CHAPTER 1

THE WAVE FUNCTION

1.1 THE SCHRÖDINGER EQUATION

Imagine a particle of mass m, constrained to move along the x-axis, subject to some specified force $F(x, t)$ (Figure 1.1). The program of *classical* mechanics is to determine the position of the particle at any given time: $x(t)$. Once we know that, we can figure out the velocity ($v = dx/dt$), the momentum ($p = mv$), the kinetic energy ($T = (1/2)mv^2$), or any other dynamical variable of interest. And how do we go about determining $x(t)$? We apply Newton's second law: $F = ma$. (For conservative systems—the only kind we shall consider, and, fortunately, the only kind that *occur* at the microscopic level—the force can be expressed as the derivative of a potential energy function,[1] $F = -\partial V/\partial x$, and Newton's law reads $m\, d^2x/dt^2 = -\partial V/\partial x$.) This, together with appropriate initial conditions (typically the position and velocity at $t = 0$), determines $x(t)$.

Quantum mechanics approaches this same problem quite differently. In this case what we're looking for is the **wave function**, $\Psi(x, t)$, of the particle, and we get it by solving the **Schrödinger equation**:

$$\boxed{i\hbar\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \Psi}{\partial x^2} + V\Psi.} \qquad [1.1]$$

[1]Magnetic forces are an exception, but let's not worry about them just yet. By the way, we shall assume throughout this book that the motion is nonrelativistic ($v \ll c$).

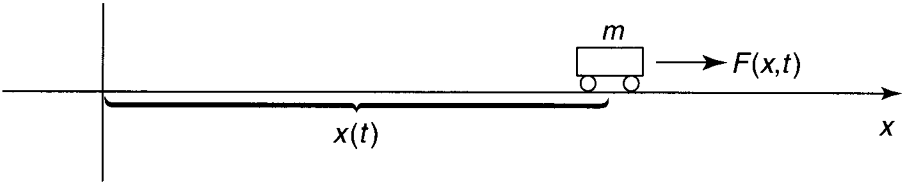

Figure 1.1: A "particle" constrained to move in one dimension under the influence of a specified force.

Here i is the square root of -1, and $\hbar$ is Planck's constant—or rather, his original constant (h) divided by 2π:

$$\hbar = \frac{h}{2\pi} = 1.054573 \times 10^{-34}\text{J s}. \qquad [1.2]$$

The Schrödinger equation plays a role logically analogous to Newton's second law: Given suitable initial conditions [typically, $\Psi(x, 0)$], the Schrödinger equation determines $\Psi(x, t)$ for all future time, just as, in classical mechanics, Newton's law determines $x(t)$ for all future time.

1.2 THE STATISTICAL INTERPRETATION

But what exactly *is* this "wave function", and what does it do for you once you've *got* it? After all, a particle, by its nature, is localized at a point, whereas the wave function (as its name suggests) is spread out in space (it's a function of x, for any given time t). How can such an object be said to describe the state of a *particle*? The answer is provided by Born's **statistical interpretation** of the wave function, which says that $|\Psi(x, t)|^2$ gives the *probability* of finding the particle at point x, at time t—or, more precisely,[2]

$$\boxed{|\Psi(x, t)|^2\, dx = \left\{ \begin{array}{l} \text{probability of finding the particle} \\ \text{between } x \text{ and } (x + dx), \text{ at time } t. \end{array} \right\}} \qquad [1.3]$$

For the wave function in Figure 1.2, you would be quite likely to find the particle in the vicinity of point A, and relatively unlikely to find it near point B.

The statistical interpretation introduces a kind of **indeterminacy** into quantum mechanics, for even if you know everything the theory has to tell you about the

[2]The wave function itself is complex, but $|\Psi|^2 = \Psi^*\Psi$ (where Ψ^* is the complex conjugate of Ψ) is real and nonnegative—as a probability, of course, must be.

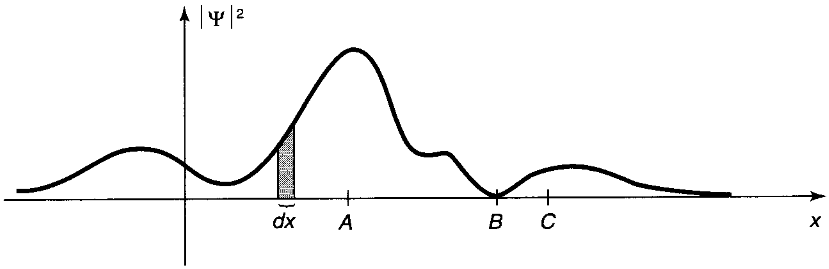

Figure 1.2: A typical wave function. The particle would be relatively likely to be found near A, and unlikely to be found near B. The shaded area represents the probability of finding the particle in the range dx.

particle (to wit: its wave function), you cannot predict with certainty the outcome of a simple experiment to measure its position—all quantum mechanics has to offer is *statistical* information about the *possible* results. This indeterminacy has been profoundly disturbing to physicists and philosophers alike. Is it a peculiarity of nature, a deficiency in the theory, a fault in the measuring apparatus, or *what*?

Suppose I *do* measure the position of the particle, and I find it to be at the point C. Question: Where was the particle just *before* I made the measurement? There are three plausible answers to this question, and they serve to characterize the main schools of thought regarding quantum indeterminacy:

1. The **realist** position: *The particle was at C.* This certainly seems like a sensible response, and it is the one Einstein advocated. Note, however, that if this is true then quantum mechanics is an **incomplete** theory, since the particle *really was* at C, and yet quantum mechanics was unable to tell us so. To the realist, indeterminacy is not a fact of nature, but a reflection of our ignorance. As d'Espagnat put it, "the position of the particle was never indeterminate, but was merely unknown to the experimenter."[3] Evidently Ψ is not the whole story—some additional information (known as a **hidden variable**) is needed to provide a complete description of the particle.

2. The **orthodox** position: *The particle wasn't really anywhere.* It was the act of measurement that forced the particle to "take a stand" (though how and why it decided on the point C we dare not ask). Jordan said it most starkly: "Observations not only *disturb* what is to be measured, they *produce* it. . . . We *compel* [the particle] to assume a definite position."[4] This view (the so-called **Copenhagen interpretation**) is associated with Bohr and his followers. Among physicists it has always been the

[3]Bernard d'Espagnat, *The Quantum Theory and Reality*, Scientific American, Nov. 1979 (Vol. 241), p. 165.

[4]Quoted in a lovely article by N. David Mermin, *Is the moon there when nobody looks*?, Physics Today, April 1985, p. 38.

most widely accepted position. Note, however, that if it is correct there is something very peculiar about the act of measurement—something that over half a century of debate has done precious little to illuminate.

3. The **agnostic** position: *Refuse to answer.* This is not quite as silly as it sounds—after all, what sense can there be in making assertions about the status of a particle *before* a measurement, when the only way of knowing whether you were right is precisely to conduct a measurement, in which case what you get is no longer "before the measurement"? It is metaphysics (in the perjorative sense of the word) to worry about something that cannot, by its nature, be tested. Pauli said, "One should no more rack one's brain about the problem of whether something one cannot know anything about exists all the same, than about the ancient question of how many angels are able to sit on the point of a needle."[5] For decades this was the "fall-back" position of most physicists: They'd try to sell you answer 2, but if you were persistent they'd switch to 3 and terminate the conversation.

Until fairly recently, all three positions (realist, orthodox, and agnostic) had their partisans. But in 1964 John Bell astonished the physics community by showing that it makes an *observable* difference if the particle had a precise (though unknown) position prior to the measurement. Bell's discovery effectively eliminated agnosticism as a viable option, and made it an *experimental* question whether 1 or 2 is the correct choice. I'll return to this story at the end of the book, when you will be in a better position to appreciate Bell's theorem; for now, suffice it to say that the experiments have confirmed decisively the orthodox interpretation[6]: A particle simply does not have a precise position prior to measurement, any more than the ripples on a pond do; it is the measurement process that insists on one particular number, and thereby in a sense *creates* the specific result, limited only by the statistical weighting imposed by the wave function.

But what if I made a *second* measurement, immediately after the first? Would I get C again, or does the act of measurement cough up some completely new number each time? On this question everyone is in agreement: A repeated measurement (on the same particle) must return the same value. Indeed, it would be tough to prove that the particle was really found at C in the first instance if this could not be confirmed by immediate repetition of the measurement. How does the orthodox interpretation account for the fact that the second measurement is bound to give the value C? Evidently the first measurement radically alters the wave function, so that it is now sharply peaked about C (Figure 1.3). We say that the wave function **collapses** upon measurement, to a spike at the point C (Ψ soon spreads out again, in accordance with the Schrödinger equation, so the second measurement must be made quickly). There

---

[5] Quoted by Mermin (previous footnote), p. 40.

[6] This statement is a little too strong: There remain a few theoretical and experimental loopholes, some of which I shall discuss in the Afterword. And there exist other formulations (such as the **many worlds** interpretation) that do not fit cleanly into any of my three categories. But I think it is wise, at least from a pedagogical point of view, to adopt a clear and coherent platform at this stage, and worry about the alternatives later.

Figure 1.3: Collapse of the wave function: graph of $|\Psi|^2$ immediately after a measurement has found the particle at point C.

are, then, two entirely distinct kinds of physical processes: "ordinary" ones, in which the wave function evolves in a leisurely fashion under the Schrödinger equation, and "measurements", in which Ψ suddenly and discontinuously collapses.[7]

1.3 PROBABILITY

Because of the statistical interpretation, **probability** plays a central role in quantum mechanics, so I digress now for a brief discussion of the theory of probability. It is mainly a question of introducing some notation and terminology, and I shall do it in the context of a simple example.

Imagine a room containing 14 people, whose ages are as follows:

one person aged 14
one person aged 15
three people aged 16
two people aged 22
two people aged 24
five people aged 25.

If we let $N(j)$ represent the number of people of age j, then

[7]The role of measurement in quantum mechanics is so critical and so bizarre that you may well be wondering what precisely *constitutes* a measurement. Does it have to do with the interaction between a microscopic (quantum) system and a macroscopic (classical) measuring apparatus (as Bohr insisted), or is it characterized by the leaving of a permanent "record" (as Heisenberg claimed), or does it involve the intervention of a conscious "observer" (as Wigner proposed)? I'll return to this thorny issue in the Afterword; for the moment let's take the naive view: A measurement is the kind of thing that a scientist does in the laboratory, with rulers, stopwatches, Geiger counters, and so on.

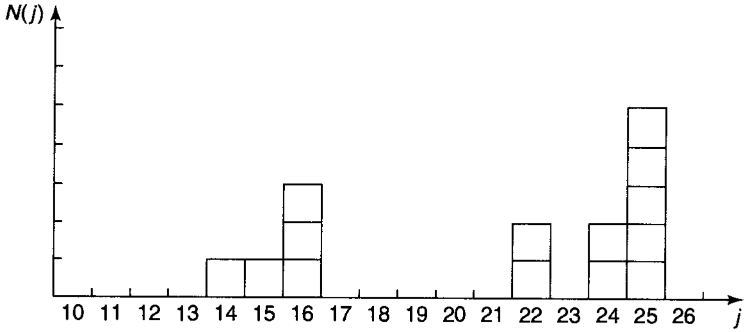

Figure 1.4: Histogram showing the number of people, $N(j)$, with age j, for the example in Section 1.3.

$$\begin{aligned} N(14) &= 1 \\ N(15) &= 1 \\ N(16) &= 3 \\ N(22) &= 2 \\ N(24) &= 2 \\ N(25) &= 5 \end{aligned}$$

while $N(17)$, for instance, is zero. The *total* number of people in the room is

$$N = \sum_{j=0}^{\infty} N(j). \qquad [1.4]$$

(In this instance, of course, $N = 14$.) Figure 1.4 is a histogram of the data. The following are some questions one might ask about this distribution.

Question 1. If you selected one individual at random from this group, what is the **probability** that this person's age would be 15? *Answer:* One chance in 14, since there are 14 possible choices, all equally likely, of whom only one has this particular age. If $P(j)$ is the probability of getting age j, then $P(14) = 1/14$, $P(15) = 1/14$, $P(16) = 3/14$, and so on. In general,

$$P(j) = \frac{N(j)}{N}. \qquad [1.5]$$

Notice that the probability of getting *either* 14 *or* 15 is the *sum* of the individual probabilities (in this case, 1/7). In particular, the sum of *all* the probabilities is 1—you're *certain* to get *some* age:

$$\sum_{j=1}^{\infty} P(j) = 1. \qquad [1.6]$$

Question 2. What is the **most probable** age? *Answer:* 25, obviously; five people share this age, whereas at most three have any other age. In general, the most probable j is the j for which $P(j)$ is a maximum.

Question 3. What is the **median** age? *Answer:* 23, for 7 people are younger than 23, and 7 are older. (In general, the median is that value of j such that the probability of getting a larger result is the same as the probability of getting a smaller result.)

Question 4. What is the **average** (or **mean**) age? *Answer:*

$$\frac{(14) + (15) + 3(16) + 2(22) + 2(24) + 5(25)}{14} = \frac{294}{14} = 21.$$

In general, the average value of j (which we shall write thus: $\langle j \rangle$) is given by

$$\langle j \rangle = \frac{\sum jN(j)}{N} = \sum_{j=0}^{\infty} jP(j). \qquad [1.7]$$

Notice that there need not be anyone with the average age or the median age—in this example nobody happens to be 21 or 23. In quantum mechanics the average is usually the quantity of interest; in that context it has come to be called the **expectation value.** It's a misleading term, since it suggests that this is the outcome you would be most likely to get if you made a single measurement (*that* would be the *most probable value*, not the average value)—but I'm afraid we're stuck with it.

Question 5. What is the average of the *squares* of the ages? *Answer:* You could get $14^2 = 196$, with probability 1/14, or $15^2 = 225$, with probability 1/14, or $16^2 = 256$, with probability 3/14, and so on. The average, then, is

$$\langle j^2 \rangle = \sum_{j=0}^{\infty} j^2 P(j). \qquad [1.8]$$

In general, the average value of some *function* of j is given by

$$\langle f(j) \rangle = \sum_{j=0}^{\infty} f(j)P(j). \qquad [1.9]$$

(Equations 1.6, 1.7, and 1.8 are, if you like, special cases of this formula.) *Beware:* The average of the squares ($\langle j^2 \rangle$) is *not* ordinarily equal to the square of the average ($\langle j \rangle^2$). For instance, if the room contains just two babies, aged 1 and 3, then $\langle x^2 \rangle = 5$, but $\langle x \rangle^2 = 4$.

Now, there is a conspicuous difference between the two histograms in Figure 1.5, even though they have the same median, the same average, the same most probable value, and the same number of elements: The first is sharply peaked about the average value, whereas the second is broad and flat. (The first might represent the age profile for students in a big-city classroom, and the second the pupils in a one-room schoolhouse.) We need a numerical measure of the amount of "spread" in a

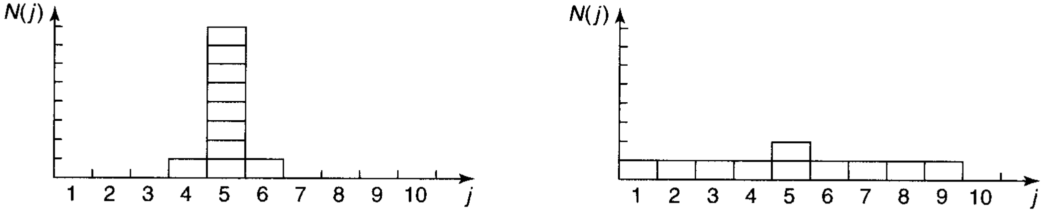

Figure 1.5: Two histograms with the same median, same average, and same most probable value, but different standard deviations.

distribution, with respect to the average. The most obvious way to do this would be to find out how far each individual deviates from the average,

$$\Delta j = j - \langle j \rangle, \qquad [1.10]$$

and compute the average of Δj. Trouble is, of course, that you get *zero*, since, by the nature of the average, Δj is as often negative as positive:

$$\begin{aligned}\langle \Delta j \rangle = \sum (j - \langle j \rangle) P(j) &= \sum j P(j) - \langle j \rangle \sum P(j) \\ &= \langle j \rangle - \langle j \rangle = 0.\end{aligned}$$

(Note that $\langle j \rangle$ is constant—it does not change as you go from one member of the sample to another—so it can be taken outside the summation.) To avoid this irritating problem, you might decide to average the *absolute value* of Δj. But absolute values are nasty to work with; instead, we get around the sign problem by *squaring* before averaging:

$$\sigma^2 \equiv \langle (\Delta j)^2 \rangle. \qquad [1.11]$$

This quantity is known as the **variance** of the distribution; σ itself (the square root of the average of the square of the deviation from the average—gulp!) is called the **standard deviation**. The latter is the customary measure of the spread about $\langle j \rangle$.

There is a useful little theorem involving standard deviations:

$$\begin{aligned}\sigma^2 = \langle (\Delta j)^2 \rangle &= \sum (\Delta j)^2 P(j) = \sum (j - \langle j \rangle)^2 P(j) \\ &= \sum (j^2 - 2j\langle j \rangle + \langle j \rangle^2) P(j) \\ &= \sum j^2 P(j) - 2\langle j \rangle \sum j P(j) + \langle j \rangle^2 \sum P(j) \\ &= \langle j^2 \rangle - 2\langle j \rangle \langle j \rangle + \langle j \rangle^2,\end{aligned}$$

or

$$\sigma^2 = \langle j^2 \rangle - \langle j \rangle^2. \qquad [1.12]$$

Equation 1.12 provides a faster method for computing σ: Simply calculate $\langle j^2 \rangle$ and $\langle j \rangle^2$, and subtract. Incidentally, I warned you a moment ago that $\langle j^2 \rangle$ is not, in general, equal to $\langle j\rangle^2$. Since σ^2 is plainly nonnegative (from its definition in Equation 1.11), Equation 1.12 implies that

$$\langle j^2\rangle \geq \langle j\rangle^2, \tag{1.13}$$

and the two are equal only when $\sigma = 0$, which is to say, for distributions with no spread at all (every member having the same value).

So far, I have assumed that we are dealing with a *discrete* variable—that is, one that can take on only certain isolated values (in the example, j had to be an integer, since I gave ages only in years). But it is simple enough to generalize to *continuous* distributions. If I select a random person off the street, the probability that her age is *precisely* 16 years, 4 hours, 27 minutes, and 3.3333 seconds is *zero*. The only sensible thing to speak about is the probability that her age lies in some *interval*—say, between 16 years, and 16 years plus one day. If the interval is sufficiently short, this probability is proportional to the length of the interval. For example, the chance that her age is between 16 and 16 plus *two* days is presumably twice the probability that it is between 16 and 16 plus *one* day. (Unless, I suppose, there was some extraordinary baby boom 16 years ago, on exactly those days—in which case we have chosen an interval too long for the rule to apply. If the baby boom lasted six hours, we'll take intervals of a second or less, to be on the safe side. Technically, we're talking about infinitesimal intervals.) Thus

$$\left\{ \begin{array}{l} \text{probability that individual (chosen at random)} \\ \text{lies between } x \text{ and } (x+dx) \end{array} \right\} = \rho(x)\,dx. \tag{1.14}$$

The proportionality factor, $\rho(x)$, is often loosely called "the probability of getting x," but this is sloppy language; a better term is **probability density**. The probability that x lies between a and b (a *finite* interval) is given by the integral of $\rho(x)$:

$$P_{ab} = \int_a^b \rho(x)\,dx, \tag{1.15}$$

and the rules we deduced for discrete distributions translate in the obvious way:

$$\int_{-\infty}^{+\infty} \rho(x)\,dx = 1, \tag{1.16}$$

$$\langle x\rangle = \int_{-\infty}^{+\infty} x\rho(x)\,dx, \tag{1.17}$$

$$\langle f(x)\rangle = \int_{-\infty}^{+\infty} f(x)\rho(x)\,dx, \tag{1.18}$$

$$\sigma^2 \equiv \langle(\Delta x)^2\rangle = \langle x^2\rangle - \langle x\rangle^2. \tag{1.19}$$

*__Problem 1.1__ For the distribution of ages in the example in Section 1.3,

(a) Compute $\langle j^2 \rangle$ and $\langle j \rangle^2$.

(b) Determine Δj for each j, and use Equation 1.11 to compute the standard deviation.

(c) Use your results in (a) and (b) to check Equation 1.12.

Problem 1.2 Consider the first 25 digits in the decimal expansion of π (3, 1, 4, 1, 5, 9, ...).

(a) If you selected one number at random from this set, what are the probabilities of getting each of the 10 digits?

(b) What is the most probable digit? What is the median digit? What is the average value?

(c) Find the standard deviation for this distribution.

Problem 1.3 The needle on a broken car speedometer is free to swing, and bounces perfectly off the pins at either end, so that if you give it a flick it is equally likely to come to rest at any angle between 0 and π.

(a) What is the probability density, $\rho(\theta)$? [$\rho(\theta)\, d\theta$ is the probability that the needle will come to rest between θ and $(\theta + d\theta)$.] Graph $\rho(\theta)$ as a function of θ, from $-\pi/2$ to $3\pi/2$. (Of course, part of this interval is excluded, so ρ is zero there.) Make sure that the total probability is 1.

(b) Compute $\langle \theta \rangle$, $\langle \theta^2 \rangle$, and σ for this distribution.

(c) Compute $\langle \sin\theta \rangle$, $\langle \cos\theta \rangle$, and $\langle \cos^2\theta \rangle$.

Problem 1.4 We consider the same device as the previous problem, but this time we are interested in the x-coordinate of the needle point—that is, the "shadow", or "projection", of the needle on the horizontal line.

(a) What is the probability density $\rho(x)$? [$\rho(x)\, dx$ is the probability that the projection lies between x and $(x + dx)$.] Graph $\rho(x)$ as a function of x, from $-2r$ to $+2r$, where r is the length of the needle. Make sure the total probability is 1. [*Hint*: You know (from Problem 1.3) the probability that θ is in a given range; the question is, what interval dx corresponds to the interval $d\theta$?]

(b) Compute $\langle x \rangle$, $\langle x^2 \rangle$, and σ for this distribution. Explain how you could have obtained these results from part (c) of Problem 1.3.

**Problem 1.5 A needle of length l is dropped at random onto a sheet of paper ruled with parallel lines a distance l apart. What is the probability that the needle will cross a line? [*Hint*: Refer to Problem 1.4.]

*Problem 1.6 Consider the **Gaussian** distribution

$$\rho(x) = Ae^{-\lambda(x-a)^2},$$

where A, a, and λ are constants. (Look up any integrals you need.)

(a) Use Equation 1.16 to determine A.

(b) Find $\langle x \rangle$, $\langle x^2 \rangle$, and σ.

(c) Sketch the graph of $\rho(x)$.

1.4 NORMALIZATION

We return now to the statistical interpretation of the wave function (Equation 1.3), which says that $|\Psi(x,t)|^2$ is the probability density for finding the particle at point x, at time t. It follows (Equation 1.16) that the integral of $|\Psi|^2$ must be 1 (the particle's got to be *some*where):

$$\boxed{\int_{-\infty}^{+\infty} |\Psi(x,t)|^2\, dx = 1.} \quad [1.20]$$

Without this, the statistical interpretation would be nonsense.

However, this requirement should disturb you: After all, the wave function is supposed to be determined by the Schrödinger equation—we can't impose an extraneous condition on Ψ without checking that the two are consistent. A glance at Equation 1.1 reveals that if $\Psi(x,t)$ is a solution, so too is $A\Psi(x,t)$, where A is any (complex) constant. What we must do, then, is pick this undetermined multiplicative factor so as to ensure that Equation 1.20 is satisfied. This process is called **normalizing** the wave function. For some solutions to the Schrödinger equation, the integral is *infinite*; in that case *no* multiplicative factor is going to make it 1. The same goes for the trivial solution $\Psi = 0$. Such **non-normalizable** solutions cannot represent particles, and must be rejected. Physically realizable states correspond to the "square-integrable" solutions to Schrödinger's equation.[8]

[8]Evidently $\Psi(x,t)$ must go to zero faster than $1/\sqrt{|x|}$, as $|x| \to \infty$. Incidentally, normalization only fixes the *modulus* of A; the *phase* remains undetermined. However, as we shall see, the latter carries no physical significance anyway.

But wait a minute! Suppose I have normalized the wave function at time $t = 0$. How do I know that it will *stay* normalized, as time goes on and Ψ evolves? (You can't keep *re*normalizing the wave function, for then A becomes a function of t, and you no longer have a solution to the Schrödinger equation.) Fortunately, the Schrödinger equation has the property that it automatically preserves the normalization of the wave function—without this crucial feature the Schrödinger equation would be incompatible with the statistical interpretation, and the whole theory would crumble. So we'd better pause for a careful proof of this point:

$$\frac{d}{dt}\int_{-\infty}^{+\infty} |\Psi(x,t)|^2\,dx = \int_{-\infty}^{+\infty} \frac{\partial}{\partial t}|\Psi(x,t)|^2\,dx. \qquad [1.21]$$

[Note that the integral is a function only of t, so I use a total derivative (d/dt) in the first term, but the integrand is a function of x as well as t, so it's a partial derivative $(\partial/\partial t)$ in the second one.] By the product rule,

$$\frac{\partial}{\partial t}|\Psi|^2 = \frac{\partial}{\partial t}(\Psi^*\Psi) = \Psi^*\frac{\partial \Psi}{\partial t} + \frac{\partial \Psi^*}{\partial t}\Psi \qquad [1.22]$$

Now the Schrödinger equation says that

$$\frac{\partial \Psi}{\partial t} = \frac{i\hbar}{2m}\frac{\partial^2 \Psi}{\partial x^2} - \frac{i}{\hbar}V\Psi, \qquad [1.23]$$

and hence also (taking the complex conjugate of Equation 1.23)

$$\frac{\partial \Psi^*}{\partial t} = -\frac{i\hbar}{2m}\frac{\partial^2 \Psi^*}{\partial x^2} + \frac{i}{\hbar}V\Psi^*, \qquad [1.24]$$

so

$$\frac{\partial}{\partial t}|\Psi|^2 = \frac{i\hbar}{2m}\left(\Psi^*\frac{\partial^2\Psi}{\partial x^2} - \frac{\partial^2\Psi^*}{\partial x^2}\Psi\right) = \frac{\partial}{\partial x}\left[\frac{i\hbar}{2m}\left(\Psi^*\frac{\partial\Psi}{\partial x} - \frac{\partial\Psi^*}{\partial x}\Psi\right)\right]. \qquad [1.25]$$

The integral (Equation 1.21) can now be evaluated explicitly:

$$\frac{d}{dt}\int_{-\infty}^{+\infty} |\Psi(x,t)|^2\,dx = \frac{i\hbar}{2m}\left(\Psi^*\frac{\partial\Psi}{\partial x} - \frac{\partial\Psi^*}{\partial x}\Psi\right)\Bigg|_{-\infty}^{+\infty}. \qquad [1.26]$$

But $\Psi(x, t)$ must go to zero as x goes to ($\pm$) infinity—otherwise the wave function would not be normalizable. It follows that

$$\frac{d}{dt}\int_{-\infty}^{+\infty} |\Psi(x,t)|^2\,dx = 0, \qquad [1.27]$$

and hence that the integral on the left is *constant* (independent of time); if Ψ is normalized at $t = 0$, it *stays* normalized for all future time. QED

Problem 1.7 At time $t = 0$ a particle is represented by the wave function

$$\Psi(x,0) = \begin{cases} Ax/a, & \text{if } 0 \leq x \leq a, \\ A(b-x)/(b-a), & \text{if } a \leq x \leq b, \\ 0, & \text{otherwise,} \end{cases}$$

where A, a, and b are constants.

(a) Normalize Ψ (that is, find A in terms of a and b).

(b) Sketch $\Psi(x, 0)$ as a function of x.

(c) Where is the particle most likely to be found, at $t = 0$?

(d) What is the probability of finding the particle to the left of a? Check your result in the limiting cases $b = a$ and $b = 2a$.

(e) What is the expectation value of x?

∗**Problem 1.8** Consider the wave function

$$\Psi(x,t) = Ae^{-\lambda|x|}e^{-i\omega t},$$

where A, λ, and ω are positive real constants. [We'll see in Chapter 2 what potential (V) actually produces such a wave function.]

(a) Normalize Ψ.

(b) Determine the expectation values of x and x^2.

(c) Find the standard deviation of x. Sketch the graph of $|\Psi|^2$, as a function of x, and mark the points $(\langle x\rangle + \sigma)$ and $(\langle x\rangle - \sigma)$ to illustrate the sense in which σ represents the "spread" in x. What is the probability that the particle would be found outside this range?

Problem 1.9 Let $P_{ab}(t)$ be the probability of finding the particle in the range $(a < x < b)$, at time t.

(a) Show that

$$\frac{dP_{ab}}{dt} = J(a,t) - J(b,t)$$

where

$$J(x,t) \equiv \frac{i\hbar}{2m}\left(\Psi\frac{\partial\Psi^*}{\partial x} - \Psi^*\frac{\partial\Psi}{\partial x}\right).$$

What are the units of $J(x, t)$? [J is called the **probability current**, because it tells you the rate at which probability is "flowing" past the point x. If $P_{ab}(t)$ is increasing, then more probability is flowing into the region at one end than flows out at the other.]

(b) Find the probability current for the wave function in the previous problem. (This is not a very pithy example, I'm afraid; we'll encounter some more substantial ones in due course.)

****Problem 1.10** Suppose you wanted to describe an **unstable particle** that spontaneously disintegrates with a "lifetime" τ. In that case the total probability of finding the particle somewhere should *not* be constant, but should decrease at (say) an exponential rate:

$$P(t) \equiv \int_{-\infty}^{+\infty} |\Psi(x,t)|^2\, dx = e^{-t/\tau}.$$

A crude way of achieving this result is as follows. In Equation 1.24 we tacitly assumed that V (the potential energy) is *real.* That is certainly reasonable, but it leads to the **conservation of probability** enshrined in Equation 1.27. What if we assign to V an imaginary part:

$$V = V_0 - i\Gamma,$$

where V_0 is the true potential energy and Γ is a positive real constant?

(a) Show that (in place of Equation 1.27) we now get

$$\frac{dP}{dt} = -\frac{2\Gamma}{\hbar}P.$$

(b) Solve for $P(t)$, and find the lifetime of the particle in terms of Γ.

1.5 MOMENTUM

For a particle in state Ψ, the expectation value of x is

$$\boxed{\langle x \rangle = \int_{-\infty}^{+\infty} x|\Psi(x,t)|^2\, dx.} \qquad [1.28]$$

What exactly does this mean? It emphatically does *not* mean that if you measure the position of one particle over and over again, $\int x|\Psi|^2 dx$ is the average of the results you'll get. On the contrary, the first measurement (whose outcome is indeterminate) will collapse the wave function to a spike at the value actually obtained, and the subsequent measurements (if they're performed quickly) will simply repeat that same result. Rather, $\langle x \rangle$ is the average of measurements performed on particles *all in the state* Ψ, which means that either you must find some way of returning the particle to its original state after each measurement, or else you prepare a whole ensemble of particles, each in the same state Ψ, and measure the positions of all of them: $\langle x \rangle$ is the average of *these* results. [I like to picture a row of bottles on a shelf, each containing a particle in the state Ψ (relative to the center of the bottle). A graduate student with a ruler is assigned to each bottle, and at a signal they all measure the positions of their respective particles. We then construct a histogram of the results, which should match $|\Psi|^2$, and compute the average, which should agree with $\langle x \rangle$. (Of course, since we're only using a finite sample, we can't expect perfect agreement, but the more bottles we use, the closer we ought to come.)] In short, the expectation value is the average of repeated measurements on an ensemble of identically prepared systems, *not* the average of repeated measurements on one and the same system.

Now, as time goes on, $\langle x \rangle$ will change (because of the time dependence of Ψ), and we might be interested in knowing how fast it moves. Referring to Equations 1.25 and 1.28, we see that[9]

$$\frac{d\langle x \rangle}{dt} = \int x \frac{\partial}{\partial t}|\Psi|^2\, dx = \frac{i\hbar}{2m} \int x \frac{\partial}{\partial x}\left(\Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x}\Psi\right) dx. \qquad [1.29]$$

This expression can be simplified using integration by parts[10]:

$$\frac{d\langle x \rangle}{dt} = -\frac{i\hbar}{2m} \int \left(\Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x}\Psi\right) dx. \qquad [1.30]$$

[I used the fact that $\partial x/\partial x = 1$, and threw away the boundary term, on the ground that Ψ goes to zero at ($\pm$) infinity.] Performing another integration by parts on the second term, we conclude that

$$\frac{d\langle x \rangle}{dt} = -\frac{i\hbar}{m} \int \Psi^* \frac{\partial \Psi}{\partial x}\, dx. \qquad [1.31]$$

What are we to make of this result? Note that we're talking about the "velocity" of the *expectation* value of x, which is not the same thing as the velocity of the *particle*. Nothing we have seen so far would enable us to calculate the velocity of a particle—it's not even clear what velocity *means* in quantum mechanics. If the particle doesn't have a determinate position (prior to measurement), neither does it have a well-defined velocity. All we could reasonably ask for is the *probability* of getting a particular value. We'll see in Chapter 3 how to construct the probability density for velocity,

[9]To keep things from getting too cluttered, I suppress the limits of integration when they are $\pm\infty$.

[10]The product rule says that

$$\frac{d}{dx}(fg) = f\frac{dg}{dx} + \frac{df}{dx}g,$$

from which it follows that

$$\int_a^b f\frac{dg}{dx}\, dx = -\int_a^b \frac{df}{dx} g\, dx + fg\Big|_a^b.$$

Under the integral sign, then, you can peel a derivative off one factor in a product and slap it onto the other one—it'll cost you a minus sign, and you'll pick up a boundary term.

given Ψ; for our present purposes it will suffice to postulate that the *expectation value of the velocity is equal to the time derivative of the expectation value of position*:

$$\langle v \rangle = \frac{d\langle x \rangle}{dt}. \tag{1.32}$$

Equation 1.31 tells us, then, how to calculate $\langle v \rangle$ directly from Ψ.

Actually, it is customary to work with momentum ($p = mv$), rather than velocity:

$$\boxed{\langle p \rangle = m\frac{d\langle x \rangle}{dt} = -i\hbar \int \left(\Psi^* \frac{\partial \Psi}{\partial x} \right) dx.} \tag{1.33}$$

Let me write the expressions for $\langle x \rangle$ and $\langle p \rangle$ in a more suggestive way:

$$\langle x \rangle = \int \Psi^* (x) \Psi \, dx, \tag{1.34}$$

$$\langle p \rangle = \int \Psi^* (\frac{\hbar}{i}\frac{\partial}{\partial x}) \Psi \, dx. \tag{1.35}$$

We say that the **operator**[11] x "represents" position, and the operator $(\hbar/i)(\partial/\partial x)$ "represents" momentum, in quantum mechanics; to calculate expectation values, we "sandwich" the appropriate operator between Ψ^* and Ψ, and integrate.

That's cute, but what about other dynamical variables? The fact is, *all* such quantities can be written in terms of position and momentum. Kinetic energy, for example, is

$$T = \frac{1}{2}mv^2 = \frac{p^2}{2m},$$

and angular momentum is

$$\mathbf{L} = \mathbf{r} \times m\mathbf{v} = \mathbf{r} \times \mathbf{p}$$

(the latter, of course, does not occur for motion in one dimension). To calculate the expectation value of such a quantity, we simply replace every p by $(\hbar/i)(\partial/\partial x)$, insert the resulting operator between Ψ^* and Ψ, and integrate:

$$\boxed{\langle Q(x, p) \rangle = \int \Psi^* Q(x, \frac{\hbar}{i}\frac{\partial}{\partial x}) \Psi \, dx.} \tag{1.36}$$

[11] An operator is an instruction to *do something* to the function that follows. The position operator tells you to *multiply* by x; the momentum operator tells you to *differentiate* with respect to x (and multiply the result by $-i\hbar$). In this book *all* operators will be derivatives (d/dt, d^2/dt^2, $\partial^2/\partial x \partial y$, etc.) or multipliers (2, i, x^2, etc.) or combinations of these.

For example,

$$\langle T \rangle = \frac{-\hbar^2}{2m} \int \Psi^* \frac{\partial^2 \Psi}{\partial x^2} \, dx. \qquad [1.37]$$

Equation 1.36 is a recipe for computing the expectation value of any dynamical quantity for a particle in state Ψ; it subsumes Equations 1.34 and 1.35 as special cases. I have tried in this section to make Equation 1.36 seem plausible, given Born's statistical interpretation, but the truth is that this equation represents such a radically new way of doing business (as compared with classical mechanics) that it's a good idea to get some practice using it before we come back (in Chapter 3) and put it on a firmer theoretical foundation. In the meantime, if you prefer to think of it as an *axiom*, that's fine with me.

Problem 1.11 Why can't you do integration by parts directly on the middle expression in Equation 1.29—pull the time derivative over onto x, note that $\partial x / \partial t = 0$, and conclude that $d\langle x \rangle / dt = 0$?

$*$**Problem 1.12** Calculate $d\langle p \rangle / dt$. *Answer*:

$$\frac{d\langle p \rangle}{dt} = \langle -\frac{\partial V}{\partial x} \rangle. \qquad [1.38]$$

(This is known as **Ehrenfest's theorem**; it tells us that *expectation values* obey Newton's second law.)

Problem 1.13 Suppose you add a constant V_0 to the potential energy (by "constant" I mean independent of x as well as t). In *classical* mechanics this doesn't change anything, but what about *quantum* mechanics? Show that the wave function picks up a time-dependent phase factor: $\exp(-iV_0 t/\hbar)$. What effect does this have on the expectation value of a dynamical variable?

1.6 THE UNCERTAINTY PRINCIPLE

Imagine that you're holding one end of a very long rope, and you generate a wave by shaking it up and down rhythmically (Figure 1.6). If someone asked you, "Precisely where *is* that wave?" you'd probably think he was a little bit nutty: The wave isn't precisely *any*where—it's spread out over 50 feet or so. On the other hand, if he asked you what its *wavelength* is, you could give him a reasonable answer: It looks like about 6 feet. By contrast, if you gave the rope a sudden jerk (Figure 1.7), you'd get a relatively narrow bump traveling down the line. This time the first question (Where precisely is the wave?) is a sensible one, and the second (What is its wavelength?) seems nutty—it isn't even vaguely periodic, so how can you assign a wavelength to it?

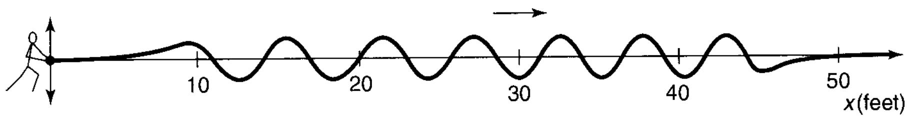

Figure 1.6: A wave with a (fairly) well-defined wavelength but an ill-defined position.

Of course, you can draw intermediate cases, in which the wave is *fairly* well localized and the wavelength is *fairly* well defined, but there is an inescapable trade-off here: The more precise a wave's position is, the less precise is its wavelength, and vice versa.[12] A theorem in Fourier analysis makes all this rigorous, but for the moment I am only concerned with the qualitative argument.

This applies, of course, to *any* wave phenomenon, and hence in particular to the quantum mechanical wave function. Now the wavelength of Ψ is related to the *momentum* of the particle by the **de Broglie formula**[13]:

$$p = \frac{h}{\lambda} = \frac{2\pi\hbar}{\lambda}. \qquad [1.39]$$

Thus a spread in *wavelength* corresponds to a spread in *momentum*, and our general observation now says that the more precisely determined a particle's position is, the less precisely its momentum is determined. Quantitatively,

$$\boxed{\sigma_x \sigma_p \geq \frac{\hbar}{2},} \qquad [1.40]$$

where σ_x is the standard deviation in x, and σ_p is the standard deviation in p. This is Heisenberg's famous **uncertainty principle**. (We'll prove it in Chapter 3, but I wanted to mention it here so you can test it out on the examples in Chapter 2.)

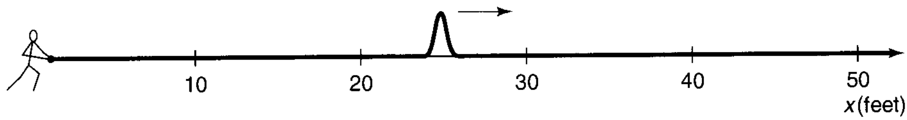

Figure 1.7: A wave with a (fairly) well-defined position but an ill-defined wavelength.

[12]That's why a piccolo player must be right on pitch, whereas a double-bass player can afford to wear garden gloves. For the piccolo, a sixty-fourth note contains many full cycles, and the frequency (we're working in the time domain now, instead of space) is well defined, whereas for the bass, at a much lower register, the sixty-fourth note contains only a few cycles, and all you hear is a general sort of "oomph," with no very clear pitch.

[13]I'll prove this in due course. Many authors take the de Broglie formula as an *axiom*, from which they then deduce the association of momentum with the operator $(\hbar/i)(\partial/\partial x)$. Although this is a conceptually cleaner approach, it involves diverting mathematical complications that I would rather save for later.

Please understand what the uncertainty principle *means*: Like position measurements, momentum measurements yield precise answers—the "spread" here refers to the fact that measurements on identical systems do not yield consistent results. You can, if you want, prepare a system such that repeated position measurements will be very close together (by making Ψ a localized "spike"), but you will pay a price: Momentum measurements on this state will be widely scattered. Or you can prepare a system with a reproducible momentum (by making Ψ a long sinusoidal wave), but in that case position measurements will be widely scattered. And, of course, if you're in a really bad mood you can prepare a system in which neither position nor momentum is well defined: Equation 1.40 is an inequality, and there's no limit on how *big* σ_x and σ_p can be—just make Ψ some long wiggly line with lots of bumps and potholes and no periodic structure.

∗**Problem 1.14** A particle of mass m is in the state

$$\Psi(x,t) = Ae^{-a[(mx^2/\hbar)+it]},$$

where A and a are positive real constants.

(a) Find A.

(b) For what potential energy function $V(x)$ does Ψ satisfy the Schrödinger equation?

(c) Calculate the expectation values of x, x^2, p, and p^2.

(d) Find σ_x and σ_p. Is their product consistent with the uncertainty principle?

CHAPTER 2

THE TIME-INDEPENDENT SCHRÖDINGER EQUATION

2.1 STATIONARY STATES

In Chapter 1 we talked a lot about the wave function and how you use it to calculate various quantities of interest. The time has come to stop procrastinating and confront what is, logically, the prior question: How do you *get* $\Psi(x, t)$ in the first place—how do you go about solving the Schrödinger equation? I shall assume for all of this chapter (and most of this book) that the potential,[1] V, is *independent of* t. In that case the Schrödinger equation can be solved by the method of **separation of variables** (the physicist's first line of attack on any partial differential equation): We look for solutions that are simple *products*,

$$\Psi(x,t) = \psi(x)\, f(t), \tag{2.1}$$

where ψ (*lower*case) is a function of x alone, and f is a function of t alone. On its face, this is an absurd restriction, and we cannot hope to get more than a tiny subset of all solutions in this way. But hang on, because the solutions we *do* obtain turn out to be of great interest. Moreover, as is typically the case with separation of variables, we will be able at the end to patch together the separable solutions in such a way as to construct the most general solution.

[1] It is tiresome to keep saying "potential energy function," so most people just call V the "potential", even though this invites occasional confusion with *electric* potential, which is actually potential energy *per unit charge*.

For separable solutions we have

$$\frac{\partial \Psi}{\partial t} = \psi \frac{df}{dt}, \quad \frac{\partial^2 \Psi}{\partial x^2} = \frac{d^2\psi}{dx^2} f$$

(*ordinary* derivatives, now), and the Schrödinger equation (Equation 1.1) reads

$$i\hbar\psi \frac{df}{dt} = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} f + V\psi f.$$

Or, dividing through by ψf:

$$i\hbar \frac{1}{f}\frac{df}{dt} = -\frac{\hbar^2}{2m}\frac{1}{\psi}\frac{d^2\psi}{dx^2} + V. \qquad [2.2]$$

Now the left side is a function of t alone, and the right side is a function of x alone.[2] The only way this can possibly be true is if both sides are in fact *constant*—otherwise, by varying t, I could change the left side without touching the right side, and the two would no longer be equal. (That's a subtle but crucial argument, so if it's new to you, be sure to pause and think it through.) For reasons that will appear in a moment, we shall call the separation constant E. Then

$$i\hbar \frac{1}{f}\frac{df}{dt} = E,$$

or

$$\frac{df}{dt} = -\frac{iE}{\hbar} f, \qquad [2.3]$$

and

$$-\frac{\hbar^2}{2m}\frac{1}{\psi}\frac{d^2\psi}{dx^2} + V = E,$$

or

$$\boxed{-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V\psi = E\psi.} \qquad [2.4]$$

Separation of variables has turned a *partial* differential equation into two *ordinary* differential equations (Equations 2.3 and 2.4). The first of these is easy to solve (just multiply through by dt and integrate); the general solution is $C\exp(-iEt/\hbar)$, but we might as well absorb the constant C into ψ (since the quantity of interest is the product ψf). Then

$$f(t) = e^{-iEt/\hbar}. \qquad [2.5]$$

The second (Equation 2.4) is called the **time-independent Schrödinger equation**; we can go no further with it until the potential $V(x)$ is specified.

[2]Note that this would *not* be true if V were a function of t as well as x.

The rest of this chapter will be devoted to solving the time-independent Schrödinger equation, for a variety of simple potentials. But before we get to that I would like to consider further the question: *What's so great about separable solutions?* After all, *most* solutions to the (time-dependent) Schrödinger equation do *not* take the form $\psi(x)f(t)$. I offer three answers—two of them physical and one mathematical:

1. They are **stationary states**. Although the wave function itself,

$$\boxed{\Psi(x,t) = \psi(x)e^{-iEt/\hbar},} \tag{2.6}$$

does (obviously) depend on t, the probability density

$$|\Psi(x,t)|^2 = \Psi^*\Psi = \psi^* e^{+iEt/\hbar}\psi e^{-iEt/\hbar} = |\psi(x)|^2 \tag{2.7}$$

does *not*—the time dependence cancels out.[3] The same thing happens in calculating the expectation value of any dynamical variable; Equation 1.36 reduces to

$$\langle Q(x,p)\rangle = \int \psi^* Q(x, \frac{\hbar}{i}\frac{d}{dx})\psi\, dx. \tag{2.8}$$

Every expectation value is constant in time; we might as well drop the factor $f(t)$ altogether, and simply use ψ in place of Ψ. (Indeed, it is common to refer to ψ as "the wave function", but this is sloppy language that can be dangerous, and it is important to remember that the *true* wave function always carries that exponential time-dependent factor.) In particular, $\langle x\rangle$ is constant, and hence (Equation 1.33) $\langle p\rangle = 0$. Nothing ever *happens* in a stationary state.

2. They are states of *definite total energy*. In classical mechanics, the total energy (kinetic plus potential) is called the **Hamiltonian**:

$$H(x,p) = \frac{p^2}{2m} + V(x). \tag{2.9}$$

The corresponding Hamiltonian *operator*, obtained by the canonical substitution $p \to (\hbar/i)(\partial/\partial x)$, is therefore[4]

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x). \tag{2.10}$$

Thus the time-independent Schrödinger equation (Equation 2.4) can be written

$$\hat{H}\psi = E\psi, \tag{2.11}$$

[3]For normalizable solutions, E must be real (see Problem 2.1a).

[4]Whenever confusion might arise, I'll put a "hat" (^) on the operator to distinguish it from the dynamical variable it represents.

and the expectation value of the total energy is

$$\langle H \rangle = \int \psi^* \hat{H} \psi \, dx = E \int |\psi|^2 \, dx = E. \qquad [2.12]$$

(Note that the normalization of Ψ entails the normalization of ψ.) Moreover,

$$\hat{H}^2 \psi = \hat{H}(\hat{H}\psi) = \hat{H}(E\psi) = E(\hat{H}\psi) = E^2 \psi,$$

and hence

$$\langle H^2 \rangle = \int \psi^* \hat{H}^2 \psi \, dx = E^2 \int |\psi|^2 \, dx = E^2.$$

So the standard deviation in H is given by

$$\sigma_H^2 = \langle H^2 \rangle - \langle H \rangle^2 = E^2 - E^2 = 0. \qquad [2.13]$$

But remember, if $\sigma = 0$, then every member of the sample must share the same value (the distribution has zero spread). Conclusion: A separable solution has the property that *every measurement of the total energy is certain to return the value E*. (That's why I chose that letter for the separation constant.)

3. The general solution is a **linear combination** of separable solutions. As we're about to discover, the time-independent Schrödinger equation (Equation 2.4) yields an infinite collection of solutions ($\psi_1(x)$, $\psi_2(x)$, $\psi_3(x)$, ...), each with its associated value of the separation constant (E_1, E_2, E_3, ...); thus there is a different wave function for each **allowed energy**:

$$\Psi_1(x, t) = \psi_1(x) e^{-iE_1 t/\hbar}, \quad \Psi_2(x, t) = \psi_2(x) e^{-iE_2 t/\hbar}, \ldots.$$

Now (as you can easily check for yourself) the (time-*de*pendent) Schrödinger equation (Equation 1.1) has the property that any linear combination[5] of solutions is itself a solution. Once we have found the separable solutions, then, we can immediately construct a much more general solution, of the form

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n(x) e^{-iE_n t/\hbar}. \qquad [2.14]$$

It so happens that *every* solution to the (time-dependent) Schrödinger equation can be written in this form—it is simply a matter of finding the right constants (c_1, c_2, ...) so as to fit the initial conditions for the problem at hand. You'll see in the following sections how all this works out in practice, and in Chapter 3 we'll put it into more elegant language, but the main point is this: Once you've solved the time-*in*dependent

[5]A linear combination of the functions $f_1(z)$, $f_2(z)$, ... is an expression of the form

$$f(z) = c_1 f_1(z) + c_2 f_2(z) + \cdots,$$

where $c_1, c_2, \ldots$ are any (complex) constants.

Schrödinger equation, you're essentially done; getting from there to the general solution of the time-*dependent* Schrödinger equation is simple and straightforward.

***Problem 2.1** Prove the following theorems:

(a) For normalizable solutions, the separation constant E must be *real*. *Hint*: Write E (in Equation 2.6) as $E_0 + i\Gamma$ (with E_0 and Γ real), and show that if Equation 1.20 is to hold for all t, Γ must be zero.

(b) ψ can always be taken to be *real* (unlike Ψ, which is necessarily complex). *Note*: This doesn't mean that every solution to the time-independent Schrödinger equation *is* real; what it says is that if you've got one that is *not*, it can always be expressed as a linear combination of solutions (with the same energy) that *are*. So in Equation 2.14 you *might as well* stick to ψ's that are real. *Hint*: If $\psi(x)$ satisfies the time-independent Schrödinger equation for a given E, so too does its complex conjugate, and hence also the real linear combinations $(\psi + \psi^*)$ and $i(\psi - \psi^*)$.

(c) If $V(x)$ is an *even* function [i.e., $V(-x) = V(x)$], then $\psi(x)$ can always be taken to be either even or odd. *Hint*: If $\psi(x)$ satisfies the time-independent Schrödinger equation for a given E, so too does $\psi(-x)$, and hence also the even and odd linear combinations $\psi(x) \pm \psi(-x)$.

***Problem 2.2** Show that E must exceed the minimum value of $V(x)$ for every normalizable solution to the time-independent Schrödinger equation. What is the classical analog to this statement? *Hint*: Rewrite Equation 2.4 in the form

$$\frac{d^2\psi}{dx^2} = \frac{2m}{\hbar^2}[V(x) - E]\psi;$$

if $E < V_{\min}$, then ψ and its second derivative always have the *same sign*—argue that such a function cannot be normalized.

2.2 THE INFINITE SQUARE WELL

Suppose

$$V(x) = \begin{cases} 0, & \text{if } 0 \leq x \leq a, \\ \infty, & \text{otherwise} \end{cases} \qquad [2.15]$$

(Figure 2.1). A particle in this potential is completely free, except at the two ends ($x = 0$ and $x = a$), where an infinite force prevents it from escaping. A classical model would be a cart on a frictionless horizontal air track, with perfectly elastic bumpers—it just keeps bouncing back and forth forever. (This potential is awfully artificial, but I urge you to treat it with respect. Despite its simplicity—or rather, orbital angular momentum L and its conjugate variable φ, the polar angle, where φ is periodic in time. That is, $J = \oint p\, dq$ is given in polar coordinates by $\int_0^{2\pi} L\, d\varphi$. In this case (1.86) becomes

$$\int_0^{2\pi} L\, d\varphi = nh. \tag{1.92}$$

For spherically symmetric potentials—as it is the case here where the electron experiences the proton's Coulomb potential—the angular momentum L is a constant of the motion. Hence (1.92) shows that angular momentum can change only in integral units of $\hbar$:

$$L \int_0^{2\pi} d\varphi = nh \qquad \Longrightarrow \qquad L = n\frac{h}{2\pi} = n\hbar, \tag{1.93}$$

which is identical with the Bohr quantization condition (1.63). This calculation also shows that the *Bohr quantization is equivalent to the quantization of action*. As stated above (1.78), the Bohr quantization condition (1.63) has the following physical meaning: while orbiting the nucleus, the electron moves only in well specified orbits, orbits with circumferences equal to integral multiples of the de Broglie wavelength.

Note that the Wilson–Sommerfeld quantization rule (1.86) does not tell us how to calculate the energy levels of non-periodic systems; it applies only to systems which are periodic. On a historical note, the quantization rules of Planck and Bohr have dominated quantum physics from 1900 to 1925; the quantum physics of this period is known as the "old quantum theory." The success of these quantization rules, as measured by the striking agreement of their results with experiment, gave irrefutable evidence for the quantization hypothesis of all material systems and constituted a triumph of the "old quantum theory." In spite of their quantitative success, these quantization conditions suffer from a serious inconsistency: they do not originate from a theory, they were postulated rather arbitrarily.

1.8 Wave Packets

At issue here is how to describe a particle within the context of quantum mechanics. As quantum particles jointly display particle and wave features, we need to look for a mathematical scheme that can embody them simultaneously.

In classical physics, a particle is *well localized* in space, for its position and velocity can be calculated simultaneously to arbitrary precision. As for quantum mechanics, it describes a material particle by a *wave function* corresponding to the matter wave associated with the particle (de Broglie's conjecture). Wave functions, however, depend on the *whole* space; hence they *cannot be localized*. If the wave function is made to vanish everywhere except in the neighborhood of the particle or the neighborhood of the "classical trajectory," it can then be used to describe the dynamics of the particle. That is, a particle which is localized within a certain region of space can be described by a matter wave whose amplitude is large in that region and zero outside it. This matter wave must then be *localized* around the region of space within which the particle is confined.

A *localized* wave function is called a *wave packet*. A wave packet therefore consists of a group of waves of slightly different wavelengths, with phases and amplitudes so chosen that they interfere constructively over a small region of space and destructively elsewhere. Not only are wave packets useful in the description of "isolated" particles that are confined to a certain spatial region, they also play a key role in understanding the connection between quantum mechanics and classical mechanics. The wave packet concept therefore represents a unifying mathematical tool that can cope with and embody nature's particle-like behavior and also its wave-like behavior.

1.8.1 Localized Wave Packets

Localized wave packets can be constructed by superposing, in the same region of space, waves of slightly different wavelengths, but with phases and amplitudes chosen to make the superposition constructive in the desired region and destructive outside it. Mathematically, we can carry out this superposition by means of *Fourier transforms*. For simplicity, we are going to consider a one-dimensional wave packet; this packet is intended to describe a "classical" particle confined to a one-dimensional region, for instance, a particle moving along the x-axis. We can construct the packet $\psi(x, t)$ by superposing plane waves (propagating along the x-axis) of different frequencies (or wavelengths):

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{i(kx - \omega t)} dk; \tag{1.94}$$

$\phi(k)$ is the amplitude of the wave packet.

In what follows we want to look at the form of the packet at a given time; we will deal with the time evolution of wave packets later. Choosing this time to be $t = 0$ and abbreviating $\psi(x, 0)$ by $\psi_0(x)$, we can reduce (1.94) to

$$\psi_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ikx} dk, \tag{1.95}$$

where $\phi(k)$ is the Fourier transform of $\psi_0(x)$,

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi_0(x) e^{-ikx} dx. \tag{1.96}$$

The relations (1.95) and (1.96) show that $\phi(k)$ determines $\psi_0(x)$ and vice versa. The packet (1.95), whose form is determined by the x-dependence of $\psi_0(x)$, does indeed have the required property of localization: $|\psi_0(x)|$ peaks at $x = 0$ and vanishes far away from $x = 0$. On the one hand, as $x \to 0$ we have $e^{ikx} \to 1$; hence the waves of different frequencies interfere constructively (i.e., the various k-integrations in (1.95) add constructively). On the other hand, far away from $x = 0$ (i.e., $|x| \gg 0$) the phase e^{ikx} goes through many periods leading to violent oscillations, thereby yielding destructive interference (i.e., the various k-integrations in (1.95) add up to zero). This implies, in the language of Born's probabilistic interpretation, that the particle has a greater probability of being found near $x = 0$ and a scant chance of being found far away from $x = 0$. The same comments apply to the amplitude $\phi(k)$ as well: $\phi(k)$ peaks at $k = 0$ and vanishes far away. Figure 1.13 displays a typical wave packet that has the required localization properties we have just discussed.

In summary, the particle is represented not by a single de Broglie wave of well-defined frequency and wavelength, but by a wave packet that is obtained by adding a large number of waves of different frequencies.

The physical interpretation of the wave packet is obvious: $\psi_0(x)$ is the wave function or probability amplitude for finding the particle at position x; hence $|\psi_0(x)|^2$ gives the probability density for finding the particle at x, and $P(x)\, dx = |\psi_0(x)|^2 dx$ gives the probability of finding

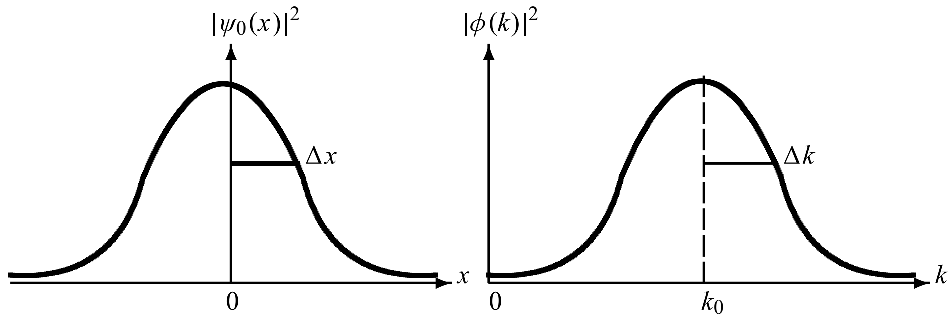

Figure 1.13 Two localized wave packets: $\psi_0(x) = (2/\pi a^2)^{1/4} e^{-x^2/a^2} e^{ik_0x}$ and $\phi(k) = (a^2/2\pi)^{1/4} e^{-a^2(k-k_0)^2/4}$; they peak at $x = 0$ and $k = k_0$, respectively, and vanish far away.

the particle between x and $x + dx$. What about the physical interpretation of $\phi(k)$? From (1.95) and (1.96) it follows that

$$\int_{-\infty}^{+\infty} |\psi_0(x)|^2 dx = \int_{-\infty}^{+\infty} |\phi(k)|^2 dk; \tag{1.97}$$

then if $\psi(x)$ is normalized so is $\phi(k)$, and vice versa. Thus, the function $\phi(k)$ can be interpreted most naturally, like $\psi_0(x)$, as a probability amplitude for measuring a wave vector k for a particle in the state $\phi(k)$. Moreover, while $|\phi(k)|^2$ represents the probability density for measuring k as the particle's wave vector, the quantity $P(k)\,dk = |\phi(k)|^2 dk$ gives the probability of finding the particle's wave vector between k and $k + dk$.

We can extract information about the particle's motion by simply expressing its corresponding matter wave in terms of the particle's energy, E, and momentum, p. Using $k = p/\hbar$, $dk = dp/\hbar$, $E = \hbar\omega$ and redefining $\tilde{\phi}(p) = \phi(k)/\sqrt{\hbar}$, we can rewrite (1.94) to (1.96) as follows:

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \tilde{\phi}(p) e^{i(px-Et)/\hbar} dp, \tag{1.98}$$

$$\psi_0(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \tilde{\phi}(p) e^{ipx/\hbar} dp, \tag{1.99}$$

$$\tilde{\phi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi_0(x) e^{-ipx/\hbar} dx, \tag{1.100}$$

where $E(p)$ is the total energy of the particle described by the wave packet $\psi(x,t)$ and $\tilde{\phi}(p)$ is the momentum amplitude of the packet.

In what follows we are going to illustrate the basic ideas of wave packets on a simple, instructive example: the Gaussian and square wave packets.

Example 1.8 (Gaussian and square wave packets)

(a) Find $\psi(x,\ 0)$ for a Gaussian wave packet $\phi(k) = A\exp\left[-a^2(k-k_0)^2/4\right]$, where A is a normalization factor to be found. Calculate the probability of finding the particle in the region $-a/2 \le x \le a/2$.

(b) Find $\phi(k)$ for a square wave packet $\psi_0(x) = \begin{cases} Ae^{ik_0x}, & |x| \leq a, \\ 0, & |x| > a. \end{cases}$
Find the factor A so that $\psi(x)$ is normalized.

Solution

(a) The normalization factor A is easy to obtain:

$$1 = \int_{-\infty}^{+\infty} |\phi(k)|^2 dk = |A|^2 \int_{-\infty}^{+\infty} \exp\left[-\frac{a^2}{2}(k-k_0)^2\right] dk, \tag{1.101}$$

which, by using a change of variable $z = k - k_0$ and using the integral $\int_{-\infty}^{+\infty} e^{-a^2z^2/2} dz = \sqrt{2\pi}/a$, leads at once to $A = \sqrt{a/\sqrt{2\pi}} = [a^2/(2\pi)]^{1/4}$. Now, the wave packet corresponding to

$$\phi(k) = \left(\frac{a^2}{2\pi}\right)^{1/4} \exp\left[-\frac{a^2}{4}(k-k_0)^2\right] \tag{1.102}$$

is

$$\psi_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ikx} dk = \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{2\pi}\right)^{1/4} \int_{-\infty}^{+\infty} e^{-a^2(k-k_0)^2/4 + ikx} dk. \tag{1.103}$$

To carry out the integration, we need simply to rearrange the exponent's argument as follows:

$$-\frac{a^2}{4}(k-k_0)^2 + ikx = -\left[\frac{a}{2}(k-k_0) - \frac{ix}{a}\right]^2 - \frac{x^2}{a^2} + ik_0x. \tag{1.104}$$

The introduction of a new variable $y = a(k-k_0)/2 - ix/a$ yields $dk = 2dy/a$, and when combined with (1.103) and (1.104), this leads to

$$\begin{aligned} \psi_0(x) &= \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{2\pi}\right)^{1/4} \int_{-\infty}^{+\infty} e^{-x^2/a^2} e^{ik_0x} e^{-y^2} \left(\frac{2}{a} dy\right) \\ &= \frac{1}{\sqrt{\pi}} \left(\frac{2}{\pi a^2}\right)^{1/4} e^{-x^2/a^2} e^{ik_0x} \int_{-\infty}^{+\infty} e^{-y^2} dy. \end{aligned} \tag{1.105}$$

Since $\int_{-\infty}^{+\infty} e^{-y^2} dy = \sqrt{\pi}$, this expression becomes

$$\psi_0(x) = \left(\frac{2}{\pi a^2}\right)^{1/4} e^{-x^2/a^2} e^{ik_0x}, \tag{1.106}$$

where e^{ik_0x} is the phase of $\psi_0(x)$; $\psi_0(x)$ is an oscillating wave with wave number k_0 modulated by a Gaussian envelope centered at the origin. We will see later that the phase factor e^{ik_0x} has real physical significance. The wave function $\psi_0(x)$ is complex, as necessitated by quantum mechanics. Note that $\psi_0(x)$, like $\phi(k)$, is normalized. Moreover, equations (1.102) and (1.106) show that the Fourier transform of a Gaussian wave packet is also a Gaussian wave packet.

The probability of finding the particle in the region $-a/2 \leq x \leq a/2$ can be obtained at once from (1.106):

$$P = \int_{-a/2}^{+a/2} |\psi_0(x)|^2 dx = \sqrt{\frac{2}{\pi a^2}} \int_{-a/2}^{+a/2} e^{-2x^2/a^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-1}^{+1} e^{-z^2/2} dz \simeq \frac{2}{3}, \tag{1.107}$$

where we have used the change of variable $z = 2x/a$.

(b) The normalization of $\psi_0(x)$ is straightforward:

$$1 = \int_{-\infty}^{+\infty} |\psi_0(x)|^2 dx = |A|^2 \int_{-a}^{a} e^{-ik_0x} e^{ik_0x} dx = |A|^2 \int_{-a}^{a} dx = 2a|A|^2; \tag{1.108}$$

hence $A = 1/\sqrt{2a}$. The Fourier transform of $\psi_0(x)$ is

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi_0(x) e^{-ikx} dx = \frac{1}{2\sqrt{\pi a}} \int_{-a}^{a} e^{ik_0x} e^{-ikx} dx = \frac{1}{\sqrt{\pi a}} \frac{\sin\,[(k-k_0)a]}{k-k_0}. \tag{1.109}$$

1.8.2 Wave Packets and the Uncertainty Relations

We want to show here that the width of a wave packet $\psi_0(x)$ and the width of its amplitude $\phi(k)$ are not independent; they are correlated by a reciprocal relationship. As it turns out, the reciprocal relationship between the widths in the x and k spaces has a direct connection to Heisenberg's uncertainty relation.

For simplicity, let us illustrate the main ideas on the Gaussian wave packet treated in the previous example (see (1.102) and (1.106)):

$$\psi_0(x) = \left(\frac{2}{\pi a^2}\right)^{1/4} e^{-x^2/a^2} e^{ik_0x}, \qquad \phi(k) = \left(\frac{a^2}{2\pi}\right)^{1/4} e^{-a^2(k-k_0)^2/4}. \tag{1.110}$$

As displayed in Figure 1.13, $|\psi_0(x)|^2$ and $|\phi(k)|^2$ are centered at $x = 0$ and $k = k_0$, respectively. It is convenient to define the half-widths Δx and Δk as corresponding to the half-maxima of $|\psi_0(x)|^2$ and $|\phi(k)|^2$. In this way, when x varies from 0 to $\pm\Delta x$ and k from k_0 to $k_0 \pm \Delta k$, the functions $|\psi_0(x)|^2$ and $|\phi(k)|^2$ drop to $e^{-1/2}$:

$$\frac{|\psi(\pm\Delta x,\ 0)|^2}{|\psi(0,\ 0)|^2} = e^{-1/2}, \qquad \frac{|\phi(k_0 \pm \Delta k)|^2}{|\phi(k_0)|^2} = e^{-1/2}. \tag{1.111}$$

These equations, combined with (1.110), lead to $e^{-2\Delta x^2/a^2} = e^{-1/2}$ and $e^{-a^2\Delta k^2/2} = e^{-1/2}$, respectively, or to

$$\Delta x = \frac{a}{2}, \qquad \Delta k = \frac{1}{a}; \tag{1.112}$$

hence

$$\Delta x \, \Delta k = \frac{1}{2}. \tag{1.113}$$

Since $\Delta k = \Delta p/\hbar$ we have

$$\boxed{\Delta x \, \Delta p = \frac{\hbar}{2}.} \tag{1.114}$$

This relation shows that if the packet's width is narrow in x-space, its width in momentum space must be very broad, and vice versa.

A comparison of (1.114) with Heisenberg's uncertainty relations (1.57) reveals that the Gaussian wave packet yields an *equality*, not an *inequality* relation. In fact, equation (1.114) is the *lowest limit* of Heisenberg's inequality. As a result, the Gaussian wave packet is called the *minimum uncertainty* wave packet. All other wave packets yield higher values for the product of the x and p uncertainties: $\Delta x \Delta p > \hbar/2$; for an illustration see Problem 1.11. In conclusion, the value of the uncertainties product $\Delta x \Delta p$ varies with the choice of ψ, but the lowest bound, $\hbar/2$, is provided by a Gaussian wave function. We have now seen how the wave packet concept offers a heuristic way of deriving Heisenberg's uncertainty relations; a more rigorous derivation is given in Chapter 2.

