)$ is replaced by $V'(x) \equiv V(x) + C$, then it follows from Eqs. (4-49) that the new Hamiltonian operator will be

$$\hat{H}' = \hat{H} + C$$

where $\hat{H}$ is the Hamiltonian for $V(x)$. If $\{\eta_n(x)\}$ and $\{E_n\}$ are the eigenvectors and eigenvalues of $\hat{H}$, then clearly

$$\hat{H}'\eta_n(x) = \hat{H}\eta_n(x) + C\eta_n(x) = E_n\eta_n(x) + C\eta_n(x) = (E_n + C)\eta_n(x)$$

from which we may conclude that the eigenvectors and eigenvalues of $\hat{H}'$ are

$$\left.\begin{aligned}\eta_n'(x) &= \eta_n(x)\\ E_n' &= E_n + C\end{aligned}\right\} n = 1,2,\ldots$$

†In Eq. (4-51) we have written "∂" instead of "d" to emphasize that the x-differentiation of $\Psi(x,t)$ is to be performed treating t as a constant, while the t-differentiation of $\Psi(x,t)$ is to be performed treating x as a constant. See footnote, p. 37.

Now with $\hat{H}'$ as the Hamiltonian, the system will evolve in time t from any given initial state $\Psi_0(x)$ to the state

$$\Psi'_t(x) = \sum_{n=1}^{\infty} (\eta'_n, \Psi_0) e^{-iE'_n t/\hbar} \eta'_n(x) = \sum_{n=1}^{\infty} (\eta_n, \Psi_0) e^{-i(E_n + C)t/\hbar} \eta_n(x)$$

Using Eq. (2-20c), this is just

$$\Psi'_t(x) = e^{-iCt/\hbar} \sum_{n=1}^{\infty} (\eta_n, \Psi_0) e^{-iE_n t/\hbar} \eta_n(x) = e^{-iCt/\hbar} \Psi_t(x)$$

where $\Psi_t(x)$ is what the state vector would have been if the system's Hamiltonian operator were $\hat{H}$ instead of $\hat{H}'$. The essential point here is that $\Psi'_t(x)$ differs from $\Psi_t(x)$ only by a scalar factor of square modulus unity; thus, according to Postulate 1, $\Psi'_t(x)$ and $\Psi_t(x)$ correspond to the same physical state. We see then that, in both classical mechanics and quantum mechanics, an additive constant C in the potential function has no effect on the motion of the system. In classical mechanics this is a consequence of the fact that C, considered as a *function of* x, has zero derivative; in quantum mechanics, on the other hand, this is evidently a consequence of the fact that C, considered as a *Hilbert space operator*, has all functions as eigenfunctions with itself as eigenvalue.

The finding of those functions $\{\eta_n(x)\}$ and numbers $\{E_n\}$ which render Eq. (4-50) an identity—i.e., the *solving* of the time-independent Schrödinger equation—is of special importance. Not only does this yield the physically important "energy levels" of the system, $E_1, E_2, \ldots$, but, as mentioned above, it also provides us with an explicit representation for the time-varying state vector through Eq. (4-42) [or equivalently, Eqs. (4-44) and (4-45)]. Now in order to solve Eq. (4–50), it is clearly necessary to specify a definite form for the potential function, $V(x)$. But having done this, one is then usually faced with a very formidable exercise in the application of the methods and techniques of differential equation theory; indeed, Eq. (4-50) has been solved *exactly* only for a very few simple forms for $V(x)$.

As an example of the kinds of results one can expect, we shall simply exhibit, but *not* derive, the energy eigenvectors and eigenvalues for the case

$$V(x) = \frac{1}{2} kx^2 \qquad (k > 0)$$

The reader will recognize this as the potential function for the "harmonic oscillator"—i.e., for a particle experiencing a spring force

$F(x) \equiv -dV/dx = -kx$, where k is the spring stiffness. Now, in classical mechanics, we know that a solution to Newton's equation (3-3a) for this force function leads to the conclusion that the particle oscillates sinusoidally about the origin with frequency $\nu = \omega/2\pi$, where ω is defined by

$$\omega \equiv \sqrt{k/m}$$

Furthermore, if A is the "amplitude" of these oscillations (i.e., the maximum value of x), then the energy of the system is

$$E_A = \frac{1}{2}kA^2 \equiv \frac{1}{2}m\omega^2 A^2 \qquad \text{(classical harmonic oscillator)} \tag{4-52a}$$

Since A can have any nonnegative value, then according to Eq. (4-52a) the energy of this system in classical mechanics can have *any value* greater than or equal to zero. To examine this problem from the standpoint of quantum mechanics, we must evidently solve the time-independent Schrödinger equation, which in this case takes the form

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\eta_n(x) + \frac{kx^2}{2}\eta_n(x) = E_n\eta_n(x)$$

In terms of the quantity ω defined above, it is found after considerable mathematical labor that the eigenvalues are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \quad n = 0,1,2,\ldots \text{ (quantum harmonic oscillator)} \tag{4-52b}$$

and the corresponding eigenvectors are given by

$$\eta_n(x) = \frac{(-1)^n}{\sqrt{\sqrt{\pi}\,2^n n!}}\left(\frac{\hbar}{m\omega}\right)^{\frac{2n-1}{4}} e^{\frac{m\omega}{2\hbar}x^2}\frac{d^n}{dx^n}\left[e^{-\frac{m\omega}{\hbar}x^2}\right]$$

We shall not offer any comment here upon the expression for the energy eigenvectors of the harmonic oscillator, except to note that they turn out to be pure real, and also that they can be shown to satisfy the requisite conditions of orthonormality and completeness in Eqs. (4-2). With regard to the energy eigenvalues in Eq. (4-52b), we see that, in contrast to the situation in classical mechanics, the allowed energy values are *discrete* rather than continuous, with a separation between the levels of $\hbar\omega \equiv \hbar\sqrt{k/m}$; moreover, the lowest energy level is *not* zero, but $\hbar\omega/2$. It is also interesting to note that these energy levels are very similar to those which were *postulated* by Planck for the radiating oscillators inside a constant temperature cavity [see Chapter 1].

We shall explore the relation between Eqs. (4-52a) and (4-52b) when we discuss the connection between classical and quantum mechanics in Sec. 4-5c. In Sec. 4-5d we shall examine more closely the details of *solving* the time-independent Schrödinger equation—but for a potential function $V(x)$ which is even simpler than that for the harmonic oscillator. For now, we turn to a consideration of the position and momentum eigenvalue equations.

The treatment of the eigenvalue equations for the position and momentum operators presents problems of a rather peculiar nature. Let x_0 denote an eigenvalue of $\hat{X}$, and let $\delta_{x_0}(x)$ denote the corresponding eigenvector; similarly, let p_0 and $\theta_{p_0}(x)$ denote an eigenvalue and eigenvector of $\hat{P}$:

$$\hat{X}\delta_{x_0}(x) = x_0\delta_{x_0}(x)$$

$$\hat{P}\theta_{p_0}(x) = p_0\theta_{p_0}(x)$$

Substituting for $\hat{X}$ and $\hat{P}$ their specific forms, these eigenvalue equations read

$$x\delta_{x_0}(x) = x_0\delta_{x_0}(x) \tag{4-53}$$

$$-i\hbar\frac{d}{dx}\theta_{p_0}(x) = p_0\theta_{p_0}(x) \tag{4-54}$$

Exercise 50.

(a) Show from Eq. (4-53) that the function $\delta_{x_0}(x)$ must have the property that $\delta_{x_0}(x) = 0$ for any $x \neq x_0$, but that $\delta_{x_0}(x_0)$ can have *any* value. [*Hint*: Write Eq. (4-53) as $(x - x_0)\delta_{x_0}(x) = 0$.]

(b) Show that the function

$$\theta_{p_0}(x) = e^{ip_0x/\hbar} \tag{4-55}$$

satisfies Eq. (4-54). [*Hint:* Recall Exercise 7 and Eq. (2-20e).]

Now, it is obvious from Eq. (4-53) that any value for x_0 is as good as any other value; similarly, we see from Eq. (4-55) that p_0 also may have any value. Therefore, the eigenvalues of $\hat{X}$ and $\hat{P}$ are *continuously distributed over the entire real axis:*

$$\left.\begin{aligned} -\infty < x_0 < +\infty \\ -\infty < p_0 < +\infty \end{aligned}\right\} \tag{4-56}$$

This is an easily obtained, if not particularly exciting, conclusion about the *eigenvalues* of $\hat{X}$ and $\hat{P}$; however, we run into difficulties when we come to consider the corresponding *eigenvectors.*

With regard to the eigenvector $\delta_{x_0}(x)$, we found in part (a) of Exercise 50 that this function must vanish everywhere except at $x = x_0$, and there it can have any value. Now such a function would normally give zero when computing an expansion coefficient of the kind in Eq. (4-6b).

$$(\delta_{x_0}, \Psi_t) = \int_{-\infty}^{\infty} \delta^*_{x_0}(x)\Psi_t(x)\,dx$$

Indeed, the only way we can *avoid* the unacceptable conclusion that all these expansion coefficients vanish, is to make $\delta_{x_0}(x_0)$ *infinite* in such a way that its product with the infinitesimal dx, $\delta_{x_0}(x_0)\,dx$, is a *finite* number. A rigorous treatment of this highly unusual function lies beyond the scope of this book, and in fact even transcends the scope of ordinary calculus. The function $\delta_{x_0}(x)$ is usually written $\delta(x - x_0)$, and is called by physicists the *Dirac delta function* (although it is not really a "function" in the strict mathematical sense). We shall discuss some of the properties of $\delta(x - x_0)$, as well as some of the important consequences of these properties, in Sec. 4-6b.

At first glance, it might seem that the momentum eigenvectors $\theta_{p_0}(x)$ in Eq. (4-55) are free from any difficulties; however, this is not so. When we compute the norm of $\theta_{p_0}(x)$, we find by virtue of Eq. (2-20d) that

$$(\theta_{p_0}, \theta_{p_0}) = \int_{-\infty}^{\infty} |\theta_{p_0}(x)|^2\,dx = \int_{-\infty}^{\infty} |e^{ip_0 x/\hbar}|^2\,dx = \int_{-\infty}^{\infty} 1 \cdot dx = \infty$$

That is, there is no way in which the functions $\theta_{p_0}(x)$ can be normalized to unity, as required by Postulates 1 and 2.

The difficulties that we are witnessing with regard to the eigenvectors of the position and momentum operators can ultimately be traced to the fact that these operators have *continuously* distributed eigenvalues [see Eq. (4-56)]. It was precisely to avoid these difficulties that we restricted our discussion in the previous sections to operators with discretely distributed eigenvalues. In order to circumvent these difficulties, it is necessary to modify the definitions of orthonormality and completeness, as these terms apply to eigenbasis vectors associated with continuously distributed eigenvalues. We shall discuss these modifications briefly in Sec. 4-6b; for now, though, all we shall need to know is the following: The *position eigenvector* $\delta_{x_0}(x)$, which corresponds to the eigenvalue x_0, is zero everywhere except at $x = x_0$, at which point it has an "infinite spike"; the *momentum eigenvector* $\theta_{p_0}(x)$, which corresponds to the eigenvalue

p_0, is, apart from some "normalization constant," equal to $\exp(ip_0x/\hbar)$.

Despite the above difficulties with the eigenvectors of the position and momentum operators, it is possible to calculate the *expectation values* of these observables in a very straightforward manner, by means of the formula in Eq. (4-16):

Exercise 51. Using Eq. (4-16), show that the expectation values of position and momentum in the state $\Psi(x,t)$ are

$$\langle \hat{X} \rangle_t = \int_{-\infty}^{\infty} x\,|\Psi(x,t)|^2\,dx \tag{4-57a}$$

$$\langle \hat{P} \rangle_t = -i\hbar \int_{-\infty}^{\infty} \Psi^*(x,t)\,\frac{\partial \Psi(x,t)}{\partial x}\,dx \tag{4-58a}$$

Moreover, the *uncertainties* in position and momentum in the state $\Psi(x,t)$ may be calculated from Eq. (4-14) if we compute, in addition, the quantities $\langle \hat{X}^2 \rangle_t$ and $\langle \hat{P}^2 \rangle_t$:

Exercise 52. Show that

$$\langle \hat{X}^2 \rangle_t = \int_{-\infty}^{\infty} x^2\,|\Psi(x,t)|^2\,dx \tag{4-57b}$$

$$\langle \hat{P}^2 \rangle_t = \hbar^2 \int_{-\infty}^{\infty} \left| \frac{\partial \Psi(x,t)}{\partial x} \right|^2 dx \tag{4-58b}$$

[*Hint:* For $\langle \hat{P}^2 \rangle_t$, use the fact that, since $\hat{P}$ is Hermitian, then $(\Psi_t, \hat{P}^2 \Psi_t) = (\hat{P}\Psi_t, \hat{P}\Psi_t)$.]

