CHAPTER

4

The Theory of Quantum Mechanics

The resumé of classical mechanics given in the last chapter was presented from a very elementary point of view: We did not attempt to critically analyze the basic concepts of length, time, force and mass; we confined ourselves to a system so simple that it precluded any discussion of such important topics as angular momentum and multiparticle phenomena; we did not present the Hamiltonian formulation in its full generality, nor even mention any other formulations; and we did not try to use classical mechanics to solve various specific "problems." In truth, our sole aim was to present, in as simple and uncomplicated a way as possible, only the barest essentials of the classical theory. And in spite of (or maybe because of) all our omissions, we did gain a concise and fairly accurate perspective of the general aims, assumptions and methodology of classical mechanics. In the present chapter we shall try to carry out an analogous program with respect to quantum mechanics.

Unfortunately, quantum mechanics is inherently abstract, and is not as easy to grasp and understand as classical mechanics. It might seem odd that, in their attempt to formulate a more truthful picture of physical phenomena, physicists came up with a theory which is so highly abstract that it seems quite remote from physical reality. Yet we must bear in mind that our personal notions of physical reality are derived from our lifelong contact with *macroscopic* phenomena, and we have no license at all to extrapolate our "macroscopic intuition" to the *microscopic* level. Indeed, the physicist has found, through carefully performed experiments, that microscopic reality is characterized by phenomena such as the wave-particle duality which seem quite unintelligible in terms of "common sense" reasoning. We should not be too surprised, then, to find that the logical system which purports to account for these seeming anomalies will itself

seem foreign to our deeply ingrained, classical point of view. As a consequence, the reader will be obliged to simply accept the abstractness of the quantum theory, because at the present time there does not seem to be any other way to get hold of it. Whether this is a consequence of "the way things are," or whether this indicates that our present understanding of quantum mechanics is in some way deficient, is a question which many physicists are still pondering.

We are going to present the theory of quantum mechanics by laying down a number of *postulates* (six, to be precise), from which we shall deduce various consequences. These postulates will undoubtedly seem strange to the reader, and completely devoid of any intuitive appeal. However, the reader must keep in mind that, together, these postulates form the simplest and most widely accepted logical basis yet devised for understanding *quantitatively* an enormous range of physical phenomena; indeed, it is this fact, and not any "reasonableness" on the part of the postulates themselves, that gives us cause to accept them. The postulates, together with their derived consequences, will constitute our "picture" of quantum mechanics.

It should be noted that both the number and content of the fundamental postulates of quantum mechanics are to some extent a matter of personal taste. In keeping with the limited aims of this book, the postulates presented here will not be as general and logically economical as they could be, and of course, we shall only attempt to derive a limited selection of their consequences.

We shall develop the theory of quantum mechanics for a nonrelativistic physical system with one degree of freedom, which we represent by the variable x. Although our treatment in Secs. 4-1 through 4-4 will be applicable to *any* kind of nonrelativistic system with one degree of freedom, the reader may find it helpful to keep in mind the specific system discussed in the preceding chapter—i.e., a particle of mass m moving on the x-axis in a potential field $V(x)$. We shall specialize our treatment to this important type of system in Sec. 4-5.

4-1 THE QUANTUM STATE

In Chapter 3 we saw that classical mechanics identifies the *state* of a physical system with the current values of certain *observables* of the system (e.g., the position x and the momentum p). Quantum mechanics, on the other hand, makes a very sharp distinction be-

tween states and observables. Concerning the *state* of a system in quantum mechanics, we have the following postulate:

> **Postulate 1.** Every possible physical state of a given system corresponds to some normed Hilbert space vector $\psi(x)$, and conversely, every normed Hilbert space vector $\psi(x)$ corresponds to a possible physical state of the system. This correspondence between physical states and normed vectors in $\mathcal{H}$ is one-to-one, except that two normed $\mathcal{H}$-vectors that differ only by an overall scalar factor of modulus unity correspond to the same physical state. The particular $\mathcal{H}$-vector to which the state of the system corresponds at time t is denoted by $\Psi_t(x)$ and is called the *state vector* of the system; the system is said to "be in the state $\Psi_t(x)$." The state of a system is *completely described* by the state vector in the sense that anything which is in principle knowable about the system at time t can be learned from the function $\Psi_t(x)$.

This postulate makes three assertions: First, it asserts that the possible physical states of a given system stand in a one-to-one correspondence with the normed $\mathcal{H}$-vectors $\psi(x)$ which are defined *up to* an overall scalar factor of modulus unity. In saying that $\psi(x)$ is "normed," we mean simply that its norm, or inner product with itself, equals 1: $(\psi,\psi) = 1$. In saying that $\psi(x)$ is defined up to an overall scalar factor of modulus unity, we mean that if c is any $\mathcal{H}$-scalar (i.e., any complex number) satisfying $|c|^2 = 1$, then the two $\mathcal{H}$-vectors $\psi(x)$ and $\phi(x) = c\psi(x)$ are "physically equivalent" in that they correspond to the same physical state.

Exercise 25. If $\psi(x)$ has unit norm, and c is a complex number satisfying $|c|^2 = 1$, show that $\phi(x) = c\psi(x)$ also has unit norm.

The second assertion of Postulate 1 is that everything that can possibly be known about the state of the system at time t can be obtained from its "state vector" $\Psi_t(x)$. However, the postulate says nothing about *what* things can be known, nor *how* they can be derived from the state vector. These questions will evidently have to be answered by subsequent postulates.

In view of these first two assertions of Postulate 1, we can conclude that the state of a system at time t is completely specified if and only if the state *vector* $\Psi_t(x)$ is given as a definite *function of* x. For comparison, we recall that in classical mechanics the state of a system at time t is completely specified if and only if the state *variables* (e.g., $x(t)$ and $p(t)$) are given as definite *real numbers.* In

this connection, the reader should *not* try to identify the x in $\Psi_t(x)$ with the classical position variable $x(t)$; that is, $\Psi_t(x)$ does *not* stand for $\Psi(x(t))$. The x in $\Psi_t(x)$ merely represents the argument of the function Ψ_t; it is a dummy variable, and does not depend in any way upon the time variable t. In the special case where the physical system under consideration is a mass particle, it turns out that there is a very intimate connection between "position" and the argument of the Hilbert space vectors; however, this connection is quite unlike anything the reader might now suppose, and it cannot be explained until after all the postulates of quantum mechanics have been introduced.

The third assertion of Postulate 1 is somewhat indirect. In writing the state vector of the system at time t as $\Psi_t(x)$, it is implied that the state vector is in some sense a function of time. To bring this out more explicitly, we shall sometimes write $\Psi_t(x)$ as $\Psi(x,t)$. It must be emphasized, however, that the functional dependence of Ψ on t is essentially different from its dependence on x: as a Hilbert space vector, $\Psi(x,t)$ is properly a function of x alone, and the parameter t serves merely to label different vectors in $\mathcal{H}$. Thus, $\Psi(x,t_1)$ and $\Psi(x,t_2)$ are to be regarded as two different $\mathcal{H}$-vectors—i.e., two different functions of x—which specify the state of the system at two different times t_1 and t_2. The behavior of $\Psi_t(x) \equiv \Psi(x,t)$ as a function of time—i.e., the "time evolution" of the state vector—will be taken up in Postulate 5 [Sec. 4-4]. Until then we shall be mainly concerned with the state of the system at a single instant t.

Finally, we should mention that the state vector $\Psi_t(x)$ is sometimes referred to as the "state function" or "wave function" of the system.

4-2 OBSERVABLES IN QUANTUM MECHANICS

The quantum mechanical specification of the *state* of a system evidently makes no reference at all to any physical *observables* of the system; this is in marked contrast to the way in which the state of a system is specified in classical mechanics [see Sec. 3-2]. However, Postulate 1 does assert that the state vector in some way contains everything we can possibly know about the system. It therefore seems reasonable to expect that, if at some time t we know the exact functional form of the state vector $\Psi_t(x)$, then we ought to be able to make some fairly definite assertions about the physical observables of the system at that instant; indeed, if we could not do this, then

the state vector would be a completely useless mathematical abstraction. Before presenting the postulate which tells us exactly what *can* be said about an observable when the state vector is known, it is first necessary to define more precisely the quantum mechanical concept of an observable. Since for the present we want to keep our discussion as general as possible, we shall refer to observables by script capital letters (e.g., the observable $\mathcal{A}$, the observable $\mathcal{B}$, etc.).

As in classical mechanics, an *observable* $\mathcal{A}$ is simply a dynamical variable that can be measured; e.g., for a mass on the x-axis, the observables are the position, the momentum, and functions of the position and momentum (of which the energy is perhaps the most useful). The *measurement* of an observable $\mathcal{A}$ is some well-defined physical operation which, when performed on the system, yields a single real number called "the value of $\mathcal{A}$." For simplicity, we shall consider all measurements to be "ideal" in the sense that the measured value has an experimental uncertainty of zero.

Now in classical mechanics, no real distinction is made between the *mathematical representation* of an observable and the *values* of the observable; however, in quantum mechanics such a distinction is of fundamental importance. Postulate 2, which we now state, is concerned with (*a*) the mathematical representation of $\mathcal{A}$, and (*b*) the possible values of $\mathcal{A}$.

Postulate 2.

(a) To each physical observable $\mathcal{A}$, there corresponds in the Hilbert space a linear Hermitian operator $\hat{A}$, which possesses a complete, orthonormal set of eigenvectors $\alpha_1(x)$, $\alpha_2(x)$, $\alpha_3(x)$, . . . and a corresponding set of real eigenvalues $A_1, A_2, A_3, \ldots$

$$\hat{A}\alpha_i(x) = A_i\alpha_i(x) \qquad i = 1,2,3 \ldots \tag{4-1}$$

Conversely, to each such operator in the Hilbert space there corresponds some physical observable.

(b) The only possible values which any measurement of $\mathcal{A}$ can yield are the eigenvalues $A_1, A_2, A_3, \ldots$.

Let us begin our discussion of this postulate by reviewing the definitions of the mathematical terms used in part (a). The fact that the $\mathcal{H}$-vectors $\{\alpha_i(x)\}$ are "eigenvectors" of the operator $\hat{A}$ with "eigenvalues" $\{A_i\}$ simply means that Eqs. (4-1) hold true for all i. To say that the eigenvectors form a "complete, orthonormal set" means that, for all i and j,

$$(\alpha_i, \alpha_j) \equiv \int_{-\infty}^{\infty} \alpha_i^*(x)\,\alpha_j(x)\,dx = \delta_{ij} \tag{4-2a}$$

and, moreover, that any $\mathcal{H}$-vector $\phi(x)$ can be written as

$$\phi(x) = \sum_{i=1}^{\infty} (\alpha_i, \phi)\alpha_i(x) \tag{4-2b}$$

In other words, the eigenvector set $\{\alpha_i(x)\}$ is an "orthonormal basis set" in the Hilbert space. For brevity, we shall refer to the set $\{\alpha_i(x)\}$ as the *eigenbasis* of $\hat{A}$. The condition that the eigenvalues $\{A_i\}$ all be real can be conveniently written

$$A_i^* = A_i \tag{4-3}$$

Finally, the fact that $\hat{A}$ is an "Hermitian" operator means that, for any two $\mathcal{H}$-vectors $\phi_1(x)$ and $\phi_2(x)$,

$$(\phi_1, \hat{A}\phi_2) = (\hat{A}\phi_1, \phi_2) \tag{4-4}$$

It should be noted that the Hermiticity of $\hat{A}$ really need not have been postulated, because we proved in Chapter 2 [see Exercise 19] that any linear operator which possesses a complete, orthonormal set of eigenvectors and a corresponding set of real eigenvalues is *necessarily* Hermitian.† Nevertheless, the Hermitian property is such a fundamental property of observable operators that we have allowed this minor redundancy to enter into our statement of Postulate 2.

Postulate 2 says, first of all, that a physical observable $\mathcal{A}$ is "mathematically represented" by a linear operator $\hat{A}$ which possesses a complete, orthonormal set of eigenvectors and a corresponding set of real eigenvalues. We shall call such an operator an *observable operator*. The consequences of mathematically representing observables by operators remain to be seen, but it is reasonable to expect that these consequences will depend strongly on the mathematical rules for manipulating operators. These rules were discussed in Sec. 2-4. Perhaps the most striking difference between these rules and the rules for manipulating ordinary numbers is that, whereas two numbers A and B always *commute* ($AB \equiv BA$), it is *not* true that two operators $\hat{A}$ and $\hat{B}$ necessarily commute [e.g., see Exercise 14]. Thus, in some sense it may be said that observables always commute in classical mechanics, but not in quantum mechanics. We shall see later that the fact that not all pairs of observable operators commute

† A linear operator can be Hermitian *without* possessing a complete, orthonormal set of eigenvectors and a corresponding set of real eigenvalues; however, such an operator does *not* correspond to an observable.

leads to some surprising (i.e., nonclassical) results—among them, the so-called "wave-particle duality" described in Chapter 1.

Although Postulate 2 ascribes no real physical significance to either the observable operator $\hat{A}$ or its eigenbasis $\{\alpha_i(x)\}$, the assertion is made in part (b) that its eigenvalues $\{A_i\}$ are the *only* numbers that can be obtained in any measurement of $\mathcal{A}$. The fact that these eigenvalues are real corresponds to the fact that the measurement operation always yields a real number. However, notice that nothing in the postulate requires that these eigenvalues form a "continuous set" (i.e., that they cover densely all or part of the real number axis); indeed, our labeling of the eigenvalues seems to suggest that they form a "discrete set" (i.e., that the difference between any two given eigenvalues is some finite, nonzero number). Now as a matter of fact, the set $\{A_i\}$ may be continuous *or* discrete *or* a combination of the two, depending entirely on the particular observable operator to which the eigenvalues belong. Nevertheless, it is highly significant that *the eigenvalues can be discretely distributed*; for this immediately opens up the possibility for allowing certain physical observables to be "quantized." We recall from Chapter 1 that it was the experimental discovery of such quantized observables that formed one of the great stumbling blocks for classical mechanics.

Until now, the main simplifying restriction imposed on our development of quantum mechanics has been the restriction to systems with only *one* degree of freedom (hence the appearance of the *single* variable x as the argument of the $\mathcal{H}$-functions). We shall now impose a second simplifying restriction: henceforth, we shall treat the general observable operator $\hat{A}$ as though its eigenvalues were *entirely discretely distributed*. We do this because a discussion of operators with continuously distributed eigenvalues involves some extraordinary mathematical manipulations, which, at this point, would tend to confuse rather than enlighten. By concentrating on operators with discretely distributed eigenvalues, we shall be able to present most of the essential points of the theory with relative clarity and simplicity; moreover, many aspects of the properties of observable operators with continuously distributed eigenvalues may be understood by drawing analogies with the results for the discrete case. A brief discussion of the mathematical techniques required to deal with continuous eigenvalues will be given later in Sec. 4-6b.

In classical mechanics, if $\mathcal{A}$ is an observable then any real function of $\mathcal{A}$, $f(\mathcal{A})$, is also an observable (e.g., $\mathcal{A}^2$ or $e^{\mathcal{A}}$); this is because if we measure a value for $\mathcal{A}$, then we have obviously measured a value for $f(\mathcal{A})$ also. Does this fact carry over into quantum mechanics? The answer to this question is yes, provided we restrict

ourselves to real functions $f(z)$ which have a power series expansion (i.e., a Taylor expansion) in z:

$$f(z) = \sum_{n=0}^{\infty} c_n z^n, \qquad \{c_n\} \text{ real} \tag{4-5a}$$

To see how this comes about, let us consider the *operator* $f(\hat{A})$, which is *defined* by formally replacing the variable z by the observable operator $\hat{A}$ in the series expression for $f(z)$:

$$f(\hat{A}) \equiv \sum_{n=0}^{\infty} c_n \hat{A}^n \tag{4-5b}$$

By $\hat{A}^n$, we mean of course $\hat{A}$ multiplied by itself n times; thus, in view of the definitions in Eqs. (2-42), $f(\hat{A})$ is a well-defined operator. We shall now show that $f(\hat{A})$ is in fact an *observable* operator, and that the particular observable to which it corresponds is $f(\mathcal{A})$. These conclusions are arrived at from the results of the following exercise.