1.8.3 Motion of Wave Packets

How do wave packets evolve in time? The answer is important, for it gives an idea not only about the motion of a quantum particle in space but also about the connection between classical and quantum mechanics. Besides studying how wave packets propagate in space, we will also examine the conditions under which packets may or may not spread.

At issue here is, knowing the initial wave packet $\psi_0(x)$ or the amplitude $\phi(k)$, how do we find $\psi(x, t)$ at any later time t? This issue reduces to calculating the integral $\int \phi(k)e^{i(kx-\omega t)}dk$ in (1.94). To calculate this integral, we need to specify the angular frequency ω and the amplitude $\phi(k)$. We will see that the spreading or nonspreading of the packet is dictated by the form of the function $\omega(k)$.

1.8.3.1 Propagation of a Wave Packet without Distortion

The simplest form of the angular frequency ω is when it is *proportional* to the wave number k; this case corresponds to a *nondispersive* propagation. Since the constant of proportionality has the dimension of a velocity[14], which we denote by v_0 (i.e., $\omega = v_0 k$), the wave packet (1.94) becomes

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k)e^{ik(x-v_0 t)}dk. \tag{1.115}$$

This relation has the same structure as (1.95), which suggests that $\psi(x,t)$ is identical with $\psi_0(x - v_0 t)$:

$$\psi(x,t) = \psi_0(x - v_0 t); \tag{1.116}$$

the form of the wave packet at time t is identical with the initial form. Therefore, when ω is proportional to k, so that $\omega = v_0 k$, the wave packet travels to the right with constant velocity v_0 *without distortion*.

However, since we are interested in wave packets that describe particles, we need to consider the more general case of *dispersive* media which transmit harmonic waves of different frequencies at different velocities. This means that ω is a *function* of k: $\omega = \omega(k)$. The form of $\omega(k)$ is determined by the requirement that the wave packet $\psi(x,t)$ describes the particle. Assuming that the amplitude $\phi(k)$ peaks at $k = k_0$, then $\phi(k) = g(k - k_0)$ is appreciably different from zero only in a narrow range $\Delta k = k - k_0$, and we can Taylor expand $\omega(k)$ about k_0:

$$\begin{aligned}
\omega(k) &= \omega(k_0) + (k-k_0)\left.\frac{d\omega(k)}{dk}\right|_{k=k_0} + \frac{1}{2}(k-k_0)^2 \left.\frac{d^2\omega(k)}{dk^2}\right|_{k=k_0} + \cdots \\
&= \omega(k_0) + (k-k_0)v_g + (k-k_0)^2\alpha + \cdots
\end{aligned} \tag{1.117}$$

[14] For propagation of light in a vacuum this constant is equal to c, the speed of light.

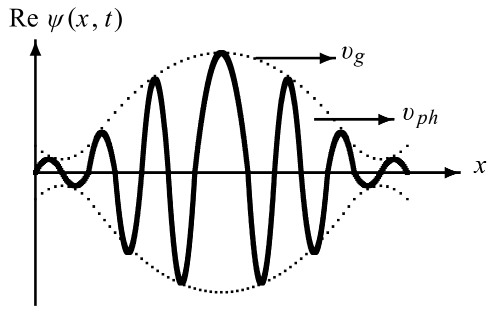

Figure 1.14 The function Re $\psi(x, t)$ of the wave packet (1.118), represented here by the solid curve contained in the dashed-curve envelope, propagates with the *group* velocity v_g along the x axis; the individual waves (not drawn here), which add up to make the solid curve, move with different *phase* velocities v_{ph}.

where $v_g = \left.\frac{d\omega(k)}{dk}\right|_{k=k_0}$ and $\alpha = \frac{1}{2}\left.\frac{d^2\omega(k)}{dk^2}\right|_{k=k_0}$.

Now, to determine $\psi(x, t)$ we need simply to substitute (1.117) into (1.94) with $\phi(k) = g(k - k_0)$. This leads to

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} e^{ik_0(x - v_{ph}t)} \int_{-\infty}^{+\infty} g(k - k_0) e^{i(k-k_0)(x - v_g t)} e^{-i(k-k_0)^2 \alpha t + \cdots} dk \qquad (1.118)$$

where[15]

$$\boxed{v_g = \frac{d\omega(k)}{dk}, \qquad v_{ph} = \frac{\omega(k)}{k};} \qquad (1.119)$$

v_{ph} and v_g are respectively the *phase* velocity and the *group* velocity. The phase velocity denotes the velocity of propagation for the phase of a single harmonic wave, $e^{ik_0(x - v_{ph}t)}$, and the group velocity represents the velocity of motion for the group of waves that make up the packet. One should not confuse the phase velocity and the group velocity; in general they are different. Only when ω is proportional to k will they be equal, as can be inferred from (1.119).

Group and phase velocities

Let us take a short detour to explain the meanings of v_{ph} and v_g. As mentioned above, when we superimpose many waves of different amplitudes and frequencies, we can obtain a wave packet or pulse which travels at the group velocity v_g; the *individual* waves that constitute the packet, however, move with different speeds; each wave moves with its own phase velocity v_{ph}. Figure 1.14 gives a qualitative illustration: the group velocity represents the velocity with which the wave packet propagates as a *whole*, where the individual waves (located inside the packet's envelope) that add up to make the packet move with different phase velocities. As shown in Figure 1.14, the wave packet has an appreciable magnitude only over a small region and falls rapidly outside this region.

The difference between the group velocity and the phase velocity can be understood quantitatively by deriving a relationship between them. A differentiation of $\omega = kv_{ph}$ (see (1.119)) with respect to k yields $d\omega/dk = v_{ph} + k(dv_{ph}/dk)$, and since $k = 2\pi/\lambda$, we have $dv_{ph}/dk =$

[15]In these equations we have omitted k_0 since they are valid for any choice of k_0.

$(dv_{ph}/d\lambda)(d\lambda/dk) = -(2\pi/k^2)(dv_{ph}/d\lambda)$ or $k(dv_{ph}/dk) = -\lambda(dv_{ph}/d\lambda)$; combining these relations, we obtain

$$\boxed{v_g = \frac{d\omega}{dk} = v_{ph} + k\frac{dv_{ph}}{dk} = v_{ph} - \lambda\frac{dv_{ph}}{d\lambda},} \tag{1.120}$$

which we can also write as

$$\boxed{v_g = v_{ph} + p\frac{dv_{ph}}{dp},} \tag{1.121}$$

since $k(dv_{ph}/dk) = (p/\hbar)(dv_{ph}/dp)(dp/dk) = p(dv_{ph}/dp)$ because $k = p/\hbar$. Equations (1.120) and (1.121) show that the group velocity may be larger or smaller than the phase velocity; it may also be equal to the phase velocity depending on the medium. If the phase velocity does not depend on the wavelength—this occurs in nondispersive media—the group and phase velocities are equal, since $dv_{ph}/d\lambda = 0$. But if v_{ph} depends on the wavelength—this occurs in dispersive media—then $dv_{ph}/d\lambda \neq 0$; hence the group velocity may be smaller or larger than the phase velocity. An example of a nondispersive medium is an inextensible string; we would expect $v_g = v_{ph}$. Water waves offer a typical dispersive medium; in Problem 1.13 we show that for deepwater waves we have $v_g = \frac{1}{2}v_{ph}$ and for surface waves we have $v_g = \frac{3}{2}v_{ph}$; see (1.212) and (1.214).

Consider the case of a particle traveling in a *constant potential* V; its total energy is $E(p) = p^2/(2m) + V$. Since the corpuscular features (energy and momentum) of a particle are connected to its wave characteristics (wave frequency and number) by the relations $E = \hbar\omega$ and $p = \hbar k$, we can rewrite (1.119) as follows:

$$\boxed{v_g = \frac{dE(p)}{dp}, \qquad v_{ph} = \frac{E(p)}{p},} \tag{1.122}$$

which, when combined with $E(p) = \frac{p^2}{2m} + V$, yield

$$v_g = \frac{d}{dp}\left(\frac{p^2}{2m} + V\right) = \frac{p}{m} = v_{particle}, \qquad v_{ph} = \frac{1}{p}\left(\frac{p^2}{2m} + V\right) = \frac{p}{2m} + \frac{V}{p}. \tag{1.123}$$

The group velocity of the wave packet is thus equal to the classical velocity of the particle, $v_g = v_{particle}$. This suggests we should view the "center" of the wave packet as traveling like a classical particle that obeys the laws of classical mechanics: the center would then follow the "classical trajectory" of the particle. We now see how the wave packet concept offers a clear connection between the classical description of a particle and its quantum mechanical description. In the case of a *free* particle, an insertion of $V = 0$ into (1.123) yields

$$v_g = \frac{p}{m}, \qquad v_{ph} = \frac{p}{2m} = \frac{1}{2}v_g. \tag{1.124}$$

This shows that, while the group velocity of the wave packet corresponding to a free particle is equal to the particle's velocity, p/m, the phase velocity is half the group velocity. The expression $v_{ph} = \frac{1}{2}v_g$ is meaningless, for it states that the wave function travels at half the speed of the particle it is intended to represent. This is unphysical indeed. The phase velocity has in general no meaningful physical significance.

Time-evolution of the packet

Having taken a short detour to discuss the phase and group velocities, let us now return to our main task of calculating the packet $\psi(x, t)$ as listed in (1.118). For this, we need to decide on where to terminate the expansion (1.117) or the exponent in the integrand of (1.118). We are going to consider two separate cases corresponding to whether we terminate the exponent in (1.118) at the linear term, $(k - k_0)v_g t$, or at the quadratic term, $(k - k_0)^2 \alpha t$. These two cases are respectively known as the *linear approximation* and the *quadratic approximation*.

In the linear approximation, which is justified when $g(k - k_0)$ is narrow enough to neglect the quadratic k^2 term, $(k - k_0)^2 \alpha t \ll 1$, the wave packet (1.118) becomes

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} e^{ik_0(x - v_{ph}t)} \int_{-\infty}^{+\infty} g(k - k_0) e^{i(k-k_0)(x - v_g t)} dk. \tag{1.125}$$

This relation can be rewritten as

$$\psi(x, t) = e^{ik_0(x - v_{ph}t)} \psi_0(x - v_g t) e^{-ik_0(x - v_g t)}, \tag{1.126}$$

where ψ_0 is the initial wave packet (see (1.95))

$$\psi_0(x - v_g t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(q) e^{i(x - v_g t)q + ik_0(x - v_g t)} dq; \tag{1.127}$$

the new variable q stands for $q = k - k_0$. Equation (1.126) leads to

$$|\psi(x, t)|^2 = \left|\psi_0(x - v_g t)\right|^2. \tag{1.128}$$

Equation (1.126) represents a wave packet whose amplitude is modulated. As depicted in Figure 1.14, the modulating wave, $\psi_0(x - v_g t)$, propagates to the right with the group velocity v_g; the modulated wave, $e^{ik_0(x - v_{ph}t)}$, represents a pure harmonic wave of constant wave number k_0 that also travels to the right with the phase velocity v_{ph}. That is, (1.126) and (1.128) represent a wave packet whose peak travels as a whole with the velocity v_g, while the individual wave propagates inside the envelope with the velocity v_{ph}. The group velocity, which gives the velocity of the packet's peak, clearly represents the velocity of the particle, since the chance of finding the particle around the packet's peak is much higher than finding it in any other region of space; the wave packet is highly localized in the neighborhood of the particle's position and vanishes elsewhere. It is therefore the group velocity, not the phase velocity, that is equal to the velocity of the particle represented by the packet. This suggests that the motion of a material particle can be described well by wave packets. By establishing a correspondence between the particle's velocity and the velocity of the wave packet's peak, we see that the wave packet concept jointly embodies the particle aspect and the wave aspect of material particles.

Now, what about the size of the wave packet in the linear approximation? Is it affected by the particle's propagation? Clearly not. This can be inferred immediately from (1.126): $\psi_0(x - v_g t)$ represents, mathematically speaking, a curve that travels to the right with a velocity v_g without deformation. This means that if the packet is initially Gaussian, it will remain Gaussian as it propagates in space without any change in its size.

To summarize, we have shown that, in the linear approximation, the wave packet propagates undistorted and undergoes a uniform translational motion. Next we are going to study the conditions under which the packet experiences deformation.

1.8.3.2 Propagation of a Wave Packet with Distortion

Let us now include the quadratic k^2 term, $(k-k_0)^2\alpha t$, in the integrand's exponent of (1.118) and drop the higher terms. This leads to

$$\psi(x,t) = e^{ik_0(x-v_{ph}t)} f(x,t), \tag{1.129}$$

where $f(x,t)$, which represents the envelope of the packet, is given by

$$f(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(q) e^{iq(x-v_g t)} e^{-iq^2\alpha t} dq, \tag{1.130}$$

with $q = k - k_0$. Were it not for the quadratic q^2 correction, $iq^2\alpha t$, the wave packet would move uniformly without any change of shape, since similarly to (1.116), $f(x,t)$ would be given by $f(x,t) = \psi_0(x - v_g t)$.

To show how α affects the width of the packet, let us consider the Gaussian packet (1.102) whose amplitude is given by $\phi(k) = (a^2/2\pi)^{1/4} \exp\left[-a^2(k-k_0)^2/4\right]$ and whose initial width is $\Delta x_0 = a/2$ and $\Delta k = \hbar/a$. Substituting $\phi(k)$ into (1.129), we obtain

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \left(\frac{a^2}{2\pi}\right)^{1/4} e^{ik_0(x-v_{ph}t)} \int_{-\infty}^{+\infty} \exp\left[iq(x-v_g t) - \left(\frac{a^2}{4} + i\alpha t\right) q^2\right] dq. \tag{1.131}$$

Evaluating the integral (the calculations are detailed in the following example, see Eq. (1.145)), we can show that the packet's density distribution is given by

$$|\psi(x,t)|^2 = \frac{1}{\sqrt{2\pi}\,\Delta x(t)} \exp\left\{-\frac{\left(x - v_g t\right)^2}{2\left[\Delta x(t)\right]^2}\right\}, \tag{1.132}$$

where $\Delta x(t)$ is the width of the packet at time t:

$$\Delta x(t) = \frac{a}{2}\sqrt{1 + \frac{16\alpha^2}{a^4} t^2} = \Delta x_0 \sqrt{1 + \frac{\alpha^2 t^2}{(\Delta x_0)^4}}. \tag{1.133}$$

We see that the packet's width, which was initially given by $\Delta x_0 = a/2$, has grown by a factor of $\sqrt{1 + \alpha^2 t^2/(\Delta x_0)^4}$ after time t. Hence the wave packet is spreading; *the spreading is due to the inclusion of the quadratic q^2 term*, $iq^2\alpha t$. Should we drop this term, the packet's width $\Delta x(t)$ would then remain constant, equal to Δx_0.

The density distribution (1.132) displays two results: (1) the center of the packet moves with the group velocity; (2) the packet's width increases linearly with time. From (1.133) we see that the packet begins to spread appreciably only when $\alpha^2 t^2/(\Delta x_0)^4 \approx 1$ or $t \approx (\Delta x_0)^2/\alpha$. In fact, if $t \ll (\Delta x_0)^2/\alpha$ the packet's spread will be negligible, whereas if $t \gg \frac{(\Delta x_0)^2}{\alpha}$ the packet's spread will be significant.

To be able to make concrete statements about the growth of the packet, as displayed in (1.133), we need to specify α; this reduces to determining the function $\omega(k)$, since $\alpha = \frac{1}{2}\left.\frac{d^2\omega}{dk^2}\right|_{k=k_0}$. For this, let us invoke an example that yields itself to explicit calculation. In fact, the example we are going to consider—a *free* particle with a Gaussian amplitude—allows the calculations to be performed *exactly*; hence there is no need to expand $\omega(k)$.

Example 1.9 (Free particle with a Gaussian wave packet)
Determine how the wave packet corresponding to a free particle, with an initial Gaussian packet, spreads in time.

Solution
The issue here is to find out how the wave packet corresponding to a free particle with $\phi(k) = (a^2/2\pi)^{1/4}e^{-a^2(k-k_0)^2/4}$ (see (1.110)) spreads in time.

First, we need to find the form of the wave packet, $\psi(x,t)$. Substituting the amplitude $\phi(k) = (a^2/2\pi)^{1/4}e^{-a^2(k-k_0)^2/4}$ into the Fourier integral (1.94), we obtain

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}}\left(\frac{a^2}{2\pi}\right)^{1/4}\int_{-\infty}^{+\infty}\exp\left[-\frac{a^2}{4}(k-k_0)^2 + i(kx-\omega t)\right]dk. \tag{1.134}$$

Since $\omega(k) = \hbar k^2/(2m)$ (the dispersion relation for a free particle), and using a change of variables $q = k - k_0$, we can write the exponent in the integrand of (1.134) as a perfect square for q:

$$\begin{aligned}
-\frac{a^2}{4}(k-k_0)^2 + i\left(kx - \frac{\hbar k^2}{2m}t\right) =& -\left(\frac{a^2}{4} + i\frac{\hbar t}{2m}\right)q^2 + i\left(x - \frac{\hbar k_0 t}{m}\right)q \\
&+ ik_0\left(x - \frac{\hbar k_0 t}{2m}\right) \\
=& -\alpha q^2 + i\left(x - \frac{\hbar k_0 t}{m}\right)q + ik_0\left(x - \frac{\hbar k_0 t}{2m}\right) \\
=& -\alpha\left[q - \frac{i}{2\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)\right]^2 - \frac{1}{4\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)^2 \\
&+ ik_0\left(x - \frac{\hbar k_0 t}{2m}\right),
\end{aligned} \tag{1.135}$$

where we have used the relation $-\alpha q^2 + iyq = -\alpha\left[q - iy/(2\alpha)\right]^2 - y^2/(4\alpha)$, with $y = x - \hbar k_0 t/m$ and

$$\alpha = \frac{a^2}{4} + i\frac{\hbar t}{2m}. \tag{1.136}$$

Substituting (1.135) into (1.134) we obtain

$$\begin{aligned}
\psi(x,t) \quad = \quad & \frac{1}{\sqrt{2\pi}}\left(\frac{a^2}{2\pi}\right)^{1/4}\exp\left[ik_0\left(x - \frac{\hbar k_0 t}{2m}\right)\right]\exp\left[-\frac{1}{4\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)^2\right] \\
& \times\int_{-\infty}^{+\infty}\exp\left\{-\alpha\left[q - \frac{i}{2\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)\right]^2\right\}dq.
\end{aligned} \tag{1.137}$$

Combined with the integral[16] $\int_{-\infty}^{+\infty}\exp\left[-\alpha\left(q - iy/(2\alpha)\right)^2\right]dq = \sqrt{\pi/\alpha}$, (1.137) leads to

$$\psi(x,t) = \frac{1}{\sqrt{\alpha}}\left(\frac{a^2}{8\pi}\right)^{1/4}\exp\left[ik_0\left(x - \frac{\hbar k_0 t}{2m}\right)\right]\exp\left[-\frac{1}{4\alpha}\left(x - \frac{\hbar k_0 t}{m}\right)^2\right]. \tag{1.138}$$

[16] If β and δ are two complex numbers and if Re $\beta > 0$, we have $\int_{-\infty}^{+\infty}e^{-\beta(q+\delta)^2}dq = \sqrt{\pi/\beta}$.

Since α is a complex number (see (1.136)), we can write it in terms of its modulus and phase

$$\alpha=\frac{a^2}{4}\left(1+i\frac{2\hbar t}{ma^2}\right)=\frac{a^2}{4}\left(1+\frac{4\hbar^2t^2}{m^2a^4}\right)^{1/2}e^{i\theta}, \tag{1.139}$$

where $\theta=\tan^{-1}\left[2\hbar t/(ma^2)\right]$; hence

$$\frac{1}{\sqrt{\alpha}}=\frac{2}{a}\left(1+\frac{4\hbar^2t^2}{m^2a^4}\right)^{-1/4}e^{-i\theta/2}. \tag{1.140}$$

Substituting (1.136) and (1.140) into (1.138), we have

$$\psi(x,t)=\left(\frac{2}{\pi a^2}\right)^{1/4}\left(1+\frac{4\hbar^2t^2}{m^2a^4}\right)^{-1/4}e^{-i\theta/2}e^{ik_0(x-\hbar k_0t/2m)}\exp\left[-\frac{(x-\hbar k_0t/m)^2}{a^2+2i\hbar t/m}\right]. \tag{1.141}$$

Since $\left|e^{-y^2/(a^2+2i\hbar t/m)}\right|^2=e^{-y^2/(a^2-2i\hbar t/m)}e^{-y^2/(a^2+2i\hbar t/m)}$, where $y=x-\hbar k_0t/m$, and since $y^2/(a^2-2i\hbar t/m)+y^2/(a^2+2i\hbar t/m)=2a^2y^2/(a^4+4\hbar^2t^2/m^2)$, we have

$$\left|\exp\left(-\frac{y^2}{a^2+2i\hbar t/m}\right)\right|^2=\exp\left(-\frac{2a^2y^2}{a^4+4\hbar^2t^2/m^2}\right); \tag{1.142}$$

hence

$$\begin{aligned}|\psi(x,t)|^2 &= \sqrt{\frac{2}{\pi a^2}}\left(1+\frac{4\hbar^2t^2}{m^2a^4}\right)^{-1/2}\left|\exp\left[-\frac{(x-\hbar k_0t/m)^2}{a^2+2i\hbar t/m}\right]\right|^2\\ &= \sqrt{\frac{2}{\pi a^2}}\frac{1}{\gamma(t)}\exp\left\{-\frac{2}{\left[a\gamma(t)\right]^2}\left(x-\frac{\hbar k_0t}{m}\right)^2\right\},\end{aligned} \tag{1.143}$$

where $\gamma(t)=\sqrt{1+4\hbar^2t^2/(m^2a^4)}$.

We see that both the wave packet (1.141) and the probability density (1.143) remain Gaussian as time evolves. This can be traced to the fact that the x-dependence of the phase, e^{ik_0x}, of $\psi_0(x)$ as displayed in (1.110) is *linear*. If the x-dependence of the phase were other than linear, say quadratic, the form of the wave packet would not remain Gaussian. So the phase factor e^{ik_0x}, which was present in $\psi_0(x)$, allows us to account for the motion of the particle.

Since the group velocity of a free particle is $v_g=d\omega/dk=\frac{d}{dk}\left(\frac{\hbar k^2}{2m}\right)\Big|_{k_0}=\hbar k_0/m$, we can rewrite (1.141) as follows[17]:

$$\psi(x,t)=\frac{1}{\sqrt{\sqrt{2\pi}\,\Delta x(t)}}e^{-i\theta/2}e^{ik_0(x-v_gt/2)}\exp\left[-\frac{\left(x-v_gt\right)^2}{a^2+2i\hbar t/m}\right], \tag{1.144}$$

$$\left|\psi(x,t)\right|^2=\frac{1}{\sqrt{2\pi}\,\Delta x(t)}\exp\left\{-\frac{\left(x-v_gt\right)^2}{2\left[\Delta x(t)\right]^2}\right\}, \tag{1.145}$$

[17] It is interesting to note that the harmonic wave $e^{ik_0(x-v_gt/2)}$ propagates with a phase velocity which is *half* the group velocity; as shown in (1.124), this is a property of *free* particles.

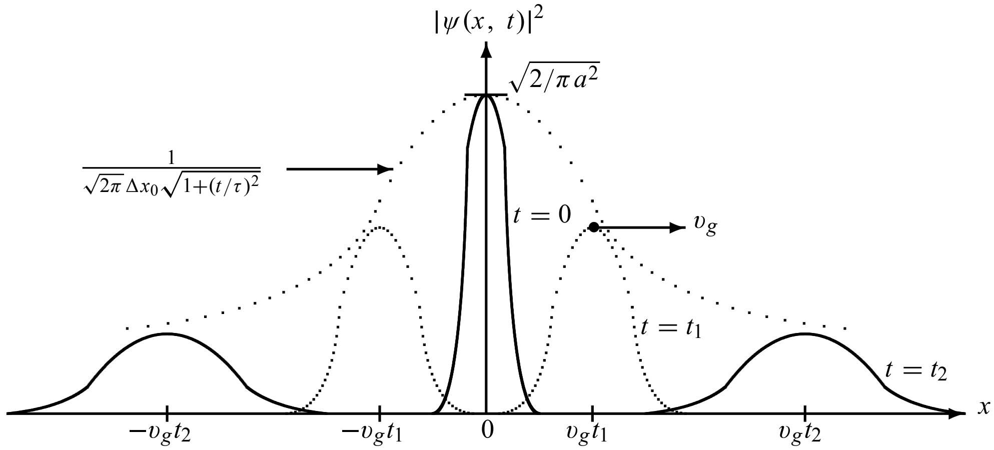

Figure 1.15 Time evolution of $|\psi(x,\ t)|^2$: the peak of the packet, which is centered at $x = v_g t$, moves with the speed v_g from left to right. The height of the packet, represented here by the dotted envelope, is modulated by the function $1/(\sqrt{2\pi}\,\Delta x(t))$, which goes to zero at $t \to \pm\infty$ and is equal to $\sqrt{2/\pi a^2}$ at $t = 0$. The width of the packet $\Delta x(t) = \Delta x_0\sqrt{1 + (t/\tau)^2}$ increases linearly with time.

where[18]

$$\Delta x(t) = \frac{a}{2}\gamma(t) = \frac{a}{2}\sqrt{1 + \frac{4\hbar^2 t^2}{m^2 a^4}} \tag{1.146}$$

represents the width of the wave packet at time t. Equations (1.144) and (1.145) describe a Gaussian wave packet that is centered at $x = v_g t$ whose peak travels with the group speed $v_g = \hbar k_0/m$ and whose width $\Delta x(t)$ increases linearly with time. So, during time t, the packet's center has moved from $x = 0$ to $x = v_g t$ and its width has expanded from $\Delta x_0 = a/2$ to $\Delta x(t) = \Delta x_0\sqrt{1 + 4\hbar^2 t^2/(m^2 a^4)}$. The wave packet therefore undergoes a distortion; although it remains Gaussian, its width broadens linearly with time whereas its height, $1/(\sqrt{2\pi}\,\Delta x(t))$, decreases with time. As depicted in Figure 1.15, the wave packet, which had a very broad width and a very small amplitude at $t \to -\infty$, becomes narrower and narrower and its amplitude larger and larger as time increases towards $t = 0$; at $t = 0$ the packet is very localized, its width and amplitude being given by $\Delta x_0 = a/2$ and $\sqrt{2/\pi a^2}$, respectively. Then, as time increases ($t > 0$), the width of the packet becomes broader and broader, and its amplitude becomes smaller and smaller.

In the rest of this section we are going to comment on several features that are relevant not only to the Gaussian packet considered above but also to more general wave packets. First, let us begin by estimating the time at which the wave packet starts to spread out appreciably. The packet, which is initially narrow, begins to grow out noticeably only when the second term, $2\hbar t/(ma^2)$, under the square root sign of (1.146) is of order unity. For convenience, let us write

[18]We can derive (1.146) also from (1.111): a combination of the half-width $|\psi(\pm\Delta x,\ t)|^2 / |\psi(0,\ 0)|^2 = e^{-1/2}$ with (1.143) yields $e^{-2[\Delta x/a\gamma(t)]^2} = e^{-1/2}$, which in turn leads to (1.146).

(1.146) in the form

$$\Delta x(t) = \Delta x_0 \sqrt{1 + \left(\frac{t}{\tau}\right)^2}, \tag{1.147}$$

where

$$\tau = \frac{2m(\Delta x_0)^2}{\hbar} \tag{1.148}$$

represents a time constant that characterizes the rate of the packet's spreading. Now we can estimate the order of magnitude of τ; it is instructive to evaluate it for *microscopic* particles as well as for *macroscopic* particles. For instance, τ for an electron whose position is defined to within 10^{-10} m is given by[19] $\tau \simeq 1.7 \times 10^{-16}$ s; on the other hand, the time constant for a macroscopic particle of mass say 1 g whose position is defined to within 1 mm is of the order[20] of $\tau \simeq 2 \times 10^{25}$ s (for an illustration see Problems 1.15 and 1.16). This crude calculation suggests that the wave packets of microscopic systems very quickly undergo significant growth; as for the packets of macroscopic systems, they begin to grow out noticeably only after the system has been in motion for an absurdly long time, a time of the order of, if not much higher than, the age of the Universe itself, which is about 4.7×10^{17} s. Having estimated the times at which the packet's spread becomes appreciable, let us now shed some light on the size of the spread. From (1.147) we see that when $t \gg \tau$ the packet's spreading is significant and, conversely, when $t \ll \tau$ the spread is negligible. As the cases $t \gg \tau$ and $t \ll \tau$ correspond to microscopic and macroscopic systems, respectively, we infer that the packet's dispersion is significant for microphysical systems and negligible for macroscopic systems. In the case of macroscopic systems, the spread is there but it is too small to detect. For an illustration see Problem 1.15 where we show that the width of a 100 g object increases by an absurdly small factor of about 10^{-29} after traveling a distance of 100 m, but the width of a 25 eV electron increases by a factor of 10^9 after traveling the same distance (in a time of 3.3×10^{-5} s). Such an immense dispersion in such a short time is indeed hard to visualize classically; this motion cannot be explained by classical physics.

So the wave packets of propagating, microscopic particles are prone to spreading out very significantly in a short time. This spatial spreading seems to generate a conceptual problem: the spreading is incompatible with our expectation that the packet should remain highly localized at all times. After all, the wave packet is supposed to represent the particle and, as such, it is expected to travel without dispersion. For instance, the charge of an electron does not spread out while moving in space; the charge should remain localized inside the corresponding wave packet. In fact, whenever microscopic particles (electrons, neutrons, protons, etc.) are observed, they are always confined to small, finite regions of space; they never spread out as suggested by equation (1.146). How do we explain this apparent contradiction? The problem here has to do with the proper interpretation of the situation: we must modify the classical concepts pertaining to the meaning of the position of a particle. The wave function (1.141) cannot be identified with a material particle. The quantity $|\psi(x, t)|^2 dx$ represents the probability (Born's interpretation) of finding the particle described by the packet $\psi(x, t)$ at time t in the spatial region located between x and $x + dx$. *The material particle does not disperse (or fuzz out); yet its position cannot be known exactly*. The spreading of the matter wave, which is accompanied by a shrinkage of its height, as indicated in Figure 1.15, corresponds to a decrease

[19] If $\Delta x_0 = 10^{-10}$ m and since the rest mass energy of an electron is $mc^2 = 0.5$ MeV and using $\hbar c \simeq 197 \times 10^{-15}$ MeV m, we have $\tau = 2mc^2(\Delta x_0)^2/((\hbar c)c) \simeq 1.7 \times 10^{-16}$ s.

[20] Since $\hbar = 1.05 \times 10^{-34}$ J s we have $\tau = 2 \times 0.001$ kg $\times$ $(0.001 \text{ m})^2/(1.05 \times 10^{-34}$ J s$) \simeq 2 \times 10^{25}$ s.

of the probability density $|\psi(x, t)|^2$ and implies in no way a growth in the size of the particle. So the wave packet gives only the probability that the particle it represents will be found at a given position. No matter how broad the packet becomes, we can show that its norm is always conserved, for it does not depend on time. In fact, as can be inferred from (1.143), the norm of the packet is equal to one:

$$\int_{-\infty}^{+\infty} |\psi(x,t)|^2 \, dx = \sqrt{\frac{2}{\pi a^2}} \frac{1}{\gamma} \int_{-\infty}^{+\infty} \exp\left\{ -\frac{2\,(x - \hbar k_0 t/m)^2}{(a\gamma)^2} \right\} dx = \sqrt{\frac{2}{\pi a^2}} \frac{1}{\gamma} \sqrt{\frac{\pi a^2 \gamma^2}{2}} = 1, \tag{1.149}$$

since $\int_{-\infty}^{+\infty} e^{-\alpha x^2} dx = \sqrt{\pi/\alpha}$. This is expected, since the probability of finding the particle somewhere along the x-axis must be equal to one. The important issue here is that the norm of the packet is time independent and that its spread does not imply that the material particle becomes bloated during its motion, but simply implies a redistribution of the probability density. So, in spite of the significant spread of the packets of microscopic particles, the norms of these packets are always conserved—normalized to unity.

Besides, we should note that the example considered here is an *idealized* case, for we are dealing with a *free* particle. If the particle is subject to a potential, as in the general case, its wave packet will not spread as dramatically as that of a free particle. In fact, a varying potential can cause the wave packet to become narrow. This is indeed what happens when a measurement is performed on a microscopic system; the interaction of the system with the measuring device makes the packet very narrow, as will be seen in Chapter 3.

Let us now study how the spreading of the wave packet affects the uncertainties product $\Delta x(t)\Delta p(t)$. First, we should point out that the average momentum of the packet $\hbar k_0$ and its uncertainty $\hbar \Delta k$ do not change in time. This can be easily inferred as follows. Rewriting (1.94) in the form

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k,0) e^{i(kx-\omega t)} dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k,t) e^{ikx} dk, \tag{1.150}$$

we have

$$\phi(k,t) = e^{-i\omega(k)t} \phi(k,0), \tag{1.151}$$

where $\phi(k,0) = (a^2/2\pi)^{1/4} e^{-a^2(k-k_0)^2/4}$; hence

$$|\phi(k,t)|^2 = |\phi(k,0)|^2 . \tag{1.152}$$

This suggests that the widths of $\phi(k,t)$ and $\phi(k,0)$ are equal; hence Δk remains constant and so must the momentum dispersion Δp (this is expected because the momentum of a free particle is a constant of the motion). Since the width of $\phi(k,0)$ is given by $\Delta k = 1/a$ (see (1.112)), we have

$$\Delta p = \hbar \Delta k = \frac{\hbar}{a}. \tag{1.153}$$

Multiplying this relation by (1.146), we have

$$\Delta x(t) \Delta p = \frac{\hbar}{2} \sqrt{1 + \frac{4\hbar^2}{m^2 a^4} t^2}, \tag{1.154}$$

which shows that $\Delta x(t)\Delta p \geq \hbar/2$ is satisfied at all times. Notably, when $t = 0$ we obtain the lower bound limit $\Delta x_0 \Delta p = \hbar/2$; this is the uncertainty relation for a stationary Gaussian packet (see (1.114)). As $|t|$ increases, however, we obtain an inequality, $\Delta x(t)\Delta p > \hbar/2$.

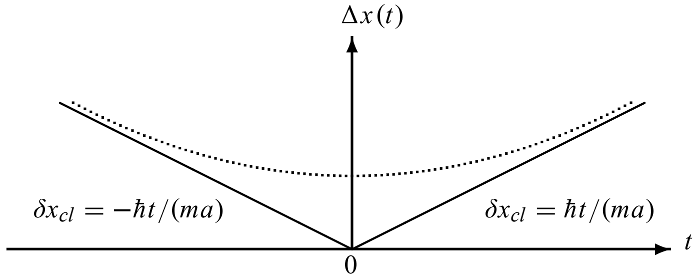

Figure 1.16 Time evolutions of the packet's width $\Delta x(t) = \Delta x_0\sqrt{1 + (\delta x_{cl}(t)/\Delta x_0)^2}$ (dotted curve) and of the classical dispersion $\delta x_{cl}(t) = \pm\hbar t/(ma)$ (solid lines). For large values of $|t|$, $\Delta x(t)$ approaches $\delta x_{cl}(t)$ and at $t = 0$, $\Delta x(0) = \Delta x_0 = a/2$.

Having shown that the width of the packet does not disperse in momentum space, let us now study the dispersion of the packet's width in x-space. Since $\Delta x_0 = a/2$ we can write (1.146) as

$$\Delta x(t) = \frac{a}{2}\sqrt{1 + \frac{4\hbar^2 t^2}{m^2 a^4}} = \Delta x_0\sqrt{1 + \left(\frac{\delta x_{cl}(t)}{\Delta x_0}\right)^2}, \qquad (1.155)$$

where the dispersion factor $\delta x_{cl}(t)/\Delta x_0$ is given by

$$\frac{\delta x_{cl}(t)}{\Delta x_0} = \pm\frac{2\hbar}{ma^2}t = \pm\frac{\hbar}{2m\,\Delta x_0^2}t; \qquad (1.156)$$

As shown in Figure 1.16, when $|t|$ is large (i.e., $t \to \pm\infty$), we have $\Delta x(t) \to \delta x_{cl}(t)$ with

$$\delta x_{cl}(t) = \pm\frac{\hbar t}{ma} = \pm\frac{\Delta p}{m}t = \pm\Delta v t, \qquad (1.157)$$

where $\Delta v = \hbar/(ma)$ represents the dispersion in velocity. This means that if a particle starts initially ($t = 0$) at $x = 0$ with a velocity dispersion equal to Δv, then Δv will remain constant but the dispersion of the particle's position will increase linearly with time: $\delta x_{cl}(t) = \hbar|t|/(ma)$ (Figure 1.16). We see from (1.155) that if $\delta x_{cl}(t)/\Delta x_0 \ll 1$, the spreading of the wave packet is negligible, but if $\delta x_{cl}(t)/\Delta x_0 \gg 1$, the wave packet will spread out without bound.

We should highlight at this level the importance of the classical limit of (1.154): in the limit $\hbar \to 0$, the product $\Delta x(t)\Delta p$ goes to zero. This means that the x and p uncertainties become negligible; that is, in the classical limit, the wave packet will propagate without spreading. In this case the center of the wave packet moves like a free particle that obeys the laws of classical mechanics. *The spread of wave packets is thus a purely quantum effect*. So when $\hbar \to 0$ all quantum effects, the spread of the packet, disappear.

We may conclude this study of wave packets by highlighting their importance:

- They provide a linkage with the Heisenberg uncertainty principle.
- They embody and unify the particle and wave features of matter waves.
- They provide a linkage between wave intensities and probabilities.
- They provide a connection between classical and quantum mechanics.

1.9 Concluding Remarks

Despite its striking success in predicting the hydrogen's energy levels and transition rates, the Bohr model suffers from a number of limitations:

- It works only for hydrogen and hydrogen-like ions such as $He^+$and Li^{2+}.
- It provides no explanation for the origin of its various assumptions. For instance, it gives no theoretical justification for the quantization condition (1.63) nor does it explain why stationary states radiate no energy.
- It fails to explain why, instead of moving *continuously* from one energy level to another, the electrons *jump* from one level to the other.

The model therefore requires considerable extension to account for the electronic properties and spectra of a wide range of atoms. Even in its present limited form, Bohr's model represents a bold and major departure from classical physics: classical physics offers no justification for the existence of discrete energy states in a system such as a hydrogen atom and no justification for the quantization of the angular momentum.

In its present form, the model not only suffers from incompleteness but also lacks the ingredients of a consistent theory. It was built upon a series of ad hoc, piecemeal assumptions. These assumptions were not derived from the first principles of a more general theory, but postulated rather arbitrarily.

The formulation of the theory of quantum mechanics was largely precipitated by the need to find a theoretical foundation for Bohr's ideas as well as to explain, from first principles, a wide variety of other microphysical phenomena such as the puzzling processes discussed in this chapter. It is indeed surprising that a single theory, quantum mechanics, is powerful and rich enough to explain accurately a wide variety of phenomena taking place at the molecular, atomic, and subatomic levels.

In this chapter we have dealt with the most important experimental facts which confirmed the failure of classical physics and subsequently led to the birth of quantum mechanics. In the rest of this text we will focus on the formalism of quantum mechanics and on its application to various microphysical processes. To prepare for this task, we need first to study the mathematical tools necessary for understanding the formalism of quantum mechanics; this is taken up in Chapter 2.

1.10 Solved Problems

Numerical calculations in quantum physics can be made simpler by using the following units. First, it is convenient to express energies in units of electronvolt (eV): one eV is defined as the energy acquired by an electron passing through a potential difference of one Volt. The electronvolt unit can be expressed in terms of joules and vice versa: 1 eV $= (1.6 \times 10^{-19}$ C) $\times$ (1 V) $= 1.6 \times 10^{-19}$ J and 1 J $= 0.625 \times 10^{19}$ eV.

It is also convenient to express the masses of subatomic particles, such as the electron, proton, and neutron, in terms of their rest mass energies: $m_e c^2 = 0.511$ MeV, $m_p c^2 = 938.27$ MeV, and $m_n c^2 = 939.56$ MeV.

In addition, the quantities $\hbar c = 197.33$ MeV fm $= 197.33 \times 10^{-15}$ MeV m or $hc = 1242.37 \times 10^{-10}$ eV m are sometimes more convenient to use than $\hbar = 1.05 \times 10^{-34}$ J s.

Additionally, instead of $1/(4\pi\varepsilon_0) = 8.9 \times 10^9$ N m^2 C^{-2}, one should sometimes use the fine structure constant $\alpha = e^2/[(4\pi\varepsilon_0)\hbar c] = 1/137$.

Problem 1.1

A 45 kW broadcasting antenna emits radio waves at a frequency of 4 MHz.

(a) How many photons are emitted per second?

(b) Is the quantum nature of the electromagnetic radiation important in analyzing the radiation emitted from this antenna?

Solution

(a) The electromagnetic energy emitted by the antenna in one second is $E = 45\,000$ J. Thus, the number of photons emitted in one second is

$$n = \frac{E}{h\nu} = \frac{45\,000 \text{ J}}{6.63 \times 10^{-34} \text{ J s} \times 4 \times 10^6 \text{ Hz}} = 1.7 \times 10^{31}. \tag{1.158}$$

(b) Since the antenna emits a huge number of photons every second, 1.7×10^{31}, the quantum nature of this radiation is unimportant. As a result, this radiation can be treated fairly accurately by the classical theory of electromagnetism.

Problem 1.2

Consider a mass–spring system where a 4 kg mass is attached to a massless spring of constant $k = 196$ N m^{-1}; the system is set to oscillate on a frictionless, horizontal table. The mass is pulled 25 cm away from the equilibrium position and then released.

(a) Use classical mechanics to find the total energy and frequency of oscillations of the system.

(b) Treating the oscillator with quantum theory, find the energy spacing between two consecutive energy levels and the total number of quanta involved. Are the quantum effects important in this system?

Solution

(a) According to classical mechanics, the frequency and the total energy of oscillations are given by

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}} = \frac{1}{2\pi}\sqrt{\frac{196}{4}} = 1.11 \text{ Hz}, \qquad E = \frac{1}{2}kA^2 = \frac{196}{2}(0.25)^2 = 6.125 \text{ J}. \tag{1.159}$$

(b) The energy spacing between two consecutive energy levels is given by

$$\Delta E = h\nu = (6.63 \times 10^{-34} \text{ J s}) \times (1.11 \text{ Hz}) = 7.4 \times 10^{-34} \text{ J} \tag{1.160}$$

and the total number of quanta is given by

$$n = \frac{E}{\Delta E} = \frac{6.125 \text{ J}}{7.4 \times 10^{-34} \text{ J}} = 8.3 \times 10^{33}. \tag{1.161}$$

We see that the energy of one quantum, 7.4×10^{-34} J, is completely negligible compared to the total energy 6.125 J, and that the number of quanta is very large. As a result, the energy levels of the oscillator can be viewed as continuous, for it is not feasible classically to measure the spacings between them. Although the quantum effects are present in the system, they are beyond human detection. So quantum effects are negligible for macroscopic systems.

Problem 1.3
When light of a given wavelength is incident on a metallic surface, the stopping potential for the photoelectrons is 3.2 V. If a second light source whose wavelength is double that of the first is used, the stopping potential drops to 0.8 V. From these data, calculate

(a) the wavelength of the first radiation and

(b) the work function and the cutoff frequency of the metal.

Solution

(a) Using (1.23) and since the wavelength of the second radiation is double that of the first one, $\lambda_2 = 2\lambda_1$, we can write

$$V_{s_1} = \frac{hc}{e\lambda_1} - \frac{W}{e}, \tag{1.162}$$

$$V_{s_2} = \frac{hc}{e\lambda_2} - \frac{W}{e} = \frac{hc}{2e\lambda_1} - \frac{W}{e}. \tag{1.163}$$

To obtain λ_1 we have only to subtract (1.163) from (1.162):

$$V_{s_1} - V_{s_2} = \frac{hc}{e\lambda_1}\left(1 - \frac{1}{2}\right) = \frac{hc}{2e\lambda_1}. \tag{1.164}$$

The wavelength is thus given by

$$\lambda_1 = \frac{hc}{2e(V_{s_1} - V_{s_2})} = \frac{6.6 \times 10^{-34}\ \text{J s} \times 3 \times 10^8 \text{m s}^{-1}}{2 \times 1.6 \times 10^{-19}\ \text{C} \times (3.2\ \text{V} - 0.8\ \text{V})} = 2.6 \times 10^{-7}\ \text{m}. \tag{1.165}$$

(b) To obtain the work function, we simply need to multiply (1.163) by 2 and subtract the result from (1.162), $V_{s_1} - 2V_{s_2} = W/e$, which leads to

$$W = e(V_{s_1} - 2V_{s_2}) = 1.6\ \text{eV} = 1.6 \times 1.6 \times 10^{-19} = 2.56 \times 10^{-19}\ \text{J}. \tag{1.166}$$

The cutoff frequency is

$$\nu = \frac{W}{h} = \frac{2.56 \times 10^{-19}\ \text{J}}{6.6 \times 10^{-34}\ \text{J s}} = 3.9 \times 10^{14}\ \text{Hz}. \tag{1.167}$$

Problem 1.4

(a) Estimate the energy of the electrons that we need to use in an electron microscope to resolve a separation of 0.27 nm.

(b) In a scattering of 2 eV protons from a crystal, the fifth maximum of the intensity is observed at an angle of 30°. Estimate the crystal's planar separation.

Solution

(a) Since the electron's momentum is $p = 2\pi\hbar/\lambda$, its kinetic energy is given by

$$E = \frac{p^2}{2m_e} = \frac{2\pi^2\hbar^2}{m_e\lambda^2}. \tag{1.168}$$

Since $m_ec^2 = 0.511$ MeV, $\hbar c = 197.33 \times 10^{-15}$ MeV m, and $\lambda = 0.27 \times 10^{-9}$ m, we have

$$E = \frac{2\pi^2(\hbar c)^2}{(m_ec^2)\lambda^2} = \frac{2\pi^2(197.33 \times 10^{-15}\ \text{MeV m})^2}{(0.511\ \text{MeV})(0.27 \times 10^{-9}\ \text{m})^2} = 20.6\ \text{eV}. \tag{1.169}$$

(b) Using Bragg's relation (1.46), $\lambda = (2d/n)\sin\phi$, where d is the crystal's planar separation, we can infer the proton's kinetic energy from (1.168):

$$E = \frac{p^2}{2m_p} = \frac{2\pi^2\hbar^2}{m_p\lambda^2} = \frac{n^2\pi^2\hbar^2}{2m_p d^2 \sin^2\phi}, \tag{1.170}$$

which leads to

$$d = \frac{n\pi\hbar}{(\sin\phi)\sqrt{2m_p E}} = \frac{n\pi\hbar c}{(\sin\phi)\sqrt{2m_p c^2 E}}. \tag{1.171}$$

Since $n = 5$ (the fifth maximum), $\phi = 30^\circ$, $E = 2$ eV, and $m_p c^2 = 938.27$ MeV, we have

$$d = \frac{5\pi \times 197.33 \times 10^{-15}\text{ MeV m}}{(\sin 30^\circ)\sqrt{2 \times 938.27\text{ MeV} \times 2 \times 10^{-6}\text{ MeV}}} = 0.101\text{ nm}. \tag{1.172}$$

Problem 1.5

A photon of energy 3 keV collides elastically with an electron initially at rest. If the photon emerges at an angle of 60°, calculate

(a) the kinetic energy of the recoiling electron and

(b) the angle at which the electron recoils.

Solution

(a) From energy conservation, we have

$$h\nu + m_e c^2 = h\nu' + (K_e + m_e c^2), \tag{1.173}$$

where $h\nu$ and $h\nu'$ are the energies of the initial and scattered photons, respectively, $m_e c^2$ is the rest mass energy of the initial electron, $(K_e + m_e c^2)$ is the total energy of the recoiling electron, and K_e is its recoil kinetic energy. The expression for K_e can immediately be inferred from (1.173):

$$K_e = h(\nu - \nu') = hc\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) = \frac{hc}{\lambda}\frac{\lambda' - \lambda}{\lambda'} = (h\nu)\frac{\Delta\lambda}{\lambda'}, \tag{1.174}$$

where the wave shift $\Delta\lambda$ is given by (1.36):

$$\begin{aligned}
\Delta\lambda &= \lambda' - \lambda = \frac{h}{m_e c}(1 - \cos\theta) = \frac{2\pi\hbar c}{m_e c^2}(1 - \cos\theta) \\
&= \frac{2\pi \times 197.33 \times 10^{-15}\text{ MeV m}}{0.511\text{ MeV}}(1 - \cos 60^\circ) \\
&= 0.0012\text{ nm}.
\end{aligned} \tag{1.175}$$

Since the wavelength of the incident photon is $\lambda = 2\pi\hbar c/(h\nu)$, we have $\lambda = 2\pi \times 197.33 \times 10^{-15}$ MeV m$/(0.003$ MeV$) = 0.414$ nm; the wavelength of the scattered photon is given by

$$\lambda' = \lambda + \Delta\lambda = 0.4152\text{ nm}. \tag{1.176}$$

Now, substituting the numerical values of λ' and $\Delta\lambda$ into (1.174), we obtain the kinetic energy of the recoiling electron

$$K_e = (h\nu)\frac{\Delta\lambda}{\lambda'} = (3\text{ keV}) \times \frac{0.0012\text{ nm}}{0.4152\text{ nm}} = 8.671\text{ eV}. \tag{1.177}$$

(b) To obtain the angle at which the electron recoils, we need simply to use the conservation of the total momentum along the $x-$ and $y-$ axes:

$$p = p_e \cos\phi + p' \cos\theta, \qquad 0 = p_e \sin\phi - p' \sin\theta. \tag{1.178}$$

These can be rewritten as

$$p_e \cos\phi = p - p' \cos\theta, \qquad p_e \sin\phi = p' \sin\theta, \tag{1.179}$$

where p and p' are the momenta of the initial and final photons, p_e is the momentum of the recoiling electron, and θ and ϕ are the angles at which the photon and electron scatter, respectively (Figure 1.4). Taking (1.179) and dividing the second equation by the first, we obtain

$$\tan\phi = \frac{\sin\theta}{p/p' - \cos\theta} = \frac{\sin\theta}{\lambda'/\lambda - \cos\theta}, \tag{1.180}$$

where we have used the momentum expressions of the incident photon $p = h/\lambda$ and of the scattered photon $p' = h/\lambda'$. Since $\lambda = 0.414$ nm and $\lambda' = 0.4152$ nm, the angle at which the electron recoils is given by

$$\phi = \tan^{-1}\left(\frac{\sin\theta}{\lambda'/\lambda - \cos\theta}\right) = \tan^{-1}\left(\frac{\sin 60^\circ}{0.4152/0.414 - \cos 60^\circ}\right) = 59.86^\circ. \tag{1.181}$$

Problem 1.6
Show that the maximum kinetic energy transferred to a proton when hit by a photon of energy $h\nu$ is $K_p = h\nu/[1 + m_p c^2/(2h\nu)]$, where m_p is the mass of the proton.

Solution
Using (1.35), we have

$$\frac{1}{\nu'} = \frac{1}{\nu} + \frac{h}{m_p c^2}(1 - \cos\theta), \tag{1.182}$$

which leads to

$$h\nu' = \frac{h\nu}{1 + (h\nu/m_p c^2)(1 - \cos\theta)}. \tag{1.183}$$

Since the kinetic energy transferred to the proton is given by $K_p = h\nu - h\nu'$, we obtain

$$K_p = h\nu - \frac{h\nu}{1 + (h\nu/m_p c^2)(1 - \cos\theta)} = \frac{h\nu}{1 + m_p c^2/[h\nu(1 - \cos\theta)]}. \tag{1.184}$$

Clearly, the maximum kinetic energy of the proton corresponds to the case where the photon scatters backwards ($\theta = \pi$),

$$K_p = \frac{h\nu}{1 + m_p c^2/(2h\nu)}. \tag{1.185}$$

Problem 1.7
Consider a photon that scatters from an electron at rest. If the Compton wavelength shift is observed to be triple the wavelength of the incident photon and if the photon scatters at 60°, calculate

(a) the wavelength of the incident photon,
(b) the energy of the recoiling electron, and
(c) the angle at which the electron scatters.

Solution

(a) In the case where the photons scatter at $\theta = 60^\circ$ and since $\Delta\lambda = 3\lambda$, the wave shift relation (1.36) yields

$$3\lambda = \frac{h}{m_e c}(1 - \cos 60^\circ), \tag{1.186}$$

which in turn leads to

$$\lambda = \frac{h}{6m_e c} = \frac{\pi\hbar c}{3m_e c^2} = \frac{3.14 \times 197.33 \times 10^{-15}\ \text{MeV m}}{3 \times 0.511\ \text{MeV}} = 4.04 \times 10^{-13}\ \text{m}. \tag{1.187}$$

(b) The energy of the recoiling electron can be obtained from the conservation of energy:

$$K_e = hc\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) = \frac{3hc}{4\lambda} = \frac{3\pi\hbar c}{2\lambda} = \frac{3 \times 3.14 \times 197.33 \times 10^{-15}\ \text{MeV m}}{2 \times 4.04 \times 10^{-13}\ \text{m}} = 2.3\ \text{MeV}. \tag{1.188}$$

In deriving this relation, we have used the fact that $\lambda' = \lambda + \Delta\lambda = 4\lambda$.

(c) Since $\lambda' = 4\lambda$ the angle ϕ at which the electron recoils can be inferred from (1.181)

$$\phi = \tan^{-1}\left(\frac{\sin\theta}{\lambda'/\lambda - \cos\theta}\right) = \tan^{-1}\left(\frac{\sin 60^\circ}{4 - \cos 60^\circ}\right) = 13.9^\circ. \tag{1.189}$$

Problem 1.8

In a double-slit experiment with a source of monoenergetic electrons, detectors are placed along a vertical screen parallel to the y-axis to monitor the diffraction pattern of the electrons emitted from the two slits. When only one slit is open, the amplitude of the electrons detected on the screen is $\psi_1(y,t) = A_1 e^{-i(ky-\omega t)}/\sqrt{1+y^2}$, and when only the other is open the amplitude is $\psi_2(y,t) = A_2 e^{-i(ky+\pi y-\omega t)}/\sqrt{1+y^2}$, where A_1 and A_2 are normalization constants that need to be found. Calculate the intensity detected on the screen when

(a) both slits are open and a light source is used to determine which of the slits the electron went through and

(b) both slits are open and no light source is used.

Plot the intensity registered on the screen as a function of y for cases (a) and (b).

Solution

Using the integral $\int_{-\infty}^{+\infty} dy/(1+y^2) = \pi$, we can obtain the normalization constants at once: $A_1 = A_2 = 1/\sqrt{\pi}$; hence ψ_1 and ψ_2 become $\psi_1(y,t) = e^{-i(ky-\omega t)}/\sqrt{\pi(1+y^2)}$, $\psi_2(y,t) = e^{-i(ky+\pi y-\omega t)}/\sqrt{\pi(1+y^2)}$.

(a) When we use a light source to observe the electrons as they exit from the two slits on their way to the vertical screen, the total intensity recorded on the screen will be determined by a simple addition of the probability densities (or of the separate intensities):

$$I(y) = |\psi_1(y,t)|^2 + |\psi_2(y,t)|^2 = \frac{2}{\pi(1+y^2)}. \tag{1.190}$$

As depicted in Figure 1.17a, the shape of the total intensity displays no interference pattern. Intruding on the electrons with the light source, we distort their motion.

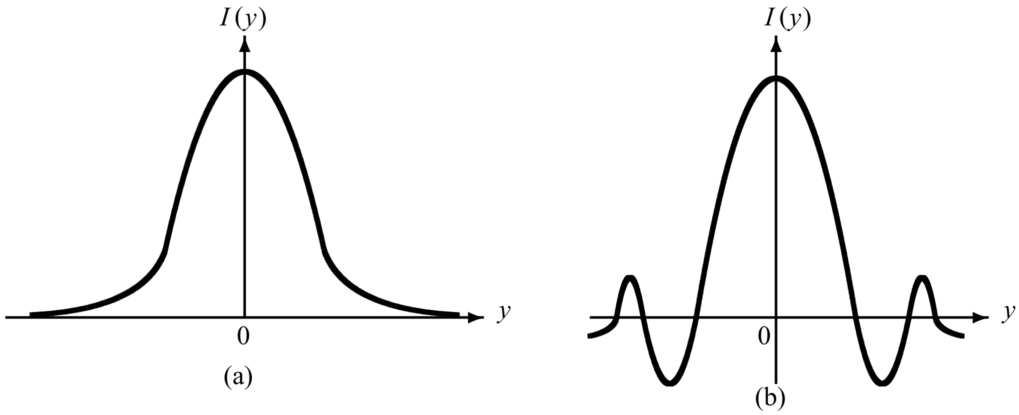

Figure 1.17 Shape of the total intensity generated in a double slit experiment when both slits are open and (a) a light source is used to observe the electrons' motion, $I(y) = 2/\pi(1 + y^2)$, and no interference is registered; (b) no light source is used, $I(y) = 4/[\pi(1+y^2)]\cos^2(\pi y/2)$, and an interference pattern occurs.

(b) When no light source is used to observe the electrons, the motion will not be distorted and the total intensity will be determined by an addition of the amplitudes, not the intensities:

$$\begin{aligned} I(y) &= |\psi_1(y,t) + \psi_2(y,t)|^2 = \frac{1}{\pi(1+y^2)}\left|e^{-i(ky-\omega t)} + e^{-i(ky+\pi y-\omega t)}\right|^2 \\ &= \frac{1}{\pi(1+y^2)}\left(1+e^{i\pi y}\right)\left(1+e^{-i\pi y}\right) \\ &= \frac{4}{\pi(1+y^2)}\cos^2\left(\frac{\pi}{2}y\right). \end{aligned} \tag{1.191}$$

The shape of this intensity does display an interference pattern which, as shown in Figure 1.17b, results from an oscillating function, $\cos^2(\pi y/2)$, modulated by $4/[\pi(1+y^2)]$.

Problem 1.9

Consider a head-on collision between an α-particle and a lead nucleus. Neglecting the recoil of the lead nucleus, calculate the distance of closest approach of a 9.0 MeV α-particle to the nucleus.

Solution

In this head-on collision the distance of closest approach r_0 can be obtained from the conservation of energy $E_i = E_f$, where E_i is the initial energy of the system, α-particle plus the lead nucleus, when the particle and the nucleus are far from each other and thus feel no electrostatic potential between them. Assuming the lead nucleus to be at rest, E_i is simply the energy of the α-particle: $E_i = 9.0$ MeV $= 9 \times 10^6 \times 1.6 \times 10^{-19}$ J.

As for E_f, it represents the energy of the system when the α-particle is at its closest distance from the nucleus. At this position, the α-particle is at rest and hence has no kinetic energy. The only energy the system has is the electrostatic potential energy between the α-particle and the lead nucleus, which has a positive charge of $82e$. Neglecting the recoil of the lead nucleus and since the charge of the α-particle is positive and equal to $2e$, we have $E_f = (2e)(82e)/(4\pi\varepsilon_0 r_0)$. The energy conservation $E_i = E_f$ or $(2e)(82e)/(4\pi\varepsilon_0 r_0) = E_i$ leads at once to

$$r_0 = \frac{(2e)(82e)}{4\pi\varepsilon_0 E_i} = 2.62 \times 10^{-14} \text{ m}, \tag{1.192}$$

where we used the values $e = 1.6 \times 10^{-19}$ C and $1/(4\pi\varepsilon_0) = 8.9 \times 10^9$ N m^2 C^{-2}.

Problem 1.10

Considering that a quintuply ionized carbon ion, C^{5+}, behaves like a hydrogen atom, calculate

(a) the radius r_n and energy E_n for a given state n and compare them with the corresponding expressions for hydrogen,

(b) the ionization energy of C^{5+} when it is in its first excited state and compare it with the corresponding value for hydrogen, and

(c) the wavelength corresponding to the transition from state $n = 3$ to state $n = 1$; compare it with the corresponding value for hydrogen.

Solution

(a) The C^{5+} ion is generated by removing five electrons from the carbon atom. To find the expressions for r_{n_C} and E_{n_C} for the C^{5+} ion (which has 6 protons), we need simply to insert $Z = 6$ into (1.76):

$$r_{n_C} = \frac{a_0}{6} n^2, \qquad E_{n_C} = -\frac{36\mathcal{R}}{n^2}, \tag{1.193}$$

where we have dropped the term m_e/M, since it is too small compared to one. Clearly, these expressions are related to their hydrogen counterparts by

$$r_{n_C} = \frac{a_0}{6} n^2 = \frac{r_{n_H}}{6}, \qquad E_{n_C} = -\frac{36\mathcal{R}}{n^2} = 36 E_{n_H}. \tag{1.194}$$

(b) The ionization energy is the one needed to remove the only remaining electron of the C^{5+} ion. When the C^{5+} ion is in its first excited state, the ionization energy is

$$E_{2_C} = -\frac{36\mathcal{R}}{4} = -9 \times 13.6 \text{ eV} = -122.4 \text{ eV}, \tag{1.195}$$

which is equal to 36 times the energy needed to ionize the hydrogen atom in its first excited state: $E_{2_H} = -3.4$ eV (note that we have taken $n = 2$ to correspond to the first excited state; as a result, the cases $n = 1$ and $n = 3$ will correspond to the ground and second excited states, respectively).

(c) The wavelength corresponding to the transition from state $n = 3$ to state $n = 1$ can be inferred from the relation $hc/\lambda = E_{3_C} - E_{1_C}$ which, when combined with $E_{1_C} = -489.6$ eV and $E_{3_C} = -54.4$ eV, leads to

$$\lambda = \frac{hc}{E_{3_C} - E_{1_C}} = \frac{2\pi\hbar c}{E_{3_C} - E_{1_C}} = \frac{2\pi\, 197.33 \times 10^{-9} \text{ eV m}}{-54.4 \text{ eV} + 489.6 \text{ eV}} = 2.85 \text{ nm}. \tag{1.196}$$

Problem 1.11

(a) Find the Fourier transform for $\phi(k) = \begin{cases} A\,(a - |k|)\,, & |k| \le a, \\ 0, & |k| > a. \end{cases}$
where a is a positive parameter and A is a normalization factor to be found.

(b) Calculate the uncertainties Δx and Δp and check whether they satisfy the uncertainty principle.

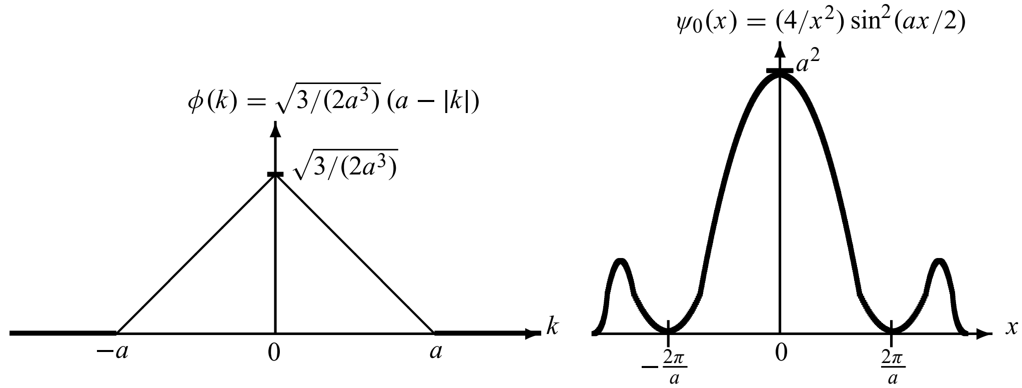

Figure 1.18 The shape of the function $\phi(k)$ and its Fourier transform $\psi_0(x)$.

Solution

(a) The normalization factor A can be found at once:

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} |\phi(k)|^2 dk = |A|^2 \int_{-a}^{0} (a+k)^2 dk + |A|^2 \int_0^a (a-k)^2 dk \\ &= 2|A|^2 \int_0^a (a-k)^2 dk = 2|A|^2 \int_0^a \left(a^2 - 2ak + k^2\right) dk \\ &= \frac{2a^3}{3}|A|^2, \end{aligned} \tag{1.197}$$

which yields $A = \sqrt{3/(2a^3)}$. The shape of $\phi(k) = \sqrt{3/(2a^3)}\,(a-|k|)$ is displayed in Figure 1.18.

Now, the Fourier transform of $\phi(k)$ is

$$\begin{aligned} \psi_0(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k) e^{ikx} dk \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{3}{2a^3}} \left[\int_{-a}^{0} (a+k)\, e^{ikx} dk + \int_0^a (a-k)\, e^{ikx} dk \right] \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{3}{2a^3}} \left[\int_{-a}^{0} k e^{ikx} dk - \int_0^a k e^{ikx} dk + a \int_{-a}^{a} e^{ikx} dk \right]. \end{aligned} \tag{1.198}$$

Using the integrations

$$\int_{-a}^{0} k e^{ikx} dk = \frac{a}{ix} e^{-iax} + \frac{1}{x^2}\left(1 - e^{-iax}\right), \tag{1.199}$$

$$\int_0^a k e^{ikx} dk = \frac{a}{ix} e^{iax} + \frac{1}{x^2}\left(e^{iax} - 1\right), \tag{1.200}$$

$$\int_{-a}^{a} e^{ikx} dk = \frac{1}{ix}\left(e^{iax} - e^{-iax}\right) = \frac{2\sin(ax)}{x}, \tag{1.201}$$

and after some straightforward calculations, we end up with

$$\psi_0(x) = \frac{4}{x^2} \sin^2\left(\frac{ax}{2}\right). \tag{1.202}$$

As shown in Figure 1.18, this wave packet is localized: it peaks at $x = 0$ and decreases gradually as x increases. We can verify that the maximum of $\psi_0(x)$ occurs at $x = 0$; writing $\psi_0(x)$ as $a^2(ax/2)^{-2}\sin^2(ax/2)$ and since $\lim_{x\to 0} \sin(bx)/(bx) \to 1$, we obtain $\psi_0(0) = a^2$.

(b) Figure 1.18a is quite suggestive in defining the half-width of $\phi(k)$: $\Delta k = a$ (hence the momentum uncertainty is $\Delta p = \hbar a$). By defining the width as $\Delta k = a$, we know with full certainty that the particle is located between $-a \le k \le a$; according to Figure 1.18a, the probability of finding the particle outside this interval is zero, for $\phi(k)$ vanishes when $|k| > a$.

Now, let us find the width Δx of $\psi_0(x)$. Since $\sin(a\pi/2a) = 1$, $\psi_0(\pi/a) = 4a^2/\pi^2$, and that $\psi_0(0) = a^2$, we can obtain from (1.202) that $\psi_0(\pi/a) = 4a^2/\pi^2 = 4/\pi^2\psi_0(0)$, or

$$\frac{\psi_0(\pi/a)}{\psi_0(0)} = \frac{4}{\pi^2}. \tag{1.203}$$

This suggests that $\Delta x = \pi/a$: when $x = \pm\Delta x = \pm\pi/a$ the wave packet $\psi_0(x)$ drops to $4/\pi^2$ from its maximum value $\psi_0(0) = a^2$. In sum, we have $\Delta x = \pi/a$ and $\Delta k = a$; hence

$$\Delta x \, \Delta k = \pi \tag{1.204}$$

or

$$\Delta x \, \Delta p = \pi\hbar, \tag{1.205}$$

since $\Delta k = \Delta p/\hbar$. In addition to satisfying Heisenberg's uncertainty principle (1.57), this relation shows that the product $\Delta x \, \Delta p$ is higher than $\hbar/2$: $\Delta x \, \Delta p > \hbar/2$. The wave packet (1.202) therefore offers a clear illustration of the general statement outlined above; namely, only Gaussian wave packets yield the *lowest* limit to Heisenberg's uncertainty principle $\Delta x \, \Delta p = \hbar/2$ (see (1.114)). All other wave packets, such as (1.202), yield higher values for the product $\Delta x \, \Delta p$.