In particular, Eqs. (4-57) can easily be generalized: If $f(x)$ is any well-behaved, real function of x, then according to Postulate 6, $f(x)$ is an *observable* with *operator* $f(\hat{X}) = f(x)$. The expectation value of this observable in the state $\Psi(x,t)$ is, according to Eq. (4-17),

$$\langle f(\hat{X}) \rangle_t = (\Psi_t, f(\hat{X})\Psi_t) = \int_{-\infty}^{\infty} \Psi^*(x,t)\,[f(x)\Psi(x,t)]\,dx$$

or

$$\langle f(\hat{X}) \rangle_t = \int_{-\infty}^{\infty} f(x)|\Psi(x,t)|^2\,dx \tag{4-59a}$$

This last equation can be given a very interesting and useful interpretation. We recall from Exercise 26 that, if f has a Taylor

series expansion, then the operator $f(\hat{A})$ has eigenvalues $\{f(A_n)\}$, where $\{A_n\}$ are the eigenvalues of $\hat{A}$. Therefore, since $\hat{X}$ has eigenvalues $\{x\}$ for all $-\infty < x < \infty$, then the operator $f(\hat{X})$ has eigenvalues $\{f(x)\}$ for all $-\infty < x < \infty$. Thus, Eq. (4-59a) expresses $\langle f(\hat{X})\rangle_t$ as a sort of "weighted sum" of the possible $f(\hat{X})$-values. To bring this out more explicitly, let us for the moment write Eq. (4-59a) as a *discrete* sum: We partition the entire x-axis into subintervals $\Delta x_1, \Delta x_2, \ldots,$ and we let x_k denote an x-value inside the subinterval Δx_k. Then if we make the lengths of these subintervals infinitesimally small, we may write by the very definition of the integral,

$$\langle f(\hat{X})\rangle_t = \sum_k f(x_k)|\Psi(x_k,t)|^2 \, \Delta x_k \tag{4-59b}$$

We now compare this with Eq. (2-8):

$$\langle f(v)\rangle = \sum_k f(v_k)p_k \tag{2-8}$$

which gives the mean value of $f(v)$ for a series of v_k-values distributed with probabilities p_k. Equation (2-8) gives $\langle f(v)\rangle$ as a weighted sum of the possible values $f(v_k)$, while Eq. (4-59b) gives $\langle f(\hat{X})\rangle_t$ as a weighted sum of the possible values $f(x_k)$. Because f is an arbitrary function, these considerations imply the correspondence

$$p_k \longleftrightarrow |\Psi(x_k,t)|^2 \, \Delta x_k \tag{4-60}$$

We note in particular that this correspondence is consistent with the condition $\Sigma_k p_k = 1$, since

$$\sum_k |\Psi(x_k,t)|^2 \, \Delta x_k = \int_{-\infty}^{\infty} |\Psi(x,t)|^2 \, dx = (\Psi_t, \Psi_t) = 1$$

Now Eq. (2-8) gives the mean value of $f(v)$ *because* p_k is the probability that the particular value v_k will be obtained in a random selection from the set of v-values; thus, the correspondence (4-60) *implies* that $|\Psi(x_k,t)|^2 \Delta x_k$ is the probability that a value in the particular interval Δx_k will be obtained in a position measurement on the state $\Psi(x,t)$. Returning to the integral expression for $\langle f(\hat{X})\rangle_t$ in Eq. (4-59a), we conclude that:

$|\Psi(x,t)|^2 \, dx$ = the probability that a position measurement on the state $\Psi(x,t)$ will yield a value between x and $x + dx$. (4-61)

It follows from this that the probability for a position measurement on the state $\Psi(x,t)$ to yield a value somewhere between x_1 and x_2 is

$$P(x_1,x_2;t) = \int_{x_1}^{x_2} |\Psi(x,t)|^2 \, dx \tag{4-62}$$

because this is just the "sum" of the probabilities for measuring a value in any of the dx-intervals between x_1 and x_2.

On account of Eq. (4-61), the quantity $|\Psi(x,t)|^2 \equiv \Psi_t^*(x)\Psi_t(x)$ is called the *position probability density function* (the term "density" is used because Eq. (4-61) implies that $|\Psi(x,t)|^2$ has units of probability *per unit* x). Although the state vector itself has no direct physical significance, we see that its square modulus has a very deep physical significance. This particular aspect of the state vector was first recognized by Max Born, and is undoubtedly one of its most important and frequently used properties. However, it must be emphasized that Eq. (4-61) by no means exhausts all the physical implications of the state vector.

In order to understand fully the significance of the position probability density function, we show in Fig. 4 a plot of $|\Psi(x,t)|^2$ versus x for some hypothetical state $\Psi(x,t)$. If one were to make a series of very many *repeated* measurements of the position on the state $\Psi(x,t)$, and if the results were plotted as a frequency bar graph over small, equal-size bins, then Eq. (4-61) implies that the shape of this graph would follow the shape of the curve in Fig. 4, to within random statistical fluctuations. Thus the curve in Fig. 4 is essentially the same as the curve in Fig. 3, taking into account the fact that the eigenvalues of $\hat{X}$ are continuously distributed over the entire real axis. Since

$$\int_{-\infty}^{\infty} |\Psi(x,t)|^2 \, dx = (\Psi_t,\Psi_t) = 1$$

then the area under the curve in Fig. 4 is unity.

Equation (4-62) implies that the area under the curve $|\Psi(x,t)|^2$ between x_1 and x_2, shown shaded in Fig. 4, is equal to $P(x_1,x_2;t)$, the probability of finding the particle between x_1 and x_2. We have used the phrase "finding the particle" in a special sense, and it is important that we understand precisely what we mean, and what we do not mean, by it. According to our discussion in Sec. 4-3b, *prior* to the position measurement we should *not* try to picture the particle as "really" being either inside or outside the interval $[x_1,x_2]$, with the position measurement "discovering" which of these two alternatives

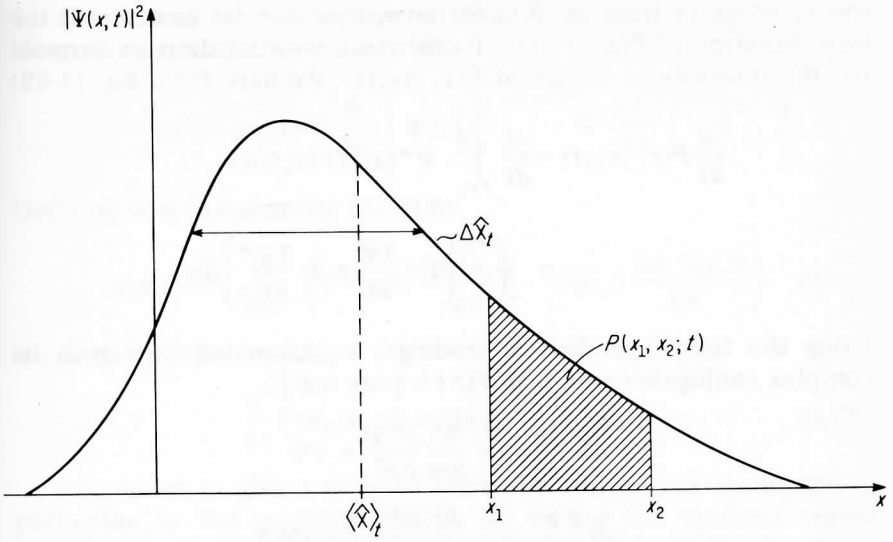

Fig. 4. A plot of the position probability density function $|\Psi(x,t)|^2$ versus x for a hypothetical state $\Psi(x,t)$. The position expectation value and uncertainty are indicated schematically. The total area under the curve is unity, by virtue of the fact that $(\Psi_t, \Psi_t) = 1$; the shaded area under the curve between x_1 and x_2 is numerically equal to $P(x_1, x_2; t)$, the probability that a position measurement at time t will yield a value between x_1 and x_2. We have drawn the curve as though $|\Psi(x,t)|^2$ were fairly well localized in one region of the x-axis; however, it is important to realize that the curve could very well consist of two or more widely separated humps. In such a case, $\Delta\hat{X}_t$ would be large, and $\langle\hat{X}\rangle_t$ would not be a particularly useful quantity.

is actually the case. For, if the state vector $\Psi_t(x)$ does *not* coincide with one of the position eigenvectors $\delta_{x_0}(x)$—as is obviously the case in Fig. 4—then we may *not* speak of the particle as "having a position" in the usual sense of this phrase. So when we say that $P(x_1, x_2; t)$ is "the probability of finding the particle somewhere between x_1 and x_2," what we really mean is that $P(x_1, x_2; t)$ is "the probability that a position measurement on the state $\Psi(x,t)$ will *develop* a position value for the particle somewhere between x_1 and x_2." It is all right to use the former phrase for brevity, provided we keep in mind the latter, more explicit interpretation.

As time evolves, the state vector $\Psi_t(x)$ changes, and the position probability density curve in Fig. 4 will change its shape and position in some more or less complicated way, subject to the condition that the area under it remain equal to unity. In Sec. 4-5c we shall examine

the behavior in time of $\langle \hat{X} \rangle_t$; for now, however, let us examine the time variation of $P(x_1, x_2; t)$. To this end, we first derive a formula for the time-rate-of-change of $P(x_1, x_2; t)$. We have from Eq. (4-62)

$$\frac{d}{dt} P(x_1, x_2; t) = \frac{d}{dt} \int_{x_1}^{x_2} \Psi^*(x,t)\Psi(x,t)\,dx$$

$$= \int_{x_1}^{x_2} \left[\Psi^* \frac{\partial \Psi}{\partial t} + \Psi \frac{\partial \Psi^*}{\partial t} \right] dx$$

Using the time-dependent Schrödinger equation together with its complex conjugate [note that $V(x)$ is pure real],

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V\Psi$$

$$-i\hbar \frac{\partial \Psi^*}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi^*}{\partial x^2} + V\Psi^*$$

it is not too difficult to eliminate the time derivatives in the above equation to obtain

$$\frac{d}{dt} P(x_1, x_2; t) = \frac{i\hbar}{2m} \int_{x_1}^{x_2} \left[\Psi^* \frac{\partial^2 \Psi}{\partial x^2} - \Psi \frac{\partial^2 \Psi^*}{\partial x^2} \right] dx \qquad (4\text{-}63)$$

Exercise 53. Carry out the steps leading to Eq. (4-63).

The integral on the right-hand side of Eq. (4-63) can be explicitly evaluated by an integration-by-parts. Remembering that "$\partial \Psi / \partial x$" just means "$d\Psi / dx$ with t treated as a constant," we have for the first term in Eq. (4-63)

$$\int_{x_1}^{x_2} \Psi^* \frac{\partial^2 \Psi}{\partial x^2} dx = \int_{x_1}^{x_2} \Psi^* \frac{\partial}{\partial x} \left(\frac{\partial \Psi}{\partial x} \right) dx = \int_{x_1}^{x_2} \Psi^* d \left(\frac{\partial \Psi}{\partial x} \right)$$

$$= \Psi^* \frac{\partial \Psi}{\partial x} \Bigg|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{\partial \Psi}{\partial x} d(\Psi^*)$$

so

$$\int_{x_1}^{x_2} \Psi^* \frac{\partial^2 \Psi}{\partial x^2} dx = \Psi^* \frac{\partial \Psi}{\partial x} \Bigg|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{\partial \Psi}{\partial x} \frac{\partial \Psi^*}{\partial x} dx$$

A similar expression holds for the second term in Eq. (4-63), except that the roles of Ψ and Ψ^* are interchanged. When the expression

for the second term is subtracted from that for the first term, the integrals cancel, and we are left with

$$\frac{d}{dt}P(x_1,x_2;t) = \frac{i\hbar}{2m}\left[\Psi^*\frac{\partial\Psi}{\partial x}\Bigg|_{x_1}^{x_2} - \Psi\frac{\partial\Psi^*}{\partial x}\Bigg|_{x_1}^{x_2}\right]$$

Defining now the quantity $S(x,t)$ by

$$S(x,t) \equiv -\frac{i\hbar}{2m}\left(\Psi^*(x,t)\frac{\partial\Psi(x,t)}{\partial x} - \Psi(x,t)\frac{\partial\Psi^*(x,t)}{\partial x}\right) \tag{4-64}$$

we obtain the final result

$$\frac{d}{dt}P(x_1,x_2;t) = S(x_1,t) - S(x_2,t) \tag{4-65}$$

In order to give a physical interpretation to this result, and in particular to the quantity $S(x,t)$, let us for the moment regard Eq. (4-65) simply as a formula for the time-rate-of-change of the shaded area in Fig. 4. Since the *total* area under the curve in Fig. 4 is always equal to unity, then any increase in the area inside $[x_1,x_2]$ must be accompanied by an equal decrease in the area outside $[x_1,x_2]$, and vice versa. Thus, it is reasonable to think of the change in the shaded area in Fig. 4 as resulting from a "flow of area" across the two boundary lines at x_1 and x_2. More specifically, suppose we let $R(x,t)$ denote the *rate* at which area is crossing the point x at time t in the positive x-direction, with the convention that an area flow in the negative x-direction is specified by a negative value for $R(x,t)$. The quantity $R(x_1,t) + (-R(x_2,t))$ would then denote the rate at which area is entering the interval $[x_1,x_2]$ at x_1 plus the rate at which area is entering this interval at x_2; clearly, this is just the *net* rate of increase of the area inside $[x_1,x_2]$, $dP(x_1,x_2;t)/dt$. We now observe that Eq. (4-65) is precisely of this form; consequently, we may interpret the quantity $S(x,t)$ as being the instantaneous rate at which area or "position probability" is crossing the point x in the positive x-direction, with negative values for $S(x,t)$ signifying a flow of position probability in the negative x-direction. We call $S(x,t)$ the *position probability current* at the point x at time t.

With this interpretation of $S(x,t)$, Eq. (4-65) is merely the statement that the instantaneous rate of change of the probability for finding the particle between x_1 and x_2, is determined solely by the instantaneous values of the position probability currents at x_1 and x_2; if the current at x_1 is larger (smaller) than the current at x_2, then the probability of finding the particle between x_1 and x_2 is increasing (decreasing).

Exercise 54.

(a) With the help of Eq. (2-13), show that the position probability current can also be written

$$S(x,t) = \frac{\hbar}{m} \operatorname{Im}\left(\Psi^*(x,t)\frac{\partial \Psi(x,t)}{\partial x}\right) \tag{4-66}$$

which, incidently, shows that $S(x,t)$ is *pure real*, as we would expect.

(b) Suppose the potential function $V(x)$ is such that the energy eigenvectors $\{\eta_n(x)\}$ turn out to be pure real. For this case, show that the position probability current vanishes identically if the system is in a stationary state, $\Psi^{(n)}(x,t)$.