Exercise 26.

(a) Prove that the operator $c\hat{A}^n$, where $n = 0,1,2, \ldots$, has eigenvectors $\{\alpha_i(x)\}$ and eigenvalues $\{cA_i^n\}$; here, $\{\alpha_i(x)\}$ and $\{A_i\}$ are the eigenvectors and eigenvalues respectively of the observable operator $\hat{A}$. [*Hint*: Establish the result for $n = 0$ and $n = 1$; then show that the result holds for any $n \geq 2$ provided it holds for $n - 1$.]

(b) Using the result of part (a), prove that the operator $f(\hat{A})$ has eigenvectors $\{\alpha_i(x)\}$ and eigenvalues $\{f(A_i)\}$.

We see then that the operator $f(\hat{A})$ has a complete orthonormal set of eigenvectors and a corresponding set of real eigenvalues; thus, by Postulate 2, $f(\hat{A})$ is indeed an *observable* operator. Furthermore, since the eigenvalues of $f(\hat{A})$ are $\{f(A_i)\}$, it seems quite reasonable to associate this observable operator with the particular observable $f(\mathcal{A})$. This establishes what we wanted to prove, and it also shows that the observable operator $f(\hat{A})$ has the same eigenbasis as $\hat{A}$ does.

In classical mechanics the fact that we identify the state of a system with certain physical observables means that *both* the state and the observables depend upon time. However, in quantum mechanics no such identification is drawn between the state and the observables, and an examination of Postulates 1 and 2 leads us to conclude that, while the state vector $\Psi_t(x)$ generally changes with time, the observable operator $\hat{A}$, along with its eigenbasis $\{\alpha_i(x)\}$ and its eigenvalues $\{A_i\}$, are all *independent* of time. One consequence of this is that, if we express the state vector $\Psi_t(x)$ as a linear combina-

tion of the eigenvectors $\{\alpha_i(x)\}$ in the manner of Eq. (4-2b),

$$\Psi_t(x) = \sum_{n=1}^{\infty} (\alpha_n, \Psi_t)\alpha_n(x) \tag{4-6a}$$

then the expansion coefficients or components of $\Psi_t(x)$ relative to the eigenbasis of $\hat{A}$ will be *time-dependent* scalars:

$$(\alpha_n, \Psi_t) \equiv \int_{-\infty}^{\infty} \alpha_n^*(x)\Psi(x,t)dx \qquad (n = 1,2,\ldots) \tag{4-6b}$$

We refer to Eq. (4-6a) as "an expansion of the state vector $\Psi_t(x)$ in the eigenbasis of the observable operator $\hat{A}$." At the moment, we have no apparent reason for ever wanting to write the state vector in such a way; however, as we shall see later, expansions of this type play a very key role in the quantum theory.

It should be remarked that the time-*in*dependence of $\hat{A}$, $\{\alpha_i(x)\}$ and $\{A_i\}$ does *not* necessarily mean that the measured values of $\mathcal{A}$ will be constant in time. For it remains to be specified how the outcome of a particular measurement depends upon the state of the system at the time of the measurement; thus, owing to the time evolution of $\Psi_t(x)$, different eigenvalues of $\hat{A}$ might be measured at different times. This matter will be clarified later.

4-3 THE QUANTUM THEORY OF MEASUREMENT

In Postulate 1 we have associated physical *states* with Hilbert space *vectors*, and in Postulate 2 we have associated physical *observables* with Hilbert space *operators*. However, we certainly cannot expect to form a meaningful physical theory merely by associating two separate physical entities with two separate mathematical entities. Clearly it is necessary to establish some sort of logical connection between the state vector of the system and the observable operators of the system. This logical connection is made, although in a somewhat indirect way, via the concept of *measurement*. The quantum theory of measurement thus forms the keystone of the theoretical structure of quantum mechanics; it is here that the respective roles of the state vector and the observable operators are clarified and interrelated.

It is perhaps appropriate to say again just what we mean by "a measurement": We regard a measurement simply as some in-principle well-defined physical operation which, when performed on a system,

yields a single, errorless, real number. By "errorless" we mean that there is no experimental uncertainty associated with the number obtained, so that it can be regarded as being infinitely precise. Clearly what we are really contemplating here is an ideal measurement—a highly simplified abstraction of what actually occurs in the laboratory. A more thoroughgoing inquiry into the nature of a physical measurement is necessary to form a truly complete picture of the quantum theory, but such a critique is quite involved and will not be attempted here.

We mentioned in Chapter 1 that the logical foundations of quantum mechanics have been and continue to be the subject of serious debate among some physicists and philosophers of science. Most of this debate has centered upon the theory of measurement. We shall not try to discuss here the intricate pros and cons of this debate; instead, we shall merely present the essential features of the quantum theory of measurement *in the form that is currently favored by most physicists.* However, it must be noted that there are a few other tenable views of the quantum theory of measurement, all of them conflicting in various degrees with the "orthodox" view. All these views, including the orthodox one, are based on various philosophical predilections which go somewhat *beyond* what the experimental evidence directly implies; indeed, the conflict between these views is confined mainly to the philosophical level, since as yet no real laboratory experiment has been devised which is capable of picking out the "correct" view. Conversely, until such an experiment comes along, the question of which view is really more legitimate has no *practical* import.

We begin in Sec. 4-3a by stating and discussing Postulate 3, which tells us exactly what can be predicted about the outcome of a measurement of an observable $\mathcal{A}$ which is performed on a system in a known state $\Psi_t(x)$. Next, in Sec. 4-3b, we present and discuss Postulate 4, which tells us how the state vector is affected by such a measurement. Finally, in Sec. 4-3c, we derive two very important consequences of these postulates—namely, the Compatibility Theorem and the celebrated Heisenberg Uncertainty Principle.

4-3a Predicting the Result of a Measurement Expectation Values and Uncertainties

We learned in Postulate 2 that the only values that can ever be measured for an observable are the eigenvalues of the corresponding observable operator. As to which one of these eigenvalues will be

obtained in any given instance, we may expect that this will be somehow determined by the particular form of the state vector at the time of the measurement. As we shall now see, the form of the state vector does indeed have an important bearing on which eigenvalue is measured, but it is in general *not possible* to predict with *absolute certainty* the outcome of a single measurement.

> **Postulate 3.** If an observable operator $\hat{A}$ has eigenbasis $\{\alpha_i(x)\}$ and eigenvalues $\{A_i\}$, and if the corresponding observable $\mathfrak{A}$ is measured on a system which, immediately prior to the measurement, is in the state $\Psi_t(x)$, then the *strongest predictive statement that can be made* concerning the result of this measurement is as follows: The *probability* that the measurement will yield the eigenvalue A_k is $|(\alpha_k,\Psi_t)|^2$.†

This postulate unquestionably marks the point at which the theory of quantum mechanics diverges most radically from the theory of classical mechanics. We recall that, in classical mechanics, if the instantaneous state of the system $[x(t),p(t)]$ is known, then it is *certain* that a measurement of some observable $\mathfrak{A} = f(x,p)$ at time t will yield the number $f(x(t),p(t))$. In contrast to this, Postulate 3 asserts that if the instantaneous state vector of the system $\Psi_t(x)$ is known, then all that can be predicted about a measurement of $\mathfrak{A}$ at time t is that $|(\alpha_1,\Psi_t)|^2$ is the probability that the number A_1 will be obtained, $|(\alpha_2,\Psi_t)|^2$ is the probability that the number A_2 will be obtained, and so on for the other eigenvalues of A. Now since Postulate 1 asserts that the state of a system is *completely defined* by the state vector $\Psi_t(x)$, and since Postulate 3 further asserts that a knowledge of the state vector suffices *only* to predict *probabilities* for obtaining various results in a measurement, then we are forced to conclude:

(i) It is often *not* possible to predict with certainty the outcome of a measurement which is performed on a system in a *completely defined state.*

(ii) If a system is subjected to two separate but identical measurements, with due care taken to insure that the system is in the *exact same state* just prior to each measurement, the results of the two measurements will *not necessarily coincide.*

In accepting Postulate 3, it is evidently incumbent upon us also to accept this "unpredictability" and "nonuniqueness" of the measurement process as being manifestations of some inherent property

†If the eigenvalues of $\hat{A}$ were *continuously* distributed, we would have to state this postulate a bit differently. We shall discuss this point later in Sec. 4-6b.

of Nature. Although this view disagrees violently with our own deeply ingrained "classical" intuition, the most we can justifiably claim in rebuttal is that, on the *macroscopic* level, this property of Nature must not be *noticeable*. We shall return to this point later in Sec. 4-5c.

Having acknowledged this negative aspect of Postulate 3, namely that a unique result of measuring $\mathfrak{A}$ on a system in a known state $\Psi_t(x)$ usually *cannot* be predicted, let us consider now its positive aspect: the *probability* for obtaining the eigenvalue A_k *can* be predicted, and is in fact equal to $|(\alpha_k,\Psi_t)|^2$.

Exercise 27. Prove that, as required by Postulate 1, the quantity $|(\alpha_k,\Psi_t)|^2$ is not changed if $\Psi_t(x)$ is replaced by the vector $c\Psi_t(x)$, where c is any $\mathcal{H}$-scalar satisfying $|c|^2 = 1$.

We recall from Eq. (4-6a) that the inner product (α_k,Ψ_t) is just the "component" of the state vector $\Psi_t(x)$ in the "direction" of the eigenbasis vector $\alpha_k(x)$; the value of this time-dependent complex number can be calculated from the functions $\alpha_k(x)$ and $\Psi_t(x)$ according to Eq. (4-6b). It is altogether fitting that the probability for measuring A_k for $\mathfrak{A}$ in the state $\Psi_t(x)$ should be determined by the inner product (α_k,Ψ_t), since this quantity depends both upon the state vector of the system and upon some property of $\hat{A}$ associated with the eigenvalue in question. However, if the square modulus of (α_k,Ψ_t) is to be a *probability*, then we must have, in analogy with Eq. (2-2a),

$$0 \leqq |(\alpha_k,\Psi_t)|^2 \leqq 1 \qquad \text{for all } k \tag{4-7a}$$

Exercise 28. Prove the above inequality. [*Hint*: Use the Schwarz inequality to prove the right-hand relation.]

In addition to satisfying Eq. (4-7a), the inner product (α_k,Ψ_t) must also be such that

$$\sum_{k=1}^{\infty} |(\alpha_k,\Psi_t)|^2 = 1 \tag{4-7b}$$

This relation, which is analogous to Eq. (2-2b), merely expresses the requirement of Postulate 2 that a measurement of $\mathfrak{A}$ is *certain* to yield either A_1 or A_2 or A_3 .... In order to prove that Eq. (4-7b) is indeed satisfied, we shall show that its left hand side is just the norm of $\Psi_t(x)$; our result then will follow from the requirement of Postulate 1 that $(\Psi_t,\Psi_t) = 1$. Using the fact that the eigenvectors $\{\alpha_i(x)\}$ form an orthonormal basis in $\mathcal{H}$, we have for any possible state vector $\Psi_t(x)$,

$$(\Psi_t,\Psi_t) = \left(\sum_{i=1}^{\infty}(\alpha_i,\Psi_t)\alpha_i,\ \sum_{j=1}^{\infty}(\alpha_j,\Psi_t)\alpha_j\right)$$

$$= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(\alpha_i,\Psi_t)^*\,(\alpha_j,\Psi_t)\,(\alpha_i,\alpha_j)$$

$$= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(\alpha_i,\Psi_t)^*\,(\alpha_j,\Psi_t)\delta_{ij}$$

$$= \sum_{i=1}^{\infty}(\alpha_i,\Psi_t)^*\,(\alpha_i,\Psi_t)$$

Thus

$$(\Psi_t,\Psi_t) = \sum_{i=1}^{\infty}|(\alpha_i,\Psi_t)|^2 \tag{4-8}$$

and Eq. (4-7b) follows at once. [Note that Eq. (4-8) is just a particular case of Eq. (2-40b).] We have proved, then, that the numbers $|(\alpha_1,\Psi_t)|^2$, $|(\alpha_2,\Psi_t)|^2$, $|(\alpha_3,\Psi_t)|^2$, . . . satisfy conditions (4-7), and so form a set of probability numbers $p_1, p_2, p_3, \ldots$ analogous to the set discussed in Sec. 2-1. We shall explore this analogy in some detail later in this section.

If it should happen that two of the eigenvalues of $\hat{A}$ are equal, say $A_3 = A_5 = A$, then the probability for measuring the value A in the state $\Psi_t(x)$ is, according to Eq. (2-3a), just the sum of the separate probabilities, $|(\alpha_3,\Psi_t)|^2 + |(\alpha_5,\Psi_t)|^2$. In this case, we would say in the terminology of quantum mechanics that the eigenvalue A is *degenerate*. In order to keep our discussion of quantum mechanics as simple as possible, we are going to restrict our attention to observables $\mathfrak{A}$ whose operators have distinct or *nondegenerate* eigenvalues:

$$\text{Require } A_i \neq A_j \text{ if } i \neq j \quad \text{(nondegenerate eigenvalues)} \tag{4-9}$$

It is with some reluctance that we impose this simplifying restriction on our development of the theory of quantum mechanics, because it turns out that the occurrence of observable operators with degenerate eigenvalues is rather common, and is the source of many interesting phenomena. Thus we feel obliged to return to this matter at the end of our development (Sec. 4-6c), and discuss briefly how some of our conclusions will be modified by the relaxation of Eq. (4-9).

In the next exercise we shall deduce the following very impor-

tant result: A measurement of $\mathfrak{A}$ on the state $\Psi_t(x)$ is *certain* to yield the eigenvalue A_k if and only if $\Psi_t(x)$ coincides with $\alpha_k(x)$. We use the term "coincide" in the loose sense allowed by Postulate 1: two *state* vectors $\Psi_1(x)$ and $\Psi_2(x)$ "coincide," or describe the same physical state, if $\Psi_1(x) = c\Psi_2(x)$, where c is any complex number satisfying $|c|^2 = 1$.

Exercise 29.

(a) Prove that, if $\Psi_t(x) = c\alpha_k(x)$, where $|c|^2 = 1$, then a measurement of $\mathfrak{A}$ at time t is *certain* to yield the value A_k. [*Hint*: Calculate the quantity $|(\alpha_i, \Psi_t)|^2$.]

(b) Prove that, if a measurement of $\mathfrak{A}$ at time t is *certain* to yield the value A_k, then $\Psi_t(x) = c\alpha_k(x)$, where $|c|^2 = 1$. [*Hint*: Prove first that $(\alpha_i, \Psi_t) = 0$ for $i \neq k$; then use Eq. (4-6a).]