Problem 1.12

Calculate the group and phase velocities for the wave packet corresponding to a relativistic particle.

Solution

Recall that the energy and momentum of a relativistic particle are given by

$$E = mc^2 = \frac{m_0 c^2}{\sqrt{1 - v^2/c^2}}, \qquad p = mv = \frac{m_0 v}{\sqrt{1 - v^2/c^2}}, \tag{1.206}$$

where m_0 is the rest mass of the particle and c is the speed of light in a vacuum. Squaring and adding the expressions of E and p, we obtain $E^2 = p^2c^2 + m_0^2c^4$; hence

$$E = c\sqrt{p^2 + m_0^2 c^2}. \tag{1.207}$$

Using this relation along with $p^2 + m_0^2 c^2 = m_0^2 c^2/(1 - \upsilon^2/c^2)$ and (1.122), we can show that the group velocity is given as follows:

$$\upsilon_g = \frac{dE}{dp} = \frac{d}{dp}\left(c\sqrt{p^2 + m_0^2 c^2}\right) = \frac{pc}{\sqrt{p^2 + m_0^2 c^2}} = \upsilon. \tag{1.208}$$

The group velocity is thus equal to the speed of the particle, $\upsilon_g = \upsilon$.

The phase velocity can be found from (1.122) and (1.207): $\upsilon_{ph} = E/p = c\sqrt{1 + m_0^2 c^2/p^2}$ which, when combined with $p = m_0\upsilon/\sqrt{1 - \upsilon^2/c^2}$, leads to $\sqrt{1 + m_0^2 c^2/p^2} = c/\upsilon$; hence

$$\upsilon_{ph} = \frac{E}{p} = c\sqrt{1 + \frac{m_0^2 c^2}{p^2}} = \frac{c^2}{\upsilon}. \tag{1.209}$$

This shows that the phase velocity of the wave corresponding to a relativistic particle with $m_0 \neq 0$ is larger than the speed of light, $\upsilon_{ph} = c^2/\upsilon > c$. This is indeed unphysical. The result $\upsilon_{ph} > c$ seems to violate the special theory of relativity, which states that the speed of material particles cannot exceed c. In fact, this principle is not violated because υ_{ph} does not represent the velocity of the particle; the velocity of the particle is represented by the group velocity (1.208). As a result, the phase speed of a relativistic particle has no meaningful physical significance.

Finally, the product of the group and phase velocities is equal to c^2, i.e., $\upsilon_g \upsilon_{ph} = c^2$.

Problem 1.13

The angular frequency of the surface waves in a liquid is given in terms of the wave number k by $\omega = \sqrt{gk + Tk^3/\rho}$, where g is the acceleration due to gravity, ρ is the density of the liquid, and T is the surface tension (which gives an upward force on an element of the surface liquid). Find the phase and group velocities for the limiting cases when the surface waves have: (a) very large wavelengths and (b) very small wavelengths.

Solution

The phase velocity can be found at once from (1.119):

$$\upsilon_{ph} = \frac{\omega}{k} = \sqrt{\frac{g}{k} + \frac{T}{\rho}k} = \sqrt{\frac{g\lambda}{2\pi} + \frac{2\pi T}{\rho\lambda}}, \tag{1.210}$$

where we have used the fact that $k = 2\pi/\lambda$, λ being the wavelength of the surface waves.

(a) If λ is very large, we can neglect the second term in (1.210); hence

$$\upsilon_{ph} = \sqrt{\frac{g\lambda}{2\pi}} = \sqrt{\frac{g}{k}}. \tag{1.211}$$

In this approximation the phase velocity does not depend on the nature of the liquid, since it depends on no parameter pertaining to the liquid such as its density or surface tension. This case corresponds, for instance, to deepwater waves, called gravity waves.

To obtain the group velocity, let us differentiate (1.211) with respect to k: $d\upsilon_{ph}/dk = -(1/2k)\sqrt{g/k} = -\upsilon_{ph}/2k$. A substitution of this relation into (1.120) shows that the group velocity is half the phase velocity:

$$\upsilon_g = \frac{d\omega}{dk} = \upsilon_{ph} + k\frac{d\upsilon_{ph}}{dk} = \upsilon_{ph} - \frac{1}{2}\upsilon_{ph} = \frac{1}{2}\upsilon_{ph} = \frac{1}{2}\sqrt{\frac{g\lambda}{2\pi}}. \tag{1.212}$$

The longer the wavelength, the faster the group velocity. This explains why a strong, steady wind will produce waves of longer wavelength than those produced by a swift wind.

(b) If λ is very small, the second term in (1.210) becomes the dominant one. So, retaining only the second term, we have

$$\upsilon_{ph} = \sqrt{\frac{2\pi T}{\rho\lambda}} = \sqrt{\frac{T}{\rho}k}, \tag{1.213}$$

which leads to $d\upsilon_{ph}/dk = \sqrt{Tk/\rho}/2k = \upsilon_{ph}/2k$. Inserting this expression into (1.120), we obtain the group velocity

$$\upsilon_g = \upsilon_{ph} + k\frac{d\upsilon_{ph}}{dk} = \upsilon_{ph} + \frac{1}{2}\upsilon_{ph} = \frac{3}{2}\upsilon_{ph}; \tag{1.214}$$

hence the smaller the wavelength, the faster the group velocity. These are called ripple waves; they occur, for instance, when a container is subject to vibrations of high frequency and small amplitude or when a gentle wind blows on the surface of a fluid.

Problem 1.14

This problem is designed to illustrate the superposition principle and the concepts of modulated and modulating functions in a wave packet. Consider two wave functions $\psi_1(y, t) = 5y \cos 7t$ and $\psi_2(y, t) = -5y \cos 9t$, where y and t are in meters and seconds, respectively. Show that their superposition generates a wave packet. Plot it and identify the modulated and modulating functions.

Solution

Using the relation $\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$, we can write the superposition of $\psi_1(y, t)$ and $\psi_2(y, t)$ as follows:

$$\begin{aligned} \psi(y,t) &= \psi_1(y,t) + \psi_2(y,t) = 5y\cos 7t - 5y\cos 9t \\ &= 5y(\cos 8t\cos t + \sin 8t\sin t) - 5y(\cos 8t\cos t - \sin 8t\sin t) \\ &= 10y\sin t\sin 8t. \end{aligned} \tag{1.215}$$

The periods of $10y\sin t$ and $\sin(8t)$ are given by 2π and $2\pi/8$, respectively. Since the period of $10y\sin t$ is larger than that of $\sin 8t$, $10y\sin t$ must be the modulating function and $\sin 8t$ the modulated function. As depicted in Figure 1.19, we see that $\sin 8t$ is modulated by $10y\sin t$.

Problem 1.15

(a) Calculate the final size of the wave packet representing a free particle after traveling a distance of 100 m for the following four cases where the particle is

(i) a 25 eV electron whose wave packet has an initial width of 10^{-6} m,

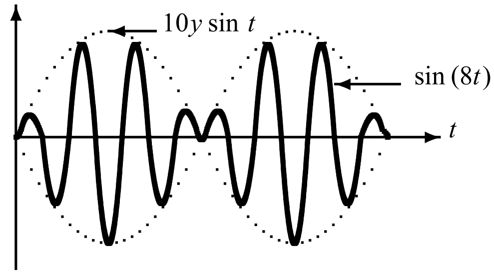

Figure 1.19 Shape of the wave packet $\psi(y, t) = 10y \sin t \sin 8t$. The function $\sin 8t$, the solid curve, is modulated by $10y \sin t$, the dashed curve.

(ii) a 25 eV electron whose wave packet has an initial width of 10^{-8} m,
(iii) a 100 MeV electron whose wave packet has an initial width of 1 mm, and
(iv) a 100 g object of size 1 cm moving at a speed of 50 m s^{-1}.
(b) Estimate the times required for the wave packets of the electron in (i) and the object in (iv) to spread to 10 mm and 10 cm, respectively. Discuss the results obtained.

Solution

(a) If the initial width of the wave packet of the particle is Δx_0, the width at time t is given by

$$\Delta x(t) = \Delta x_0 \sqrt{1 + \left(\frac{\delta x}{\Delta x_0}\right)^2}, \tag{1.216}$$

where the dispersion factor is given by

$$\frac{\delta x}{\Delta x_0} = \frac{2\hbar t}{ma^2} = \frac{\hbar t}{2m(a/2)^2} = \frac{\hbar t}{2m\,(\Delta x_0)^2}. \tag{1.217}$$

(i) For the 25 eV electron, which is clearly not relativistic, the time to travel the $L = 100$ m distance is given by $t = L/v = L\sqrt{mc^2/2E}/c$, since $E = \frac{1}{2}mv^2 = \frac{1}{2}mc^2(v^2/c^2)$ or $v = c\sqrt{2E/(mc^2)}$. We can therefore write the dispersion factor as

$$\frac{\delta x}{\Delta x_0} = \frac{\hbar}{2m\Delta x_0^2}t = \frac{\hbar}{2m\Delta x_0^2}\frac{L}{c}\sqrt{\frac{mc^2}{2E}} = \frac{\hbar c L}{2mc^2\Delta x_0^2}\sqrt{\frac{mc^2}{2E}}. \tag{1.218}$$

The numerics of this expression can be made easy by using the following quantities: $\hbar c \simeq 197 \times 10^{-15}$ MeV m, the rest mass energy of an electron is $mc^2 = 0.5$ MeV, $\Delta x_0 = 10^{-6}$ m, $E = 25$ eV $= 25 \times 10^{-6}$ MeV, and $L = 100$ m. Inserting these quantities into (1.218), we obtain

$$\frac{\delta x}{\Delta x_0} \simeq \frac{197 \times 10^{-15}\text{ MeV m} \times 100\text{ m}}{2 \times 0.5\text{ MeV} \times 10^{-12}\text{ m}^2}\sqrt{\frac{0.5\text{ MeV}}{2 \times 25 \times 10^{-6}\text{ MeV}}} \simeq 2 \times 10^3; \tag{1.219}$$

the time it takes the electron to travel the 100 m distance is given, as shown above, by

$$t = \frac{L}{c}\sqrt{\frac{mc^2}{2E}} = \frac{100\text{ m}}{3 \times 10^8\text{ m s}^{-1}}\sqrt{\frac{0.5\text{ MeV}}{2 \times 25 \times 10^{-6}\text{ MeV}}} = 3.3 \times 10^{-5}\text{ s}. \tag{1.220}$$

Using $t = 3.3 \times 10^{-5}$ s and substituting (1.219) into (1.216), we obtain

$$\Delta x(t = 3.3 \times 10^{-5}\ \text{s}) = 10^{-6}\ \text{m} \times \sqrt{1 + 4 \times 10^6} \simeq 2 \times 10^{-3}\ \text{m} = 2\ \text{mm}. \tag{1.221}$$

The width of the wave packet representing the electron has increased from an initial value of 10^{-6} m to 2×10^{-3} m, i.e., by a factor of about 10^3. The spread of the electron's wave packet is thus quite large.

(ii) The calculation needed here is identical to that of part (i), except the value of Δx_0 is now 10^{-8} m instead of 10^{-6} m. This leads to $\delta x/\Delta x_0 \simeq 2 \times 10^7$ and hence the width is $\Delta x(t) = 20$ cm; the width has therefore increased by a factor of about 10^7. This calculation is intended to show that *the narrower the initial wave packet, the larger the final spread*. In fact, starting in part (i) with an initial width of 10^{-6} m, the final width has increased to 2×10^{-3} m by a factor of about 10^3; but in part (ii) we started with an initial width of 10^{-8} m, and the final width has increased to 20 cm by a factor of about 10^7.

(iii) The motion of a 100 MeV electron is relativistic; hence to good approximation, its speed is equal to the speed of light, $\upsilon \simeq c$. Therefore the time it takes the electron to travel a distance of $L = 100$ m is $t \simeq L/c = 3.3 \times 10^{-7}$ s. The dispersion factor for this electron can be obtained from (1.217) where $\Delta x_0 = 10^{-3}$ m:

$$\frac{\delta x}{\Delta x_0} = \frac{\hbar L}{2mc\Delta x_0^2} = \frac{\hbar c L}{2mc^2\Delta x_0^2} \simeq \frac{197 \times 10^{-15}\ \text{MeV m} \times 100\ \text{m}}{2 \times 0.5\ \text{MeV} \times 10^{-6}\ \text{m}^2} \simeq 2 \times 10^{-5}. \tag{1.222}$$

The increase in the width of the wave packet is relatively small:

$$\Delta x(t = 3.3 \times 10^{-7}\ \text{s}) = 10^{-3}\ \text{m} \times \sqrt{1 + 4 \times 10^{-10}} \simeq 10^{-3}\ \text{m} = \Delta x_0. \tag{1.223}$$

So the width did not increase appreciably. We can conclude from this calculation that, when the motion of a microscopic particle is relativistic, the width of the corresponding wave packet increases by a relatively small amount.

(iv) In the case of a macroscopic object of mass $m = 0.1$ kg, the time to travel the distance $L = 100$ m is $t = L/\upsilon = 100\ \text{m}/50\ \text{m s}^{-1} = 2$ s. Since the size of the system is about $\Delta x_0 = 1$ cm $= 0.01$ m and $\hbar = 1.05 \times 10^{-34}$ J s, the dispersion factor for the object can be obtained from (1.217):

$$\frac{\delta x}{\Delta x_0} = \frac{\hbar t}{2m\Delta x_0^2} \simeq \frac{1.05 \times 10^{-34}\ \text{J s} \times 2\ \text{s}}{2 \times 0.1\ \text{kg} \times 10^{-4}\ \text{m}^2} \simeq 10^{-29}. \tag{1.224}$$

Since $\delta x/\Delta x_0 = 10^{-29} \ll 1$, the increase in the width of the wave packet is utterly undetectable:

$$\Delta x(2s) = 10^{-2}\ \text{m} \times \sqrt{1 + 10^{-58}} \simeq 10^{-2}\ \text{m} = \Delta x_0. \tag{1.225}$$

(b) Using (1.216) and (1.217) we obtain the expression for the time t in which the wave packet spreads to $\Delta x(t)$:

$$t = \tau\sqrt{\left(\frac{\Delta x(t)}{\Delta x_0}\right)^2 - 1}, \tag{1.226}$$

where τ represents a time constant $\tau = 2m(\Delta x_0)^2/\hbar$ (see (1.148)). The time constant for the electron of part (i) is given by

$$\tau = \frac{2mc^2(\Delta x_0)^2}{\hbar c^2} \simeq \frac{2 \times 0.5\ \text{MeV} \times 10^{-12}\ \text{m}^2}{197 \times 10^{-15}\ \text{MeV m} \times 3 \times 10^8\text{m s}^{-1}} = 1.7 \times 10^{-8}\ \text{s}, \tag{1.227}$$

and the time constant for the object of part (iv) is given by

$$\tau = \frac{2m(\Delta x_0)^2}{\hbar} \simeq \frac{2 \times 0.1 \text{ kg} \times 10^{-4} \text{ m}^2}{1.05 \times 10^{-34} \text{ J s}} = 1.9 \times 10^{29} \text{ s}. \tag{1.228}$$

Note that the time constant, while very small for a microscopic particle, is exceedingly large for macroscopic objects.

On the one hand, a substitution of the time constant (1.227) into (1.226) yields the time required for the electron's packet to spread to 10 mm:

$$t = 1.7 \times 10^{-8} \text{ s}\sqrt{\left(\frac{10^{-2}}{10^{-6}}\right)^2 - 1} \simeq 1.7 \times 10^{-4} \text{ s}. \tag{1.229}$$

On the other hand, a substitution of (1.228) into (1.226) gives the time required for the object to spread to 10 cm:

$$t = 1.9 \times 10^{29} \text{ s}\sqrt{\left(\frac{10^{-1}}{10^{-2}}\right)^2 - 1} \simeq 1.9 \times 10^{30} \text{ s}. \tag{1.230}$$

The result (1.229) shows that the size of the electron's wave packet grows in a matter of 1.7×10^{-4} s from 10^{-6} m to 10^{-2} m, a very large spread in a very short time. As for (1.230), it shows that the object has to be constantly in motion for about 1.9×10^{30} s for its wave packet to grow from 1 cm to 10 cm, a small spread for such an absurdly large time; this time is absurd because it is much larger than the age of the Universe, which is about 4.7×10^{17} s. We see that the spread of macroscopic objects becomes appreciable only if the motion lasts for a long, long time. However, the spread of microscopic objects is fast and large.

We can summarize these ideas in three points:

- The width of the wave packet of a nonrelativistic, microscopic particle increases substantially and quickly. The narrower the wave packet at the start, the further and the quicker it will spread.
- When the particle is microscopic and relativistic, the width corresponding to its wave packet does not increase appreciably.
- For a nonrelativistic, macroscopic particle, the width of its corresponding wave packet remains practically constant. The spread becomes appreciable only after absurdly long times, times that are larger than the lifetime of the Universe itself!

Problem 1.16

A neutron is confined in space to 10^{-14} m. Calculate the time its packet will take to spread to

(a) four times its original size,

(b) a size equal to the Earth's diameter, and

(c) a size equal to the distance between the Earth and the Moon.

Solution

Since the rest mass energy of a neutron is equal to $m_n c^2 = 939.6$ MeV, we can infer the time constant for the neutron from (1.227):

$$\tau = \frac{2m_n c^2(\Delta x_0)^2}{\hbar c^2} \simeq \frac{2 \times 939.6 \text{ MeV} \times (10^{-14} \text{ m})^2}{197 \times 10^{-15} \text{ MeV m} \times 3 \times 10^8 \text{ m s}^{-1}} = 3.2 \times 10^{-21} \text{ s}. \tag{1.231}$$

Inserting this value in (1.226) we obtain the time it takes for the neutron's packet to grow from an initial width Δx_0 to a final size $\Delta x(t)$:

$$t = \tau\sqrt{\left(\frac{\Delta x(t)}{\Delta x_0}\right)^2 - 1} = 3.2 \times 10^{-21}\ \mathrm{s}\sqrt{\left(\frac{\Delta x(t)}{\Delta x_0}\right)^2 - 1}. \tag{1.232}$$

The calculation of t reduces to simple substitutions.

(a) Substituting $\Delta x(t) = 4\Delta x_0$ into (1.232), we obtain the time needed for the neutron's packet to expand to four times its original size:

$$t = 3.2 \times 10^{-21}\ \mathrm{s}\sqrt{16 - 1} = 1.2 \times 10^{-20}\ \mathrm{s}. \tag{1.233}$$

(b) The neutron's packet will expand from an initial size of 10^{-14} m to 12.7×10^6 m (the diameter of the Earth) in a time of

$$t = 3.2 \times 10^{-21}\ \mathrm{s}\sqrt{\left(\frac{12.7 \times 10^6\ \mathrm{m}}{10^{-14}\ \mathrm{m}}\right)^2 - 1} = 4.1\ \mathrm{s}. \tag{1.234}$$

(c) The time needed for the neutron's packet to spread from 10^{-14} m to 3.84×10^8 m (the distance between the Earth and the Moon) is

$$t = 3.2 \times 10^{-21}\ \mathrm{s}\sqrt{\left(\frac{3.84 \times 10^8\ \mathrm{m}}{10^{-14}\ \mathrm{m}}\right)^2 - 1} = 12.3\ \mathrm{s}. \tag{1.235}$$

The calculations carried out in this problem show that the spread of the packets of microscopic particles is significant and occurs very fast: the size of the packet for an earthly neutron can expand to reach the Moon in a mere 12.3 s! Such an immense expansion in such a short time is indeed hard to visualize classically. One should not confuse the packet's expansion with a growth in the size of the system. As mentioned above, the spread of the wave packet does not mean that the material particle becomes bloated. It simply implies a redistribution of the probability density. In spite of the significant spread of the wave packet, the packet's norm is always conserved; as shown in (1.149) it is equal to 1.

Problem 1.17

Use the uncertainty principle to estimate: (a) the ground state radius of the hydrogen atom and (b) the ground state energy of the hydrogen atom.

Solution

(a) According to the uncertainty principle, the electron's momentum and the radius of its orbit are related by $rp \sim \hbar$; hence $p \sim \hbar/r$. To find the ground state radius, we simply need to minimize the electron–proton energy

$$E(r) = \frac{p^2}{2m_e} - \frac{e^2}{4\pi\varepsilon_0 r} = \frac{\hbar^2}{2m_e r^2} - \frac{e^2}{4\pi\varepsilon_0 r} \tag{1.236}$$

with respect to r:

$$0 = \frac{dE}{dr} = -\frac{\hbar^2}{m_e r_0^3} + \frac{e^2}{4\pi\varepsilon_0 r_0^2}. \tag{1.237}$$

This leads to the Bohr radius

$$r_0 = \frac{4\pi\varepsilon_0\hbar^2}{m_e e^2} = 0.053 \text{ nm}. \tag{1.238}$$

(b) Inserting (1.238) into (1.236), we obtain the Bohr energy:

$$E(r_0) = \frac{\hbar^2}{2mr_0^2} - \frac{e^2}{4\pi\varepsilon_0 r_0} = -\frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 = -13.6 \text{ eV}. \tag{1.239}$$

The results obtained for r_0 and $E(r_0)$, as shown in (1.238) and (1.239), are indeed impressively accurate given the crudeness of the approximation.

Problem 1.18

Consider the bound state of two quarks having the same mass m and interacting via a potential energy $V(r) = kr$ where k is a constant.

(a) Using the Bohr model, find the speed, the radius, and the energy of the system in the case of circular orbits. Determine also the angular frequency of the radiation generated by a transition of the system from energy state n to energy state m.

(b) Obtain numerical values for the speed, the radius, and the energy for the case of the ground state, $n = 1$, by taking a quark mass of $mc^2 = 2$ GeV and $k = 0.5$ GeV fm^{-1}.

Solution

(a) Consider the two quarks to move circularly, much like the electron and proton in a hydrogen atom; then we can write the force between them as

$$\mu\frac{v^2}{r} = \frac{dV(r)}{dr} = k, \tag{1.240}$$

where $\mu = m/2$ is the reduced mass and $V(r)$ is the potential. From the Bohr quantization condition of the orbital angular momentum, we have

$$L = \mu v r = n\hbar. \tag{1.241}$$

Multiplying (1.240) by (1.241), we end up with $\mu^2 v^3 = n\hbar k$, which yields the (quantized) speed of the relative motion for the two-quark system:

$$v_n = \left(\frac{\hbar k}{\mu^2}\right)^{1/3} n^{1/3}. \tag{1.242}$$

The radius can be obtained from (1.241), $r_n = n\hbar/(\mu v_n)$; using (1.242), this leads to

$$r_n = \left(\frac{\hbar^2}{\mu k}\right)^{1/3} n^{2/3}. \tag{1.243}$$

We can obtain the total energy of the relative motion by adding the kinetic and potential energies:

$$E_n = \frac{1}{2}\mu v_n^2 + kr_n = \frac{3}{2}\left(\frac{\hbar^2 k^2}{\mu}\right)^{1/3} n^{2/3}. \tag{1.244}$$

In deriving this relation, we have used the relations for υ_n and r_n as given by (1.242) by (1.243), respectively.

The angular frequency of the radiation generated by a transition from n to m is given by

$$\omega_{nm} = \frac{E_n - E_m}{\hbar} = \frac{3}{2}\left(\frac{k^2}{\mu\hbar}\right)^{1/3}\left(n^{2/3} - m^{2/3}\right). \tag{1.245}$$

(b) Inserting $n = 1$, $\hbar c \simeq 0.197$ GeV fm, $\mu c^2 = mc^2/2 = 1$ GeV, and $k = 0.5$ GeV fm^{-1} into (1.242) to (1.244), we have

$$\upsilon_1 = \left(\frac{\hbar ck}{(\mu c^2)^2}\right)^{1/3} c \simeq \left(\frac{0.197\text{ GeV fm} \times 0.5\text{ GeV fm}^{-1}}{(1\text{ GeV})^2}\right)^{1/3} c = 0.46c, \tag{1.246}$$

where c is the speed of light and

$$r_1 = \left(\frac{(\hbar c)^2}{\mu c^2 k}\right)^{1/3} \simeq \left(\frac{(0.197\text{ GeV fm})^2}{1\text{ GeV} \times 0.5\text{ GeV fm}^{-1}}\right)^{1/3} = 0.427\text{ fm}, \tag{1.247}$$

$$E_1 = \frac{3}{2}\left(\frac{(\hbar c)^2 k^2}{\mu c^2}\right)^{1/3} \simeq \frac{3}{2}\left(\frac{(0.197\text{ GeV fm})^2(0.5\text{ GeV fm}^{-1})^2}{1\text{ GeV}}\right)^{1/3} = 0.32\text{ GeV}. \tag{1.248}$$

1.11 Exercises

Exercise 1.1
Consider a metal that is being welded.

(a) How hot is the metal when it radiates most strongly at 490 nm?

(b) Assuming that it radiates like a blackbody, calculate the intensity of its radiation.

Exercise 1.2
Consider a star, a light bulb, and a slab of ice; their respective temperatures are 8500 K, 850 K, and 273.15 K.

(a) Estimate the wavelength at which their radiated energies peak.

(b) Estimate the intensities of their radiation.

Exercise 1.3
Consider a 75 W light bulb and an 850 W microwave oven. If the wavelengths of the radiation they emit are 500 nm and 150 mm, respectively, estimate the number of photons they emit per second. Are the quantum effects important in them?

Exercise 1.4
Assuming that a given star radiates like a blackbody, estimate

(a) the temperature at its surface and

(b) the wavelength of its strongest radiation,

when it emits a total intensity of 575 MW m^{-2}.

Chapter 2

Mathematical Tools of Quantum Mechanics

2.1 Introduction

We deal here with the mathematical machinery needed to study quantum mechanics. Although this chapter is mathematical in scope, no attempt is made to be mathematically complete or rigorous. We limit ourselves to those practical issues that are relevant to the formalism of quantum mechanics.

The Schrödinger equation is one of the cornerstones of the theory of quantum mechanics; it has the structure of a *linear* equation. The formalism of quantum mechanics deals with operators that are linear and wave functions that belong to an abstract Hilbert space. The mathematical properties and structure of Hilbert spaces are essential for a proper understanding of the formalism of quantum mechanics. For this, we are going to review briefly the properties of Hilbert spaces and those of linear operators. We will then consider Dirac's *bra-ket* notation.

Quantum mechanics was formulated in two different ways by Schrödinger and Heisenberg. Schrödinger's wave mechanics and Heisenberg's matrix mechanics are the representations of the general formalism of quantum mechanics in *continuous* and *discrete* basis systems, respectively. For this, we will also examine the mathematics involved in representing kets, bras, bra-kets, and operators in discrete and continuous bases.

2.2 The Hilbert Space and Wave Functions

2.2.1 The Linear Vector Space

A linear vector space consists of two sets of elements and two algebraic rules:

- a set of *vectors* $\psi,\ \phi,\ \chi,\ \ldots$ and a set of *scalars* $a, b, c, \ldots$;
- a rule for vector *addition* and a rule for scalar *multiplication*.

(a) Addition rule
The addition rule has the properties and structure of an *abelian* group:

- If ψ and ϕ are vectors (elements) of a space, their sum, $\psi + \phi$, is also a vector of the same space.
- Commutativity: $\psi + \phi = \phi + \psi$.
- Associativity: $(\psi + \phi) + \chi = \psi + (\phi + \chi)$.
- Existence of a zero or neutral vector: for each vector ψ, there must exist a zero vector O such that: $O + \psi = \psi + O = \psi$.
- Existence of a symmetric or inverse vector: each vector ψ must have a symmetric vector $(-\psi)$ such that $\psi + (-\psi) = (-\psi) + \psi = O$.

(b) Multiplication rule
The multiplication of vectors by scalars (scalars can be real or complex numbers) has these properties:

- The product of a scalar with a vector gives another vector. In general, if ψ and ϕ are two vectors of the space, any linear combination $a\psi + b\phi$ is also a vector of the space, a and b being scalars.
- Distributivity with respect to addition:

$$a(\psi + \phi) = a\psi + a\phi, \qquad (a+b)\psi = a\psi + b\psi, \tag{2.1}$$

- Associativity with respect to multiplication of scalars:

$$a(b\psi) = (ab)\psi \tag{2.2}$$

- For each element ψ there must exist a unitary scalar I and a zero scalar "o" such that

$$I\psi = \psi I = \psi \quad \text{and} \quad o\psi = \psi o = o. \tag{2.3}$$

2.2.2 The Hilbert Space

A Hilbert space $\mathcal{H}$ consists of a set of vectors $\psi, \phi, \chi, \ldots$ and a set of *scalars* $a, b, c, \ldots$ which satisfy the following *four* properties:

(a) $\mathcal{H}$ is a linear space

The properties of a linear space were considered in the previous section.

(b) $\mathcal{H}$ has a defined scalar product that is strictly positive

The scalar product of an element ψ with another element ϕ is in general a complex number, denoted by (ψ, ϕ), where (ψ, ϕ) = complex number. **Note:** Watch out for the order! Since the scalar product is a complex number, the quantity (ψ, ϕ) is generally not equal to (ϕ, ψ): $(\psi, \phi) = \psi^*\phi$ while $(\phi, \psi) = \phi^*\psi$. The scalar product satisfies the following properties:

- The scalar product of ψ with ϕ is equal to the complex conjugate of the scalar product of ϕ with ψ:

$$(\psi, \phi) = (\phi, \psi)^*. \tag{2.4}$$

- The scalar product of ϕ with ψ is linear with respect to the second factor if $\psi = a\psi_1 + b\psi_2$:
$$(\phi, a\psi_1 + b\psi_2) = a(\phi, \psi_1) + b(\phi, \psi_2), \tag{2.5}$$
and antilinear with respect to the first factor if $\phi = a\phi_1 + b\phi_2$:
$$(a\phi_1 + b\phi_2, \psi) = a^*(\phi_1, \psi) + b^*(\phi_2, \psi). \tag{2.6}$$

- The scalar product of a vector ψ with itself is a positive real number:
$$(\psi, \psi) = \| \psi \|^2 \geq 0, \tag{2.7}$$
where the equality holds only for $\psi = O$.

(c) $\mathcal{H}$ is separable

There exists a Cauchy sequence $\psi_n \in \mathcal{H}$ $(n = 1, 2, \ldots)$ such that for every ψ of $\mathcal{H}$ and $\varepsilon > 0$, there exists at least one ψ_n of the sequence for which
$$\| \psi - \psi_n \| < \varepsilon. \tag{2.8}$$

(d) $\mathcal{H}$ is complete

Every Cauchy sequence $\psi_n \in \mathcal{H}$ converges to an element of $\mathcal{H}$. That is, for any ψ_n, the relation
$$\lim_{n,m\to\infty} \| \psi_n - \psi_m \| = 0, \tag{2.9}$$
defines a unique limit ψ of $\mathcal{H}$ such that
$$\lim_{n\to\infty} \| \psi - \psi_n \| = 0. \tag{2.10}$$

Remark

We should note that in a scalar product (ϕ, ψ), the second factor, ψ, belongs to the Hilbert space $\mathcal{H}$, while the first factor, ϕ, belongs to its dual Hilbert space $\mathcal{H}_d$. The distinction between $\mathcal{H}$ and $\mathcal{H}_d$ is due to the fact that, as mentioned above, the scalar product is not commutative: $(\phi, \psi) \neq (\psi, \phi)$; the order matters! From linear algebra, we know that every vector space can be associated with a dual vector space.

2.2.3 Dimension and Basis of a Vector Space

A set of N nonzero vectors $\phi_1, \phi_2, \ldots, \phi_N$ is said to be *linearly independent* if and only if the solution of the equation
$$\sum_{i=1}^{N} a_i \phi_i = 0 \tag{2.11}$$
is $a_1 = a_2 = \cdots = a_N = 0$. But if there exists a set of scalars, which are not all zero, so that one of the vectors (say ϕ_n) can be expressed as a linear combination of the others,
$$\phi_n = \sum_{i=1}^{n-1} a_i \phi_i + \sum_{i=n+1}^{N} a_i \phi_i, \tag{2.12}$$

the set $\{\phi_i\}$ is said to be *linearly dependent*.
Dimension: The *dimension* of a vector space is given by the *maximum number* of linearly independent vectors the space can have. For instance, if the maximum number of linearly independent vectors a space has is N (i.e., $\phi_1, \phi_2, \ldots, \phi_N$), this space is said to be N-dimensional. In this N-dimensional vector space, any vector ψ can be expanded as a linear combination:

$$\psi = \sum_{i=1}^{N} a_i \phi_i. \tag{2.13}$$

Basis: The *basis* of a vector space consists of a set of the maximum possible number of linearly independent vectors belonging to that space. This set of vectors, $\phi_1, \phi_2, \ldots, \phi_N$, to be denoted in short by $\{\phi_i\}$, is called the basis of the vector space, while the vectors $\phi_1, \phi_2, \ldots, \phi_N$ are called the base vectors. Although the set of these linearly independent vectors is arbitrary, it is convenient to choose them *orthonormal*; that is, their scalar products satisfy the relation $(\phi_i, \phi_j) = \delta_{ij}$ (we may recall that $\delta_{ij} = 1$ whenever $i = j$ and zero otherwise). The basis is said to be *orthonormal* if it consists of a set of orthonormal vectors. Moreover, the basis is said to be *complete* if it spans the entire space; that is, there is no need to introduce any additional base vector. The expansion coefficients a_i in (2.13) are called the *components* of the vector ψ in the basis. Each component is given by the scalar product of ψ with the corresponding base vector, $a_j = (\phi_j, \psi)$.

Examples of linear vector spaces
Let us give two examples of linear spaces that are Hilbert spaces: one having a *finite (discrete)* set of base vectors, the other an *infinite (continuous)* basis.

- The first one is the three-dimensional Euclidean vector space; the basis of this space consists of three linearly independent vectors, usually denoted by $\vec{i}, \vec{j}, \vec{k}$. Any vector of the Euclidean space can be written in terms of the base vectors as $\vec{A} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$, where a_1, a_2, and a_3 are the components of $\vec{A}$ in the basis; each component can be determined by taking the scalar product of $\vec{A}$ with the corresponding base vector: $a_1 = \vec{i} \cdot \vec{A}$, $a_2 = \vec{j} \cdot \vec{A}$, and $a_3 = \vec{k} \cdot \vec{A}$. Note that the scalar product in the Euclidean space is real and hence symmetric. The norm in this space is the usual length of vectors $\| \vec{A} \| = A$. Note also that whenever $a_1\vec{i} + a_2\vec{j} + a_3\vec{k} = 0$ we have $a_1 = a_2 = a_3 = 0$ and that none of the unit vectors $\vec{i}, \vec{j}, \vec{k}$ can be expressed as a linear combination of the other two.

- The second example is the space of the entire complex functions $\psi(x)$; the dimension of this space is infinite for it has an infinite number of linearly independent basis vectors.

Example 2.1
Check whether the following sets of functions are linearly independent or dependent on the real x-axis.
(a) $f(x) = 4$, $g(x) = x^2$, $h(x) = e^{2x}$
(b) $f(x) = x$, $g(x) = x^2$, $h(x) = x^3$
(c) $f(x) = x$, $g(x) = 5x$, $h(x) = x^2$
(d) $f(x) = 2 + x^2$, $g(x) = 3 - x + 4x^3$, $h(x) = 2x + 3x^2 - 8x^3$

Solution

(a) The first set is clearly linearly independent since $a_1 f(x) + a_2 g(x) + a_3 h(x) = 4a_1 + a_2 x^2 + a_3 e^{2x} = 0$ implies that $a_1 = a_2 = a_3 = 0$ for any value of x.

(b) The functions $f(x) = x$, $g(x) = x^2$, $h(x) = x^3$ are also linearly independent since $a_1 x + a_2 x^2 + a_3 x^3 = 0$ implies that $a_1 = a_2 = a_3 = 0$ no matter what the value of x. For instance, taking $x = -1, 1, 3$, the following system of three equations

$$-a_1 + a_2 - a_3 = 0, \qquad a_1 + a_2 + a_3 = 0, \qquad 3a_1 + 9a_2 + 27a_3 = 0 \tag{2.14}$$

yields $a_1 = a_2 = a_3 = 0$.

(c) The functions $f(x) = x$, $g(x) = 5x$, $h(x) = x^2$ are not linearly independent, since $g(x) = 5f(x) + 0 \times h(x)$.

(d) The functions $f(x) = 2 + x^2$, $g(x) = 3 - x + 4x^3$, $h(x) = 2x + 3x^2 - 8x^3$ are not linearly independent since $h(x) = 3f(x) - 2g(x)$.

Example 2.2

Are the following sets of vectors (in the three-dimensional Euclidean space) linearly independent or dependent?

(a) $\vec{A} = (3, 0, 0)$, $\vec{B} = (0, -2, 0)$, $\vec{C} = (0, 0, -1)$

(b) $\vec{A} = (6, -9, 0)$, $\vec{B} = (-2, 3, 0)$

(c) $\vec{A} = (2, 3, -1)$, $\vec{B} = (0, 1, 2)$, $\vec{C} = (0, 0, -5)$

(d) $\vec{A} = (1, -2, 3)$, $\vec{B} = (-4, 1, 7)$, $\vec{C} = (0, 10, 11)$, and $\vec{D} = (14, 3, -4)$

Solution

(a) The three vectors $\vec{A} = (3, 0, 0)$, $\vec{B} = (0, -2, 0)$, $\vec{C} = (0, 0, -1)$ are linearly independent, since

$$a_1\vec{A} + a_2\vec{B} + a_3\vec{C} = 0 \Longrightarrow 3a_1\vec{i} - 2a_2\vec{j} - a_3\vec{k} = 0 \tag{2.15}$$

leads to

$$3a_1 = 0, \qquad -2a_2 = 0, \qquad -a_3 = 0, \tag{2.16}$$

which yields $a_1 = a_2 = a_3 = 0$.

(b) The vectors $\vec{A} = (6, -9, 0)$, $\vec{B} = (-2, 3, 0)$ are linearly dependent, since the solution to

$$a_1\vec{A} + a_2\vec{B} = 0 \quad \Longrightarrow \quad (6a_1 - 2a_2)\vec{i} + (-9a_1 + 3a_2)\vec{j} = 0 \tag{2.17}$$

is $a_1 = a_2/3$. The first vector is equal to -3 times the second one: $\vec{A} = -3\vec{B}$.

(c) The vectors $\vec{A} = (2, 3, -1)$, $\vec{B} = (0, 1, 2)$, $\vec{C} = (0, 0, -5)$ are linearly independent, since

$$a_1\vec{A} + a_2\vec{B} + a_3\vec{C} = 0 \Longrightarrow 2a_1\vec{i} + (3a_1 + a_2)\vec{j} + (-a_1 + 2a_2 - 5a_3)\vec{k} = 0 \tag{2.18}$$

leads to

$$2a_1 = 0, \qquad 3a_1 + a_2 = 0, \qquad -a_1 + 2a_2 - 5a_3 = 0. \tag{2.19}$$

The only solution of this system is $a_1 = a_2 = a_3 = 0$.

(d) The vectors $\vec{A} = (1, -2, 3)$, $\vec{B} = (-4, 1, 7)$, $\vec{C} = (0, 10, 11)$, and $\vec{D} = (14, 3, -4)$ are not linearly independent, because $\vec{D}$ can be expressed in terms of the other vectors:

$$\vec{D} = 2\vec{A} - 3\vec{B} + \vec{C}. \tag{2.20}$$

2.2.4 Square-Integrable Functions: Wave Functions

In the case of function spaces, a "vector" element is given by a *complex function* and the *scalar product* by *integrals*. That is, the scalar product of two functions $\psi(x)$ and $\phi(x)$ is given by

$$(\psi, \phi) = \int \psi^*(x)\phi(x)\, dx. \tag{2.21}$$

If this integral *diverges*, the scalar product *does not exist*. As a result, if we want the function space to possess a scalar product, we must select only those functions for which (ψ, ϕ) is *finite*. In particular, a function $\psi(x)$ is said to be *square integrable* if the scalar product of ψ with itself,

$$(\psi, \psi) = \int |\psi(x)|^2\, dx, \tag{2.22}$$

is *finite*.

It is easy to verify that the space of square-integrable functions possesses the properties of a Hilbert space. For instance, any linear combination of square-integrable functions is also a square-integrable function and (2.21) satisfies all the properties of the scalar product of a Hilbert space.

Note that the dimension of the Hilbert space of square-integrable functions is infinite, since each wave function can be expanded in terms of an infinite number of linearly independent functions. The dimension of a space is given by the maximum number of linearly independent basis vectors required to span that space.

A good example of square-integrable functions is the *wave function* of quantum mechanics, $\psi(\vec{r}, t)$. We have seen in Chapter 1 that, according to Born's probabilistic interpretation of $\psi(\vec{r}, t)$, the quantity $|\ \psi(\vec{r}, t)\ |^2\ d^3r$ represents the probability of finding, at time t, the particle in a volume d^3r, centered around the point $\vec{r}$. The probability of finding the particle somewhere in space must then be equal to 1:

$$\int |\ \psi(\vec{r}, t)\ |^2\ d^3r = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} |\ \psi(\vec{r}, t)\ |^2\ dz = 1; \tag{2.23}$$

hence the wave functions of quantum mechanics are square-integrable. Wave functions satisfying (2.23) are said to be normalized or square-integrable. As wave mechanics deals with square-integrable functions, any wave function which is not square-integrable has no physical meaning in quantum mechanics.

2.3 Dirac Notation

The physical state of a system is represented in quantum mechanics by elements of a Hilbert space; these elements are called state vectors. We can represent the state vectors in different bases by means of function expansions. This is analogous to specifying an ordinary (Euclidean) vector by its components in various coordinate systems. For instance, we can represent equivalently a vector by its components in a Cartesian coordinate system, in a spherical coordinate system, or in a cylindrical coordinate system. *The meaning of a vector is, of course, independent of the coordinate system chosen to represent its components*. Similarly, the state of a microscopic system has a meaning independent of the basis in which it is expanded.

To free state vectors from coordinate meaning, Dirac introduced what was to become an invaluable notation in quantum mechanics; it allows one to manipulate the formalism of quantum mechanics with ease and clarity. He introduced the concepts of kets, bras, and bra-kets, which will be explained below.

Kets: elements of a vector space
Dirac denoted the state vector ψ by the symbol $|\ \psi\rangle$, which he called a *ket* vector, or simply a ket. Kets belong to the Hilbert (vector) space $\mathcal{H}$, or, in short, to the ket-space.

Bras: elements of a dual space
As mentioned above, we know from linear algebra that a dual space can be associated with every vector space. Dirac denoted the elements of a dual space by the symbol $\langle\ |$, which he called a bra vector, or simply a bra; for instance, the element $\langle\psi\ |$ represents a bra. *Note: For every ket $|\ \psi\rangle$ there exists a unique bra $\langle\psi\ |$ and vice versa.* Again, while kets belong to the Hilbert space $\mathcal{H}$, the corresponding bras belong to its dual (Hilbert) space $\mathcal{H}_d$.

Bra-ket: Dirac notation for the scalar product
Dirac denoted the scalar (inner) product by the symbol $\langle\ |\ \rangle$, which he called a a *bra-ket*. For instance, the scalar product (ϕ, ψ) is denoted by the bra-ket $\langle\phi\ |\ \psi\rangle$:

$$(\phi, \psi) \quad \longrightarrow \quad \langle\phi\ |\ \psi\rangle. \tag{2.24}$$

Note: When a ket (or bra) is multiplied by a complex number, we also get a ket (or bra).

Remark: In wave mechanics we deal with wave functions $\psi(\vec{r}, t)$, but in the more general formalism of quantum mechanics we deal with abstract kets $|\ \psi\rangle$. Wave functions, like kets, are elements of a Hilbert space. We should note that, like a wave function, a ket represents the system completely, and hence knowing $|\ \psi\rangle$ means knowing all its amplitudes in all possible representations. As mentioned above, kets are independent of any particular representation. There is no reason to single out a particular representation basis such as the representation in the position space. Of course, if we want to know the probability of finding the particle at some position in space, we need to work out the formalism within the coordinate representation. The state vector of this particle at time t will be given by the spatial wave function $\langle\vec{r}, t\ |\ \psi\rangle = \psi(\vec{r}, t)$. In the coordinate representation, the scalar product $\langle\phi\ |\ \psi\rangle$ is given by

$$\langle\phi\ |\ \psi\rangle = \int \phi^*(\vec{r}, t)\psi(\vec{r}, t)\, d^3r. \tag{2.25}$$

Similarly, if we are considering the three-dimensional momentum of a particle, the ket $|\ \psi\rangle$ will have to be expressed in momentum space. In this case the state of the particle will be described by a wave function $\psi(\vec{p}, t)$, where $\vec{p}$ is the momentum of the particle.

Properties of kets, bras, and bra-kets

- **Every ket has a corresponding bra**

  To every *ket* $|\ \psi\rangle$, there corresponds a unique *bra* $\langle\psi\ |$ and vice versa:

$$|\ \psi\rangle \quad \longleftrightarrow \quad \langle\psi\ |\ . \tag{2.26}$$

  There is a one-to-one correspondence between bras and kets:

$$a\ |\ \psi\rangle + b\ |\ \phi\rangle \quad \longleftrightarrow \quad a^*\langle\psi\ | + b^*\langle\phi\ |, \tag{2.27}$$

  where a and b are complex numbers. The following is a common notation:

$$|\ a\psi\rangle\ = a\ |\ \psi\rangle, \qquad\qquad \langle a\psi\ | = a^*\langle\psi\ |\ . \tag{2.28}$$

- **Properties of the scalar product**

  In quantum mechanics, since the scalar product is a complex number, the ordering matters a lot. We must be careful to distinguish a scalar product from its complex conjugate; $\langle\psi \mid \phi\rangle$ is not the same thing as $\langle\phi \mid \psi\rangle$:

$$\langle\phi \mid \psi\rangle^* = \langle\psi \mid \phi\rangle. \tag{2.29}$$

  This property becomes clearer if we apply it to (2.21):

$$\langle\phi \mid \psi\rangle^* = \left(\int \phi^*(\vec{r}, t)\psi(\vec{r}, t)\, d^3r\right)^* = \int \psi^*(\vec{r}, t)\phi(\vec{r}, t)\, d^3r = \langle\psi \mid \phi\rangle. \tag{2.30}$$

  When $\mid \psi\rangle$ and $\mid \phi\rangle$ are real, we would have $\langle\psi \mid \phi\rangle = \langle\phi \mid \psi\rangle$. Let us list some additional properties of the scalar product:

$$\langle\psi \mid a_1\psi_1 + a_2\psi_2\rangle = a_1\langle\psi \mid \psi_1\rangle + a_2\langle\psi \mid \psi_2\rangle, \tag{2.31}$$

$$\langle a_1\phi_1 + a_2\phi_2 \mid \psi\rangle = a_1^*\langle\phi_1 \mid \psi\rangle + a_2^*\langle\phi_2 \mid \psi\rangle, \tag{2.32}$$

$$\begin{aligned}\langle a_1\phi_1 + a_2\phi_2 \mid b_1\psi_1 + b_2\psi_2\rangle &= a_1^*b_1\langle\phi_1 \mid \psi_1\rangle + a_1^*b_2\langle\phi_1 \mid \psi_2\rangle \\ &\quad + a_2^*b_1\langle\phi_2 \mid \psi_1\rangle + a_2^*b_2\langle\phi_2 \mid \psi_2\rangle.\end{aligned} \tag{2.33}$$

- **The norm is real and positive**

  For any state vector $\mid \psi\rangle$ of the Hilbert space $\mathcal{H}$, the norm $\langle\psi \mid \psi\rangle$ is real and positive; $\langle\psi \mid \psi\rangle$ is equal to zero only for the case where $\mid \psi\rangle = O$, where O is the zero vector. If the state $\mid \psi\rangle$ is normalized then $\langle\psi \mid \psi\rangle = 1$.

- **Schwarz inequality**

  For any two states $\mid \psi\rangle$ and $\mid \phi\rangle$ of the Hilbert space, we can show that

$$|\langle\psi \mid \phi\rangle|^2 \leq \langle\psi \mid \psi\rangle\langle\phi \mid \phi\rangle. \tag{2.34}$$

  If $\mid \psi\rangle$ and $\mid \phi\rangle$ are linearly dependent (i.e., proportional: $\mid \psi\rangle = \alpha \mid \phi\rangle$, where α is a scalar), this relation becomes an equality. The Schwarz inequality (2.34) is analogous to the following relation of the real Euclidean space

$$|\vec{A} \cdot \vec{B}|^2 \leq |\vec{A}|^2 |\vec{B}|^2 . \tag{2.35}$$

- **Triangle inequality**

$$\sqrt{\langle\psi + \phi \mid \psi + \phi\rangle} \leq \sqrt{\langle\psi \mid \psi\rangle} + \sqrt{\langle\phi \mid \phi\rangle}. \tag{2.36}$$

  If $\mid \psi\rangle$ and $\mid \phi\rangle$ are linearly dependent, $\mid \psi\rangle = \alpha \mid \phi\rangle$, and if the proportionality scalar α is real and positive, the triangle inequality becomes an equality. The counterpart of this inequality in Euclidean space is given by $|\vec{A} + \vec{B}| \leq |\vec{A}| + |\vec{B}|$.

- **Orthogonal states**

  Two *kets*, $\mid \psi\rangle$ and $\mid \phi\rangle$, are said to be orthogonal if they have a vanishing scalar product:

$$\langle\psi \mid \phi\rangle = 0. \tag{2.37}$$

- **Orthonormal states**

  Two *kets*, $| \psi\rangle$ and $| \phi\rangle$, are said to be orthonormal if they are orthogonal and if each one of them has a unit norm:

$$\langle\psi \mid \phi\rangle = 0, \qquad \langle\psi \mid \psi\rangle = 1, \qquad \langle\phi \mid \phi\rangle = 1. \tag{2.38}$$

- **Forbidden quantities**

  If $| \psi\rangle$ and $| \phi\rangle$ belong to the same vector (Hilbert) space, products of the type $| \psi\rangle \, | \phi\rangle$ and $\langle\psi \mid \langle\phi \mid$ are forbidden. They are nonsensical, since $| \psi\rangle \, | \phi\rangle$ and $\langle\psi \mid \langle\phi \mid$ are neither kets nor bras (an explicit illustration of this will be carried out in the example below and later on when we discuss the representation in a discrete basis). If $| \psi\rangle$ and $| \phi\rangle$ belong, however, to different vector spaces (e.g., $| \psi\rangle$ belongs to a spin space and $| \phi\rangle$ to an orbital angular momentum space), then the product $| \psi\rangle \, | \phi\rangle$, written as $| \psi\rangle \otimes | \phi\rangle$, represents a tensor product of $| \psi\rangle$ and $| \phi\rangle$. Only in these typical cases are such products meaningful.

Example 2.3

(**Note:** We will see later in this chapter that kets are represented by column matrices and bras by row matrices; this example is offered earlier than it should because we need to show some concrete illustrations of the formalism.) Consider the following two kets:

$$| \psi\rangle = \begin{pmatrix} -3i \\ 2+i \\ 4 \end{pmatrix}, \qquad | \phi\rangle = \begin{pmatrix} 2 \\ -i \\ 2-3i \end{pmatrix}.$$

(a) Find the bra $\langle\phi \mid$.
(b) Evaluate the scalar product $\langle\phi \mid \psi\rangle$.
(c) Examine why the products $| \psi\rangle \, | \phi\rangle$ and $\langle\phi \mid \langle\psi \mid$ do not make sense.

Solution

(a) As will be explained later when we introduce the Hermitian adjoint of kets and bras, we want to mention that the bra $\langle\phi \mid$ can be obtained by simply taking the complex conjugate of the transpose of the ket $| \phi\rangle$:

$$\langle\phi \mid = (2 \quad i \quad 2+3i). \tag{2.39}$$

(b) The scalar product $\langle\phi \mid \psi\rangle$ can be calculated as follows:

$$\begin{aligned} \langle\phi \mid \psi\rangle &= (2 \quad i \quad 2+3i) \begin{pmatrix} -3i \\ 2+i \\ 4 \end{pmatrix} \\ &= 2(-3i) + i(2+i) + 4(2+3i) \\ &= 7+8i. \end{aligned} \tag{2.40}$$

(c) First, the product $| \psi\rangle \, | \phi\rangle$ cannot be performed because, from linear algebra, the product of two column matrices cannot be performed. Similarly, since two row matrices cannot be multiplied, the product $\langle\phi \mid \langle\psi \mid$ is meaningless.

Physical meaning of the scalar product
The scalar product can be interpreted in two ways. First, by analogy with the scalar product of ordinary vectors in the Euclidean space, where $\vec{A}\cdot\vec{B}$ represents the projection of $\vec{B}$ on $\vec{A}$, the product $\langle\phi \mid \psi\rangle$ also represents the projection of $\mid \psi\rangle$ onto $\mid \phi\rangle$. Second, in the case of normalized states and according to Born's probabilistic interpretation, the quantity $\langle\phi \mid \psi\rangle$ represents the probability amplitude that the system's state $\mid \psi\rangle$ will, after a measurement is performed on the system, be found to be in another state $\mid \phi\rangle$.

Example 2.4 (Bra-ket algebra)
Consider the states $\mid \psi\rangle = 3i \mid \phi_1\rangle - 7i \mid \phi_2\rangle$ and $\mid \chi\rangle = -\mid \phi_1\rangle + 2i \mid \phi_2\rangle$, where $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ are orthonormal.

(a) Calculate $\mid \psi + \chi\rangle$ and $\langle \psi + \chi \mid$.

(b) Calculate the scalar products $\langle\psi \mid \chi\rangle$ and $\langle\chi \mid \psi\rangle$. Are they equal?

(c) Show that the states $\mid \psi\rangle$ and $\mid \chi\rangle$ satisfy the Schwarz inequality.

(d) Show that the states $\mid \psi\rangle$ and $\mid \chi\rangle$ satisfy the triangle inequality.

Solution

(a) The calculation of $\mid \psi + \chi\rangle$ is straightforward:

$$\begin{aligned} \mid \psi + \chi\rangle &= \mid \psi\rangle + \mid \chi\rangle = (3i \mid \phi_1\rangle - 7i \mid \phi_2\rangle) + (-\mid \phi_1\rangle + 2i \mid \phi_2\rangle) \\ &= (-1+3i) \mid \phi_1\rangle - 5i \mid \phi_2\rangle. \end{aligned} \tag{2.41}$$

This leads at once to the expression of $\langle \psi + \chi \mid$:

$$\langle \psi + \chi \mid = (-1+3i)^* \langle\phi_1 \mid + (-5i)^* \langle\phi_2 \mid = (-1-3i)\langle\phi_1 \mid + 5i\langle\phi_2 \mid . \tag{2.42}$$

(b) Since $\langle\phi_1 \mid \phi_1\rangle = \langle\phi_2 \mid \phi_2\rangle = 1$, $\langle\phi_1 \mid \phi_2\rangle = \langle\phi_2 \mid \phi_1\rangle = 0$, and since the bras corresponding to the kets $\mid \psi\rangle = 3i \mid \phi_1\rangle - 7i \mid \phi_2\rangle$ and $\mid \chi\rangle = -\mid \phi_1\rangle + 2i \mid \phi_2\rangle$ are given by $\langle\psi \mid = -3i\langle\phi_1 \mid + 7i\langle\phi_2 \mid$ and $\langle\chi \mid = -\langle\phi_1 \mid - 2i\langle\phi_2 \mid$, the scalar products are

$$\begin{aligned} \langle\psi \mid \chi\rangle &= (-3i\langle\phi_1 \mid + 7i\langle\phi_2 \mid)(-\mid \phi_1\rangle + 2i \mid \phi_2\rangle) \\ &= (-3i)(-1)\langle\phi_1 \mid \phi_1\rangle + (7i)(2i)\langle\phi_2 \mid \phi_2\rangle \\ &= -14 + 3i, \end{aligned} \tag{2.43}$$

$$\begin{aligned} \langle\chi \mid \psi\rangle &= (-\langle\phi_1 \mid - 2i\langle\phi_2 \mid)(3i \mid \phi_1\rangle - 7i \mid \phi_2\rangle) \\ &= (-1)(3i)\langle\phi_1 \mid \phi_1\rangle + (-2i)(-7i)\langle\phi_2 \mid \phi_2\rangle \\ &= -14 - 3i. \end{aligned} \tag{2.44}$$

We see that $\langle\psi \mid \chi\rangle$ is equal to the complex conjugate of $\langle\chi \mid \psi\rangle$.

(c) Let us first calculate $\langle\psi \mid \psi\rangle$ and $\langle\chi \mid \chi\rangle$:

$$\langle\psi \mid \psi\rangle = (-3i\langle\phi_1 \mid + 7i\langle\phi_2 \mid)(3i \mid \phi_1\rangle - 7i \mid \phi_2\rangle) = (-3i)(3i) + (7i)(-7i) = 58, \tag{2.45}$$

$$\langle\chi \mid \chi\rangle = (-\langle\phi_1 \mid - 2i\langle\phi_2 \mid)(-\mid \phi_1\rangle + 2i \mid \phi_2\rangle) = (-1)(-1) + (-2i)(2i) = 5. \tag{2.46}$$

Since $\langle\psi \mid \chi\rangle = -14 + 3i$ we have $\mid \langle\psi \mid \chi\rangle \mid^2 = 14^2 + 3^2 = 205$. Combining the values of $\mid \langle\psi \mid \chi\rangle \mid^2$, $\langle\psi \mid \psi\rangle$, and $\langle\chi \mid \chi\rangle$, we see that the Schwarz inequality (2.34) is satisfied:

$$205 < (58)(5) \Longrightarrow \mid \langle\psi \mid \chi\rangle \mid^2 < \langle\psi \mid \psi\rangle\langle\chi \mid \chi\rangle. \tag{2.47}$$

(d) First, let us use (2.41) and (2.42) to calculate $\langle\psi+\chi \mid \psi+\chi\rangle$:

$$\begin{aligned}\langle\psi+\chi \mid \psi+\chi\rangle &= [(-1-3i)\langle\phi_1 \mid +5i\langle\phi_2 \mid][(-1+3i)\mid\phi_1\rangle - 5i \mid \phi_2\rangle] \\ &= (-1-3i)(-1+3i)+(5i)(-5i) \\ &= 35. \qquad (2.48)\end{aligned}$$

Since $\langle\psi \mid \psi\rangle = 58$ and $\langle\chi \mid \chi\rangle = 5$, we infer that the triangle inequality (2.36) is satisfied:

$$\sqrt{35} < \sqrt{58}+\sqrt{5} \Longrightarrow \sqrt{\langle\psi+\chi \mid \psi+\chi\rangle} < \sqrt{\langle\psi \mid \psi\rangle}+\sqrt{\langle\chi \mid \chi\rangle}. \qquad (2.49)$$

Example 2.5
Consider two states $|\psi_1\rangle = 2i|\phi_1\rangle+|\phi_2\rangle-a|\phi_3\rangle+4|\phi_4\rangle$ and $|\psi_2\rangle = 3|\phi_1\rangle-i|\phi_2\rangle+5|\phi_3\rangle-|\phi_4\rangle$, where $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$, and $|\phi_4\rangle$ are orthonormal kets, and where a is a constant. Find the value of a so that $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal.

Solution
For the states $|\psi_1\rangle$ and $|\psi_2\rangle$ to be orthogonal, the scalar product $\langle\psi_2 \mid \psi_1\rangle$ must be zero. Using the relation $\langle\psi_2 \mid = 3\langle\phi_1| + i\langle\phi_2| + 5\langle\phi_3| - \langle\phi_4|$, we can easily find the scalar product

$$\begin{aligned}\langle\psi_2 \mid \psi_1\rangle &= (3\langle\phi_1| + i\langle\phi_2| + 5\langle\phi_3| - \langle\phi_4|)\,(2i|\phi_1\rangle + |\phi_2\rangle - a|\phi_3\rangle + 4|\phi_4\rangle) \\ &= 7i - 5a - 4. \qquad (2.50)\end{aligned}$$

Since $\langle\psi_2 \mid \psi_1\rangle = 7i - 5a - 4 = 0$, the value of a is $a = (7i-4)/5$.

2.4 Operators

2.4.1 General Definitions

Definition of an operator: An operator[1] $\hat{A}$ is a *mathematical rule* that when applied to a ket $\mid\psi\rangle$ transforms it into another ket $\mid\psi'\rangle$ of the same space and when it acts on a bra $\langle\phi\mid$ transforms it into another bra $\langle\phi'\mid$:

$$\hat{A}\mid\psi\rangle = \mid\psi'\rangle, \qquad \langle\phi\mid\hat{A} = \langle\phi'\mid. \qquad (2.51)$$

A similar definition applies to wave functions:

$$\hat{A}\psi(\vec{r}) = \psi'(\vec{r}), \qquad \phi(\vec{r})\hat{A} = \phi'(\vec{r}). \qquad (2.52)$$

Examples of operators
Here are some of the operators that we will use in this text:

- Unity operator: it leaves any ket unchanged, $\hat{I}\mid\psi\rangle = \mid\psi\rangle$.
- The gradient operator: $\vec{\nabla}\psi(\vec{r}) = (\partial\psi(\vec{r})/\partial x)\vec{i} + (\partial\psi(\vec{r})/\partial y)\vec{j} + (\partial\psi(\vec{r})/\partial z)\vec{k}$.

[1] The hat on $\hat{A}$ will be used throughout this text to distinguish an operator $\hat{A}$ from a complex number or a matrix A.

- The linear momentum operator: $\vec{P}\psi(\vec{r}) = -i\hbar\vec{\nabla}\psi(\vec{r})$.
- The Laplacian operator: $\nabla^2\psi(\vec{r}) = \partial^2\psi(\vec{r})/\partial x^2 + \partial^2\psi(\vec{r})/\partial y^2 + \partial^2\psi(\vec{r})/\partial z^2$.
- The parity operator: $\hat{\mathcal{P}}\psi(\vec{r}) = \psi(-\vec{r})$.

Products of operators
The product of two operators is generally not commutative:

$$\hat{A}\hat{B} \neq \hat{B}\hat{A}. \tag{2.53}$$

The product of operators is, however, associative:

$$\hat{A}\hat{B}\hat{C} = \hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C}. \tag{2.54}$$

We may also write $\hat{A}^n\hat{A}^m = \hat{A}^{n+m}$. When the product $\hat{A}\hat{B}$ operates on a ket $\mid \psi\rangle$ (the order of application is important), the operator $\hat{B}$ acts first on $\mid \psi\rangle$ and then $\hat{A}$ acts on the new ket $(B \mid \psi\rangle)$:

$$\hat{A}\hat{B} \mid \psi\rangle = \hat{A}(\hat{B} \mid \psi\rangle). \tag{2.55}$$

Similarly, when $\hat{A}\hat{B}\hat{C}\hat{D}$ operates on a ket $\mid \psi\rangle$, $\hat{D}$ acts first, then $\hat{C}$, then $\hat{B}$, and then $\hat{A}$.

When an operator $\hat{A}$ is sandwiched between a bra $\langle\phi \mid$ and a ket $\mid \psi\rangle$, it yields in general a complex number: $\langle\phi \mid \hat{A} \mid \psi\rangle$ = complex number. The quantity $\langle\phi \mid \hat{A} \mid \psi\rangle$ can also be a purely real or a purely imaginary number. **Note:** In evaluating $\langle\phi \mid \hat{A} \mid \psi\rangle$ it does not matter if one first applies $\hat{A}$ to the ket and then takes the bra-ket or one first applies $\hat{A}$ to the bra and then takes the bra-ket; that is $(\langle\phi \mid \hat{A}) \mid \psi\rangle = \langle\phi \mid (\hat{A} \mid \psi\rangle)$.

Linear operators
An operator $\hat{A}$ is said to be *linear* if it obeys the distributive law and, like all operators, it commutes with constants. That is, an operator $\hat{A}$ is linear if, for any vectors $\mid \psi_1\rangle$ and $\mid \psi_2\rangle$ and any complex numbers a_1 and a_2, we have

$$\hat{A}\left(a_1 \mid \psi_1\rangle + a_2 \mid \psi_2\rangle\right) = a_1\hat{A} \mid \psi_1\rangle + a_2\hat{A} \mid \psi_2\rangle, \tag{2.56}$$

and

$$(\langle\psi_1 \mid a_1 + \langle\psi_2 \mid a_2)\hat{A} = a_1\langle\psi_1 \mid \hat{A} + a_2\langle\psi_2 \mid \hat{A}. \tag{2.57}$$

Remarks

- The *expectation* or *mean* value $\langle\hat{A}\rangle$ of an operator $\hat{A}$ with respect to a state $\mid \psi\rangle$ is defined by
$$\langle\hat{A}\rangle = \frac{\langle\psi \mid \hat{A} \mid \psi\rangle}{\langle\psi \mid \psi\rangle}. \tag{2.58}$$
- The quantity $\mid \phi\rangle\langle\psi \mid$ (i.e., the product of a ket with a bra) is a linear operator in Dirac's notation. To see this, when $\mid \phi\rangle\langle\psi \mid$ is applied to a ket $\mid \psi'\rangle$, we obtain another ket:
$$\mid \phi\rangle\langle\psi \mid \psi'\rangle = \langle\psi \mid \psi'\rangle \mid \phi\rangle, \tag{2.59}$$
since $\langle\psi \mid \psi'\rangle$ is a complex number.
- Products of the type $\mid \psi\rangle\hat{A}$ and $\hat{A}\langle\psi \mid$ (i.e., when an operator stands on the right of a ket or on the left of a bra) are forbidden. They are not operators, or kets, or bras; they have no mathematical or physical meanings (see equation (2.219) for an illustration).

2.4.2 Hermitian Adjoint

The Hermitian adjoint or conjugate[2], $\alpha^\dagger$, of a complex number α is the complex conjugate of this number: $\alpha^\dagger = \alpha^*$. The Hermitian adjoint, or simply the adjoint, $\hat{A}^\dagger$, of an operator $\hat{A}$ is defined by this relation:

$$\langle \psi \mid \hat{A}^\dagger \mid \phi \rangle = \langle \phi \mid \hat{A} \mid \psi \rangle^*. \tag{2.60}$$

Properties of the Hermitian conjugate rule
To obtain the Hermitian adjoint of any expression, we must cyclically reverse the order of the factors and make three replacements:

- Replace constants by their complex conjugates: $\alpha^\dagger = \alpha^*$.
- Replace kets (bras) by the corresponding bras (kets): $(\mid \psi\rangle)^\dagger = \langle \psi \mid$ and $(\langle \psi \mid)^\dagger = \mid \psi\rangle$.
- Replace operators by their adjoints.

Following these rules, we can write

$$\begin{aligned}
(\hat{A}^\dagger)^\dagger &= \hat{A}, & (2.61)\\
(a\hat{A})^\dagger &= a^*\hat{A}^\dagger, & (2.62)\\
(\hat{A}^n)^\dagger &= (\hat{A}^\dagger)^n, & (2.63)\\
(\hat{A}+\hat{B}+\hat{C}+\hat{D})^\dagger &= \hat{A}^\dagger+\hat{B}^\dagger+\hat{C}^\dagger+\hat{D}^\dagger, & (2.64)\\
(\hat{A}\hat{B}\hat{C}\hat{D})^\dagger &= \hat{D}^\dagger\hat{C}^\dagger\hat{B}^\dagger\hat{A}^\dagger, & (2.65)\\
(\hat{A}\hat{B}\hat{C}\hat{D} \mid \psi\rangle)^\dagger &= \langle \psi \mid D^\dagger C^\dagger B^\dagger A^\dagger. & (2.66)
\end{aligned}$$

The Hermitian adjoint of the operator $\mid \psi\rangle\langle\phi \mid$ is given by

$$(\mid \psi\rangle\langle \phi \mid)^\dagger = \mid \phi\rangle\langle \psi \mid . \tag{2.67}$$

Operators act inside kets and bras, respectively, as follows:

$$\mid \alpha\hat{A}\psi\rangle = \alpha\hat{A} \mid \psi\rangle, \qquad \langle \alpha\hat{A}\psi \mid = \alpha^*\langle\psi \mid \hat{A}^\dagger. \tag{2.68}$$

Note also that $\langle \alpha\hat{A}^\dagger\psi \mid = \alpha^*\langle\psi \mid (\hat{A}^\dagger)^\dagger = \alpha^*\langle\psi \mid \hat{A}$. Hence, we can also write:

$$\langle \psi \mid \hat{A} \mid \phi\rangle = \langle \hat{A}^\dagger\psi \mid \phi\rangle = \langle \psi \mid \hat{A}\phi\rangle. \tag{2.69}$$

Hermitian and skew-Hermitian operators
An operator $\hat{A}$ is said to be *Hermitian* if it is equal to its adjoint $\hat{A}^\dagger$:

$$\boxed{\hat{A} = \hat{A}^\dagger \quad \text{or} \quad \langle \psi \mid \hat{A} \mid \phi\rangle = \langle \phi \mid \hat{A} \mid \psi\rangle^*.} \tag{2.70}$$

[2] The terms "adjoint" and "conjugate" are used indiscriminately.