In most practical situations, the physicist deals with a large number N of noninteracting particles, all subject to the same potential and all in the same state $\Psi(x,t)$. In such a case, he often speaks of there being $N \cdot P(x_1,x_2;t)$ particles inside the interval $[x_1,x_2]$, and $N \cdot S(x,t)$ particles per second crossing the point x in the positive x-direction (or negative x-direction if $S(x,t) < 0$). Although these statements are *not* literally correct within the framework of orthodox quantum mechanics—for example, they would be meaningless for $N = 1$—no *practical* difficulty is encountered if one does not try to single out *specific ones* of the particles as really being inside $[x_1,x_2]$, or *specific ones* of the particles as really crossing the point x at time t. But strictly speaking, $|\Psi(x,t)|^2$ is a *position probability* density, not a *particle* density, and $S(x,t)$ is a position probability current, not a particle current. These considerations are illustrative of the dramatic revision which quantum mechanics has effected with respect to the familiar, classical concept of a "physical particle." In the next section we shall pursue this conceptual revision to what might be called its logical extreme; in the section following that, we shall show how quantum mechanics allows us to regain, in the *macroscopic limit*, our familiar classical description of a "particle" and its dynamical behavior.

4-5b The Position-Momentum Uncertainty Relation The Wave-Particle Duality

We shall now demonstrate that the postulates of quantum mechanics imply that a particle moving along the x-axis can exhibit the attributes of either a *particle* or a *wave*. In order to avoid confus-

ing the "particle" of our system with the particle "property" or "attribute," we shall adopt for our system an *electron* on the x-axis.

Suppose first that the state vector of the electron coincides with one of the eigenvectors $\delta_{x_0}(x)$ of the position operator $\hat{X}$. According to our discussion in Sec. 4-3b, we may say in this case—and *only* in this case—that the observable *position* "has the value x_0," or more simply that the electron "is at the point x_0." Note that this conclusion is entirely consistent with the *spatially localized* form of the function $\delta_{x_0}(x)$: since this function has an infinite spike at $x = x_0$ and vanishes everywhere else, then the position probability density function $|\delta_{x_0}(x)|^2$ clearly implies that a measurement of position will necessarily find the electron at the point $x = x_0$. Now it must be emphasized that this property of "having a position" or of "being spatially localized" is essentially the *defining* property of a *particle*. And it is *only* when the state vector of the electron coincides with one of these infinitely localized position eigenvectors $\delta_{x_0}(x)$ that we can meaningfully assert that the electron "has a position" and therefore "is a particle."

Suppose, on the other hand, that the state vector of the electron coincides with one of the eigenvectors $\theta_{p_0}(x)$ of the momentum operator $\hat{P}$. According to our discussion in Sec. 4-3b, we may say in this case—and *only* in this case—that the observable *momentum* "has the value p_0," or more simply that the electron "is moving with momentum p_0." Now, if we write down the explicit form of the function $\theta_{p_0}(x)$ in Eq. (4-55) [see Eq. (2-20a)],

$$\theta_{p_0}(x) = e^{ip_0x/\hbar} \equiv \cos\left[\frac{p_0x}{\hbar}\right] + i \sin\left[\frac{p_0x}{\hbar}\right]$$

we can see that this function is definitely *not* localized on the x-axis. Instead, $\theta_{p_0}(x)$ is seen to have an *infinite, periodic, spatial extension*, with fundamental period or "wavelength"

$$\lambda_0 = h/p_0 \tag{4-67}$$

Exercise 55. Show that $\theta_{p_0}(x)$ is periodic in x with period $\lambda_0 = h/p_0$.
[*Hint*: Prove that $\theta_{p_0}(x + \lambda_0) \equiv \theta_{p_0}(x)$.]

Now the property of *periodic spatial extension* is essentially the defining attribute of a *wave*, just as the property of *sharp spatial localization* is the defining attribute of a *particle*. Therefore, when the state vector of the electron coincides with the momentum eigenvector $\theta_{p_0}(x)$—i.e., when the electron has momentum p_0—then the electron is in a certain sense a *wave*, and can be said to "have a wave-

length" whose numerical value is given by Eq. (4-67). We note in passing that this is essentially the same conclusion that physicists were *experimentally* led to *before* the invention of quantum mechanics [recall our discussion of Eq. (1-2)].†

We may summarize these results by saying that, when the state vector of the electron coincides with an eigenvector of $\hat{X}$, then the electron has the attributes of (and therefore "is") a *particle*; on the other hand, when the state vector of the electron coincides with an eigenvector of $\hat{P}$, then the electron has the attributes of (and therefore "is") a *wave*. The questions now arise, if the electron is in the state $\delta_{x_0}(x)$, does it have any wave attributes, and if the electron is in the state $\theta_{p_0}(x)$, does it have any particle attributes? Logically, we expect a negative answer to both questions, since the properties of "sharp spatial localization" and "periodic spatial extension" are mutually exclusive properties.

To verify this conjecture, suppose first that the electron is in the state $\Psi_t(x) = \delta_{x_0}(x)$, and suppose we wish to calculate the expectation value of the momentum. Following Eq. (4-58a), we write

$$\langle \hat{P} \rangle_t = -i\hbar \int_{-\infty}^{\infty} \delta^*_{x_0}(x) \frac{d\delta_{x_0}(x)}{dx} dx$$

$$= -i\hbar \int_{-\infty}^{\infty} \left(\frac{d\delta_{x_0}(x)}{dx} \right) (\delta_{x_0}(x) \cdot dx)$$

where we have used the fact that $\delta_{x_0}(x)$ is real. Now, since $\delta_{x_0}(x) = 0$ for all $x \neq x_0$, the only contribution to the integral comes at $x = x_0$; here, the product $\delta_{x_0} \cdot dx$ is *finite*, but $d\delta_{x_0}/dx$ is *undefined* owing to the radical discontinuity in $\delta_{x_0}(x)$ at $x = x_0$. Thus, $\langle P \rangle_t$ is essentially *undefined* in the state $\Psi_t(x) = \delta_{x_0}(x)$. Let us suppose next that the electron is in the state $\Psi_t(x) = \theta_{p_0}(x)$, and suppose we wish to calculate the position probability density function. According to Eqs. (4-61) and (2-20d), we have

$$|\Psi_t(x)|^2 = |\theta_{p_0}(x)|^2 = |e^{ip_0 x/\hbar}|^2 = 1$$

†It should be pointed out that, although $\theta_{p_0}(x)$ is periodic in x, the corresponding position probability density function, $|\theta_{p_0}(x)|^2$, is *not* periodic: as we shall see shortly, the square modulus of $\theta_{p_0}(x)$ is simply a constant. However, the periodic nature of $\theta_{p_0}(x)$ can be rendered "physically observable" by causing a beam of electrons with momentum p_0 to *interact* with a suitable apparatus, such as the crystal diffraction grating of the Davisson-Germer experiment. An analysis of this "interaction," which in itself is essentially a "measurement" of momentum or wavelength, is too complicated for us to consider here.

But this implies that a position measurement is *equally* likely to yield *any* value for x.

These unusual conclusions may best be understood with the help of the Compatibility Theorem and the Heisenberg Uncertainty Principle. To this end, we first make the following simple calculation:

Exercise 56. Prove that the operators $\hat{X}$ and $\hat{P}$, as defined in Postulate 6, satisfy the relation

$$\hat{X}\hat{P} - \hat{P}\hat{X} = i\hbar \tag{4-68}$$

[*Hint*: Prove that, for any $\mathcal{H}$-vector $\phi(x)$, $\hat{X}[\hat{P}\phi(x)] - \hat{P}[\hat{X}\phi(x)] = i\hbar\phi(x)$.]

So the position and momentum operators *do not commute.* Therefore, the Compatibility Theorem tells us that position and momentum are *not* "compatible" or "simultaneously measureable." Moreover, upon substituting Eq. (4-68) into the Heisenberg Uncertainty Relation, we find

$$\Delta\hat{X}_t \cdot \Delta\hat{P}_t \geqq \frac{\hbar}{2} \tag{4-69}$$

Exercise 57. Derive Eq. (4-69).

The above inequality is known as the position-momentum uncertainty relation; it is obviously very similar in form to the time-energy uncertainty relation in Eq. (4-36), although it must be noted that the quantity T_a in Eq. (4-36) is *not* to be regarded as an uncertainty in some sort of "time operator." According to Eq. (4-69), the more precisely the position of the electron is defined (i.e., the smaller $\Delta\hat{X}_t$ is), the less precisely the momentum of the electron is defined (i.e., the larger ΔP_t must be)—and vice versa. Indeed, if the electron can be said to "have a position" (i.e., if $\Delta\hat{X}_t = 0$), then it *cannot* be said to "have a momentum" (i.e., $\Delta\hat{P}_t$ must be infinite)—and of course vice versa. Therefore we see that, although quantum mechanics allows an electron to possess "particle" attributes *and* "wave" attributes, it expressly *forbids* the electron from being a particle and a wave *simultaneously*.

In light of these conclusions, we can understand the so-called wave-particle duality of Nature by the following chain of reasoning:

If we measure the *position* of an electron, then regardless of the state vector of the electron just prior to the measurement, immediately *after* the measurement it will coincide with one of the position eigenvectors, $\delta_{x_0}(x)$ [by Postulate 4]. We will then have $\Delta\hat{X} = 0$ [by Exercise 33], so that the position of the electron is well-defined and

can be said to "have a value." Evidently, the position measurement has *endowed* the state vector of the electron with the property of *sharp spatial localization*, so that the electron may truly be regarded as a "particle." However, the electron then has *no* wavelike attributes: for since $\Delta\hat{X} = 0$ then Eq. (4-69) requires that $\Delta\hat{P} = \infty$, which means that the electron cannot be said to "have a momentum or wavelength." Thus, the *measurement* of the observable "position" has *developed* the particle nature of the electron, but it has at the same time *destroyed* the wave nature of the electron.

Exercise 58. Rewrite the preceding paragraph, except start out with the phrase, "If we measure the *momentum* or *wavelength* of an electron, then . . ."

The wave-particle duality of Nature, when viewed from the standpoint of classical mechanics, represented a genuine paradox; however, it should be clear from our discussion here that this phenomenon emerges as a very logical consequence of the basic tenets of quantum mechanics. Historically, the duality was one of the chief motivations for seeking an alternative to classical mechanics. From a modern point of view, the existence in Nature of the wave-particle duality provides strong evidence for the validity of the entire quantum theory.

The real source of the wave-particle duality is evidently the strict incompatibility of position and momentum. A full appreciation of this incompatibility is essential for a proper understanding of many results in quantum mechanics. The following exercise provides a case in point.

Exercise 59. Consider a particle of mass m in an harmonic oscillator potential, $V(x) = kx^2/2 \equiv m\omega^2 x^2/2$.

(a) Suppose the particle is a *classical* particle with total energy E. Show that a position measurement *cannot* find the particle outside the interval $-\sqrt{2E/m\omega^2} \leqq x \leqq \sqrt{2E/m\omega^2}$. [*Hint*: Recall the discussion of Fig. 2.]

(b) Suppose the particle is a *quantum* particle with total energy E_n. Show that a position measurement *can* find the particle outside the interval $-\sqrt{2E_n/m\omega^2} \leqq x \leqq \sqrt{2E_n/m\omega^2}$. [*Hint*: Since the particle has total energy E_n, it must be in the stationary state $\eta_n(x) \exp(-iE_n t/\hbar)$, where $\eta_n(x)$ and E_n were written down for the harmonic oscillator in Sec. 4-5a. Show, without performing any detailed calculations, that the position probability density function for this stationary state does not vanish identically outside the interval in question.]

The phenomenon described in part (b) of the above exercise is not peculiar to the harmonic oscillator potential alone; it is found to occur for nearly all potentials of a similar "concave-up" shape. From a classical point of view, however, this phenomenon seems quite paradoxical: if the particle can be found at points where $V(x) > E$, does not this imply the patently untenable conclusion that the kinetic energy is negative? To understand why this reasoning is fallacious from a quantum viewpoint, let us consider the five observables, position, momentum, kinetic energy, potential energy, and total energy. Of the five corresponding operators, $\hat{X}$, $\hat{P}$, $\hat{P}^2/2m$, $V(\hat{X})$, and $\hat{P}^2/2m + V(\hat{X})$, straightforward calculations utilizing Eq. (4-68) reveal that two and *only* two pairs commute; more specifically the only commuting pairs of operators are $\hat{P}$ and $\hat{P}^2/2m$, and, for any reasonable function V, $\hat{X}$ and $V(\hat{X})$ [see Exercise 36]. It follows from the Compatibility Theorem that momentum and kinetic energy are compatible, as are also position and potential energy; however, position and kinetic energy are *not* compatible, momentum and potential energy are *not* compatible, and the total energy is *not* compatible with *any* of the other observables.† To verify these incompatibilities, we need only observe the great dissimilarities among the position/potential energy eigenvectors $\{\delta_{x_0}(x)\}$, the momentum/kinetic energy eigenvectors $\{\theta_{p_0}(x)\}$, and, for the harmonic oscillator, the total energy eigenvectors $\{\eta_n(x)\}$. In view of these incompatibilities we may evidently refute the classical objection, that an harmonic oscillator would have to have a negative kinetic energy if its position were such that its potential energy exceeded its total energy, simply by observing that the oscillator *cannot* be said to "have values" for all these variables simultaneously. If the oscillator "has energy E_n," it *cannot* sensibly be said to "have values" of position and kinetic energy, because $\eta_n(x)$ is not an eigenvector of either the position operator or the kinetic energy operator; moreover, if the position is *measured*, the state vector will thereby be forced into one of the position eigenvectors $\delta_{x_0}(x)$, and it will *then* be impossible to ascribe to the oscillator either a kinetic energy or a total energy. Thus we see that, because of the fundamental incompatibility between position and momentum, an harmonic oscillator can have *either* a position *or* a kinetic energy *or* a total energy, but *not* any two of these simultaneously; moreover, whatever value may actually be realized for one of these observables at any particular time will necessarily be one of the legitimate eigenvalues of the respective operator—namely, x_0 or $p_0^2/2m$ or $(n + 1/2)\hbar\omega$.