We see, then, that a necessary and sufficient condition for a measurement of $\mathfrak{A}$ to yield a unique, predictable value is that the state vector of the system coincide with some eigenvector of $\hat{A}$. This is obviously a very important and useful result. However, it should *not* be taken to imply that the physical state corresponding to the normed $\mathcal{H}$-vector $\alpha_k(x)$ is in some way more precisely defined than the physical state corresponding to some normed, *linear combination* of two or more of the vectors $\{\alpha_i(x)\}$; for, Postulate 1 stipulates that *any* normed $\mathcal{H}$-vector *completely defines* a physical state, regardless of whether or not the vector in question happens to coincide with an eigenvector of some observable operator. Moreover, while $\alpha_k(x)$ may be a "nice" state vector for predicting the result of a measurement of $\mathfrak{A}$, it will *not* necessarily coincide with one of the eigenvectors $\{\beta_i(x)\}$ of some other observable operator $\hat{B}$, in which case the state vector $\alpha_k(x)$ will *not* respond "nicely" to a measurement of $\mathfrak{B}$. In a sense, we are merely reemphasizing here the basic point that, in quantum mechanics, the state of a system is solely and completely determined by the functional form of its state vector, and *not*, as in classical mechanics, by the expected result of any measurement which might be performed on the system.

In Sec. 2-1 we considered the problem of randomly selecting a ball from a box of N identical balls, with n_1 of the balls bearing the number v_1, n_2 bearing the number v_2, etc., and with $\Sigma_k n_k = N$. We observed that the probability p_k that the randomly selected ball will show the number v_k is n_k/N. Suppose we let

$$\left.\begin{aligned} v_k &= A_k \\ n_k &= |(\alpha_k, \Psi_t)|^2 N \end{aligned}\right\} \qquad (4\text{-}10a)$$

We note that the second of these equalities is indeed legitimate, since Eq. (4-7b) insures that $\Sigma_k n_k = N$; in addition, all of the numbers $|(\alpha_k, \Psi_t)|^2 N$ can be made as close to integer values as desired simply by taking N large enough. The probability that a randomly selected ball will show the number A_k is evidently

$$p_k = n_k/N = |(\alpha_k, \Psi_t)|^2 \tag{4-10b}$$

which is just the probability for measuring A_k in the state $\Psi_t(x)$. Therefore, we see that this hypothetical ball-drawing experiment *simulates* the process of measuring $\mathfrak{A}$ on a system in the state $\Psi_t(x)$. The major difference is that, in the ball-drawing experiment, it is in principle possible to eliminate the unpredictability and nonuniqueness of a drawing merely by ascertaining the exact positions of all the balls in the box and drawing accordingly; however, as we have previously emphasized, there is *no* possible way of eliminating the element of uncertainty in the measurement of an observable on a given state. Another difference is that, owing to the time-dependence of the inner products (α_k, Ψ_t), the numbers $\{n_i\}$ and the probabilities $\{p_i\}$ in Eqs. (4-10) will in general change with time (but note that the values $\{v_i\} = \{A_i\}$ will evidently remain fixed). However, we shall postpone until Sec. 4-4 a consideration of this time-dependence, and for now continue to confine our discussion to a single instant t.

In view of the foregoing analogy between the measurement process in quantum mechanics and the ball-drawing experiment of Sec. 2-1, it is clear that we can define for a *series* of measurements an "average value" and an "rms deviation" analogous to $\langle v \rangle$ and Δv in Eqs. (2-4) and (2-5). However we recall that, in the multiple drawing procedure used to define $\langle v \rangle$ and Δv, we were careful to return to the box each ball drawn before making the next random drawing; in other words, we wanted each drawing to be made with the box of balls in the same "state." Therefore, in order to define quantities analogous to $\langle v \rangle$ and Δv for a series of measurements, we must take care to insure that the state vector of the system is the *same* for each measurement of the series. We shall not concern ourselves here with how this can be accomplished; however, we should remark that this is not a trivial requirement, because, as we shall find in Postulate 4, the state vector of the system usually suffers a drastic alteration as a result of a measurement. For this reason, it is necessary to distinguish in our subsequent discussion two different types of multiple measurements:

(i) A series of M *repeated measurements* on the state $\Psi_t(x)$ is a series of M measurements which are performed with the system always in the state $\Psi_t(x)$ just prior to each measurement.

(ii) A series of M *successive measurements* is a series of M measurements performed in rapid succession, such that the state vector of the system for the nth measurement is the state which results from the $(n\text{-}1)$th measurement.

We shall discuss the results obtained in a series of *successive* measurements after we have introduced Postulate 4. For now, we consider a series of very many *repeated* measurements of $\mathfrak{A}$ on the state $\Psi_t(x)$: We denote by the sumbol $\langle \hat{A} \rangle_t$ the *average* of the values obtained in these repeated measurements, and by $\Delta \hat{A}_t$ the *rms deviation* of these values. Therefore, putting $v_k = A_k$ and $p_k = |(\alpha_k, \Psi_t)|^2$ into the expressions for $\langle v \rangle$ and Δv in Eqs. (2-7) and (2-9), we have at once

$$\langle \hat{A} \rangle_t = \sum_{k=1}^{\infty} |(\alpha_k, \Psi_t)|^2 A_k \tag{4-11}$$

and

$$\Delta \hat{A}_t = \sqrt{\left(\sum_{k=1}^{\infty} |(\alpha_k, \Psi_t)|^2 A_k^2 \right) - \left(\sum_{k=1}^{\infty} |(\alpha_k, \Psi_t)|^2 A_k \right)^2} \tag{4-12}$$

In the terminology of quantum mechanics, $\langle \hat{A} \rangle_t$ is called "the *expectation value* of $\mathfrak{A}$ in the state $\Psi_t(x)$," and $\Delta \hat{A}_t$ is called "the *uncertainty* in $\mathfrak{A}$ in the state $\Psi_t(x)$." We must be careful, though, not to read any incorrect implications into these two names. Thus, we should not necessarily "expect" to obtain the value $\langle \hat{A} \rangle_t$ in *any* measurement of $\mathfrak{A}$ on the state $\Psi_t(x)$, because $\langle \hat{A} \rangle_t$ will not necessarily coincide with one of the eigenvalues of $\hat{A}$. In addition, the "uncertainty" described by $\Delta \hat{A}_t$ is not due to a less-than-perfect measuring technique, and measured values which differ significantly from the expectation value are no less legitimate than measured values which are nearly equal to the expectation value. In the final analysis, $\langle \hat{A} \rangle_t$ and $\Delta \hat{A}_t$ are best understood by analogy with $\langle v \rangle$ and Δv in Sec. 2-1: $\langle \hat{A} \rangle_t$, being by definition the average of the values obtained in a series of very many repeated measurements of $\mathfrak{A}$ on the state $\Psi_t(x)$, is a sort of "representative number" for all these values; $\Delta \hat{A}_t$, being by definition the rms deviation of these values, provides a quantitative estimate of their "dispersion," and hence a quantitative estimate of *how adequately* these values are "represented" by the single value $\langle \hat{A} \rangle_t$.

Exercise 30. If $f(z)$ is any real function expandable in a Taylor series, show that $\langle f(\hat{A}) \rangle_t$, the expectation value of $f(\mathfrak{A})$ in the state $\Psi_t(x)$, is given by

$$\langle f(\hat{A}) \rangle_t = \sum_{k=1}^{\infty} |(\alpha_k, \Psi_t)|^2 f(A_k) \tag{4-13}$$

where $\{\alpha_k(x)\}$ and $\{A_k\}$ are the eigenvectors and eigenvalues respectively of the observable operator $\hat{A}$. Compare the form of this result with Eq. (2-8). [*Hint*: To prove Eq. (4-13), apply Eq. (4-11) to the operator $f(\hat{A})$ as defined in Eq. (4-5b) and discussed in Exercise 26.]

It is significant that, according to Eq. (4-11), $\langle\hat{A}\rangle_t$ is uniquely determined by the state vector $\Psi_t(x)$. That is, although we *cannot* generally predict what we might call the "value" of $\mathcal{A}$ in the state $\Psi_t(x)$, we *can* predict the "expectation value" of A in the state $\Psi_t(x)$. Expectation values of observable operators are of practical importance for the following reason: In making measurements in the laboratory, the physicist will often effect a similtaneous measurement of many identical systems (e.g., atoms), each of which is in the *same* state $\Psi_t(x)$. Clearly, if M separate but identical systems, all in the same state $\Psi_t(x)$, are each measured once, the results will be essentially the same as if one of the systems were subjected to M *repeated* measurements in that state. Thus, the expectation value is often a very useful single number to characterize an experimental result.

Equations (4-11) and (4-12) for $\langle\hat{A}\rangle_t$ and $\Delta\hat{A}_t$ are not the most convenient forms for these two quantities. In the next two exercises, we shall derive more useful expressions for the expectation value and the uncertainty.

Exercise 31. By applying Eq. (4-13) to Eq. (4-12), prove that

$$\Delta\hat{A}_t = \sqrt{\langle\hat{A}^2\rangle_t - \langle\hat{A}\rangle_t^2} \tag{4-14}$$

The foregoing expression for the uncertainty is analogous to Eq. (2-6) for Δv, and the remarks made there are applicable here, too. The significance of Eq. (4-14) is that it expresses the uncertainty completely in terms of expectation values. This renders all the more interesting the result of the following exercise.

Exercise 32. Prove that

$$\langle\hat{A}\rangle_t = (\Psi_t, \hat{A}\Psi_t) \tag{4-15}$$

[*Hint*: Expand the inner product $(\Psi_t, \hat{A}\Psi_t)$ as we expanded (Ψ_t, Ψ_t) in our derivation of Eq. (4-8), and so obtain the right-hand side of Eq. (4-11).]

According to Eq. (4-15), we can calculate the expectation value of $\mathcal{A}$ in the state $\Psi_t(x)$ by first forming the vector $\hat{A}\Psi_t(x)$, and then taking the inner product of this vector with the vector $\Psi_t(x)$:

$$\langle\hat{A}\rangle_t = \int_{-\infty}^{\infty} \Psi_t^*(x)[\hat{A}\Psi_t(x)]\,dx \tag{4-16}$$

The virtue of Eq. (4-15) (or Eq. (4-16)) as opposed to Eq. (4-11) is that the former expresses $\langle \hat{A} \rangle_t$ in terms of the operator $\hat{A}$ and the state vector $\Psi_t(x)$ *only*, and does not involve the eigenvectors and eigenvalues of $\hat{A}$. It should also be remarked that Eqs. (4-11) and (4-12) must be modified slightly to accommodate *continuously* distributed eigenvalues, whereas Eqs. (4-15) and (4-14) are valid irrespective of the mode of distribution of the eigenvalues.

It is interesting to note that the reality of $\langle \hat{A} \rangle_t$, which of course is obvious from Eq. (4-11), can be seen from Eq. (4-15) to be a direct consequence of the fact that $\hat{A}$ is Hermitian: thus, using first Eq. (2-34a) and then Eq. (4-4), we find

$$\langle \hat{A} \rangle_t^* = (\Psi_t, \hat{A}\Psi_t)^* = (\hat{A}\Psi_t, \Psi_t) = (\Psi_t, \hat{A}\Psi_t) = \langle \hat{A} \rangle_t$$

which proves that $\langle \hat{A} \rangle_t$ is real.

We can replace $\hat{A}$ in Eq. (4-15) by $f(\hat{A})$, and so obtain

$$\langle f(\hat{A}) \rangle_t = (\Psi_t, f(\hat{A})\Psi_t) \tag{4-17}$$

Applying this to Eq. (4-14), we can write $\Delta \hat{A}_t$ more explicitly as

$$\Delta \hat{A}_t = \sqrt{(\Psi_t, \hat{A}^2 \Psi_t) - (\Psi_t, \hat{A}\Psi_t)^2} \tag{4-18}$$

To summarize the main features of a series of *repeated* measurements, we show in Fig. 3 the sort of results which might be expected if M repeated measurements of $\mathfrak{A}$ are performed on some state $\Psi_t(x)$. On the horizontal axis, we plot the eigenvalues of $\hat{A}$, and above each eigenvalue we draw a bar whose height is equal to the number of times that eigenvalue was measured. Ideally, we expect the eigenvalue A_i to be obtained $|(\alpha_i, \Psi_t)|^2 M$ times, but the actual number of times A_i is obtained will usually differ somewhat from this number owing to the randomness involved. In Fig. 3 we have connected the points $(A_i, |(\alpha_i, \Psi_t)|^2 M)$ with a smooth curve; we call this curve "the *distribution curve* for $\mathfrak{A}$ in the state $\Psi_t(x)$." Clearly, a *complete* specification of the expected results of these repeated measurements requires a specification of all the points $(A_i, |(\alpha_i, \Psi_t)|^2 M)$—i.e., the full distribution curve. A *less detailed* but still very useful description of these results is obtained simply by specifying the values $\langle \hat{A} \rangle_t$ and $\Delta \hat{A}_t$: these two numbers evidently characterize, insofar as it is possible, the "center" and "width," respectively, of the distribution curve.

Exercise 33. Suppose $\Psi_t(x)$ coincides with the eigenvector $\alpha_k(x)$—i.e., $\Psi_t(x) = c\alpha_k(x)$ where $|c|^2 = 1$.

(a) Show that it follows from *both* Eqs. (4-13) and (4-17) that $\langle f(\hat{A}) \rangle_t = f(A_k)$.

(b) Using the result of part (a), show that in this case,

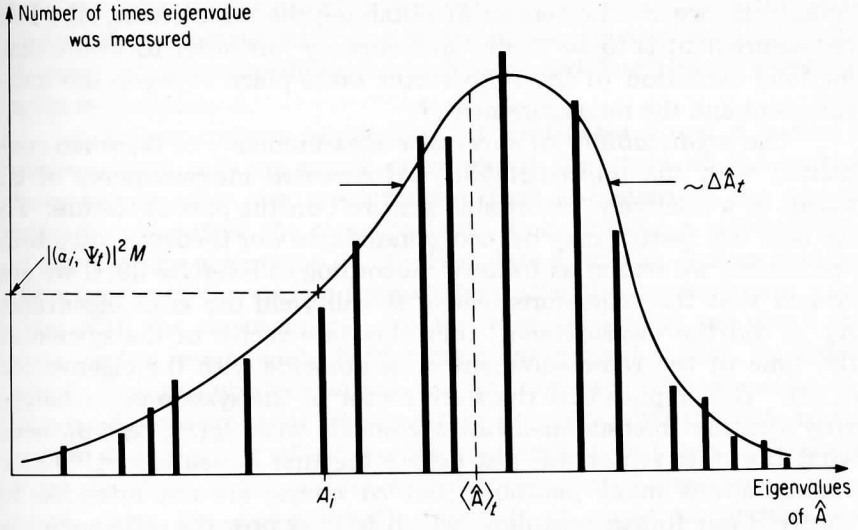

Fig. 3. A bar graph showing the eigenvalues of $\hat{A}$ versus the number of times each eigenvalue was obtained in a hypothetical series of M *repeated* measurements of $\mathcal{A}$ on the state $\Psi_t(x)$. The smooth curve is the "distribution curve for $\mathcal{A}$ in the state $\Psi_t(x)$," and by definition connects the points $(A_i, |(\alpha_i, \Psi_t)|^2 M)$ in a smooth but otherwise arbitrary way. The quantities $\langle\hat{A}\rangle_t$ and $\Delta\hat{A}_t$ measure the mean and rms spread of the distribution curve.

$\langle\hat{A}\rangle_t = A_k$ and $\Delta\hat{A}_t = 0$. Describe the shape of the distribution curve for $\mathcal{A}$ in the state $\alpha_k(x)$.