On the other hand, an operator $\hat{B}$ is said to be *skew-Hermitian* or *anti-Hermitian* if

$$\boxed{\hat{B}^\dagger = -\hat{B} \qquad \text{or} \quad \langle\psi \mid \hat{B} \mid \phi\rangle = -\langle\phi \mid \hat{B} \mid \psi\rangle^*.} \tag{2.71}$$

Remark

The Hermitian adjoint of an operator is not, in general, equal to its complex conjugate: $\hat{A}^\dagger \neq \hat{A}^*$.

Example 2.6

(a) Discuss the hermiticity of the operators $(\hat{A} + \hat{A}^\dagger)$, $i(\hat{A} + \hat{A}^\dagger)$, and $i(\hat{A} - \hat{A}^\dagger)$.

(b) Find the Hermitian adjoint of $f(\hat{A}) = (1 + i\hat{A} + 3\hat{A}^2)(1 - 2i\hat{A} - 9\hat{A}^2)/(5 + 7\hat{A})$.

(c) Show that the expectation value of a Hermitian operator is real and that of an anti-Hermitian operator is imaginary.

Solution

(a) The operator $\hat{B} = \hat{A} + \hat{A}^\dagger$ is Hermitian regardless of whether or not $\hat{A}$ is Hermitian, since

$$\hat{B}^\dagger = (\hat{A} + \hat{A}^\dagger)^\dagger = \hat{A}^\dagger + \hat{A} = \hat{B}. \tag{2.72}$$

Similarly, the operator $i(\hat{A} - \hat{A}^\dagger)$ is also Hermitian; but $i(\hat{A} + \hat{A}^\dagger)$ is anti-Hermitian, since $[i(\hat{A} + \hat{A}^\dagger)]^\dagger = -i(\hat{A} + \hat{A}^\dagger)$.

(b) Since the Hermitian adjoint of an operator function $f(\hat{A})$ is given by $f^\dagger(\hat{A}) = f^*(\hat{A}^\dagger)$, we can write

$$\left(\frac{(1 + i\hat{A} + 3\hat{A}^2)(1 - 2i\hat{A} - 9\hat{A}^2)}{5 + 7\hat{A}}\right)^\dagger = \frac{(1 + 2i\hat{A}^\dagger - 9\hat{A}^{\dagger 2})(1 - i\hat{A}^\dagger + 3\hat{A}^{\dagger 2})}{5 + 7\hat{A}^\dagger}. \tag{2.73}$$

(c) From (2.70) we immediately infer that the expectation value of a Hermitian operator is real, for it satisfies the following property:

$$\langle\psi \mid \hat{A} \mid \psi\rangle = \langle\psi \mid \hat{A} \mid \psi\rangle^*; \tag{2.74}$$

that is, if $\hat{A}^\dagger = \hat{A}$ then $\langle\psi \mid \hat{A} \mid \psi\rangle$ is real. Similarly, for an anti-Hermitian operator, $\hat{B}^\dagger = -\hat{B}$, we have

$$\langle\psi \mid \hat{B} \mid \psi\rangle = -\langle\psi \mid \hat{B} \mid \psi\rangle^*, \tag{2.75}$$

which means that $\langle\psi \mid \hat{B} \mid \psi\rangle$ is a purely imaginary number.

2.4.3 Projection Operators

An operator $\hat{P}$ is said to be a *projection operator* if it is Hermitian and equal to its own square:

$$\hat{P}^\dagger = \hat{P}, \qquad \hat{P}^2 = \hat{P}. \tag{2.76}$$

The unit operator $\hat{I}$ is a simple example of a projection operator, since $\hat{I}^\dagger = \hat{I}$, $\quad \hat{I}^2 = \hat{I}$.

Properties of projection operators

- The product of two commuting projection operators, $\hat{P}_1$ and $\hat{P}_2$, is also a projection operator, since

$$(\hat{P}_1\hat{P}_2)^\dagger = \hat{P}_2^\dagger\hat{P}_1^\dagger = \hat{P}_2\hat{P}_1 = \hat{P}_1\hat{P}_2 \text{ and } (\hat{P}_1\hat{P}_2)^2 = \hat{P}_1\hat{P}_2\hat{P}_1\hat{P}_2 = \hat{P}_1^2\hat{P}_2^2 = \hat{P}_1\hat{P}_2. \tag{2.77}$$

- The sum of two projection operators is generally not a projection operator.
- Two projection operators are said to be orthogonal if their product is zero.
- For a sum of projection operators $\hat{P}_1 + \hat{P}_2 + \hat{P}_3 + \cdots$ to be a projection operator, it is necessary and sufficient that these projection operators be mutually orthogonal (i.e., the cross-product terms must vanish).

Example 2.7
Show that the operator $|\psi\rangle\langle\psi|$ is a projection operator only when $|\psi\rangle$ is normalized.

Solution
It is easy to ascertain that the operator $|\psi\rangle\langle\psi|$ is Hermitian, since $(|\psi\rangle\langle\psi|)^\dagger = |\psi\rangle\langle\psi|$. As for the square of this operator, it is given by

$$(|\psi\rangle\langle\psi|)^2 = (|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|\psi\rangle\langle\psi|. \tag{2.78}$$

Thus, if $|\psi\rangle$ is normalized, we have $(|\psi\rangle\langle\psi|)^2 = |\psi\rangle\langle\psi|$. In sum, if the state $|\psi\rangle$ is normalized, the product of the ket $|\psi\rangle$ with the bra $\langle\psi|$ is a projection operator.

2.4.4 Commutator Algebra

The *commutator* of two operators $\hat{A}$ and $\hat{B}$, denoted by $[\hat{A},\ \hat{B}]$, is defined by

$$\boxed{[\hat{A},\ \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A},} \tag{2.79}$$

and the *anticommutator* $\{\hat{A},\ \hat{B}\}$ is defined by

$$\boxed{\{\hat{A},\ \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}.} \tag{2.80}$$

Two operators are said to commute if their commutator is equal to zero and hence $\hat{A}\hat{B} = \hat{B}\hat{A}$. Any operator commutes with itself:

$$[\hat{A},\ \hat{A}] = 0. \tag{2.81}$$

Note that if two operators are Hermitian and their product is also Hermitian, these operators commute:

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger = \hat{B}\hat{A}, \tag{2.82}$$

and since $(\hat{A}\hat{B})^\dagger = \hat{A}\hat{B}$ we have $\hat{A}\hat{B} = \hat{B}\hat{A}$.

As an example, we may mention the commutators involving the x-position operator, $\hat{X}$, and the x-component of the momentum operator, $\hat{P}_x = -i\hbar\partial/\partial x$, as well as the y and the z components

$$\boxed{[\hat{X},\ \hat{P}_x] = i\hbar\hat{I}, \qquad [\hat{Y},\ \hat{P}_y] = i\hbar\hat{I}, \qquad [\hat{Z},\ \hat{P}_z] = i\hbar\hat{I},} \tag{2.83}$$

where $\hat{I}$ is the unit operator.

Properties of commutators

Using the commutator relation (2.79), we can establish the following properties:

- Antisymmetry:

$$[\hat{A},\ \hat{B}] = -[\hat{B},\ \hat{A}] \tag{2.84}$$

- Linearity:

$$[\hat{A},\ \hat{B} + \hat{C} + \hat{D} + \cdots] = [\hat{A},\ \hat{B}] + [\hat{A},\ \hat{C}] + [\hat{A},\ \hat{D}] + \cdots \tag{2.85}$$

- Hermitian conjugate of a commutator:

$$[\hat{A},\ \hat{B}]^\dagger = [\hat{B}^\dagger,\ \hat{A}^\dagger] \tag{2.86}$$

- Distributivity:

$$[\hat{A},\ \hat{B}\hat{C}] = [\hat{A},\ \hat{B}]\hat{C} + \hat{B}[\hat{A},\ \hat{C}] \tag{2.87}$$

$$[\hat{A}\hat{B},\ \hat{C}] = \hat{A}[\hat{B},\ \hat{C}] + [\hat{A},\ \hat{C}]\hat{B} \tag{2.88}$$

- Jacobi identity:

$$[\hat{A}, [\hat{B},\ \hat{C}]] + [\hat{B},\ [\hat{C},\ \hat{A}]] + [\hat{C},\ [\hat{A},\ \hat{B}]] = 0 \tag{2.89}$$

- By repeated applications of (2.87), we can show that

$$[\hat{A},\ \hat{B}^n] = \sum_{j=0}^{n-1} \hat{B}^j [\hat{A},\ \hat{B}] \hat{B}^{n-j-1} \tag{2.90}$$

$$[\hat{A}^n,\ \hat{B}] = \sum_{j=0}^{n-1} \hat{A}^{n-j-1} [\hat{A},\ \hat{B}] \hat{A}^j \tag{2.91}$$

- Operators commute with scalars: an operator $\hat{A}$ commutes with any scalar b:

$$[\hat{A},\ b] = 0 \tag{2.92}$$

Example 2.8

(a) Show that the commutator of two Hermitian operators is anti-Hermitian.
(b) Evaluate the commutator $[\hat{A},\ [\hat{B},\ \hat{C}]\hat{D}]$.

Solution

(a) If $\hat{A}$ and $\hat{B}$ are Hermitian, we can write

$$[\hat{A},\ \hat{B}]^\dagger = (\hat{A}\hat{B} - \hat{B}\hat{A})^\dagger = \hat{B}^\dagger\hat{A}^\dagger - \hat{A}^\dagger\hat{B}^\dagger = \hat{B}\hat{A} - \hat{A}\hat{B} = -[\hat{A},\ \hat{B}]; \tag{2.93}$$

that is, the commutator of $\hat{A}$ and $\hat{B}$ is anti-Hermitian: $[\hat{A},\ \hat{B}]^\dagger = -[\hat{A},\ \hat{B}]$.

(b) Using the distributivity relation (2.87), we have

$$\begin{aligned}
[\hat{A},\ [\hat{B},\ \hat{C}]\hat{D}] &= [\hat{B},\ \hat{C}][\hat{A},\ \hat{D}] + [\hat{A},\ [\hat{B},\ \hat{C}]]\hat{D} \\
&= (\hat{B}\hat{C} - \hat{C}\hat{B})(\hat{A}\hat{D} - \hat{D}\hat{A}) + \hat{A}(\hat{B}\hat{C} - \hat{C}\hat{B})\hat{D} - (\hat{B}\hat{C} - \hat{C}\hat{B})\hat{A}\hat{D} \\
&= \hat{C}\hat{B}\hat{D}\hat{A} - \hat{B}\hat{C}\hat{D}\hat{A} + \hat{A}\hat{B}\hat{C}\hat{D} - \hat{A}\hat{C}\hat{B}\hat{D}.
\end{aligned} \tag{2.94}$$

2.4.5 Uncertainty Relation between Two Operators

An interesting application of the commutator algebra is to derive a general relation giving the uncertainties product of two operators, $\hat{A}$ and $\hat{B}$. In particular, we want to give a formal derivation of Heisenberg's uncertainty relations.

Let $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ denote the expectation values of two Hermitian operators $\hat{A}$ and $\hat{B}$ with respect to a normalized state vector $\mid\psi\rangle$: $\langle\hat{A}\rangle = \langle\psi\mid\hat{A}\mid\psi\rangle$ and $\langle\hat{B}\rangle = \langle\psi\mid\hat{B}\mid\psi\rangle$. Introducing the operators $\Delta\hat{A}$ and $\Delta\hat{B}$,

$$\Delta\hat{A} = \hat{A} - \langle\hat{A}\rangle, \qquad \Delta\hat{B} = \hat{B} - \langle\hat{B}\rangle, \tag{2.95}$$

we have $(\Delta\hat{A})^2 = \hat{A}^2 - 2\hat{A}\langle\hat{A}\rangle + \langle\hat{A}\rangle^2$ and $(\Delta\hat{B})^2 = \hat{B}^2 - 2\hat{B}\langle\hat{B}\rangle + \langle\hat{B}\rangle^2$, and hence

$$\langle\psi\mid(\Delta\hat{A})^2\mid\psi\rangle = \langle(\Delta\hat{A})^2\rangle = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2, \qquad \langle(\Delta\hat{B})^2\rangle = \langle\hat{B}^2\rangle - \langle\hat{B}\rangle^2, \tag{2.96}$$

where $\langle\hat{A}^2\rangle = \langle\psi\mid\hat{A}^2\mid\psi\rangle$ and $\langle\hat{B}^2\rangle = \langle\psi\mid\hat{B}^2\mid\psi\rangle$. The *uncertainties* ΔA and ΔB are defined by

$$\boxed{\Delta A = \sqrt{\langle(\Delta\hat{A})^2\rangle} = \sqrt{\langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2}, \qquad \Delta B = \sqrt{\langle(\Delta\hat{B})^2\rangle} = \sqrt{\langle\hat{B}^2\rangle - \langle\hat{B}\rangle^2}.} \tag{2.97}$$

Let us write the action of the operators (2.95) on any state $\mid\psi\rangle$ as follows:

$$\mid\chi\rangle = \Delta\hat{A}\mid\psi\rangle = \left(\hat{A} - \langle\hat{A}\rangle\right)\mid\psi\rangle, \qquad \mid\phi\rangle = \Delta\hat{B}\mid\psi\rangle = \left(\hat{B} - \langle\hat{B}\rangle\right)\mid\psi\rangle. \tag{2.98}$$

The Schwarz inequality for the states $\mid\chi\rangle$ and $\mid\phi\rangle$ is given by

$$\langle\chi\mid\chi\rangle\langle\phi\mid\phi\rangle \ \geq\ |\langle\chi\mid\phi\rangle|^2. \tag{2.99}$$

Since $\hat{A}$ and $\hat{B}$ are Hermitian, $\Delta\hat{A}$ and $\Delta\hat{B}$ must also be Hermitian: $\Delta\hat{A}^\dagger = \hat{A}^\dagger - \langle\hat{A}\rangle = \hat{A} - \langle\hat{A}\rangle = \Delta\hat{A}$ and $\Delta\hat{B}^\dagger = \hat{B} - \langle\hat{B}\rangle = \Delta\hat{B}$. Thus, we can show the following three relations:

$$\langle\chi\mid\chi\rangle = \langle\psi\mid(\Delta\hat{A})^2\mid\psi\rangle, \quad \langle\phi\mid\phi\rangle = \langle\psi\mid(\Delta\hat{B})^2\mid\psi\rangle, \quad \langle\chi\mid\phi\rangle = \langle\psi\mid\Delta\hat{A}\Delta\hat{B}\mid\psi\rangle. \tag{2.100}$$

For instance, since $\Delta\hat{A}^\dagger = \Delta\hat{A}$ we have $\langle\chi \mid \chi\rangle = \langle\psi \mid \Delta\hat{A}^\dagger\Delta\hat{A} \mid \psi\rangle = \langle\psi \mid (\Delta\hat{A})^2 \mid \psi\rangle = \langle(\Delta\hat{A})^2\rangle$. Hence, the Schwarz inequality (2.99) becomes

$$\langle(\Delta\hat{A})^2\rangle\langle(\Delta\hat{B})^2\rangle \;\geq\; \left|\langle\Delta\hat{A}\Delta\hat{B}\rangle\right|^2. \tag{2.101}$$

Notice that the last term $\Delta\hat{A}\Delta\hat{B}$ of this equation can be written as

$$\Delta\hat{A}\Delta\hat{B} = \frac{1}{2}[\Delta\hat{A},\ \Delta\hat{B}] + \frac{1}{2}\{\Delta\hat{A},\ \Delta\hat{B}\} = \frac{1}{2}[\hat{A},\ \hat{B}] + \frac{1}{2}\{\Delta\hat{A},\ \Delta\hat{B}\}, \tag{2.102}$$

where we have used the fact that $[\Delta\hat{A},\ \Delta\hat{B}] = [\hat{A},\ \hat{B}]$. Since $[\hat{A},\ \hat{B}]$ is anti-Hermitian and $\{\Delta\hat{A},\ \Delta\hat{B}\}$ is Hermitian and since the expectation value of a Hermitian operator is real and that the expectation value of an anti-Hermitian operator is imaginary (see Example 2.6), the expectation value $\langle\Delta\hat{A}\Delta\hat{B}\rangle$ of (2.102) becomes equal to the sum of a real part $\langle\{\Delta\hat{A},\ \Delta\hat{B}\}\rangle/2$ and an imaginary part $\langle[\hat{A},\ \hat{B}]\rangle/2$; hence

$$\left|\langle\Delta\hat{A}\Delta\hat{B}\rangle\right|^2 = \frac{1}{4}\left|\langle[\hat{A},\ \hat{B}]\rangle\right|^2 + \frac{1}{4}\left|\langle\{\Delta\hat{A},\ \Delta\hat{B}\}\rangle\right|^2. \tag{2.103}$$

Since the last term is a positive real number, we can infer the following relation:

$$\left|\langle\Delta\hat{A}\Delta\hat{B}\rangle\right|^2 \;\geq\; \frac{1}{4}\left|\langle[\hat{A},\ \hat{B}]\rangle\right|^2. \tag{2.104}$$

Comparing equations (2.101) and (2.104), we conclude that

$$\langle(\Delta\hat{A})^2\rangle\langle(\Delta\hat{B})^2\rangle \;\geq\; \frac{1}{4}\left|\langle[\hat{A},\ \hat{B}]\rangle\right|^2, \tag{2.105}$$

which (by taking its square root) can be reduced to

$$\boxed{\Delta A\Delta B \geq \frac{1}{2}\left|\langle[\hat{A},\ \hat{B}]\rangle\right|.} \tag{2.106}$$

This uncertainty relation plays an important role in the formalism of quantum mechanics. Its application to position and momentum operators leads to the Heisenberg uncertainty relations, which represent one of the cornerstones of quantum mechanics; see the next example.

Example 2.9 (Heisenberg uncertainty relations)
Find the uncertainty relations between the components of the position and the momentum operators.

Solution
By applying (2.106) to the x-components of the position operator $\hat{X}$, and the momentum operator $\hat{P}_x$, we obtain $\Delta x\Delta p_x \geq \frac{1}{2} \mid \langle[\hat{X},\ \hat{P}_x]\rangle \mid$. But since $[\hat{X},\ \hat{P}_x] = i\hbar\hat{I}$, we have $\Delta x\Delta p_x \geq \hbar/2$; the uncertainty relations for the $y-$ and $z-$ components follow immediately:

$$\boxed{\Delta x\Delta p_x \geq \frac{\hbar}{2}, \qquad \Delta y\Delta p_y \geq \frac{\hbar}{2}, \qquad \Delta z\Delta p_z \geq \frac{\hbar}{2}.} \tag{2.107}$$

These are the Heisenberg uncertainty relations.

2.4.6 Functions of Operators

Let $F(\hat{A})$ be a function of an operator $\hat{A}$. If $\hat{A}$ is a linear operator, we can Taylor expand $F(\hat{A})$ in a power series of $\hat{A}$:

$$F(\hat{A}) = \sum_{n=0}^{\infty} a_n \hat{A}^n, \tag{2.108}$$

where a_n is just an expansion coefficient. As an illustration of an operator function, consider $e^{a\hat{A}}$, where a is a scalar which can be complex or real. We can expand it as follows:

$$e^{a\hat{A}} = \sum_{n=0}^{\infty} \frac{a^n}{n!}\hat{A}^n = \hat{I} + a\hat{A} + \frac{a^2}{2!}\hat{A}^2 + \frac{a^3}{3!}\hat{A}^3 + \cdots. \tag{2.109}$$

Commutators involving function operators

If $\hat{A}$ commutes with another operator $\hat{B}$, then $\hat{B}$ commutes with any operator function that depends on $\hat{A}$:

$$[\hat{A},\ \hat{B}] = 0 \quad \Longrightarrow \quad [\hat{B},\ F(\hat{A})] = 0; \tag{2.110}$$

in particular, $F(\hat{A})$ commutes with $\hat{A}$ and with any other function, $G(\hat{A})$, of $\hat{A}$:

$$[\hat{A},\ F(\hat{A})] = 0, \qquad [\hat{A}^n,\ F(\hat{A})] = 0, \qquad [F(\hat{A}),\ G(\hat{A})] = 0. \tag{2.111}$$

Hermitian adjoint of function operators

The adjoint of $F(\hat{A})$ is given by

$$[F(\hat{A})]^\dagger = F^*(\hat{A}^\dagger). \tag{2.112}$$

Note that if $\hat{A}$ is Hermitian, $F(\hat{A})$ is not necessarily Hermitian; $F(\hat{A})$ will be Hermitian only if F is a real function and $\hat{A}$ is Hermitian. An example is

$$(e^{\hat{A}})^\dagger = e^{\hat{A}^\dagger}, \qquad (e^{i\hat{A}})^\dagger = e^{-i\hat{A}^\dagger}, \qquad (e^{i\alpha\hat{A}})^\dagger = e^{-i\alpha^*\hat{A}^\dagger}, \tag{2.113}$$

where α is a complex number. So if $\hat{A}$ is Hermitian, an operator function which can be expanded as $F(\hat{A}) = \sum_{n=0}^{\infty} a_n \hat{A}^n$ will be Hermitian only if the expansion coefficients a_n are real numbers. But in general, $F(\hat{A})$ is not Hermitian even if $\hat{A}$ is Hermitian, since

$$F^*(\hat{A}^\dagger) = \sum_{n=0}^{\infty} a_n^* (\hat{A}^\dagger)^n. \tag{2.114}$$

Relations involving function operators

Note that

$$[\hat{A},\ \hat{B}] \neq 0 \quad \Longrightarrow \quad [\hat{B},\ F(\hat{A})] \neq 0; \tag{2.115}$$

in particular, $e^{\hat{A}}e^{\hat{B}} \neq e^{\hat{A}+\hat{B}}$. Using (2.109) we can ascertain that

$$e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}}e^{[\hat{A},\ \hat{B}]/2}, \tag{2.116}$$

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A},\ \hat{B}] + \frac{1}{2!}[\hat{A},\ [\hat{A},\ \hat{B}]] + \frac{1}{3!}[\hat{A},\ [\hat{A},\ [\hat{A},\ \hat{B}]]] + \cdots. \tag{2.117}$$

2.4.7 Inverse and Unitary Operators

Inverse of an operator: Assuming it exists[3] the *inverse* $\hat{A}^{-1}$ of a linear operator $\hat{A}$ is defined by the relation

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{I}, \tag{2.118}$$

where $\hat{I}$ is the unit operator, the operator that leaves any state $|\psi\rangle$ unchanged.

Quotient of two operators: Dividing an operator $\hat{A}$ by another operator $\hat{B}$ (provided that the inverse $\hat{B}^{-1}$ exists) is equivalent to multiplying $\hat{A}$ by $\hat{B}^{-1}$:

$$\frac{\hat{A}}{\hat{B}} = \hat{A}\hat{B}^{-1}. \tag{2.119}$$

The side on which the quotient is taken matters:

$$\frac{\hat{A}}{\hat{B}} = \hat{A}\frac{\hat{I}}{\hat{B}} = \hat{A}\hat{B}^{-1} \quad \text{and} \quad \frac{\hat{I}}{\hat{B}}\hat{A} = \hat{B}^{-1}\hat{A}. \tag{2.120}$$

In general, we have $\hat{A}\hat{B}^{-1} \neq \hat{B}^{-1}\hat{A}$. For an illustration of these ideas, see Problem 2.12. We may mention here the following properties about the inverse of operators:

$$\left(\hat{A}\hat{B}\hat{C}\hat{D}\right)^{-1} = \hat{D}^{-1}\hat{C}^{-1}\hat{B}^{-1}\hat{A}^{-1}, \qquad \left(\hat{A}^{n}\right)^{-1} = \left(\hat{A}^{-1}\right)^{n}. \tag{2.121}$$

Unitary operators: A linear operator $\hat{U}$ is said to be *unitary* if its inverse $\hat{U}^{-1}$ is equal to its adjoint $\hat{U}^{\dagger}$:

$$\hat{U}^{\dagger} = \hat{U}^{-1} \qquad \text{or} \qquad \hat{U}\hat{U}^{\dagger} = \hat{U}^{\dagger}\hat{U} = \hat{I}. \tag{2.122}$$

The product of two unitary operators is also unitary, since

$$(\hat{U}\hat{V})(\hat{U}\hat{V})^{\dagger} = (\hat{U}\hat{V})(\hat{V}^{\dagger}\hat{U}^{\dagger}) = \hat{U}(\hat{V}\hat{V}^{\dagger})\hat{U}^{\dagger} = \hat{U}\hat{U}^{\dagger} = \hat{I}, \tag{2.123}$$

or $(\hat{U}\hat{V})^{\dagger} = (\hat{U}\hat{V})^{-1}$. This result can be generalized to any number of operators; the product of a number of unitary operators is also unitary, since

$$\begin{aligned}(\hat{A}\hat{B}\hat{C}\hat{D}\cdots)(\hat{A}\hat{B}\hat{C}\hat{D}\cdots)^{\dagger} &= \hat{A}\hat{B}\hat{C}\hat{D}(\cdots)\hat{D}^{\dagger}\hat{C}^{\dagger}\hat{B}^{\dagger}\hat{A}^{\dagger} = \hat{A}\hat{B}\hat{C}(\hat{D}\hat{D}^{\dagger})\hat{C}^{\dagger}\hat{B}^{\dagger}\hat{A}^{\dagger} \\ &= \hat{A}\hat{B}(\hat{C}\hat{C}^{\dagger})\hat{B}^{\dagger}\hat{A}^{\dagger} = \hat{A}(\hat{B}\hat{B}^{\dagger})\hat{A}^{\dagger} \\ &= \hat{A}\hat{A}^{\dagger} = \hat{I}, \end{aligned} \tag{2.124}$$

or $(\hat{A}\hat{B}\hat{C}\hat{D}\cdots)^{\dagger} = (\hat{A}\hat{B}\hat{C}\hat{D}\cdots)^{-1}$.

Example 2.10 (Unitary operator)

What conditions must the parameter ε and the operator $\hat{G}$ satisfy so that the operator $\hat{U} = e^{i\varepsilon\hat{G}}$ is unitary?

[3] Not every operator has an inverse, just as in the case of matrices. The inverse of a matrix exists only when its determinant is nonzero.

Solution
Clearly, if ε is real and $\hat{G}$ is Hermitian, the operator $e^{i\varepsilon\hat{G}}$ would be unitary. Using the property $[F(\hat{A})]^\dagger = F^*(\hat{A}^\dagger)$, we see that

$$(e^{i\varepsilon\hat{G}})^\dagger = e^{-i\varepsilon\hat{G}} = (e^{i\varepsilon\hat{G}})^{-1}, \tag{2.125}$$

that is, $\hat{U}^\dagger = \hat{U}^{-1}$.

2.4.8 Eigenvalues and Eigenvectors of an Operator

Having studied the properties of operators and states, we are now ready to discuss how to find the eigenvalues and eigenvectors of an operator.

A state vector $\mid \psi\rangle$ is said to be an *eigenvector* (also called an eigenket or eigenstate) of an operator $\hat{A}$ if the application of $\hat{A}$ to $\mid \psi\rangle$ gives

$$\hat{A} \mid \psi\rangle \;=\; a \mid \psi\rangle, \tag{2.126}$$

where a is a complex number, called an *eigenvalue* of $\hat{A}$. This equation is known as the *eigenvalue equation*, or *eigenvalue problem*, of the operator $\hat{A}$. Its solutions yield the eigenvalues and eigenvectors of $\hat{A}$. In Section 2.5.3 we will see how to solve the eigenvalue problem in a discrete basis.

A simple example is the eigenvalue problem for the unity operator $\hat{I}$:

$$\hat{I} \mid \psi\rangle = \mid \psi\rangle. \tag{2.127}$$

This means that all vectors are eigenvectors of $\hat{I}$ with one eigenvalue, 1. Note that

$$\hat{A} \mid \psi\rangle \;=\; a \mid \psi\rangle \;\Longrightarrow\; \hat{A}^n \mid \psi\rangle \;=\; a^n \mid \psi\rangle \quad \text{and} \quad F(\hat{A}) \mid \psi\rangle = F(a) \mid \psi\rangle. \tag{2.128}$$

For instance, we have

$$\hat{A} \mid \psi\rangle \;=\; a \mid \psi\rangle \quad \Longrightarrow \quad e^{i\hat{A}} \mid \psi\rangle \;=\; e^{ia} \mid \psi\rangle. \tag{2.129}$$

Example 2.11 (Eigenvalues of the inverse of an operator)
Show that if $\hat{A}^{-1}$ exists, the eigenvalues of $\hat{A}^{-1}$ are just the inverses of those of $\hat{A}$.

Solution
Since $\hat{A}^{-1}\hat{A} = \hat{I}$ we have on the one hand

$$\hat{A}^{-1}\hat{A} \mid \psi\rangle \;= \mid \psi\rangle, \tag{2.130}$$

and on the other hand

$$\hat{A}^{-1}\hat{A} \mid \psi\rangle \;=\; \hat{A}^{-1}(\hat{A} \mid \psi\rangle) \;=\; a\hat{A}^{-1} \mid \psi\rangle. \tag{2.131}$$

Combining the previous two equations, we obtain

$$a\hat{A}^{-1} \mid \psi\rangle \;= \mid \psi\rangle, \tag{2.132}$$

hence

$$\hat{A}^{-1} \mid \psi\rangle = \frac{1}{a} \mid \psi\rangle. \tag{2.133}$$

This means that $\mid \psi\rangle$ is also an eigenvector of $\hat{A}^{-1}$ with eigenvalue $1/a$. That is, if $\hat{A}^{-1}$ exists, then

$$\boxed{\hat{A} \mid \psi\rangle = a \mid \psi\rangle \quad \Longrightarrow \quad \hat{A}^{-1} \mid \psi\rangle = \frac{1}{a} \mid \psi\rangle.} \tag{2.134}$$

Some useful theorems pertaining to the eigenvalue problem

Theorem 2.1 *For a Hermitian operator, all of its eigenvalues are real and the eigenvectors corresponding to different eigenvalues are orthogonal.*

$$\text{If } \hat{A}^\dagger = \hat{A}, \quad \hat{A} \mid \phi_n\rangle = a_n \mid \phi_n\rangle \quad \Longrightarrow \quad a_n = \text{real number, and } \langle\phi_m \mid \phi_n\rangle = \delta_{mn}. \tag{2.135}$$

Proof of Theorem 2.1
Note that

$$\hat{A} \mid \phi_n\rangle = a_n \mid \phi_n\rangle \quad \Longrightarrow \quad \langle\phi_m \mid \hat{A} \mid \phi_n\rangle = a_n\langle\phi_m \mid \phi_n\rangle, \tag{2.136}$$

and

$$\langle\phi_m \mid \hat{A}^\dagger = a_m^*\langle\phi_m \mid \quad \Longrightarrow \quad \langle\phi_m \mid \hat{A}^\dagger \mid \phi_n\rangle = a_m^*\langle\phi_m \mid \phi_n\rangle. \tag{2.137}$$

Subtracting (2.137) from (2.136) and using the fact that $\hat{A}$ is Hermitian, $\hat{A} = \hat{A}^\dagger$, we have

$$(a_n - a_m^*)\langle\phi_m \mid \phi_n\rangle = 0. \tag{2.138}$$

Two cases must be considered separately:

- Case $m = n$: since $\langle\phi_n \mid \phi_n\rangle > 0$, we must have $a_n = a_n^*$; hence the eigenvalues a_n must be real.
- Case $m \neq n$: since in general $a_n \neq a_m^*$, we must have $\langle\phi_m \mid \phi_n\rangle = 0$; that is, $\mid \phi_m\rangle$ and $\mid \phi_n\rangle$ must be orthogonal.

Theorem 2.2 *The eigenstates of a Hermitian operator define a complete set of mutually orthonormal basis states. The operator is diagonal in this eigenbasis with its diagonal elements equal to the eigenvalues. This basis set is unique if the operator has no degenerate eigenvalues and not unique (in fact it is infinite) if there is any degeneracy.*

Theorem 2.3 *If two Hermitian operators, $\hat{A}$ and $\hat{B}$, commute and if $\hat{A}$ has no degenerate eigenvalue, then each eigenvector of $\hat{A}$ is also an eigenvector of $\hat{B}$. In addition, we can construct a common orthonormal basis that is made of the joint eigenvectors of $\hat{A}$ and $\hat{B}$.*

Proof of Theorem 2.3
Since $\hat{A}$ is Hermitian with no degenerate eigenvalue, to each eigenvalue of $\hat{A}$ there corresponds only one eigenvector. Consider the equation

$$\hat{A} \mid \phi_n\rangle = a_n \mid \phi_n\rangle. \tag{2.139}$$

Since $\hat{A}$ commutes with $\hat{B}$ we can write

$$\hat{B}\hat{A} \mid \phi_n\rangle = \hat{A}\hat{B} \mid \phi_n\rangle \quad \text{or} \quad \hat{A}(\hat{B} \mid \phi_n\rangle) = a_n(\hat{B} \mid \phi_n\rangle); \tag{2.140}$$

that is, $(\hat{B} \mid \phi_n\rangle)$ is an eigenvector of $\hat{A}$ with eigenvalue a_n. But since this eigenvector is unique (apart from an arbitrary phase constant), the ket $\mid \phi_n\rangle$ must also be an eigenvector of $\hat{B}$:

$$\hat{B} \mid \phi_n\rangle = b_n \mid \phi_n\rangle. \tag{2.141}$$

Since each eigenvector of $\hat{A}$ is also an eigenvector of $\hat{B}$ (and vice versa), both of these operators must have a common basis. This basis is unique; it is made of the joint eigenvectors of $\hat{A}$ and $\hat{B}$. This theorem also holds for any number of mutually commuting Hermitian operators.

Now, if a_n is a degenerate eigenvalue, we can only say that $\hat{B} \mid \phi_n\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue a_n; $\mid \phi_n\rangle$ is not necessarily an eigenvector of $\hat{B}$. If one of the operators is degenerate, there exist an infinite number of orthonormal basis sets that are common to these two operators; that is, the joint basis does exist and it is not unique.

Theorem 2.4 *The eigenvalues of an anti-Hermitian operator are either purely imaginary or equal to zero.*

Theorem 2.5 *The eigenvalues of a unitary operator are complex numbers of moduli equal to one; the eigenvectors of a unitary operator that has no degenerate eigenvalues are mutually orthogonal.*

Proof of Theorem 2.5

Let $\mid \phi_n\rangle$ and $\mid \phi_m\rangle$ be eigenvectors to the unitary operator $\hat{U}$ with eigenvalues a_n and a_m, respectively. We can write

$$(\langle\phi_m \mid \hat{U}^\dagger)(\hat{U} \mid \phi_n\rangle) = a_m^* a_n \langle\phi_m \mid \phi_n\rangle. \tag{2.142}$$

Since $\hat{U}^\dagger\hat{U} = \hat{I}$ this equation can be rewritten as

$$(a_m^* a_n - 1)\langle\phi_m \mid \phi_n\rangle = 0, \tag{2.143}$$

which in turn leads to the following two cases:

- Case $n = m$: since $\langle\phi_n \mid \phi_n\rangle > 0$ then $a_n^* a_n = \mid a_n \mid^2 = 1$, and hence $\mid a_n \mid = 1$.
- Case $n \neq m$: the only possibility for this case is that $\mid \phi_m\rangle$ and $\mid \phi_n\rangle$ are orthogonal, $\langle\phi_m \mid \phi_n\rangle = 0$.

2.4.9 Infinitesimal and Finite Unitary Transformations

We want to study here how quantities such as kets, bras, operators, and scalars transform under unitary transformations. A unitary transformation is the application of a unitary operator $\hat{U}$ to one of these quantities.

2.4.9.1 Unitary Transformations

Kets $\mid \psi\rangle$ and bras $\langle\psi \mid$ transform as follows:

$$\mid \psi'\rangle = \hat{U} \mid \psi\rangle, \qquad \langle\psi' \mid = \langle\psi \mid \hat{U}^\dagger. \tag{2.144}$$

Let us now find out how operators transform under unitary transformations. Since the transform of $\hat{A} \mid \psi\rangle = \mid \phi\rangle$ is $\hat{A}' \mid \psi'\rangle = \mid \phi'\rangle$, we can rewrite $\hat{A}' \mid \psi'\rangle = \mid \phi'\rangle$ as $\hat{A}'\hat{U} \mid \psi\rangle = \hat{U} \mid \phi\rangle = \hat{U}\hat{A} \mid \psi\rangle$ which, in turn, leads to $\hat{A}'\hat{U} = \hat{U}\hat{A}$. Multiplying both sides of $\hat{A}'\hat{U} = \hat{U}\hat{A}$ by $\hat{U}^\dagger$ and since $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}$, we have

$$\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger, \qquad \hat{A} = \hat{U}^\dagger\hat{A}'\hat{U}. \tag{2.145}$$

The results reached in (2.144) and (2.145) may be summarized as follows:

$$\boxed{\mid \psi'\rangle = \hat{U} \mid \psi\rangle, \qquad \langle\psi' \mid = \langle\psi \mid \hat{U}^\dagger, \qquad \hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger,} \tag{2.146}$$

$$\boxed{\mid \psi\rangle = \hat{U}^\dagger \mid \psi'\rangle, \qquad \langle\psi \mid = \langle\psi' \mid \hat{U}, \qquad \hat{A} = \hat{U}^\dagger\hat{A}'\hat{U}.} \tag{2.147}$$

Properties of unitary transformations

- If an operator $\hat{A}$ is Hermitian, its transformed $\hat{A}'$ is also Hermitian, since

$$\hat{A}'^\dagger = (\hat{U}\hat{A}\hat{U}^\dagger)^\dagger = \hat{U}\hat{A}^\dagger\hat{U}^\dagger = \hat{U}\hat{A}\hat{U}^\dagger = \hat{A}'. \tag{2.148}$$

- The eigenvalues of $\hat{A}$ and those of its transformed $\hat{A}'$ are the same:

$$\hat{A} \mid \psi_n\rangle = a_n \mid \psi_n\rangle \quad \Longrightarrow \quad \hat{A}' \mid \psi'_n\rangle = a_n \mid \psi'_n\rangle, \tag{2.149}$$

  since

$$\begin{aligned} \hat{A}' \mid \psi'_n\rangle &= (\hat{U}\hat{A}\hat{U}^\dagger)(\hat{U} \mid \psi_n\rangle) = \hat{U}\hat{A}(\hat{U}^\dagger\hat{U}) \mid \psi_n\rangle \\ &= \hat{U}\hat{A} \mid \psi_n\rangle = a_n(\hat{U} \mid \psi_n\rangle) = a_n \mid \psi'_n\rangle. \end{aligned} \tag{2.150}$$

- Commutators that are equal to (complex) numbers remain unchanged under unitary transformations, since the transformation of $[\hat{A}, \hat{B}] = a$, where a is a complex number, is given by

$$\begin{aligned} [\hat{A}', \hat{B}'] &= [\hat{U}\hat{A}\hat{U}^\dagger, \hat{U}\hat{B}\hat{U}^\dagger] = (\hat{U}\hat{A}\hat{U}^\dagger)(\hat{U}\hat{B}\hat{U}^\dagger) - (\hat{U}\hat{B}\hat{U}^\dagger)(\hat{U}\hat{A}\hat{U}^\dagger) \\ &= \hat{U}[\hat{A}, \hat{B}]\hat{U}^\dagger = \hat{U}a\hat{U}^\dagger = a\hat{U}\hat{U}^\dagger = a \\ &= [\hat{A}, \hat{B}]. \end{aligned} \tag{2.151}$$

- We can also verify the following general relations:

$$\hat{A} = \beta\hat{B} + \gamma\hat{C} \quad \Longrightarrow \quad \hat{A}' = \beta\hat{B}' + \gamma\hat{C}', \tag{2.152}$$

$$\hat{A} = \alpha\hat{B}\hat{C}\hat{D} \quad \Longrightarrow \quad \hat{A}' = \alpha\hat{B}'\hat{C}'\hat{D}', \tag{2.153}$$

  where $\hat{A}'$, $\hat{B}'$, $\hat{C}'$, and $\hat{D}'$ are the transforms of $\hat{A}$, $\hat{B}$, $\hat{C}$, and $\hat{D}$, respectively.

- Since the result (2.151) is valid for any complex number, we can state that complex numbers, such as $\langle \psi \mid \hat{A} \mid \chi \rangle$, remain unchanged under unitary transformations, since

$$\langle \psi' \mid \hat{A}' \mid \chi' \rangle = (\langle \psi \mid \hat{U}^\dagger)(\hat{U}\hat{A}\hat{U}^\dagger)(\hat{U} \mid \chi \rangle) = \langle \psi \mid (\hat{U}^\dagger \hat{U})\hat{A}(\hat{U}^\dagger U) \mid \chi \rangle = \langle \psi \mid \hat{A} \mid \chi \rangle. \tag{2.154}$$

  Taking $\hat{A} = \hat{I}$ we see that scalar products of the type

$$\langle \psi' \mid \chi' \rangle = \langle \psi \mid \chi \rangle \tag{2.155}$$

  are invariant under unitary transformations; notably, the norm of a state vector is conserved:

$$\langle \psi' \mid \psi' \rangle = \langle \psi \mid \psi \rangle. \tag{2.156}$$

- We can also verify that $\left(\hat{U}\hat{A}\hat{U}^\dagger\right)^n = \hat{U}\hat{A}^n\hat{U}^\dagger$ since

$$\begin{aligned} \left(\hat{U}\hat{A}\hat{U}^\dagger\right)^n &= \left(\hat{U}\hat{A}\hat{U}^\dagger\right)\left(\hat{U}\hat{A}\hat{U}^\dagger\right)\cdots\left(\hat{U}\hat{A}\hat{U}^\dagger\right) = \hat{U}\hat{A}(\hat{U}^\dagger\hat{U})\hat{A}(\hat{U}^\dagger\hat{U})\cdots(\hat{U}^\dagger\hat{U})\hat{A}\hat{U}^\dagger \\ &= \hat{U}\hat{A}^n\hat{U}^\dagger. \end{aligned} \tag{2.157}$$

- We can generalize the previous result to obtain the transformation of any operator function $f(\hat{A})$:

$$\hat{U} f(\hat{A})\hat{U}^\dagger = f(\hat{U}\hat{A}\hat{U}^\dagger) = f(\hat{A}'), \tag{2.158}$$

  or more generally

$$\hat{U} f(\hat{A}, \hat{B}, \hat{C}, \cdots)\hat{U}^\dagger = f(\hat{U}\hat{A}\hat{U}^\dagger, \hat{U}\hat{B}\hat{U}^\dagger, \hat{U}\hat{C}\hat{U}^\dagger, \cdots) = f(\hat{A}', \hat{B}', \hat{C}', \cdots). \tag{2.159}$$

A unitary transformation does not change the physics of a system; it merely transforms one description of the system to another physically equivalent description.

In what follows we want to consider two types of unitary transformations: infinitesimal transformations and finite transformations.

2.4.9.2 Infinitesimal Unitary Transformations

Consider an operator $\hat{U}$ which depends on an infinitesimally small real parameter ε and which varies only slightly from the unity operator $\hat{I}$:

$$\hat{U}_\varepsilon(\hat{G}) = \hat{I} + i\varepsilon\hat{G}, \tag{2.160}$$

where $\hat{G}$ is called the *generator* of the infinitesimal transformation. Clearly, $\hat{U}_\varepsilon$ is a unitary transformation only when the parameter ε is real and $\hat{G}$ is Hermitian, since

$$\hat{U}_\varepsilon\hat{U}_\varepsilon^\dagger = (\hat{I} + i\varepsilon\hat{G})(\hat{I} - i\varepsilon\hat{G}^\dagger) \simeq \hat{I} + i\varepsilon(\hat{G} - \hat{G}^\dagger) = \hat{I}, \tag{2.161}$$

where we have neglected the quadratic terms in ε.

The transformation of a state vector $\mid \psi \rangle$ is

$$\mid \psi' \rangle = (\hat{I} + i\varepsilon\hat{G}) \mid \psi \rangle = \mid \psi \rangle + \delta \mid \psi \rangle, \tag{2.162}$$

where

$$\delta \mid \psi\rangle = i\varepsilon\hat{G} \mid \psi\rangle. \tag{2.163}$$

The transformation of an operator $\hat{A}$ is given by

$$\boxed{\hat{A}' = (\hat{I} + i\varepsilon\hat{G})\hat{A}(\hat{I} - i\varepsilon\hat{G}) \simeq \hat{A} + i\varepsilon[\hat{G}, \hat{A}].} \tag{2.164}$$

If $\hat{G}$ commutes with $\hat{A}$, the unitary transformation will leave $\hat{A}$ unchanged, $\hat{A}' = \hat{A}$:

$$[\hat{G}, \hat{A}] = 0 \quad \Longrightarrow \quad \hat{A}' = (\hat{I} + i\varepsilon\hat{G})\hat{A}(\hat{I} - i\varepsilon\hat{G}) = \hat{A}. \tag{2.165}$$

2.4.9.3 Finite Unitary Transformations

We can construct a *finite* unitary transformation from (2.160) by performing a succession of infinitesimal transformations in steps of ε; the application of a series of successive unitary transformations is equivalent to the application of a single unitary transformation. Denoting $\varepsilon = \alpha/N$, where N is an integer and α is a finite parameter, we can apply the same unitary transformation N times; in the limit $N \to +\infty$ we obtain

$$\hat{U}_\alpha(\hat{G}) = \lim_{N\to\infty} \prod_{k=1}^{N} \left(1 + i\frac{\alpha}{N}\hat{G}\right) = \lim_{N\to+\infty} \left(1 + i\frac{\alpha}{N}\hat{G}\right)^N = e^{i\alpha\hat{G}}, \tag{2.166}$$

where $\hat{G}$ is now the generator of the finite transformation and α is its parameter.

As shown in (2.125), $\hat{U}$ is unitary only when the parameter α is real and $\hat{G}$ is Hermitian, since

$$(e^{i\alpha\hat{G}})^\dagger = e^{-i\alpha\hat{G}} = (e^{i\alpha\hat{G}})^{-1}. \tag{2.167}$$

Using the commutation relation (2.117), we can write the transformation $\hat{A}'$ of an operator $\hat{A}$ as follows:

$$\boxed{e^{i\alpha\hat{G}}\hat{A}e^{-i\alpha\hat{G}} = \hat{A} + i\alpha[\hat{G}, \hat{A}] + \frac{(i\alpha)^2}{2!}\left[\hat{G},\ [\hat{G},\ \hat{A}]\right] + \frac{(i\alpha)^3}{3!}\left[\hat{G},\ [\hat{G},\ [\hat{G},\ \hat{A}]]\right] + \cdots.} \tag{2.168}$$

If $\hat{G}$ commutes with $\hat{A}$, the unitary transformation will leave $\hat{A}$ unchanged, $\hat{A}' = \hat{A}$:

$$[\hat{G}, \hat{A}] = 0 \Longrightarrow \hat{A}' = e^{i\alpha\hat{G}}\hat{A}e^{-i\alpha\hat{G}} = \hat{A}. \tag{2.169}$$

In Chapter 3, we will consider some important applications of infinitesimal unitary transformations to study time translations, space translations, space rotations, and conservation laws.

2.5 Representation in Discrete Bases

By analogy with the expansion of Euclidean space vectors in terms of the basis vectors, we need to express any ket $\mid \psi\rangle$ of the Hilbert space in terms of a complete set of mutually orthonormal base kets. State vectors are then represented by their components in this basis.

2.5.1 Matrix Representation of Kets, Bras, and Operators

Consider a discrete, complete, and orthonormal basis which is made of an infinite[4] set of kets $\mid \phi_1\rangle, \mid \phi_2\rangle, \mid \phi_3\rangle, \ldots, \mid \phi_n\rangle$ and denote it by $\{\mid \phi_n\rangle\}$. Note that the basis $\{\mid \phi_n\rangle\}$ is discrete, yet it has an infinite number of unit vectors. In the limit $n \to \infty$, the ordering index n of the unit vectors $\mid \phi_n\rangle$ is *discrete* or *countable*; that is, the sequence $\mid \phi_1\rangle, \mid \phi_2\rangle, \mid \phi_3\rangle, \ldots$ is countably infinite. As an illustration, consider the special functions, such as the Hermite, Legendre, or Laguerre polynomials, $H_n(x)$, $P_n(x)$, and $L_n(x)$. These polynomials are identified by a discrete index n and by a continuous variable x; although n varies discretely, it can be infinite.

In Section 2.6, we will consider bases that have a continuous and infinite number of base vectors; in these bases the index n increases continuously. Thus, each basis has a continuum of base vectors.

In this section the notation $\{\mid \phi_n\rangle\}$ will be used to abbreviate an infinitely countable set of vectors (i.e., $\mid \phi_1\rangle, \mid \phi_2\rangle, \mid \phi_3\rangle, \ldots$) of the Hilbert space $\mathcal{H}$. The orthonormality condition of the base kets is expressed by

$$\langle \phi_n \mid \phi_m\rangle = \delta_{nm}, \tag{2.170}$$

where δ_{nm} is the *Kronecker delta* symbol defined by

$$\delta_{nm} = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases} \tag{2.171}$$

The completeness, or closure, relation for this basis is given by

$$\sum_{n=1}^{\infty} \mid \phi_n\rangle\langle \phi_n \mid = \hat{I}, \tag{2.172}$$

where $\hat{I}$ is the unit operator; when the unit operator acts on any ket, it leaves the ket unchanged.

2.5.1.1 Matrix Representation of Kets and Bras

Let us now examine how to represent the vector $\mid \psi\rangle$ within the context of the basis $\{\mid \phi_n\rangle\}$. The completeness property of this basis enables us to expand any state vector $\mid \psi\rangle$ in terms of the base kets $\mid \phi_n\rangle$:

$$\mid \psi\rangle = \hat{I} \mid \psi\rangle = \left(\sum_{n=1}^{\infty} \mid \phi_n\rangle\langle \phi_n \mid\right) \mid \psi\rangle = \sum_{n=1}^{\infty} a_n \mid \phi_n\rangle, \tag{2.173}$$

where the coefficient a_n, which is equal to $\langle \phi_n \mid \psi\rangle$, represents the projection of $\mid \psi\rangle$ onto $\mid \phi_n\rangle$; a_n is the component of $\mid \psi\rangle$ along the vector $\mid \phi_n\rangle$. Recall that the coefficients a_n are complex numbers. So, within the basis $\{\mid \phi_n\rangle\}$, the ket $\mid \psi\rangle$ is represented by the set of its components, $a_1, a_2, a_3, \ldots$ along $\mid \phi_1\rangle, \mid \phi_2\rangle, \mid \phi_3\rangle, \ldots$, respectively. Hence $\mid \psi\rangle$ can be represented by a *column* vector which has a countably infinite number of components:

$$\mid \psi\rangle \longrightarrow \begin{pmatrix} \langle \phi_1 \mid \psi\rangle \\ \langle \phi_2 \mid \psi\rangle \\ \vdots \\ \langle \phi_n \mid \psi\rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix}. \tag{2.174}$$

[4] Kets are elements of the Hilbert space, and the dimension of a Hilbert space is infinite.

The bra $\langle \psi \mid$ can be represented by a *row* vector:

$$\begin{aligned}\langle \psi \mid \longrightarrow\ & (\langle \psi \mid \phi_1 \rangle\ \ \langle \psi \mid \phi_2 \rangle\ \ \cdots\ \langle \psi \mid \phi_n \rangle\ \cdots) \\ =\ & (\langle \phi_1 \mid \psi \rangle^*\ \ \langle \phi_2 \mid \psi \rangle^*\ \ \cdots\ \ \langle \phi_n \mid \psi \rangle^*\ \cdots) \\ =\ & (a_1^*\ \ a_2^*\ \ \cdots\ a_n^*\ \cdots). \end{aligned} \tag{2.175}$$

Using this representation, we see that a bra-ket $\langle \psi \mid \phi \rangle$ is a complex number equal to the matrix product of the row matrix corresponding to the bra $\langle \psi \mid$ with the column matrix corresponding to the ket $\mid \phi \rangle$:

$$\langle \psi \mid \phi \rangle = (a_1^*\ \ a_2^*\ \ \cdots\ \ a_n^*\ \cdots) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \\ \vdots \end{pmatrix} = \sum_n a_n^* b_n, \tag{2.176}$$

where $b_n = \langle \phi_n \mid \phi \rangle$. We see that, within this representation, the matrices representing $\mid \psi \rangle$ and $\langle \psi \mid$ are Hermitian adjoints of each other.

Remark

A ket $|\psi\rangle$ is normalized if $\langle \psi \mid \psi \rangle = \sum_n |a_n|^2 = 1$. If $|\psi\rangle$ is not normalized and we want to normalized it, we need simply to multiply it by a constant α so that $\langle \alpha\psi \mid \alpha\psi \rangle = |\alpha|^2 \langle \psi \mid \psi \rangle = 1$, and hence $\alpha = 1/\sqrt{\langle \psi \mid \psi \rangle}$.

Example 2.12

Consider the following two kets:

$$\mid \psi \rangle = \begin{pmatrix} 5i \\ 2 \\ -i \end{pmatrix}, \qquad \mid \phi \rangle = \begin{pmatrix} 3 \\ 8i \\ -9i \end{pmatrix}.$$

(a) Find $\mid \psi \rangle^*$ and $\langle \psi \mid$.
(b) Is $\mid \psi \rangle$ normalized? If not, normalize it.
(c) Are $\mid \psi \rangle$ and $\mid \phi \rangle$ orthogonal?

Solution

(a) The expressions of $\mid \psi \rangle^*$ and $\langle \psi \mid$ are given by

$$\mid \psi \rangle^* = \begin{pmatrix} -5i \\ 2 \\ i \end{pmatrix}, \qquad \langle \psi \mid\ =\ (-5i \quad 2 \quad i), \tag{2.177}$$

where we have used the fact that $\langle \psi \mid$ is equal to the complex conjugate of the transpose of the ket $\mid \psi \rangle$. Hence, we should reiterate the important fact that $\mid \psi \rangle^* \neq \langle \psi \mid$.

(b) The norm of $\mid \psi \rangle$ is given by

$$\langle \psi \mid \psi \rangle = (-5i \quad 2 \quad i) \begin{pmatrix} 5i \\ 2 \\ -i \end{pmatrix} = (-5i)(5i) + (2)(2) + (i)(-i) = 30. \tag{2.178}$$

Thus, $\mid \psi\rangle$ is not normalized. By multiplying it with $1/\sqrt{30}$, it becomes normalized:

$$\mid \chi\rangle = \frac{1}{\sqrt{30}} \mid \psi\rangle = \frac{1}{\sqrt{30}} \begin{pmatrix} 5i \\ 2 \\ -i \end{pmatrix} \qquad \Longrightarrow \qquad \langle \chi \mid \chi\rangle = 1. \tag{2.179}$$

(c) The kets $\mid \psi\rangle$ and $\mid \phi\rangle$ are not orthogonal since their scalar product is not zero:

$$\langle \psi \mid \phi\rangle = (-5i \quad 2 \quad i) \begin{pmatrix} 3 \\ 8i \\ -9i \end{pmatrix} = (-5i)(3) + (2)(8i) + (i)(-9i) = 9 + i. \tag{2.180}$$

2.5.1.2 Matrix Representation of Operators

For each linear operator $\hat{A}$, we can write

$$\hat{A} = \hat{I}\hat{A}\hat{I} = \left(\sum_{n=1}^{\infty} \mid \phi_n\rangle\langle\phi_n \mid\right) \hat{A} \left(\sum_{m=1}^{\infty} \mid \phi_m\rangle\langle\phi_m \mid\right) = \sum_{nm} A_{nm} \mid \phi_n\rangle\langle\phi_m \mid, \tag{2.181}$$

where A_{nm} is the nm matrix element of the operator $\hat{A}$:

$$A_{nm} \;=\; \langle\phi_n \mid \hat{A} \mid \phi_m\rangle. \tag{2.182}$$

We see that the operator $\hat{A}$ is represented, within the basis $\{\mid \phi_n\rangle\}$, by a *square* matrix A (A without a hat designates a matrix), which has a countably infinite number of columns and a countably infinite number of rows:

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.183}$$

For instance, the *unit* operator $\hat{I}$ is represented by the unit matrix; when the unit matrix is multiplied with another matrix, it leaves that unchanged:

$$I = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.184}$$

In summary, *kets are represented by column vectors, bras by row vectors, and operators by square matrices*.

2.5.1.3 Matrix Representation of Some Other Operators

(a) Hermitian adjoint operation

Let us now look at the matrix representation of the Hermitian adjoint operation of an operator. First, recall that the *transpose* of a matrix A, denoted by A^T, is obtained by interchanging the rows with the columns:

$$(A^T)_{nm} = A_{mn} \quad \text{or} \quad \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} & A_{31} & \cdots \\ A_{12} & A_{22} & A_{32} & \cdots \\ A_{13} & A_{23} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.185}$$

Similarly, the transpose of a column matrix is a row matrix, and the transpose of a row matrix is a column matrix:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix}^T = \begin{pmatrix} a_1 & a_2 & \cdots & a_n & \cdots \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a_1 & a_2 & \cdots & a_n & \cdots \end{pmatrix}^T = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix}. \tag{2.186}$$

So a square matrix A is symmetric if it is equal to its transpose, $A^T = A$. A skew-symmetric matrix is a square matrix whose transpose equals the negative of the matrix, $A^T = -A$.

The *complex conjugate* of a matrix is obtained by simply taking the complex conjugate of all its elements: $(A^*)_{nm} = (A_{nm})^*$.

The matrix which represents the operator $\hat{A}^\dagger$ is obtained by taking the complex conjugate of the matrix transpose of A:

$$A^\dagger = (A^T)^* \quad \text{or} \quad (\hat{A}^\dagger)_{nm} = \langle \phi_n \mid \hat{A}^\dagger \mid \phi_m \rangle = \langle \phi_m \mid \hat{A} \mid \phi_n \rangle^* = A^*_{mn}; \tag{2.187}$$

that is,

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}^\dagger = \begin{pmatrix} A^*_{11} & A^*_{21} & A^*_{31} & \cdots \\ A^*_{12} & A^*_{22} & A^*_{32} & \cdots \\ A^*_{13} & A^*_{23} & A^*_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.188}$$

If an operator $\hat{A}$ is Hermitian, its matrix satisfies this condition:

$$(A^T)^* = A \quad \text{or} \quad A^*_{mn} = A_{nm}. \tag{2.189}$$

The diagonal elements of a Hermitian matrix therefore must be real numbers. Note that *a Hermitian matrix must be square*.

(b) Inverse and unitary operators

A matrix has an inverse only if it is square and its determinant is nonzero; a matrix that has an inverse is called a nonsingular matrix and a matrix that has no inverse is called a singular matrix. The elements A^{-1}_{nm} of the inverse matrix A^{-1}, representing an operator $\hat{A}^{-1}$, are given by the relation

$$A^{-1}_{nm} = \frac{\text{cofactor of } A_{mn}}{\text{determinant of } A} \quad \text{or} \quad A^{-1} = \frac{B^T}{\text{determinant of } A}, \tag{2.190}$$

where B is the matrix of *cofactors* (also called the minor); the cofactor of element A_{mn} is equal to $(-1)^{m+n}$ times the determinant of the submatrix obtained from A by removing the mth row and the nth column. Note that when the matrix, representing an operator, has a determinant equal to zero, this operator does not possess an inverse. Note that $A^{-1}A = AA^{-1} = I$ where I is the unit matrix.

The inverse of a product of matrices is obtained as follows:

$$(ABC\cdots PQ)^{-1} = Q^{-1}P^{-1}\cdots C^{-1}B^{-1}A^{-1}. \tag{2.191}$$

The inverse of the inverse of a matrix is equal to the matrix itself, $\left(A^{-1}\right)^{-1} = A$.

A *unitary* operator $\hat{U}$ is represented by a unitary matrix. A matrix U is said to be unitary if its inverse is equal to its adjoint:

$$U^{-1} = U^{\dagger} \quad \text{or} \quad U^{\dagger}U = I, \tag{2.192}$$

where I is the unit matrix.

Example 2.13 (Inverse of a matrix)

Calculate the inverse of the matrix $A = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}$. Is this matrix unitary?

Solution

Since the determinant of A is $\det(A) = -4 + 16i$, we have $A^{-1} = B^T/(-4+16i)$, where the elements of the cofactor matrix B are given by $B_{nm} = (-1)^{n+m}$ times the determinant of the submatrix obtained from A by removing the nth row and the mth column. In this way, we have

$$B_{11} = (-1)^{1+1}\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} = (-1)^2\begin{vmatrix} 1 & 5 \\ -i & -2 \end{vmatrix} = -2 + 5i, \tag{2.193}$$

$$B_{12} = (-1)^{1+2}\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} = (-1)^3\begin{vmatrix} 3 & 5 \\ 0 & -2 \end{vmatrix} = 6, \tag{2.194}$$

$$B_{13} = (-1)^{1+3}\begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} = (-1)^4\begin{vmatrix} 3 & 1 \\ 0 & -i \end{vmatrix} = -3i, \tag{2.195}$$

$$B_{21} = (-1)^3\begin{vmatrix} i & 0 \\ -i & -2 \end{vmatrix} = 2i, \qquad B_{22} = (-1)^4\begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix} = -4, \tag{2.196}$$

$$B_{23} = (-1)^5\begin{vmatrix} 2 & i \\ 0 & -i \end{vmatrix} = 2i, \qquad B_{31} = (-1)^4\begin{vmatrix} i & 0 \\ 1 & 5 \end{vmatrix} = 5i, \tag{2.197}$$

$$B_{32} = (-1)^5\begin{vmatrix} 2 & 0 \\ 3 & 5 \end{vmatrix} = -10, \qquad B_{33} = (-1)^6\begin{vmatrix} 2 & i \\ 3 & 1 \end{vmatrix} = 2 - 3i, \tag{2.198}$$

and hence

$$B = \begin{pmatrix} -2+5i & 6 & -3i \\ 2i & -4 & 2i \\ 5i & -10 & 2-3i \end{pmatrix}. \tag{2.199}$$

Taking the transpose of B, we obtain

$$\begin{aligned} A^{-1} &= \frac{1}{-4+16i} B^T = \frac{-1-4i}{68} \begin{pmatrix} -2+5i & 2i & 5i \\ 6 & -4 & -10 \\ -3i & 2i & 2-3i \end{pmatrix} \\ &= \frac{1}{68} \begin{pmatrix} 22+3i & 8-2i & 20-5i \\ -6-24i & 4+16i & 10+40i \\ -12+3i & 8-2i & -14-5i \end{pmatrix}. \end{aligned} \tag{2.200}$$

Clearly, this matrix is not unitary since its inverse is not equal to its Hermitian adjoint: $A^{-1} \neq A^\dagger$.

(c) Matrix representation of $\mid \psi\rangle\langle\psi \mid$

It is now easy to see that the product $\mid \psi\rangle\langle\psi \mid$ is indeed an operator, since its representation within $\{\mid \phi_n\rangle\}$ is a square matrix:

$$\mid \psi\rangle\langle\psi \mid = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} \begin{pmatrix} a_1^* & a_2^* & a_3^* & \cdots \end{pmatrix} = \begin{pmatrix} a_1a_1^* & a_1a_2^* & a_1a_3^* & \cdots \\ a_2a_1^* & a_2a_2^* & a_2a_3^* & \cdots \\ a_3a_1^* & a_3a_2^* & a_3a_3^* & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.201}$$

(d) Trace of an operator

The trace $\mathrm{Tr}(\hat{A})$ of an operator $\hat{A}$ is given, within an orthonormal basis $\{\mid \phi_n\rangle\}$, by the expression

$$\mathrm{Tr}(\hat{A}) = \sum_n \langle\phi_n \mid \hat{A} \mid \phi_n\rangle = \sum_n A_{nn}; \tag{2.202}$$

we will see later that the trace of an operator does not depend on the basis. The trace of a matrix is equal to the sum of its diagonal elements:

$$\mathrm{Tr} \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = A_{11} + A_{22} + A_{33} + \cdots. \tag{2.203}$$

Properties of the trace

We can ascertain that

$$\mathrm{Tr}(\hat{A}^\dagger) = (\mathrm{Tr}(\hat{A}))^*, \tag{2.204}$$

$$\mathrm{Tr}(\alpha\hat{A} + \beta\hat{B} + \gamma\hat{C} + \cdots) = \alpha\mathrm{Tr}(\hat{A}) + \beta\mathrm{Tr}(\hat{B}) + \gamma\mathrm{Tr}(\hat{C}) + \cdots, \tag{2.205}$$

and the trace of a product of operators is invariant under the cyclic permutations of these operators:

$$\mathrm{Tr}(\hat{A}\hat{B}\hat{C}\hat{D}\hat{E}) = \mathrm{Tr}(\hat{E}\hat{A}\hat{B}\hat{C}\hat{D}) = \mathrm{Tr}(\hat{D}\hat{E}\hat{A}\hat{B}\hat{C}) = \mathrm{Tr}(\hat{C}\hat{D}\hat{E}\hat{A}\hat{B}) = \cdots. \tag{2.206}$$

Example 2.14

(a) Show that $\mathrm{Tr}(\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{A})$.

(b) Show that the trace of a commutator is always zero.

(c) Illustrate the results shown in (a) and (b) on the following matrices:

$$A = \begin{pmatrix} 8-2i & 4i & 0 \\ 1 & 0 & 1-i \\ -8 & i & 6i \end{pmatrix}, \qquad B = \begin{pmatrix} -i & 2 & 1-i \\ 6 & 1+i & 3i \\ 1 & 5+7i & 0 \end{pmatrix}.$$

Solution

(a) Using the definition of the trace,

$$\mathrm{Tr}(\hat{A}\hat{B}) = \sum_n \langle \phi_n \mid \hat{A}\hat{B} \mid \phi_n \rangle, \tag{2.207}$$

and inserting the unit operator between $\hat{A}$ and $\hat{B}$ we have

$$\begin{aligned} \mathrm{Tr}(\hat{A}\hat{B}) &= \sum_n \langle \phi_n \mid \hat{A} \left(\sum_m \mid \phi_m \rangle\langle \phi_m \mid \right) \hat{B} \mid \phi_n \rangle = \sum_{nm} \langle \phi_n \mid \hat{A} \mid \phi_m \rangle\langle \phi_m \mid \hat{B} \mid \phi_n \rangle \\ &= \sum_{nm} A_{nm} B_{mn}. \end{aligned} \tag{2.208}$$

On the other hand, since $\mathrm{Tr}(\hat{A}\hat{B}) = \sum_n \langle \phi_n \mid \hat{A}\hat{B} \mid \phi_n \rangle$, we have

$$\begin{aligned} \mathrm{Tr}(\hat{B}\hat{A}) &= \sum_m \langle \phi_m \mid \hat{B} \left(\sum_n \mid \phi_n \rangle\langle \phi_n \mid \right) \hat{A} \mid \phi_m \rangle = \sum_m \langle \phi_m \mid \hat{B} \mid \phi_n \rangle\langle \phi_n \mid \hat{A} \mid \phi_m \rangle \\ &= \sum_{nm} B_{mn} A_{nm}. \end{aligned} \tag{2.209}$$

Comparing (2.208) and (2.209), we see that $\mathrm{Tr}(\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{A})$.