†An exception to this last statement arises if $V(x)$ = constant, in which case the kinetic and total energies are essentially the same.

As the reader pursues his study of quantum mechanics he will constantly feel the urge to try to understand or "see through" quantum phenomena in terms of common-sense classical concepts. Most of the nonorthodox views of quantum mechanics try, in varying degrees, to do just this. The orthodox view, which we have taken here, must be regarded as rather radical in this respect: it asserts, or at least strongly suggests, that there *is no* adequate interpretation of quantum mechanics via purely classical notions, and that the message which Nature is trying to get across to us is simply that such classical ways of thinking do not apply to the microscopic physical world. If we accept this view, then we would evidently be better employed in trying to "see through" classical mechanics in terms of the concepts of quantum mechanics. This, in short, will be our goal in the next section.

4-5c The Ehrenfest Equations
The Classical Limit of Quantum Mechanics

We have seen that quantum mechanics must be used instead of classical mechanics when one deals with the *microscopic* physical world. However, this does not mean that we should discard the classical theory altogether; for if we can be certain of anything at all in physics, it is that classical mechanics correctly and efficiently describes many aspects of the *macroscopic* physical world. Therefore, if quantum mechanics is indeed a more comprehensive theory of physical phenomena than classical mechanics, then it is incumbent upon the former to *reduce to* rather than replace the latter in the macroscopic limit.

In order to demonstrate that quantum mechanics does satisfy this requirement, it is sufficient to show that, for any observable $\mathcal{A}$ which has a classical analogue:

(i) In the macroscopic limit, any discreteness in the eigenvalues of $\hat{A}$ is not noticeable.

(ii) In the macroscopic limit, the uncertainty in $\mathcal{A}$, $\Delta\hat{A}_t$, is in practice so small in comparison with the expectation value of $\mathcal{A}$, $\langle\hat{A}\rangle_t$, that we can for all practical purposes say that "$\mathcal{A}$ has the value $\langle\hat{A}\rangle_t$."

(iii) In the macroscopic limit, the time evolution equation for $\langle\hat{A}\rangle_t$ coincides with the classical equation of motion for $\mathcal{A}(t)$.

Clearly, if these three requirements are fulfilled, then a measurement of $\mathcal{A}$ will in effect *always* yield the value $\langle\hat{A}\rangle_t$, and moreover this

"value of $\mathcal{A}$" will evolve continuously with time according to the laws of classical mechanics.

Unfortunately, we shall not be able to give rigorous, general proofs of the first two requirements; however, we will at least be able to see that they are quite plausible. Let us consider first requirement (i). We have seen in Sec. 4-5a that the position and momentum operators already possess continuously distributed eigenvalues; however, the energy operator is often found to have discrete eigenvalues. For example, in Sec. 4-5a we mentioned (without proof) that for the linear harmonic oscillator potential, $V(x) = kx^2/2 \equiv m\omega^2 x^2/2$, the energy eigenvalues are given by [see Eq. (4-52b)]

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \qquad n = 0,1,2,\ldots$$

Therefore, the *relative spacing* of the energy levels around the value E_n is

$$\frac{E_{n+1} - E_n}{E_n} = \frac{\hbar\omega}{\left(n + \frac{1}{2}\right)\hbar\omega} = \frac{1}{n + \frac{1}{2}}$$

Evidently, this relative spacing will be small if n is large, in which case it is approximately $1/n$. Now, *classically* such a particle would oscillate sinusoidally about the origin with frequency

$$\nu = \omega/2\pi$$

and could do so with any of the energies [see Eq. (4-52a)]

$$E_A = \frac{1}{2}m\omega^2 A^2 \qquad A \geq 0$$

Now in the *macroscopic limit* we know that this last equation is essentially valid; since the quantum expression is presumed to be universally valid, then the quantum number n must be such that

$$\left.\begin{aligned} &E_n \cong E_A \\ \text{or} \quad &\left(n + \frac{1}{2}\right) \cong \frac{1}{2}\left(\frac{m}{\hbar}\right)\omega A^2 \end{aligned}\right\} \text{ macroscopic limit}$$

Exercise 60.

(a) Show that, for a 1 gram particle oscillating with frequency 3 cycles per second and amplitude 1 centimeter, n is roughly 10^{28}. Thus, conclude that the relative spacing between neighboring quantum energy levels is so small that

it could never be detected experimentally. [Use $\hbar = 1.054 \times 10^{-27}$ erg · sec.]

(b) Repeat this calculation for an electron ($m \simeq 10^{-27}$ gram) oscillating with a frequency on the order of visible light ($\nu \simeq 10^{15}$ cps) and amplitude on the order of an atomic diameter ($A \simeq 10^{-8}$ cm). Would quantum effects be noticeable in this case?

(c) For the system of part (a), show that the lowest energy level, corresponding to $n = 0$, is so small on the macroscopic scale that it could not be experimentally distinguished from the classical minimum of zero.

The results of the preceding exercise are in many respects typical of most potentials encountered in quantum mechanics: Owing to the smallness of $\hbar$, macroscopic energies necessarily correspond to large values of the quantum number n, and this in turn implies a relative spacing between the nearby levels which is so small that the discrete values *appear* to be continuous. Only when one enters the atomic or subatomic realms does the spacing between the levels become significant.

We consider next requirement (ii). We found in the preceding section that position and momentum are *not* compatible observables; in fact, according to Eq. (4-69),

$$\Delta\hat{X}_t \cdot \Delta\hat{P}_t \geqq \frac{\hbar}{2}$$

so it impossible for position and momentum to simultaneously have *exactly* defined values. However, let us get some idea of just how stringent this limitation is from a *macroscopic* point of view. Note first that if the particle has a momentum uncertainty of $\Delta\hat{P}_t$, then it will have a velocity uncertainty of $\Delta\hat{V}_t = \Delta\hat{P}_t/m$.† Therefore the uncertainties in position and velocity must always satisfy

$$\Delta\hat{X}_t \cdot \Delta\hat{V}_t \geqq \frac{\hbar}{2m} \tag{4-70}$$

Exercise 61.

(a) Show that a 1-gram particle can have its position defined to within 0.001 micron and its velocity defined to within 0.001 micron per century, and yet the uncertainty relation in Eq. (4-70) would *not* be violated. [*Note*: 1 micron = 10^{-4} cm, and 1 year $\cong 3 \times 10^{7}$ sec.]

†Since $v = p/m$ in classical mechanics, then Postulate 6 implies that the quantum velocity operator is just $\hat{V} = \frac{1}{m}\hat{P}$.

(b) For an electron (mass $\simeq 10^{-27}$ gram) confined to an interval on the order of an atomic diameter (distance $\simeq 10^{-8}$ cm), what is the *minimum* value for the uncertainty in the velocity? Would it make sense in this case to speak of the velocity of the electron as "having a value?" [Note for comparison that 1 cm/sec is roughly equal to 2×10^{-2} mile per hour.]

We can see from part (a) of this exercise that it is quite possible, at least from the standpoint of the Uncertainty Principle alone, for a *macroscopic* particle to have its position and its momentum *simultaneously* defined with great precision. Again, we see that the reason for this is the extreme smallness of the constant $\hbar$: Since $\hbar$ is practically zero on the macroscopic scale, then Eq. (4-68) implies that $\hat{X}$ and $\hat{P}$ approximately commute, so that position and momentum are approximately compatible. It is thus possible for the state of a macroscopic particle to be such that the "widths" of the position and momentum distribution curves, $\Delta\hat{X}_t$ and $\Delta\hat{P}_t$, are *simultaneously so small* in comparison with macroscopic values that there is a nil probability of measuring for x and p values which differ *significantly* from the "peak" values, $\langle\hat{X}\rangle_t$ and $\langle\hat{P}\rangle_t$. In such a case, we would be quite justified in saying that the particle at time t is at the point $x(t) \equiv \langle X\rangle_t$ and is moving with momentum $p(t) \equiv \langle P\rangle_t$. This, of course, is just requirement (ii).

The preceding observations have not "proved" that quantum mechanics satisfies requirements (i) and (ii), but they do illustrate fairly well what the general situation is: The observable operators which seem to be required to account for the experimentally observed behavior of real microscopic systems, are found to be such that the universal constant $\hbar$ controls both the spacing between adjacent eigenvalues,

$$A_{n+1} - A_n \simeq \hbar$$

as well as the amount by which two incompatible observables "miss" being compatible,

$$\hat{A}\hat{B} - \hat{B}\hat{A} \simeq \hbar, \qquad \Delta\hat{A} \cdot \Delta\hat{B} \gtrsim \hbar$$

Thus it may be said that the nonclassical features of quantum mechanics owe their existance to the fact that $\hbar$ is *finite* and not zero.† But since from a *macroscopic* point of view $\hbar$ *appears to be*

†The reader who has had some contact with Special Relativity will notice an intriguing analogy here: Relativistic effects are pronounced only when one deals with velocities that are on the order of the velocity of light, c; thus it may be said that the unique features of relativity owe their existence to the fact that c is *finite* and not infinite.

zero, then here all eigenvalues *appear to be continuous* and all observables *appear to be compatible*—and these are just the requirements (i) and (ii).

We turn finally to consider requirement (iii). We have previously shown that $\langle \hat{H} \rangle_t$ is *always* constant in time [see Exercise 43], so it is obviously true that the expectation value of the *energy* satisfies its classical equation of motion [see Exercise 22]. Let us examine the time-dependence of the expectation values of *position* and *momentum*.

The general equation of motion for the expectation value of any observable was derived in Sec. 4-4b and is given by Eq. (4-34). From this equation, it is clear that

$$\frac{d}{dt} \langle \hat{X} \rangle_t = \frac{i}{\hbar} (\Psi_t, [\hat{H}\hat{X} - \hat{X}\hat{H}] \Psi_t) \tag{4-71a}$$

and

$$\frac{d}{dt} \langle \hat{P} \rangle_t = \frac{i}{\hbar} (\Psi_t, [\hat{H}\hat{P} - \hat{P}\hat{H}] \Psi_t) \tag{4-71b}$$

Now the right-hand sides of these equations can be evaluated explicitly by making use of the expression for $\hat{H}$ in Eq. (4-49a):

Exercise 62. Prove that

$$\hat{H}\hat{X} - \hat{X}\hat{H} = -i\hbar \frac{1}{m} \hat{P} \tag{4-72a}$$

and

$$\hat{H}\hat{P} - \hat{P}\hat{H} = -i\hbar F(\hat{X}) \tag{4-72b}$$

where, in the last equation, the function $F(x)$ is defined by

$$F(x) \equiv -\frac{d}{dx} V(x)$$

so that, by Eqs. (3-2) and (4-48), $F(\hat{X})$ is the *force* operator. [*Hint:* First show that $\hat{X}$ commutes with $V(\hat{X})$ while $\hat{P}$ commutes with $\hat{P}^2/2m$, so that

$$\hat{H}\hat{X} - \hat{X}\hat{H} = \frac{1}{2m} (\hat{P}^2 \hat{X} - \hat{X}\hat{P}^2)$$

and

$$\hat{H}\hat{P} - \hat{P}\hat{H} = V(\hat{X})\hat{P} - \hat{P}V(\hat{X})$$

The right-hand side of the first relation is most easily reduced by applying the operator equation (4-68); the right-hand side of the second relation may be reduced by inserting $V(\hat{X}) = V(x)$

and $\hat{P} = -i\hbar(d/dx)$, and then calculating the effect of the resulting operator on some arbitrary function $\phi(x)$.]

When Eqs. (4-72) are inserted into Eqs. (4-71), we immediately obtain the results

$$\frac{d}{dt}\langle\hat{X}\rangle_t = \frac{1}{m}\langle\hat{P}\rangle_t \tag{4-73a}$$

$$\frac{d}{dt}\langle\hat{P}\rangle_t = \langle F(\hat{X})\rangle_t \tag{4-73b}$$

Equations (4-73) are known as *Ehrenfest's equations*. It must be emphasized that these equations are quite general and involve *no* approximations. The first Ehrenfest equation says that $\langle\hat{P}\rangle_t$ and $d\langle\hat{X}\rangle_t/dt$ are related in precisely the same way as the classical momentum p and velocity dx/dt [see Eq. (3-3b)]. The second equation, however, is a bit subtle; by means of the first equation, we can write it as

$$\frac{d}{dt}\left(m\frac{d}{dt}\langle\hat{X}\rangle_t\right) = \langle F(\hat{X})\rangle_t$$

or

$$\frac{d^2}{dt^2}\langle\hat{X}\rangle_t = \frac{1}{m}\langle F(\hat{X})\rangle_t \tag{4-74}$$

Now this equation is *almost* identical to Newton's second law, Eq. (3-3a):

$$\frac{d^2x}{dt^2} = \frac{1}{m}F(x) \tag{3-3a}$$

It would be *exactly* identical to Newton's second law *if and only if*

$$\langle F(\hat{X})\rangle_t = F(\langle\hat{X}\rangle_t) \tag{4-75a}$$

for in this case Eq. (4-74) would read

$$\frac{d^2}{dt^2}\langle\hat{X}\rangle_t = \frac{1}{m}F(\langle\hat{X}\rangle_t) \tag{4-75b}$$

which is the same as Newton's second law provided we identify $\langle\hat{X}\rangle_t$ with $x(t)$. In other words, *if* Eq. (4-75a) holds, then $\langle\hat{X}\rangle_t$ will evolve with time in exactly the same way as the position function $x(t)$ does in classical mechanics. Since we have also shown that $\langle\hat{P}\rangle_t$ is related to $\langle\hat{X}\rangle_t$ in the same way that $p(t)$ is related to $x(t)$, then we could conclude that quantum mechanics "corresponds to"

classical mechanics in the sense that the time evolutions of $\langle \hat{X} \rangle_t$ and $\langle P \rangle_t$ coincide with the time evolutions of $x(t)$ and $p(t)$, respectively. We repeat, though, that this conclusion hinges on the validity of Eq. (4-75a).