4-3b The Effect of a Measurement upon the State The "Value of an Observable"

According to Postulate 3, if two *repeated* measurements of some observable $\mathcal{A}$ are made on a system in a given state, the results of these two measurements will *not* usually coincide. If we were actually accustomed to making measurements on the microscopic level, this fact would seem commonplace to us. Equally commonplace would be the following fact: if two *successive* measurements of $\mathcal{A}$ are made on a system, the results *would* always coincide. More specifically, suppose the system is in some state $\Psi_t(x)$, not necessarily an eigenvector of $\hat{A}$, and suppose that a measurement of $\mathcal{A}$ yields the eigenvalue A_k. If we then make an immediate remeasurement of $\mathcal{A}$, without "readjusting" the state vector as we did for repeated mea-

surements, we can be certain of obtaining the same result, A_k. The remeasurement is to be made "immediately" in order to insure that no time evolution of the state vector takes place between the measurement and the remeasurement.†

The *predictability* of *successive* measurements of $\mathfrak{A}$, when contrasted with the *unpredictability* of *repeated* measurements of $\mathfrak{A}$, stands as a relatively "charitable gesture" on the part of Nature. To see how this feature may be incorporated into our theory of quantum mechanics, we reason as follows: According to Exercise 29, if we are *certain* that the remeasurement of $\mathfrak{A}$ will yield the same eigenvalue A_k as did the measurement, then the state vector of the system at the time of the remeasurement *must* coincide with the eigenvector $\alpha_k(x)$. This implies that the state vector of the system immediately *after* the first measurement must coincide with $\alpha_k(x)$, *regardless* of what the state vector was just *before* the first measurement. These considerations make plausible (but of course are not intended to "prove") our fourth postulate, which tells us how the state vector is affected by a measurement.

> **Postulate 4.** A measurement of an observable generally causes a drastic, uncontrollable alteration in the state vector of the system; specifically, regardless of the form of the state vector just *before* the measurement, immediately *after* the measurement it will coincide with the eigenvector corresponding to the eigenvalue obtained in the measurement.‡

Exercise 34. Prove that, as a consequence of Postulates 3 and 4, two *successive* measurements of $\mathfrak{A}$ will necessarily yield identical results, assuming the measurements are performed sufficiently close together in time.

Postulate 4 asserts that a measurement of an observable $\mathfrak{A}$ essentially forces the state vector of the system into an eigenvector of $\hat{A}$. However, it makes no claims regarding the details of the process by which this change in the state vector occurs. Indeed, since it is generally not possible to predict with certainty which eigenvalue will be obtained in a measurement, it follows that it is generally not possible to predict with certainty which eigenvector the state of the system will be forced into by the measurement. All we can say is that the *probability* that a measurement of $\mathfrak{A}$ will force the system

†The time evolution of the state vector is the subject of Postulate 5, and will be discussed in Sec. 4-4.

‡If our development allowed for observable operators with *degenerate* eigenvalues, we would have to state this postulate a bit differently. We shall discuss this point later in Sec. 4-6c.

into the eigenvector $\alpha_k(x)$ is $|(\alpha_k, \Psi_t)|^2$, where $\Psi_t(x)$ is the state vector of the system just prior to the measurement. Thus, the same element of "indeterminism" which pervaded Postulate 3 greets us again in Postulate 4.

A rather curious implication of Postulates 3 and 4 is that a measurement tells us much more about the state of the system immediately *after* the measurement than the state of the system immediately *before* the measurement. For, suppose a measurement of $\mathfrak{A}$ yields the eigenvalue A_k: by Postulate 4, we can deduce that the state vector of the system immediately *after* the measurement coincides with the eigenvector $\alpha_k(x)$; however, all we can say about the state vector immediately *before* the measurement is that, by Postulate 3, its inner product with $\alpha_k(x)$ was nonzero. Thus the outcome of a single measurement tells us rather precisely what the state of the system is *as a result of our having measured it*, but very little about what the state of the system was when we *started* measuring it. In a sense, then, the measurement operation in quantum mechanics is more in the nature of a "preparation" of a state, rather than an "observation" of a state (although we evidently have no real control over which state will be "prepared" in a given measurement).

The reader should now begin to appreciate the really profound difference between the classical and quantum theories of measurement. This difference can be further illuminated by a careful examination of the meaning of the phrase "the value of an observable."

In classical mechanics, we regard an observable as *always* "having a value," and a measurement of an observable simply amounts to taking an unobtrusive peek at what its current value really is. In particular, the value of an observable is presumed to *exist* regardless of whether or not it is *perceived* in a measurement by some observer.

In quantum mechanics, for the particular case in which the state vector $\Psi_t(x)$ coincides with some eigenvector $\alpha_k(x)$ of $\hat{A}$, we can evidently adopt without difficulty the viewpoint of classical mechanics and say that $\mathfrak{A}$ "has the value" A_k. The justification for this is simply that, by Postulate 3, it is absolutely certain that an ideal measurement of $\mathfrak{A}$ will yield the number A_k. We run into difficulties, however, when we consider the more general case in which $\Psi_t(x)$ is a linear combination of two or more eigenvectors of $\hat{A}$:

$$\Psi_t(x) = \sum_i (\alpha_i, \Psi_t)\alpha_i(x) \tag{4-19}$$

In this case, a measurement of $\mathfrak{A}$ can clearly yield *any* value A_i for which the corresponding component (α_i, Ψ_t) is nonzero; therefore the value obtained in a measurement of $\mathfrak{A}$ on this state will be *neither*

predictable nor unique. For this reason, it seems rather pointless, if not meaningless, to try to ascribe some particular "value" to $\mathfrak{A}$ in this state. In fact, the prevailing view among physicists today is that *if the state vector of the system does not coincide with an eigenvector of* $\hat{A}$, *then the corresponding observable* $\mathfrak{A}$ *cannot be said to "have a value" in the generally accepted sense of this phrase.* There are other tenable points of view on this somewhat philosophical matter; however, the foregoing is the "orthodox" view, and therefore is the one that we shall adopt in this introductory treatment of quantum mechanics.

The above conclusion is such a radical departure from our classical way of thinking that we cannot help but look for some easy way of avoiding it. For example, it is tempting to try to interpret the expansion of $\Psi_t(x)$ in Eq. (4-19) as being merely a shorthand way of saying that "the state vector of the system *really* coincides with *one particular* eigenvector $\alpha_i(x)$, but owing to a lack of information we can say only that $|(\alpha_j,\Psi_t)|^2$ is the *probability* that this particular eigenvector is $\alpha_j(x)$." Now, if this interpretation of Eq. (4-19) were legitimate, then $\mathfrak{A}$ could indeed be said to "have a value"; we just wouldn't know for sure what the value really is, since we don't know for sure with which eigenvector $\Psi_t(x)$ really coincides. However, according to the orthodox view of quantum mechanics, this is *not* a correct interpretation of the expansion in Eq. (4-19). For, this linear combination is a perfectly legitimate normed $\mathcal{H}$-vector, and so by Postulate 1 defines a physical state of the system just as precisely and completely as does any one of the eigenvectors of $\hat{A}$. Moreover, if it so happened that this linear combination coincided with some eigenvector $\beta_k(x)$ of some other observable operator $\hat{B}$ (in which case $\mathfrak{B}$ could be said to "have the value B_k"), then it would be *inconsistent* to say that the state vector *really* coincides with one of the eigenvectors $\alpha_i(x)$. We see, then, that within the framework of the four postulates which we have laid down here, there seems to be no simple, satisfactory way around the conclusion that an observable does *not always* "have a value."

Generally speaking, then, a measurement of $\mathfrak{A}$ in quantum mechanics is *not* simply a matter of "taking an unobtrusive peek at the value of $\mathfrak{A}$." For, if the state vector of the system does *not* coincide with one of the eigenvectors of $\hat{A}$, then (i) the measurement cannot be "unobtrusive," since the state vector will necessarily be altered by the measurement, and (ii) a "value of $\mathfrak{A}$" does not even exist in the usual sense of the phrase, since the result of the measurement is not uniquely predetermined. Perhaps a more accurate de-

scription of the measurement process is to say that *the very act of measuring* $\mathfrak{A}$ *essentially develops a value of* $\mathfrak{A}$; it evidently accomplishes this by the simple expedient of forcing the state vector into one of the eigenvectors of $\hat{A}$, so that *then* $\mathfrak{A}$ will indeed have a value. In this view, $|(\alpha_k, \Psi_t)|^2$ is the probability that a measurement of $\mathfrak{A}$ on the state $\Psi_t(x)$ will *develop* the value A_k. But, to repeat ourselves, this is *not* the same as saying that $|(\alpha_k, \Psi_t)|^2$ is the probability that $\mathfrak{A}$ *has* the value A_k in the state $\Psi_t(x)$. Regardless of one's point of view on this somewhat philosophical matter, it clearly is usually safer to speak of the "value obtained in a measurement of $\mathfrak{A}$" rather than the "value of $\mathfrak{A}$."

We shall agree, then, that it is *strictly* legitimate to say that $\mathfrak{A}$ "has a value" in the state $\Psi_t(x)$ if and only if a measurement of $\mathfrak{A}$ on this state is certain to yield a definite result—i.e., if and only if $\Psi_t(x)$ coincides with an eigenvector of $\hat{A}$. Suppose it happens that $\Psi_t(x)$ "almost" coincides with an eigenvector of $\hat{A}$; more specifically, suppose that in the expansion of $\Psi_t(x)$ in the eigenbasis of $\hat{A}$, [Eq. (4-19)], the coefficient of one particular eigenvector $\alpha_k(x)$ strongly dominates the coefficients of all the other eigenvectors, so that in accordance with Eq. (4-7b),

$$|(\alpha_i, \Psi_t)|^2 \begin{cases} \ll 1 & \text{for } i \neq k \\ \lesssim 1 & \text{for } i \neq k \end{cases}$$

Exercise 35. If $\Psi_t(x)$ is such that the above conditions are satisfied prove that:

(a) The distribution curve [see Fig. 3] for $\mathfrak{A}$ in this state has a sharp, narrow peak at the eigenvalue A_k.

(b) A measurement of $\mathfrak{A}$ on this state is *almost* certain to yield a definite result.

In view of conclusion (b) in the above exercise, it is tempting to say that $\mathfrak{A}$ *almost* has a value in the state $\Psi_t(x)$. The strength of the qualifier "almost" is evidently governed by the width of the distribution curve for $\mathfrak{A}$ in this state. Since this width is proportional to $\Delta\hat{A}_t$, we may therefore regard the uncertainty in $\mathfrak{A}$ in the state $\Psi_t(x)$ as indicating the extent to which $\mathfrak{A}$ can be said to "have a value" in this state: the *smaller* $\Delta\hat{A}_t$ is, the *more* sense it makes to say $\mathfrak{A}$ "has a value," and the *larger* $\Delta\hat{A}_t$ is, the *less* sense it makes to say this. This rather loose way of speaking will prove useful in interpreting some of our subsequent results.

We shall conclude our general discussion of the quantum theory

of measurement in the next section, where we shall derive two important theorems pertaining to the measurement of *two* observables, $\mathfrak{A}$ and $\mathfrak{B}$.

4-3c The Compatibility Theorem and the Heisenberg Uncertainty Principle

We come now to consider one of the most interesting and important topics in quantum mechanics, namely, the problem of the "simultaneous measurability" or "compatibility" of two observables. Let us begin by explaining precisely what we mean by these terms.

Suppose a given system is subjected to three *successive* measurements involving two observables, $\mathfrak{A}$ and $\mathfrak{B}$: the first measurement, denoted by $M_{\mathfrak{A}}$, measures $\mathfrak{A}$; the second measurement, denoted by $M_{\mathfrak{B}}$, measures $\mathfrak{B}$; and finally the third measurement, denoted by $M'_{\mathfrak{A}}$, measures $\mathfrak{A}$ again. It is of course understood that these measurements are to be performed in very rapid succession so that there is no significant time evolution of the state of the system between $M_{\mathfrak{A}}$ and $M_{\mathfrak{B}}$, and between $M_{\mathfrak{B}}$ and $M'_{\mathfrak{A}}$. With respect to these measurements, we now make the following definition: the observables $\mathfrak{A}$ and $\mathfrak{B}$ are said to be *simultaneously measurable* or *compatible* if and only if the result of $M'_{\mathfrak{A}}$ is *certain* to coincide with the result of $M_{\mathfrak{A}}$, regardless of what the state of the system was just prior to $M_{\mathfrak{A}}$.

From the standpoint of classical mechanics this definition is rather useless, because in classical mechanics *all* pairs of observables are "compatible": $\mathfrak{A}$ and $\mathfrak{B}$ each "have a value" at all times, and since the (ideal) measurement $M_{\mathfrak{B}}$ will have no effect upon the value of $\mathfrak{A}$, it follows that $M_{\mathfrak{A}}$ and $M'_{\mathfrak{A}}$ will always yield the same number.

From the standpoint of quantum mechanics, though, we can easily see that $\mathfrak{A}$ and $\mathfrak{B}$ might very well fail to satisfy our condition for being simultaneously measurable: We denote by $\hat{A}$ and $\hat{B}$ the observable operators for $\mathfrak{A}$ and $\mathfrak{B}$, and we denote by $\{\alpha_i(x)\}$, $\{\beta_i(x)\}$ and $\{A_i\}$, $\{B_i\}$ the associated eigenbases and eigenvalues of these operators.

$$\left.\begin{aligned} \hat{A}\alpha_i(x) &= A_i\alpha_i(x) \quad i = 1,2,\ldots \\ \hat{B}\beta_i(x) &= B_i\beta_i(x) \quad i = 1,2,\ldots \end{aligned}\right\} \tag{4-20}$$

Suppose that $M_{\mathfrak{A}}$ yields the value A_n, and $M_{\mathfrak{B}}$ then yields the value B_m. By Postulate 4, the state vector just after $M_{\mathfrak{B}}$, or just before $M'_{\mathfrak{A}}$, coincides with $\beta_m(x)$. Now, it is not necessarily the case that $\beta_m(x)$

coincides with $\alpha_n(x)$; if it does not, then by Postulate 3 (or Exercise 29) $M'_{\mathfrak{A}}$ will not *necessarily* yield the value A_n, and therefore $\mathfrak{A}$ and $\mathfrak{B}$ are not "compatible."

If, in the foregoing example, the second measurement $M_{\mathfrak{B}}$ had not been performed, then the results of $M_{\mathfrak{A}}$ and $M'_{\mathfrak{A}}$ would obviously have been the same [see Exercise 34]. This implies that $M_{\mathfrak{B}}$ always has the potential of "spoiling" the remeasurement, $M'_{\mathfrak{A}}$. It must be emphasized that this spoilage, if it occurs, is *not* the result of an imperfect measuring technique, but rather follows as a simple, direct consequence of Postulates 3 and 4.

We shall now state and prove one of the fundamental theorems of quantum mechanics, which we shall call the "Compatibility Theorem." This theorem essentially provides us with *two* conditions, *either* of which is both necessary and sufficient for $\mathfrak{A}$ and $\mathfrak{B}$ to be compatible observables. The proof of this theorem is in itself an illuminating exercise in the application of Postulates 3 and 4. Our proof here will be subject to our continuing restriction that $\hat{A}$ and $\hat{B}$ have nondegenerate eigenvalues [see Eq. (4-9)]; the theorem is actually valid without this restriction, but certain parts of the proof are more complicated.†

The Compatibility Theorem. Given two observables $\mathfrak{A}$ and $\mathfrak{B}$ with corresponding operators $\hat{A}$ and $\hat{B}$, then any one of the following three conditions implies the other two:

(i) $\mathfrak{A}$ and $\mathfrak{B}$ are compatible observables.
(ii) $\hat{A}$ and $\hat{B}$ possess a common eigenbasis.
(iii) $\hat{A}$ and $\hat{B}$ commute.

Proof: Our proof will consist in showing that (i) and (ii) imply each other, and (ii) and (iii) imply each other. The fact that (i) and (iii) imply each other will then follow trivially.