(b) Since $\mathrm{Tr}(\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{A})$ we can infer at once that the trace of any commutator is always zero:

$$\mathrm{Tr}([\hat{A},\ \hat{B}]) = \mathrm{Tr}(\hat{A}\hat{B}) - \mathrm{Tr}(\hat{B}\hat{A}) = 0. \tag{2.210}$$

(c) Let us verify that the traces of the products AB and BA are equal. Since

$$AB = \begin{pmatrix} -2+16i & 12 & -6-10i \\ 1-2i & 14+2i & 1-i \\ 20i & -59+31i & -11+8i \end{pmatrix}, \quad BA = \begin{pmatrix} -8 & 5+i & 8+4i \\ 49-35i & -3+24i & -16 \\ 13+5i & 4i & 12+2i \end{pmatrix}, \tag{2.211}$$

we have

$$\mathrm{Tr}(AB) = \mathrm{Tr} \begin{pmatrix} -2+16i & 12 & -6-10i \\ 1-2i & 14+2i & 1-i \\ 20i & -59+31i & -11+8i \end{pmatrix} = 1+26i, \tag{2.212}$$

$$\mathrm{Tr}(BA) = \mathrm{Tr} \begin{pmatrix} -8 & 5+i & 8+4i \\ 49-35i & -3+24i & -16 \\ 13+5i & 4i & 12+2i \end{pmatrix} = 1+26i = \mathrm{Tr}(AB). \tag{2.213}$$

This leads to $\mathrm{Tr}(AB) - \mathrm{Tr}(BA) = (1+26i) - (1+26i) = 0$ or $\mathrm{Tr}([A,\ B]) = 0$.

2.5.1.4 Matrix Representation of Several Other Quantities

(a) Matrix representation of $|\,\phi\rangle = \hat{A}\,|\,\psi\rangle$
The relation $|\,\phi\rangle = \hat{A}\,|\,\psi\rangle$ can be cast into the algebraic form $\hat{I}\,|\,\phi\rangle = \hat{I}\hat{A}\hat{I}\,|\,\psi\rangle$ or

$$\left(\sum_n |\,\phi_n\rangle\langle\phi_n\,|\right)|\,\phi\rangle = \left(\sum_n |\,\phi_n\rangle\langle\phi_n\,|\right)\hat{A}\left(\sum_m |\,\phi_m\rangle\langle\phi_m\,|\right)|\,\psi\rangle, \tag{2.214}$$

which in turn can be written as

$$\sum_n b_n\,|\,\phi_n\rangle = \sum_{nm} a_m\,|\,\phi_n\rangle\langle\phi_n\,|\,\hat{A}\,|\,\phi_m\rangle = \sum_{nm} a_m A_{nm}\,|\,\phi_n\rangle, \tag{2.215}$$

where $b_n = \langle\phi_n\,|\,\phi\rangle$, $A_{nm} = \langle\phi_n\,|\,\hat{A}\,|\,\phi_m\rangle$, and $a_m = \langle\phi_m\,|\,\psi\rangle$. It is easy to see that (2.215) yields $b_n = \sum_m A_{nm}a_m$; hence the matrix representation of $|\,\phi\rangle = \hat{A}\,|\,\psi\rangle$ is given by

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix}. \tag{2.216}$$

(b) Matrix representation of $\langle\phi\,|\,\hat{A}\,|\,\psi\rangle$
As for $\langle\phi\,|\,\hat{A}\,|\,\psi\rangle$ we have

$$\begin{aligned} \langle\phi\,|\,\hat{A}\,|\,\psi\rangle &= \langle\phi\,|\,\hat{I}\hat{A}\hat{I}\,|\,\psi\rangle = \langle\phi\,|\left(\sum_{n=1}^{\infty} |\,\phi_n\rangle\langle\phi_n\,|\right)\hat{A}\left(\sum_{m=1}^{\infty} |\,\phi_m\rangle\langle\phi_m\,|\right)|\,\psi\rangle \\ &= \sum_{nm}\langle\phi\,|\,\phi_n\rangle\langle\phi_n\,|\,\hat{A}\,|\,\phi_m\rangle\langle\phi_m\,|\,\psi\rangle \\ &= \sum_{nm} b_n^* A_{nm} a_m. \end{aligned} \tag{2.217}$$

This is a complex number; its matrix representation goes as follows:

$$\langle\phi\,|\,\hat{A}\,|\,\psi\rangle \longrightarrow (b_1^* \quad b_2^* \quad b_3^* \quad \cdots)\begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix}. \tag{2.218}$$

Remark
It is now easy to see explicitly why products of the type $|\,\psi\rangle\,|\,\phi\rangle$, $\langle\psi\,|\,\langle\phi\,|$, $\hat{A}\langle\psi\,|$, or $|\,\psi\rangle\hat{A}$ are forbidden. They cannot have matrix representations; they are nonsensical. For instance, $|\,\psi\rangle\,|\,\phi\rangle$ is represented by the product of two column matrices:

$$|\,\psi\rangle\,|\,\phi\rangle \longrightarrow \begin{pmatrix} \langle\phi_1\,|\,\psi\rangle \\ \langle\phi_2\,|\,\psi\rangle \\ \vdots \end{pmatrix}\begin{pmatrix} \langle\phi_1\,|\,\phi\rangle \\ \langle\phi_2\,|\,\phi\rangle \\ \vdots \end{pmatrix}. \tag{2.219}$$

This product is clearly not possible to perform, for the product of two matrices is possible only when the number of columns of the first is equal to the number of rows of the second; in (2.219) the first matrix has one single column and the second an infinite number of rows.

2.5.1.5 Properties of a Matrix A

- Real if $A = A^*$ or $A_{mn} = A^*_{mn}$
- Imaginary if $A = -A^*$ or $A_{mn} = -A^*_{mn}$
- Symmetric if $A = A^T$ or $A_{mn} = A_{nm}$
- Antisymmetric if $A = -A^T$ or $A_{mn} = -A_{nm}$ with $A_{mm} = 0$
- Hermitian if $A = A^\dagger$ or $A_{mn} = A^*_{nm}$
- Anti-Hermitian if $A = -A^\dagger$ or $A_{mn} = -A^*_{nm}$
- Orthogonal if $A^T = A^{-1}$ or $AA^T = I$ or $(AA^T)_{mn} = \delta_{mn}$
- Unitary if $A^\dagger = A^{-1}$ or $AA^\dagger = I$ or $(AA^\dagger)_{mn} = \delta_{mn}$

Example 2.15
Consider a matrix A (which represents an operator $\hat{A}$), a ket $\mid\psi\rangle$, and a bra $\langle\phi\mid$:

$$A = \begin{pmatrix} 5 & 3+2i & 3i \\ -i & 3i & 8 \\ 1-i & 1 & 4 \end{pmatrix}, \quad \mid\psi\rangle = \begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix}, \quad \langle\phi\mid = \begin{pmatrix} 6 & -i & 5 \end{pmatrix}.$$

(a) Calculate the quantities $A\mid\psi\rangle$, $\langle\phi\mid A$, $\langle\phi\mid A\mid\psi\rangle$, and $\mid\psi\rangle\langle\phi\mid$.

(b) Find the complex conjugate, the transpose, and the Hermitian conjugate of A, $\mid\psi\rangle$, and $\langle\phi\mid$.

(c) Calculate $\langle\phi\mid\psi\rangle$ and $\langle\psi\mid\phi\rangle$; are they equal? Comment on the differences between the complex conjugate, Hermitian conjugate, and transpose of kets and bras.

Solution

(a) The calculations are straightforward:

$$A\mid\psi\rangle = \begin{pmatrix} 5 & 3+2i & 3i \\ -i & 3i & 8 \\ 1-i & 1 & 4 \end{pmatrix}\begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix} = \begin{pmatrix} -5+17i \\ 17+34i \\ 11+14i \end{pmatrix}, \tag{2.220}$$

$$\langle\phi\mid A = \begin{pmatrix} 6 & -i & 5 \end{pmatrix}\begin{pmatrix} 5 & 3+2i & 3i \\ -i & 3i & 8 \\ 1-i & 1 & 4 \end{pmatrix} = \begin{pmatrix} 34-5i & 26+12i & 20+10i \end{pmatrix}, \tag{2.221}$$

$$\langle\phi\mid A\mid\psi\rangle = \begin{pmatrix} 6 & -i & 5 \end{pmatrix}\begin{pmatrix} 5 & 3+2i & 3i \\ -i & 3i & 8 \\ 1-i & 1 & 4 \end{pmatrix}\begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix} = 59+155i, \tag{2.222}$$

$$\mid\psi\rangle\langle\phi\mid = \begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix}\begin{pmatrix} 6 & -i & 5 \end{pmatrix} = \begin{pmatrix} -6+6i & 1+i & -5+5i \\ 18 & -3i & 15 \\ 12+18i & 3-2i & 10+15i \end{pmatrix}. \tag{2.223}$$

(b) To obtain the complex conjugate of A, $|\ \psi\rangle$, and $\langle\phi\ |$, we need simply to take the complex conjugate of their elements:

$$A^* = \begin{pmatrix} 5 & 3-2i & -3i \\ i & -3i & 8 \\ 1+i & 1 & 4 \end{pmatrix}, \quad |\ \psi\rangle^* = \begin{pmatrix} -1-i \\ 3 \\ 2-3i \end{pmatrix}, \quad \langle\phi\ |^* = \begin{pmatrix} 6 & i & 5 \end{pmatrix}. \tag{2.224}$$

For the transpose of A, $|\ \psi\rangle$, and $\langle\phi\ |$, we simply interchange columns with rows:

$$A^T = \begin{pmatrix} 5 & -i & 1-i \\ 3+2i & 3i & 1 \\ 3i & 8 & 4 \end{pmatrix}, \quad |\ \psi\rangle^T = \begin{pmatrix} -1+i & 3 & 2+3i \end{pmatrix}, \quad \langle\phi\ |^T = \begin{pmatrix} 6 \\ -i \\ 5 \end{pmatrix}. \tag{2.225}$$

The Hermitian conjugate can be obtained by taking the complex conjugates of the transpose expressions calculated above: $A^\dagger = (A^T)^*$, $|\ \psi\rangle^\dagger = \left(|\ \psi\rangle^T\right)^* = \langle\psi\ |$, $\langle\phi\ |^\dagger = \left(\langle\phi\ |^T\right)^* = |\ \phi\rangle$:

$$A^\dagger = \begin{pmatrix} 5 & i & 1+i \\ 3-2i & -3i & 1 \\ -3i & 8 & 4 \end{pmatrix}, \quad \langle\psi\ | = \begin{pmatrix} -1-i & 3 & 2-3i \end{pmatrix}, \quad |\ \phi\rangle = \begin{pmatrix} 6 \\ i \\ 5 \end{pmatrix}. \tag{2.226}$$

(c) Using the kets and bras above, we can easily calculate the needed scalar products:

$$\langle\phi\ |\ \psi\rangle = \begin{pmatrix} 6 & -i & 5 \end{pmatrix} \begin{pmatrix} -1+i \\ 3 \\ 2+3i \end{pmatrix} = 6(-1+i)+(-i)(3)+5(2+3i) = 4+18i, \tag{2.227}$$

$$\langle\psi\ |\ \phi\rangle = \begin{pmatrix} -1-i & 3 & 2-3i \end{pmatrix} \begin{pmatrix} 6 \\ i \\ 5 \end{pmatrix} = 6(-1-i)+(i)(3)+5(2-3i) = 4-18i. \tag{2.228}$$

We see that $\langle\phi\ |\ \psi\rangle$ and $\langle\psi\ |\ \phi\rangle$ are not equal; they are complex conjugates of each other:

$$\langle\psi\ |\ \phi\rangle = \langle\phi\ |\ \psi\rangle^* = 4 - 18i. \tag{2.229}$$

Remark

We should underscore the importance of the differences between $|\ \psi\rangle^*$, $|\ \psi\rangle^T$, and $|\ \psi\rangle^\dagger$. Most notably, we should note (from equations (2.224)–(2.226)) that $|\ \psi\rangle^*$ is a ket, while $|\ \psi\rangle^T$ and $|\ \psi\rangle^\dagger$ are bras. Additionally, we should note that $\langle\phi\ |^*$ is a bra, while $\langle\phi\ |^T$ and $\langle\phi\ |^\dagger$ are kets.

2.5.2 Change of Bases and Unitary Transformations

In a Euclidean space, a vector $\vec{A}$ may be represented by its components in different coordinate systems or in different bases. The transformation from one basis to the other is called a change of basis. The components of $\vec{A}$ in a given basis can be expressed in terms of the components of $\vec{A}$ in another basis by means of a *transformation* matrix.

Similarly, state vectors and operators of quantum mechanics may also be represented in different bases. In this section we are going to study how to transform from one basis to another. That is, knowing the components of kets, bras, and operators in a basis $\{|\ \phi_n\rangle\}$, how does one determine the corresponding components in a different basis $\{| \phi'_n\rangle\}$? Assuming that $\{| \phi_n\rangle\}$ and $\{| \phi'_n\rangle\}$ are two different bases, we can expand each ket $| \phi_n\rangle$ of the old basis in terms of the new basis $\{| \phi'_n\rangle\}$ as follows:

$$| \phi_n\rangle = \left(\sum_m | \phi'_m\rangle\langle\phi'_m |\right) | \phi_n\rangle = \sum_m U_{mn} | \phi'_m\rangle, \qquad (2.230)$$

where

$$U_{mn} = \langle\phi'_m | \phi_n\rangle. \qquad (2.231)$$

The matrix U, providing the transformation from the old basis $\{| \phi_n\rangle\}$ to the new basis $\{| \phi'_n\rangle\}$, is given by

$$U = \begin{pmatrix} \langle\phi'_1 | \phi_1\rangle & \langle\phi'_1 | \phi_2\rangle & \langle\phi'_1 | \phi_3\rangle \\ \langle\phi'_2 | \phi_1\rangle & \langle\phi'_2 | \phi_2\rangle & \langle\phi'_2 | \phi_3\rangle \\ \langle\phi'_3 | \phi_1\rangle & \langle\phi'_3 | \phi_2\rangle & \langle\phi'_3 | \phi_3\rangle \end{pmatrix}. \qquad (2.232)$$

Example 2.16 (Unitarity of the transformation matrix)
Let U be a transformation matrix which connects two *complete* and *orthonormal* bases $\{| \phi_n\rangle\}$ and $\{| \phi'_n\rangle\}$. Show that U is *unitary*.

Solution
For this we need to prove that $\hat{U}\hat{U}^\dagger = \hat{I}$, which reduces to showing that $\langle\phi_m | \hat{U}\hat{U}^\dagger | \phi_n\rangle = \delta_{mn}$. This goes as follows:

$$\langle\phi_m | \hat{U}\hat{U}^\dagger | \phi_n\rangle = \langle\phi_m | \hat{U}\left(\sum_l | \phi_l\rangle\langle\phi_l |\right)\hat{U}^\dagger | \phi_n\rangle = \sum_l U_{ml}U^*_{nl}, \qquad (2.233)$$

where $U_{ml} = \langle\phi_m | \hat{U} | \phi_l\rangle$ and $U^*_{nl} = \langle\phi_l | \hat{U}^\dagger | \phi_n\rangle = \langle\phi_n | \hat{U} | \phi_l\rangle^*$. According to (2.231), $U_{ml} = \langle\phi'_m | \phi_l\rangle$ and $U^*_{nl} = \langle\phi_l | \phi'_n\rangle$; we can thus rewrite (2.233) as

$$\sum_l U_{ml}U^*_{nl} = \sum_l \langle\phi'_m | \phi_l\rangle\langle\phi_l | \phi'_n\rangle = \langle\phi'_m | \phi'_n\rangle = \delta_{mn}. \qquad (2.234)$$

Combining (2.233) and (2.234), we infer $\langle\phi_m | \hat{U}\hat{U}^\dagger | \phi_n\rangle = \delta_{mn}$, or $\hat{U}\hat{U}^\dagger = \hat{I}$.

2.5.2.1 Transformations of Kets, Bras, and Operators

The components $\langle\phi'_n | \psi\rangle$ of a state vector $| \psi\rangle$ in a new basis $\{| \phi'_n\rangle\}$ can be expressed in terms of the components $\langle\phi_n | \psi\rangle$ of $| \psi\rangle$ in an old basis $\{| \phi_n\rangle\}$ as follows:

$$\langle\phi'_m | \psi\rangle = \langle\phi'_m | \hat{I} | \psi\rangle = \langle\phi'_m | \left(\sum_n | \phi_n\rangle\langle\phi_n |\right) | \psi\rangle = \sum_n U_{mn}\langle\phi_n | \psi\rangle. \qquad (2.235)$$

This relation, along with its complex conjugate, can be generalized into

$$| \psi_{new}\rangle = \hat{U} | \psi_{old}\rangle, \qquad \langle\psi_{new} | = \langle\psi_{old} | \hat{U}^\dagger. \qquad (2.236)$$

Let us now examine how operators transform when we change from one basis to another. The matrix elements $A'_{mn} = \langle \phi'_m \mid \hat{A} \mid \phi'_n \rangle$ of an operator $\hat{A}$ in the new basis can be expressed in terms of the old matrix elements, $A_{jl} = \langle \phi_j \mid \hat{A} \mid \phi_l \rangle$, as follows:

$$A'_{mn} = \langle \phi'_m \mid \left(\sum_j \mid \phi_j \rangle \langle \phi_j \mid \right) \hat{A} \left(\sum_l \mid \phi_l \rangle \langle \phi_l \mid \right) \mid \phi'_n \rangle = \sum_{jl} U_{mj} A_{jl} U^*_{nl}; \tag{2.237}$$

that is,

$$\hat{A}_{new} = \hat{U} \hat{A}_{old} \hat{U}^\dagger \qquad \text{or} \qquad \hat{A}_{old} = \hat{U}^\dagger \hat{A}_{new} \hat{U}. \tag{2.238}$$

We may summarize the results of the change of basis in the following relations:

$$\boxed{\mid \psi_{new} \rangle = \hat{U} \mid \psi_{old} \rangle, \qquad \langle \psi_{new} \mid = \langle \psi_{old} \mid \hat{U}^\dagger, \qquad \hat{A}_{new} = \hat{U} \hat{A}_{old} \hat{U}^\dagger,} \tag{2.239}$$

or

$$\boxed{\mid \psi_{old} \rangle = \hat{U}^\dagger \mid \psi_{new} \rangle, \qquad \langle \psi_{old} \mid = \langle \psi_{new} \mid \hat{U}, \qquad \hat{A}_{old} = \hat{U}^\dagger \hat{A}_{new} \hat{U}.} \tag{2.240}$$

These relations are similar to the ones we derived when we studied unitary transformations; see (2.146) and (2.147).

Example 2.17
Show that the operator $\hat{U} = \sum_n \mid \phi'_n \rangle \langle \phi_n \mid$ satisfies all the properties discussed above.

Solution
First, note that $\hat{U}$ is unitary:

$$\hat{U}\hat{U}^\dagger = \sum_{nl} \mid \phi'_n \rangle \langle \phi_n \mid \phi_l \rangle \langle \phi'_l \mid = \sum_{nl} \mid \phi'_n \rangle \langle \phi'_l \mid \delta_{nl} = \sum_n \mid \phi'_n \rangle \langle \phi'_n \mid = \hat{I}. \tag{2.241}$$

Second, the action of $\hat{U}$ on a ket of the old basis gives the corresponding ket from the new basis:

$$\hat{U} \mid \phi_m \rangle = \sum_n \mid \phi'_n \rangle \langle \phi_n \mid \phi_m \rangle = \sum_n \mid \phi'_n \rangle \delta_{nm} = \mid \phi'_m \rangle. \tag{2.242}$$

We can also verify that the action $\hat{U}^\dagger$ on a ket of the new basis gives the corresponding ket from the old basis:

$$\hat{U}^\dagger \mid \phi'_m \rangle = \sum_l \mid \phi_l \rangle \langle \phi'_l \mid \phi'_m \rangle = \sum_l \mid \phi_l \rangle \delta_{lm} = \mid \phi_m \rangle. \tag{2.243}$$

How does a trace transform under unitary transformations? Using the cyclic property of the trace, $\text{Tr}(\hat{A}\hat{B}\hat{C}) = \text{Tr}(\hat{C}\hat{A}\hat{B}) = \text{Tr}(\hat{B}\hat{C}\hat{A})$, we can ascertain that

$$\text{Tr}(\hat{A}') = \text{Tr}(\hat{U}\hat{A}\hat{U}^\dagger) = \text{Tr}(\hat{U}^\dagger \hat{U} \hat{A}) = \text{Tr}(\hat{A}), \tag{2.244}$$

$$\begin{aligned}\text{Tr}\,(|\,\phi_n\rangle\langle\phi_m\,|) &= \sum_l \langle\phi_l \mid \phi_n\rangle\langle\phi_m \mid \phi_l\rangle = \sum_l \langle\phi_m \mid \phi_l\rangle\langle\phi_l \mid \phi_n\rangle \\ &= \langle\phi_m \mid \left(\sum_l \mid \phi_l\rangle\langle\phi_l \mid\right) \mid \phi_n\rangle = \langle\phi_m \mid \phi_n\rangle = \delta_{mn}, \end{aligned} \tag{2.245}$$

$$\text{Tr}\left(|\,\phi'_m\rangle\langle\phi_n\,|\right) = \langle\phi_n \mid \phi'_m\rangle. \tag{2.246}$$

Example 2.18 (The trace is base independent)
Show that the trace of an operator does not depend on the basis in which it is expressed.

Solution
Let us show that the trace of an operator $\hat{A}$ in a basis $\{|\,\phi_n\rangle\}$ is equal to its trace in another basis $\{|\,\phi'_n\rangle\}$. First, the trace of $\hat{A}$ in the basis $\{|\,\phi_n\rangle\}$ is given by

$$\text{Tr}(\hat{A}) = \sum_n \langle\phi_n \mid \hat{A} \mid \phi_n\rangle \tag{2.247}$$

and in $\{|\,\phi'_n\rangle\}$ by

$$\text{Tr}(\hat{A}) = \sum_n \langle\phi'_n \mid \hat{A} \mid \phi'_n\rangle. \tag{2.248}$$

Starting from (2.247) and using the completeness of the other basis, $\{|\,\phi'_n\rangle\}$, we have

$$\begin{aligned}\text{Tr}(\hat{A}) &= \sum_n \langle\phi_n \mid \hat{A} \mid \phi_n\rangle = \sum_n \langle\phi_n \mid \left(\sum_m \mid \phi'_m\rangle\langle\phi'_m \mid\right)\hat{A} \mid \phi_n\rangle \\ &= \sum_{nm} \langle\phi_n \mid \phi'_m\rangle\langle\phi'_m \mid \hat{A} \mid \phi_n\rangle. \end{aligned} \tag{2.249}$$

All we need to do now is simply to interchange the positions of the numbers (scalars) $\langle\phi_n \mid \phi'_m\rangle$ and $\langle\phi'_m \mid \hat{A} \mid \phi_n\rangle$:

$$\text{Tr}(\hat{A}) = \sum_m \langle\phi'_m \mid \hat{A}\left(\sum_n \mid \phi_n\rangle\langle\phi_n \mid\right) \mid \phi'_m\rangle = \sum_m \langle\phi'_m \mid \hat{A} \mid \phi'_m\rangle. \tag{2.250}$$

From (2.249) and (2.250) we see that

$$\text{Tr}(\hat{A}) = \sum_n \langle\phi_n \mid \hat{A} \mid \phi_n\rangle = \sum_n \langle\phi'_n \mid \hat{A} \mid \phi'_n\rangle. \tag{2.251}$$

2.5.3 Matrix Representation of the Eigenvalue Problem

At issue here is to work out the matrix representation of the eigenvalue problem (2.126) and then solve it. That is, we want to find the eigenvalues a and the eigenvectors $|\,\psi\rangle$ of an operator $\hat{A}$ such that

$$\hat{A} \mid \psi\rangle = a \mid \psi\rangle, \tag{2.252}$$

where a is a complex number. Inserting the unit operator between $\hat{A}$ and $\mid \psi\rangle$ and multiplying by $\langle\phi_m \mid$, we can cast the eigenvalue equation in the form

$$\langle\phi_m \mid \hat{A}\left(\sum_n \mid \phi_n\rangle\langle\phi_n \mid\right) \mid \psi\rangle = a\langle\phi_m \mid \left(\sum_n \mid \phi_n\rangle\langle\phi_n \mid\right) \mid \psi\rangle, \tag{2.253}$$

or

$$\sum_n A_{mn}\langle\phi_n \mid \psi\rangle = a\sum_n \langle\phi_n \mid \psi\rangle\delta_{nm}, \tag{2.254}$$

which can be rewritten as

$$\sum_n \left[A_{mn} - a\delta_{nm}\right]\langle\phi_n \mid \psi\rangle = 0, \tag{2.255}$$

with $A_{mn} = \langle\phi_m \mid \hat{A} \mid \phi_n\rangle$.

This equation represents an infinite, homogeneous system of equations for the coefficients $\langle\phi_n \mid \psi\rangle$, since the basis $\{\mid \phi_n\rangle\}$ is made of an infinite number of base kets. This system of equations can have nonzero solutions only if its determinant vanishes:

$$\det\left(A_{mn} - a\delta_{nm}\right) = 0. \tag{2.256}$$

The problem that arises here is that this determinant corresponds to a matrix with an infinite number of columns and rows. To solve (2.256) we need to truncate the basis $\{\mid \phi_n\rangle\}$ and assume that it contains only N terms, where N must be large enough to guarantee convergence. In this case we can reduce (2.256) to the following Nth degree determinant:

$$\begin{vmatrix} A_{11} - a & A_{12} & A_{13} & \cdots & A_{1N} \\ A_{21} & A_{22} - a & A_{23} & \cdots & A_{2N} \\ A_{31} & A_{32} & A_{33} - a & \cdots & A_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & A_{N3} & \cdots & A_{NN} - a \end{vmatrix} = 0. \tag{2.257}$$

This is known as the *secular* or *characteristic equation*. The solutions of this equation yield the N *eigenvalues* $a_1, a_2, a_3, \ldots, a_N$, since it is an Nth order equation in a. The set of these N eigenvalues is called the spectrum of $\hat{A}$. Knowing the set of eigenvalues $a_1, a_2, a_3, \ldots, a_N$, we can easily determine the corresponding set of *eigenvectors* $\mid \phi_1\rangle, \mid \phi_2\rangle, \ldots, \mid \phi_N\rangle$. For each eigenvalue a_m of $\hat{A}$, we can obtain from the "secular" equation (2.257) the N components $\langle\phi_1 \mid \psi\rangle, \langle\phi_2 \mid \psi\rangle, \langle\phi_3 \mid \psi\rangle, \ldots, \langle\phi_N \mid \psi\rangle$ of the corresponding eigenvector $\mid \phi_m\rangle$.

If a number of different eigenvectors (two or more) have the same eigenvalue, this eigenvalue is said to be *degenerate*. The order of degeneracy is determined by the number of linearly independent eigenvectors that have the same eigenvalue. For instance, if an eigenvalue has five different eigenvectors, it is said to be fivefold degenerate.

In the case where the set of eigenvectors $\mid \phi_n\rangle$ of $\hat{A}$ is complete and orthonormal, this set can be used as a basis. In this basis the matrix representing the operator $\hat{A}$ is diagonal,

$$A = \begin{pmatrix} a_1 & 0 & 0 & \cdots \\ 0 & a_2 & 0 & \cdots \\ 0 & 0 & a_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.258}$$

the diagonal elements being the eigenvalues a_n of $\hat{A}$, since

$$\langle \phi_m \mid \hat{A} \mid \phi_n \rangle = a_n \langle \phi_m \mid \phi_n \rangle = a_n \delta_{mn}. \tag{2.259}$$

Note that the trace and determinant of a matrix are given, respectively, by the sum and product of the *eigenvalues*:

$$\mathrm{Tr}(A) = \sum_n a_n = a_1 + a_2 + a_3 + \cdots, \tag{2.260}$$

$$\det(A) = \prod_n a_n = a_1 a_2 a_3 \cdots. \tag{2.261}$$

Properties of determinants

Let us mention several useful properties that pertain to determinants. The determinant of a product of matrices is equal to the product of their determinants:

$$\det(ABCD\cdots) = \det(A)\cdot\det(B)\cdot\det(C)\cdot\det(D)\cdots, \tag{2.262}$$

$$\det(A^*) = (\det(A))^*, \qquad \det(A^\dagger) = (\det(A))^*, \tag{2.263}$$

$$\det(A^T) = \det(A), \qquad \det(A) = e^{\mathrm{Tr}(\ln A)}. \tag{2.264}$$

Some theorems pertaining to the eigenvalue problem

Here is a list of useful theorems (the proofs are left as exercises):

- The eigenvalues of a symmetric matrix are real; the eigenvectors form an orthonormal basis.
- The eigenvalues of an antisymmetric matrix are purely imaginary or zero.
- The eigenvalues of a Hermitian matrix are real; the eigenvectors form an orthonormal basis.
- The eigenvalues of a skew-Hermitian matrix are purely imaginary or zero.
- The eigenvalues of a unitary matrix have absolute value equal to one.
- If the eigenvalues of a square matrix are not degenerate (distinct), the corresponding eigenvectors form a basis (i.e., they form a linearly independent set).

Example 2.19 (Eigenvalues and eigenvectors of a matrix)

Find the eigenvalues and the normalized eigenvectors of the matrix

$$A = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}.$$

Solution

To find the eigenvalues of A, we simply need to solve the secular equation $\det(A - aI) = 0$:

$$0 = \begin{vmatrix} 7-a & 0 & 0 \\ 0 & 1-a & -i \\ 0 & i & -1-a \end{vmatrix} = (7-a)\left[-(1-a)(1+a) + i^2\right] = (7-a)(a^2-2). \tag{2.265}$$

The eigenvalues of A are thus given by

$$a_1 = 7, \quad a_2 = \sqrt{2}, \quad a_3 = -\sqrt{2}. \tag{2.266}$$

Let us now calculate the eigenvectors of A. To find the eigenvector corresponding to the first eigenvalue, $a_1 = 7$, we need to solve the matrix equation

$$\begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 7\begin{pmatrix} x \\ y \\ z \end{pmatrix} \Longrightarrow \begin{array}{rcl} 7x & = & 7x \\ y - iz & = & 7y \\ iy - z & = & 7z \end{array}; \tag{2.267}$$

this yields $x = 1$ (because the eigenvector is normalized) and $y = z = 0$. So the eigenvector corresponding to $a_1 = 7$ is given by the column matrix

$$|\, a_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{2.268}$$

This eigenvector is normalized since $\langle a_1 \mid a_1\rangle = 1$.

The eigenvector corresponding to the second eigenvalue, $a_2 = \sqrt{2}$, can be obtained from the matrix equation

$$\begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \sqrt{2}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \Longrightarrow \begin{array}{rcl} (7-\sqrt{2})x & = & 0 \\ (1-\sqrt{2})y - iz & = & 0 \\ iy - (1+\sqrt{2})z & = & 0 \end{array}; \tag{2.269}$$

this yields $x = 0$ and $z = i(\sqrt{2} - 1)y$. So the eigenvector corresponding to $a_2 = \sqrt{2}$ is given by the column matrix

$$|\, a_2\rangle = \begin{pmatrix} 0 \\ y \\ i(\sqrt{2}-1)y \end{pmatrix}. \tag{2.270}$$

The value of the variable y can be obtained from the normalization condition of $|\, a_2\rangle$:

$$1 = \langle a_2 \mid a_2\rangle = \begin{pmatrix} 0 & y^* & -i(\sqrt{2}-1)y^* \end{pmatrix}\begin{pmatrix} 0 \\ y \\ i(\sqrt{2}-1)y \end{pmatrix} = 2(2-\sqrt{2}) \mid y \mid^2 . \tag{2.271}$$

Taking only the positive value of y (a similar calculation can be performed easily if one is interested in the negative value of y), we have $y = 1/\sqrt{2(2-\sqrt{2})}$; hence the eigenvector (2.270) becomes

$$|\, a_2\rangle = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2(2-\sqrt{2})}} \\ \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} \end{pmatrix}. \tag{2.272}$$

Following the same procedure that led to (2.272), we can show that the third eigenvector is given by

$$|\, a_3\rangle = \begin{pmatrix} 0 \\ y \\ -i(1+\sqrt{2})y \end{pmatrix}; \tag{2.273}$$

its normalization leads to $y = 1/\sqrt{2(2+\sqrt{2})}$ (we have considered only the positive value of y); hence

$$|\, a_3\rangle = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix}. \tag{2.274}$$

2.6 Representation in Continuous Bases

In this section we are going to consider the representation of state vectors, bras, and operators in *continuous* bases. After presenting the general formalism, we will consider two important applications: representations in the *position* and *momentum* spaces.

In the previous section we saw that the representations of kets, bras, and operators in a discrete basis are given by discrete matrices. We will show here that these quantities are represented in a *continuous* basis by *continuous matrices*, that is, by noncountable infinite matrices.

2.6.1 General Treatment

The orthonormality condition of the base *kets* of the continuous basis $|\, \chi_k\rangle$ is expressed not by the usual discrete Kronecker delta as in (2.170) but by Dirac's *continuous delta* function:

$$\langle \chi_k \,|\, \chi_{k'}\rangle = \delta(k' - k), \tag{2.275}$$

where k and k' are continuous parameters and where $\delta(k'-k)$ is the Dirac delta function (see Appendix A), which is defined by

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ikx}\, dk. \tag{2.276}$$

As for the completeness condition of this continuous basis, it is not given by a discrete sum as in (2.172), but by an integral over the continuous variable

$$\int_{-\infty}^{+\infty} dk\ |\, \chi_k\rangle\langle \chi_k \,| = \hat{I}, \tag{2.277}$$

where $\hat{I}$ is the unit operator.

Every state vector $|\, \psi\rangle$ can be expanded in terms of the complete set of basis *kets* $|\, \chi_k\rangle$:

$$|\, \psi\rangle = \hat{I}\,|\, \psi\rangle = \left(\int_{-\infty}^{+\infty} dk\ |\, \chi_k\rangle\langle \chi_k \,|\right)|\, \psi\rangle = \int_{-\infty}^{+\infty} dk\, b(k)\,|\, \chi_k\rangle, \tag{2.278}$$

where $b(_k)$, which is equal to $\langle \chi_k \mid \psi \rangle$, represents the projection of $\mid \psi \rangle$ on $\mid \chi_k \rangle$.

The norm of the discrete base *kets* is finite ($\langle \phi_n \mid \phi_n \rangle = 1$), but the norm of the continuous base *kets* is infinite; a combination of (2.275) and (2.276) leads to

$$\langle \chi_k \mid \chi_k \rangle = \delta(0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \quad \longrightarrow \quad \infty. \tag{2.279}$$

This implies that the kets $\mid \chi_k \rangle$ are not square integrable and hence are not elements of the Hilbert space; recall that the space spanned by square-integrable functions is a Hilbert space. Despite the divergence of the norm of $\mid \chi_k \rangle$, the set $\mid \chi_k \rangle$ does constitute a valid basis of vectors that span the Hilbert space, since for any state vector $\mid \psi \rangle$, the scalar product $\langle \chi_k \mid \psi \rangle$ is finite.

The Dirac delta function

Before dealing with the representation of kets, bras, and operators, let us make a short detour to list some of the most important properties of the Dirac delta function (for a more detailed presentation, see Appendix A):

$$\delta(x) = 0, \qquad \text{for} \qquad x \neq 0, \tag{2.280}$$

$$\int_a^b f(x)\delta(x - x_0)\, dx = \begin{cases} f(x_0) & \text{if} \quad a < x_0 < b, \\ 0 & \text{elsewhere,} \end{cases} \tag{2.281}$$

$$\int_{-\infty}^{\infty} f(x) \frac{d^n \delta(x-a)}{dx^n}\, dx = (-1)^n \left. \frac{d^n f(x)}{dx^n} \right|_{x=a}, \tag{2.282}$$

$$\delta(\vec{r} - \vec{r}\,') = \delta(x - x\,')\delta(y - y\,')\delta(z - z\,') = \frac{1}{r^2 \sin\theta} \delta(r - r')\delta(\theta - \theta')\delta(\varphi - \varphi'). \tag{2.283}$$

Representation of kets, bras, and operators

The representation of kets, bras, and operators can be easily inferred from the study that was carried out in the previous section, for the case of a discrete basis. For instance, the ket $\mid \psi \rangle$ is represented by a single column matrix which has a continuous (noncountable) and infinite number of components (rows) $b(k)$:

$$\mid \psi \rangle \longrightarrow \begin{pmatrix} \vdots \\ \langle \chi_k \mid \psi \rangle \\ \vdots \end{pmatrix}. \tag{2.284}$$

The bra $\langle \psi \mid$ is represented by a single row matrix which has a continuous (noncountable) and infinite number of components (columns):

$$\langle \psi \mid \longrightarrow (\cdots\cdots \quad \langle \psi \mid \chi_k \rangle \quad \cdots\cdots). \tag{2.285}$$

Operators are represented by square continuous matrices whose rows and columns have continuous and infinite numbers of components:

$$\hat{A} \longrightarrow \begin{pmatrix} \ddots & \vdots & \ddots \\ \cdots & A(k, k') & \cdots \\ \ddots & \vdots & \ddots \end{pmatrix}. \tag{2.286}$$

As an application, we are going to consider the representations in the position and momentum bases.

2.6.2 Position Representation

In the position representation, the basis consists of an infinite set of vectors $\{|\ \vec{r}\rangle\}$ which are eigenkets to the position operator $\hat{\vec{R}}$:

$$\hat{\vec{R}}\ |\ \vec{r}\rangle\ =\ \vec{r}\ |\ \vec{r}\rangle, \tag{2.287}$$

where $\vec{r}$ (without a hat), the position vector, is the eigenvalue of the operator $\hat{\vec{R}}$. The orthonormality and completeness conditions are respectively given by

$$\langle\vec{r}\ |\ \vec{r}\ '\rangle = \delta(\vec{r}-\vec{r}\ ') \ = \ \delta(x-x')\delta(y-y')\delta(z-z'), \tag{2.288}$$

$$\int d^3 r\ |\ \vec{r}\rangle\langle\vec{r}\ | \ = \ \hat{I}, \tag{2.289}$$

since, as discussed in Appendix A, the three-dimensional delta function is given by

$$\delta(\vec{r}-\vec{r}\ ') = \frac{1}{(2\pi)^3}\int d^3k\, e^{i\vec{k}\cdot(\vec{r}-\vec{r}\ ')}. \tag{2.290}$$

So every state vector $|\ \psi\rangle$ can be expanded as follows:

$$|\ \psi\rangle\ =\int d^3 r\ |\ \vec{r}\rangle\langle\vec{r}\ |\ \psi\rangle\ \equiv\int d^3 r\ \psi(\vec{r})\ |\ \vec{r}\rangle, \tag{2.291}$$

where $\psi(\vec{r})$ denotes the components of $|\ \psi\rangle$ in the $\{|\ \vec{r}\rangle\}$ basis:

$$\langle\vec{r}\ |\ \psi\rangle\ =\ \psi(\vec{r}). \tag{2.292}$$

This is known as the *wave function* for the state vector $|\ \psi\rangle$. Recall that, according to the probabilistic interpretation of Born, the quantity $|\ \langle\vec{r}\ |\ \psi\rangle\ |^2\ d^3r$ represents the probability of finding the system in the volume element d^3r.

The scalar product between two state vectors, $|\ \psi\rangle$ and $|\ \phi\rangle$, can be expressed in this form:

$$\langle\phi\ |\ \psi\rangle\ =\ \langle\phi\ |\left(\int d^3r\ \ |\ \vec{r}\rangle\langle\vec{r}\ |\right)|\ \psi\rangle = \int d^3r\ \phi^*(\vec{r})\psi(\vec{r}). \tag{2.293}$$

Since $\hat{\vec{R}}\ |\ \vec{r}\rangle = \vec{r}\ |\ \vec{r}\rangle$ we have

$$\langle\vec{r}\ '\ |\ \hat{\vec{R}}^{\,n}\ |\ \vec{r}\rangle = \vec{r}^{\ n}\delta(\vec{r}\ '-\vec{r}). \tag{2.294}$$

Note that the operator $\hat{\vec{R}}$ is Hermitian, since

$$\begin{aligned}\langle\phi\ |\ \hat{\vec{R}}\ |\ \psi\rangle\ &=\ \int d^3r\ \vec{r}\langle\phi\ |\ \vec{r}\rangle\langle\vec{r}\ |\ \psi\rangle\ =\ \left[\int d^3r\ \vec{r}\langle\psi\ |\ \vec{r}\rangle\langle\vec{r}\ |\ \phi\rangle\right]^*\\ &=\ \langle\psi\ |\ \hat{\vec{R}}\ |\ \phi\rangle^*.\end{aligned} \tag{2.295}$$

2.6.3 Momentum Representation

The basis $\{|\ \vec{p}\rangle\}$ of the momentum representation is obtained from the eigenkets of the momentum operator $\hat{\vec{P}}$:

$$\hat{\vec{P}}\ |\ \vec{p}\rangle\ =\ \vec{p}\ |\ \vec{p}\rangle, \tag{2.296}$$

where $\vec{p}$ is the momentum vector. The algebra relevant to this representation can be easily inferred from the position representation. The orthonormality and completeness conditions of the momentum space basis $|\ \vec{p}\rangle$ are given by

$$\langle\vec{p}\ |\ \vec{p}\,'\rangle\ =\ \delta(\vec{p}-\vec{p}\,') \qquad \text{and} \qquad \int \mathrm{d}^3\mathrm{p}\ \ |\ \vec{\mathrm{p}}\rangle\langle\vec{\mathrm{p}}\ | = \hat{\mathrm{I}}. \tag{2.297}$$

Expanding $|\ \psi\rangle$ in this basis, we obtain

$$|\ \psi\rangle = \int d^3p\ \ |\ \vec{p}\rangle\langle\vec{p}\ |\ \psi\rangle = \int d^3p\ \Psi(\vec{p})\ |\ \vec{p}\rangle, \tag{2.298}$$

where the expansion coefficient $\Psi(\vec{p})$ represents the *momentum space wave function*. The quantity $|\ \Psi(\vec{p})\ |^2\ d^3p$ is the probability of finding the system's momentum in the volume element d^3p located between $\vec{p}$ and $\vec{p}+d\vec{p}$.

By analogy with (2.293) the scalar product between two states is given in the momentum space by

$$\langle\phi\ |\ \psi\rangle\ =\ \langle\phi\ |\left(\int d^3p\ \ |\ \vec{p}\rangle\langle\vec{p}\ |\right)|\ \psi\rangle = \int d^3p\ \Phi^*(\vec{p})\Psi(\vec{p}). \tag{2.299}$$

Since $\hat{\vec{P}}\ |\ \vec{p}\rangle\ =\ \vec{p}\ |\ \vec{p}\rangle$ we have

$$\langle\vec{p}\,'\ |\ \hat{\vec{P}}^{\,n}\ |\ \vec{p}\rangle = \vec{p}^{\,n}\delta(\vec{p}\,'-\vec{p}). \tag{2.300}$$

2.6.4 Connecting the Position and Momentum Representations

Let us now study how to establish a connection between the position and the momentum representations. By analogy with the foregoing study, when changing from the $\{|\ \vec{r}\rangle\}$ basis to the $\{|\ \vec{p}\rangle\}$ basis, we encounter the *transformation* function $\langle\vec{r}\ |\ \vec{p}\rangle$.

To find the expression for the transformation function $\langle\vec{r}\ |\ \vec{p}\rangle$, let us establish a connection between the position and momentum representations of the state vector $|\ \psi\rangle$:

$$\langle\vec{r}\ |\ \psi\rangle\ =\ \langle\vec{r}\ |\left(\int d^3p\ \ |\ \vec{p}\rangle\langle\vec{p}\ |\right)|\ \psi\rangle\ =\ \int d^3p\ \ \langle\vec{r}\ |\ \vec{p}\rangle\Psi(\vec{p}); \tag{2.301}$$

that is,

$$\psi(\vec{r}) = \int d^3p\ \langle\vec{r}\ |\ \vec{p}\rangle\Psi(\vec{p}). \tag{2.302}$$

Similarly, we can write

$$\Psi(\vec{p})\ =\ \langle\vec{p}\ |\ \psi\rangle\ =\ \langle\vec{p}\ |\int d^3r\ |\ \vec{r}\rangle\langle\vec{r}\ |\ \psi\rangle\ =\ \int d^3r\langle\vec{p}\ |\ \vec{r}\rangle\psi(\vec{r}). \tag{2.303}$$

The last two relations imply that $\psi(\vec{r})$ and $\Psi(\vec{p})$ are to be viewed as Fourier transforms of each other. In quantum mechanics the Fourier transform of a function $f(\vec{r})$ is given by

$$f(\vec{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3p\, e^{i\vec{p}\cdot\vec{r}/\hbar} g(\vec{p}); \tag{2.304}$$

notice the presence of Planck's constant. Hence the function $\langle \vec{r} \mid \vec{p} \rangle$ is given by

$$\boxed{\langle \vec{r} \mid \vec{p} \rangle \;=\; \frac{1}{(2\pi\hbar)^{3/2}} e^{i\vec{p}\cdot\vec{r}/\hbar}.} \tag{2.305}$$

This function transforms from the momentum to the position representation. The function corresponding to the inverse transformation, $\langle \vec{p} \mid \vec{r} \rangle$, is given by

$$\boxed{\langle \vec{p} \mid \vec{r} \rangle \;=\; \langle \vec{r} \mid \vec{p} \rangle^* = \frac{1}{(2\pi\hbar)^{3/2}} e^{-i\vec{p}\cdot\vec{r}/\hbar}.} \tag{2.306}$$

The quantity $|\langle \vec{r} \mid \vec{p} \rangle|^2$ represents the probability density of finding the particle in a region around $\vec{r}$ where its momentum is equal to $\vec{p}$.

Remark

If the position wave function

$$\psi(\vec{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3p\, e^{i\vec{p}\cdot\vec{r}/\hbar} \Psi(\vec{p}) \tag{2.307}$$

is normalized (i.e., $\int d^3r\; \psi(\vec{r})\psi^*(\vec{r}) = 1$), its Fourier transform

$$\Psi(\vec{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3r\; e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}) \tag{2.308}$$

must also be normalized, since

$$\begin{aligned}
\int d^3p\, \Psi^*(\vec{p})\Psi(\vec{p}) &= \int d^3p \Psi^*(\vec{p}) \left[\frac{1}{(2\pi\hbar)^{3/2}} \int d^3r\; e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}) \right] \\
&= \int d^3r\; \psi(\vec{r}) \left[\frac{1}{(2\pi\hbar)^{3/2}} \int d^3p\, \Psi^*(\vec{p}) e^{-i\vec{p}\cdot\vec{r}/\hbar} \right] \\
&= \int d^3r\; \psi(\vec{r})\psi^*(\vec{r}) \\
&= 1.
\end{aligned} \tag{2.309}$$

This result is known as *Parseval's theorem*.

2.6.4.1 Momentum Operator in the Position Representation

To determine the form of the momentum operator $\hat{\vec{P}}$ in the position representation, let us calculate $\langle \vec{r} \mid \hat{\vec{P}} \mid \psi \rangle$:

$$\begin{aligned}
\langle \vec{r} \mid \hat{\vec{P}} \mid \psi \rangle &= \int \langle \vec{r} \mid \hat{\vec{P}} \mid \vec{p} \rangle \langle \vec{p} \mid \psi \rangle d^3p = \int \vec{p} \langle \vec{r} \mid \vec{p} \rangle \langle \vec{p} \mid \psi \rangle d^3p \\
&= \frac{1}{(2\pi\hbar)^{3/2}} \int \vec{p}\, e^{i\vec{p}\cdot\vec{r}/\hbar} \Psi(\vec{p}) d^3p,
\end{aligned} \tag{2.310}$$

where we have used the relation $\int \mid \vec{p}\rangle\langle\vec{p} \mid d^3p = \hat{I}$ along with Eq. (2.305). Now, since $\vec{p}\, e^{i\vec{p}\cdot\vec{r}/\hbar} = -i\hbar\vec{\nabla}e^{i\vec{p}\cdot\vec{r}/\hbar}$, and using Eq. (2.305) again, we can rewrite (2.310) as

$$\begin{aligned}\langle\vec{r} \mid \hat{\vec{P}} \mid \psi\rangle &= -i\hbar\vec{\nabla}\left(\frac{1}{(2\pi\hbar)^{3/2}}\int e^{i\vec{p}\cdot\vec{r}/\hbar}\Psi(\vec{p})d^3p\right) \\ &= -i\hbar\vec{\nabla}\left(\int \langle\vec{r} \mid \vec{p}\rangle\langle\vec{p} \mid \psi\rangle d^3p\right) \\ &= -i\hbar\vec{\nabla}\langle\vec{r} \mid \psi\rangle. \end{aligned} \tag{2.311}$$

Thus, $\hat{\vec{P}}$ is given in the position representation by

$$\boxed{\hat{\vec{P}} = -i\hbar\vec{\nabla}.} \tag{2.312}$$

Its Cartesian components are

$$\boxed{\hat{P}_x = -i\hbar\frac{\partial}{\partial x}, \qquad \hat{P}_y = -i\hbar\frac{\partial}{\partial y}, \qquad \hat{P}_z = -i\hbar\frac{\partial}{\partial z}.} \tag{2.313}$$

Note that the form of the momentum operator (2.312) can be derived by simply applying the gradient operator $\vec{\nabla}$ on a *plane* wave function $\psi(\vec{r}, t) = Ae^{i(\vec{p}\cdot\vec{r}-Et)/\hbar}$:

$$-i\hbar\vec{\nabla}\psi(\vec{r}, t) = \vec{p}\psi(\vec{r}, t) = \hat{\vec{P}}\psi(\vec{r}, t). \tag{2.314}$$

It is easy to verify that $\hat{\vec{P}}$ is Hermitian (see equation (2.378)).

Now, since $\hat{\vec{P}} = -i\hbar\vec{\nabla}$, we can write the Hamiltonian operator $\hat{H} = \hat{\vec{P}}^{\,2}/(2m) + \hat{V}$ in the position representation as follows:

$$\boxed{\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(\vec{r}) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + \hat{V}(\vec{r}),} \tag{2.315}$$

where ∇^2 is the Laplacian operator; it is given in Cartesian coordinates by $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$.

2.6.4.2 Position Operator in the Momentum Representation

The form of the position operator $\hat{\vec{R}}$ in the momentum representation can be easily inferred from the representation of $\hat{\vec{P}}$ in the position space. In momentum space the position operator can be written as follows:

$$\hat{R}_j = i\hbar\frac{\partial}{\partial p_j} \qquad (j = x, y, z) \tag{2.316}$$

or

$$\boxed{\hat{X} = i\hbar\frac{\partial}{\partial p_x}, \qquad \hat{Y} = i\hbar\frac{\partial}{\partial p_y}, \qquad \hat{Z} = i\hbar\frac{\partial}{\partial p_z}.} \tag{2.317}$$

2.6.4.3 Important Commutation Relations

Let us now calculate the commutator $[\hat{R}_j, \hat{P}_k]$ in the position representation. As the separate actions of $\hat{X}\hat{P}_x$ and $\hat{P}_x\hat{X}$ on the wave function $\psi(\vec{r})$ are given by

$$\hat{X}\hat{P}_x\psi(\vec{r}) = -i\hbar x\frac{\partial\psi(\vec{r})}{\partial x}, \tag{2.318}$$

$$\hat{P}_x\hat{X}\psi(\vec{r}) = -i\hbar\frac{\partial}{\partial x}(x\psi(\vec{r})) = -i\hbar\psi(\vec{r}) - i\hbar x\frac{\partial\psi(\vec{r})}{\partial x}, \tag{2.319}$$

we have

$$\begin{aligned}[\hat{X}, \hat{P}_x]\psi(\vec{r}) &= \hat{X}\hat{P}_x\psi(\vec{r}) - \hat{P}_x\hat{X}\psi(\vec{r}) = -i\hbar x\frac{\partial\psi(\vec{r})}{\partial x} + i\hbar\psi(\vec{r}) + i\hbar x\frac{\partial\psi(\vec{r})}{\partial x} \\ &= i\hbar\psi(\vec{r}) \end{aligned} \tag{2.320}$$

or

$$[\hat{X}, \hat{P}_x] = i\hbar. \tag{2.321}$$

Similar relations can be derived at once for the y and the z components:

$$\boxed{[\hat{X}, \hat{P}_x] = i\hbar, \qquad [\hat{Y}, \hat{P}_Y] = i\hbar, \qquad [\hat{Z}, \hat{P}_Z] = i\hbar.} \tag{2.322}$$

We can verify that

$$[\hat{X}, \hat{P}_y] = [\hat{X}, \hat{P}_z] = [\hat{Y}, \hat{P}_x] = [\hat{Y}, \hat{P}_z] = [\hat{Z}, \hat{P}_x] = [\hat{Z}, \hat{P}_y] = 0, \tag{2.323}$$

since the x, y, z degrees of freedom are independent; the previous two relations can be grouped into

$$\boxed{[\hat{R}_j, \hat{P}_k] = i\hbar\delta_{jk}, \qquad [\hat{R}_j, \hat{R}_k] = 0, \qquad [\hat{P}_j, \hat{P}_k] = 0 \qquad (j,\ k = x, y, z).} \tag{2.324}$$

These relations are often called the *canonical commutation relations*.

Now, from (2.321) we can show that (for the proof see Problem 2.8 on page 139)

$$\boxed{[\hat{X}^n, \hat{P}_x] = i\hbar n\hat{X}^{n-1}, \qquad [\hat{X}, \hat{P}_x^n] = i\hbar n\hat{P}_x^{n-1}.} \tag{2.325}$$

Following the same procedure that led to (2.320), we can obtain a more general commutation relation of $\hat{P}_x$ with an arbitrary function $f(\hat{X})$:

$$\boxed{[f(\hat{X}), \hat{P}_x] = i\hbar\frac{df(\hat{X})}{d\hat{X}} \quad\Longrightarrow\quad \left[\hat{\vec{P}},\ F(\hat{\vec{R}})\right] = -i\hbar\vec{\nabla}F(\hat{\vec{R}})),} \tag{2.326}$$

where F is a function of the operator $\hat{\vec{R}}$.

The explicit form of operators thus depends on the representation adopted. We have seen, however, that the *commutation relations for operators are representation independent*. In particular, the commutator $[\hat{R}_j, \hat{P}_k]$ is given by $i\hbar\delta_{jk}$ in the position and the momentum representations; see the next example.

Example 2.20 (Commutators are representation independent)
Calculate the commutator $[\hat{X}, \hat{P}]$ in the momentum representation and verify that it is equal to $i\hbar$.

Solution
As the operator $\hat{X}$ is given in the momentum representation by $\hat{X} = i\hbar\partial/\partial p$, we have

$$\begin{aligned} [\hat{X}, \hat{P}]\psi(p) &= \hat{X}\hat{P}\psi(p) - \hat{P}\hat{X}\psi(p) = i\hbar\frac{\partial}{\partial p}\left(p\psi(p)\right) - i\hbar p\frac{\partial\psi(p)}{\partial p} \\ &= i\hbar\psi(p) + i\hbar p\frac{\partial\psi(p)}{\partial p} - i\hbar p\frac{\partial\psi(p)}{\partial p} = i\hbar\psi(p). \end{aligned} \tag{2.327}$$

Thus, the commutator $[\hat{X}, \hat{P}]$ is given in the *momentum representation* by

$$[\hat{X}, \hat{P}] = \left[i\hbar\frac{\partial}{\partial p},\ \hat{P}\right] = i\hbar. \tag{2.328}$$

The commutator $[\hat{X}, \hat{P}]$ was also shown to be equal to $i\hbar$ in the *position representation* (see equation (2.321):

$$[\hat{X}, \hat{P}] = -\left[\hat{X},\ i\hbar\frac{\partial}{\partial p_x}\right] = i\hbar. \tag{2.329}$$

2.6.5 Parity Operator

The *space reflection* about the origin of the coordinate system is called an *inversion* or a *parity* operation. This transformation is *discrete*. The parity operator $\hat{\mathcal{P}}$ is defined by its action on the kets $\mid \vec{r}\rangle$ of the position space:

$$\hat{\mathcal{P}} \mid \vec{r}\rangle = \mid -\vec{r}\rangle, \qquad \langle\vec{r} \mid \hat{\mathcal{P}}^\dagger = \langle -\vec{r} \mid, \tag{2.330}$$

such that

$$\hat{\mathcal{P}}\psi(\vec{r}) = \psi(-\vec{r}). \tag{2.331}$$

The parity operator is Hermitian, $\hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}$, since

$$\begin{aligned} \int d^3r\ \phi^*(\vec{r})\left[\hat{\mathcal{P}}\psi(\vec{r})\right] &= \int d^3r\phi^*(\vec{r})\psi(-\vec{r}) = \int d^3r\ \phi^*(-\vec{r})\psi(\vec{r}) \\ &= \int d^3r\ \left[\hat{\mathcal{P}}\phi(\vec{r})\right]^*\psi(\vec{r}). \end{aligned} \tag{2.332}$$

From the definition (2.331), we have

$$\hat{\mathcal{P}}^2\psi(\vec{r}) = \hat{\mathcal{P}}\psi(-\vec{r}) = \psi(\vec{r}); \tag{2.333}$$

hence $\hat{\mathcal{P}}^2$ is equal to the unity operator:

$$\hat{\mathcal{P}}^2 = \hat{I} \qquad \text{or} \qquad \hat{\mathcal{P}} = \hat{\mathcal{P}}^{-1}. \tag{2.334}$$

The parity operator is therefore *unitary*, since its Hermitian adjoint is equal to its inverse:

$$\hat{\mathcal{P}}^\dagger = \hat{\mathcal{P}}^{-1}. \tag{2.335}$$

Now, since $\hat{\mathcal{P}}^2 = \hat{I}$, the eigenvalues of $\hat{\mathcal{P}}$ are $+1$ or -1 with the corresponding eigenstates

$$\hat{\mathcal{P}}\psi_+(\vec{r}) = \psi_+(-\vec{r}) = \psi_+(\vec{r}), \qquad \hat{\mathcal{P}}\psi_-(\vec{r}) = \psi_-(-\vec{r}) = -\psi_-(\vec{r}). \tag{2.336}$$

The eigenstate $\mid \psi_+\rangle$ is said to be *even* and $\mid \psi_-\rangle$ is *odd*. Therefore, the eigenfunctions of the parity operator have *definite parity*: they are either even or odd.

Since $\mid \psi_+\rangle$ and $\mid \psi_-\rangle$ are joint eigenstates of the same Hermitian operator $\hat{\mathcal{P}}$ but with different eigenvalues, these eigenstates must be orthogonal:

$$\langle\psi_+ \mid \psi_-\rangle = \int d^3r\ \psi_+^*(-\vec{r})\psi_-(-\vec{r}) \equiv -\int d^3r\ \psi_+^*(\vec{r})\psi_-(\vec{r}) = -\langle\psi_+ \mid \psi_-\rangle; \tag{2.337}$$

hence $\langle\psi_+ \mid \psi_-\rangle$ is zero. The states $\mid \psi_+\rangle$ and $\mid \psi_-\rangle$ form a complete set since any function can be written as $\psi(\vec{r}) = \psi_+(\vec{r}) + \psi_-(\vec{r})$, which leads to

$$\psi_+(\vec{r}) = \frac{1}{2}\left[\psi(\vec{r}) + \psi(-\vec{r})\right], \qquad \psi_-(\vec{r}) = \frac{1}{2}\left[\psi(\vec{r}) - \psi(-\vec{r})\right]. \tag{2.338}$$

Since $\hat{\mathcal{P}}^2 = I$ we have

$$\hat{\mathcal{P}}^n = \begin{cases} \hat{\mathcal{P}} & \text{when } n \text{ is odd,} \\ \hat{I} & \text{when } n \text{ is even.} \end{cases} \tag{2.339}$$

Even and odd operators

An operator $\hat{A}$ is said to be *even* if it obeys the condition

$$\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}} = \hat{A} \tag{2.340}$$

and an operator $\hat{B}$ is *odd* if

$$\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}} = -\hat{B}. \tag{2.341}$$

We can easily verify that even operators commute with the parity operator $\hat{\mathcal{P}}$ and that odd operators anticommute with $\hat{\mathcal{P}}$:

$$\hat{A}\hat{\mathcal{P}} = (\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}})\hat{\mathcal{P}} = \hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}}^2 = \hat{\mathcal{P}}\hat{A}, \tag{2.342}$$

$$\hat{B}\hat{\mathcal{P}} = -(\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}})\hat{\mathcal{P}} = -\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}}^2 = -\hat{\mathcal{P}}\hat{B}. \tag{2.343}$$

The fact that even operators commute with the parity operator has very useful consequences. Let us examine the following two important cases depending on whether an even operator has nondegenerate or degenerate eigenvalues:

- If an even operator is Hermitian and none of its eigenvalues is degenerate, then this operator has the same eigenvectors as those of the parity operator. And since the eigenvectors of the parity operator are either even or odd, the eigenvectors of an even, Hermitian, and nondegenerate operator must also be either even or odd; they are said to have a *definite parity*. This property will have useful applications when we solve the Schrödinger equation for even Hamiltonians.

- If the even operator has a degenerate spectrum, its eigenvectors do not necessarily have a definite parity.

What about the parity of the position and momentum operators, $\hat{\vec{R}}$ and $\hat{\vec{P}}$? We can easily show that both of them are odd, since they anticommute with the parity operator:

$$\hat{\mathcal{P}}\hat{\vec{R}} = -\hat{\vec{R}}\hat{\mathcal{P}}, \qquad \hat{\mathcal{P}}\hat{\vec{P}} = -\hat{\vec{P}}\hat{\mathcal{P}}; \tag{2.344}$$

hence

$$\hat{\mathcal{P}}\hat{\vec{R}}\hat{\mathcal{P}}^\dagger = -\hat{\vec{R}}, \qquad \hat{\mathcal{P}}\hat{\vec{P}}\hat{\mathcal{P}}^\dagger = -\hat{\vec{P}}, \tag{2.345}$$

since $\hat{\mathcal{P}}\hat{\mathcal{P}}^\dagger = 1$. For instance, to show that $\hat{\vec{R}}$ anticommutes with $\hat{\mathcal{P}}$, we need simply to look at the following relations:

$$\hat{\mathcal{P}}\hat{\vec{R}} \mid \vec{r}\rangle = \vec{r}\hat{\mathcal{P}} \mid \vec{r}\rangle = \vec{r} \mid -\vec{r}\rangle, \tag{2.346}$$

$$\hat{\vec{R}}\hat{\mathcal{P}} \mid \vec{r}\rangle = \hat{\vec{R}} \mid -\vec{r}\rangle = -\vec{r} \mid -\vec{r}\rangle. \tag{2.347}$$

If the operators $\hat{A}$ and $\hat{B}$ are even and odd, respectively, we can verify that

$$\hat{\mathcal{P}}\hat{A}^n\hat{\mathcal{P}} = \hat{A}^n, \qquad \hat{\mathcal{P}}\hat{B}^n\hat{\mathcal{P}} = (-1)^n\hat{B}^n. \tag{2.348}$$

These relations can be shown as follows:

$$\hat{\mathcal{P}}\hat{A}^n\hat{\mathcal{P}} = (\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}})\,(\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}})\cdots(\hat{\mathcal{P}}\hat{A}\hat{\mathcal{P}}) = \hat{A}^n, \tag{2.349}$$

$$\hat{\mathcal{P}}\hat{B}^n\hat{\mathcal{P}} = (\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}})\,(\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}})\cdots(\hat{\mathcal{P}}\hat{B}\hat{\mathcal{P}}) = (-1)^n\hat{B}^n. \tag{2.350}$$

2.7 Matrix and Wave Mechanics

In this chapter we have so far worked out the mathematics pertaining to quantum mechanics in two different representations: *discrete* basis systems and *continuous* basis systems. The theory of quantum mechanics deals in essence with solving the following eigenvalue problem:

$$\hat{H} \mid \psi\rangle = E \mid \psi\rangle, \tag{2.351}$$

where $\hat{H}$ is the Hamiltonian of the system. This equation is general and does not depend on any coordinate system or representation. But to solve it, we need to represent it in a given basis system. The complexity associated with solving this eigenvalue equation will then vary from one basis to another.

In what follows we are going to examine the representation of this eigenvalue equation in a *discrete* basis and then in a *continuous* basis.

2.7.1 Matrix Mechanics

The representation of quantum mechanics in a *discrete* basis yields a *matrix* eigenvalue problem. That is, the representation of (2.351) in a discrete basis $\{\mid \phi_n\rangle\}$ yields the following matrix eigenvalue equation (see (2.257)):

$$\begin{vmatrix} H_{11}-E & H_{12} & H_{13} & \cdots & H_{1N} \\ H_{21} & H_{22}-E & H_{23} & \cdots & H_{2N} \\ H_{31} & H_{32} & H_{33}-E & \cdots & H_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{N1} & H_{N2} & H_{N3} & \cdots & H_{NN}-E \end{vmatrix} = 0. \tag{2.352}$$

This is an Nth order equation in E; its solutions yield the energy spectrum of the system: E_1, E_2, E_3, ..., E_N. Knowing the set of eigenvalues E_1, E_2, E_3, ..., E_N, we can easily determine the corresponding set of eigenvectors $\mid \phi_1\rangle$, $\mid \phi_2\rangle$, ..., $\mid \phi_N\rangle$.

The diagonalization of the Hamiltonian matrix (2.352) of a system yields the energy spectrum as well as the state vectors of the system. This procedure, which was worked out by Heisenberg, involves only matrix quantities and matrix eigenvalue equations. This formulation of quantum mechanics is known as *matrix mechanics*.

The starting point of Heisenberg, in his attempt to find a theoretical foundation to Bohr's ideas, was the atomic transition relation, $\nu_{mn} = (E_m - E_n)/h$, which gives the frequencies of the radiation associated with the electron's transition from orbit m to orbit n. The frequencies ν_{mn} can be arranged in a square matrix, where the mn element corresponds to the transition from the mth to the nth quantum state.

We can also construct matrices for other dynamical quantities related to the transition $m \to n$. In this way, every physical quantity is represented by a matrix. For instance, we represent the energy levels by an energy matrix, the position by a position matrix, the momentum by a momentum matrix, the angular momentum by an angular momentum matrix, and so on. In calculating the various physical magnitudes, one has thus to deal with the algebra of matrix quantities. So, within the context of matrix mechanics, one deals with noncommuting quantities, for the product of matrices does not commute. This is an essential feature that distinguishes matrix mechanics from classical mechanics, where all the quantities commute. Take, for instance, the position and momentum quantities. While commuting in classical mechanics, $px = xp$, they do not commute within the context of matrix mechanics; they are related by the commutation relation $[\hat{X},\ \hat{P}_x] = i\hbar$. The same thing applies for the components of angular momentum. We should note that the role played by the commutation relations within the context of matrix mechanics is similar to the role played by Bohr's quantization condition in atomic theory. Heisenberg's matrix mechanics therefore requires the introduction of some mathematical machinery—linear vector spaces, Hilbert space, commutator algebra, and matrix algebra—that is entirely different from the mathematical machinery of classical mechanics. Here lies the justification for having devoted a somewhat lengthy section, Section 2.5, to study the matrix representation of quantum mechanics.

2.7.2 Wave Mechanics

Representing the formalism of quantum mechanics in a *continuous* basis yields an eigenvalue problem not in the form of a matrix equation, as in Heisenberg's formulation, but in the form of a *differential equation*. The representation of the eigenvalue equation (2.351) in the *position* space yields

$$\langle \vec{r} \mid \hat{H} \mid \psi\rangle = E\langle \vec{r} \mid \psi\rangle. \tag{2.353}$$

As shown in (2.315), the Hamiltonian is given in the position representation by $-\hbar^2\nabla^2/(2m)+\hat{V}(\vec{r})$, so we can rewrite (2.353) in a more familiar form:

$$\boxed{-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r})+\hat{V}(\vec{r})\psi(\vec{r})=E\psi(\vec{r}),} \tag{2.354}$$

where $\langle\vec{r}\mid\psi\rangle=\psi(\vec{r})$ is the *wave function* of the system. This differential equation is known as the *Schrödinger equation* (its origin will be discussed in Chapter 3). Its solutions yield the energy spectrum of the system as well as its wave function. This formulation of quantum mechanics in *the position representation* is called *wave mechanics*.