Exercise 63. Show that Eq. (4-75a) *does* hold for the special cases $F(x) = 0$, $F(x) = k$, $F(x) = kx$, where k is any real number; thus conclude that, for these three cases, $\langle \hat{X} \rangle_t$ and $\langle \hat{P} \rangle_t$ obey the usual classical equations of motion. On the other hand, show that Eq. (4-75a) is *not* an identity for the case $F(x) = x^2$.

Now, except for the three cases noted in the above exercise, Eq. (4-75a) is *not* generally valid. The essential reason for this is the same as the reason why $\langle \hat{X}^2 \rangle_t$ is not generally equal to $\langle \hat{X} \rangle_t^2$—namely, the "width" of the position distribution curve

$$\Delta \hat{X}_t = \sqrt{\langle \hat{X}^2 \rangle_t - \langle \hat{X} \rangle_t^2}$$

is *not* always zero [see Fig. 4]. Thus, for state vectors which have a significant dispersion in position, Eq. (4-75b) is not valid for arbitrary force fields $F(x)$, and we must be content with Eq. (4-74).

Suppose, however, we pass to the *macroscopic limit*; here, according to requirement (ii), the width of the position distribution curve will be so small that a series of repeated measurements of the position will yield x-values which are all virtually indistinguishable from the value $\langle \hat{X} \rangle_t$. In such a case, it clearly makes no difference whether we evaluate the average of the $F(x)$-values—i.e., calculate $\langle F(\hat{X}) \rangle_t$—or evaluate F for the average of the x-values—i.e., calculate $F(\langle \hat{X} \rangle_t)$—since we will obtain the same result either way. Therefore, in the *macroscopic limit* Eqs. (4-75) *do* hold for arbitrary force functions $F(x)$, so that, by the foregoing arguments, requirement (iii) is indeed satisfied for the position and momentum observables.

The general relationship which exists between classical mechanics and quantum mechanics is usually referred to as the *Correspondence Principle*. Like so many other things in quantum mechanics, the Correspondence Principle is a very deep and many faceted subject, and we certainly have not exhausted it here. Much of the original thinking on this matter was done by Niels Bohr.

The fundamental connection between the time evolutions of the classical and quantum states was evidently established via the Ehrenfest equations. The fact that these equations fell out of our formalism so simply is certainly a satisfying result, and one which the reader may have found rather surprising. However, it should be noted that we really built this result into quantum mechanics in Postulate 6, when we in effect *used classical mechanics* to tell us what to write down for the Hamiltonian operator $\hat{H}$. In a sense, then,

we have *not* completely relegated classical mechanics to the status of a mere "special case" of quantum mechanics. For this reason, it is not altogether clear just what the precise logical relationship is between classical and quantum mechanics. We shall terminate our discussion of this matter with the broad observation that *we ourselves*, as the ultimate observers of any physical system, are essentially classical objects, in that our senses can directly perceive only *macroscopic phenomena* (e.g., dial readings, instrument settings, etc.). This fact probably places severe restrictions not only on what things we can perceive about a *microscopic* system, but also on how we interpret what we perceive. As a result, it is highly questionable if we shall ever be able to discard completely the "crutch" of classical mechanics on the microscopic level.

4-5d A Problem

For a given physical system, it is always important to find the eigenvectors $\{\eta_n(x)\}$ and eigenvalues $\{E_n\}$ of the Hamiltonian operator $\hat{H}$. There are two reasons for this. First, the numbers E_1, E_2, ... provide us with the *allowed energy levels* of the system, and these are always of great practical use in describing the system. Secondly, a knowledge of the functions $\{\eta_n(x)\}$ and the numbers $\{E_n\}$ permit us to write down at once the *solution to the time-evolution problem*; thus, from the given initial state vector $\Psi_0(x)$, we merely calculate the set of complex numbers

$$(\eta_n, \Psi_0) \equiv \int_{-\infty}^{\infty} \eta_n^*(x)\Psi_0(x)dx \qquad n = 1, 2, \ldots \tag{4-76a}$$

which determines the expansion of $\Psi_0(x)$ in the energy eigenbasis,

$$\Psi_0(x) = \sum_{n=1}^{\infty} (\eta_n, \Psi_0)\eta_n(x) \tag{4-76b}$$

and we then have at once an expression for the state vector $\Psi_t(x)$ at any time $t > 0$ [see Eq. (4-42)]:

$$\Psi_t(x) = \sum_{n=1}^{\infty} (\eta_n, \Psi_0)e^{-iE_n t/\hbar}\,\eta_n(x) \tag{4-76c}$$

For a quantum system which has a classical analogue, it is clear that the finding of the energy eigenvectors and eigenvalues is equivalent to solving the time-independent Schrödinger equation for the

system's potential function $V(x)$ [see Eq. (4-50)]:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\eta_n(x) + V(x)\eta_n(x) = E_n\eta_n(x) \tag{4-77}$$

It would therefore seem a logical next step in our discussion of quantum mechanics to undertake a detailed examination of the general properties of this equation, and to obtain the specific solutions for various "physically important" potential functions. However, this usually turns out to be a complicated and laborious enterprise: Exact solutions to the Schrödinger equation can be found for only a few simple forms for $V(x)$, and even these usually require considerable mathematical expertise. For more complicated potentials, it is necessary to resort to various *approximation techniques*, of which there are a wide variety with varying conditions of applicability. Some of these approximation techniques are quite involved, and almost constitute separate disciplines in their own right.

For these reasons, we shall not delve into the broad problem of solving, either exactly or approximately, the Schrödinger equation for various physical systems. The serious student of physics will find this problem treated exhaustively in existing textbooks on quantum mechanics; indeed, the solution of the Schrödinger equation is usually the major concern of the standard textbooks. In a sense, the aim of this book has not been to *solve* the Schrödinger equation, but rather to place it in the context of a broad (if somewhat simplified) theoretical framework. It is hoped that this will provide the student with an over-all *perspective* of quantum mechanics before he becomes immersed in the complexities of its many applications.

Nevertheless, we cannot in good conscience refrain from working out at least *one* "quantum mechanics problem"—if only to demonstrate that the rather abstract formalism which we have developed in these pages can indeed by applied to a concrete situation. To this end, we shall consider the relatively simple system of a particle of mass m in a one-dimensional "infinite square well," which is the name given by physicists to the potential field defined by

$$V(x) = \begin{cases} 0 & \text{for } |x| \leq L/2 \\ \infty & \text{for } |x| > L/2 \end{cases} \tag{4-78}$$

From the standpoint of classical mechanics, the motion of a particle in this potential field can be understood as follows: Inside the well, $-L/2 < x < L/2$, the particle experiences no force since $F \equiv -dV/dx \equiv 0$, so it moves with a constant momentum, p_0. When the particle strikes the "impenetrable wall" at $x = +L/2$, it experiences an infinite force in the negative x-direction,

$$F(+L/2) = -\left.\frac{dV}{dx}\right|_{L/2} = -\infty$$

which, however, acts for only an infinitesimal time interval; the effect of this impulse is simply to turn the particle around so that its momentum becomes $-p_0$, for this is the only way for the energy of the particle to be conserved. Similarly, when the particle strikes the back wall at a time $L/v_0 = Lm/p_0$ later, its direction of motion is again reversed. Clearly, if the initial state $[x_0, p_0]$ is specified, then it is possible to find the state $[x(t), p(t)]$ at any later time t. Essentially, though, the particle just bounces back and forth between the walls of the well with a period of $2Lm/p_0$, its momentum assuming only the two values $+p_0$ and $-p_0$. The energy of the particle is just

$$E = p_0^2/2m \tag{4-79}$$

and can clearly assume *any* value greater than or equal to zero, depending entirely upon the initial value p_0.

To treat this problem from the standpoint of quantum mechanics, we must first solve the time-independent Schrödinger equation (4-77) for the potential function given in Eq. (4-78). Owing to the way in which this potential function is defined, the Schrödinger equation takes different forms for the two regions $|x| > L/2$ and $|x| \leqq L/2$:

$$|x| > \frac{L}{2}: \quad -\frac{\hbar^2}{2m}\eta_n''(x) + \infty \cdot \eta_n(x) = E_n \eta_n(x) \tag{4-80a}$$

$$|x| \leqq \frac{L}{2}: \quad -\frac{\hbar^2}{2m}\eta_n''(x) = E_n \eta_n(x) \tag{4-80b}$$

The function $\eta_n(x)$ will thus have two pieces, one for "outside" the well, which satisfies Eq. (4-80a), and one for "inside" the well, which satisfies Eq. (4-80b). Clearly, the only solutions to the outside equation are

$$\eta_n(x) \equiv 0 \quad \text{for} \quad |x| > \frac{L}{2} \tag{4-81}$$

Therefore we need consider only the "inside" equation, (4-80b). Now, it can be shown that the infinite jump-discontinuities in $V(x)$ at $x = -L/2$ and $x = +L/2$ render the solutions to the Schrödinger equation *nondifferentiable* but *still continuous* at these two points. Thus, the inside and outside pieces of $\eta_n(x)$ must join at $x = \pm L/2$, but will generally do so with a "kink." In view of Eq. (4-81), the re-

quirement that the inside solutions must connect with the outside solutions implies that these inside solutions must satisfy the *boundary conditions*

$$\eta_n(-L/2) = \eta_n(+L/2) = 0 \tag{4-82}$$

Our problem, then, is to find those functions $\eta_n(x)$ and numbers E_n which satisfy the differential equation (4-80b) and the boundary conditions (4-82). In accordance with Eq. (4-2a), we may expect the functions $\{\eta_n(x)\}$ to form an orthonormal set:

$$(\eta_m, \eta_n) \equiv \int_{-L/2}^{L/2} \eta_m^*(x)\eta_n(x)dx = 0, \qquad m \neq n \tag{4-83a}$$

$$(\eta_n, \eta_n) \equiv \int_{-L/2}^{L/2} |\eta_n(x)|^2 \, dx = 1 \tag{4-83b}$$

Exercise 64.

(a) Show that the two functions $A_n \cos k_n x$ and $B_n \sin k_n x$ satisfy Eq. (4-80b), provided $k_n \equiv \sqrt{2mE_n}/\hbar$.

(b) Show that the boundary conditions (4-82) permit solutions of the form $A_n \cos k_n x$ only if $k_n L/2 = n\pi/2$, where $n = 1, 3, 5, \ldots$. Show that solutions of the form $B_n \sin k_n x$ are admissible only if $k_n L/2 = n\pi/2$, where $n = 2, 4, 6, \ldots$.

(c) On the basis of parts (a) and (b), show that the energy eigenvalues for the infinite square well are

$$E_n = \frac{\pi^2 \hbar^2}{2mL^2} n^2 \qquad n = 1, 2, 3, \ldots \tag{4-84}$$

and the corresponding energy eigenvectors are

$$\eta_n(x) = \begin{cases} \sqrt{\dfrac{2}{L}} \cos \dfrac{n\pi}{L} x & n = 1,3,5, \ldots \quad (4\text{-}85a) \\ \sqrt{\dfrac{2}{L}} \sin \dfrac{n\pi}{L} x & n = 2,4,6, \ldots \quad (4\text{-}85b) \end{cases}$$

where the constants A_n and B_n have been chosen in such a way that Eq. (4-83b) is satisfied. Verify that the functions in Eqs. (4-85) satisfy Eqs. (4-83a). *Hint*: For checking the orthonormality conditions, the following identities will be helpful:

$$\int_{-a/2}^{a/2} \sin\frac{n\pi x}{a} \cdot \sin\frac{m\pi x}{a}\,dx = \int_{-a/2}^{a/2} \cos\frac{n\pi x}{a} \cdot \cos\frac{m\pi x}{a}\,dx$$

$$= \begin{cases} 0 & \text{for } m \neq n \\ a/2 & \text{for } m = n \end{cases}$$

$$\int_{-a/2}^{a/2} \sin\frac{n\pi x}{a} \cdot \cos\frac{m\pi x}{a}\,dx = 0$$

(d) Show that the relative spacing of the energy levels around the value E_n is

$$\frac{E_{n+1} - E_n}{E_n} = \frac{2n+1}{n^2} \simeq \frac{2}{n}$$

(e) For a one gram particle moving with velocity 1 cm/sec in a well of length 1 cm, we expect the classical formula for the energy to be correct. Calculate the quantum number n for this case. Would the nearby relative spacings make the energy levels appear discrete or continuous? Would the minimum energy level E_1 be significantly different from zero on this macroscopic scale?

(f) An electron inside a medium weight atom can have energies on the order of 1 kev = 1.6×10^{-9} erg. Calculate the quantum number n for an electron ($m \simeq 10^{-27}$ gram) with an approximate energy of 1 kev, inside a 1 Angstrom well ($L = 10^{-8}$ cm). Would the nearby relative spacings make the energy levels appear discrete or continuous? Would the minimum energy level E_1 be significantly different from zero on this microscopic scale?

(g) Of the three observables, position, momentum, and total energy, if the particle can be said to "have a value" for one, can it be said to "have a value" for either of the other two? [*Hint*: Contrast the forms of the eigenbases $\{\delta_{x_0}(x)\}$, $\{\theta_{p_0}(x)\}$ and $\{\eta_n(x)\}$.]

(h) Show that the uncertainty in the momentum of any mass particle in an infinite square well can never be smaller than $\hbar/2L$.

Exercise 65. Suppose a measurement of the energy at time $t = 0$ yields the number E_n.