(i) *implies* (ii): Suppose that, just prior to $M_{\mathfrak{A}}$, the state vector of the system coincides with *any* eigenvector $\alpha_i(x)$ of $\hat{A}$. Then $M_{\mathfrak{A}}$ will yield the value A_i [by Exercise 29]. When $M_{\mathfrak{B}}$ is performed, the state vector will become coincident with *some* eigenvector $\beta_j(x)$ of $\hat{B}$ [by Postulate 4]. Now since $\mathfrak{A}$ and $\mathfrak{B}$ are given to be compatible, then it is certain that the third measurement $M'_{\mathfrak{A}}$ must yield the same value A_i that was obtained in the first measurement. But by Exercise 29, if a measurement of $\mathfrak{A}$ performed on the state $\beta_j(x)$ is certain to yield the eigenvalue A_i, then $\beta_j(x)$ must coincide with $\alpha_i(x)$. We have proved then that *any* vector of the set $\{\alpha_i(x)\}$ coincides with *some*

†For a general proof of the Compatibility Theorem, see Chapter IV of F. Mandl, *Quantum Mechanics*, cited in footnote in Preface.

vector of the set $\{\beta_j(x)\}$. Since these sets are orthonormal basis sets, the correspondence must be one-to-one;† thus, simply by rearranging indices we can put $\{\alpha_n(x)\} = \{\beta_n(x)\} = \{\phi_n(x)\}$, where $\{\phi_n(x)\}$ is a "common eigenbasis" for $\hat{A}$ and $\hat{B}$:

$$\left.\begin{aligned} \hat{A}\phi_n(x) &= A_n\phi_n(x) \qquad n = 1,2,\ldots \\ \hat{B}\phi_n(x) &= B_n\phi_n(x) \qquad n = 1,2,\ldots \end{aligned}\right\} \tag{4-21}$$

(ii) *implies* (i): Given the common eigenbasis $\{\phi_n(x)\}$ as in Eqs. (4-21), measurement $M_{\mathfrak{A}}$, with any result A_n, leaves the system in the state $\phi_n(x)$, by Postulate 4. Then by Exercise 29, $M_{\mathfrak{B}}$ must yield the value B_n, and what is more important, will *leave* the system in the state $\phi_n(x)$. Hence, by Exercise 29, $M'_{\mathfrak{A}}$ must yield the value A_n again. Since these arguments are independent of which eigenvalue A_n was obtained in the first measurement, it therefore holds regardless of what state the system was in just prior to that measurement; consequently, $\mathfrak{A}$ and $\mathfrak{B}$ are compatible observables.

(ii) *implies* (iii): Given the common eigenbasis $\{\phi_n(x)\}$ as in Eqs. (4-21), then using the linearity of $\hat{A}$ and $\hat{B}$ we have

$$\hat{A}\hat{B}\phi_n(x) = \hat{A}B_n\phi_n(x) = B_n\hat{A}\phi_n(x) = B_nA_n\phi_n(x)$$

$$\hat{B}\hat{A}\phi_n(x) = \hat{B}A_n\phi_n(x) = A_n\hat{B}\phi_n(x) = A_nB_n\phi_n(x)$$

Thus, $(\hat{A}\hat{B} - \hat{B}\hat{A})\phi_n(x) = 0$. Now we are not yet done, for in order to show that $\hat{A}$ and $\hat{B}$ commute, we must show that $(\hat{A}\hat{B} - \hat{B}\hat{A})\psi(x) = 0$ for *any* $\mathcal{H}$-vector $\psi(x)$. To this end, we first expand the given $\mathcal{H}$-vector $\psi(x)$ in the eigenbasis $\{\phi_n(x)\}$:

$$\psi(x) = \sum_n c_n\phi_n(x)$$

where $c_n \equiv (\phi_n, \psi)$ [see Eq. (2-39)]. Then, using the fact that both the product and the sum of two linear operators are also linear opera-

† This is "almost obvious." We have shown that *each* α-vector coincides with *some* β-vector. Now, two α-vectors cannot coincide with the *same* β-vector, since any two α-vectors are orthogonal; similarly, two β-vectors cannot coincide with the same α-vector. It remains only to show that no β-vector is "missed" in the α-to-β correspondence. To prove this, assume the contrary: some β-vector, $\beta_k(x)$, does not coincide with any of the α-vectors. Then the expansion of $\beta_k(x)$ in the α-basis, $\beta_k(x) = \Sigma_i(\alpha_i,\beta_k)\alpha_i(x)$, must contain *at least two* non-vanishing terms, thus implying that $\beta_k(x)$ is nonorthogonal to at least two α-vectors. But since each α-vector coincides with a unique β-vector, then we must conclude that $\beta_k(x)$ is nonorthogonal to at least two β-vectors—a conclusion which clearly contradicts the orthonormality of the β-vectors. Therefore the assumption is false, and the α-β correspondence is indeed one-to-one.

tors, we find

$$(\hat{A}\hat{B} - \hat{B}\hat{A})\psi(x) = \sum_n c_n(\hat{A}\hat{B} - \hat{B}\hat{A})\phi_n(x) = \sum_n c_n \cdot 0 = 0$$

Thus the operator $\hat{A}\hat{B}$ acting on any $\mathcal{H}$-vector $\psi(x)$ produces the same vector as does the operator $\hat{B}\hat{A}$ acting on $\psi(x)$. This means that $\hat{A}$ and $\hat{B}$ commute.

(iii) *implies* (ii): Given that $\hat{A}$ and $\hat{B}$ commute, then for *any* eigenvector $\alpha_i(x)$ of $\hat{A}$, we have

$$\hat{A}\hat{B}\alpha_i(x) = \hat{B}\hat{A}\alpha_i(x) = \hat{B}A_i\alpha_i(x) = A_i\hat{B}\alpha_i(x)$$

In other words, $\hat{A}$ operating on the $\mathcal{H}$-vector $\hat{B}\alpha_i(x)$ has the effect of simply multiplying this vector by the number A_i; this implies that the vector $\hat{B}\alpha_i(x)$ is an eigenvector of $\hat{A}$ belonging to the eigenvalue A_i. Since the eigenvalues of $\hat{A}$ are taken to be nondegenerate, then the vector $\hat{B}\alpha_i(x)$ can differ from the vector $\alpha_i(x)$ by at most a scalar factor c, which merely takes account of the fact that the norms of $\alpha_i(x)$ and $\hat{B}\alpha_i(x)$ need not be equal:

$$\hat{B}\alpha_i(x) = c\alpha_i(x)$$

But this equation implies that $\alpha_i(x)$ is a (normed) eigenvector of $\hat{B}$ belonging to the eigenvalue c. Since the eigenvalues of $\hat{B}$ are nondegenerate, then we have for some j, $c = B_j$ and $\alpha_i(x) = \beta_j(x)$. We have proved then that *any* vector of the set $\{\alpha_i(x)\}$ coincides with *some* vector of the set $\{\beta_j(x)\}$. Since these sets are orthonormal basis sets, the correspondence must be one-to-one,† so simply by rearranging indices we can put $\{\alpha_n(x)\} = \{\beta_n(x)\} = \{\phi_n(x)\}$, where $\{\phi_n(x)\}$ is the "common eigenbasis" set in Eq. (4-21).

Q.E.D.

As a simple illustration of this theorem, we recall from Exercise 26 that, for any reasonable function f, the operator $f(\hat{A})$ is an observable operator which has the same eigenbasis as the observable operator $\hat{A}$. Therefore, the Compatibility Theorem tells us that the operators $\hat{A}$ and $f(\hat{A})$ must *commute*, as may easily be verified from Eq. (4-5b), and also that the observables $\mathfrak{A}$ and $f(\mathfrak{A})$ must be *compatible*, which seems eminently reasonable.

Exercise 36.

(a) Using the power series expansion of $f(\hat{A})$ in Eq. (4-5b), prove directly that $\hat{A}$ and $f(\hat{A})$ commute.

(b) Using the fact that $\hat{A}$ and $f(\hat{A})$ have the same eigenbasis $\{\alpha_i(x)\}$, prove directly that $\mathfrak{A}$ and $f(\mathfrak{A})$ are compatible.

†See preceding footnote.

If $\hat{A}$ and $\hat{B}$ do *not* possess a common eigenbasis, then according to the Compatibility Theorem two successive measurements $M_{\mathfrak{A}}$ and $M'_{\mathfrak{A}}$ of $\mathfrak{A}$, when separated by a measurement $M_{\mathfrak{B}}$ of $\mathfrak{B}$, will *not always* yield identical results. However, for a given result of $M_{\mathfrak{A}}$, it *is* possible to make *probabilistic* predictions about the outcome of $M'_{\mathfrak{A}}$. To do this, one first writes down the expansion of the $\hat{A}$-eigenbasis in terms of the $\hat{B}$-eigenbasis and vice versa:

$$\alpha_k(x) = \sum_{i=1}^{\infty} (\beta_i, \alpha_k)\beta_i(x)$$

$$\beta_k(x) = \sum_{i=1}^{\infty} (\alpha_i, \beta_k)\alpha_i(x)$$

Using these expansions, along with Postulates 3 and 4, one can calculate the probabilities for all possible "routes" which lead from the given result of $M_{\mathfrak{A}}$ to any particular result of $M'_{\mathfrak{A}}$. Thus, suppose $M_{\mathfrak{A}}$ yields A_n, and we wish to know the probability that $M'_{\mathfrak{A}}$ will yield A_m. We reason as follows: After $M_{\mathfrak{A}}$ yields A_n, the system is in the state $\alpha_n(x)$, and the probability that $M_{\mathfrak{B}}$ will yield some result B_i is $|(\beta_i, \alpha_n)|^2$. If B_i is obtained on the second measurement, the system will be in the state $\beta_i(x)$ for $M'_{\mathfrak{A}}$, so the probability that $M'_{\mathfrak{A}}$ will then yield the desired result A_m is $|(\alpha_m, \beta_i)|^2$. Thus, according to Eq. (2-3b), the product $|(\beta_i, \alpha_n)|^2 \, |(\alpha_m, \beta_i)|^2$ gives the probability that the value A_m will be obtained in $M'_{\mathfrak{A}}$ *via* the result B_i for $M_{\mathfrak{B}}$. Using Eq. (2-3a), we conclude that the sum $\Sigma_i |(\beta_i, \alpha_n)|^2 \, |(\alpha_m, \beta_i)|^2$ is the probability that, given the result A_n of $M_{\mathfrak{A}}$, $M'_{\mathfrak{A}}$ will yield the value A_m *regardless* of the result of $M_{\mathfrak{B}}$. The following exercise will illustrate this procedure more explicitly.

Exercise 37. Suppose $\hat{A}$ and $\hat{B}$ "almost" possess a common eigenbasis; more specifically, suppose that when the eigenvectors of $\hat{B}$ are expanded in terms of the eigenvectors of $\hat{A}$, one has

$$\left.\begin{aligned} \beta_1(x) &= \frac{\sqrt{3}}{2}\alpha_1(x) + \frac{1}{2}\alpha_2(x) \\ \beta_2(x) &= \frac{1}{2}\alpha_1(x) - \frac{\sqrt{3}}{2}\alpha_2(x) \\ \beta_n(x) &= \alpha_n(x), \qquad n \geq 3 \end{aligned}\right\}$$

(a) Verify that this expansion is consistent with the orthornormality of $\{\alpha_i(x)\}$ and $\{\beta_i(x)\}$; i.e., prove that if $(\alpha_i, \alpha_j) = \delta_{ij}$, then $(\beta_i, \beta_j) = \delta_{ij}$ also.

(b) Expand the eigenvectors of $\hat{A}$ in terms of the eigenvectors of $\hat{B}$.

(c) Prove that if $M_{\mathfrak{A}}$ yields any one of the values $A_3, A_4, \ldots,$ then $M'_{\mathfrak{A}}$ will necessarily yield the same result.

(d) Prove that if $M_{\mathfrak{A}}$ yields the value A_1, then there is a 5/8 probability that $M'_{\mathfrak{A}}$ will yield A_1 and a 3/8 probability that $M'_{\mathfrak{A}}$ will yield A_2.

If two observables $\mathfrak{A}$ and $\mathfrak{B}$ are incompatible, then the Compatibility Theorem tells us that their corresponding operators $\hat{A}$ and $\hat{B}$ do not commute; that is, for at least one vector $\psi(x)$ in $\mathcal{H}$, $(\hat{A}\hat{B} - \hat{B}\hat{A})\,\psi(x) \neq 0$. Now it turns out that the noncommutability of many pairs of noncommuting observable operators can be expressed by an equation of the form

$$\hat{A}\hat{B} - \hat{B}\hat{A} = c \tag{4-22}$$

where c is some nonzero scalar; Eq. (4-22) means simply that, for *any* $\mathcal{H}$-vector $\psi(x)$, $(\hat{A}\hat{B} - \hat{B}\hat{A})\psi(x) = c\psi(x)$. For such a case, the inherent incompatibility of $\mathfrak{A}$ and $\mathfrak{B}$ is strikingly illustrated by the famous "Uncertainty Principle," which was first ennunciated by W. Heisenberg:

The Heisenberg Uncertainty Principle. If $\hat{A}$ and $\hat{B}$ are such that $\hat{A}\hat{B} - \hat{B}\hat{A} = c$, where c is a scalar, then the uncertainties in $\mathfrak{A}$ and $\mathfrak{B}$ in *any* state $\Psi_t(x)$ satisfy

$$\Delta\hat{A}_t \cdot \Delta\hat{B}_t \geqq \frac{1}{2}\,|c| \tag{4-23}$$

Before presenting the proof of this theorem (the proof is purely mathematical and involves no physical arguments), let us first point out its profound physical implications.

According to the Heisenberg Uncertainty Principle, if two observable operators $\hat{A}$ and $\hat{B}$ satisfy Eq. (4-22) with $c \neq 0$, then the product of the uncertainties in $\mathfrak{A}$ and $\mathfrak{B}$ in any state $\Psi_t(x)$ is strictly bounded away from zero. Thus, if we somehow contrive to force the system into states having smaller and smaller uncertainties in $\mathfrak{A}$, then these states will necessarily have larger and larger uncertainties in $\mathfrak{B}$—and vice versa. In view of our discussion at the end of Sec. 4-3b, we can also express these conclusions in the following way: If $\hat{A}$ and $\hat{B}$ satisfy Eq. (4-22) with $c \neq 0$, then Eq. (4-23) implies that the *more* sense it makes to say that $\mathfrak{A}$ "has a value" in a given state, the *less* sense it makes to say that $\mathfrak{B}$ "has a value" in that state—and vice versa. Furthermore, if it so happens that $\Psi_t(x)$ coincides with one of the eigenvectors of $\hat{A}$, so that $\Delta\hat{A}_t = 0$ and $\mathfrak{A}$ therefore *really* has a value, then Eq. (4-23) evidently requires that $\Delta\hat{B}_t = \infty$, in which case it would be *completely meaningless* to speak of $\mathfrak{B}$ as having a value.

In view of these implications, it is not surprising that the Heisenberg Uncertainty Principle occupies a prominent place in the overall structure of quantum mechanics. For example, we shall see in Sec. 4-5c that it leads to a satisfactory resolution of the "wave-particle paradox."

We shall now conclude our development of the quantum theory of measurement by proving the Heisenberg Uncertainty Principle. Although the proof entails a considerable amount of mathematical manipulation, the reader should note that it relies mainly upon (i) the definition of Hermiticity in Eq. (4-4), (ii) the mathematical expression for the uncertainty in Eq. (4-14), and (iii) the properties of the inner product of two $\mathcal{H}$-vectors in Eqs. (2-34) (particularly the Schwarz inequality).