Unlike Heisenberg, Schödinger took an entirely different starting point in his quest to find a theoretical justification for Bohr's ideas. He started from the de Broglie particle–wave hypothesis and extended it to the electrons orbiting around the nucleus. Schrödinger aimed at finding an equation that describes the motion of the electron within an atom. Here the focus is on the *wave* aspect of the electron. We can show, as we did in Chapter 1, that the Bohr quantization condition, $L=n\hbar$, is equivalent to the de Broglie relation, $\lambda=2\pi\hbar/p$. To establish this connection, we need simply to make three assumptions: (a) the wavelength of the wave associated with the orbiting electron is connected to the electron's linear momentum p by $\lambda=2\pi\hbar/p$, (b) the electron's orbit is circular, and (c) the circumference of the electron's orbit is an integer multiple of the electron's wavelength, i.e., $2\pi r=n\lambda$. This leads at once to $2\pi r=n\times(2\pi\hbar/p)$ or $n\hbar=rp\equiv L$. This means that, for every orbit, there is only one wavelength (or one wave) associated with the electron while revolving in that orbit. This wave can be described by means of a *wave function*. So Bohr's quantization condition implies, in essence, a uniqueness of the wave function for each orbit of the electron. In Chapter 3 we will show how Schrödinger obtained his differential equation (2.354) to describe the motion of an electron in an atom.

2.8 Concluding Remarks

Historically, the matrix formulation of quantum mechanics was worked out by Heisenberg shortly before Schrödinger introduced his wave theory. The equivalence between the matrix and wave formulations was proved a few years later by using the theory of unitary transformations. Different in form, yet identical in contents, wave mechanics and matrix mechanics achieve the same goal: finding the energy spectrum and the states of quantum systems.

The matrix formulation has the advantage of greater (formal) generality, yet it suffers from a number of disadvantages. On the conceptual side, it offers no visual idea about the structure of the atom; it is less intuitive than wave mechanics. On the technical side, it is difficult to use in some problems of relative ease such as finding the stationary states of atoms. Matrix mechanics, however, becomes powerful and practical in solving problems such as the harmonic oscillator or in treating the formalism of angular momentum.

But most of the efforts of quantum mechanics focus on solving the Schrödinger equation, not the Heisenberg matrix eigenvalue problem. So in the rest of this text we deal mostly with wave mechanics. Matrix mechanics is used only in a few problems, such as the harmonic oscillator, where it is more suitable than Schrödinger's wave mechanics.

In wave mechanics we need only to specify the potential in which the particle moves; the Schrödinger equation takes care of the rest. That is, knowing $\hat{V}(\vec{r})$, we can in principle solve equation (2.354) to obtain the various energy levels of the particle and their corresponding wave functions. The complexity we encounter in solving the differential equation depends entirely on the form of the potential; the simpler the potential the easier the solution. Exact solutions of the Schrödinger equation are possible only for a few idealized systems; we deal with such systems in Chapters 4 and 6. However, exact solutions are generally not possible, for real systems do not yield themselves to exact solutions. In such cases one has to resort to approximate solutions. We deal with such approximate treatments in Chapters 9 and 10; Chapter 9 deals with time-independent potentials and Chapter 10 with time-dependent potentials.

Before embarking on the applications of the Schrödinger equation, we need first to lay down the theoretical foundations of quantum mechanics. We take up this task in Chapter 3, where we deal with the postulates of the theory as well as their implications; the postulates are the bedrock on which the theory is built.

2.9 Solved Problems

Problem 2.1

Consider the states $\mid \psi\rangle = 9i \mid \phi_1\rangle + 2 \mid \phi_2\rangle$ and $\mid \chi\rangle = -\frac{i}{\sqrt{2}} \mid \phi_1\rangle + \frac{1}{\sqrt{2}} \mid \phi_2\rangle$, where the two vectors $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ form a complete and orthonormal basis.

(a) Calculate the operators $\mid \psi\rangle\langle\chi \mid$ and $\mid \chi\rangle\langle\psi \mid$. Are they equal?

(b) Find the Hermitian conjugates of $\mid \psi\rangle$, $\mid \chi\rangle$, $\mid \psi\rangle\langle\chi \mid$, and $\mid \chi\rangle\langle\psi \mid$.

(c) Calculate $\mathrm{Tr}(\mid \psi\rangle\langle\chi \mid)$ and $\mathrm{Tr}(\mid \chi\rangle\langle\psi \mid)$. Are they equal?

(d) Calculate $\mid \psi\rangle\langle\psi \mid$ and $\mid \chi\rangle\langle\chi \mid$ and the traces $\mathrm{Tr}(\mid \psi\rangle\langle\psi \mid)$ and $\mathrm{Tr}(\mid \chi\rangle\langle\chi \mid)$. Are they projection operators?

Solution

(a) The bras corresponding to $\mid \psi\rangle = 9i \mid \phi_1\rangle + 2 \mid \phi_2\rangle$ and $\mid \chi\rangle = -i \mid \phi_1\rangle/\sqrt{2} + \mid \phi_2\rangle/\sqrt{2}$ are given by $\langle\psi \mid = -9i\langle\phi_1 \mid + 2\langle\phi_2 \mid$ and $\langle\chi \mid = \frac{i}{\sqrt{2}}\langle\phi_1 \mid + \frac{1}{\sqrt{2}}\langle\phi_2 \mid$, respectively. Hence we have

$$\begin{aligned} \mid \psi\rangle\langle\chi \mid &= \frac{1}{\sqrt{2}}\left(9i \mid \phi_1\rangle + 2 \mid \phi_2\rangle\right)\left(i\langle\phi_1 \mid + \langle\phi_2 \mid\right) \\ &= \frac{1}{\sqrt{2}}\left(-9 \mid \phi_1\rangle\langle\phi_1 \mid + 9i \mid \phi_1\rangle\langle\phi_2 \mid + 2i \mid \phi_2\rangle\langle\phi_1 \mid + 2 \mid \phi_2\rangle\langle\phi_2 \mid\right), \end{aligned} \tag{2.355}$$

$$\mid \chi\rangle\langle\psi \mid = \frac{1}{\sqrt{2}}\left(-9 \mid \phi_1\rangle\langle\phi_1 \mid - 2i \mid \phi_1\rangle\langle\phi_2 \mid - 9i \mid \phi_2\rangle\langle\phi_1 \mid + 2 \mid \phi_2\rangle\langle\phi_2 \mid\right). \tag{2.356}$$

As expected, $\mid \psi\rangle\langle\chi \mid$ and $\mid \chi\rangle\langle\psi \mid$ are not equal; they would be equal only if the states $\mid \psi\rangle$ and $\mid \chi\rangle$ were proportional and the proportionality constant real.

(b) To find the Hermitian conjugates of $\mid \psi\rangle$, $\mid \chi\rangle$, $\mid \psi\rangle\langle\chi \mid$, and $\mid \chi\rangle\langle\psi \mid$, we need simply to replace the factors with their respective complex conjugates, the bras with kets, and the kets with bras:

$$\mid \psi\rangle^\dagger = \langle\psi \mid = -9i\langle\phi_1 \mid + 2\langle\phi_2 \mid, \qquad \mid \chi\rangle^\dagger = \langle\chi \mid = \frac{1}{\sqrt{2}}\left(i\langle\phi_1 \mid + \langle\phi_2 \mid\right), \tag{2.357}$$

$$(\mid \psi\rangle\langle\chi \mid)^\dagger = \mid \chi\rangle\langle\psi \mid \;=\; \frac{1}{\sqrt{2}}\left(-9 \mid \phi_1\rangle\langle\phi_1 \mid -2i \mid \phi_1\rangle\langle\phi_2 \mid \right. \\ \left. -\, 9i \mid \phi_2\rangle\langle\phi_1 \mid +2 \mid \phi_2\rangle\langle\phi_2 \mid\right), \qquad (2.358)$$

$$(\mid \chi\rangle\langle\psi \mid)^\dagger = \mid \psi\rangle\langle\chi \mid \;=\; \frac{1}{\sqrt{2}}\left(-9 \mid \phi_1\rangle\langle\phi_1 \mid +9i \mid \phi_1\rangle\langle\phi_2 \mid \right. \\ \left. +\, 2i \mid \phi_2\rangle\langle\phi_1 \mid +2 \mid \phi_2\rangle\langle\phi_2 \mid\right). \qquad (2.359)$$

(c) Using the property $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ and since $\langle\phi_1 \mid \phi_1\rangle = \langle\phi_2 \mid \phi_2\rangle = 1$ and $\langle\phi_1 \mid \phi_2\rangle = \langle\phi_2 \mid \phi_1\rangle = 0$, we obtain

$$\begin{aligned} \mathrm{Tr}(\mid \psi\rangle\langle\chi \mid) &= \mathrm{Tr}(\langle\chi \mid \psi\rangle) = \langle\chi \mid \psi\rangle \\ &= \left(\frac{i}{\sqrt{2}}\langle\phi_1 \mid + \frac{1}{\sqrt{2}}\langle\phi_2 \mid\right)(9i \mid \phi_1\rangle + 2 \mid \phi_2\rangle) = -\frac{7}{\sqrt{2}}, \end{aligned} \qquad (2.360)$$

$$\begin{aligned} \mathrm{Tr}(\mid \chi\rangle\langle\psi \mid) &= \mathrm{Tr}(\langle\psi \mid \chi\rangle) = \langle\psi \mid \chi\rangle \\ &= (-9i\langle\phi_1 \mid +2\langle\phi_2 \mid)\left(-\frac{i}{\sqrt{2}} \mid \phi_1\rangle + \frac{1}{\sqrt{2}} \mid \phi_2\rangle\right) = -\frac{7}{\sqrt{2}} \\ &= \mathrm{Tr}(\mid \psi\rangle\langle\chi \mid). \end{aligned} \qquad (2.361)$$

The traces $\mathrm{Tr}(\mid \psi\rangle\langle\chi \mid)$ and $\mathrm{Tr}(\mid \chi\rangle\langle\psi \mid)$ are equal only because the scalar product of $\mid \psi\rangle$ and $\mid \chi\rangle$ is a real number. Were this product a complex number, the traces would be different; in fact, they would be the complex conjugate of one another.

(d) The expressions $\mid \psi\rangle\langle\psi \mid$ and $\mid \chi\rangle\langle\chi \mid$ are

$$\begin{aligned} \mid \psi\rangle\langle\psi \mid &= (9i \mid \phi_1\rangle + 2 \mid \phi_2\rangle)(-9i\langle\phi_1 \mid +2\langle\phi_2 \mid) \\ &= 81 \mid \phi_1\rangle\langle\phi_1 \mid +18i \mid \phi_1\rangle\langle\phi_2 \mid -18i \mid \phi_2\rangle\langle\phi_1 \mid +4 \mid \phi_2\rangle\langle\phi_2 \mid, \end{aligned} \qquad (2.362)$$

$$\begin{aligned} \mid \chi\rangle\langle\chi \mid) &= \frac{1}{2}(\mid \phi_1\rangle\langle\phi_1 \mid -i \mid \phi_1\rangle\langle\phi_2 \mid +i \mid \phi_2\rangle\langle\phi_1 \mid + \mid \phi_2\rangle\langle\phi_2 \mid) \\ &= \frac{1}{2}(1 - i \mid \phi_1\rangle\langle\phi_2 \mid +i \mid \phi_2\rangle\langle\phi_1 \mid). \end{aligned} \qquad (2.363)$$

In deriving (2.363) we have used the fact that the basis is complete, $\mid \phi_1\rangle\langle\phi_1 \mid + \mid \phi_2\rangle\langle\phi_2 \mid = 1$. The traces $\mathrm{Tr}(\mid \psi\rangle\langle\psi \mid)$ and $\mathrm{Tr}(\mid \chi\rangle\langle\chi \mid)$ can then be calculated immediately:

$$\mathrm{Tr}(\mid \psi\rangle\langle\psi \mid) = \langle\psi \mid \psi\rangle = (-9i\langle\phi_1 \mid +2\langle\phi_2 \mid)(9i \mid \phi_1\rangle + 2 \mid \phi_2\rangle) = 85, \qquad (2.364)$$

$$\mathrm{Tr}(\mid \chi\rangle\langle\chi \mid) = \langle\chi \mid \chi\rangle = \frac{1}{2}(i\langle\phi_1 \mid +\langle\phi_2 \mid)(-i \mid \phi_1\rangle + \mid \phi_2\rangle) = 1. \qquad (2.365)$$

So $\mid \chi\rangle$ is normalized but $\mid \psi\rangle$ is not. Since $\mid \chi\rangle$ is normalized, we can easily ascertain that $\mid \chi\rangle\langle\chi \mid$ is a projection operator, because it is Hermitian, $(\mid \chi\rangle\langle\chi \mid)^\dagger = \mid \chi\rangle\langle\chi \mid$, and equal to its own square:

$$(\mid \chi\rangle\langle\chi \mid)^2 = \mid \chi\rangle\langle\chi \mid \chi\rangle\langle\chi \mid = (\langle\chi \mid \chi\rangle) \mid \chi\rangle\langle\chi \mid = \mid \chi\rangle\langle\chi \mid. \qquad (2.366)$$

As for $\mid \psi\rangle\langle\psi \mid$, although it is Hermitian, it cannot be a projection operator since $\mid \psi\rangle$ is not normalized. That is, $\mid \psi\rangle\langle\psi \mid$ is not equal to its own square:

$$(\mid \psi\rangle\langle\psi \mid)^2 = \mid \psi\rangle\langle\psi \mid \psi\rangle\langle\psi \mid = (\langle\psi \mid \psi\rangle) \mid \psi\rangle\langle\psi \mid = 85 \mid \psi\rangle\langle\psi \mid. \qquad (2.367)$$

Problem 2.2

(a) Find a complete and orthonormal basis for a space of the trigonometric functions of the form $\psi(\theta) = \sum_{n=0}^{N} a_n \cos(n\theta)$.

(b) Illustrate the results derived in (a) for the case $N = 5$; find the basis vectors.

Solution

(a) Since $\cos(n\theta) = \frac{1}{2}\left(e^{in\theta} + e^{-in\theta}\right)$, we can write $\sum_{n=0}^{N} a_n \cos(n\theta)$ as

$$\frac{1}{2}\sum_{n=0}^{N} a_n \left(e^{in\theta} + e^{-in\theta}\right) = \frac{1}{2}\left[\sum_{n=0}^{N} a_n e^{in\theta} + \sum_{n=-N}^{0} a_{-n} e^{in\theta}\right] = \sum_{n=-N}^{N} C_n e^{in\theta}, \qquad (2.368)$$

where $C_n = a_n/2$ for $n > 0$, $C_n = a_{-n}/2$ for $n < 0$, and $C_0 = a_0$. Since any trigonometric function of the form $\psi(x) = \sum_{n=0}^{N} a_n \cos(n\theta)$ can be expressed in terms of the functions $\phi_n(\theta) = e^{in\theta}/\sqrt{2\pi}$, we can try to take the set $\phi_n(\theta)$ as a basis. As this set is complete, let us see if it is orthonormal. The various functions $\phi_n(\theta)$ are indeed orthonormal, since their scalar products are given by

$$\langle \phi_m \mid \phi_n \rangle = \int_{-\pi}^{\pi} \phi_m^*(\theta)\phi_n(\theta)d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)\theta}\, d\theta = \delta_{nm}. \qquad (2.369)$$

In deriving this result, we have considered two cases: $n = m$ and $n \neq m$. First, the case $n = m$ is obvious, since $\langle \phi_n \mid \phi_n \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta = 1$. On the other hand, when $n \neq m$ we have

$$\langle \phi_m \mid \phi_n \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)\theta} d\theta = \frac{1}{2\pi}\frac{e^{i(n-m)\pi} - e^{-i(n-m)\pi}}{i(n-m)} = \frac{2i\sin((n-m)\pi)}{2i\pi(n-m)} = 0, \qquad (2.370)$$

since $\sin((n-m)\pi) = 0$. So the functions $\phi_n(\theta) = e^{in\theta}/\sqrt{2\pi}$ form a complete and orthonormal basis. From (2.368) we see that the basis has $2N+1$ functions $\phi_n(\theta)$; hence the dimension of this space of functions is equal to $2N+1$.

(b) In the case where $N = 5$, the dimension of the space is equal to 11, for the basis has 11 vectors: $\phi_{-5}(\theta) = e^{-5i\theta}/\sqrt{2\pi}$, $\phi_{-4}(\theta) = e^{-4i\theta}/\sqrt{2\pi}$, ..., $\phi_0(\theta) = 1/\sqrt{2\pi}$, ..., $\phi_4(\theta) = e^{4i\theta}/\sqrt{2\pi}$, $\phi_5(\theta) = e^{5i\theta}/\sqrt{2\pi}$.

Problem 2.3

(a) Show that the sum of two projection operators cannot be a projection operator unless their product is zero.

(b) Show that the product of two projection operators cannot be a projection operator unless they commute.

Solution

Recall that an operator $\hat{P}$ is a projection operator if it satisfies $\hat{P}^\dagger = \hat{P}$ and $\hat{P}^2 = \hat{P}$.

(a) If two operators $\hat{A}$ and $\hat{B}$ are projection operators and if $\hat{A}\hat{B} = \hat{B}\hat{A}$, we want to show that $(\hat{A}+\hat{B})^\dagger = \hat{A}+\hat{B}$ and that $(\hat{A}+\hat{B})^2 = \hat{A}+\hat{B}$. First, the hermiticity is easy to ascertain since $\hat{A}$ and $\hat{B}$ are both Hermitian: $(\hat{A}+\hat{B})^\dagger = \hat{A}+\hat{B}$. Let us now look at the square of $(\hat{A}+\hat{B})$; since $\hat{A}^2 = \hat{A}$ and $\hat{B}^2 = \hat{B}$, we can write

$$(\hat{A}+\hat{B})^2 = \hat{A}^2 + \hat{B}^2 + (\hat{A}\hat{B} + \hat{B}\hat{A}) = \hat{A} + \hat{B} + (\hat{A}\hat{B} + \hat{B}\hat{A}). \qquad (2.371)$$

Clearly, only when the product of $\hat{A}$ and $\hat{B}$ is zero will their sum be a projection operator.

(b) At issue here is to show that if two operators $\hat{A}$ and $\hat{B}$ are projection operators and if they commute, $[\hat{A},\ \hat{B}] = 0$, their product is a projection operator. That is, we need to show that $(\hat{A}\hat{B})^\dagger = \hat{A}\hat{B}$ and $(\hat{A}\hat{B})^2 = \hat{A}\hat{B}$. Again, since $\hat{A}$ and $\hat{B}$ are Hermitian and since they commute, we see that $(\hat{A}\hat{B})^\dagger = \hat{B}\hat{A} = \hat{A}\hat{B}$. As for the square of $\hat{A}\hat{B}$, we have

$$(\hat{A}\hat{B})^2 = (\hat{A}\hat{B})(\hat{A}\hat{B}) = \hat{A}(\hat{B}\hat{A})\hat{B} = \hat{A}(\hat{A}\hat{B})\hat{B} = \hat{A}^2\hat{B}^2 = \hat{A}\hat{B}, \tag{2.372}$$

hence the product $\hat{A}\hat{B}$ is a projection operator.

Problem 2.4

Consider a state $|\psi\rangle = \frac{1}{\sqrt{2}}|\phi_1\rangle + \frac{1}{\sqrt{5}}|\phi_2\rangle + \frac{1}{\sqrt{10}}|\phi_3\rangle$ which is given in terms of three orthonormal eigenstates $|\phi_1\rangle$, $|\phi_2\rangle$ and $|\phi_3\rangle$ of an operator $\hat{B}$ such that $\hat{B}|\phi_n\rangle = n^2|\phi_n\rangle$. Find the expectation value of $\hat{B}$ for the state $|\psi\rangle$.

Solution

Using Eq (2.58), we can write the expectation value of $\hat{B}$ for the state $|\psi\rangle$ as $\langle\hat{B}\rangle = \langle\psi \mid \hat{B} \mid \psi\rangle/\langle\psi \mid \psi\rangle$ where

$$\begin{aligned}\langle\psi \mid \psi\rangle &= \left(\frac{1}{\sqrt{2}}\langle\phi_1 \mid + \frac{1}{\sqrt{5}}\langle\phi_2 \mid + \frac{1}{\sqrt{10}}\langle\phi_3 \mid\right)\left(\frac{1}{\sqrt{2}}|\phi_1\rangle + \frac{1}{\sqrt{5}}|\phi_2\rangle + \frac{1}{\sqrt{10}}|\phi_3\rangle\right) \\ &= \frac{8}{10}\end{aligned} \tag{2.373}$$

and

$$\begin{aligned}\langle\psi \mid \hat{B} \mid \psi\rangle &= \left(\frac{1}{\sqrt{2}}\langle\phi_1 \mid + \frac{1}{\sqrt{5}}\langle\phi_2 \mid + \frac{1}{\sqrt{10}}\langle\phi_3 \mid\right)\hat{B}\left(\frac{1}{\sqrt{2}}|\phi_1\rangle + \frac{1}{\sqrt{5}}|\phi_2\rangle + \frac{1}{\sqrt{10}}|\phi_3\rangle\right) \\ &= \frac{1}{2} + \frac{2^2}{5} + \frac{3^2}{10} \\ &= \frac{22}{10}.\end{aligned} \tag{2.374}$$

Hence, the expectation value of $\hat{B}$ is given by

$$\langle\hat{B}\rangle = \frac{\langle\psi \mid \hat{B} \mid \psi\rangle}{\langle\psi \mid \psi\rangle} = \frac{22/10}{8/10} = \frac{11}{4}. \tag{2.375}$$

Problem 2.5

(a) Study the hermiticity of these operators: $\hat{X}$, d/dx, and id/dx. What about the complex conjugate of these operators? Are the Hermitian conjugates of the position and momentum operators equal to their complex conjugates?

(b) Use the results of (a) to discuss the hermiticity of the operators $e^{\hat{X}}$, $e^{d/dx}$, and $e^{id/dx}$.

(c) Find the Hermitian conjugate of the operator $\hat{X}d/dx$.

(d) Use the results of (a) to discuss the hermiticity of the components of the angular momentum operator (Chapter 5): $\hat{L}_x = -i\hbar\left(\hat{Y}\partial/\partial z - \hat{Z}\partial/\partial y\right)$, $\hat{L}_y = -i\hbar\left(\hat{Z}\partial/\partial x - \hat{X}\partial/\partial z\right)$, $\hat{L}_z = -i\hbar\left(\hat{X}\partial/\partial y - \hat{Y}\partial/\partial x\right)$.

Solution

(a) Using (2.69) and (2.70), and using the fact that the eigenvalues of $\hat{X}$ are real (i.e., $\hat{X}^* = \hat{X}$ or $x^* = x$), we can verify that $\hat{X}$ is Hermitian (i.e., $\hat{X}^\dagger = \hat{X}$) since

$$\begin{aligned}\langle\psi \mid \hat{X}\psi\rangle &= \int_{-\infty}^{+\infty} \psi^*(x)\,(x\psi(x))\,dx = \int_{-\infty}^{+\infty} \left(x\psi(x)^*\right)\psi(x)\,dx \\ &= \int_{-\infty}^{+\infty} (x\psi(x))^*\,\psi(x)\,dx = \langle\hat{X}\psi \mid \psi\rangle. \end{aligned} \tag{2.376}$$

Now, since $\psi(x)$ vanishes as $x \to \pm\infty$, an integration by parts leads to

$$\begin{aligned}\langle\psi \mid \frac{d}{dx}\psi\rangle &= \int_{-\infty}^{+\infty} \psi^*(x)\left(\frac{d\psi(x)}{dx}\right)dx = \psi^*(x)\psi(x)\Big|_{x=-\infty}^{x=+\infty} - \int_{-\infty}^{+\infty}\left(\frac{d\psi^*(x)}{dx}\right)\psi(x)\,dx \\ &= -\int_{-\infty}^{+\infty}\left(\frac{d\psi(x)}{dx}\right)^*\psi(x)\,dx = -\langle\frac{d}{dx}\psi \mid \psi\rangle. \end{aligned} \tag{2.377}$$

So, d/dx is anti-Hermitian: $(d/dx)^\dagger = -d/dx$. Since d/dx is anti-Hermitian, id/dx must be Hermitian, since $(id/dx)^\dagger = -i\,(-d/dx) = id/dx$. The results derived above are

$$\boxed{\hat{X}^\dagger = \hat{X}, \qquad \left(\frac{d}{dx}\right)^\dagger = -\frac{d}{dx}, \qquad \left(i\frac{d}{dx}\right)^\dagger = i\frac{d}{dx}.} \tag{2.378}$$

From this relation we see that the momentum operator $\hat{P} = -i\hbar d/dx$ is Hermitian: $\hat{P}^\dagger = \hat{P}$. We can also infer that, although the momentum operator is Hermitian, its complex conjugate is not equal to $\hat{P}$, since $\hat{P}^* = (-i\hbar d/dx)^* = i\hbar d/dx = -\hat{P}$. We may group these results into the following relation:

$$\boxed{\hat{X}^\dagger = \hat{X}, \quad \hat{X}^* = \hat{X}, \qquad \hat{P}^\dagger = \hat{P}, \quad \hat{P}^* = -\hat{P}.} \tag{2.379}$$

(b) Using the relations $(e^{\hat{A}})^\dagger = e^{\hat{A}^\dagger}$ and $(e^{i\hat{A}})^\dagger = e^{-i\hat{A}^\dagger}$ derived in (2.113), we infer

$$\boxed{(e^{\hat{X}})^\dagger = e^{\hat{X}}, \qquad (e^{d/dx})^\dagger = e^{-d/dx}, \qquad (e^{id/dx})^\dagger = e^{id/dx}.} \tag{2.380}$$

(c) Since $\hat{X}$ is Hermitian and d/dx is anti-Hermitian, we have

$$\left(\hat{X}\frac{d}{dx}\right)^\dagger = \left(\frac{d}{dx}\right)^\dagger (\hat{X})^\dagger = -\frac{d}{dx}\hat{X}, \tag{2.381}$$

where $d\hat{X}/dx$ is given by

$$\frac{d}{dx}\left(\hat{X}\psi(x)\right) = \left(1 + x\frac{d}{dx}\right)\psi(x); \tag{2.382}$$

hence

$$\left(\hat{X}\frac{d}{dx}\right)^\dagger = -\hat{X}\frac{d}{dx} - 1. \tag{2.383}$$

(d) From the results derived in (a), we infer that the operators $\hat{Y}$, $\hat{Z}$, $i\partial/\partial x$, and $i\partial/\partial y$ are Hermitian. We can verify that $\hat{L}_x$ is also Hermitian:

$$\hat{L}_x^\dagger = -i\hbar\left(\frac{\partial}{\partial z}\hat{Y} - \frac{\partial}{\partial y}\hat{Z}\right) = -i\hbar\left(\hat{Y}\frac{\partial}{\partial z} - \hat{Z}\frac{\partial}{\partial y}\right) = \hat{L}_x; \tag{2.384}$$

in deriving this relation, we used the fact that the y and z degrees of freedom commute (i.e., $\partial\hat{Y}/\partial z = \hat{Y}\partial/\partial z$ and $\partial\hat{Z}/\partial y = \hat{Z}\partial/\partial y$), for they are independent. Similarly, the hermiticity of $\hat{L}_y = -i\hbar\left(\hat{Z}\partial/\partial x - \hat{X}\partial/\partial z\right)$ and $\hat{L}_z = -i\hbar\left(\hat{X}\partial/\partial y - \hat{Y}\partial/\partial x\right)$ is obvious.

Problem 2.6

(a) Show that the operator $\hat{A} = i(\hat{X}^2 + 1)d/dx + i\hat{X}$ is Hermitian.

(b) Find the state $\psi(x)$ for which $\hat{A}\psi(x) = 0$ and normalize it.

(c) Calculate the probability of finding the particle (represented by $\psi(x)$) in the region: $-1 \leq x \leq 1$.

Solution

(a) From the previous problem we know that $\hat{X}^\dagger = \hat{X}$ and $(d/dx)^\dagger = -d/dx$. We can thus infer the Hermitian conjugate of $\hat{A}$:

$$\begin{aligned}\hat{A}^\dagger &= -i\left(\frac{d}{dx}\right)^\dagger(\hat{X}^2)^\dagger - i\left(\frac{d}{dx}\right)^\dagger - i\hat{X}^\dagger = i\left(\frac{d}{dx}\right)(\hat{X}^2) + i\left(\frac{d}{dx}\right) - i\hat{X} \\ &= i\hat{X}^2\frac{d}{dx} + i\left[\frac{d}{dx},\ \hat{X}^2\right] + i\frac{d}{dx} - i\hat{X}.\end{aligned} \tag{2.385}$$

Using the relation $[\hat{B},\ \hat{C}^2] = \hat{C}[\hat{B},\ \hat{C}] + [\hat{B},\ \hat{C}]\hat{C}$ along with $[d/dx,\ \hat{X}] = 1$, we can easily evaluate the commutator $[d/dx,\ \hat{X}^2]$:

$$\left[\frac{d}{dx},\ \hat{X}^2\right] = \hat{X}\left[\frac{d}{dx},\ \hat{X}\right] + \left[\frac{d}{dx},\ \hat{X}\right]\hat{X} = 2\hat{X}. \tag{2.386}$$

A combination of (2.385) and (2.386) shows that $\hat{A}$ is Hermitian:

$$\hat{A}^\dagger = i(\hat{X}^2 + 1)\frac{d}{dx} + i\hat{X} = \hat{A}. \tag{2.387}$$

(b) The state $\psi(x)$ for which $\hat{A}\psi(x) = 0$, i.e.,

$$i(\hat{X}^2 + 1)\frac{d\psi(x)}{dx} + i\hat{X}\psi(x) = 0, \tag{2.388}$$

corresponds to

$$\frac{d\psi(x)}{dx} = -\frac{x}{x^2 + 1}\psi(x). \tag{2.389}$$

The solution to this equation is given by

$$\psi(x) = \frac{B}{\sqrt{x^2 + 1}}. \tag{2.390}$$

Since $\int_{-\infty}^{+\infty} dx/(x^2+1) = \pi$ we have

$$1 = \int_{-\infty}^{+\infty} |\psi(x)|^2\, dx = B^2 \int_{-\infty}^{+\infty} \frac{dx}{x^2+1} = B^2\pi, \tag{2.391}$$

which leads to $B = 1/\sqrt{\pi}$ and hence $\psi(x) = \frac{1}{\sqrt{\pi(x^2+1)}}$.

(c) Using the integral $\int_{-1}^{+1} dx/(x^2+1) = \pi/2$, we can obtain the probability immediately:

$$P = \int_{-1}^{+1} |\psi(x)|^2\, dx = \frac{1}{\pi}\int_{-1}^{+1} \frac{dx}{x^2+1} = \frac{1}{2}. \tag{2.392}$$

Problem 2.7
Discuss the conditions for these operators to be unitary: (a) $(1+i\hat{A})/(1-i\hat{A})$,
(b) $(\hat{A}+i\hat{B})/\sqrt{\hat{A}^2+\hat{B}^2}$.

Solution
An operator $\hat{U}$ is unitary if $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}$ (see (2.122)).

(a) Since

$$\left(\frac{1+i\hat{A}}{1-i\hat{A}}\right)^\dagger = \frac{1-i\hat{A}^\dagger}{1+i\hat{A}^\dagger}, \tag{2.393}$$

we see that if $\hat{A}$ is Hermitian, the expression $(1+i\hat{A})/(1-i\hat{A})$ is unitary:

$$\left(\frac{1+i\hat{A}}{1-i\hat{A}}\right)^\dagger \frac{1+i\hat{A}}{1-i\hat{A}} = \frac{1-i\hat{A}}{1+i\hat{A}}\frac{1+i\hat{A}}{1-i\hat{A}} = \hat{I}. \tag{2.394}$$

(b) Similarly, if $\hat{A}$ and $\hat{B}$ are Hermitian and commute, the expression $(\hat{A}+i\hat{B})/\sqrt{\hat{A}^2+\hat{B}^2}$ is unitary:

$$\begin{aligned}\left(\frac{\hat{A}+i\hat{B}}{\sqrt{\hat{A}^2+\hat{B}^2}}\right)^\dagger \frac{\hat{A}+i\hat{B}}{\sqrt{\hat{A}^2+\hat{B}^2}} &= \frac{\hat{A}-i\hat{B}}{\sqrt{\hat{A}^2+\hat{B}^2}}\frac{\hat{A}+i\hat{B}}{\sqrt{\hat{A}^2+\hat{B}^2}} = \frac{\hat{A}^2+\hat{B}^2+i(\hat{A}\hat{B}-\hat{B}\hat{A})}{\hat{A}^2+\hat{B}^2} \\ &= \frac{\hat{A}^2+\hat{B}^2}{\hat{A}^2+\hat{B}^2} = \hat{I}. \end{aligned}\tag{2.395}$$

Problem 2.8

(a) Using the commutator $[\hat{X},\ \hat{p}] = i\hbar$, show that $[\hat{X}^m,\ \hat{P}] = im\hbar\hat{X}^{m-1}$, with $m > 1$. Can you think of a direct way to get to the same result?

(b) Use the result of (a) to show the general relation $[F(\hat{X}),\ \hat{P}] = i\hbar dF(\hat{X})/d\hat{X}$, where $F(\hat{X})$ is a differentiable operator function of $\hat{X}$.

Solution

(a) Let us attempt a proof by induction. Assuming that $[\hat{X}^m,\ \hat{P}] = im\hbar\hat{X}^{m-1}$ is valid for $m = k$ (note that it holds for $n = 1$; i.e., $[\hat{X},\ \hat{P}] = i\hbar$),

$$[\hat{X}^k,\ \hat{P}] = ik\hbar\hat{X}^{k-1}, \tag{2.396}$$

let us show that it holds for $m = k+1$:

$$[\hat{X}^{k+1},\ \hat{P}] = [\hat{X}^k\hat{X},\ \hat{P}] = \hat{X}^k[\hat{X},\ \hat{P}] + [\hat{X}^k,\ \hat{P}]\hat{X}, \tag{2.397}$$

where we have used the relation $[\hat{A}\hat{B},\ \hat{C}] = \hat{A}[\hat{B},\ \hat{C}] + [\hat{A},\ \hat{C}]\hat{B}$. Now, since $[\hat{X},\ \hat{P}] = i\hbar$ and $[\hat{X}^k,\ \hat{P}] = ik\hbar\hat{X}^{k-1}$, we rewrite (2.397) as

$$[\hat{X}^{k+1},\ \hat{P}] = i\hbar\hat{X}^k + (ik\hbar\hat{X}^{k-1})\hat{X} = i\hbar(k+1)\hat{X}^k. \tag{2.398}$$

So this relation is valid for any value of k, notably for $k = m-1$:

$$\boxed{[\hat{X}^m,\ \hat{P}] = im\hbar\hat{X}^{m-1}.} \tag{2.399}$$

In fact, it is easy to arrive at this result *directly* through brute force as follows. Using the relation $[\hat{A}^n, \hat{B}] = \hat{A}^{n-1}[\hat{A}, \hat{B}] + [\hat{A}^{n-1}, \hat{B}]\hat{A}$ along with $[\hat{X}, \hat{P}_x] = i\hbar$, we can obtain

$$[\hat{X}^2, \hat{P}_x] = \hat{X}[\hat{X}, \hat{P}_x] + [\hat{X}, \hat{P}_x]\hat{X} = 2i\hbar\hat{X}, \tag{2.400}$$

which leads to

$$[\hat{X}^3, \hat{P}_x] = \hat{X}^2[\hat{X}, \hat{P}_x] + [\hat{X}^2, \hat{P}_x]\hat{X} = 3i\hat{X}^2\hbar; \tag{2.401}$$

this in turn leads to

$$[\hat{X}^4, \hat{P}_x] = \hat{X}^3[\hat{X}, \hat{P}_x] + [\hat{X}^3, \hat{P}_x]\hat{X} = 4i\hat{X}^3\hbar. \tag{2.402}$$

Continuing in this way, we can get to any power of $\hat{X}$: $[\hat{X}^m,\ \hat{P}] = im\hbar\hat{X}^{m-1}$.

A more direct and simpler method is to apply the commutator $[\hat{X}^m,\ \hat{P}]$ on some wave function $\psi(x)$:

$$\begin{aligned}
[\hat{X}^m, \hat{P}_x]\psi(x) &= \left(\hat{X}^m\hat{P}_x - \hat{P}_x\hat{X}^m\right)\psi(x) \\
&= x^m\left(-i\hbar\frac{d\psi(x)}{dx}\right) + i\hbar\frac{d}{dx}\left(x^m\psi(x)\right) \\
&= x^m\left(-i\hbar\frac{d\psi(x)}{dx}\right) + im\hbar x^{m-1}\psi(x) - x^m\left(-i\hbar\frac{d\psi(x)}{dx}\right) \\
&= im\hbar x^{m-1}\psi(x).
\end{aligned} \tag{2.403}$$

Since $[\hat{X}^m, \hat{P}_x]\psi(x) = im\hbar x^{m-1}\psi(x)$ we see that $[\hat{X}^m,\ \hat{P}] = im\hbar\hat{X}^{m-1}$.

(b) Let us Taylor expand $F(\hat{X})$ in powers of $\hat{X}$, $F(\hat{X}) = \sum_k a_k\hat{X}^k$, and insert this expression into $[F(\hat{X}),\ \hat{P}]$:

$$\left[F(\hat{X}),\ \hat{P}\right] = \left[\sum_k a_k\hat{X}^k,\ \hat{P}\right] = \sum_k a_k[\hat{X}^k,\ \hat{P}], \tag{2.404}$$

where the commutator $[\hat{X}^k,\ \hat{P}]$ is given by (2.396). Thus, we have

$$\left[F(\hat{X}),\ \hat{P}\right] = i\hbar\sum_k k a_k \hat{X}^{k-1} = i\hbar\frac{d(\sum_k a_k\hat{X}^k)}{d\hat{X}} = i\hbar\frac{dF(\hat{X})}{d\hat{X}}. \tag{2.405}$$

A much simpler method again consists in applying the commutator $\left[F(\hat{X}),\ \hat{P}\right]$ on some wave function $\psi(x)$. Since $F(\hat{X})\psi(x) = F(x)\psi(x)$, we have

$$\begin{aligned}
\left[F(\hat{X}),\ \hat{P}\right]\psi(x) &= F(\hat{X})\hat{P}\psi(x) + i\hbar\frac{d}{dx}\left(F(x)\psi(x)\right) \\
&= F(\hat{X})\hat{P}\psi(x) - \left(-i\hbar\frac{d\psi(x)}{dx}\right)F(x) + i\hbar\frac{dF(x)}{dx}\psi(x) \\
&= F(\hat{X})\hat{P}\psi(x) - F(\hat{X})\hat{P}\psi(x) + i\hbar\frac{dF(x)}{dx}\psi(x) \\
&= i\hbar\frac{dF(x)}{dx}\psi(x).
\end{aligned} \tag{2.406}$$

Since $\left[F(\hat{X}),\ \hat{P}\right]\psi(x) = i\hbar\frac{dF(x)}{dx}\psi(x)$ we see that $\left[F(\hat{X}),\ \hat{P}\right] = i\hbar\frac{dF(\hat{X})}{d\hat{X}}$.

Problem 2.9

Consider the matrices $A = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2i & 0 \\ i & 0 & -5i \end{pmatrix}$.

(a) Are A and B Hermitian? Calculate AB and BA and verify that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$; then calculate $[A,\ B]$ and verify that $\mathrm{Tr}([A,\ B]) = 0$.

(b) Find the eigenvalues and the normalized eigenvectors of A. Verify that the sum of the eigenvalues of A is equal to the value of $\mathrm{Tr}(A)$ calculated in (a) and that the three eigenvectors form a basis.

(c) Verify that $U^\dagger AU$ is diagonal and that $U^{-1} = U^\dagger$, where U is the matrix formed by the normalized eigenvectors of A.

(d) Calculate the inverse of $A' = U^\dagger AU$ and verify that A'^{-1} is a diagonal matrix whose eigenvalues are the inverse of those of A'.

Solution

(a) Taking the Hermitian adjoints of the matrices A and B (see (2.188))

$$A^\dagger = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix}, \qquad B^\dagger = \begin{pmatrix} 1 & 0 & -i \\ 0 & -2i & 0 \\ 3 & 0 & 5i \end{pmatrix}, \tag{2.407}$$

we see that A is Hermitian and B is not. Using the products

$$AB = \begin{pmatrix} 7 & 0 & 21 \\ 1 & 2i & -5 \\ -i & -2 & 5i \end{pmatrix}, \quad BA = \begin{pmatrix} 7 & 3i & -3 \\ 0 & 2i & 2 \\ 7i & 5 & 5i \end{pmatrix}, \tag{2.408}$$

we can obtain the commutator

$$[A, B] = \begin{pmatrix} 0 & -3i & 24 \\ 1 & 0 & -7 \\ -8i & -7 & 0 \end{pmatrix}. \tag{2.409}$$

From (2.408) we see that

$$\mathrm{Tr}(AB) = 7 + 2i + 5i = 7 + 7i = \mathrm{Tr}(BA). \tag{2.410}$$

That is, the cyclic permutation of matrices leaves the trace unchanged; see (2.206). On the other hand, (2.409) shows that the trace of the commutator $[A,\ B]$ is zero: $\mathrm{Tr}([A,\ B]) = 0+0+0 = 0$.

(b) The eigenvalues and eigenvectors of A were calculated in Example 2.19 (see (2.266), (2.268), (2.272), (2.274)). We have $a_1 = 7$, $a_2 = \sqrt{2}$, and $a_3 = -\sqrt{2}$:

$$\mid a_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mid a_2\rangle = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2(2-\sqrt{2})}} \\ \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} \end{pmatrix}, \quad \mid a_3\rangle = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix}. \tag{2.411}$$

One can easily verify that the eigenvectors $\mid a_1\rangle$, $\mid a_2\rangle$, and $\mid a_3\rangle$ are mutually orthogonal: $\langle a_i \mid a_j\rangle = \delta_{ij}$ where $i, j = 1, 2, 3$. Since the set of $\mid a_1\rangle$, $\mid a_2\rangle$, and $\mid a_3\rangle$ satisfy the completeness condition

$$\sum_{j=1}^{3} \mid a_j\rangle\langle a_j \mid = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.412}$$

and since they are orthonormal, they form a complete and orthonormal basis.

(c) The columns of the matrix U are given by the eigenvectors (2.411):

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ 0 & \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix}. \tag{2.413}$$

We can show that the product $U^\dagger AU$ is diagonal where the diagonal elements are the eigenvalues of the matrix A; $U^\dagger AU$ is given by

$$\begin{aligned} &\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} \\ 0 & \frac{1}{\sqrt{2(2+\sqrt{2})}} & \frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix} \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ 0 & \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix} \\ &= \begin{pmatrix} 7 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}. \end{aligned} \tag{2.414}$$

We can also show that $U^\dagger U = 1$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} \\ 0 & \frac{1}{\sqrt{2(2+\sqrt{2})}} & \frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2(2-\sqrt{2})}} & \frac{1}{\sqrt{2(2+\sqrt{2})}} \\ 0 & \frac{i(\sqrt{2}-1)}{\sqrt{2(2-\sqrt{2})}} & -\frac{i(1+\sqrt{2})}{\sqrt{2(2+\sqrt{2})}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.415}$$

This implies that the matrix U is unitary: $U^\dagger = U^{-1}$. Note that, from (2.413), we have $\mid \det(U) \mid = \mid -i \mid = 1$.

(d) Using (2.414) we can verify that the inverse of $A' = U^\dagger A U$ is a diagonal matrix whose elements are given by the inverse of the diagonal elements of A':

$$A' = \begin{pmatrix} 7 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix} \quad \Longrightarrow \quad A'^{-1} = \begin{pmatrix} \frac{1}{7} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}. \tag{2.416}$$

Problem 2.10

Consider a particle whose Hamiltonian matrix is $H = \begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

(a) Is $\mid \lambda \rangle = \begin{pmatrix} i \\ 7i \\ -2 \end{pmatrix}$ an eigenstate of H? Is H Hermitian?

(b) Find the energy eigenvalues, a_1, a_2, and a_3, and the normalized energy eigenvectors, $\mid a_1\rangle$, $\mid a_2\rangle$, and $\mid a_3\rangle$, of H.

(c) Find the matrix corresponding to the operator obtained from the ket-bra product of the first eigenvector $P = \mid a_1\rangle\langle a_1 \mid$. Is P a projection operator? Calculate the commutator $[P,\ H]$ firstly by using commutator algebra and then by using matrix products.

Solution

(a) The ket $\mid \lambda\rangle$ is an eigenstate of H only if the action of the Hamiltonian on $\mid \lambda\rangle$ is of the form $H \mid \lambda\rangle = b \mid \lambda\rangle$, where b is constant. This is not the case here:

$$H \mid \lambda\rangle = \begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} i \\ 7i \\ -2 \end{pmatrix} = \begin{pmatrix} -7+2i \\ -1+7i \\ 7i \end{pmatrix}. \tag{2.417}$$

Using the definition of the Hermitian adjoint of matrices (2.188), it is easy to ascertain that H is Hermitian:

$$H^\dagger = \begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} = H. \tag{2.418}$$

(b) The energy eigenvalues can be obtained by solving the secular equation

$$\begin{aligned} 0 &= \begin{vmatrix} 2-a & i & 0 \\ -i & 1-a & 1 \\ 0 & 1 & -a \end{vmatrix} = (2-a)\left[(1-a)(-a)-1\right] - i(-i)(-a) \\ &= -(a-1)(a-1-\sqrt{3})(a-1+\sqrt{3}), \end{aligned} \tag{2.419}$$

which leads to

$$a_1 = 1, \quad a_2 = 1 - \sqrt{3}, \quad a_3 = 1 + \sqrt{3}. \tag{2.420}$$

To find the eigenvector corresponding to the first eigenvalue, $a_1 = 1$, we need to solve the matrix equation

$$\begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Longrightarrow \begin{array}{rcl} x + iy & = & 0 \\ -ix + z & = & 0 \\ y - z & = & 0 \end{array} \tag{2.421}$$

which yields $x = 1, y = z = i$. So the eigenvector corresponding to $a_1 = 1$ is

$$| a_1\rangle = \begin{pmatrix} 1 \\ i \\ i \end{pmatrix}. \tag{2.422}$$

This eigenvector is not normalized since $\langle a_1 \mid a_1\rangle = 1 + (i^*)(i) + (i^*)(i) = 3$. The normalized $| a_1\rangle$ is therefore

$$| a_1\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ i \\ i \end{pmatrix}. \tag{2.423}$$

Solving (2.421) for the other two energy eigenvalues, $a_2 = 1 - \sqrt{3}$, $a_3 = 1 + \sqrt{3}$, and normalizing, we end up with

$$| a_2\rangle = \frac{1}{\sqrt{6(2-\sqrt{3})}} \begin{pmatrix} i(2-\sqrt{3}) \\ 1-\sqrt{3} \\ 1 \end{pmatrix}, \qquad | a_3\rangle = \frac{1}{\sqrt{6(2+\sqrt{3})}} \begin{pmatrix} i(2+\sqrt{3}) \\ 1+\sqrt{3} \\ 1 \end{pmatrix}. \tag{2.424}$$

(c) The operator P is given by

$$P = | a_1\rangle\langle a_1 | = \frac{1}{3}\begin{pmatrix} 1 \\ i \\ i \end{pmatrix} \begin{pmatrix} 1 & -i & -i \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix}. \tag{2.425}$$

Since this matrix is Hermitian and since the square of P is equal to P,

$$P^2 = \frac{1}{9}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix} = P, \tag{2.426}$$

so P is a projection operator. Using the relations $H \mid a_1\rangle = \mid a_1\rangle$ and $\langle a_1 \mid H = \langle a_1 \mid$ (because H is Hermitian), and since $P = \mid a_1\rangle\langle a_1 \mid$, we can evaluate algebraically the commutator $[P,\ H]$ as follows:

$$[P,\ H] = PH - HP = | a_1\rangle\langle a_1 | H - H | a_1\rangle\langle a_1 | = | a_1\rangle\langle a_1 | - | a_1\rangle\langle a_1 | = 0. \tag{2.427}$$

We can reach the same result by using the matrices of H and P:

$$\begin{aligned} [P,\ H] &= \frac{1}{3}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 2 & i & 0 \\ -i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -i & -i \\ i & 1 & 1 \\ i & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{2.428}$$

Problem 2.11

Consider the matrices $A = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}$.

(a) Check if A and B are Hermitian and find the eigenvalues and eigenvectors of A. Any degeneracies?

(b) Verify that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, $\det(AB) = \det(A)\det(B)$, and $\det(B^\dagger) = (\det(B))^*$.

(c) Calculate the commutator $[A, B]$ and the anticommutator $\{A, B\}$.

(d) Calculate the inverses A^{-1}, B^{-1}, and $(AB)^{-1}$. Verify that $(AB)^{-1} = B^{-1}A^{-1}$.

(e) Calculate A^2 and infer the expressions of A^{2n} and A^{2n+1}. Use these results to calculate the matrix of e^{xA}.

Solution

(a) The matrix A is Hermitian but B is not. The eigenvalues of A are $a_1 = -1$ and $a_2 = a_3 = 1$ and its normalized eigenvectors are

$$|\, a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix}, \qquad |\, a_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -i \end{pmatrix}, \qquad |\, a_3\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{2.429}$$

Note that the eigenvalue 1 is doubly degenerate, since the two eigenvectors $|\, a_2\rangle$ and $|\, a_3\rangle$ correspond to the same eigenvalue $a_2 = a_3 = 1$.

(b) A calculation of the products (AB) and (BA) reveals that the traces $\mathrm{Tr}(AB)$ and $\mathrm{Tr}(BA)$ are equal:

$$\begin{aligned} \mathrm{Tr}(AB) &= \mathrm{Tr}\begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ -2i & 1 & 0 \end{pmatrix} = 1, \\ \mathrm{Tr}(BA) &= \mathrm{Tr}\begin{pmatrix} 0 & i & 2i \\ -5i & 1 & 3i \\ 2i & -i & 0 \end{pmatrix} = 1 = \mathrm{Tr}(AB). \end{aligned} \tag{2.430}$$

From the matrices A and B, we have $\det(A) = i(i) = -1$, $\det(B) = -4 + 16i$. We can thus write

$$\det(AB) = \det\begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ -2i & 1 & 0 \end{pmatrix} = 4 - 16i = (-1)(-4 + 16i) = \det(A)\det(B). \tag{2.431}$$

On the other hand, since $\det(B) = -4 + 16i$ and $\det(B^\dagger) = -4 - 16i$, we see that $\det(B^\dagger) = -4 - 16i = (-4 + 16i)^* = (\det(B))^*$.

(c) The commutator $[A,\ B]$ is given by

$$AB - BA = \begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ -2i & 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & i & 2i \\ -5i & 1 & 3i \\ 2i & -i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 - i & -4i \\ 3 + 5i & 0 & 5 - 3i \\ -4i & 1 + i & 0 \end{pmatrix} \tag{2.432}$$

and the anticommutator $\{A, B\}$ by

$$AB + BA = \begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ -2i & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & i & 2i \\ -5i & 1 & 3i \\ 2i & -i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1+i & 0 \\ 3-5i & 2 & 5+3i \\ 0 & 1-i & 0 \end{pmatrix}. \tag{2.433}$$

(d) A calculation similar to (2.200) leads to the inverses of A, B, and AB:

$$A^{-1} = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}, \qquad B^{-1} = \frac{1}{68}\begin{pmatrix} 22+3i & 8-2i & 20-5i \\ -6-24i & 4+16i & 10+40i \\ -12+3i & 8-2i & -14-5i \end{pmatrix}, \tag{2.434}$$

$$(AB)^{-1} = \frac{1}{68}\begin{pmatrix} -5-20i & 8-2i & -3+22i \\ 40-10i & 4+16i & 24-6i \\ -5+14i & 8-2i & -3-12i \end{pmatrix}. \tag{2.435}$$

From (2.434) it is now easy to verify that the product $B^{-1}A^{-1}$ is equal to $(AB)^{-1}$:

$$B^{-1}A^{-1} = \frac{1}{68}\begin{pmatrix} -5-20i & 8-2i & -3+22i \\ 40-10i & 4+16i & 24-6i \\ -5+14i & 8-2i & -3-12i \end{pmatrix} = (AB)^{-1}. \tag{2.436}$$

(e) Since

$$A^2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I, \tag{2.437}$$

we can write $A^3 = A$, $A^4 = I$, $A^5 = A$, and so on. We can generalize these relations to any value of n: $A^{2n} = I$ and $A^{2n+1} = A$:

$$A^{2n} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I, \qquad A^{2n+1} = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix} = A. \tag{2.438}$$

Since $A^{2n} = I$ and $A^{2n+1} = A$, we can write

$$e^{xA} = \sum_{n=0}^{\infty} \frac{x^n A^n}{n!} = \sum_{n=0}^{\infty} \frac{x^{2n} A^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{x^{2n+1} A^{2n+1}}{(2n+1)!} = I \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} + A \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}. \tag{2.439}$$

The relations

$$\sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} = \cosh x, \qquad \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} = \sinh x, \tag{2.440}$$

lead to

$$\begin{aligned} e^{xA} &= I\cosh x + A\sinh x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\cosh x + \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}\sinh x \\ &= \begin{pmatrix} \cosh x & 0 & i\sinh x \\ 0 & \cosh x + \sinh x & 0 \\ -i\sinh x & 0 & \cosh x \end{pmatrix}. \end{aligned} \tag{2.441}$$

Problem 2.12

Consider two matrices: $A = \begin{pmatrix} 0 & i & 2 \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}$. Calculate $A^{-1}\ B$ and $B\ A^{-1}$. Are they equal?

Solution

As mentioned above, a calculation similar to (2.200) leads to the inverse of A:

$$A^{-1} = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ 1/2 & -i/2 & 0 \end{pmatrix}. \tag{2.442}$$

The products $A^{-1}\ B$ and $B\ A^{-1}$ are given by

$$A^{-1}B = \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ 1/2 & -i/2 & 0 \end{pmatrix} \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -2i \\ 3 & 1 & 5 \\ 1-3i/2 & 0 & -5i/2 \end{pmatrix}, \tag{2.443}$$

$$BA^{-1} = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix} \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ 1/2 & -i/2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i & 2i \\ 5/2 & 1-5i/2 & 3i \\ -1 & 0 & 0 \end{pmatrix}. \tag{2.444}$$

We see that $A^{-1}\ B$ and $B\ A^{-1}$ are not equal.

Remark

We should note that the quotient B/A of two matrices A and B is equal to the product BA^{-1} and not $A^{-1}B$; that is:

$$\frac{B}{A} = BA^{-1} = \frac{\begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}}{\begin{pmatrix} 0 & i & 2 \\ 0 & 1 & 0 \\ -i & 0 & 0 \end{pmatrix}} = \begin{pmatrix} 0 & i & 2i \\ 5/2 & 1-5i/2 & 3i \\ -1 & 0 & 0 \end{pmatrix}. \tag{2.445}$$

Problem 2.13

Consider the matrices $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

(a) Find the eigenvalues and normalized eigenvectors of A and B. Denote the eigenvectors of A by $\mid a_1\rangle$, $\mid a_2\rangle$, $\mid a_3\rangle$ and those of B by $\mid b_1\rangle$, $\mid b_2\rangle$, $\mid b_3\rangle$. Are there any degenerate eigenvalues?

(b) Show that each of the sets $\mid a_1\rangle, \mid a_2\rangle, \mid a_3\rangle$ and $\mid b_1\rangle, \mid b_2\rangle, \mid b_3\rangle$ forms an orthonormal and complete basis, i.e., show that $\langle a_j \mid a_k\rangle = \delta_{jk}$ and $\sum_{j=1}^{3} \mid a_j\rangle\langle a_j \mid = I$, where I is the 3×3 unit matrix; then show that the same holds for $\mid b_1\rangle, \mid b_2\rangle, \mid b_3\rangle$.

(c) Find the matrix U of the transformation from the basis $\{\mid a\rangle\}$ to $\{\mid b\rangle\}$. Show that $U^{-1} = U^{\dagger}$. Verify that $U^{\dagger}U = I$. Calculate how the matrix A transforms under U, i.e., calculate $A' = UAU^{\dagger}$.

Solution

(a) It is easy to verify that the eigenvalues of A are $a_1 = 0$, $a_2 = \sqrt{2}$, $a_3 = -\sqrt{2}$ and their corresponding normalized eigenvectors are

$$| a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \qquad | a_2\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}, \qquad | a_3\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}. \tag{2.446}$$

The eigenvalues of B are $b_1 = 1$, $b_2 = 0$, $b_3 = -1$ and their corresponding normalized eigenvectors are

$$| b_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad | b_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad | b_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.447}$$

None of the eigenvalues of A and B are degenerate.

(b) The set $| a_1\rangle$, $| a_2\rangle$, $| a_3\rangle$ is indeed complete because the sum of $| a_1\rangle\langle a_1 |$, $| a_2\rangle\langle a_2 |$, and $| a_3\rangle\langle a_3 |$ as given by

$$| a_1\rangle\langle a_1 | = \frac{1}{2}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \tag{2.448}$$

$$| a_2\rangle\langle a_2 | = \frac{1}{4}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}\begin{pmatrix} 1 & \sqrt{2} & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 2 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}, \tag{2.449}$$

$$| a_3\rangle\langle a_3 | = \frac{1}{4}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}\begin{pmatrix} 1 & -\sqrt{2} & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & -\sqrt{2} & 1 \\ -\sqrt{2} & 2 & -\sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix}, \tag{2.450}$$

is equal to unity:

$$\begin{aligned} \sum_{j=1}^{3} | a_j\rangle\langle a_j | &= \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 2 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix} \\ &\quad + \frac{1}{4}\begin{pmatrix} 1 & -\sqrt{2} & 1 \\ -\sqrt{2} & 2 & -\sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{2.451}$$

The states $| a_1\rangle$, $| a_2\rangle$, $| a_3\rangle$ are orthonormal, since $\langle a_1 | a_2\rangle = \langle a_1 | a_3\rangle = \langle a_3 | a_2\rangle = 0$ and $\langle a_1 | a_1\rangle = \langle a_2 | a_2\rangle = \langle a_3 | a_3\rangle = 1$. Following the same procedure, we can ascertain that

$$| b_1\rangle\langle b_1 | + | b_2\rangle\langle b_2 | + | b_3\rangle\langle b_3 | = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.452}$$

We can verify that the states $\mid b_1\rangle$, $\mid b_2\rangle$, $\mid b_3\rangle$ are orthonormal, since $\langle b_1 \mid b_2\rangle = \langle b_1 \mid b_3\rangle = \langle b_3 \mid b_2\rangle = 0$ and $\langle b_1 \mid b_1\rangle = \langle b_2 \mid b_2\rangle = \langle b_3 \mid b_3\rangle = 1$.

(c) The elements of the matrix U, corresponding to the transformation from the basis $\{\mid a\rangle\}$ to $\{\mid b\rangle\}$, are given by $U_{jk} = \langle b_j \mid a_k\rangle$ where $j,\ k = 1, 2, 3$:

$$U = \begin{pmatrix} \langle b_1 \mid a_1\rangle & \langle b_1 \mid a_2\rangle & \langle b_1 \mid a_3\rangle \\ \langle b_2 \mid a_1\rangle & \langle b_2 \mid a_2\rangle & \langle b_2 \mid a_3\rangle \\ \langle b_3 \mid a_1\rangle & \langle b_3 \mid a_2\rangle & \langle b_3 \mid a_3\rangle \end{pmatrix}, \tag{2.453}$$

where the elements $\langle b_j \mid a_k\rangle$ can be calculated from (2.446) and (2.447):

$$U_{11} = \langle b_1 \mid a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = -\frac{\sqrt{2}}{2}, \tag{2.454}$$

$$U_{12} = \langle b_1 \mid a_2\rangle = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} = \frac{1}{2}, \tag{2.455}$$

$$U_{13} = \langle b_1 \mid a_3\rangle = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} = \frac{1}{2}, \tag{2.456}$$

$$U_{21} = \langle b_2 \mid a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = 0, \tag{2.457}$$

$$U_{22} = \langle b_2 \mid a_2\rangle = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} = \frac{\sqrt{2}}{2}, \tag{2.458}$$

$$U_{23} = \langle b_2 \mid a_3\rangle = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} = -\frac{\sqrt{2}}{2}, \tag{2.459}$$

$$U_{31} = \langle b_3 \mid a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \frac{\sqrt{2}}{2}, \tag{2.460}$$

$$U_{32} = \langle b_3 \mid a_2\rangle = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} = \frac{1}{2}, \tag{2.461}$$

$$U_{33} = \langle b_3 \mid a_3\rangle = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} = \frac{1}{2}. \tag{2.462}$$

Collecting these elements, we obtain

$$U = \frac{1}{2}\begin{pmatrix} -\sqrt{2} & 1 & 1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 & 1 \end{pmatrix}. \tag{2.463}$$

Calculating the inverse of U as we did in (2.200), we see that it is equal to its Hermitian adjoint:

$$U^{-1} = \frac{1}{2}\begin{pmatrix} -\sqrt{2} & 0 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \\ 1 & -\sqrt{2} & 1 \end{pmatrix} = U^{\dagger}. \tag{2.464}$$

This implies that the matrix U is unitary. The matrix A transforms as follows:

$$\begin{aligned} A' &= UAU^{\dagger} = \frac{1}{4}\begin{pmatrix} -\sqrt{2} & 1 & 1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} -\sqrt{2} & 0 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \\ 1 & -\sqrt{2} & 1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1-\sqrt{2} & -1 & 1 \\ -1 & -2 & 1 \\ 1 & 1 & 1+\sqrt{2} \end{pmatrix}. \end{aligned} \tag{2.465}$$

Problem 2.14

Calculate the following expressions involving Dirac's delta function:

(a) $\int_{-5}^{5} \cos(3x)\delta(x-\pi/3)\,dx$

(b) $\int_0^{10} \left[e^{2x-7}+4\right]\delta(x+3)\,dx$

(c) $\left[2\cos^2(3x)-\sin(x/2)\right]\delta(x+\pi)$

(d) $\int_0^{\pi} \cos(3\theta)\delta'''(\theta-\pi/2)\,d\theta$

(e) $\int_2^9 \left(x^2-5x+2\right)\delta[2(x-4)]\,dx$.

Solution

(a) Since $x=\pi/3$ lies within the interval $(-5,\ 5)$, equation (2.281) yields

$$\int_{-5}^{5}\cos(3x)\delta(x-\pi/3)\,dx = \cos\left(3\frac{\pi}{3}\right) = -1. \tag{2.466}$$

(b) Since $x=-3$ lies outside the interval $(0,\ 10)$, Eq (2.281) yields at once

$$\int_0^{10}\left[e^{2x-7}+4\right]\delta(x+3)\,dx = 0. \tag{2.467}$$

(c) Using the relation $f(x)\delta(x-a) = f(a)\delta(x-a)$ which is listed in Appendix A, we have

$$\begin{aligned} \left[2\cos^2(3x)-\sin(x/2)\right]\delta(x+\pi) &= \left[2\cos^2(3(-\pi))-\sin((-\pi)/2)\right]\delta(x+\pi) \\ &= 3\delta(x+\pi). \end{aligned} \tag{2.468}$$

(d) Inserting $n=3$ into Eq (2.282) and since $\cos'''(3\theta) = 27\sin(3\theta)$, we obtain

$$\begin{aligned} \int_0^{\pi}\cos(3\theta)\delta'''(\theta-\pi/2)\,d\theta &= (-1)^3\cos'''(3\pi/2) = (-1)^3\,27\sin(3\pi/2) \\ &= 27. \end{aligned} \tag{2.469}$$

(e) Since $\delta[2(x-4)] = (1/2)\delta(x-4)$, we have

$$\begin{aligned}\int_2^9 \left(x^2 - 5x + 2\right)\delta[2(x-4)], dx &= \frac{1}{2}\int_2^9 \left(x^2 - 5x + 2\right)\delta(x-4)\,dx \\ &= \frac{1}{2}\left(4^2 - 5\times 4 + 2\right) = -1. \end{aligned} \tag{2.470}$$

Problem 2.15

Consider a system whose Hamiltonian is given by $\hat{H} = \alpha\,(\mid \phi_1\rangle\langle\phi_2 \mid + \mid \phi_2\rangle\langle\phi_1 \mid)$, where α is a real number having the dimensions of energy and $\mid \phi_1\rangle$, $\mid \phi_2\rangle$ are normalized eigenstates of a Hermitian operator $\hat{A}$ that has no degenerate eigenvalues.

(a) Is $\hat{H}$ a projection operator? What about $\alpha^{-2}\hat{H}^2$?

(b) Show that $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ are not eigenstates of $\hat{H}$.

(c) Calculate the commutators $[\hat{H},\ \mid \phi_1\rangle\langle\phi_1 \mid]$ and $[\hat{H},\ \mid \phi_2\rangle\langle\phi_2 \mid]$ then find the relation that may exist between them.

(d) Find the normalized eigenstates of $\hat{H}$ and their corresponding energy eigenvalues.

(e) Assuming that $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ form a complete and orthonormal basis, find the matrix representing $\hat{H}$ in the basis. Find the eigenvalues and eigenvectors of the matrix and compare the results with those derived in (d).