(a) Write down the expression for $\Psi_t(x)$ for $t > 0$. What is the position probability density function at any time t? Sketch rough graphs of the position probability density function

for the three cases $n = 1$, 2 and 3, and notice that, for a given value of n, there are places in the well where the particle can never be found.

(b) Calculate $\langle \hat{X} \rangle_t$ and $\langle \hat{P} \rangle_t$. Discuss your answers from the standpoint of the concept of the "stationary state." Describe the shape of the energy distribution curve at time t.

(c) Show that when the particle "has energy E_n" then the uncertainties in position and momentum are

$$\Delta \hat{X}_t = \frac{L}{2\sqrt{3}} \sqrt{1 - \frac{6}{\pi^2 n^2}} \quad \text{and} \quad \Delta \hat{P}_t = \frac{\hbar \pi n}{L} = \sqrt{2mE_n}$$

Is the position-momentum uncertainty relation satisfied? Is the time-energy uncertainty relation satisfied?

Exercise 66. Suppose the initial state vector is

$$\Psi_0(x) = \frac{\sqrt{3}}{2} \eta_1(x) + \frac{1}{2} \eta_2(x)$$

(a) What is $\Psi_t(x)$? If the energy of the system is measured at time t, what values can be obtained, and what are their respective probabilities? Using Eqs. (4-11) and (4-12), show that

$$\langle \hat{H} \rangle_t = \left[\frac{\pi^2 \hbar^2}{2mL^2} \right] \frac{7}{4} \quad \text{and} \quad \Delta \hat{H}_t = \left[\frac{\pi^2 \hbar^2}{2mL^2} \right] \frac{3\sqrt{3}}{4}$$

Prove that the evolution time of any observable cannot be less than $(4mL^2/3\sqrt{3}\pi^2\hbar)$.

(b) Show that the position probability density and the position probability current are given respectively by

$$|\Psi_t(x)|^2 = \frac{3}{4} \eta_1^2(x) + \frac{1}{4} \eta_2^2(x) + \frac{\sqrt{3}}{2} \eta_1(x) \eta_2(x) \cos \left[\frac{E_2 - E_1}{\hbar} t \right]$$

$$S(x,t) = \frac{\hbar}{m} \frac{\sqrt{3}}{4} \left(\eta_1'(x) \eta_2(x) - \eta_1(x) \eta_2'(x) \right) \sin \left[\frac{E_2 - E_1}{\hbar} \right] t$$

Show that the period of oscillation of the density and current are on the order of the evolution time minimum as estimated in part (a). Evaluate the position probability density and current at the points $x = 0$, $x = L/4$, $x = L/3$, $x = L/2$.

(c) Show that the expectation value of the position at time t is

$$\langle \hat{X} \rangle_t = \frac{\sqrt{3}}{2} \left\{ \int_{-L/2}^{L/2} x \eta_1(x) \eta_2(x) dx \right\} \cos \left[\frac{E_2 - E_1}{\hbar} t \right]$$

$$= \left(\frac{16\sqrt{3}}{9\pi^2} \right) \frac{L}{2} \cos \left[\frac{3\pi^2 \hbar}{2mL^2} \right] t$$

Notice that $\langle \hat{X} \rangle_t$ in this case is *sinusoidal* in time with amplitude $\cong 0.3(L/2)$; by contrast, the classical position function $x(t)$ has a *sawtooth* shape when plotted against t, with amplitude $L/2$.

(d) From either Eq. (4-58a) or (more easily) Eq. (4-73a), show that the expectation value of the momentum at time t is

$$\langle \hat{P} \rangle_t = \frac{-4}{\sqrt{3}} \frac{\hbar}{L} \sin \left[\frac{3\pi^2 \hbar}{2mL^2} t \right] = \left(\frac{-8}{\sqrt{21}\pi} \right) \sqrt{2m \langle \hat{H} \rangle} \sin \left[\frac{3\pi^2 \hbar}{2mL^2} t \right]$$

where, in the last step, we have made use of the expression for $\langle H \rangle$ in part (a). Notice that $\langle \hat{P} \rangle_t$ in this case is *sinusoidal* in time with amplitude $\cong 0.56\sqrt{2m\langle \hat{H} \rangle}$; by contrast, the classical momentum function $p(t)$ has a *square-wave* shape when plotted against t, with amplitude $\sqrt{2mE}$.

4-6 EXTENSIONS OF THE THEORY

At various stages in our presentation of the theory of quantum mechanics we imposed certain simplifying restrictions so that the main ideas would not be obscured by considerations of slightly lesser importance. It seems appropriate to conclude our development by discussing very briefly how the *removal* of some of these restrictions will affect our simplified picture of quantum mechanics. It must be emphasized that the following discussion is not meant to be as detailed as that in the previous sections. Our purpose now is merely to get a rough idea of what is involved in obtaining a more general theory, and to thereby establish a bridge to the reader's future studies in quantum mechanics.

4-6a Systems with More Than One Degree of Freedom

The first major restriction which we imposed was that the physical system have only one degree of freedom. We labeled this degree of freedom by the variable or "coordinate" x. Now, most systems of interest in physics have *more* than one degree of freedom. For example, a particle in real space will have three degrees of freedom corresponding to its position (i.e., x,y,z or r,θ,ϕ), and it may also have additional degrees of freedom due to its orientation or to some internal structure; again, a system consisting of two particles with no other attributes will generally have six degrees of freedom, which can be labelled by the six coordinates $x_1, y_1, z_1, x_2, y_2, z_2$.

It is not difficult, at least from a formal standpoint, to adapt our treatment of a system with one degree of freedom to a system with n degrees of freedom. To do this, we first associate with each degree of freedom a so-called "generalized coordinate" q_i. The set of generalized coordinates $q_1, q_2, \ldots, q_n$ may consist of cartesian coordinates, angles, and in general any group of variables which, when taken together, specify the "configuration" of the system in the same way that x specifies the configuration of a particle in one dimension. Having done this, we then set up a Hilbert space for the system. The *vectors* in this Hilbert space are all those complex functions ψ of n real variables $q_1, q_2, \ldots, q_n$ which satisfy [see Eq. (2-35)]

$$\int\!\!-\!\!\int |\psi(q_1, \ldots, q_n)|^2 \, dq_1 \cdots dq_n < \infty \tag{4-86}$$

Here, the integrations are to be carried out over the full ranges of the various coordinates q_i. From this point on, the development of the properties of the Hilbert space is *identical* to that given in Secs. 2-3 and 2-4, except that the single variable x is everywhere replaced by the n variables $q_1, \ldots, q_n$. Thus, for example, the definition of the inner product in Eq. (2-32) becomes

$$(\psi_1,\psi_2) \equiv \int\!\!-\!\!\int \psi_1^*(q_1, \ldots, q_n)\psi_2(q_1, \ldots, q_n)dq_1 \cdots dq_n \tag{4-87}$$

The possible state vectors of the system are, as before, all the normed vectors in $\mathcal{H}$; we write the state vector at time t as $\Psi_t(q_1, \ldots, q_n)$ or $\Psi(q_1, \ldots, q_n; t)$.

Now, in advanced treatments of classical mechanics, there is developed a well-defined procedure for associating with any "generalized coordinate" q_i a "generalized momentum" p_i. We shall not go

into this procedure here, but we might just mention by way of example that, if $q_i = x$ then p_i turns out to be $p_x = mv_x$, and if $q_i = \theta$ then p_i turns out to be the angular momentum associated with the rotation of θ. In any case, the two sets of variables $\{q_i\}$ and $\{p_i\}$ form the basic *observables* of both the classical description and the quantum description. In direct analogy with Eqs. (4-46) and (4-47), we postulate that the quantum operators corresponding to q_i and p_i are

$$\hat{Q}_i = q_i \qquad \text{and} \qquad \hat{P}_i = -i\hbar \frac{\partial}{\partial q_i} \tag{4-88}$$

and, moreover, that the operator corresponding to any observable $f(q_1, p_1, \ldots, q_n, p_n)$ is [see Eq. (4-48)]

$$f\left(q_1, -i\hbar \frac{\partial}{\partial q_1}, \ldots, q_n, -i\hbar \frac{\partial}{\partial q_n}\right) \tag{4-89}$$

As in Eqs. (4-68) and (4-69), the above definitions for $\hat{Q}_i$ and $\hat{P}_i$ imply that

$$\hat{Q}_i \hat{P}_i - \hat{P}_i \hat{Q}_i = i\hbar \tag{4-90}$$

so that, by the Heisenberg Uncertainty Principle,

$$\Delta \hat{Q}_i \cdot \Delta \hat{P}_i \geq \frac{\hbar}{2} \tag{4-91}$$

Although $\hat{Q}_i$ and $\hat{P}_i$ do *not* commute with each other, they *do* commute with all the rest of the $\hat{Q}_j$'s and $\hat{P}_j$'s. In particular, this implies by the Compatibility Theorem that the generalized coordinates are all *simultaneously measureable*. By the same sort of arguments used to derive Eq. (4-61), it can be shown that the quantity

$$|\Psi_t(q_1, \ldots, q_n)|^2 \, dq_1 \cdots dq_n \tag{4-92}$$

is the probability that a simultaneous measurement of all the coordinates at time t will yield a value for q_1 in the interval $(q_1, q_1 + dq_1)$, *and* a value for q_2 in the interval $(q_2, q_2 + dq_2)$, . . . , *and* a value for q_n in the interval $(q_n, q_n + dq_n)$.

For the important case of a simple particle in three dimensions, we can write the state vector as $\Psi_t(x,y,z) \equiv \Psi(x,y,z;t)$. Since the classical Hamiltonian function is now

$$H(x,y,z,p_x,p_y,p_z) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V(x,y,z)$$

then replacing the variable x by the operator x, the variable p_x by the operator $-i\hbar\,(\partial/\partial x)$, etc., we obtain in analogy with Eqs. (4-50)

and (4-51) the time-independent and the time-dependent Schrödinger equations:

$$-\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right]\eta_n(x,y,z)+V(x,y,z)\eta_n(x,y,z)=E_n\eta_n(x,y,z) \tag{4-93}$$

$$-\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right]\Psi(x,y,z;t)+V(x,y,z)\Psi(x,y,z;t)=i\hbar\frac{\partial}{\partial t}\Psi(x,y,z;t) \tag{4-94}$$

Thus, for example, if one solves Eq. (4-93) for the Coulomb potential, $V(x,y,z)=\frac{-e^2}{4\pi\epsilon_0}[x^2+y^2+z^2]^{-\frac{1}{2}}$, one obtains for the eigenvalues $\{E_n\}$ just the "Bohr energy levels" for the hydrogen atom. Of course, one *also* obtains, through the associated eigenvectors $\{\eta_n(x)\}$, *much more* information about the hydrogen atom than was available via the old Bohr theory.

When we extend our treatment of a particle from one dimension to three dimensions, we find that another important observable comes into play: In classical mechanics, a particle at the position **r** with momentum **p** has an "angular momentum about the origin"—or, as we shall prefer to say, an *orbital angular momentum*—given by $\boldsymbol{\ell}=\mathbf{r}\times\mathbf{p}$. In component form, this means that

$$\ell_x=yp_z-zp_y,\quad \ell_y=zp_x-xp_z,\quad \ell_z=xp_y-yp_x \tag{4-95}$$

Therefore, according to the generalized form of Postulate 6, the quantum *operator* for the observable ℓ_x is

$$\hat{L}_x=\hat{Y}\hat{P}_z-\hat{Z}\hat{P}_y=-i\hbar\left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right) \tag{4-96}$$

Analogous expressions follow for $\hat{L}_y$ and $\hat{L}_z$. Now from these definitions of $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$, along with the commutation relations $\hat{X}\hat{P}_x-\hat{P}_x\hat{X}=i\hbar$, $\hat{Y}\hat{P}_y-\hat{P}_y\hat{Y}=i\hbar$, and $\hat{Z}\hat{P}_z-\hat{P}_z\hat{Z}=i\hbar$, it is fairly straightforward to establish the following commutation relations among the three orbital angular momentum component operators:

$$\left.\begin{aligned}\hat{L}_x\hat{L}_y-\hat{L}_y\hat{L}_x&=i\hbar\hat{L}_z\\ \hat{L}_y\hat{L}_z-\hat{L}_z\hat{L}_y&=i\hbar\hat{L}_x\\ \hat{L}_z\hat{L}_x-\hat{L}_x\hat{L}_z&=i\hbar\hat{L}_y\end{aligned}\right\} \tag{4-97}$$

Exercise 67. Derive the first of Eqs. (4-97). Note that the other two can be obtained by cyclically permuting the indices.

One then generalizes and calls *any* three operators which satisfy the above commutation relations the "components of an *angular momentum operator.*" The general treatment of angular momentum in quantum mechanics is somewhat involved. However, a great many important physical phenomena find their interpretations in terms of the various properties of angular momentum operators.

4-6b Some Remarks on Continuous Eigenvalues

Our development of the theory of quantum mechanics in the first four sections of this chapter was restricted to observables with *discretely* distributed eigenvalues. However, in Sec. 4-5a we had to abandon this restriction to some extent, since the position and momentum operators of Postulate 6 were found to have *continuously* distributed eigenvalues. In connection with this, we encountered some unusual mathematical difficulties associated with the eigenbasis vectors $\{\delta_{x_0}(x)\}$ and $\{\theta_{p_0}(x)\}$ of these two operators [see the discussion following Exercise 50]. Similar difficulties arise in the *general* treatment of operators with continuously distributed eigenvalues. The handling of these difficulties is somewhat involved, but pivots mainly upon the so-called *Dirac delta function*, $\delta(x - x_0)$, a highly unusual entity which just happens to coincide with the position eigenvector $\delta_{x_0}(x)$.