To simplify our notation, we shall omit the subscript t in what follows, keeping in mind the fact that all our calculations hold at any one instant of time. Using the expressions

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

for the uncertainties in $\mathcal{A}$ and $\mathcal{B}$ in the state $\Psi(x)$, we shall first prove the *generalized uncertainty relation:*

$$\Delta\hat{A} \cdot \Delta\hat{B} \geqq \frac{1}{2} |(\Psi, [\hat{A}\hat{B} - \hat{B}\hat{A}]\ \Psi)| \tag{4-24}$$

From this generally valid relation, we may easily obtain Eq. (4-23) by inserting Eq. (4-22) and then using the fact that $(\Psi, \Psi) = 1$.

To prove Eq. (4-24), we first define two operators $\hat{A}'$ and $\hat{B}'$ by

$$\hat{A}' \equiv \hat{A} - \langle\hat{A}\rangle \quad \text{and} \quad \hat{B}' \equiv \hat{B} - \langle\hat{B}\rangle$$

Concerning these operators, we next prove three lemmas.

Lemma 1. $\hat{A}'$ and $\hat{B}'$ are Hermitian operators.

Exercise 38. Prove Lemma 1. [*Hint*: Using the fact that $\hat{A}'$ is the difference of two Hermitian operators, show that $(\psi_1, \hat{A}'\psi_2) = (\hat{A}'\psi_1, \psi_2)$ for any two Hilbert space vectors $\psi_1(x)$ and $\psi_2(x)$.]

Lemma 2. $\hat{A}'\hat{B}' - \hat{B}'\hat{A}' = \hat{A}\hat{B} - \hat{B}\hat{A}$

Exercise 39. Prove Lemma 2. [*Hint:* Note, for example, that $\hat{A}\langle\hat{B}\rangle = \langle\hat{B}\rangle\hat{A}$, since $\hat{A}$ is a *linear* operator and $\langle\hat{B}\rangle$ is a *scalar*.]

Lemma 3. $(\hat{A}'\Psi, \hat{A}'\Psi) = (\Delta\hat{A})^2$

The proof of Lemma 3 goes as follows:

$$\begin{aligned}(\hat{A}'\Psi,\hat{A}'\Psi) &= (\Psi,\hat{A}'^2\Psi)\\ &= (\Psi,[\hat{A}-\langle\hat{A}\rangle]^2\Psi)\\ &= (\Psi,[\hat{A}^2-2\langle\hat{A}\rangle\hat{A}+\langle\hat{A}\rangle^2]\Psi)\\ &= (\Psi,\hat{A}^2\Psi)-2\langle\hat{A}\rangle(\Psi,\hat{A}\Psi)+\langle\hat{A}\rangle^2(\Psi,\Psi)\\ &= \langle\hat{A}^2\rangle-2\langle\hat{A}\rangle\langle\hat{A}\rangle+\langle\hat{A}\rangle^2\cdot 1\\ &= \langle\hat{A}^2\rangle-\langle\hat{A}\rangle^2\\ \therefore(\hat{A}'\Psi,\hat{A}'\Psi) &= (\Delta\hat{A})^2\end{aligned}$$

The first step in the above proof makes use of the Hermiticity of $\hat{A}'$, which was established in Lemma 1. The remaining steps invoke the definitions of $\hat{A}'$, $\langle\hat{A}\rangle$ and $\Delta\hat{A}$, as well as the mathematical properties expressed in Eqs. (2-34) and (2-42).†

With these three lemmas, the proof of the generalized uncertainty relation in Eq. (4-24) goes as follows:

$$\begin{aligned}(\Psi,[\hat{A}\hat{B}-\hat{B}\hat{A}]\Psi) &= (\Psi,[\hat{A}'\hat{B}'-\hat{B}'\hat{A}']\Psi) && \text{[by Lemma 2]}\\ &= (\Psi,\hat{A}'\hat{B}'\Psi)-(\Psi,\hat{B}'\hat{A}'\Psi)\\ &= (\hat{A}'\Psi,\hat{B}'\Psi)-(\hat{B}'\Psi,\hat{A}'\Psi) && \text{[by Lemma 1]}\\ &= (\hat{A}'\Psi,\hat{B}'\Psi)-(\hat{A}'\Psi,\hat{B}'\Psi)^*\end{aligned}$$

so

$$(\Psi,[\hat{A}\hat{B}-\hat{B}\hat{A}]\Psi) = 2i\,\mathrm{Im}\,(\hat{A}'\Psi,\hat{B}'\Psi)$$

In the last two steps we have made use of Eqs. (2-34a) and (2-13), respectively. We now take the modulus of both sides of this equation. Using the fact that $|i| = 1$, along with Eq. (2-18a), we obtain

$$|(\Psi,[\hat{A}\hat{B}-\hat{B}\hat{A}]\Psi)| = 2|\mathrm{Im}\,(\hat{A}'\Psi,\hat{B}'\Psi)|$$

According to Eqs. (2-17), the magnitude of the imaginary part of a complex number is never greater than the modulus of the complex number, so we may write

$$|(\Psi,[\hat{A}\hat{B}-\hat{B}\hat{A}]\Psi)| \leqq 2|(\hat{A}'\Psi,\hat{B}'\Psi)|$$

†The expression for $\Delta\hat{A}$ in Lemma 3 allows a very simple proof of the fact that *the vanishing of* $\Delta\hat{A}_t$ *is a necessary and sufficient condition for* $\Psi_t(x)$ *to be coincident with an eigenvector of* $\hat{A}$. For, according to Lemma 3, $\Delta\hat{A}_t$ vanishes if and only if the vector $\hat{A}'\Psi_t(x) \equiv (\hat{A}-\langle\hat{A}\rangle_t)\Psi_t(x)$ has zero norm. But the only vector with zero norm is the null vector, $f(x) \equiv 0$. Therefore, $\Delta\hat{A}_t = 0$ if and only if $\hat{A}\Psi_t(x) \equiv \langle\hat{A}\rangle_t\Psi_t(x)$—i.e., if and only if $\Psi_t(x)$ is an eigenvector of $\hat{A}$. We had previously inferred the *sufficiency* condition only indirectly through Exercises 29 and 33.

We now write the Schwarz inequality, Eq. (2-34d), with $\psi_1(x) = \hat{A}'\Psi(x)$ and $\psi_2(x) = \hat{B}'\Psi(x)$:

$$|(\hat{A}'\Psi, \hat{B}'\Psi)| \leqq \sqrt{(\hat{A}'\Psi, \hat{A}'\Psi)} \cdot \sqrt{(\hat{B}'\Psi, \hat{B}'\Psi)}$$

Applying this to the previous inequality, we obtain

$$|(\Psi, [\hat{A}\hat{B} - \hat{B}\hat{A}]\Psi)| \leqq 2\sqrt{(\hat{A}'\Psi, \hat{A}'\Psi)} \cdot \sqrt{(\hat{B}'\Psi, \hat{B}'\Psi)}$$

or, by virtue of Lemma 3,

$$|(\Psi, [\hat{A}\hat{B} - \hat{B}\hat{A}]\Psi)| \leqq 2\Delta\hat{A} \cdot \Delta\hat{B}$$

This last equation is just Eq. (4-24), and so the proof is complete.

4-4 TIME EVOLUTION OF THE QUANTUM STATE

In our review of classical mechanics in Chapter 3 we considered at some length the problem of the *time evolution* of the classical state, but we did not deem it necessary to dwell very long on the rather obvious *definition* of the classical state. By contrast, our comparatively lengthy sojourn into quantum mechanics thus far has been almost exclusively concerned with describing precisely what is meant by the "state of a system" in quantum mechanics—a task which we have found to be anything but trivial. For, since quantum mechanics makes a radical, seemingly perverse distinction between the *state* of a system and the physical *observables*, it was necessary first to define each of these concepts separately, and then, using the quantum theory of measurement, to carefully delineate the subtle relationship between the two. Fortunately, however, once we arrive at a reasonably good understanding of these concepts, it is not too difficult to comprehend what the theory of quantum mechanics has to say about how a system behaves with time. As we shall see, in this respect quantum mechanics has much in common with the classical approach, in that it provides us with well-defined "equations of motion" for the state vector $\Psi_t(x)$, and also for the expectation value $\langle\hat{A}\rangle_t$ and the probability coefficients (α_k, Ψ_t) for a given observable $\mathfrak{A}$. In this section we shall obtain these equations of motion and discuss some of their important implications.

In order that our development will be valid for *any* physical system with one degree of freedom, we shall continue to refer to observables "in general" by $\mathfrak{A}, \mathfrak{B}$, etc. In Sec. 4-5 we shall exhibit the specific forms of the relevant observable operators for the particular case of a mass m on the x-axis (this will constitute our Postulate 6), and we shall then apply our general theory to that particular problem in some detail.

4-4a Energy and the Hamiltonian Operator The Time Evolution Equation for $\Psi_t(x)$

In our discussion of the time-evolution problem in classical mechanics, we found that there was a very intimate connection between *energy* and *time*. Thus, while on the one hand the observable "energy" was unique in that it maintained a constant value in time [see Exercise 22], it turned out that it was the energy written as the Hamiltonian function $H(x,p)$ that actually governed the time development of the system [see Eqs. (3-10)]. Keeping this in mind, we begin our discussion of the time-evolution problem in quantum mechanics by presenting the postulate which tells us exactly how the state vector of a system changes with time.

Postulate 5. For every physical system there exists a linear, Hermitian operator $\hat{H}$, called the *Hamiltonian operator*, which has the following properties:

(a) The Hamiltonian operator $\hat{H}$ is the observable operator corresponding to the total energy of the system. Hence, $\hat{H}$ possesses a complete, orthonormal set of eigenvectors $\{\eta_k(x)\}$ and a corresponding set of real eigenvalues $\{E_k\}$,

$$\hat{H}\eta_k(x) = E_k\eta_k(x) \qquad k = 1, 2, \ldots \tag{4-25}$$

where the numbers $\{E_k\}$ are the allowed values of the total energy of the system.

(b) The Hamiltonian operator $\hat{H}$ determines the time evolution of the state vector of the system, $\Psi_t(x) \equiv \Psi(x,t)$, through the differential equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x,t) = \hat{H}\Psi(x,t) \tag{4-26}$$

provided the system is not disturbed. The constant $\hbar$ is called "h-bar," and has the value

$$\hbar \equiv h/2\pi = 1.054 \times 10^{-34} \text{ joule} \cdot \text{sec} \tag{4-27}$$

According to this postulate, if $\Psi_0(x)$ is the state vector of the system at time $t = 0$, then (provided the system is not intruded upon by some external agency) as t assumes successive values $t_1, t_2, \ldots,$ the state vector successively coincides with the $\mathcal{H}$-vectors $\Psi_{t_1}(x) \equiv \Psi(x,t_1)$, $\Psi_{t_2}(x) \equiv \Psi(x,t_2), \ldots,$ where $\Psi(x,t)$ is that solution of Eq. (4-26) which satisfies the "initial condition" $\Psi(x,0) \equiv \Psi_0(x)$. Consequently, the state vector of the system evolves with time in a completely deterministic way, just as $x(t)$ and $p(t)$ do in classical me-

chanics. This is true so long as the system is not disturbed, and in this connection we must emphasize that, according to Postulate 4, *the measurement process represents such a disturbance* which alters the otherwise orderly time development of the state vector.

We may now trace the following interesting parallelism between classical mechanics and quantum mechanics: In {classical | quantum} mechanics, once the Hamiltonian {function, $H(x,p)$ | operator, $\hat{H}$} is specified for a given system, then the undisturbed time evolution of the state of the system $\{[x(t),p(t)] \mid \Psi(x,t)\}$ can be uniquely determined by solving the time-evolution equation {Eqs. (3-10)| Eq. (4-26)}, subject to the specified initial condition $\{[x(0),p(0)] = [x_0, p_0] \mid \Psi(x,0) = \Psi_0(x)\}$. Pursuing this parallelism a bit further, we recall that the postulates of classical mechanics did *not* specify the exact *form* which the Hamiltonian function assumes for a given physical system; for example, Newton's second law does not tell us that the Hamiltonian function appropriate to a mass m attached to a spring is $H(x,p) = p^2/2m + kx^2/2$. Similarly, Postulate 5 does *not* tell us what to write down for the Hamiltonian operator $\hat{H}$ for a given physical system. Indeed, it is the task of the physicist as a "clever observer of Nature" to *discover* or *invent* an appropriate Hamiltonian {function, $H(x,p)$| operator, $\hat{H}$} for a given {classical | quantum} system.

We see then that, given the explicit form of the Hamiltonian operator, we can in principle calculate the state vector of the system at time t *if* we know the state vector at time 0. An obvious question at this point is, how can we ever know what the state vector is at $t = 0$? Unless we are just "given" $\Psi_0(x)$ outright, the only way we can know what it is for sure is to *make a measurement* of some observable $\mathcal{A}$ at time $t = 0$: By Postulate 4, this measurement will force the system into one of the eigenvectors of $\hat{A}$, and merely by taking cognizance of which eigenvalue was obtained in the measurement, we will know *which* eigenvector the system was forced into. Thus, if a measurement of $\mathcal{A}$ at time 0 yields the result A_k, then immediately after the measurement the state vector will begin to evolve from the vector $\Psi_0(x) = \alpha_k(x)$ in the manner prescribed by Eq. (4-26). This orderly time evolution will continue until such time as another measurement is performed on the system, at which time the foregoing process will be repeated. It will be observed that this method of "preparing" the state $\Psi_0(x)$ depends crucially upon the fact that the measurement operation tells us what the state vector is immediately *after* a measurement, rather than immediately *before*.

One of the requirements of Postulate 1 was that the state vector of the system always have unit norm: $(\Psi_t, \Psi_t) = 1$. It is reasonable to ask whether or not the time behavior of $\Psi_t(x)$, as specified by

Eq. (4-26), is such that, if $(\Psi_0,\Psi_0) = 1$, then it will be true that $(\Psi_t,\Psi_t) = 1$ for all $t > 0$. To show that this is indeed the case, we shall prove that the time derivative of the norm of $\Psi_t(x)$ vanishes identically. We have

$$\frac{d}{dt}(\Psi_t,\Psi_t) = \left(\frac{\partial\Psi_t}{\partial t},\Psi_t\right) + \left(\Psi_t,\frac{\partial\Psi_t}{\partial t}\right) \tag{4-28}$$

Exercise 40. By writing out the inner product (Ψ_t,Ψ_t) in its integral form, $\int\Psi^*(x,t)\Psi(x,t)dx$, and using the fact that the integration variable x is independent of t, prove Eq. (4-28).

Now according to Eq. (4-26),

$$\frac{\partial\Psi_t}{\partial t} = \frac{1}{i\hbar}\hat{H}\Psi_t = -\frac{i}{\hbar}\hat{H}\Psi_t$$

where we have used the fact that, since $i^2 = -1$, then $1/i = -i$. Inserting this into the right-hand side of Eq. (4-28), we obtain

$$\begin{aligned}\frac{d}{dt}(\Psi_t,\Psi_t) &= \left(-\frac{i}{\hbar}\hat{H}\Psi_t,\Psi_t\right) + \left(\Psi_t,-\frac{i}{\hbar}\hat{H}\Psi_t\right)\\ &= \left[-\frac{i}{\hbar}\right]^*(\hat{H}\Psi_t,\Psi_t) + \left[-\frac{i}{\hbar}\right](\Psi_t,\hat{H}\Psi_t)\\ &= \left[\frac{i}{\hbar}\right][(\hat{H}\Psi_t,\Psi_t) - (\Psi_t,\hat{H}\Psi_t)]\end{aligned}$$

Since $\hat{H}$ is an Hermitian operator, then for any state $\Psi_t(x)$, we have $(\hat{H}\Psi_t,\Psi_t) = (\Psi_t,\hat{H}\Psi_t)$; therefore, we conclude that

$$\frac{d}{dt}(\Psi_t,\Psi_t) = 0 \tag{4-29}$$

This proves that the time evolution of the state vector, as dictated by Postulate 5, is consistent with the requirement of Postulate 1 that the state vector always have unit norm.