Solution

(a) Since $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ are eigenstates of $\hat{A}$ and since $\hat{A}$ is Hermitian, they must be orthogonal, $\langle\phi_1 \mid \phi_2\rangle = 0$ (instance of Theorem 2.1). Now, since $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ are both normalized and since $\langle\phi_1 \mid \phi_2\rangle = 0$, we can reduce $\hat{H}^2$ to

$$\begin{aligned}\hat{H}^2 &= \alpha^2\,(\mid \phi_1\rangle\langle\phi_2 \mid + \mid \phi_2\rangle\langle\phi_1 \mid)\,(\mid \phi_1\rangle\langle\phi_1 \mid + \mid \phi_2\rangle\langle\phi_2 \mid) \\ &= \alpha^2\,(\mid \phi_1\rangle\langle\phi_2 \mid + \mid \phi_2\rangle\langle\phi_1 \mid)\,, \end{aligned} \tag{2.471}$$

which is different from $\hat{H}$; hence $\hat{H}$ is not a projection operator. The operator $\alpha^{-2}\hat{H}^2$ is a projection operator since it is both Hermitian and equal to its own square. Using (2.471) we can write

$$\begin{aligned}(\alpha^{-2}\hat{H}^2)^2 &= (\mid \phi_1\rangle\langle\phi_2 \mid + \mid \phi_2\rangle\langle\phi_1 \mid)\,(\mid \phi_1\rangle\langle\phi_2 \mid + \mid \phi_2\rangle\langle\phi_1 \mid) \\ &= \mid \phi_1\rangle\langle\phi_1 \mid + \mid \phi_2\rangle\langle\phi_2 \mid = \alpha^{-2}\hat{H}^2. \end{aligned} \tag{2.472}$$

(b) Since $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ are both normalized, and since $\langle\phi_1 \mid \phi_2\rangle = 0$, we have

$$\hat{H} \mid \phi_1\rangle = \alpha \mid \phi_1\rangle\langle\phi_2 \mid \phi_1\rangle + \alpha \mid \phi_2\rangle\langle\phi_1 \mid \phi_1\rangle = \alpha \mid \phi_2\rangle, \tag{2.473}$$

$$\hat{H} \mid \phi_2\rangle = \alpha \mid \phi_1\rangle; \tag{2.474}$$

hence $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ are not eigenstates of $\hat{H}$. In addition, we have

$$\langle\phi_1 \mid \hat{H} \mid \phi_1\rangle = \langle\phi_2 \mid \hat{H} \mid \phi_2\rangle = 0. \tag{2.475}$$

(c) Using the relations derived above, $\hat{H} \mid \phi_1\rangle = \alpha \mid \phi_2\rangle$ and $\hat{H} \mid \phi_2\rangle = \alpha \mid \phi_1\rangle$, we can write

$$[\hat{H}, \mid \phi_1\rangle\langle\phi_1 \mid] = \alpha\,(\mid \phi_2\rangle\langle\phi_1 \mid - \mid \phi_1\rangle\langle\phi_2 \mid)\,, \tag{2.476}$$

$$[\hat{H}, \mid \phi_2\rangle\langle\phi_2 \mid] = \alpha \left(\mid \phi_1\rangle\langle\phi_2 \mid - \mid \phi_2\rangle\langle\phi_1 \mid\right); \tag{2.477}$$

hence

$$[\hat{H}, \mid \phi_1\rangle\langle\phi_1 \mid] = -[\hat{H}, \mid \phi_2\rangle\langle\phi_2 \mid]. \tag{2.478}$$

(d) Consider a general state $\mid \psi\rangle = \lambda_1 \mid \phi_1\rangle + \lambda_2 \mid \phi_2\rangle$. Applying $\hat{H}$ to this state, we get

$$\begin{aligned} \hat{H} \mid \psi\rangle &= \alpha \left(\mid \phi_1\rangle\langle\phi_2 \mid + \mid \phi_2\rangle\langle\phi_1 \mid\right) \left(\lambda_1 \mid \phi_1\rangle + \lambda_2 \mid \phi_2\rangle\right) \\ &= \alpha \left(\lambda_2 \mid \phi_1\rangle + \lambda_1 \mid \phi_2\rangle\right). \end{aligned} \tag{2.479}$$

Now, since $\mid \psi\rangle$ is normalized, we have

$$\langle\psi \mid \psi\rangle = \mid \lambda_1 \mid^2 + \mid \lambda_2 \mid^2 = 1. \tag{2.480}$$

The previous two equations show that $\mid \lambda_1 \mid = \mid \lambda_2 \mid = 1/\sqrt{2}$ and that $\lambda_1 = \pm\lambda_2$. Hence the eigenstates of the system are:

$$\mid \psi_\pm\rangle = \frac{1}{\sqrt{2}} \left(\mid \phi_1\rangle \pm \mid \phi_2\rangle\right). \tag{2.481}$$

The corresponding eigenvalues are $\pm\alpha$:

$$\hat{H} \mid \psi_\pm\rangle = \pm\alpha \mid \psi_\pm\rangle. \tag{2.482}$$

(e) Since $\langle\phi_1 \mid \phi_2\rangle = \langle\phi_2 \mid \phi_1\rangle = 0$ and $\langle\phi_1 \mid \phi_1\rangle = \langle\phi_2 \mid \phi_2\rangle = 1$, we can verify that $H_{11} = \langle\phi_1 \mid \hat{H} \mid \phi_1\rangle = 0$, $H_{22} = \langle\phi_2 \mid \hat{H} \mid \phi_2\rangle = 0$, $H_{12} = \langle\phi_1 \mid \hat{H} \mid \phi_2\rangle = \alpha$, $H_{21} = \langle\phi_2 \mid \hat{H} \mid \phi_1\rangle = \alpha$. The matrix of $\hat{H}$ is thus given by

$$H = \alpha \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.483}$$

The eigenvalues of this matrix are equal to $\pm\alpha$ and the corresponding eigenvectors are $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$. These results are indeed similar to those derived in (d).

Problem 2.16

Consider the matrices $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & -3i \\ 0 & 3i & 5 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & -i & 3i \\ -i & 0 & i \\ 3i & i & 0 \end{pmatrix}$.

(a) Check the hermiticity of A and B.

(b) Find the eigenvalues of A and B; denote the eigenvalues of A by a_1, a_2, and a_3. Explain why the eigenvalues of A are real and those of B are imaginary.

(c) Calculate $\mathrm{Tr}(A)$ and $\det(A)$. Verify $\mathrm{Tr}(A) = a_1 + a_2 + a_3$, $\det(A) = a_1 a_2 a_3$.

Solution

(a) Matrix A is Hermitian but B is anti-Hermitian:

$$A^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & -3i \\ 0 & 3i & 5 \end{pmatrix} = A, \qquad B^\dagger = \begin{pmatrix} 0 & i & -3i \\ i & 0 & -i \\ -3i & -i & 0 \end{pmatrix} = -B. \tag{2.484}$$

(b) The eigenvalues of A are $a_1 = 6 - \sqrt{10}$, $a_2 = 1$, and $a_3 = 6 + \sqrt{10}$ and those of B are $b_1 = -i\left(3 + \sqrt{17}\right)/2$, $b_2 = 3i$, and $b_3 = i\left(-3 + \sqrt{17}\right)/2$. The eigenvalues of A are real and those of B are imaginary. This is expected since, as shown in (2.74) and (2.75), the expectation values of Hermitian operators are real and those of anti-Hermitian operators are imaginary.

(c) A direct calculation of the trace and the determinant of A yields $\mathrm{Tr}(A) = 1+7+5 = 13$ and $\det(A) = (7)(5)-(3i)(-3i) = 26$. Adding and multiplying the eigenvalues $a_1 = 6-\sqrt{10}$, $a_2 = 1$, $a_3 = 6 + \sqrt{10}$, we have $a_1 + a_2 + a_3 = 6 - \sqrt{10} + 1 + 6 + \sqrt{10} = 13$ and $a_1 a_2 a_3 = (6 - \sqrt{10})(1)(6 + \sqrt{10}) = 26$. This confirms the results (2.260) and (2.261):

$$\mathrm{Tr}(A) = a_1 + a_2 + a_3 = 13, \qquad \det(A) = a_1 a_2 a_3 = 26. \tag{2.485}$$

Problem 2.17

Consider a one-dimensional particle which moves along the x-axis and whose Hamiltonian is $\hat{H} = -\mathcal{E} d^2/dx^2 + 16\mathcal{E}\hat{X}^2$, where $\mathcal{E}$ is a real constant having the dimensions of energy.

(a) Is $\psi(x) = Ae^{-2x^2}$, where A is a normalization constant that needs to be found, an eigenfunction of $\hat{H}$? If yes, find the energy eigenvalue.

(b) Calculate the probability of finding the particle anywhere along the negative x-axis.

(c) Find the energy eigenvalue corresponding to the wave function $\phi(x) = 2x\psi(x)$.

(d) Specify the parities of $\phi(x)$ and $\psi(x)$. Are $\phi(x)$ and $\psi(x)$ orthogonal?

Solution

(a) The integral $\int_{-\infty}^{+\infty} e^{-4x^2}\,dx = \sqrt{\pi}/2$ allows us to find the normalization constant:

$$1 = \int_{-\infty}^{+\infty} |\psi(x)|^2\,dx = A^2 \int_{-\infty}^{+\infty} e^{-4x^2}\,dx = A^2\frac{\sqrt{\pi}}{2}; \tag{2.486}$$

this leads to $A = \sqrt{2/\sqrt{\pi}}$ and hence $\psi(x) = \sqrt{2/\sqrt{\pi}}e^{-2x^2}$. Since the first and second derivatives of $\psi(x)$ are given by

$$\psi'(x) = \frac{d\psi(x)}{dx} = -4x\psi(x), \qquad \psi''(x) = \frac{d^2\psi(x)}{dx^2} = (16x^2 - 4)\psi(x), \tag{2.487}$$

we see that $\psi(x)$ is an eigenfunction of $\hat{H}$ with an energy eigenvalue equal to $4\mathcal{E}$:

$$\hat{H}\psi(x) = -\mathcal{E}\frac{d^2\psi(x)}{dx^2} + 16\mathcal{E}x^2\psi(x) = -\mathcal{E}(16x^2-4)\psi(x) + 16\mathcal{E}x^2\psi(x) = 4\mathcal{E}\psi(x). \tag{2.488}$$

(b) Since $\int_{-\infty}^{0} e^{-4x^2}\,dx = \sqrt{\pi}/4$, the probability of finding the particle anywhere along the negative x-axis is equal to $\frac{1}{2}$:

$$\int_{-\infty}^{0} |\psi(x)|^2\,dx = \frac{2}{\sqrt{\pi}}\int_{-\infty}^{0} e^{-4x^2}\,dx = \frac{1}{2}. \tag{2.489}$$

This is expected, since this probability is half the total probability, which in turn is equal to one.

(c) Since the second derivative of $\phi(x) = 2x\psi(x)$ is $\phi''(x) = 4\psi'(x) + 2x\psi''(x) = 8x(-3+4x^2)\psi(x) = 4(-3+4x^2)\phi(x)$, we see that $\phi(x)$ is an eigenfunction of $\hat{H}$ with an energy eigenvalue equal to $12\mathcal{E}$:

$$\hat{H}\phi(x) = -\mathcal{E}\frac{d^2\phi(x)}{dx^2} + 16\mathcal{E}x^2\phi(x) = -4\mathcal{E}(-3+4x^2)\phi(x) + 16\mathcal{E}x^2\phi(x) = 12\mathcal{E}\phi(x). \tag{2.490}$$

(d) The wave functions $\psi(x)$ and $\phi(x)$ are even and odd, respectively, since $\psi(-x) = \psi(x)$ and $\phi(-x) = -\phi(x)$; hence their product is an odd function. Therefore, they are orthogonal, since the symmetric integration of an odd function is zero:

$$\begin{aligned}\langle \phi \mid \psi \rangle &= \int_{-\infty}^{+\infty} \phi^*(x)\psi(x)\,dx = \int_{-\infty}^{+\infty} \phi(x)\psi(x)\,dx = \int_{+\infty}^{-\infty} \phi(-x)\psi(-x)(-dx) \\ &= -\int_{-\infty}^{+\infty} \phi(x)\psi(x)\,dx = 0.\end{aligned} \tag{2.491}$$

Problem 2.18

(a) Find the eigenvalues and the eigenfunctions of the operator $\hat{A} = -d^2/dx^2$; restrict the search for the eigenfunctions to those complex functions that vanish everywhere except in the region $0 < x < a$.

(b) Normalize the eigenfunction and find the probability in the region $0 < x < a/2$.

Solution

(a) The eigenvalue problem for $-d^2/dx^2$ consists of solving the differential equation

$$-\frac{d^2\psi(x)}{dx^2} = \alpha\psi(x) \tag{2.492}$$

and finding the eigenvalues α and the eigenfunction $\psi(x)$. The most general solution to this equation is

$$\psi(x) = Ae^{ibx} + Be^{-ibx}, \tag{2.493}$$

with $\alpha = b^2$. Using the boundary conditions of $\psi(x)$ at $x = 0$ and $x = a$, we have

$$\psi(0) = A + B = 0 \implies B = -A, \qquad \psi(a) = Ae^{iba} + Be^{-iba} = 0. \tag{2.494}$$

A substitution of $B = -A$ into the second equation leads to $A\left(e^{iba} - e^{-iba}\right) = 0$ or $e^{iba} = e^{-iba}$ which leads to $e^{2iba} = 1$. Thus, we have $\sin 2ba = 0$ and $\cos 2ba = 1$, so $ba = n\pi$. The eigenvalues are then given by $\alpha_n = n^2\pi^2/a^2$ and the corresponding eigenvectors by $\psi_n(x) = A\left(e^{in\pi x/a} - e^{-in\pi x/a}\right)$; that is,

$$\alpha_n = \frac{n^2\pi^2}{a^2}, \qquad \psi_n(x) = C_n \sin\left(\frac{n\pi x}{a}\right). \tag{2.495}$$

So the eigenvalue spectrum of the operator $\hat{A} = -d^2/dx^2$ is discrete, because the eigenvalues and eigenfunctions depend on a discrete number n.

(b) The normalization of $\psi_n(x)$,

$$1 = C_n^2 \int_0^a \sin^2\left(\frac{n\pi x}{a}\right) dx = \frac{C_n^2}{2}\int_0^a \left[1 - \cos\left(\frac{2n\pi x}{a}\right)\right] dx = \frac{C_n^2}{2}a, \tag{2.496}$$

yields $C_n = \sqrt{2/a}$ and hence $\psi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$. The probability in the region $0 < x < a/2$ is given by

$$\frac{2}{a}\int_0^{a/2} \sin^2\left(\frac{n\pi x}{a}\right) dx = \frac{1}{a}\int_0^{a/2}\left[1-\cos\left(\frac{2n\pi x}{a}\right)\right] dx = \frac{1}{2}. \tag{2.497}$$

This is expected since the total probability is 1: $\int_0^a \mid \psi_n(x) \mid^2 dx = 1$.

2.10 Exercises

Exercise 2.1
Consider the two states $\mid \psi\rangle = i \mid \phi_1\rangle + 3i \mid \phi_2\rangle - \mid \phi_3\rangle$ and $\mid \chi\rangle = \mid \phi_1\rangle - i \mid \phi_2\rangle + 5i \mid \phi_3\rangle$, where $\mid \phi_1\rangle$, $\mid \phi_2\rangle$ and $\mid \phi_3\rangle$ are orthonormal.

(a) Calculate $\langle\psi \mid \psi\rangle$, $\langle\chi \mid \chi\rangle$, $\langle\psi \mid \chi\rangle$, $\langle\chi \mid \psi\rangle$, and infer $\langle\psi+\chi \mid \psi+\chi\rangle$. Are the scalar products $\langle\psi \mid \chi\rangle$ and $\langle\chi \mid \psi\rangle$ equal?

(b) Calculate $\mid \psi\rangle\langle\chi \mid$ and $\mid \chi\rangle\langle\psi \mid$. Are they equal? Calculate their traces and compare them.

(c) Find the Hermitian conjugates of $\mid \psi\rangle$, $\mid \chi\rangle$, $\mid \psi\rangle\langle\chi \mid$, and $\mid \chi\rangle\langle\psi \mid$.

Exercise 2.2
Consider two states $|\psi_1\rangle = |\phi_1\rangle + 4i|\phi_2\rangle + 5|\phi_3\rangle$ and $|\psi_2\rangle = b|\phi_1\rangle + 4|\phi_2\rangle - 3i|\phi_3\rangle$, where $|\phi_1\rangle$, $|\phi_2\rangle$, and$|\phi_3\rangle$ are orthonormal kets, and where b is a constant. Find the value of b so that $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal.

Exercise 2.3
If $\mid \phi_1\rangle$, $\mid \phi_2\rangle$, and $\mid \phi_3\rangle$ are orthonormal, show that the states $\mid \psi\rangle = i \mid \phi_1\rangle + 3i \mid \phi_2\rangle - \mid \phi_3\rangle$ and $\mid \chi\rangle = \mid \phi_1\rangle - i \mid \phi_2\rangle + 5i \mid \phi_3\rangle$ satisfy

(a) the triangle inequality and

(b) the Schwarz inequality.

Exercise 2.4
Find the constant α so that the states $\mid \psi\rangle = \alpha \mid \phi_1\rangle + 5 \mid \phi_2\rangle$ and $\mid \chi\rangle = 3\alpha \mid \phi_1\rangle - 4 \mid \phi_2\rangle$ are orthogonal; consider $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ to be orthonormal.

Exercise 2.5
If $\mid \psi\rangle = \mid \phi_1\rangle + \mid \phi_2\rangle$ and $\mid \chi\rangle = \mid \phi_1\rangle - \mid \phi_2\rangle$, prove the following relations (note that $\mid \phi_1\rangle$ and $\mid \phi_2\rangle$ are not orthonormal):

(a) $\langle\psi \mid \psi\rangle + \langle\chi \mid \chi\rangle = 2\langle\phi_1 \mid \phi_1\rangle + 2\langle\phi_2 \mid \phi_2\rangle$,

(b) $\langle\psi \mid \psi\rangle - \langle\chi \mid \chi\rangle = 2\langle\phi_1 \mid \phi_2\rangle + 2\langle\phi_2 \mid \phi_1\rangle$.

Exercise 2.6
Consider a state which is given in terms of three orthonormal vectors $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ as follows:

$$|\psi\rangle = \frac{1}{\sqrt{15}}|\phi_1\rangle + \frac{1}{\sqrt{3}}|\phi_2\rangle + \frac{1}{\sqrt{5}}|\phi_3\rangle,$$

where $|\phi_n\rangle$ are eigenstates to an operator $\hat{B}$ such that: $\hat{B}|\phi_n\rangle = (3n^2-1)|\phi_n\rangle$ with $n = 1, 2, 3$.

(a) Find the norm of the state $|\psi\rangle$.

(b) Find the expectation value of $\hat{B}$ for the state $|\psi\rangle$.

(c) Find the expectation value of $\hat{B}^2$ for the state $|\psi\rangle$.

Chapter 3

Postulates of Quantum Mechanics

3.1 Introduction

The formalism of quantum mechanics is based on a number of postulates. These postulates are in turn based on a wide range of experimental observations; the underlying physical ideas of these experimental observations have been briefly mentioned in Chapter 1. In this chapter we present a formal discussion of these postulates, and how they can be used to extract quantitative information about microphysical systems.

These postulates cannot be derived; they result from experiment. They represent the minimal set of assumptions needed to develop the theory of quantum mechanics. But how does one find out about the validity of these postulates? Their validity cannot be determined directly; only an indirect inferential statement is possible. For this, one has to turn to the theory built upon these postulates: if the theory works, the postulates will be valid; otherwise they will make no sense. Quantum theory not only works, but works extremely well, and this represents its experimental justification. It has a very penetrating qualitative as well as quantitative prediction power; this prediction power has been verified by a rich collection of experiments. So the accurate prediction power of quantum theory gives irrefutable evidence to the validity of the postulates upon which the theory is built.

3.2 The Basic Postulates of Quantum Mechanics

According to classical mechanics, the state of a particle is specified, at any time t, by two fundamental dynamical variables: the position $\vec{r}(t)$ and the momentum $\vec{p}(t)$. Any other physical quantity, relevant to the system, can be calculated in terms of these two dynamical variables. In addition, knowing these variables at a time t, we can predict, using for instance Hamilton's equations $dx/dt = \partial H/\partial p$ and $dp/dt = -\partial H/\partial x$, the values of these variables at any later time t'.

The quantum mechanical counterparts to these ideas are specified by postulates, which enable us to understand:

- how a quantum state is described mathematically at a given time t,
- how to calculate the various physical quantities from this quantum state, and

- knowing the system's state at a time t, how to find the state at any later time t'; that is, how to describe the time evolution of a system.

The answers to these questions are provided by the following set of five postulates.

Postulate 1: State of a system
The state of any physical system is specified, at each time t, by a state vector $|\psi(t)\rangle$ in a Hilbert space $\mathcal{H}$; $|\psi(t)\rangle$ contains (and serves as the basis to extract) all the needed information about the system. Any superposition of state vectors is also a state vector.

Postulate 2: Observables and operators
To every physically measurable quantity A, called an observable or dynamical variable, there corresponds a linear Hermitian operator $\hat{A}$ whose eigenvectors form a complete basis.

Postulate 3: Measurements and eigenvalues of operators
The measurement of an observable A may be represented formally by the action of $\hat{A}$ on a state vector $|\psi(t)\rangle$. The only possible result of such a measurement is one of the eigenvalues a_n (which are real) of the operator $\hat{A}$. If the result of a measurement of A on a state $|\psi(t)\rangle$ is a_n, the state of the system *immediately after* the measurement changes to $|\psi_n\rangle$:

$$\hat{A}|\psi(t)\rangle = a_n|\psi_n\rangle, \tag{3.1}$$

where $a_n = \langle\psi_n|\psi(t)\rangle$. **Note:** a_n is the component of $|\psi(t)\rangle$ when projected[1] onto the eigenvector $|\psi_n\rangle$.

Postulate 4: Probabilistic outcome of measurements

- **Discrete spectra**: When measuring an observable A of a system in a state $|\psi\rangle$, the probability of obtaining one of the nondegenerate eigenvalues a_n of the corresponding operator $\hat{A}$ is given by

$$P_n(a_n) = \frac{|\langle\psi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \frac{|a_n|^2}{\langle\psi|\psi\rangle}, \tag{3.2}$$

  where $|\psi_n\rangle$ is the eigenstate of $\hat{A}$ with eigenvalue a_n. If the eigenvalue a_n is m-degenerate, P_n becomes

$$P_n(a_n) = \frac{\sum_{j=1}^{m} |\langle\psi_n^j|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \frac{\sum_{j=1}^{m} |a_n^{(j)}|^2}{\langle\psi|\psi\rangle}. \tag{3.3}$$

  The act of measurement changes the state of the system from $|\psi\rangle$ to $|\psi_n\rangle$. If the system is already in an eigenstate $|\psi_n\rangle$ of $\hat{A}$, a measurement of A yields with certainty the corresponding eigenvalue a_n: $\hat{A}|\psi_n\rangle = a_n|\psi_n\rangle$.

- **Continuous spectra:** The relation (3.2), which is valid for discrete spectra, can be extended to determine the probability density that a measurement of $\hat{A}$ yields a value between a and $a + da$ on a system which is initially in a state $|\psi\rangle$:

$$\frac{dP(a)}{da} = \frac{|\psi(a)|^2}{\langle\psi|\psi\rangle} = \frac{|\psi(a)|^2}{\int_{-\infty}^{+\infty} |\psi(a')|^2 da'}; \tag{3.4}$$

  for instance, the probability density for finding a particle between x and $x + dx$ is given by $dP(x)/dx = |\psi(x)|^2/\langle\psi|\psi\rangle$.

[1] To see this, we need only to expand $|\psi(t)\rangle$ in terms of the eigenvectors of $\hat{A}$ which form a complete basis: $|\psi(t)\rangle = \sum_n |\psi_n\rangle\langle\psi_n|\psi(t)\rangle = \sum_n a_n|\psi_n\rangle$.

Postulate 5: Time evolution of a system
The time evolution of the state vector $|\psi(t)\rangle$ of a system is governed by the time-dependent *Schrödinger equation*

$$i\hbar \frac{\partial|\psi(t)\rangle}{\partial t} = \hat{H}|\psi(t)\rangle, \tag{3.5}$$

where $\hat{H}$ is the Hamiltonian operator corresponding to the total energy of the system.

Remark
These postulates fall into two categories:

- The first four describe the system at a given time.
- The fifth shows how this description evolves in time.

In the rest of this chapter we are going to consider the physical implications of each one of the four postulates. Namely, we shall look at the state of a quantum system and its interpretation, the physical observables, measurements in quantum mechanics, and finally the time evolution of quantum systems.

3.3 The State of a System

To describe a system in quantum mechanics, we use a mathematical entity (a complex function) belonging to a Hilbert space, the state vector $|\psi(t)\rangle$, which contains all the information we need to know about the system and from which all needed physical quantities can be computed. As discussed in Chapter 2, the state vector $|\psi(t)\rangle$ may be represented in two ways:

- A wave function $\psi(\vec{r}, t)$ in the position space: $\psi(\vec{r}, t) = \langle \vec{r}|\psi(t)\rangle$.
- A momentum wave function $\Psi(\vec{p}, t)$ in the momentum space: $\Psi(\vec{p}, t) = \langle \vec{p}|\psi(t)\rangle$.

So, for instance, to describe the state of a one-dimensional particle in quantum mechanics we use a complex function $\psi(x, t)$ instead of two real real numbers (x, p) in classical physics.

The wave functions to be used are only those that correspond to physical systems. What are the mathematical requirements that a wave function must satisfy to represent a physical system? Wave functions $\psi(x)$ that are physically acceptable must, along with their first derivatives $d\psi(x)/dx$, be *finite*, *continuous*, and *single-valued everywhere*. As will be discussed in Chapter 4, we will examine the underlying physics behind the continuity conditions of $\psi(x)$ and $d\psi(x)/dx$ (we will see that $\psi(x)$ and $d\psi(x)/dx$ must be be continuous because the probability density and the linear momentum are continuous functions of x).

3.3.1 Probability Density

What about the physical meaning of a wave function? Only the square of its norm, $|\psi(\vec{r}, t)|^2$, has meaning. According to Born's probabilistic interpretation, the square of the norm of $\psi(\vec{r}, t)$,

$$P(\vec{r}, t) = |\psi(\vec{r}, t)|^2, \tag{3.6}$$

represents a position probability density; that is, the quantity $|\psi(\vec{r},t)|^2 d^3r$ represents the probability of finding the particle at time t in a volume element d^3r located between $\vec{r}$ and $\vec{r}+d\vec{r}$. Therefore, the total probability of finding the system somewhere in space is equal to 1:

$$\int |\psi(\vec{r},t)|^2 d^3r = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} |\psi(\vec{r},t)|^2 dz = 1. \tag{3.7}$$

A wave function $\psi(\vec{r},t)$ satisfying this relation is said to be *normalized*. We may mention that $\psi(\vec{r})$ has the physical dimensions of $1/\sqrt{L^3}$, where L is a length. Hence, the physical dimensions of $|\psi(\vec{r})|^2$ is $1/L^3$: $\left[|\psi(\vec{r})|^2\right] = 1/L^3$.

Note that the wave functions $\psi(\vec{r},t)$ and $e^{i\alpha}\psi(\vec{r},t)$, where α is a real number, represent the same state.

Example 3.1 (Physical and unphysical wave functions)
Which among the following functions represent physically acceptable wave functions: $f(x) = 3\sin\pi x$, $g(x) = 4-|x|$, $h^2(x) = 5x$, and $e(x) = x^2$.

Solution
Among these functions only $f(x) = 3\sin\pi x$ represents a physically acceptable wave function, since $f(x)$ and its derivative are finite, continuous, single-valued everywhere, and integrable.

The other functions cannot be wave functions, since $g(x) = 4-|x|$ is not continuous, not finite, and not square integrable; $h^2(x) = 5x$ is neither finite nor square integrable; and $e(x) = x^2$ is neither finite nor square integrable.

3.3.2 The Superposition Principle

The state of a system does not have to be represented by a *single* wave function; it can be represented by a *superposition* of two or more wave functions. An example from the macroscopic world is a vibrating string; its state can be represented by a single wave or by the superposition (linear combination) of many waves.

If $\psi_1(\vec{r},t)$ and $\psi_2(\vec{r},t)$ separately satisfy the Schrödinger equation, then the wave function $\psi(\vec{r},t) = \alpha_1\psi_1(\vec{r},t) + \alpha_2\psi_2(\vec{r},t)$ also satisfies the Schrödinger equation, where α_1 and α_2 are complex numbers. The Schrödinger equation is a linear equation. So in general, according to the superposition principle, the linear superposition of many wave functions (which describe the various permissible physical states of a system) gives a new wave function which represents a possible physical state of the system:

$$|\psi\rangle = \sum_i \alpha_i |\psi_i\rangle, \tag{3.8}$$

where the α_i are complex numbers. The quantity

$$P = \left|\sum_i \alpha_i |\psi_i\rangle\right|^2, \tag{3.9}$$

represents the probability for this superposition. If the states $|\psi_i\rangle$ are mutually *orthonormal*, the probability will be equal to the sum of the individual probabilities:

$$P = \left|\sum_i \alpha_i |\psi_i\rangle\right|^2 = \sum_i |\alpha_i|^2 = P_1 + P_2 + P_3 + \cdots, \tag{3.10}$$

where $P_i = |\alpha_i|^2$; P_i is the probability of finding the system in the state $|\psi_i\rangle$.

Example 3.2

Consider a system whose state is given in terms of an orthonormal set of three vectors: $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$ as

$$|\psi\rangle = \frac{\sqrt{3}}{3}|\phi_1\rangle + \frac{2}{3}|\phi_2\rangle + \frac{\sqrt{2}}{3}|\phi_3\rangle.$$

(a) Verify that $|\psi\rangle$ is normalized. Then, calculate the probability of finding the system in any one of the states $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$. Verify that the total probability is equal to one.

(b) Consider now an ensemble of 810 identical systems, each one of them in the state $|\psi\rangle$. If measurements are done on all of them, how many systems will be found in each of the states $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$?

Solution

(a) Using the orthonormality condition $\langle\phi_j|\phi_k\rangle = \delta_{jk}$ where $j, k = 1, 2, 3$, we can verify that $|\psi\rangle$ is normalized:

$$\langle\psi|\psi\rangle = \frac{1}{3}\langle\phi_1|\phi_1\rangle + \frac{4}{9}\langle\phi_2|\phi_2\rangle + \frac{2}{9}\langle\phi_3|\phi_3\rangle = \frac{1}{3} + \frac{4}{9} + \frac{2}{9} = 1. \tag{3.11}$$

Since $|\psi\rangle$ is normalized, the probability of finding the system in $|\phi_1\rangle$ is given by

$$P_1 = |\langle\phi_1|\psi\rangle|^2 = \left|\frac{\sqrt{3}}{3}\langle\phi_1|\phi_1\rangle + \frac{2}{3}\langle\phi_1|\phi_2\rangle + \frac{\sqrt{2}}{3}\langle\phi_1|\phi_3\rangle\right|^2 = \frac{1}{3}, \tag{3.12}$$

since $\langle\phi_1|\phi_1\rangle = 1$ and $\langle\phi_1|\phi_2\rangle = \langle\phi_1|\phi_3\rangle = 0$.

Similarly, from the relations $\langle\phi_2|\phi_2\rangle = 1$ and $\langle\phi_2|\phi_1\rangle = \langle\phi_2|\phi_3\rangle = 0$, we obtain the probability of finding the system in $|\phi_2\rangle$:

$$P_2 = |\langle\phi_2|\psi\rangle|^2 = \left|\frac{2}{3}\langle\phi_2|\phi_2\rangle\right|^2 = \frac{4}{9}. \tag{3.13}$$

As for $\langle\phi_3|\phi_3\rangle = 1$ and $\langle\phi_3|\phi_1\rangle = \langle\phi_3|\phi_2\rangle = 0$, they lead to the probability of finding the system in $|\phi_3\rangle$:

$$P_3 = |\langle\phi_3|\psi\rangle|^2 = \left|\frac{\sqrt{2}}{3}\langle\phi_3|\phi_3\rangle\right|^2 = \frac{2}{9}. \tag{3.14}$$

As expected, the total probability is equal to one:

$$P = P_1 + P_2 + P_3 = \frac{1}{3} + \frac{4}{9} + \frac{2}{9} = 1. \tag{3.15}$$

(b) The number of systems that will be found in the state $|\phi_1\rangle$ is

$$N_1 = 810 \times P_1 = \frac{810}{3} = 270. \tag{3.16}$$

Likewise, the number of systems that will be found in states $|\phi_2\rangle$ and $|\phi_3\rangle$ are given, respectively, by

$$N_2 = 810 \times P_2 = \frac{810 \times 4}{9} = 360, \qquad N_3 = 810 \times P_3 = \frac{810 \times 2}{9} = 180. \tag{3.17}$$

3.4 Observables and Operators

An observable is a dynamical variable that can be measured; the dynamical variables encountered most in classical mechanics are the position, linear momentum, angular momentum, and energy. How do we mathematically represent these and other variables in quantum mechanics?

According to the second postulate, a *Hermitian operator* is associated with every *physical observable*. In the preceding chapter, we have seen that the position representation of the linear momentum operator is given in one-dimensional space by $\hat{P} = -i\hbar\partial/\partial x$ and in three-dimensional space by $\hat{\vec{P}} = -i\hbar\vec{\nabla}$.

In general, any function, $f(\vec{r}, \vec{p})$, which depends on the position and momentum variables, $\vec{r}$ and $\vec{p}$, can be "*quantized*" or made into a function of operators by replacing $\vec{r}$ and $\vec{p}$ with their corresponding operators:

$$f(\vec{r}, \vec{p}) \quad \longrightarrow \quad F(\hat{\vec{R}}, \hat{\vec{P}}) = f(\hat{\vec{R}}, -i\hbar\vec{\nabla}), \tag{3.18}$$

or $f(x, p) \to F(\hat{X}, \ -i\hbar\partial/\partial x)$. For instance, the operator corresponding to the Hamiltonian

$$H = \frac{1}{2m}\vec{p}^{\,2} + V(\vec{r}, t) \tag{3.19}$$

is given in the position representation by

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\hat{\vec{R}}, t), \tag{3.20}$$

where ∇^2 is the Laplacian operator; it is given in Cartesian coordinates by: $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$.

Since the momentum operator $\hat{\vec{P}}$ is Hermitian, and if the potential $V(\hat{\vec{R}}, t)$ is a real function, the Hamiltonian (3.19) is Hermitian. We saw in Chapter 2 that the eigenvalues of Hermitian operators are real. Hence, the spectrum of the Hamiltonian, which consists of the entire set of its eigenvalues, is real. This spectrum can be discrete, continuous, or a mixture of both. In the case of *bound* states, the Hamiltonian has a *discrete* spectrum of values and a *continuous* spectrum for *unbound* states. In general, an operator will have bound or unbound spectra in the same manner that the corresponding classical variable has bound or unbound orbits. As for $\hat{\vec{R}}$ and $\hat{\vec{P}}$, they have continuous spectra, since r and p may take a continuum of values.

Table 3.1 Some observables and their corresponding operators.

Observable	Corresponding operator
$\vec{r}$	$\hat{\vec{R}}$
$\vec{p}$	$\hat{\vec{P}} = -i\hbar\vec{\nabla}$
$T = \frac{p^2}{2m}$	$\hat{T} = -\frac{\hbar^2}{2m}\nabla^2$
$E = \frac{p^2}{2m} + V(\vec{r}, t)$	$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(\hat{\vec{R}}, t)$
$\vec{L} = \vec{r} \times \vec{p}$	$\hat{\vec{L}} = -i\hbar\hat{\vec{R}} \times \vec{\nabla}$

According to Postulate 5, the total energy E for time-dependent systems is associated to the operator

$$\hat{H} = i\hbar\frac{\partial}{\partial t}. \tag{3.21}$$

This can be seen as follows. The wave function of a free particle of momentum $\vec{p}$ and total energy E is given by $\psi(\vec{r}, t) = Ae^{i(\vec{p}\cdot\vec{r}-Et)/\hbar}$, where A is a constant. The time derivative of $\psi(\vec{r}, t)$ yields

$$i\hbar\frac{\partial\psi(\vec{r}, t)}{\partial t} = E\psi(\vec{r}, t). \tag{3.22}$$

Let us look at the eigenfunctions and eigenvalues of the momentum operator $\hat{\vec{P}}$. The eigenvalue equation

$$-i\hbar\vec{\nabla}\psi(\vec{r}) = \vec{p}\psi(\vec{r}) \tag{3.23}$$

yields the eigenfunction $\psi(\vec{r})$ corresponding to the eigenvalue $\vec{p}$ such that $|\psi(\vec{r})|^2 d^3r$ is the probability of finding the particle with a momentum $\vec{p}$ in the volume element d^3r centered about $\vec{r}$. The solution to the eigenvalue equation (3.23) is

$$\psi(\vec{r}) = Ae^{i\vec{p}\cdot\vec{r}/\hbar}, \tag{3.24}$$

where A is a normalization constant. Since $\vec{p} = \hbar\vec{k}$ is the eigenvalue of the operator $\hat{\vec{P}}$, the eigenfunction (3.24) reduces to $\psi(\vec{r}) = Ae^{i\vec{k}\cdot\vec{r}}$; hence the eigenvalue equation (3.23) becomes

$$\hat{\vec{P}}\psi(\vec{r}) = \hbar\vec{k}\psi(\vec{r}). \tag{3.25}$$

To summarize, there is a one-to-one correspondence between observables and operators (Table 3.1).

Example 3.3 (Orbital angular momentum)
Find the operator representing the classical orbital angular momentum.

Solution
The classical expression for the orbital angular momentum of a particle whose position and linear momentum are $\vec{r}$ and $\vec{p}$ is given by $\vec{L} = \vec{r} \times \vec{p} = l_x\vec{i} + l_y\vec{j} + l_z\vec{k}$, where $l_x = yp_z - zp_y$, $l_y = zp_x - xp_z$, $l_z = xp_y - yp_x$.

To find the operator representing the classical angular momentum, we need simply to replace $\vec{r}$ and $\vec{p}$ with their corresponding operators $\hat{\vec{R}}$ and $\hat{\vec{P}} = -i\hbar\vec{\nabla}$: $\hat{\vec{L}} = -i\hbar\hat{\vec{R}} \times \vec{\nabla}$. This leads to

$$\hat{L}_x = \hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y = -i\hbar\left(\hat{Y}\frac{\partial}{\partial z} - \hat{Z}\frac{\partial}{\partial y}\right), \tag{3.26}$$

$$\hat{L}_y = \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z = -i\hbar\left(\hat{Z}\frac{\partial}{\partial x} - \hat{X}\frac{\partial}{\partial Z}\right), \tag{3.27}$$

$$\hat{L}_z = \hat{X}\hat{P}_y - \hat{Y}\hat{P}_x = -i\hbar\left(\hat{X}\frac{\partial}{\partial y} - \hat{Y}\frac{\partial}{\partial x}\right). \tag{3.28}$$

Recall that in classical mechanics the position and momentum components commute, $xp_x = p_x x$, and so do the components of the angular momentum, $l_x l_y = l_y l_x$. In quantum mechanics, however, this is not the case, since $\hat{X}\hat{P}_x = \hat{P}_x\hat{X} + i\hbar$ and, as will be shown in Chapter 5, $\hat{L}_x\hat{L}_y = \hat{L}_y\hat{L}_x + i\hbar\hat{L}_z$, and so on.

3.5 Measurement in Quantum Mechanics

Quantum theory is about the results of measurement; it says nothing about what might happen in the physical world outside the context of measurement. So the emphasis is on measurement.

3.5.1 How Measurements Disturb Systems

In classical physics it is possible to perform measurements on a system without disturbing it significantly. In quantum mechanics, however, the measurement process perturbs the system significantly. While carrying out measurements on classical systems, this perturbation does exist, but it is small enough that it can be neglected. In atomic and subatomic systems, however, the act of measurement induces nonnegligible or significant disturbances.

As an illustration, consider an experiment that measures the position of a hydrogenic electron. For this, we need to bombard the electron with electromagnetic radiation (photons). If we want to determine the position accurately, the wavelength of the radiation must be sufficiently short. Since the electronic orbit is of the order of 10^{-10} m, we must use a radiation whose wavelength is smaller than 10^{-10} m. That is, we need to bombard the electron with photons of energies higher than

$$h\nu = h\frac{c}{\lambda} = h\frac{3 \times 10^8}{10^{-10}} \sim 10^4 \text{ eV}. \tag{3.29}$$

When such photons strike the electron, not only will they perturb it, they will knock it completely off its orbit; recall that the ionization energy of the hydrogen atom is about 13.5 eV. Thus, the mere act of measuring the position of the electron disturbs it appreciably.

Let us now discuss the general concept of measurement in quantum mechanics. *The act of measurement generally changes the state of the system.* In theory we can represent the measuring device by an operator so that, after carrying out the measurement, the system will be in one of the eigenstates of the operator. Consider a system which is in a state $|\psi\rangle$. Before measuring an observable A, the state $|\psi\rangle$ can be represented by a linear superposition of eigenstates $|\psi_n\rangle$

of the corresponding operator $\hat{A}$:

$$|\psi\rangle = \sum_n |\psi_n\rangle\langle\psi_n|\psi\rangle = \sum_n a_n|\psi_n\rangle. \tag{3.30}$$

According to Postulate 4, the act of measuring A changes the state of the system from $|\psi\rangle$ to one of the eigenstates $|\psi_n\rangle$ of the operator $\hat{A}$, and the result obtained is the eigenvalue a_n. The only *exception* to this rule is when the *system is already in one of the eigenstates of the observable being measured.* For instance, if the system is in the eigenstate $|\psi_n\rangle$, a measurement of the observable A yields with certainty (i.e., with probability = 1) the value a_n without changing the state $|\psi_n\rangle$.

Before a measurement, we do not know in advance with certainty in which eigenstate, among the various states $|\psi_n\rangle$, a system will be after the measurement; only a probabilistic outcome is possible. Postulate 4 states that the probability of finding the system in one particular nondegenerate eigenstate $|\psi_n\rangle$ is given by

$$P_n = \frac{|\langle\psi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle}. \tag{3.31}$$

Note that the wave function does not predict the results of individual measurements; it instead determines the probability distribution, $P = |\psi|^2$, over measurements on many identical systems in the same state.

Finally, we may state that quantum mechanics is the mechanics applicable to objects for which measurements necessarily interfere with the state of the system. Quantum mechanically, we cannot ignore the effects of the measuring equipment on the system, for they are important. In general, certain measurements cannot be performed without major disturbances to other properties of the quantum system. In conclusion, *it is the effects of the interference by the equipment on the system which is the essence of quantum mechanics*.

3.5.2 Expectation Values

The expectation value $\langle\hat{A}\rangle$ of $\hat{A}$ with respect to a state $|\psi\rangle$ is defined by

$$\langle\hat{A}\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}. \tag{3.32}$$

For instance, the energy of a system is given by the expectation value of the Hamiltonian: $E = \langle\hat{H}\rangle = \langle\psi|\hat{H}|\psi\rangle/\langle\psi|\psi\rangle$.

In essence, the expectation value $\langle\hat{A}\rangle$ represents the average result of measuring $\hat{A}$ on the state $|\ \psi\rangle$. To see this, using the complete set of eigenvectors $|\psi_n\rangle$ of $\hat{A}$ as a basis (i.e., $\hat{A}$ is diagonal in ψ_n), we can rewrite $\langle\hat{A}\rangle$ as follows:

$$\langle\hat{A}\rangle = \frac{1}{\langle\psi|\psi\rangle}\sum_{nm}\langle\psi|\psi_m\rangle\langle\psi_m|\hat{A}|\psi_n\rangle\langle\psi_n|\psi\rangle = \sum_n a_n\frac{|\langle\psi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle}, \tag{3.33}$$

where we have used $\langle\psi_m|\hat{A}|\psi_n\rangle = a_n\delta_{nm}$. Since the quantity $|\langle\psi_n|\psi\rangle|^2/\langle\psi|\psi\rangle$ gives the probability P_n of finding the value a_n after measuring the observable A, we can indeed interpret $\langle\hat{A}\rangle$ as an *average* of a series of measurements of A:

$$\boxed{\langle\hat{A}\rangle = \sum_n a_n\frac{|\langle\psi_n|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \sum_n a_n P_n.} \tag{3.34}$$

That is, the expectation value of an observable is obtained by adding all permissible eigenvalues a_n, with each a_n multiplied by the corresponding probability P_n.

The relation (3.34), which is valid for *discrete* spectra, can be extended to a *continuous* distribution of probabilities $P(a)$ as follows:

$$\langle \hat{A} \rangle = \frac{\int_{-\infty}^{+\infty} a\, |\psi(a)|^2 da}{\int_{-\infty}^{+\infty} |\psi(a)|^2 da} = \int_{-\infty}^{+\infty} a\, dP(a). \tag{3.35}$$

The expectation value of an observable can be obtained physically as follows: prepare a very large number of *identical* systems each in the *same* state $|\psi\rangle$. The observable A is then measured on all these identical systems; the results of these measurements are $a_1, a_2, \ldots, a_n, \ldots$; the corresponding probabilities of occurrence are $P_1, P_2, \ldots, P_n, \ldots$. The average value of all these repeated measurements is called the expectation value of $\hat{A}$ with respect to the state $|\psi\rangle$.

Note that the process of obtaining different results when measuring the same observable on many identically prepared systems is contrary to classical physics, where these measurements must give the same outcome. In quantum mechanics, however, we can predict only the probability of obtaining a certain value for an observable.

Example 3.4

Consider a system whose state is given in terms of a complete and orthonormal set of five vectors $|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle, |\phi_4\rangle, |\phi_5\rangle$ as follows:

$$|\psi\rangle = \frac{1}{\sqrt{19}}|\phi_1\rangle + \frac{2}{\sqrt{19}}|\phi_2\rangle + \sqrt{\frac{2}{19}}|\phi_3\rangle + \sqrt{\frac{3}{19}}|\phi_4\rangle + \sqrt{\frac{5}{19}}|\phi_5\rangle,$$

where $|\phi_n\rangle$ are eigenstates to the system's Hamiltonian, $\hat{H}|\phi_n\rangle = n\varepsilon_0|\phi_n\rangle$ with $n = 1, 2, 3, 4, 5$, and where ε_0 has the dimensions of energy.

(a) If the energy is measured on a large number of identical systems that are all initially in the same state $|\psi\rangle$, what values would one obtain and with what probabilities?

(b) Find the average energy of one such system.

Solution

First, note that $|\psi\rangle$ is not normalized:

$$\langle\psi|\psi\rangle = \sum_{n=1}^{5} a_n^2 \langle\phi_n|\phi_n\rangle = \sum_{n=1}^{5} a_n^2 = \frac{1}{19} + \frac{4}{19} + \frac{2}{19} + \frac{3}{19} + \frac{5}{19} = \frac{15}{19}, \tag{3.36}$$

since $\langle\phi_j|\phi_k\rangle = \delta_{jk}$ with $j, k = 1, 2, 3, 4, 5$.

(a) Since $E_n = \langle\phi_n|\hat{H}|\phi_n\rangle = n\varepsilon_0$ $(n = 1, 2, 3, 4, 5)$, the various measurements of the energy of the system yield the values $E_1 = \varepsilon_0$, $E_2 = 2\varepsilon_0$, $E_3 = 3\varepsilon_0$, $E_4 = 4\varepsilon_0$, $E_5 = 5\varepsilon_0$ with the following probabilities:

$$P_1(E_1) = \frac{|\langle\phi_1|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \left|\frac{1}{\sqrt{19}}\langle\phi_1|\phi_1\rangle\right|^2 \times \frac{19}{15} = \frac{1}{15}, \tag{3.37}$$

$$P_2(E_2) = \frac{|\langle\phi_2|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \left|\frac{2}{\sqrt{19}}\langle\phi_2|\phi_2\rangle\right|^2 \times \frac{19}{15} = \frac{4}{15}, \tag{3.38}$$

$$P_3(E_3) = \frac{|\langle\phi_3|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \left|\sqrt{\frac{2}{19}}\langle\phi_3|\phi_3\rangle\right|^2 \times \frac{19}{15} = \frac{2}{15}, \tag{3.39}$$

$$P_4(E_4) = \frac{|\langle\phi_4|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \left|\sqrt{\frac{3}{19}}\langle\phi_4|\phi_4\rangle\right|^2 \times \frac{19}{15} = \frac{3}{15}, \tag{3.40}$$

and

$$P_5(E_5) = \frac{|\langle\phi_5|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \left|\sqrt{\frac{5}{19}}\langle\phi_5|\phi_5\rangle\right|^2 \times \frac{19}{15} = \frac{5}{15}, \tag{3.41}$$

(b) The average energy of a system is given by

$$E = \sum_{j=1}^{5} P_j E_j = \frac{1}{15}\varepsilon_0 + \frac{8}{15}\varepsilon_0 + \frac{6}{15}\varepsilon_0 + \frac{12}{15}\varepsilon_0 + \frac{25}{15}\varepsilon_0 = \frac{52}{15}\varepsilon_0. \tag{3.42}$$

This energy can also be obtained from the expectation value of the Hamiltonian:

$$\begin{aligned} E &= \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{19}{15}\sum_{n=1}^{5} a_n^2 \langle\phi_n|\hat{H}|\phi_n\rangle = \frac{19}{15}\left(\frac{1}{19} + \frac{8}{19} + \frac{6}{19} + \frac{12}{19} + \frac{25}{19}\right)\varepsilon_0 \\ &= \frac{52}{15}\varepsilon_0, \end{aligned} \tag{3.43}$$

where the values of the coefficients a_n^2 are listed in (3.36).

3.5.3 Complete Sets of Commuting Operators (CSCO)

Two observables A and B are said to be *compatible* when their corresponding operators commute, $[\hat{A}, \hat{B}] = 0$; observables corresponding to noncommuting operators are said to be *non-compatible*.

In what follows we are going to consider the task of measuring two observables A and B on a given system. Since the act of measurement generally *perturbs* the system, the result of measuring A and B therefore depends on the *order* in which they are carried out. Measuring A first and then B leads[2] in general to results that are different from those obtained by measuring B first and then A. How does this take place?

If $\hat{A}$ and $\hat{B}$ do not commute and if the system is in an eigenstate $|\psi_n^{(a)}\rangle$ of $\hat{A}$, a measurement of A yields with certainty a value a_n, since $\hat{A}|\psi_n^{(a)}\rangle = a_n|\psi_n^{(a)}\rangle$. Then, when we measure B, the state of the system will be left in one of the eigenstates of B. If we measure A again, we will find a value which will be different from a_n. What is this new value? We cannot answer this question with certainty: only a probabilistic outcome is possible. For this, we need to expand the eigenstates of B in terms of those of A, and thus provide a probabilistic answer as to the value of measuring A. *So if $\hat{A}$ and $\hat{B}$ do not commute, they cannot be measured simultaneously; the order in which they are measured matters.*

[2]The act of measuring A first and then B is represented by the action of product $\hat{B}\hat{A}$ of their corresponding operators on the state vector.

What happens when A and B commute? We can show that the results of their measurements will not depend on the order in which they are carried out. Before showing this, let us mention a useful theorem.

Theorem 3.1 *If two observables are compatible, their corresponding operators possess a set of common (or simultaneous) eigenstates (this theorem holds for both degenerate and nondegenerate eigenstates).*

Proof

We provide here a proof for the nondegenerate case only. If $|\psi_n\rangle$ is a nondegenerate eigenstate of $\hat{A}$, $\hat{A}|\psi_n\rangle = a_n|\psi_n\rangle$, we have

$$\langle\psi_m|[\hat{A}, \hat{B}]|\psi_n\rangle = (a_m - a_n)\langle\psi_m|\hat{B}|\psi_n\rangle = 0, \tag{3.44}$$

since $\hat{A}$ and $\hat{B}$ commute. So $\langle\psi_m|\hat{B}|\psi_n\rangle$ must vanish unless $a_n = a_m$. That is,

$$\langle\psi_m|\hat{B}|\psi_n\rangle = \langle\psi_n|\hat{B}|\psi_n\rangle \propto \delta_{nm}. \tag{3.45}$$

Hence the $|\psi_n\rangle$ are *joint* or *simultaneous* eigenstates of $\hat{A}$ and $\hat{B}$ (this completes the proof).

Denoting the simultaneous eigenstate of $\hat{A}$ and $\hat{B}$ by $|\psi_{n_1}^{(a)}, \psi_{n_2}^{(b)}\rangle$, we have

$$\hat{A}|\psi_{n_1}^{(a)}, \psi_{n_2}^{(b)}\rangle = a_{n_1}|\psi_{n_1}^{(a)}, \psi_{n_2}^{(b)}\rangle, \tag{3.46}$$

$$\hat{B}|\psi_{n_1}^{(a)}, \psi_{n_2}^{(b)}\rangle = b_{n_2}|\psi_{n_1}^{(a)}, \psi_{n_2}^{(b)}\rangle. \tag{3.47}$$

Theorem 3.1 can be generalized to the case of many mutually compatible observables A, B, C, ... These compatible observables possess a *complete* set of *joint eigenstates*

$$|\psi_n\rangle = |\psi_{n_1}^{(a)}, \psi_{n_2}^{(b)}, \psi_{n_3}^{(c)}, \ldots\rangle. \tag{3.48}$$

The completeness and orthonormality conditions of this set are

$$\sum_{n_1}\sum_{n_2}\sum_{n_3}\cdots|\psi_{n_1}^{(a)}, \psi_{n_2}^{(b)}, \psi_{n_3}^{(c)}, \ldots\rangle\langle\psi_{n_1}^{(a)}, \psi_{n_2}^{(b)}, \psi_{n_3}^{(c)}, \ldots| = 1; \tag{3.49}$$

$$\langle\psi_{n'}|\psi_n\rangle = \delta_{n'n} = \delta_{n_1'n_1}\delta_{n_2'n_2}\delta_{n_3'n_3}\cdots. \tag{3.50}$$

Let us now show why, when two observables A and B are compatible, the order in which we carry out their measurements is irrelevant. Measuring A first, we would find a value a_n and would leave the system in an eigenstate of A. According to Theorem 3.1, this eigenstate is also an eigenstate of B. Thus a measurement of B yields with certainty b_n without affecting the state of the system. In this way, if we measure A again, we obtain with certainty the same initial value a_n. Similarly, another measurement of B will yield b_n and will leave the system in the same joint eigenstate of A and B. Thus, if two observables A and B are compatible, and if the system is initially in an eigenstate of one of their operators, their measurements not only yield precise values (eigenvalues) but they will not depend on the order in which the measurements were performed. In this case, A and B are said to be *simultaneously measurable*. *So compatible observables can be measured simultaneously with arbitrary accuracy; noncompatible observables cannot*.

What happens if an operator, say $\hat{A}$, has *degenerate* eigenvalues? The specification of one eigenvalue does not uniquely determine the state of the system. Among the degenerate eigenstates of $\hat{A}$, only a subset of them are also eigenstates of $\hat{B}$. Thus, the set of states that are joint eigenstates of both $\hat{A}$ and $\hat{B}$ is not complete. To resolve the degeneracy, we can introduce a third operator $\hat{C}$ which commutes with both $\hat{A}$ and $\hat{B}$; then we can construct a set of joint eigenstates of $\hat{A}$, $\hat{B}$, and $\hat{C}$ that is complete. If the degeneracy persists, we may introduce a fourth operator $\hat{D}$ that commutes with the previous three and then look for their joint eigenstates which form a complete set. Continuing in this way, we will ultimately exhaust all the operators (that is, there are no more independent operators) which commute with each other. When that happens, we have then obtained a *complete set of commuting operators* (CSCO). Only then will the state of the system be specified unambiguously, for the joint eigenstates of the CSCO are determined uniquely and will form a complete set (recall that a complete set of eigenvectors of an operator is called a basis). We should, at this level, state the following definition.

Definition: A set of Hermitian operators, $\hat{A}$, $\hat{B}$, $\hat{C}$, ..., is called a CSCO if the operators mutually commute and if the set of their common eigenstates is complete and not degenerate (i.e., unique).

The complete commuting set may sometimes consist of only one operator. *Any operator with nondegenerate eigenvalues constitutes, all by itself, a CSCO.* For instance, the position operator $\hat{X}$ of a one-dimensional, spinless particle provides a complete set. Its momentum operator $\hat{P}$ is also a complete set; together, however, $\hat{X}$ and $\hat{P}$ cannot form a CSCO, for they do not commute. In three-dimensional problems, the three-coordinate position operators $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ form a CSCO; similarly, the components of the momentum operator $\hat{P}_x$, $\hat{P}_y$, and $\hat{P}_z$ also form a CSCO. In the case of spherically symmetric three-dimensional potentials, the set $\hat{H}$, $\hat{\vec{L}}^2$, $\hat{L}_z$ forms a CSCO. Note that in this case of spherical symmetry, we need three operators to form a CSCO because $\hat{H}$, $\hat{\vec{L}}^2$, and $\hat{L}_z$ are all degenerate; hence the complete and unique determination of the wave function cannot be achieved with one operator or with two.

In summary, when a given operator, say $\hat{A}$, is degenerate, the wave function cannot be determined uniquely unless we introduce one or more additional operators so as to form a complete commuting set.

3.5.4 Measurement and the Uncertainty Relations

We have seen in Chapter 2 that the uncertainty condition pertaining to the measurement of any two observables A and B is given by

$$\Delta A \Delta B \geq \frac{1}{2}|\langle[\hat{A},\ \hat{B}]\rangle|, \tag{3.51}$$

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

Let us illustrate this on the joint measurement of the position and momentum observables. Since these observables are not compatible, their simultaneous measurement with infinite accuracy is not possible; that is, since $[\hat{X}, \hat{P}] = i\hbar$ there exists no state which is a simultaneous eigenstate of $\hat{X}$ and $\hat{P}$. For the case of the position and momentum operators, the relation (3.51) yields

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \tag{3.52}$$

This condition shows that the position and momentum of a microscopic system cannot be measured with infinite accuracy both at once. If the position is measured with an uncertainty Δx, the uncertainty associated with its momentum measurement cannot be smaller than $\hbar/2\Delta x$. This is due to *the interference between the two measurements*. If we measure the position first, we perturb the system by changing its state to an eigenstate of the position operator; then the measurement of the momentum throws the system into an eigenstate of the momentum operator.

Another interesting application of the uncertainty relation (3.51) is to the orbital angular momentum of a particle. Since its components satisfy the commutator $[\hat{L}_x, \ \hat{L}_y] = i\hbar\hat{L}_z$, we obtain

$$\Delta L_x \Delta L_y \geq \frac{1}{2}\hbar|\langle\hat{L}_z\rangle|. \tag{3.53}$$

We can obtain the other two inequalities by means of a cyclic permutation of x, y, and z. If $\langle\hat{L}_z\rangle = 0$, $\hat{L}_x$ and $\hat{L}_y$ will have sharp values simultaneously. This occurs when the particle is in an s state. In fact, when a particle is in an s state, we have $\langle\hat{L}_x\rangle = \langle\hat{L}_y\rangle = \langle\hat{L}_z\rangle = 0$; hence all the components of orbital angular momentum will have sharp values simultaneously.

3.6 Time Evolution of the System's State

3.6.1 Time Evolution Operator

We want to examine here how quantum states evolve in time. That is, given the initial state $|\psi(t_0)\rangle$, how does one find the state $|\psi(t)\rangle$ at any later time t? The two states can be related by means of a linear operator $\hat{U}(t, t_0)$ such that

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle \qquad (t > t_0); \tag{3.54}$$

$\hat{U}(t, t_0)$ is known as the *time evolution operator* or *propagator*. From (3.54), we infer that

$$\hat{U}(t_0, t_0) = \hat{I}, \tag{3.55}$$

where $\hat{I}$ is the unit (identity) operator.

The issue now is to find $\hat{U}(t, t_0)$. For this, we need simply to substitute (3.54) into the time-dependent Schrödinger equation (3.5):

$$i\hbar\frac{\partial}{\partial t}\left(\hat{U}(t, t_0)|\psi(t_0)\rangle\right) = \hat{H}\left(\hat{U}(t, t_0)|\psi(t_0)\rangle\right) \tag{3.56}$$

or

$$\frac{\partial\hat{U}(t, t_0)}{\partial t} = -\frac{i}{\hbar}\hat{H}\hat{U}(t, t_0). \tag{3.57}$$

The integration of this differential equation depends on whether or not the Hamiltonian depends on time. If it does not depend on time, and taking into account the initial condition (3.55), we can easily ascertain that the integration of (3.57) leads to

$$\boxed{\hat{U}(t, t_0) = e^{-i(t-t_0)\hat{H}/\hbar} \qquad \text{and} \qquad |\psi(t)\rangle = e^{-i(t-t_0)\hat{H}/\hbar}|\psi(t_0)\rangle.} \tag{3.58}$$

We will show in Section 3.7 that the operator $\hat{U}(t, t_0) = e^{-i(t-t_0)\hat{H}/\hbar}$ represents a finite time translation.

If, on the other hand, $\hat{H}$ depends on time the integration of (3.57) becomes less trivial. We will deal with this issue in Chapter 10 when we look at time-dependent potentials or at the time-dependent perturbation theory. In this chapter, and in all chapters up to Chapter 10, we will consider only Hamiltonians that do not depend on time.

Note that $\hat{U}(t, t_0)$ is a unitary operator, since

$$\hat{U}(t, t_0)\hat{U}^\dagger(t, t_0) = \hat{U}(t, t_0)\hat{U}^{-1}(t, t_0) = e^{-i(t-t_0)\hat{H}/\hbar}e^{i(t-t_0)\hat{H}/\hbar} = \hat{I} \tag{3.59}$$

or $\hat{U}^\dagger = \hat{U}^{-1}$.

3.6.2 Stationary States: Time-Independent Potentials

In the position representation, the time-dependent Schrödinger equation (3.5) for a particle of mass m moving in a time-dependent potential $\hat{V}(\vec{r}, t)$ can be written as follows:

$$\boxed{i\hbar\frac{\partial\Psi(\vec{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{r}, t) + \hat{V}(\vec{r}, t)\Psi(\vec{r}, t).} \tag{3.60}$$

Now, let us consider the particular case of *time-independent* potentials: $\hat{V}(\vec{r}, t) = \hat{V}(\vec{r})$. In this case the Hamiltonian operator will also be time independent, and hence the Schrödinger equation will have solutions that are *separable*, i.e., solutions that consist of a product of two functions, one depending only on $\vec{r}$ and the other only on time:

$$\Psi(\vec{r}, t) = \psi(\vec{r})f(t). \tag{3.61}$$

Substituting (3.61) into (3.60) and dividing both sides by $\psi(\vec{r})f(t)$, we obtain

$$i\hbar\frac{1}{f(t)}\frac{df(t)}{dt} = \frac{1}{\psi(\vec{r})}\left[-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r}) + \hat{V}(\vec{r})\psi(\vec{r})\right]. \tag{3.62}$$

Since the left-hand side depends only on time and the right-hand side depends only on $\vec{r}$, both sides must be equal to a constant; this constant, which we denote by E, has the dimensions of energy. We can therefore break (3.62) into two separate differential equations, one depending on time only,

$$i\hbar\frac{df(t)}{dt} = Ef(t), \tag{3.63}$$

and the other on the space variable $\vec{r}$,

$$\boxed{\left[-\frac{\hbar^2}{2m}\nabla^2 + \hat{V}(\vec{r})\right]\psi(\vec{r}) = E\psi(\vec{r}).} \tag{3.64}$$

This equation is known as the *time-independent* Schrödinger equation for a particle of mass m moving in a time-independent potential $\hat{V}(\vec{r})$.

The solutions to (3.63) can be written as $f(t) = e^{-iEt/\hbar}$; hence the state (3.61) becomes

$$\Psi(\vec{r}, t) = \psi(\vec{r})e^{-iEt/\hbar}. \tag{3.65}$$

This particular solution of the Schrödinger equation (3.60) for a *time-independent potential* is called a *stationary state*. Why is this state called *stationary*? The reason is obvious: the probability density is stationary, i.e., it does not depend on time:

$$|\Psi(\vec{r}, t)|^2 = |\psi(\vec{r})e^{-iEt/\hbar}|^2 = |\psi(\vec{r})|^2. \tag{3.66}$$

Note that such a state has a precise value for the energy, $E = \hbar\omega$.

In summary, stationary states, which are given by the solutions of (3.64), exist only for time-independent potentials. The set of energy levels that are solutions to this equation are called the *energy spectrum* of the system. The states corresponding to discrete and continuous spectra are called *bound* and *unbound* states, respectively. We will consider these questions in detail in Chapter 4.

The most general solution to the time-dependent Schrödinger equation (3.60) can be written as an expansion in terms of the stationary states $\psi_n(\vec{r})\exp(-iE_nt/\hbar)$:

$$\Psi(\vec{r}, t) = \sum_n c_n \psi_n(\vec{r}) \exp\left(-\frac{iE_nt}{\hbar}\right), \tag{3.67}$$

where $c_n = \langle\psi_n|\Psi(t=0)\rangle = \int \psi_n^*(\vec{r})\psi(\vec{r})\, d^3r$. The general solution (3.67) is not a stationary state, because a linear superposition of stationary states is not necessarily a stationary state.

Remark

The time-dependent and time-independent Schrödinger equations are given in one dimension by (see (3.60) and (3.64))

$$\boxed{i\hbar\frac{\partial\Psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + \hat{V}(x,t)\Psi(x,t),} \tag{3.68}$$

$$\boxed{-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + \hat{V}(x)\psi(x) = E\psi(x).} \tag{3.69}$$

3.6.3 Schrödinger Equation and Wave Packets

Can we derive the Schrödinger equation (3.5) formally from first principles? No, we cannot; we can only postulate it. What we can do, however, is to provide an educated guess on the formal steps leading to it. *Wave packets* offer the formal tool to achieve that. We are going to show how to start from a wave packet and end up with the Schrödinger equation.

As seen in Chapter 1, the wave packet representing a particle of energy E and momentum p moving in a potential V is given by

$$\begin{aligned}\Psi(x,t) &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty}\tilde{\phi}(p)\exp\left[\frac{i}{\hbar}(px - Et)\right]dp \\ &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty}\tilde{\phi}(p)\exp\left[\frac{i}{\hbar}\left(px - \left(\frac{p^2}{2m} + V\right)t\right)\right]dp;\end{aligned} \tag{3.70}$$

recall that wave packets unify the corpuscular (E and p) and the wave (k and ω) features of particles: $k = p/\hbar$, $\hbar\omega = E = p^2/(2m) + V$. A partial time derivative of (3.70) yields

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty}\tilde{\phi}(p)\left(\frac{p^2}{2m} + V\right)\exp\left[\frac{i}{\hbar}\left(px - \left(\frac{p^2}{2m} + V\right)t\right)\right]dp. \tag{3.71}$$

Since $p^2/(2m) = -(\hbar^2/2m)\partial^2/\partial x^2$ and assuming that V is constant, we can take the term $-(\hbar^2/2m)\partial^2/\partial x^2 + V$ outside the integral sign, for it does not depend on p:

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V\right)\frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty}\tilde{\phi}(p)\exp\left[\frac{i}{\hbar}\left(px - \left(\frac{p^2}{2m} + V\right)t\right)\right]dp. \tag{3.72}$$

This can be written as

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V\right]\Psi(x,t). \tag{3.73}$$

Now, since this equation is valid for spatially varying potentials $V = V(x)$, we see that we have ended up with the Schrödinger equation (3.68).

3.6.4 The Conservation of Probability

Since the Hamiltonian operator is Hermitian, we can show that the norm $\langle\Psi(t)|\Psi(t)\rangle$, which is given by

$$\langle\Psi(t)|\Psi(t)\rangle = \int |\Psi(\vec{r},t)|^2\, d^3r, \tag{3.74}$$

is time independent. This means, if $|\Psi(t)\rangle$ is normalized, it stays normalized for all subsequent times. This is a direct consequence of the hermiticity of $\hat{H}$.

To prove that $\langle\Psi(t)|\Psi(t)\rangle$ is constant, we need simply to show that its time derivative is zero. First, the time derivative of $\langle\Psi(t)|\Psi(t)\rangle$ is

$$\frac{d}{dt}\langle\Psi(t)|\Psi(t)\rangle = \left(\frac{d}{dt}\langle\Psi(t)|\right)|\Psi(t)\rangle + \langle\Psi(t)|\left(\frac{d|\Psi(t)\rangle}{dt}\right), \tag{3.75}$$

where $d|\Psi(t)\rangle/dt$ and $d\langle\Psi(t)|/dt$ can be obtained from (3.5):

$$\frac{d}{dt}|\Psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\Psi(t)\rangle, \tag{3.76}$$

$$\frac{d}{dt}\langle\Psi(t)| = \frac{i}{\hbar}\langle\Psi(t)|\hat{H}^\dagger = \frac{i}{\hbar}\langle\Psi(t)|\hat{H}. \tag{3.77}$$

Inserting these two equations into (3.75), we end up with

$$\frac{d}{dt}\langle\Psi(t)|\Psi(t)\rangle = \left(\frac{i}{\hbar} - \frac{i}{\hbar}\right)\langle\Psi(t)|\hat{H}|\Psi(t)\rangle = 0. \tag{3.78}$$

Thus, the probability density $\langle\Psi|\Psi\rangle$ does not evolve in time.

In what follows we are going to calculate the probability density in the position representation. For this, we need to invoke the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi(\vec{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{r},t) + \hat{V}(\vec{r},t)\Psi(\vec{r},t) \tag{3.79}$$

and its complex conjugate

$$-i\hbar\frac{\partial\Psi^*(\vec{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi^*(\vec{r},t) + \hat{V}(\vec{r},t)\Psi^*(\vec{r},t). \tag{3.80}$$

Multiplying both sides of (3.79) by $\Psi^*(\vec{r}, t)$ and both sides of (3.80) by $\Psi(\vec{r}, t)$, and subtracting the two resulting equations, we obtain

$$i\hbar\frac{\partial}{\partial t}\left[\Psi^*(\vec{r}, t)\Psi(\vec{r}, t)\right] = -\frac{\hbar^2}{2m}\left[\Psi^*(\vec{r}, t)\nabla^2\Psi(\vec{r}, t) - \Psi\nabla^2\Psi^*\right]. \tag{3.81}$$

We can rewrite this equation as

$$\boxed{\frac{\partial\rho(\vec{r}, t)}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0,} \tag{3.82}$$

where $\rho(\vec{r}, t)$ and $\vec{J}$ are given by

$$\boxed{\rho(\vec{r}, t) = \Psi^*(\vec{r}, t)\Psi(\vec{r}, t), \qquad \vec{J}(\vec{r}, t) = \frac{i\hbar}{2m}\left(\Psi\vec{\nabla}\Psi^* - \Psi^*\vec{\nabla}\Psi\right);} \tag{3.83}$$

$\rho(\vec{r}, t)$ is called the *probability density*, while $\vec{J}(\vec{r}, t)$ is the *probability current density*, or simply the *current density*, or even the particle density flux. By analogy with charge conservation in electrodynamics, equation (3.82) is interpreted as the *conservation of probability*.