We remarked in our discussion of $\delta_{x_0}(x)$ in Sec. 4-5a that this function should vanish for $x \neq x_0$, but that $\delta_{x_0}(x_0)$ should be *infinite* in such a way that its product with the infinitesimal dx be *finite*. In fact, the actual definition of the Dirac delta function is such that

$$\delta_{x_0}(x_0)dx \equiv \delta(0)dx = 1$$

More specifically, the Dirac delta function $\delta(x - x_0)$ is *defined* by the following two relations:

$$\begin{cases} \delta(x - x_0) = 0 \quad \text{for } x \neq x_0 & \text{(4-98a)} \\ \displaystyle\int_a^b \delta(x - x_0)dx = 1 \quad \text{for any } a < x_0 < b & \text{(4-98b)} \end{cases}$$

Exercise 68. Using these defining properties of the Dirac delta function, derive the following two properties:

$$\delta(x - x_0) = \delta(x_0 - x) \tag{4-99}$$

$$f(x) = \int_{-\infty}^{\infty} f(x')\delta(x' - x)dx' \tag{4-100}$$

[*Hint:* To establish Eq. (4-100), multiply the arbitrary function value $f(x)$ by $1 = \int \delta(x' - x)dx'$, move $f(x)$ inside the x'-integral, and then given an argument for replacing the $f(x)$ by $f(x')$.]

We recall that, for an observable operator $\hat{A}$ whose eigenvalues and eigenvectors can be labeled by a *discrete* index n,

$$\hat{A}\alpha_n(x) = A_n\alpha_n(x) \tag{4-101a}$$

the eigenvectors $\{\alpha_n(x)\}$ must form an orthonormal basis in the Hilbert space:

$$(\alpha_n, \alpha_{n'}) = \delta_{nn'} \quad \text{for all } n,n' \tag{4-101b}$$

$$\phi(x) = \sum_n (\alpha_n, \phi)\alpha_n(x) \quad \text{for all } \phi(x) \text{ in } \mathcal{H} \tag{4-101c}$$

Now for the case in which the eigenvalues and eigenvectors of $\hat{A}$ must be labeled by a *continuous* index ν,

$$\hat{A}\alpha_\nu(x) = A(\nu)\alpha_\nu(x) \tag{4-102a}$$

it is necessary to modify somewhat our definition of an orthonormal basis set. Specifically, Eq. (4-101b) is modified by replacing the Kronecker delta symbol δ_{ab} by the Dirac delta function $\delta(a - b)$,

$$(\alpha_\nu, \alpha_{\nu'}) = \delta(\nu - \nu') \quad \text{for all } \nu,\nu' \tag{4-102b}$$

and Eq. (4-101c) is modified by replacing the sum over the discrete index n by an integral over the continuous index ν,

$$\phi(x) = \int (\alpha_\nu, \phi)\alpha_\nu(x)d\nu \quad \text{for all } \phi(x) \text{ in } \mathcal{H} \tag{4-102c}$$

Finally, the rule in Postulate 3, that $|(\alpha_n, \Psi_t)|^2$ gives the probability of measuring the discrete value A_n in the state $\Psi_t(x)$, takes the following form for the case of continuous eigenvalues: *If* $\mathcal{A}$ *is measured on the state* $\Psi_t(x)$, *the probability of obtaining a value between* $A(\nu)$ *and* $A(\nu + d\nu)$ *is* $|(\alpha_\nu, \Psi_t)|^2\, d\nu$.

The mathematical transition from Eq. (4-101c) to Eq. (4-102c) is a rather natural one; however, the transition from Eq. (4-101b) to Eq. (4-102b) is not altogether obvious at first sight. In order to understand why we write $\delta(\nu - \nu')$ in Eq. (4-102b) rather than just $\delta_{\nu\nu'}$, let us make the following calculation: let us write out ex-

plicitly the definition of the quantity $(\alpha_{\nu'},\phi)$, and then insert for $\phi(x)$ the expansion in Eq. (4-102c). We have

$$(\alpha_{\nu'},\phi) \equiv \int_{-\infty}^{\infty} \alpha_{\nu'}^{*}(x)\phi(x)dx = \int_{-\infty}^{\infty} \alpha_{\nu'}^{*}(x) \left[\int (\alpha_{\nu},\phi)\alpha_{\nu}(x)d\nu \right] dx$$

Interchanging the order of the x- and ν-integrations, we obtain

$$(\alpha_{\nu'},\phi) = \int (\alpha_{\nu},\phi) \left[\int_{-\infty}^{\infty} \alpha_{\nu'}^{*}(x)\alpha_{\nu}(x)dx \right] d\nu \equiv \int (\alpha_{\nu},\phi)\,(\alpha_{\nu'},\alpha_{\nu})d\nu$$

Now the quantity (α_{ν},ϕ) is a complex number which depends on the index ν, or in other words, (α_{ν},ϕ) is some arbitrary complex function of the real variable ν; writing this function $c(\nu)$, the last equation becomes

$$c(\nu') = \int c(\nu)\,(\alpha_{\nu'},\alpha_{\nu})\,d\nu$$

Clearly, if $(\alpha_{\nu'},\alpha_{\nu}) = \delta_{\nu\nu'}$, then the integral on the right would vanish, thereby rendering the equation incorrect. Indeed, the *only* way for this equation to hold true for any arbitrary function $c(\nu)$ is for $(\alpha_{\nu'},\alpha_{\nu})$ to be the Dirac delta function $\delta(\nu' - \nu)$, in which case the equation is simply an instance of Eq. (4-100). We see then that, if we want Eq. (4-102c) to be valid for all functions $\phi(x)$, we are essentially *forced* to require the functions $\{\alpha_{\nu}(x)\}$ to satisfy Eq. (4-102b). As an example, the position and momentum eigenvectors are supposed to satisfy

$$(\delta_{x_1},\delta_{x_2}) = \delta(x_1 - x_2) \qquad -\infty < x_1,x_2 < \infty \qquad (4\text{-}103a)$$

and

$$(\theta_{p_1},\theta_{p_2}) = \delta(p_1 - p_2) \qquad -\infty < p_1,p_2 < \infty \qquad (4\text{-}103b)$$

Exercise 69. Using the definition of the inner product, along with Eq. (4-100), prove that the position eigenvectors do indeed satisfy the modified orthonormality relation in Eq. (4-103a). [*Hint:* Remember that $\delta_{x_0}(x) \equiv \delta(x - x_0)$ is pure real.]

We cannot discuss here all the ramifications of these generalizations for continuous eigenvalues. In particular, Eq. (4-103b), and the precise relationship between the functions $\delta_{x_0}(x)$ and $\theta_{p_0}(x)$, can be fully appreciated only in the context of an area of mathematics known as Fourier analysis. However, there is one important consequence of all this that perhaps should be brought out. It will be re-

called that any $\mathcal{H}$-vector $\phi(x)$ may be said to have "components"

$$(\alpha_n,\phi) \qquad n = 1,2,\ldots \tag{4-104a}$$

relative to a given or orthonormal basis set $\{\alpha_n(x)\}$, just as any $\mathcal{E}_3$-vector v has components

$$\mathbf{e}_n \cdot \mathbf{v} \qquad n = 1, 2, \text{ and } 3 \tag{4-104b}$$

relative to a given orthonormal basis set $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ [see Eqs. (2-29c) and (2-39c)]. Thus any vector in $\mathcal{H}$ can be "represented" relative to a given orthonormal basis by an ∞-*tuple of complex numbers*, in the same sense that any vector in $\mathcal{E}_3$ can be represented relative to a given orthonormal basis by a triplet of real numbers. Evidently, the components of $\phi(x)$ relative to the continuous eigenbasis $\{\alpha_\nu(x)\}$ will be labelled by the continuous index ν, (α_ν,ϕ). In particular, let us calculate the components of $\phi(x)$ relative to the *position* eigenbasis, $\{\delta_\nu(x)\}$.

Exercise 70. Using the definition of the inner product, along with Eq. (4-100), prove that

$$(\delta_\nu,\phi) = \phi(\nu) \qquad -\infty < \nu < \infty \tag{4-105}$$

Therefore, relative to the eigenbasis of the *position* operator $\hat{X}$, a given *vector* $\phi(x)$—i.e., a given *function* ϕ of x—has *components* which are just the set of all the *values* of this function $c_\nu \equiv \phi(\nu)$. In this sense we may say that any vector $\phi(x)$ in $\mathcal{H}$ "represents itself" with respect to the *position* eigenbasis. Indeed, the expansion of $\phi(x)$ in the position eigenbasis $\{\delta_\nu(x)\}$ is, according to Eqs. (4-102c) and (4-105),

$$\phi(x) = \int_{-\infty}^{\infty} (\delta_\nu,\phi)\delta_\nu(x)d\nu = \int_{-\infty}^{\infty} \phi(\nu)\delta(x-\nu)d\nu \tag{4-106}$$

which is evidently nothing more than Eq. (4-100).

Equation (4-105) provides us with an interesting interpretation of the definition of the inner product, Eq. (2-32), as the following exercise demonstrates.

Exercise 71. In Exercise 13 we derived Eq. (2-40a), which gives the inner product of two $\mathcal{H}$-vectors in terms of their *components* relative to a given orthonormal basis $\{\epsilon_i(x)\}$. If we take for $\{\epsilon_i(x)\}$ the position eigenbasis $\{\delta_\nu(x)\}$, and if we replace the sum over the discrete index n by an integral over the continuous index ν, show that Eq. (2-40a) takes the form of our original *definition* of the inner product in Eq. (2-32). [*Hint*: In Eq. (2-40a), recall that $c_i = (\epsilon_i, \psi)$ and $d_i = (\epsilon_i,\phi)$.]

Finally, we should note that, in accordance with the previously mentioned interpretation of $|(\alpha_\nu,\Psi_t)|^2\,d\nu$ as being a measurement probability, the quantity $|(\delta_\nu,\Psi_t)|^2\,d\nu$ is evidently supposed to represent the probability of measuring for the *position* a value between ν and $\nu+d\nu$. But, by Eq. (4-105), we see that

$$|(\delta_\nu,\Psi_t)|^2\,d\nu = |\Psi_t(\nu)|^2\,d\nu$$

so we have recovered the "Born interpretation" of the state vector in Eq. (4-61).

As the reader may have sensed, our definition and use of the symbol $\delta(x - x_0)$ plays "fast and loose" with the laws of calculus. A rigorous treatment of the Dirac delta function requires a rather lengthy sojourn into an area of mathematics known as Distribution Theory; our treatment here is illustrative of the more "intuitive" approach taken by most standard textbooks on quantum mechanics, and the reader is referred to any of these books for a more detailed analysis of the Dirac delta function.

4-6c The Problem of Degeneracy

We wish now to discuss the effects of removing the restriction in Eq. (4-9) that the eigenvalues of an observable operator be unequal or *nondegenerate*. To this end, let us examine the simple case in which a certain eigenvalue of an observable operator $\hat{A}$ is *doubly degenerate;* more specifically, we suppose that the eigenvectors $\alpha_1(x)$ and $\alpha_2(x)$ correspond to the same eigenvalue A_{12}, but that no other eigenvector of $\hat{A}$ has this eigenvalue:

$$A_1 = A_2 = A_{12}, \qquad \text{but} \quad A_i \neq A_{12} \text{ for } i \geqq 3 \tag{4-107}$$

This circumstance will necessitate modifications in both Postulates 3 and 4. As indicated in our discussion just preceding Eq. (4-9), the required modification of Postulate 3 is merely a straightforward application of the "addition rule" for probabilities in Eq. (2-3a):

> **Postulate 3′.** If the eigenvalues of $\hat{A}$ satisfy Eq. (4-107), and if $\mathfrak{A}$ is measured on the state $\Psi_t(x)$, then the probability for obtaining the eigenvalue A_{12} is $|(\alpha_1,\Psi_t)|^2 + |(\alpha_2,\Psi_t)|^2$.

Of somewhat more interest is the modified form of Postulate 4, which reads as follows:

> **Postulate 4′.** If the eigenvalues of $\hat{A}$ satisfy Eq. (4-107), and if a measurement of $\mathfrak{A}$ on the state $\Psi_t(x)$ yields the eigenvalue A_{12},

then the state vector of the system immediately after the measurement is given by

$$\Psi'_{A_{12}}(x) = \frac{(\alpha_1, \Psi_t)\alpha_1(x) + (\alpha_2, \Psi_t)\alpha_2(x)}{\sqrt{|(\alpha_1, \Psi_t)|^2 + |(\alpha_2, \Psi_t)|^2}} \tag{4-108}$$

This last equation may appear strange at first, but it actually has a very simple interpretation: We can write the state vector immediately *before* the measurement as [see Eq. (4-6a)]

$$\Psi_t(x) = \left\{(\alpha_1, \Psi_t)\alpha_1(x) + (\alpha_2, \Psi_t)\alpha_2(x)\right\} + \sum_{i=3}^{\infty} (\alpha_i, \Psi_t)\alpha_i(x)$$

In view of this, we see that Postulate 4′ merely asserts that a measurement of $\mathfrak{A}$ *with the result* A_{12} essentially "wipes out" that portion of $\Psi_t(x)$ corresponding to eigenvalues *other than* A_{12}, but "passes undistorted" the parts of the state vector which *belong* to the eigenvalue A_{12}. The denominator in the above expression for $\Psi'_{A_{12}}(x)$ is merely to make the vector properly normalized.

Exercise 72. Prove that the vector $\Psi'_{A_{12}}(x)$ has unit norm.

It should be noted that the form of Postulate 4 which we presented in Sec. 4-3b is just a special case of Postulate 4′; for if the eigenvalues of $\hat{A}$ are nondegenerate, then by Postulate 4′, a measurement of $\mathfrak{A}$ with the result A_1 will leave the system in the state

$$\Psi'_{A_1}(x) = \frac{(\alpha_1, \Psi_t)}{|(\alpha_1, \Psi_t)|}\alpha_1(x)$$

But since the complex number multiplying $\alpha_1(x)$ obviously has unit modulus, then $\Psi'_{A_1}(x)$ may be said to coincide with $\alpha_1(x)$ in the sense allowed by Postulate 1. This is just the statement of Postulate 4 in Sec. 4-3b.