Although the time evolution of the state vector is completely specified by the differential equation (4-26), there is another way of describing this time evolution which has a formal elegance, and occasional usefulness, that Eq. (4-26) lacks. According to Postulate 5, if we place the system in some particular state $\Psi_0(x)$ at time 0, then the passage of a time t "transforms" this state vector into some new vector $\Psi_t(x)$, which can be found by solving Eq. (4-26). However, we recall from our discussion in Sec. 2-4 that the transformation of a given $\mathcal{H}$-vector into another $\mathcal{H}$-vector may generally be regarded as the result of an *operator* acting on the given vector. In this spirit, it

is tempting to write

$$\Psi_t(x) = \hat{U}(t)\Psi_0(x) \tag{4-30}$$

where $\hat{U}(t)$ is some operator that transforms the state vector at time 0 into the state vector at time t. Now, we know from Eq. (4-26) that the time evolution of the state vector depends crucially upon the system's Hamiltonian operator $\hat{H}$, so it is reasonable to expect that $\hat{U}(t)$ will be some function of $\hat{H}$ as well as t. It is also possible that $\hat{U}(t)$ might depend upon $\Psi_0(x)$ as well, but if this were the case then Eq. (4-30) would not be very interesting or useful. However, by using the definition of $\hat{U}(t)$ in Eq. (4-30) together with the time-evolution equation (4-26), we can show that the operator $\hat{U}(t)$ is a well-defined function of $\hat{H}$ and t alone, and is independent of the initial state $\Psi_0(x)$. We call $\hat{U}(t)$ the *time-evolution operator* of the system.

To deduce the form of the time-evolution operator, we first substitute Eq. (4-30) into Eq. (4-26):

$$i\hbar \frac{\partial \hat{U}(t)}{\partial t} \Psi_0(x) = \hat{H}\hat{U}(t)\Psi_0(x)$$

This equation says that the operator $i\hbar(\partial \hat{U}/\partial t)$ acting on the $\mathcal{H}$-vector $\Psi_0(x)$ must produce the same $\mathcal{H}$-vector as does the operator $\hat{H}\hat{U}$ acting on $\Psi_0(x)$. Since we want this to be true for *any* choice of $\Psi_0(x)$, then we must have

$$i\hbar \frac{\partial \hat{U}(t)}{\partial t} = \hat{H}\hat{U}(t) \tag{4-31}$$

This is a differential equation for the operator $\hat{U}(t)$. The time-evolution operator must satisfy this differential equation for all $\hat{H}$ and t, and it must *also* satisfy the "initial condition"

$$\hat{U}(0) = 1 \tag{4-32}$$

Exercise 41. Derive Eq. (4-32) from the definition of $\hat{U}(t)$ in Eq. (4-30), and explain its meaning.

We shall now demonstrate that the required form for the operator $\hat{U}(t)$ is

$$\hat{U}(t) = 1 + \sum_{n=1}^{\infty} \frac{\left(-\frac{i}{\hbar}\hat{H}t\right)^n}{n!} \equiv e^{-i\hat{H}t/\hbar} \tag{4-33}$$

This equation asserts that $\hat{U}(t)$ is a certain power series in the operator $\hat{H}$ (the power series has a well-defined meaning by virtue of Eqs. (2-42)), and we have chosen to denote this power series by the

symbol $\exp(-i\hat{H}t/\hbar)$. In adopting this symbolic representation for the power series in Eq. (4-33), we are merely following the procedure described in Sec. 4-2 for defining functions of operators [see Eqs. (4-5)]; for, it will be recalled that the function $f(x) = e^x$ has the Taylor series expansion

$$e^x = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!}$$

and we have merely replaced x by $(-i\hat{H}t/\hbar)$ in obtaining the right-hand side of Eq. (4-33).

It is easy to see that the power series in Eq. (4-33) satisfies the initial condition Eq. (4-32), since all the terms under the summation sign vanish for $t = 0$. To see that this power series also satisfies the differential equation (4-31), we calculate its time derivative. Remembering that the operator $\hat{H}$ is linear, so that $(\hat{H}t)^n = \hat{H}^n t^n$, and also that $\hat{H}$ is independent of t, we have from Eq. (4-33)

$$\frac{\partial \hat{U}(t)}{\partial t} = 0 + \sum_{n=1}^{\infty} \frac{\left(-\frac{i}{\hbar}\hat{H}\right)^n (nt^{n-1})}{n!} = \left(-\frac{i}{\hbar}\hat{H}\right) \sum_{n=1}^{\infty} \frac{\left(-\frac{i}{\hbar}\hat{H}t\right)^{n-1}}{(n-1)!}$$

$$= \left(-\frac{i}{\hbar}\hat{H}\right)\left[1 + \sum_{n=1}^{\infty} \frac{\left(-\frac{i}{\hbar}\hat{H}t\right)^n}{n!}\right] = \frac{1}{i\hbar}\hat{H}\hat{U}(t)$$

which is identical to Eq. (4-31). Consequently, the operator $U(t)$ in Eq. (4-33) is indeed the time-evolution operator of the system; it satisfies Eqs. (4-31) and (4-32), and therefore it satisfies Eq. (4-30) for all $t \geq 0$ and all choices of the initial-state vector $\Psi_0(x)$.

Exercise 42. Show that if one ignores the operator character of $\hat{H}$ and $\hat{U}(t)$, then the "symbol" $\exp(-i\hat{H}t/\hbar)$ satisfies Eqs. (4-31) and (4-32) in a purely formal sense.

Since $\hat{U}(t)$ in Eq. (4-33) is a linear combination of products of the linear operator $\hat{H}$, it follows that $\hat{U}(t)$ is a *linear* operator. However, the coefficients in this linear combination are obviously not real, so it does not follow that $\hat{U}(t)$ is an Hermitian operator [see Exercise 16]. In fact, it is easy to see that $\hat{U}(t)$ is *not* Hermitian: We recall from Exercise 26 that the operator $f(\hat{A})$ has eigenvectors $\{\alpha_n(x)\}$ and eigenvalues $\{f(A_n)\}$, where $\{\alpha_n(x)\}$ and $\{A_n\}$ are the eigenvectors and eigenvalues of $\hat{A}$. Therefore, the operator $\hat{U}(t)$ has eigenvectors $\{\eta_n(x)\}$ and eigenvalues $\{\exp(-iE_n t/\hbar)\}$. In

particular, since the eigenvalues of $U(t)$ are not pure real [see Eq. (2-20a)], then $\hat{U}(t)$ cannot be Hermitian [see the discussion preceding Exercise 18]. Therefore, $\hat{U}(t)$ is *not* to be regarded as an *observable* operator; its character and function are altogether different from the operators which we have discussed so far.

The time-evolution operator plays an important role in more advanced treatments of the time behavior of systems; however, in our subsequent work in this book we shall use the time-evolution *equation*, Eq. (4-26), rather that the time-evolution *operator*, Eq. (4-33). We have introduced the time-evolution operator primarily because the picture conveyed by Eq. (4-30), of $\Psi_0(x)$ being "carried into" $\Psi_t(x)$ by a linear operator which depends only on $\hat{H}$ and t, is conceptually very important, and significantly enhances our appreciation of the time-evolution process in quantum mechanics.

It was pointed out earlier that the general observable operator $\hat{A}$, along with its eigenbasis $\{\alpha_n(x)\}$ and eigenvalues $\{A_n\}$, do *not* depend upon time. However, the time-dependence of $\Psi_t(x)$ prescribed by Postulate 5 clearly implies a time-dependence for the expectation value $\langle\hat{A}\rangle_t = (\Psi_t, \hat{A}\Psi_t)$, and the probability coefficients (α_i, Ψ_t). In the next two sections we shall derive from Eq. (4-26) the time-evolution equations for these two quantities.

4-4b The Time Evolution Equation for $\langle\hat{A}\rangle_t$ The Time-Energy Uncertainty Relation

We found in Sec. 4-3a that the results of many repeated measurements of an observable $\mathfrak{A}$ on a state $\Psi_t(x)$ can often be adequately described by the expectation value $\langle\hat{A}\rangle_t$ and the uncertainty $\Delta\hat{A}_t$ [see Fig. 3]. The behavior of these quantities with time is therefore of some interest. We shall not examine here the time evolution of $\Delta\hat{A}_t$, but we note that, by virtue of Eq. (4-14), the time behavior of $\Delta\hat{A}_t$ can in principle be determined if one knows how to determine the time behavior of expectation values. It is this latter problem that we shall examine in this section.

To investigate the time dependence of $\langle\hat{A}\rangle_t$, let us simply calculate its time derivative. This calculation is formally very similar to the one carried out in the last section in proving Eq. (4-29). Using the fact that the inner-product integration variable x and the observable operator $\hat{A}$ are both independent of time, we have in analogy with Eq. (4-28),

$$\frac{d}{dt}\langle\hat{A}\rangle_t = \frac{d}{dt}(\Psi_t, \hat{A}\Psi_t) = \left(\frac{\partial\Psi_t}{\partial t}, \hat{A}\Psi_t\right) + \left(\Psi_t, \hat{A}\frac{\partial\Psi_t}{\partial t}\right)$$

Then, inserting the expression for $\partial \Psi_t / \partial t$ given in Postulate 5,

$$\frac{d}{dt}\langle \hat{A} \rangle_t = \left(-\frac{i}{\hbar}\hat{H}\Psi_t, \hat{A}\Psi_t \right) + \left(\Psi_t, \hat{A}\left[-\frac{i}{\hbar}\hat{H}\Psi_t \right] \right)$$

$$= \left[-\frac{i}{\hbar} \right]^* (\hat{H}\Psi_t, \hat{A}\Psi_t) + \left[-\frac{i}{\hbar} \right] (\Psi_t, \hat{A}\hat{H}\Psi_t)$$

$$= \left[\frac{i}{\hbar} \right] [(\hat{H}\Psi_t, \hat{A}\Psi_t) - (\Psi_t, \hat{A}\hat{H}\Psi_t)]$$

Finally, invoking the Hermiticity of $\hat{H}$, we obtain

$$\frac{d}{dt}\langle \hat{A} \rangle_t = \left[\frac{i}{\hbar} \right] [(\Psi_t, \hat{H}\hat{A}\Psi_t) - (\Psi_t, \hat{A}\hat{H}\Psi_t)]$$

or

$$\frac{d}{dt}\langle \hat{A} \rangle_t = \frac{i}{\hbar}(\Psi_t, [\hat{H}\hat{A} - \hat{A}\hat{H}]\Psi_t) \tag{4-34}$$

This is the fundamental time-evolution equation for the expectation value of an observable $\mathcal{A}$. It evidently gives the instantaneous time-rate-of-change of $\langle \hat{A} \rangle_t$ in terms of the instantaneous state vector $\Psi_t(x)$ and the operator $(\hat{H}\hat{A} - \hat{A}\hat{H})$.

Exercise 43. Using Eq. (4-34), prove that if $\hat{A}$ commutes with the Hamiltonian operator of the system, then both the expectation value $\langle \hat{A} \rangle_t$ and the uncertainty $\Delta \hat{A}_t$ are constant in time. [*Hint*: To prove that $\Delta \hat{A}_t$ is constant, show that both $\langle \hat{A}^2 \rangle_t$ and $\langle \hat{A} \rangle_t^2$ are constants.]

The foregoing exercise provides us with a rule for determining the "constants of the motion": If $\hat{A}$ commutes with $\hat{H}$, or equivalently, if $\mathcal{A}$ is compatible with the total energy of the system, then $\mathcal{A}$ is a constant of the motion in the sense that $\langle \hat{A} \rangle_t$ and $\Delta \hat{A}_t$ are independent of time. In particular, since $\hat{H}$ certainly commutes with itself, then $\langle \hat{H} \rangle_t$ and $\Delta \hat{H}_t$ are *always* constant in time. This is analogous to the result in classical mechanics that the energy is a constant of the motion [see Exercise 22]. We shall discuss the concept of a "constant of the motion" in greater detail in the next section.

One aspect of the relationship between time and energy which has no direct analogue in classical mechanics is the so-called "time-energy uncertainty principle." We define for a given observable $\mathcal{A}$ its *evolution time* $T_{\mathcal{A}}$ by

$$T_{\mathcal{A}} \equiv \Delta \hat{A}_t \Big/ \left| \frac{d\langle \hat{A} \rangle_t}{dt} \right| \tag{4-35}$$

To understand the physical significance of $T_{\mathcal{A}}$, let us first suppose that $d\langle\hat{A}\rangle_t/dt$ is constant in time. Then, in a given time interval Δt, the expectation value of $\mathcal{A}$ would change by an amount $|d\langle\hat{A}\rangle_t/dt|\cdot\Delta t$; in particular, we see from Eq. (4-35) that in a time interval equal to $T_{\mathcal{A}}$, $\langle\hat{A}\rangle_t$ would change by an amount equal to $\Delta\hat{A}_t$. In general, even if $d\langle\hat{A}\rangle_t/dt$ is *not* a constant, it is obvious that $T_{\mathcal{A}}$, as defined by Eq. (4-35), provides a very reasonable estimate of the amount of of time which must elapse before the expectation value of $\mathcal{A}$ changes by an amount equal to the uncertainty in $\mathcal{A}$. In other words, $T_{\mathcal{A}}$ is the time which must elapse before the *average* of the values measured for $\mathcal{A}$ in a series of repeated measurements, $\langle\hat{A}\rangle_t$, changes or evolves enough to be *noticeable* over the *intrinsic spread* in these values, $\Delta\hat{A}_t$. It is in this sense that $T_{\mathcal{A}}$ specifies the "evolution time" of $\mathcal{A}$.

Now, by combining the time-evolution equation for $\langle\hat{A}\rangle_t$, Eq. (4-34), with the generalized uncertainty relation, Eq. (4-24), we can easily derive the following inequality:

$$T_{\mathcal{A}}\cdot\Delta\hat{H} \geq \frac{\hbar}{2} \tag{4-36}$$

Exercise 44. Derive Eq. (4-36). [*Hint;* First obtain an expression for $|d\langle\hat{A}\rangle_t/dt|$ from Eq. (4-34); then make use of Eqs. (4-24) and (4-35).]

The above inequality is called the *time-energy uncertainty relation*. It states that the more precisely the energy of a system is defined (i.e., the smaller $\Delta\hat{H}$ is), then the more slowly will *any* observable $\mathcal{A}$ change "noticeably" with time (i.e., the larger any $T_{\mathcal{A}}$ must be); conversely, if *any* observable plainly exhibits a rapid variation with time (i.e., if any $T_{\mathcal{A}}$ is small), then the system *cannot* have a well-defined energy (i.e., $\Delta\hat{H}$ must be large). These predictions of Eq. (4-36), although obviously "nonclassical" in character, have been amply confirmed in the laboratory by spectroscopic studies comparing the "widths" of excited atomic energy levels with the corresponding "lifetimes" of these levels: it is found that narrow, sharp energy levels have long lifetimes, while broad, diffuse energy levels have short lifetimes.

4-4c The Time Evolution Equation for (α_n, Ψ_t) Constants of the Motion and Stationary States

There are two reasons why it is useful to know the time-dependence of the quantities

$$a_n(t) \equiv (\alpha_n, \Psi_t) \qquad n = 1,2,\ldots \tag{4-37}$$

First, the square moduli of these quantities determine the shape of the distribution curve for $\mathfrak{A}$ [see Fig. 3], and it is clearly of interest to know how this curve changes with time. Second, since the state vector of the system can always be written in the form

$$\Psi_t(x) = \sum_{n=1}^{\infty} (\alpha_n, \Psi_t)\alpha_n(x) \equiv \sum_{n=1}^{\infty} a_n(t)\alpha_n(x) \tag{4-38}$$

then if a simple expression for $a_n(t)$ can be found, we will have an explicit representation for the time-varying state vector.

To see what we can discover about the time-dependence of $a_n(t)$, let us calculate its time derivative. Since $\alpha_n(x)$ is independent of t, we have

$$\frac{d}{dt} a_n(t) \equiv \frac{d}{dt} (\alpha_n, \Psi_t) = \left(\alpha_n, \frac{\partial \Psi_t}{\partial t}\right)$$

Inserting the expression for $\partial \Psi_t / \partial t$ given in Postulate 5,

$$\frac{d}{dt} a_n(t) = \left(\alpha_n, -\frac{i}{\hbar}\hat{H}\Psi_t\right) = -\frac{i}{\hbar}(\alpha_n, \hat{H}\Psi_t)$$

We now replace $\Psi_t(x)$ by its expansion in Eq. (4-38), and make use of the linearity of $\hat{H}$:

$$\frac{d}{dt} a_n(t) = -\frac{i}{\hbar}\left(\alpha_n, \hat{H} \sum_{m=1}^{\infty} a_m(t)\alpha_m\right) = -\frac{i}{\hbar}\left(\alpha_n, \sum_{m=1}^{\infty} a_m(t)\hat{H}\alpha_m\right)$$

Finally, using Eqs. (2-34b) and (2-34c), we obtain the result

$$\frac{d}{dt} a_n(t) = -\frac{i}{\hbar} \sum_{m=1}^{\infty} a_m(t)\,(\alpha_n, \hat{H}\alpha_m) \qquad n = 1,2,\ldots \tag{4-39}$$

For an arbitrary eigenbasis $\{\alpha_i(x)\}$, this is about as far as we can go. Recognizing that the quantities $(\alpha_n, H\alpha_m)$ in Eqs. (4-39) are ordinary complex numbers, we see that Eqs. (4-39) express the time derivative of *each* $a_n(t)$ as a linear combination of *all* the $a_i(t)$. In principle, one could solve this infinite set of "coupled, linear differential equations," and so obtain an explicit expression for each $a_n(t)$; in practice, however, this is usually far too difficult to do for an arbitrary eigenbasis $\{\alpha_i(x)\}$.

There is, however, one important case for which Eqs. (4-39) can be solved fairly easily. Suppose the eigenbasis $\{\alpha_i(x)\}$ coincides with the *energy* eigenbasis $\{\eta_i(x)\}$; this will be true if the observable operator $\hat{A}$ under consideration coincides with $\hat{H}$, or more generally, by the Compatibility Theorem, if $\hat{A}$ *commutes* with $\hat{H}$. In this case,

it is easy to show that Eqs. (4-39) "uncouple" and take the simple forms

$$\frac{d}{dt}(\eta_n,\Psi_t) = -\frac{i}{\hbar}E_n(\eta_n,\Psi_t) \qquad n = 1,2,\ldots \tag{4-40}$$

Exercise 45. Show that Eqs. (4-39) reduce to Eqs. (4-40) when the eigenvectors $\{\alpha_i(x)\}$ coincide with the energy eigenvectors $\{\eta_i(x)\}$ as defined in Eq. (4-25).

Eqs. (4-40) can be immediately integrated: Writing it as

$$\frac{d(\eta_n,\Psi_t)}{(\eta_n,\Psi_t)} = -\frac{iE_n}{\hbar}\,dt$$

then an elementary integration yields

$$\log(\eta_n,\Psi_t) = -\frac{iE_n}{\hbar}t + \log C$$

where we have written the integrating constant as $\log C$. Using the properties of the exponential, we can write this last equation as

$$(\eta_n,\Psi_t) = Ce^{-iE_nt/\hbar}$$

where the complex exponential has been defined and discussed in Exercise 7. Finally since $e^0 = 1$, we see that we must set $C = (\eta_n,\Psi_0)$; consequently, we conclude that†

$$(\eta_n,\Psi_t) = (\eta_n,\Psi_0)e^{-iE_nt/\hbar} \qquad n = 1,2,\ldots \tag{4-41}$$

Thus while it is not possible to say very much about the time dependence of (α_n,Ψ_t) for an *arbitrary* eigenbasis $\{\alpha_i(x)\}$, Eqs. (4-41) give the *explicit* time dependence for the case in which the eigenbasis is the *energy* eigenbasis. For the remainder of this section, we shall examine some of the consequences of Eqs. (4-41).

Let us first discuss Eqs. (4-41) from the viewpoint that $|(\eta_n,\Psi_t)|^2$ represents a *probability*. We recall from Exercise 43 that any observable $\mathfrak{A}$ whose corresponding operator $\hat{A}$ commutes with $\hat{H}$ is a sort of "constant of the motion" in the sense that $\langle\hat{A}\rangle_t$ and $\Delta\hat{A}_t$ do not change with time. Now if $\hat{A}$ commutes with $\hat{H}$, then according to the Compatibility Theorem, $\hat{A}$ and $\hat{H}$ must share the same

†Our derivation of Eq. (4-41) may leave the reader a bit uneasy, since it involved the complex logarithm, a concept which we have not discussed. However, the reader can easily verify that the formula for (η_n,Ψ_t) given in Eqs. (4-41) does indeed satisfy the differential equations (4-40), simply by invoking Eq. (2-20e). Since Eqs. (4-41) satisfy *both* the differential equations and the initial conditions, it is therefore *the* solution.

eigenbasis; thus, by making at most a rearrangement of indices, we can put

$$\alpha_n(x) = \eta_n(x) \qquad n = 1, 2, \ldots$$

This being the case, Eqs. (4-41) imply that

$$(\alpha_n, \Psi_t) = (\eta_n, \Psi_t) = (\eta_n, \Psi_0)\, e^{-iE_n t/\hbar} = (\alpha_n, \Psi_0)\, e^{-iE_n t/\hbar}$$

Therefore, the probability of measuring for $\mathfrak{A}$ the eigenvalue A_n at time t is

$$|(\alpha_n, \Psi_t)|^2 = |(\alpha_n, \Psi_0)|^2 |e^{-iE_n t/\hbar}|^2 = |(\alpha_n, \Psi_0)|^2$$

where we have made use of Eqs. (2-18a) and (2-20d). We have thus shown that, if $\hat{A}$ commutes with $\hat{H}$, then the probability of measuring for $\mathfrak{A}$ the eigenvalue A_n is the same at time t as at time 0. Since the quantities $|(\alpha_n, \Psi_t)|^2$ determine the distribution curve for $\mathfrak{A}$ at time t [see Fig. 3], then we see that, if $\hat{A}$ commutes with $\hat{H}$, not only are the "center" and "width" of the distribution curve constant in time [see Exercise 43], but indeed the *entire curve* does not change with time. We are therefore quite justified in calling $\mathfrak{A}$ a *constant of the motion* whenever $\hat{A}$ commutes with $\hat{H}$.

Let us next examine the consequences of Eqs. (4-41) with regard to the *representation of the time-varying state vector*. Since we can always expand $\Psi_t(x)$ in the eigenbasis of $\hat{H}$,

$$\Psi_t(x) = \sum_{n=1}^{\infty} (\eta_n, \Psi_t)\eta_n(x)$$

then substituting Eqs. (4-41) yields at once

$$\Psi_t(x) = \sum_{n=1}^{\infty} (\eta_n, \Psi_0) e^{-iE_n t/\hbar} \eta_n(x) \tag{4-42}$$

Inasmuch as the complex numbers (η_n, Ψ_0) do not depend upon time, Eq. (4-42) shows *explicitly* the time dependence of the state vector $\Psi_t(x)$, and therefore represents a *general solution to the fundamental time-evolution equation* (4-26). We note in particular that, since $e^0 = 1$, then for $t = 0$ Eq. (4-42) reduces to

$$\Psi_0(x) = \sum_{n=1}^{\infty} (\eta_n, \Psi_0)\eta_n(x) \tag{4-43}$$

which of course is an identity for *any* $\mathcal{H}$-vector $\Psi_0(x)$ [see Eq. (4-2b)]. Indeed, an easy way to remember Eq. (4-42) is to first write down Eq. (4-43), and then simply insert the factor $\exp(-iE_n t/\hbar)$ in

the nth term of the sum.† We have in Eq. (4-42) yet another instance of the intimate connection between time and energy: if we can find the energy eigenvectors $\{\eta_i(x)\}$ and eigenvalues $\{E_i\}$—i.e., if we can solve the energy eigenvalue equation (4-25)—then we can write down at once a complete solution to the time-evolution equation for the state vector, Eq. (4-26).

Exercise 46. By directly substituting Eq. (4-42) into Eq. (4-26), show that this expression for $\Psi_t(x)$ does indeed satisfy the time-evolution equation of Postulate 5. [*Hint*: Make use of Eqs. (2-20e) and (4-25), and remember that neither $\partial/\partial t$ nor $\hat{H}$ has any effect upon the complex constants (η_n, Ψ_0).]

It is interesting to consider the special case in which the initial state vector, $\Psi_0(x)$, coincides with one of the energy eigenvectors, say $\eta_k(x)$. In this case, the coefficients in Eq. (4-42) are

$$(\eta_n, \Psi_0) = (\eta_n, \eta_k) = \delta_{nk}$$

and Eq. (4-42) reduces to

$$\Psi_t(x) = e^{-iE_k t/\hbar}\, \eta_k(x)$$

This says that $\Psi_t(x)$ differs from $\Psi_0(x) = \eta_k(x)$ only by a scalar factor whose square modulus is *unity* [see Eq. (2-20d)]:

$$|e^{-iE_k t/\hbar}|^2 = 1$$

Consequently, by Postulate 1, the state of the system at time t is *physically the same* as the state at time 0. In such a case, it is clear that *all* observables will behave like "constants of the motion."

Exercise 47.

(a) Prove directly that if $\Psi_t(x) = e^{-iE_k t/\hbar}\eta_k(x)$, then for *any* observable $\mathcal{A}$, the expectation value $\langle \hat{A} \rangle_t$, the uncertainty $\Delta \hat{A}_t$, and the probabilities $|(\alpha_n, \Psi_t)|^2$ are all constant in time—regardless of whether or not $\hat{A}$ commutes with $\hat{H}$.

†Equation (4-42) can also be *derived* from Eq. (4-43) by using the time evolution operator $\hat{U}(t) = \exp(-i\hat{H}t/\hbar)$, which was discussed in Sec. 4-4a. We recall that this operator is a *linear* operator, and has eigenvectors $\{\eta_n(x)\}$ and eigenvalues $\{\exp(-iE_n t/\hbar)\}$. Therefore,

$$\Psi_t(x) = \hat{U}(t)\Psi_0(x) = \hat{U}(t) \sum_{n=1}^{\infty} (\eta_n, \Psi_0)\eta_n(x)$$

$$= \sum_{n=1}^{\infty} (\eta_n, \Psi_0)\hat{U}(t)\eta_n(x) = \sum_{n=1}^{\infty} (\eta_n, \Psi_0)\, e^{-iE_n t/\hbar}\, \eta_n(x)$$

which is evidently Eq. (4-42).

(b) Show that in this case we will have at any time t, $\Delta\hat{H} = 0$ and $T_a = \infty$. Discuss this result in the light of the Time-Energy Uncertainty Relation, Eq. (4-36). [*Hint*: Use Exercise 33 and the definition in Eq. (4-35).]

The foregoing considerations motivate us to define, for a given system with a given Hamiltonian, the set of *time-dependent* $\mathcal{H}$-vectors $\{\Psi_t^{(n)}(x)\}$:

$$\Psi_t^{(n)}(x) \equiv \Psi^{(n)}(x,t) \equiv e^{-iE_n t/\hbar}\eta_n(x) \qquad n = 1, 2, \ldots \quad (4\text{-}44)$$

According to the above arguments, if $\Psi_0(x)$ coincides with $\eta_n(x)$, then $\Psi_t(x)$ will be given by $\Psi_t^{(n)}(x)$. Moreover, by Exercise 47, if the state vector of the system coincides with $\Psi_t^{(n)}(x)$, then *all* observables behave like constants of the motion. For this reason, the vectors $\{\Psi_t^{(n)}(x)\}$ are called the *stationary states* of the system. In terms of these stationary states, Eq. (4-42) can be written

$$\Psi_t(x) = \sum_{n=1}^{\infty} (\Psi_0^{(n)}, \Psi_0)\Psi_t^{(n)}(x) \quad (4\text{-}45)$$

Exercise 48. Prove that Eq. (4-45) follows from Eqs. (4-42) and (4-44).

Thus the stationary states of a system are important because, given any initial state vector $\Psi_0(x)$, the state vector at any subsequent time t can be written as a linear combination of the stationary states. The coefficients in this linear combination are the *time-independent* complex numbers $(\Psi_0^{(n)}, \Psi_0)$, which can evidently be evaluated if the stationary states and the initial state are known as functions of x.

We see then that $\Psi_t(x)$ can in principle always be found if the stationary states can be determined. According to Eq. (4-42), the determination of the stationary states is tantamount to finding the eigenvectors and eigenvalues of the energy operator, $\hat{H}$. For this reason, most of the actual "problem solving" in quantum mechanics is concerned with solving the energy eigenvalue equation (4-25) for various Hamiltonian operators—an enterprise which is usually very difficult from a mathematical point of view.

4-4d "Determinism" in Quantum Mechanics

We have now essentially completed the task of erecting the main conceptual framework of quantum mechanics. In order to *apply* the theory to a particular situation, it remains only to specify the *exact*

forms of the relevant observable operators, and then to find their eigenvectors and eigenvalues. In the next section we shall discuss how this is done for one particularly important class of physical systems. But before doing so, it seems appropriate now to reconsider, from the viewpoint of quantum mechanics, our brief remarks in Sec. 3-3 on the matter of "determinism."

By "determinism" we mean here the general possibility of predicting exactly how the state of a system will *change* in any given circumstance. In classical mechanics, the change in the state of any system with time is in principle completely predictable (barring, of course, any unwarranted disturbance of the system by some external agent), and on this basis we concluded that classical mechanics implies a "deterministic" universe.

With regard to quantum mechanics, the situation from one point of view is very similar: In our discussion of Postulate 5 we saw that the state vector of a quantum system evolves with time in a completely predictable manner, and in this sense it may be said that quantum mechanics, like classical mechanics, is a "deterministic theory." However, in quantum mechanics the state of a system not only changes with the passage of time, but it also changes as a result of being measured. In our discussion of Postulates 3 and 4 we found that the change induced in the state vector by a measurement is *in principle* neither controllable nor predictable. That is, if we decide to measure $\mathcal{A}$ on a given state, it is usually not possible to know precisely which eigenvector of $\hat{A}$ the state will be forced into by the measurement. From this point of view, it is evident that quantum mechanics is *not* a completely deterministic theory.

If we choose to regard the entire universe as a single system, governed by one super-Hamiltonian operator, then since there is nothing "external" which can make a measurement on this system, we may justifiably assert that the state of the whole universe evolves with time in a completely deterministic way. If, however, we wish to consider only a *portion* of the universe as our system, omitting for example ourselves and our measuring apparatus, then we must evidently contend with a certain amount of indeterministic behavior every time we make a measurement upon the system.

The general problem of assessing the impact of quantum mechanics upon the concepts of determinism and causality is obviously a very intriguing and many-faceted one, and is as much in the domain of philosophy as physics. Without demeaning the importance of this complex problem, we shall not try to discuss it any further here. We shall simply point out that any serious discussion of these matters must also consider the question of whether or not the so-called "orthodox" interpretation of quantum mechanics, which we have

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