In developing the various consequences of these more general forms of the two "measurement postulates," the following simple theorem plays a key role.

Exercise 73. Prove that, if $\alpha_1(x)$ and $\alpha_2(x)$ are eigenvectors of $\hat{A}$ belonging to the *same* eigenvalue, then any linear combination of these two vectors, $c_1\alpha_1(x) + c_2\alpha_2(x)$, is *also* an eigenvector of $\hat{A}$ belonging to this eigenvalue. [*Hint:* Using the linearity of $\hat{A}$, examine the effect of $\hat{A}$ acting on the linear combination.]

Since $\Psi'_{A_{12}}(x)$ in Postulate 4′ is evidently a linear combination of $\alpha_1(x)$ and $\alpha_2(x)$, then one consequence of the foregoing theorem is that the following statement is *generally valid*: A measurement of

$\mathfrak{A}$ forces the state vector of the system into *an* eigenvector of $\hat{A}$ belonging to the eigenvalue obtained in the measurement. If the eigenvalue measured is nondegenerate, then there is only one corresponding eigenvector, and it is with this eigenvector that the state vector will coincide after the measurement. However, if the eigenvalue measured is degenerate, then there are infinitely many corresponding eigenvectors (i.e., all possible linear combinations of the "standard" eigenvectors), and it is necessary to know the state vector of the system *before* the measurement in order to determine unambiguously the state vector *after* the measurement.

A second consequence of Postulates 3′ and 4′, and the theorem of Exercise 73, is that the Compatibility Theorem which we presented in Sec. 4-3c remains valid as stated. However, the possibility of degenerate eigenvalues confers upon the Compatibility Theorem a new importance, which we shall now attempt to explain:

Suppose again that the eigenvectors $\alpha_1(x)$ and $\alpha_2(x)$ of the observable operator $\hat{A}$ correspond to the same eigenvalue A_{12}. Then according to Exercise 73, the two vectors $\tilde{\alpha}_1(x)$ and $\tilde{\alpha}_2(x)$ defined by

$$\tilde{\alpha}_1(x) = c_1\alpha_1(x) + c_2\alpha_2(x)$$

and

$$\tilde{\alpha}_2(x) = c_3\alpha_1(x) + c_4\alpha_2(x)$$

where c_1, c_2, c_3 and c_4 are any complex numbers, are also eigenvectors of $\hat{A}$ belonging to the eigenvalue A_{12}; in addition, it is clear that these two eigenvectors are orthogonal to all the other eigenvectors $\alpha_3(x)$, $\alpha_4(x)$, ... Now, it is possible to choose the c_i-numbers in such a way that $\tilde{\alpha}_1(x)$ and $\tilde{\alpha}_2(x)$ have unit norms and are orthogonal to each other; in fact, there are infinitely many ways of doing this. We can see this most easily if we take as an analogy a three-dimensional vector space in which the basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are respectively identified with the eigenvectors $\alpha_1(x)$ and $\alpha_2(x)$, while the basis vector $\mathbf{e}_3$ represents all the other eigenvectors $\alpha_3(x)$, $\alpha_4(x)$, . . . , all of which are orthogonal to the first two vectors. In this analogy, $\mathbf{e}_1$ and $\mathbf{e}_2$ are eigenvectors belonging to the eigenvalue A_{12}; but, according to Exercise 73, any vector in the plane formed by $\mathbf{e}_1$ and $\mathbf{e}_2$ is also an eigenvector of $\hat{A}$ with this same eigenvalue. This being the case, we need not tie ourselves down to the orthonormal pair $\mathbf{e}_1$ and $\mathbf{e}_2$, but we may evidently use any pair $\tilde{\mathbf{e}}_1$ and $\tilde{\mathbf{e}}_2$ which differs from $\mathbf{e}_1$ and $\mathbf{e}_2$ by a simple rotation about the $\mathbf{e}_3$-axis; for such a pair would belong to the same eigenvalue of $\hat{A}$ as $\mathbf{e}_1$ and $\mathbf{e}_2$, and moreover, the vectors $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$ and $\mathbf{e}_3$ would clearly constitute an orthonormal basis set. In a similar way, we can always replace $\alpha_1(x)$ and $\alpha_2(x)$ with any one of an infinite number of orthonormal pairs

$\tilde{\alpha}_1(x)$ and $\tilde{\alpha}_2(x)$ in the "plane" of $\alpha_1(x)$ and $\alpha_2(x)$; from the point of view of the operator $\hat{A}$, it makes no difference at all whether we choose for its eigenbasis the orthonormal set $\alpha_1(x), \alpha_2(x), \alpha_3(x), \ldots$ or the orthonormal set $\tilde{\alpha}_1(x), \tilde{\alpha}_2(x), \alpha_3(x), \ldots$.

With this point understood, suppose we now introduce a second observable operator $\hat{B}$ which commutes with $\hat{A}$. According to the Compatibility Theorem, $\hat{A}$ and $\hat{B}$ possess a common eigenbasis. However, this is *not* to say that *every* eigenvector of $\hat{A}$ in the "plane" of $\alpha_1(x)$ and $\alpha_2(x)$ is necessarily an eigenvector of $\hat{B}$ as well; all we can conclude from the Compatibility Theorem is that *at least two* eigenvectors in this plane, say $\tilde{\alpha}_1(x)$ and $\tilde{\alpha}_2(x)$, are orthonormal eigenvectors of $\hat{B}$. Now, if it should happen that $\tilde{\alpha}_1(x)$ and $\tilde{\alpha}_2(x)$ do not correspond to the same eigenvalue of $\hat{B}$, then they will in fact be the *only* eigenvectors of $\hat{B}$ in the plane of $\alpha_1(x)$ and $\alpha_2(x)$.† Consequently, the common eigenbasis $\{\phi_n(x)\}$ will necessarily contain $\tilde{\alpha}_1(x)$ and $\tilde{\alpha}_2(x)$ and *not* $\alpha_1(x)$ and $\alpha_2(x)$. In a manner of speaking, the introduction of the compatible observable $\mathcal{B}$ has "resolved" the ambiguity of which of the infinitely many pairs of A_{12}-eigenvectors "ought" to be used; at the same time, of course, it is likely that other ambiguities in either the $\hat{A}$ eigenvectors or the $\hat{B}$ eigenvectors become similarly resolved.

Now, it is clear that the introduction of the compatible observable $\mathcal{B}$ may not completely resolve all degeneracies; that is, there may still be "subspaces" of two or even more dimensions such that all vectors in any one of these subspaces are eigenvectors belonging to the same eigenvalue of $\hat{A}$ *and* the same eigenvalue of $\hat{B}$. In such a case, we must search for a third observable $\mathcal{C}$ whose operator $\hat{C}$ commutes with *both* $\hat{A}$ and $\hat{B}$, and which further reduces the number of degeneracies.‡ We continue in this way until we obtain what is called a *complete set of compatible observables*, $\mathcal{A}, \mathcal{B}, \ldots, \mathcal{F}$. The corresponding operators, $\hat{A}, \hat{B}, \ldots, \hat{F}$ all commute in pairs, and possess a common eigenbasis $\{\phi_n(x)\}$ which has the following special property: each eigenvector of the set $\{\phi_n(x)\}$ corresponds to a unique set of eigenvalues of the operators $\hat{A}, \hat{B}, \ldots, \hat{F}$. In other words, any two of the eigenvectors $\phi_m(x)$ and $\phi_n(x)$ belong to different eigenvalues of at least one of these operators. This, in fact, is just the defining condition for a complete set of compatible observables. It is a basic

†This follows from the theorem proved in Exercise 18: Two eigenvectors of a Hermitian operator belonging to *unequal* eigenvalues must be orthogonal.

‡We emphasize "both" here, because given that $\hat{A}$ and $\hat{B}$ commute, then the fact that $\hat{C}$ commutes with $\hat{A}$ does *not* guarantee that $\hat{C}$ commutes with $\hat{B}$ as well. As an example, one can show from Eqs. (4-88) that the operator $\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$ commutes with both $\hat{L}_x$ and $\hat{L}_y$, even though $\hat{L}_x$ and $\hat{L}_y$ clearly do not commute with each other.

assumption of quantum mechanics that every physical system possesses at least one such set of observables.

The concept of a complete set of compatible observables leads us to another important concept, namely, that of a "maximal measurement." We recall that, for the case of nondegenerate eigenvalues, Postulate 4 tells us that a measurement of a *single* observable is sufficient to "prepare a state"—i.e., to force the system into a state vector which is *known*, even though the state vector just prior to the measurement was *unknown*. However, when degenerate eigenvalues are involved, we cannot make this statement; for, if the measurement should yield a degenerate eigenvalue, then Postulate 4′ clearly implies that the resulting state vector cannot be uniquely determined unless the initial state vector is known. How, then, can we "prepare a state" when degenerate eigenvalues are involved? The answer to this question is simply that we must simultaneously measure *all* the observables of a complete, compatible set. The performance of these simultaneous or successive measurements, which will not "interfere" with one another since the observables are by definition compatible, will necessarily leave the state vector of the system coincident with a vector which is at once an eigenvector of *all* the operators $\hat{A}$, $\hat{B}, \ldots, \hat{F}$—i.e., with one of the simultaneous eigenbasis vectors $\{\phi_n(x)\}$. Since the observables are "complete," we then need only examine the eigenvalues $A_i, B_j, \ldots, F_k$ obtained in these measurements to pinpoint precisely *which one* of the common eigenvectors the state vector of the system has been forced into. The preparation of a state by simultaneously measuring all of the observables of a complete, compatible set is called a *maximal measurement* of the system. It is clear that a single measurement of an observable whose eigenvalues are completely nondegenerate is already a maximal measurement.

The presence of degenerate eigenvalues can be directly observed in a variety of laboratory experiments. For example, the Hamiltonian operator for a number of electrons in the pure Coulomb field of an atomic nucleus is found to commute with the angular momentum operator, but to be highly degenerate; in other words, it is found that many of the common energy/angular momentum eigenbasis vectors will correspond to the *same* eigenvalue of the energy operator, but *different* eigenvalues of the angular momentum operator. If an external magnetic field is impressed upon the atom, the Hamiltonian operator will be altered, because the classical Hamiltonian function will now contain terms describing the interaction energy between the electrons and the external field. Now, under certain conditions, the new Hamiltonian operator will *still* commute with the angular mo-

mentum operator, but it will *not* be so highly degenerate. Thus a set of angular momentum eigenvalues which previously corresponded to the *same* energy eigenvalue will now correspond to *different* energy eigenvalues. The result is termed a "splitting of the energy levels," and the effects can be observed and quantitatively accounted for in spectroscopic studies in the laboratory.

Let us summarize now the important modifications to our theory which are necessitated by the presence of degenerate eigenvalues: Postulate 3 is modified, but in a very obvious way. Postulate 4 is modified in such a way as to suggest that an ideal measurement is analogous to sending the state vector of the system through a "filter," which passes undistorted all those components of the state vector which belong to the measured eigenvalue, but which completely cuts out all the other components. The Compatibility Theorem remains intact, but a unique choice for the vectors of the common eigenbasis set $\{\phi_n(x)\}$ requires the specification of a "complete set of compatible observables." In order for the measurement process to yield a state vector which is known even though the state vector just prior to the measurement is not known—i.e., in order to "prepare a state"—it is necessary to make a "maximal measurement"; this is simply a set of simultaneous measurements of all the members of a complete set of compatible observables, and it results in a precisely known state, irrespective of the initial state, simply because each of the common eigenbasis vectors $\{\phi_n(x)\}$ corresponds to a unique set of eigenvalues for these observables.

4-6d Concluding Remarks

In closing, we might recall that the theory of quantum mechanics given here is restricted to systems which are *nonrelativistic*. In other words, we have presented a theory which is valid only for systems which contain velocities that are much less than the velocity of light, c, or equivalently, for systems which contain particles whose energies are much less than their rest-mass energies, mc^2. We remarked in Sec. 4-5c that our "common sense" ideas about the physical world are for a world in which $\hbar = 0$ and $c = \infty$. In point of fact, the *real* world is a world in which $\hbar$ is finite but very small, while c is finite but very large. Of course, the terms "very small" and "very large" are relative to our customary units of length, mass and time (e.g., the mks units). Now, by simply redefining these units, it is actually possible to obtain a system of units in which

$$\hbar = c = 1 \tag{4-109}$$

The set of units defined by Eqs. (4-109) are called the *natural units*. Recent experiments in the physics of "elementary particles" (e.g., protons, electrons, photons, muons, kaons, neutrinos, etc.) have indicated that the fundamental processes in the universe seem to take place on the scale of the natural units. These processes must therefore be treated within the framework of *relativistic quantum mechanics*, or quantum field theory as it is also called. Even the most cursory discussion of this presumably "ultimate theory" is far beyond the scope of this book. This is partly because the theory is very complicated, and partly because it has not yet been put into a completely consistent and fully understood form.

Broadly speaking, ordinary nonrelativistic quantum mechanics has provided us with a reasonably complete conceptual and quantitative understanding of atomic physics and chemistry; that is, it has been able to explain with remarkable precision the properties of the elements of the periodic table, and thus the properties of all substances which can be formed from these elements. It is hoped that someday relativistic quantum mechanics will be able to afford a similar understanding of the many elementary particles, and of the fundamental mechanisms by which these elementary particles interact with each other to form the various atomic nuclei. It may well be that, in its final form, relativistic quantum theory will be as different from ordinary quantum mechanics, as ordinary quantum mechanics is from classical mechanics. Nevertheless, the nonrelativistic theory of quantum mechanics has and will continue to have a very large range of validity, and it will undoubtedly long stand as one of the most inventive and successful achievements of the human intellect.

Index