Let us find the relationship between the density operators $\hat{\rho}(t)$ and $\hat{\rho}(t_0)$. Since $|\Psi(t)\rangle = \hat{U}(t, t_0)|\Psi(t_0)\rangle$ and $\langle\Psi(t)| = \langle\Psi(t_0)|\hat{U}^\dagger(t, t_0)$, we have

$$\hat{\rho}(t) = |\Psi(t)\rangle\langle\Psi(t)| = \hat{U}(t, t_0)|\Psi(0)\rangle\langle\Psi(0)|\hat{U}^\dagger(t, t_0). \tag{3.84}$$

This is known as the *density operator* for the state $|\Psi(t)\rangle$. Hence knowing $\hat{\rho}(t_0)$ we can calculate $\hat{\rho}(t)$ as follows:

$$\hat{\rho}(t) = \hat{U}(t, t_0)\hat{\rho}(t_0)\hat{U}^\dagger(t, t_0). \tag{3.85}$$

3.6.5 Time Evolution of Expectation Values

We want to look here at the time dependence of the expectation value of a linear operator; if the state $|\Psi(t)\rangle$ is normalized, the expectation value is given by

$$\langle\hat{A}\rangle = \langle\Psi(t)|\hat{A}|\Psi(t)\rangle. \tag{3.86}$$

Using (3.76) and (3.77), we can write $d\langle\hat{A}\rangle/dt$ as follows:

$$\frac{d}{dt}\langle\hat{A}\rangle = \frac{1}{i\hbar}\langle\Psi(t)|\hat{A}\hat{H} - \hat{H}\hat{A}|\Psi(t)\rangle + \langle\Psi(t)|\frac{\partial A}{\partial t}|\Psi(t)\rangle \tag{3.87}$$

or

$$\boxed{\frac{d}{dt}\langle\hat{A}\rangle = \frac{1}{i\hbar}\langle[\hat{A}, \hat{H}]\rangle + \langle\frac{\partial\hat{A}}{\partial t}\rangle.} \tag{3.88}$$

Two important results stem from this relation. First, if the observable A does not depend explicitly on time, the term $\partial\hat{A}/\partial t$ will vanish, so the rate of change of the expectation value of $\hat{A}$ is given by $\langle[\hat{A}, \hat{H}]\rangle/i\hbar$. Second, besides not depending explicitly on time, if the observable A commutes with the Hamiltonian, the quantity $d\langle\hat{A}\rangle/dt$ will then be zero; hence the expectation value $\langle \hat{A} \rangle$ will be constant in time. So if $\hat{A}$ commutes with the Hamiltonian and is not dependent on time, the observable A is said to be a *constant of the motion*; that is, *the expectation value of an operator that does not depend on time and that commutes with the Hamiltonian is constant in time*:

$$\boxed{\text{If} \quad [\hat{H}, \hat{A}] = 0 \quad \text{and} \quad \frac{\partial \hat{A}}{\partial t} = 0 \implies \frac{d\langle \hat{A} \rangle}{dt} = 0 \implies \langle \hat{A} \rangle = \text{constant.}} \tag{3.89}$$

For instance, we can verify that the energy, the linear momentum, and the angular momentum of an isolated system are conserved: $d\langle \hat{H} \rangle/dt = 0$, $d\langle \hat{\vec{P}} \rangle/dt = 0$, and $d\langle \hat{\vec{L}} \rangle/dt = 0$. This implies that the expectation values of $\hat{H}$, $\hat{\vec{P}}$, and $\hat{\vec{L}}$ are constant. Recall from classical physics that the conservation of energy, linear momentum, and angular momentum are consequences of the following symmetries, respectively: homogeneity of time, homogeneity of space, and isotropy of space. We will show in the following section that these symmetries are associated, respectively, with invariances in time translation, space translation, and space rotation.

As an example, let us consider the time evolution of the expectation value of the density operator $\hat{\rho}(t) = |\Psi(t)\rangle\langle\Psi(t)|$; see (3.84). From (3.5), which leads to $\partial|\Psi(t)\rangle/\partial t = (1/i\hbar)\hat{H}|\Psi(t)\rangle$ and $\partial\langle\Psi(t)|/\partial t = -(1/i\hbar)\langle\Psi(t)|\hat{H}$, we have

$$\frac{\partial \hat{\rho}(t)}{\partial t} = \frac{1}{i\hbar}\hat{H}|\Psi(t)\rangle\langle\Psi(t)| - \frac{1}{i\hbar}|\Psi(t)\rangle\langle\Psi(t)|\hat{H} = -\frac{1}{i\hbar}[\hat{\rho}(t), \hat{H}]. \tag{3.90}$$

A substitution of this relation into (3.88) leads to

$$\frac{d}{dt}\langle \hat{\rho}(t) \rangle = \frac{1}{i\hbar}\langle [\hat{\rho}(t), \hat{H}] \rangle + \langle \frac{\partial \hat{\rho}(t)}{\partial t} \rangle = \frac{1}{i\hbar}\langle [\hat{\rho}(t), \hat{H}] \rangle - \frac{1}{i\hbar}\langle [\hat{\rho}(t), \hat{H}] \rangle = 0. \tag{3.91}$$

So the density operator is a constant of the motion. In fact, we can easily show that

$$\begin{aligned} \langle [\hat{\rho}(t), \hat{H}] \rangle &= \langle\Psi(t)|[|\Psi(t)\rangle\langle\Psi(t)|, \hat{H}]|\Psi(t)\rangle \\ &= \langle\Psi(t)|\Psi(t)\rangle\langle\Psi(t)|\hat{H}|\Psi(t)\rangle - \langle\Psi(t)|\hat{H}|\Psi(t)\rangle\langle\Psi(t)|\Psi(t)\rangle \\ &= 0, \end{aligned} \tag{3.92}$$

which, when combined with (3.90), yields $\langle \partial\hat{\rho}(t)/\partial t \rangle = 0$.

Finally, we should note that the constants of motion are nothing but observables that can be measured simultaneously with the energy to arbitrary accuracy. If a system has a complete set of commuting operators (CSCO), the number of these operators is given by the total number of constants of the motion.

3.7 Symmetries and Conservation Laws

We are interested here in symmetries that leave the Hamiltonian of an *isolated* system invariant. We will show that for each such symmetry there corresponds an observable which is a constant of the motion. The invariance principles relevant to our study are the time translation invariance and the space translation invariance. We may recall from classical physics that whenever a system is invariant under space translations, its total momentum is conserved; and whenever it is invariant under rotations, its total angular momentum is also conserved.

To prepare the stage for symmetries and conservation laws in quantum mechanics, we are going to examine the properties of infinitesimal and finite unitary transformations that are most essential to these invariance principles.

3.7.1 Infinitesimal Unitary Transformations

In Chapter 2 we saw that the transformations of a state vector $|\psi\rangle$ and an operator $\hat{A}$ under an infinitesimal unitary transformation $U_\varepsilon(\hat{G}) = \hat{I} + i\varepsilon\hat{G}$ are given by

$$
\begin{aligned}
|\psi'\rangle &= (\hat{I} + i\varepsilon\hat{G})|\psi\rangle = |\psi\rangle + \delta|\psi\rangle, && (3.93)\\
\hat{A}' &= (\hat{I} + i\varepsilon\hat{G})\hat{A}(\hat{I} - i\varepsilon\hat{G}) \simeq \hat{A} + i\varepsilon[\hat{G}, \hat{A}], && (3.94)
\end{aligned}
$$

where ε and $\hat{G}$ are called the parameter and the generator of the transformation, respectively.

Let us consider two important applications of infinitesimal unitary transformations: time and space translations.

3.7.1.1 Time Translations: $\hat{G} = \hat{H}/\hbar$

The application of $\hat{U}_{\delta t}(\hat{H}) = \hat{I} + (i/\hbar)\delta t\,\hat{H}$ on a state $|\psi(t)\rangle$ gives

$$\left(\hat{I} + \frac{i}{\hbar}\,\delta t\,\hat{H}\right)|\psi(t)\rangle = |\psi(t)\rangle + \left(\frac{i}{\hbar}\delta t\right)\hat{H}|\psi(t)\rangle. \tag{3.95}$$

Since $\hat{H}|\psi(t)\rangle = i\hbar\partial|\psi(t)\rangle/\partial t$ we have

$$\boxed{\left(\hat{I} + \frac{i}{\hbar}\,\delta t\,\hat{H}\right)|\psi(t)\rangle = |\psi(t)\rangle - \delta t\frac{\partial|\psi(t)\rangle}{\partial t} \simeq |\psi(t-\delta t)\rangle,} \tag{3.96}$$

because $|\psi(t)\rangle - \delta t\,\partial|\psi(t)\rangle/\partial t$ is nothing but the first-order Taylor expansion of $|\psi(t-\delta t)\rangle$. We conclude from (3.96) that the application of $\hat{U}_{\delta t}(\hat{H})$ to $|\psi(t)\rangle$ generates a state $|\psi(t-\delta t)\rangle$ which consists simply of a *time translation* of $|\psi(t)\rangle$ by an amount equal to δt. *The Hamiltonian in* $(\hat{I}+(i/\hbar)\delta t\,\hat{H})$ *is thus the generator of infinitesimal time translations*. Note that this translation preserves the shape of the state $|\psi(t)\rangle$, for its overall shape is merely translated in time by δt.

3.7.1.2 Spatial Translations: $\hat{G} = \hat{P}_x/\hbar$

The application of $\hat{U}_\varepsilon(\hat{P}_x) = \hat{I} + (i/\hbar)\varepsilon\hat{P}_x$ to $\psi(x)$ gives

$$\left(\hat{I} + \frac{i}{\hbar}\varepsilon\hat{P}_x\right)\psi(x) = \psi(x) + \left(\frac{i}{\hbar}\varepsilon\right)\hat{P}_x\psi(x). \tag{3.97}$$

Since $\hat{P}_x = -i\hbar\partial/\partial x$ and since the first-order Taylor expansion of $\psi(x+\varepsilon)$ is given by $\psi(x+\varepsilon) = \psi(x) + \varepsilon\partial\psi(x)/\partial x$, we have

$$\boxed{\left(\hat{I} + \frac{i}{\hbar}\varepsilon\hat{P}_x\right)\psi(x) = \psi(x) + \varepsilon\frac{\partial\psi(x)}{\partial x} \simeq \psi(x+\varepsilon).} \tag{3.98}$$

So, when $\hat{U}_\varepsilon(\hat{P}_x)$ acts on a wave function, it translates it spatially by an amount equal to ε.

Using $[\hat{X},\ \hat{P}_x] = i\hbar$ we infer from (3.94) that the position operator $\hat{X}$ transforms as follows:

$$\hat{X}' = \left(\hat{I} + \frac{i}{\hbar}\varepsilon\hat{P}_x\right)\hat{X}\left(\hat{I} - \frac{i}{\hbar}\varepsilon\hat{P}_x\right) \simeq \hat{X} + \frac{i}{\hbar}\varepsilon[\hat{P}_x, \hat{X}] = \hat{X} + \varepsilon. \tag{3.99}$$

The relations (3.98) and (3.99) show that the *linear momentum operator in* $(\hat{I} + (i/\hbar)\varepsilon\hat{P}_x)$ *is a generator of infinitesimal spatial translations*.

3.7.2 Finite Unitary Transformations

In Chapter 2 we saw that a *finite* unitary transformation can be constructed by performing a succession of infinitesimal transformations. For instance, by applying a single infinitesimal time translation N times in steps of τ/N, we can generate a finite time translation

$$\hat{U}_\tau(\hat{H}) = \lim_{N\to+\infty}\prod_{k=1}^{N}\left(\hat{I} + \frac{i}{\hbar}\frac{\tau}{N}\hat{H}\right) = \lim_{N\to+\infty}\left(\hat{I} + \frac{i}{\hbar}\tau\hat{H}\right)^N = \exp\left(\frac{i}{\hbar}\tau\hat{H}\right), \quad (3.100)$$

where the Hamiltonian is the generator of finite time translations. We should note that the time evolution operator $\hat{U}(t, t_0) = e^{-i(t-t_0)\hat{H}/\hbar}$, displayed in (3.58), represents a finite unitary transformation where $\hat{H}$ is the generator of the time translation.

By analogy with (3.96) we can show that the application of $\hat{U}_\tau(\hat{H})$ to $|\psi(t)\rangle$ yields

$$\boxed{\hat{U}_\tau(\hat{H})|\psi(t)\rangle = \exp\left(\frac{i}{\hbar}\tau\hat{H}\right)|\psi(t)\rangle = |\psi(t-\tau)\rangle,} \quad (3.101)$$

where $|\psi(t-\tau)\rangle$ is merely a time translation of $|\psi(t)\rangle$.

Similarly, we can infer from (3.98) that the application of $\hat{U}_{\vec{a}}(\hat{\vec{P}}) = \exp(i\vec{a}\cdot\hat{\vec{P}}/\hbar)$ to a wave function causes it to be translated in space by a vector $\vec{a}$:

$$\boxed{\hat{U}_{\vec{a}}(\hat{\vec{P}})\psi(\vec{r}) = \exp\left(\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\psi(\vec{r}) = \psi(\vec{r}+\vec{a}).} \quad (3.102)$$

To calculate the transformed position vector operator $\hat{\vec{R}}'$, let us invoke a relation we derived in Chapter 2:

$$\hat{A}' = e^{i\alpha\hat{G}}\hat{A}e^{-i\alpha\hat{G}} = \hat{A} + i\alpha[\hat{G}, \hat{A}] + \frac{(i\alpha)^2}{2!}[\hat{G}, [\hat{G}, \hat{A}]] + \frac{(i\alpha)^3}{3!}[\hat{G}, [\hat{G}, [\hat{G}, \hat{A}]]] + \cdots. \quad (3.103)$$

An application of this relation to the spatial translation operator $\hat{U}_{\vec{a}}(\hat{\vec{P}})$ yields

$$\hat{\vec{R}}' = \exp\left(\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\hat{\vec{R}}\exp\left(-\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right) = \hat{\vec{R}} + \frac{i}{\hbar}[\vec{a}\cdot\hat{\vec{P}}, \hat{\vec{R}}] = \hat{\vec{R}} + \vec{a}. \quad (3.104)$$

In deriving this, we have used the fact that $[\vec{a}\cdot\hat{\vec{P}}, \hat{\vec{R}}] = -i\hbar\vec{a}$ and that the other commutators are zero, notably $[\vec{a}\cdot\hat{\vec{P}}, [\vec{a}\cdot\hat{\vec{P}}, \hat{\vec{R}}]] = 0$. From (3.102) and (3.104), we see that *the linear momentum in* $\exp(i\vec{a}\cdot\hat{\vec{P}}/\hbar)$ *is a generator of finite spatial translations*.

3.7.3 Symmetries and Conservation Laws

We want to show here that every invariance principle of $\hat{H}$ is connected with a conservation law.

The Hamiltonian of a system transforms under a unitary transformation $e^{i\alpha\hat{G}}$ as follows; see (3.103):

$$\hat{H}' = e^{i\alpha\hat{G}}\hat{H}e^{-i\alpha\hat{G}} = \hat{H} + i\alpha[\hat{G}, \hat{H}] + \frac{(i\alpha)^2}{2!}[\hat{G}, [\hat{G}, \hat{H}]] + \frac{(i\alpha)^3}{3!}[\hat{G}, [\hat{G}, [\hat{G}, \hat{H}]]] + \cdots. \quad (3.105)$$

If $\hat{H}$ commutes with $\hat{G}$, it also commutes with the unitary transformation $\hat{U}_\alpha(\hat{G}) = e^{i\alpha\hat{G}}$. In this case we may infer two important conclusions. On the one hand, there is an *invariance principle*: the Hamiltonian is invariant under the transformation $\hat{U}_\alpha(\hat{G})$, since

$$\hat{H}' = e^{i\alpha\hat{G}}\hat{H}e^{-i\alpha\hat{G}} = e^{i\alpha\hat{G}}e^{-i\alpha\hat{G}}\hat{H} = \hat{H}. \tag{3.106}$$

On the other hand, if in addition to $[\hat{G},\ \hat{H}] = 0$, the operator $\hat{G}$ does not depend on time explicitly, there is a *conservation law*: equation (3.88) shows that $\hat{G}$ is a *constant of the motion*, since

$$\frac{d}{dt}\langle\hat{G}\rangle = \frac{1}{i\hbar}\langle[\hat{G}, \hat{H}]\rangle + \langle\frac{\partial\hat{G}}{\partial t}\rangle = 0. \tag{3.107}$$

We say that $\hat{G}$ is conserved.

So whenever the Hamiltonian is invariant under a unitary transformation, the generator of the transformation is conserved. We may say, in general, that *for every invariance symmetry of the Hamiltonian, there corresponds a conservation law.*

3.7.3.1 Conservation of Energy and Linear Momentum

Let us consider two interesting applications pertaining to the invariance of the Hamiltonian of an *isolated* system with respect to time translations and to space translations. First, let us consider time translations. As shown in (3.58), time translations are generated in the case of time-independent Hamiltonians by the evolution operator $\hat{U}(t, t_0) = e^{-i(t-t_0)\hat{H}/\hbar}$. Since $\hat{H}$ commutes with the generator of the time translation (which is given by $\hat{H}$ itself), it is invariant under time translations. *As $\hat{H}$ is invariant under time translations, the energy of an isolated system is conserved*. We should note that if the system is invariant under time translations, this means there is a symmetry of time homogeneity. Time homogeneity implies that the time-displaced state $\psi(t-\tau)$, like $\psi(t)$, satisfies the Schrödinger equation.

The second application pertains to the spatial translations, or to transformations under $\hat{U}_{\vec{a}}(\hat{\vec{P}}) = \exp(i\vec{a}\cdot\hat{\vec{P}}/\hbar)$, of an isolated system. The linear momentum is invariant under $\hat{U}_{\vec{a}}(\hat{\vec{P}})$ and the position operator transforms according to (3.104):

$$\hat{\vec{P}}\,' = \hat{\vec{P}}, \qquad \hat{\vec{R}}\,' = \hat{\vec{R}} + \vec{a}. \tag{3.108}$$

For instance, since the Hamiltonian of a free particle does not depend on the coordinates, it commutes with the linear momentum $[\hat{H}, \hat{\vec{P}}] = 0$. The Hamiltonian is then invariant under spatial translations, since

$$\hat{H}' = \exp\left(\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\hat{H}\exp\left(-\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right) = \exp\left(\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\exp\left(-\frac{i}{\hbar}\vec{a}\cdot\hat{\vec{P}}\right)\hat{H} = \hat{H}. \tag{3.109}$$

Since $[\hat{H}, \hat{\vec{P}}] = 0$ and since the linear momentum operator does not depend explicitly on time, we infer from (3.88) that $\hat{\vec{P}}$ is a constant of the motion, since

$$\frac{d}{dt}\langle\hat{\vec{P}}\rangle = \frac{1}{i\hbar}\langle[\hat{\vec{P}}, \hat{H}]\rangle + \langle\frac{\partial\hat{\vec{P}}}{\partial t}\rangle = 0. \tag{3.110}$$

So if $[\hat{H},\ \hat{\vec{P}}] = 0$ the *Hamiltonian will be invariant under spatial translations and the linear momentum will be conserved*. A more general case where the linear momentum is a constant of the motion is provided by an isolated system, for its total linear momentum is conserved. Note that the invariance of the system under spatial translations means there is a symmetry of spatial homogeneity. The requirement for the homogeneity of space implies that the spatially displaced wave function $\psi(\vec{r}+\vec{a})$, much like $\psi(\vec{r})$, satisfies the Schrödinger equation.

In summary, the symmetry of time homogeneity gives rise to the conservation of energy, whereas the symmetry of space homogeneity gives rise to the conservation of linear momentum.

In Chapter 7 we will see that the symmetry of space isotropy, or the invariance of the Hamiltonian with respect to space rotations, leads to conservation of the angular momentum.

Parity operator

The unitary transformations we have considered so far, time translations and space translations, are *continuous*. We may consider now a *discrete* unitary transformation, the *parity*. As seen in Chapter 2, the parity transformation consists of an inversion or reflection through the origin of the coordinate system:

$$\hat{\mathcal{P}}\psi(\vec{r}) = \psi(-\vec{r}). \tag{3.111}$$

If the parity operator commutes with the system's Hamiltonian,

$$[\hat{H},\ \hat{\mathcal{P}}] = 0, \tag{3.112}$$

the *parity will be conserved*, and hence a constant of the motion. In this case the Hamiltonian and the parity operator have simultaneous eigenstates. For instance, we will see in Chapter 4 that the wave functions of a particle moving in a symmetric potential, $\hat{V}(\vec{r}) = \hat{V}(-\vec{r})$, have definite parities: they can be only even or odd. Similarly, we can ascertain that the parity of an isolated system is a constant of the motion.

3.8 Connecting Quantum to Classical Mechanics

3.8.1 Poisson Brackets and Commutators

To establish a connection between quantum mechanics and classical mechanics, we may look at the time evolution of observables.

Before describing the time evolution of a dynamical variable within the context of classical mechanics, let us review the main ideas of the mathematical tool relevant to this description, the *Poisson bracket*. The Poisson bracket between two dynamical variables A and B is defined in terms of the generalized coordinates q_i and the momenta p_i of the system:

$$\{A, B\} = \sum_j \left(\frac{\partial A}{\partial q_j}\frac{\partial B}{\partial p_j} - \frac{\partial A}{\partial p_j}\frac{\partial B}{\partial q_j} \right). \tag{3.113}$$

Since the variables q_i are independent of p_i, we have $\partial q_j/\partial p_k = 0$, $\partial p_j/\partial q_k = 0$; thus we can show that

$$\{q_j, q_k\} = \{p_j, p_k\} = 0, \qquad \{q_j, p_k\} = \delta_{jk}. \tag{3.114}$$

Using (3.113) we can easily infer the following properties of the Poisson brackets:

- Antisymmetry

$$\{A,\ B\} = -\{B,\ A\} \tag{3.115}$$

- Linearity

$$\{A,\ \alpha B + \beta C + \gamma D + \cdots\} = \alpha\{A,\ B\} + \beta\{A,\ C\} + \gamma\{A,\ D\} + \cdots \tag{3.116}$$

- Complex conjugate

$$\{A,\ B\}^* = \{A^*,\ B^*\} \tag{3.117}$$

- Distributivity

$$\{A,\ BC\} = \{A,\ B\}C + B\{A,\ C\}, \qquad \{AB,\ C\} = A\{B,\ C\} + \{A,\ C\}B \tag{3.118}$$

- Jacobi identity

$$\{A, \{B,\ C\}\} + \{B,\ \{C,\ A\}\} + \{C,\ \{A,\ B\}\} = 0 \tag{3.119}$$

- Using $df^n(x)/dx = nf^{n-1}(x)df(x)/dx$, we can show that

$$\{A,\ B^n\} = nB^{n-1}\{A,\ B\}, \qquad \{A^n,\ B\} = nA^{n-1}\{A,\ B\} \tag{3.120}$$

These properties are similar to the properties of the quantum mechanical commutators seen in Chapter 2.

The total time derivative of a dynamical variable A is given by

$$\frac{dA}{dt} = \sum_j \left(\frac{\partial A}{\partial q_j}\frac{\partial q_j}{\partial t} + \frac{\partial A}{\partial p_j}\frac{\partial p_j}{\partial t}\right) + \frac{\partial A}{\partial t} = \sum_j \left(\frac{\partial A}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial A}{\partial p_j}\frac{\partial H}{\partial p_j}\right) + \frac{\partial A}{\partial t}; \tag{3.121}$$

in deriving this relation we have used the Hamilton equations of classical mechanics:

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}, \qquad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}, \tag{3.122}$$

where H is the Hamiltonian of the system. The total time evolution of a dynamical variable A is thus given by the following equation of motion:

$$\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t}. \tag{3.123}$$

Note that if A does not depend explicitly on time, its time evolution is given simply by $dA/dt = \{A, H\}$. If $dA/dt = 0$ or $\{A, H\} = 0$, A is said to be a *constant of the motion*.

Comparing the classical relation (3.123) with its quantum mechanical counterpart (3.88),

$$\frac{d}{dt}\langle \hat{A}\rangle = \frac{1}{i\hbar}\langle[\hat{A}, \hat{H}]\rangle + \langle\frac{\partial \hat{A}}{\partial t}\rangle, \tag{3.124}$$

we see that they are identical only if we identify the Poisson bracket $\{A, H\}$ with the commutator $[\hat{A}, \hat{H}]/(i\hbar)$. We may thus infer the following general rule. The Poisson bracket of any pair of classical variables can be obtained from the commutator between the corresponding pair of quantum operators by dividing it by $i\hbar$:

$$\boxed{\frac{1}{i\hbar}[\hat{A}, \hat{B}] \longrightarrow \{A, B\}_{classical}.} \tag{3.125}$$

Note that the expressions of classical mechanics can be derived from their quantum counterparts, but the opposite is not possible. That is, dividing quantum mechanical expressions by $i\hbar$ leads to their classical analog, but multiplying classical mechanical expressions by $i\hbar$ doesn't necessarily lead to their quantum counterparts.

Example 3.5

(a) Evaluate the Poisson bracket $\{x, p\}$ between the position, x, and momentum, p, variables.

(b) Compare the commutator $\left[\hat{X}, \hat{P}\right]$ with Poisson bracket $\{x, p\}$ calculated in Part (a).

Solution

(a) Applying the general relation

$$\{A, B\} = \sum_j \left(\frac{\partial A}{\partial x_j}\frac{\partial B}{\partial p_j} - \frac{\partial A}{\partial p_j}\frac{\partial B}{\partial x_j} \right) \tag{3.126}$$

to x and p, we can readily evaluate the given Poisson bracket:

$$\begin{aligned} \{x, p\} &= \frac{\partial(x)}{\partial x}\frac{\partial(p)}{\partial p} - \frac{\partial(x)}{\partial p}\frac{\partial(p)}{\partial x} \\ &= \frac{\partial(x)}{\partial x}\frac{\partial(p)}{\partial p} \\ &= 1. \end{aligned} \tag{3.127}$$

(b) Using the fact that $[\hat{X}, \hat{P}] = i\hbar$, we see that

$$\frac{1}{i\hbar}[\hat{X}, \hat{P}] = 1, \tag{3.128}$$

which is equal to the Poisson bracket (3.127); that is,

$$\frac{1}{i\hbar}[\hat{X}, \hat{P}] = \{x, p\}_{classical} = 1. \tag{3.129}$$

This result is in agreement with Eq. (3.125).

3.8.2 The Ehrenfest Theorem

If quantum mechanics is to be more general than classical mechanics, it must contain classical mechanics as a limiting case. To illustrate this idea, let us look at the time evolution of the expectation values of the position and momentum operators, $\hat{\vec{R}}$ and $\hat{\vec{P}}$, of a particle moving in a potential $\hat{V}(\vec{r})$, and then compare these relations with their classical counterparts.

Since the position and the momentum observables do not depend explicitly on time, within the context of wave mechanics, the terms $\langle \partial\hat{\vec{R}}/\partial t\rangle$ and $\langle \partial\hat{\vec{P}}/\partial t\rangle$ are zero. Hence, inserting

$\hat{H} = \hat{\vec{P}}^2/(2m) + \hat{V}(\hat{\vec{R}}, t)$ into (3.88) and using the fact that $\hat{\vec{R}}$ commutes with $\hat{V}(\hat{\vec{R}}, t)$, we can write

$$\frac{d}{dt}\langle\hat{\vec{R}}\rangle = \frac{1}{i\hbar}\langle[\hat{\vec{R}}, \hat{H}]\rangle = \frac{1}{i\hbar}\langle[\hat{\vec{R}}, \frac{\hat{\vec{P}}^2}{2m} + \hat{V}(\hat{\vec{R}}, t)]\rangle = \frac{1}{2im\hbar}\langle[\hat{\vec{R}}, \hat{\vec{P}}^2]\rangle. \tag{3.130}$$

Since

$$[\hat{\vec{R}}, \hat{\vec{P}}^2] = 2i\hbar\hat{\vec{P}}, \tag{3.131}$$

we have

$$\boxed{\frac{d}{dt}\langle\hat{\vec{R}}\rangle = \frac{1}{m}\langle\hat{\vec{P}}\rangle.} \tag{3.132}$$

As for $d\langle\hat{\vec{P}}\rangle/dt$, we can infer its expression from a treatment analogous to $d\langle\hat{\vec{R}}\rangle/dt$. Using

$$[\hat{\vec{P}}, \hat{V}(\hat{\vec{R}}, t)] = -i\hbar\vec{\nabla}\hat{V}(\vec{R}, t), \tag{3.133}$$

we can write

$$\boxed{\frac{d}{dt}\langle\hat{\vec{P}}\rangle = \frac{1}{i\hbar}\langle[\hat{\vec{P}}, \hat{V}(\hat{\vec{R}}, t)]\rangle = -\langle\vec{\nabla}\hat{V}(\hat{\vec{R}}, t)\rangle.} \tag{3.134}$$

The two relations (3.132) and (3.134), expressing the time evolution of the expectation values of the position and momentum operators, are known as the *Ehrenfest theorem*, or Ehrenfest equations. Their respective forms are reminiscent of the Hamilton–Jacobi equations of classical mechanics,

$$\frac{d\vec{r}}{dt} = \frac{\vec{p}}{m}, \qquad \frac{d\vec{p}}{dt} = -\vec{\nabla}V(\vec{r}), \tag{3.135}$$

which reduce to Newton's equation of motion for a *classical* particle of mass m, position $\vec{r}$, and momentum $\vec{p}$:

$$\frac{d\vec{p}}{dt} = m\frac{d^2\vec{r}}{dt^2} = -\vec{\nabla}V(\vec{r}). \tag{3.136}$$

Notice $\hbar$ has completely disappeared in the Ehrenfest equations (3.132) and (3.134). These two equations certainly establish a connection between quantum mechanics and classical mechanics. We can, within this context, view the center of the wave packet as moving like a classical particle when subject to a potential $V(\vec{r})$.

3.8.3 Quantum Mechanics and Classical Mechanics

In Chapter 1 we focused mainly on those experimental observations which confirm the failure of classical physics at the microscopic level. We should bear in mind, however, that classical physics works perfectly well within the realm of the macroscopic world. Thus, if the theory of quantum mechanics is to be considered more general than classical physics, it must yield accurate results not only on the microscopic scale but at the classical limit as well.

How does one decide on when to use classical or quantum mechanics to describe the motion of a given system? That is, how do we know when a classical description is good enough or when a quantum description becomes a must? The answer is provided by comparing the size of those quantities of the system that have the dimensions of an action with the Planck constant, h. Since, as shown in (3.125), the quantum relations are characterized by h, we can state that if the value of the action of a system is too large compared to h, this system can be accurately described by means of classical physics. Otherwise, the use of a quantal description becomes unavoidable. One should recall that, for microscopic systems, the size of action variables is of the order of h; for instance, the angular momentum of the hydrogen atom is $L = n\hbar$, where n is finite.

Another equivalent way of defining the classical limit is by means of "*length*." Since $\lambda = h/p$ the classical domain can be specified by the limit $\lambda \to 0$. This means that, when the de Broglie wavelength of a system is too small compared to its size, the system can be described accurately by means of classical physics.

In summary, the classical limit can be described as the limit $h \to 0$ or, equivalently, as the limit $\lambda \to 0$. In these limits the results of quantum mechanics should be similar to those of classical physics:

$$\lim_{h\to 0} \text{Quantum Mechanics} \longrightarrow \text{Classical Mechanics}, \tag{3.137}$$

$$\lim_{\lambda\to 0} \text{Quantum Mechanics} \longrightarrow \text{Classical Mechanics}. \tag{3.138}$$

Classical mechanics can thus be regarded as the short wavelength limit of quantum mechanics. In this way, quantum mechanics contains classical mechanics as a limiting case. So, in the limit of $h \to 0$ or $\lambda \to 0$, quantum dynamical quantities should have, as proposed by Bohr, a one-to-one correspondence with their classical counterparts. This is the essence of the *correspondence principle*.

But how does one reconcile, in the classical limit, the probabilistic nature of quantum mechanics with the determinism of classical physics? The answer is quite straightforward: quantum *fluctuations* must become negligible or even vanish when $h \to 0$, for Heisenberg's uncertainty principle would acquire the status of *certainty*; when $h \to 0$, the fluctuations in the position and momentum will vanish, $\Delta x \to 0$ and $\Delta p \to 0$. Thus, the position and momentum can be measured simultaneously with arbitrary accuracy. This implies that the probabilistic assessments of dynamical quantities by quantum mechanics must give way to exact calculations (these ideas will be discussed further when we study the WKB method in Chapter 9).

So, for those cases where the action variables of a system are too large compared to h (or, equivalently, when the lengths of this system are too large compared to its de Broglie wavelength), quantum mechanics gives the same results as classical mechanics.

In the rest of this text, we will deal with the various applications of the Schrödinger equation. We start, in Chapter 4, with the simple case of one-dimensional systems and later on consider more realistic systems.

3.9 Solved Problems

Problem 3.1

A particle of mass m, which moves freely inside an infinite potential well of length a, has the following initial wave function at $t = 0$:

$$\psi(x, 0) = \frac{A}{\sqrt{a}} \sin\left(\frac{\pi x}{a}\right) + \sqrt{\frac{3}{5a}} \sin\left(\frac{3\pi x}{a}\right) + \frac{1}{\sqrt{5a}} \sin\left(\frac{5\pi x}{a}\right),$$

where A is a real constant.

(a) Find A so that $\psi(x, 0)$ is normalized.

(b) If measurements of the energy are carried out, what are the values that will be found and what are the corresponding probabilities? Calculate the average energy.

(c) Find the wave function $\psi(x, t)$ at any later time t.

(d) Determine the probability of finding the system at a time t in the state $\varphi(x, t) = \sqrt{2/a}\sin(5\pi x/a)\exp(-iE_5 t/\hbar)$; then determine the probability of finding it in the state $\chi(x, t) = \sqrt{2/a}\sin(2\pi x/a)\exp(-iE_2 t/\hbar)$.

Solution

Since the functions

$$\phi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi x}{a}\right) \tag{3.139}$$

are orthonormal,

$$\langle\phi_n|\phi_m\rangle = \int_0^a \phi_n^*(x)\phi_m(x)\,dx = \frac{2}{a}\int_0^a \sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi x}{a}\right)dx = \delta_{nm}, \tag{3.140}$$

it is more convenient to write $\psi(x, 0)$ in terms of $\phi_n(x)$:

$$\begin{aligned}\psi(x, 0) &= \frac{A}{\sqrt{a}}\sin\left(\frac{\pi x}{a}\right) + \sqrt{\frac{3}{5a}}\sin\left(\frac{3\pi x}{a}\right) + \frac{1}{\sqrt{5a}}\sin\left(\frac{5\pi x}{a}\right) \\ &= \frac{A}{\sqrt{2}}\phi_1(x) + \sqrt{\frac{3}{10}}\phi_3(x) + \frac{1}{\sqrt{10}}\phi_5(x).\end{aligned} \tag{3.141}$$

(a) Since $\langle\phi_n|\phi_m\rangle = \delta_{nm}$ the normalization of $\psi(x, 0)$ yields

$$1 = \langle\psi|\psi\rangle = \frac{A^2}{2} + \frac{3}{10} + \frac{1}{10}, \tag{3.142}$$

or $A = \sqrt{6/5}$; hence

$$\psi(x, 0) = \sqrt{\frac{3}{5}}\phi_1(x) + \sqrt{\frac{3}{10}}\phi_3(x) + \frac{1}{\sqrt{10}}\phi_5(x). \tag{3.143}$$

(b) Since the second derivative of (3.139) is given by $d^2\phi_n(x)/dx^2 = -(n^2\pi^2/a^2)\phi_n(x)$, and since the Hamiltonian of a free particle is $\hat{H} = -(\hbar^2/2m)d^2/dx^2$, the expectation value of $\hat{H}$ with respect to $\phi_n(x)$ is

$$E_n = \langle\phi_n|\hat{H}|\phi_n\rangle = -\frac{\hbar^2}{2m}\int_0^a \phi_n^*(x)\frac{d^2\phi_n(x)}{dx^2}dx = \frac{n^2\pi^2\hbar^2}{2ma^2}. \tag{3.144}$$

If a measurement is carried out on the system, we would obtain $E_n = n^2\pi^2\hbar^2/(2ma^2)$ with a corresponding probability of $P_n(E_n) = |\langle\phi_n|\psi\rangle|^2$. Since the initial wave function (3.143) contains only three eigenstates of $\hat{H}$, $\phi_1(x)$, $\phi_3(x)$, and $\phi_5(x)$, the results of the energy measurements along with the corresponding probabilities are

$$E_1 = \langle\phi_1|\hat{H}|\phi_1\rangle = \frac{\pi^2\hbar^2}{2ma^2}, \qquad P_1(E_1) = |\langle\phi_1|\psi\rangle|^2 = \frac{3}{5}, \tag{3.145}$$

$$E_3 = \langle\phi_3|\hat{H}|\phi_3\rangle = \frac{9\pi^2\hbar^2}{2ma^2}, \qquad P_3(E_3) = |\langle\phi_3|\psi\rangle|^2 = \frac{3}{10}, \tag{3.146}$$

$$E_5 = \langle\phi_5|\hat{H}|\phi_5\rangle = \frac{25\pi^2\hbar^2}{2ma^2}, \qquad P_5(E_5) = |\langle\phi_5|\psi\rangle|^2 = \frac{1}{10}. \tag{3.147}$$

The average energy is

$$E = \sum_n P_n E_n = \frac{3}{5}E_1 + \frac{3}{10}E_3 + \frac{1}{10}E_5 = \frac{29\pi^2\hbar^2}{10ma^2}. \tag{3.148}$$

(c) As the initial state $\psi(x, 0)$ is given by (3.143), the wave function $\psi(x, t)$ at any later time t is

$$\psi(x,t) = \sqrt{\frac{3}{5}}\phi_1(x)e^{-iE_1t/\hbar} + \sqrt{\frac{3}{10}}\phi_3(x)e^{-iE_3t/\hbar} + \frac{1}{\sqrt{10}}\phi_5(x)e^{-iE_5t/\hbar}, \tag{3.149}$$

where the expressions of E_n are listed in (3.144) and $\phi_n(x)$ in (3.139).

(d) First, let us express $\varphi(x, t)$ in terms of $\phi_n(x)$:

$$\varphi(x,t) = \sqrt{\frac{2}{a}}\sin\left(\frac{5\pi x}{a}\right)e^{-iE_5t/\hbar} = \phi_5(x)e^{-iE_5t/\hbar}. \tag{3.150}$$

The probability of finding the system at a time t in the state $\varphi(x, t)$ is

$$P = |\langle\varphi|\psi\rangle|^2 = \left|\int_0^a \varphi^*(x,t)\psi(x,t)\,dx\right|^2 = \frac{1}{10}\left|\int_0^a \phi_5^*(x)\phi_5(x)\,dx\right|^2 = \frac{1}{10}, \tag{3.151}$$

since $\langle\varphi|\phi_1\rangle = \langle\varphi|\phi_3\rangle = 0$ and $\langle\varphi|\phi_5\rangle = \exp(iE_5t/\hbar)$.

Similarly, since $\chi(x,t) = \sqrt{2/a}\sin(2\pi x/a)\exp(-iE_2t/\hbar) = \phi_2(x)\exp(-iE_2t/\hbar)$, we can easily show that the probability for finding the system in the state $\chi(x, t)$ is zero:

$$P = |\langle\chi|\psi\rangle|^2 = \left|\int_0^a \chi^*(x,t)\psi(x,t)\,dx\right|^2 = 0, \tag{3.152}$$

since $\langle\chi|\phi_1\rangle = \langle\chi|\phi_3\rangle = \langle\chi|\phi_5\rangle = 0$.

Problem 3.2

A particle of mass m, which moves freely inside an infinite potential well of length a, is initially in the state $\psi(x, 0) = \sqrt{3/5a}\sin(3\pi x/a) + (1/\sqrt{5a})\sin(5\pi x/a)$.

(a) Find $\psi(x, t)$ at any later time t.

(b) Calculate the probability density $\rho(x, t)$ and the current density, $\vec{J}(x, t)$.

(c) Verify that the probability is conserved, i.e., $\partial\rho/\partial t + \vec{\nabla}\cdot\vec{J}(x, t) = 0$.

Solution

(a) Since $\psi(x, 0)$ can be expressed in terms of $\phi_n(x) = \sqrt{2/a}\sin(n\pi x/a)$ as

$$\psi(x,0) = \sqrt{\frac{3}{5a}}\sin\left(\frac{3\pi x}{a}\right) + \frac{1}{\sqrt{5a}}\sin\left(\frac{5\pi x}{a}\right) = \sqrt{\frac{3}{10}}\phi_3(x) + \frac{1}{\sqrt{10}}\phi_5(x), \tag{3.153}$$

we can write

$$\begin{aligned}\psi(x,t) &= \sqrt{\frac{3}{5a}}\sin\left(\frac{3\pi x}{a}\right)e^{-iE_3t/\hbar} + \frac{1}{\sqrt{5a}}\sin\left(\frac{5\pi x}{a}\right)e^{-iE_5t/\hbar} \\ &= \sqrt{\frac{3}{10}}\phi_3(x)e^{-iE_3t/\hbar} + \frac{1}{\sqrt{10}}\phi_5(x)e^{-iE_5t/\hbar},\end{aligned} \tag{3.154}$$

where the expressions for E_n are listed in (3.144): $E_n = n^2\pi^2\hbar^2/(2ma^2)$.

(b) Since $\rho(x,t) = \psi^*(x,t)\psi(x,t)$, where $\psi(x,t)$ is given by (3.154), we can write

$$\rho(x,t) = \frac{3}{10}\phi_3^2(x) + \frac{\sqrt{3}}{10}\phi_3(x)\phi_5(x)\left[e^{i(E_3-E_5)t/\hbar} + e^{-i(E_3-E_5)t/\hbar}\right] + \frac{1}{10}\phi_5^2(x). \quad (3.155)$$

From (3.144) we have $E_3 - E_5 = 9E_1 - 25E_1 = -16E_1 = -8\pi^2\hbar^2/(ma^2)$. Thus, $\rho(x,t)$ becomes

$$\begin{aligned}\rho(x,t) &= \frac{3}{10}\phi_3^2(x) + \frac{\sqrt{3}}{5}\phi_3(x)\phi_5(x)\cos\left(\frac{16E_1t}{\hbar}\right) + \frac{1}{10}\phi_5^2(x)\\ &= \frac{3}{5a}\sin^2\left(\frac{3\pi x}{a}\right) + \frac{2\sqrt{3}}{5a}\sin\left(\frac{3\pi x}{a}\right)\sin\left(\frac{5\pi x}{a}\right)\cos\left(\frac{16E_1t}{\hbar}\right)\\ &\quad + \frac{1}{5a}\sin^2\left(\frac{5\pi x}{a}\right). \end{aligned} \quad (3.156)$$

Since the system is one-dimensional, the action of the gradient operator on $\psi(x,t)$ and $\psi^*(x,t)$ is given by $\vec{\nabla}\psi(x,t) = (d\psi(x,t)/dx)\vec{i}$ and $\vec{\nabla}\psi^*(x,t) = (d\psi^*(x,t)/dx)\vec{i}$. We can thus write the current density $\vec{J}(x,t) = (i\hbar/2m)\left(\psi(x,t)\vec{\nabla}\psi^*(x,t) - \psi^*(x,t)\vec{\nabla}\psi(x,t)\right)$ as

$$\vec{J}(x,t) = \frac{i\hbar}{2m}\left(\psi(x,t)\frac{d\psi^*(x,t)}{dx} - \psi^*(x,t)\frac{d\psi(x,t)}{dx}\right)\vec{i}. \quad (3.157)$$

Using (3.154) we have

$$\frac{d\psi(x,t)}{dx} = \frac{3\pi}{a}\sqrt{\frac{3}{5a}}\cos\left(\frac{3\pi x}{a}\right)e^{-iE_3t/\hbar} + \frac{5\pi}{a}\frac{1}{\sqrt{5a}}\cos\left(\frac{5\pi x}{a}\right)e^{-iE_5t/\hbar}, \quad (3.158)$$

$$\frac{d\psi^*(x,t)}{dx} = \frac{3\pi}{a}\sqrt{\frac{3}{5a}}\cos\left(\frac{3\pi x}{a}\right)e^{iE_3t/\hbar} + \frac{5\pi}{a}\frac{1}{\sqrt{5a}}\cos\left(\frac{5\pi x}{a}\right)e^{iE_5t/\hbar}. \quad (3.159)$$

A straightforward calculation yields

$$\begin{aligned}\psi\frac{d\psi^*}{dx} - \psi^*\frac{d\psi}{dx} &= -2i\pi\frac{\sqrt{3}}{5a^2}\left[5\sin\left(\frac{3\pi x}{a}\right)\cos\left(\frac{5\pi x}{a}\right) - 3\sin\left(\frac{5\pi x}{a}\right)\cos\left(\frac{3\pi x}{a}\right)\right]\\ &\quad\times\sin\left(\frac{E_3-E_5}{\hbar}t\right). \end{aligned} \quad (3.160)$$

Inserting this into (3.157) and using $E_3 - E_5 = -16E_1$, we have

$$\vec{J}(x,t) = -\frac{\pi\hbar}{m}\frac{\sqrt{3}}{5a^2}\left[5\sin\left(\frac{3\pi x}{a}\right)\cos\left(\frac{5\pi x}{a}\right) - 3\sin\left(\frac{5\pi x}{a}\right)\cos\left(\frac{3\pi x}{a}\right)\right]\sin\left(\frac{16E_1t}{\hbar}\right)\vec{i}. \quad (3.161)$$

(c) Performing the time derivative of (3.156) and using the expression $32\sqrt{3}E_1/(5a\hbar) = 16\pi^2\hbar\sqrt{3}/(5ma^3)$, since $E_1 = \pi^2\hbar^2/(2ma^2)$, we obtain

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= -\frac{32\sqrt{3}E_1}{5a\hbar}\sin\left(\frac{3\pi x}{a}\right)\sin\left(\frac{5\pi x}{a}\right)\sin\left(\frac{16E_1t}{\hbar}\right)\\ &= -\frac{16\pi^2\hbar\sqrt{3}}{5ma^3}\sin\left(\frac{3\pi x}{a}\right)\sin\left(\frac{5\pi x}{a}\right)\sin\left(\frac{16E_1t}{\hbar}\right), \end{aligned} \quad (3.162)$$

Now, taking the divergence of (3.161), we end up with

$$\vec{\nabla} \cdot \vec{J}(x,t) = \frac{dJ(x,t)}{dx} = \frac{16\pi^2\hbar\sqrt{3}}{5ma^3}\sin\left(\frac{3\pi x}{a}\right)\sin\left(\frac{5\pi x}{a}\right)\sin\left(\frac{16E_1 t}{\hbar}\right). \tag{3.163}$$

The addition of (3.162) and (3.163) confirms the conservation of probability:

$$\frac{\partial\rho}{\partial t} + \vec{\nabla} \cdot \vec{J}(x,t) = 0. \tag{3.164}$$

Problem 3.3

Consider a one-dimensional particle which is confined within the region $0 \leq x \leq a$ and whose wave function is $\Psi(x,t) = \sin(\pi x/a)\exp(-i\omega t)$.

(a) Find the potential $V(x)$.

(b) Calculate the probability of finding the particle in the interval $a/4 \leq x \leq 3a/4$.

Solution

(a) Since the first time derivative and the second x derivative of $\Psi(x,t)$ are given by $\partial\Psi(x,t)/\partial t = -i\omega\Psi(x,t)$ and $\partial^2\Psi(x,t)/\partial x^2 = -(\pi^2/a^2)\Psi(x,t)$, the Schrödinger equation (3.68) yields

$$i\hbar(-i\omega)\Psi(x,t) = \frac{\hbar^2}{2m}\frac{\pi^2}{a^2}\Psi(x,t) + \hat{V}(x,t)\Psi(x,t). \tag{3.165}$$

Hence $V(x,t)$ is time independent and given by $V(x) = \hbar\omega - \hbar^2\pi^2/(2ma^2)$.

(b) The probability of finding the particle in the interval $a/4 \leq x \leq 3a/4$ can be obtained from (3.4):

$$P = \frac{\int_{a/4}^{3a/4} |\psi(x)|^2 dx}{\int_0^a |\psi(x)|^2 dx} = \frac{\int_{a/4}^{3a/4} \sin^2(\pi x/a)\, dx}{\int_0^a \sin^2(\pi x/a)\, dx} = \frac{2+\pi}{2\pi} = 0.82 \tag{3.166}$$

Problem 3.4

A system is initially in the state $|\psi_0\rangle = [\sqrt{2}|\phi_1\rangle + \sqrt{3}|\phi_2\rangle + |\phi_3\rangle + |\phi_4\rangle]/\sqrt{7}$, where $|\phi_n\rangle$ are eigenstates of the system's Hamiltonian such that $\hat{H}|\phi_n\rangle = n^2\mathcal{E}_0|\phi_n\rangle$.

(a) If energy is measured, what values will be obtained and with what probabilities?

(b) Consider an operator $\hat{A}$ whose action on $|\phi_n\rangle$ is defined by $\hat{A}|\phi_n\rangle = (n+1)a_0|\phi_n\rangle$. If A is measured, what values will be obtained and with what probabilities?

(c) Suppose that a measurement of the energy yields $4\mathcal{E}_0$. If we measure A immediately afterwards, what value will be obtained?

Solution

(a) A measurement of the energy yields $E_n = \langle\phi_n|\hat{H}|\phi_n\rangle = n^2\mathcal{E}_0$, that is

$$E_1 = \mathcal{E}_0, \quad E_2 = 4\mathcal{E}_0, \quad E_3 = 9\mathcal{E}_0, \quad E_4 = 16\mathcal{E}_0. \tag{3.167}$$

Since $|\ \psi_0\rangle$ is normalized, $\langle\psi_0 \mid \psi_0\rangle = (2+3+1+1)/7 = 1$, and using (3.2), we can write the probabilities corresponding to (3.167) as $P(E_n) = |\langle\phi_n|\psi_0\rangle|^2 / \langle\psi_0 \mid \psi_0\rangle = |\langle\phi_n|\psi_0\rangle|^2$; hence, using the fact that $\langle\phi_n|\phi_m\rangle = \delta_{nm}$, we have

$$P(E_1) = \left|\sqrt{\frac{2}{7}}\langle\phi_1|\phi_1\rangle\right|^2 = \frac{2}{7}, \qquad P(E_2) = \left|\sqrt{\frac{3}{7}}\langle\phi_2|\phi_2\rangle\right|^2 = \frac{3}{7}, \tag{3.168}$$

$$P(E_3) = \left|\frac{1}{\sqrt{7}}\langle\phi_3|\phi_3\rangle\right|^2 = \frac{1}{7}, \qquad P(E_4) = \left|\frac{1}{\sqrt{7}}\langle\phi_4|\phi_4\rangle\right|^2 = \frac{1}{7}. \tag{3.169}$$

(b) Similarly, a measurement of the observable $\hat{A}$ yields $a_n = \langle\phi_n|\hat{A}|\phi_n\rangle = (n+1)a_0$; that is,

$$a_1 = 2a_0, \quad a_2 = 3a_0, \quad a_3 = 4a_0, \quad a_4 = 5a_0. \tag{3.170}$$

Again, using (3.2) and since $|\ \psi_0\rangle$ is normalized, we can ascertain that the probabilities corresponding to the values (3.170) are given by $P(a_n) = |\langle\phi_n|\psi_0\rangle|^2 / \langle\psi_0 \mid \psi_0\rangle = |\langle\phi_n|\psi_0\rangle|^2$, or

$$P(a_1) = \left|\sqrt{\frac{2}{7}}\langle\phi_1|\phi_1\rangle\right|^2 = \frac{2}{7}, \qquad P(a_2) = \left|\sqrt{\frac{3}{7}}\langle\phi_2|\phi_2\rangle\right|^2 = \frac{3}{7}, \tag{3.171}$$

$$P(a_3) = \left|\frac{1}{\sqrt{7}}\langle\phi_3|\phi_3\rangle\right|^2 = \frac{1}{7}, \qquad P(a_4) = \left|\frac{1}{\sqrt{7}}\langle\phi_4|\phi_4\rangle\right|^2 = \frac{1}{7}. \tag{3.172}$$

(c) An energy measurement that yields $4\mathcal{E}_0$ implies that the system is left in the state $|\phi_2\rangle$. A measurement of the observable A immediately afterwards leads to

$$\langle\phi_2|\hat{A}|\phi_2\rangle = 3a_0\langle\phi_2|\phi_2\rangle = 3a_0. \tag{3.173}$$

Problem 3.5

(a) Assuming that the system of Problem 3.4 is initially in the state $|\phi_3\rangle$, what values for the energy and the observable A will be obtained if we measure: (i)H first then A, (ii) A first then H?

(b) Compare the results obtained in (i) and (ii) and infer whether $\hat{H}$ and $\hat{A}$ are compatible. Calculate $[\hat{H},\ \hat{A}]|\phi_3\rangle$.

Solution

(a) (i) The measurement of H first then A is represented by $\hat{A}\hat{H}|\phi_3\rangle$. Using the relations $\hat{H}|\phi_n\rangle = n^2\mathcal{E}_0|\phi_n\rangle$ and $\hat{A}|\phi_n\rangle = na_0|\phi_{n+1}\rangle$, we have

$$\hat{A}\hat{H}|\phi_3\rangle = 9\mathcal{E}_0\hat{A}|\phi_3\rangle = 27\mathcal{E}_0 a_0|\phi_4\rangle. \tag{3.174}$$

(ii) Measuring A first and then H, we will obtain

$$\hat{H}\hat{A}|\phi_3\rangle = 3a_0\hat{H}|\phi_4\rangle = 48\mathcal{E}_0 a_0|\phi_4\rangle. \tag{3.175}$$

(b) Equations (3.174) and (3.175) show that the actions of $\hat{A}\hat{H}$ and $\hat{H}\hat{A}$ yield different results. This means that $\hat{H}$ and $\hat{A}$ do not commute; hence they are not compatible. We can thus write

$$[\hat{H},\ \hat{A}]|\phi_3\rangle = (48-27)\mathcal{E}_0 a_0|\phi_4\rangle = 17\mathcal{E}_0 a_0|\phi_4\rangle. \tag{3.176}$$

Problem 3.6
Consider a physical system whose Hamiltonian H and initial state $|\psi_0\rangle$ are given by

$$H = \mathcal{E}\begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad |\psi_0\rangle = \frac{1}{\sqrt{5}}\begin{pmatrix} 1-i \\ 1-i \\ 1 \end{pmatrix},$$

where $\mathcal{E}$ has the dimensions of energy.

(a) What values will we obtain when measuring the energy and with what probabilities?

(b) Calculate $\langle \hat{H}\rangle$, the expectation value of the Hamiltonian.

Solution

(a) The results of the energy measurement are given by the eigenvalues of H. A diagonalization of H yields a nondegenerate eigenenergy $E_1 = \mathcal{E}$ and a doubly degenerate value $E_2 = E_3 = -\mathcal{E}$ whose respective eigenvectors are given by

$$|\phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \qquad |\phi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix}, \qquad |\phi_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \tag{3.177}$$

these eigenvectors are orthogonal since H is Hermitian. Note that the initial state $|\psi_0\rangle$ can be written in terms of $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ as follows:

$$|\psi_0\rangle = \frac{1}{\sqrt{5}}\begin{pmatrix} 1-i \\ 1-i \\ 1 \end{pmatrix} = \sqrt{\frac{2}{5}}|\phi_1\rangle + \sqrt{\frac{2}{5}}|\phi_2\rangle + \frac{1}{\sqrt{5}}|\phi_3\rangle. \tag{3.178}$$

Since $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ are orthonormal, the probability of measuring $E_1 = \mathcal{E}$ is given by

$$P_1(E_1) = |\langle\phi_1|\psi_0\rangle|^2 = \left|\sqrt{\frac{2}{5}}\langle\phi_1|\phi_1\rangle\right|^2 = \frac{2}{5}. \tag{3.179}$$

Now, since the other eigenvalue is doubly degenerate, $E_2 = E_3 = -\mathcal{E}$, the probability of measuring $-\mathcal{E}$ can be obtained from (3.3):

$$P_2(E_2) = |\langle\phi_2|\psi_0\rangle|^2 + |\langle\phi_3|\psi_0\rangle|^2 = \frac{2}{5} + \frac{1}{5} = \frac{3}{5}. \tag{3.180}$$

(b) From (3.179) and (3.180), we have

$$\langle \hat{H}\rangle = P_1 E_1 + P_2 E_2 = \frac{2}{5}\mathcal{E} - \frac{3}{5}\mathcal{E} = -\frac{1}{5}\mathcal{E}. \tag{3.181}$$

We can obtain the same result by calculating the expectation value of $\hat{H}$ with respect to $|\psi_0\rangle$. Since $\langle\psi_0|\psi_0\rangle = 1$, we have $\langle \hat{H}\rangle = \langle\psi_0|\hat{H}|\psi_0\rangle/\langle\psi_0|\psi_0\rangle = \langle\psi_0|\hat{H}|\psi_0\rangle$:

$$\langle \hat{H}\rangle = \langle\psi_0|\hat{H}|\psi_0\rangle = \frac{\mathcal{E}}{5}\begin{pmatrix} 1+i & 1+i & 1 \end{pmatrix}\begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1-i \\ 1-i \\ 1 \end{pmatrix} = -\frac{1}{5}\mathcal{E}. \tag{3.182}$$

Problem 3.7
Consider a system whose Hamiltonian H and an operator A are given by the matrices

$$H = \mathcal{E}_0 \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad A = a \begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

where $\mathcal{E}_0$ has the dimensions of energy.

(a) If we measure the energy, what values will we obtain?

(b) Suppose that when we measure the energy, we obtain a value of $-\mathcal{E}_0$. Immediately afterwards, we measure A. What values will we obtain for A and what are the probabilities corresponding to each value?

(c) Calculate the uncertainty ΔA.

Solution

(a) The possible energies are given by the eigenvalues of H. A diagonalization of H yields three nondegenerate eigenenergies $E_1 = 0$, $E_2 = -\mathcal{E}_0$, and $E_3 = 2\mathcal{E}_0$. The respective eigenvectors are

$$|\phi_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad |\phi_2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad |\phi_3\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}; \tag{3.183}$$

these eigenvectors are orthonormal.

(b) If a measurement of the energy yields $-\mathcal{E}_0$, this means that the system is left in the state $|\phi_2\rangle$. When we measure the next observable, A, the system is in the state $|\phi_2\rangle$. The result we obtain for A is given by any of the eigenvalues of A. A diagonalization of A yields three nondegenerate values: $a_1 = -\sqrt{17}a$, $a_2 = 0$, and $a_3 = \sqrt{17}a$; their respective eigenvectors are given by

$$|a_1\rangle = \frac{1}{\sqrt{34}} \begin{pmatrix} 4 \\ -\sqrt{17} \\ 1 \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{\sqrt{17}} \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}, \qquad |a_3\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 4 \\ \sqrt{17} \\ 1 \end{pmatrix}. \tag{3.184}$$

Thus, when measuring A on a system which is in the state $|\phi_2\rangle$, the probability of finding $-\sqrt{17}a$ is given by

$$P_1(a_1) = |\langle a_1|\phi_2\rangle|^2 = \left| \frac{1}{\sqrt{34}} \begin{pmatrix} 4 & -\sqrt{17} & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right|^2 = \frac{1}{34}. \tag{3.185}$$

Similarly, the probabilities of measuring 0 and $\sqrt{17}a$ are

$$P_2(a_2) = |\langle a_2|\phi_2\rangle|^2 = \left| \frac{1}{\sqrt{17}} \begin{pmatrix} 1 & 0 & -4 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right|^2 = \frac{16}{17}, \tag{3.186}$$

$$P_3(a_3) = |\langle a_3|\phi_2\rangle|^2 = \left| \frac{1}{\sqrt{34}} \begin{pmatrix} 4 & \sqrt{17} & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right|^2 = \frac{1}{34}. \tag{3.187}$$

(c) Since the system, when measuring A is in the state $|\phi_2\rangle$, the uncertainty ΔA is given by $\Delta A = \sqrt{\langle\phi_2|A^2|\phi_2\rangle - \langle\phi_2|A|\phi_2\rangle^2}$, where

$$\langle\phi_2|A|\phi_2\rangle = a\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0, \tag{3.188}$$

$$\langle\phi_2|A^2|\phi_2\rangle = a^2\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = a^2. \tag{3.189}$$

Thus we have $\Delta A = a$.

Problem 3.8

Consider a system whose state and two observables are given by

$$|\psi(t)\rangle = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \qquad A = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(a) What is the probability that a measurement of A at time t yields -1?

(b) Let us carry out a set of two measurements where B is measured first and then, immediately afterwards, A is measured. Find the probability of obtaining a value of 0 for B and a value of 1 for A.

(c) Now we measure A first then, immediately afterwards, B. Find the probability of obtaining a value of 1 for A and a value of 0 for B.

(d) Compare the results of (b) and (c). Explain.

(e) Which among the sets of operators $\{\hat{A}\}$, $\{\hat{B}\}$, and $\{\hat{A}, \hat{B}\}$ form a complete set of commuting operators (CSCO)?

Solution

(a) A measurement of A yields any of the eigenvalues of A which are given by $a_1 = -1$, $a_2 = 0$, $a_3 = 1$; the respective (normalized) eigenstates are

$$|a_1\rangle = \frac{1}{2}\begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \qquad |a_3\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}. \tag{3.190}$$

The probability of obtaining $a_1 = -1$ is

$$P(-1) = \frac{|\langle a_1|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} = \frac{1}{6}\left|\frac{1}{2}\begin{pmatrix} -1 & \sqrt{2} & -1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}\right|^2 = \frac{1}{3}, \tag{3.191}$$

where we have used the fact that $\langle\psi(t)|\psi(t)\rangle = \begin{pmatrix} -1 & 2 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = 6$.

(b) A measurement of B yields a value which is equal to any of the eigenvalues of B: $b_1 = -1$, $b_2 = 0$, and $b_3 = 1$; their corresponding eigenvectors are

$$|b_1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad |b_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |b_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3.192}$$

Since the system was in the state $|\psi(t)\rangle$, the probability of obtaining the value $b_2 = 0$ for B is

$$P(b_2) = \frac{|\langle b_2|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} = \frac{1}{6}\left|\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}\right|^2 = \frac{2}{3}. \tag{3.193}$$

We deal now with the measurement of the other observable, A. The observables A and B do not have common eigenstates, since they do not commute. After measuring B (the result is $b_2 = 0$), the system is left, according to Postulate 3, in a state $|\phi\rangle$ which can be found by projecting $|\psi(t)\rangle$ onto $|b_2\rangle$:

$$|\phi\rangle = |b_2\rangle\langle b_2|\psi(t)\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}. \tag{3.194}$$

The probability of finding 1 when we measure A is given by

$$P(a_3) = \frac{|\langle a_3|\phi\rangle|^2}{\langle\phi|\phi\rangle} = \frac{1}{4}\left|\frac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}\right|^2 = \frac{1}{2}, \tag{3.195}$$

since $\langle\phi|\phi\rangle = 4$. In summary, when measuring B then A, the probability of finding a value of 0 for B and 1 for A is given by the product of the probabilities (3.193) and (3.195):

$$P(b_2, a_3) = P(b_2)P(a_3) = \frac{2}{3}\frac{1}{2} = \frac{1}{3}. \tag{3.196}$$

(c) Next we measure A first then B. Since the system is in the state $|\psi(t)\rangle$, the probability of measuring $a_3 = 1$ for A is given by

$$P'(a_3) = \frac{|\langle a_3|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} = \frac{1}{6}\left|\frac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}\right|^2 = \frac{1}{3}, \tag{3.197}$$

where we have used the expression (3.190) for $|a_3\rangle$.

We then proceed to the measurement of B. The state of the system just after measuring A (with a value $a_3 = 1$) is given by a projection of $|\psi(t)\rangle$ onto $|a_3\rangle$:

$$|\chi\rangle = |a_3\rangle\langle a_3|\psi(t)\rangle = \frac{1}{4}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}\begin{pmatrix} 1 & \sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}. \tag{3.198}$$

So the probability of finding a value of $b_2 = 0$ when measuring B is given by

$$P'(b_2) = \frac{|\langle b_2|\chi\rangle|^2}{\langle\chi|\chi\rangle} = \frac{1}{2}\left|\frac{\sqrt{2}}{2}\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}\right|^2 = \frac{1}{2}, \tag{3.199}$$

since $\langle\chi|\chi\rangle = 2$.

So when measuring A then B, the probability of finding a value of 1 for A and 0 for B is given by the product of the probabilities (3.199) and (3.197):

$$P(a_3, b_2) = P'(a_3)P'(b_2) = \frac{1}{3}\frac{1}{2} = \frac{1}{6}. \tag{3.200}$$

(d) The probabilities $P(b_2, a_3)$ and $P(a_3, b_2)$, as shown in (3.196) and (3.200), are different. This is expected, since A and B do not commute. The result of the successive measurements of A and B therefore depends on the order in which they are carried out. The probability of obtaining 0 for B then 1 for A is equal to $\frac{1}{3}$. On the other hand, the probability of obtaining 1 for A then 0 for B is equal to $\frac{1}{6}$. However, if the observables A and B commute, the result of the measurements will not depend on the order in which they are carried out (this idea is illustrated in the following solved problem).

(e) As stated in the text, any operator with non-degenerate eigenvalues constitutes, all by itself, a CSCO. Hence each of $\{\hat{A}\}$ and $\{\hat{B}\}$ forms a CSCO, since their eigenvalues are not degenerate. However, the set $\{\hat{A}, \hat{B}\}$ does not form a CSCO since the opertators $\{\hat{A}\}$ and $\{\hat{B}\}$ do not commute.

Problem 3.9
Consider a system whose state and two observables A and B are given by

$$|\psi(t)\rangle = \frac{1}{6}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}, \qquad A = \frac{1}{\sqrt{2}}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}.$$

(a) We perform a measurement where A is measured first and then, immediately afterwards, B is measured. Find the probability of obtaining a value of 0 for A and a value of 1 for B.

(b) Now we measure B first then, immediately afterwards, A. Find the probability of obtaining a value of 1 for B and a value of 0 for A.

(c) Compare the results of (b) and (c). Explain.

(d) Which among the sets of operators $\{\hat{A}\}$, $\{\hat{B}\}$, and $\{\hat{A}, \hat{B}\}$ form a complete set of commuting operators (CSCO)?

Solution

(a) A measurement of A yields any of the eigenvalues of A which are given by $a_1 = 0$ (not degenerate) and $a_2 = a_3 = 2$ (doubly degenerate); the respective (normalized) eigenstates are

$$|a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}, \qquad |a_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3.201}$$

The probability that a measurement of A yields $a_1 = 0$ is given by

$$P(a_1) = \frac{|\langle a_1|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} = \frac{36}{17}\left|\frac{1}{\sqrt{2}}\frac{1}{6}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}\right|^2 = \frac{8}{17}, \tag{3.202}$$

where we have used the fact that $\langle\psi(t)|\psi(t)\rangle = \frac{1}{36}\begin{pmatrix} 1 & 0 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} = \frac{17}{36}$.

Since the system was initially in the state $|\psi(t)\rangle$, after a measurement of A yields $a_1 = 0$, the system is left, as mentioned in Postulate 3, in the following state:

$$|\phi\rangle = |a_1\rangle\langle a_1|\psi(t)\rangle = \frac{1}{2}\frac{1}{6}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}. \tag{3.203}$$

As for the measurement of B, we obtain any of the eigenvalues $b_1 = -1$, $b_2 = b_3 = 1$; their corresponding eigenvectors are

$$|b_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}, \qquad |b_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}, \qquad |b_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3.204}$$

Since the system is now in the state $|\phi\rangle$, the probability of obtaining the (doubly degenerate) value $b_2 = b_3 = 1$ for B is

$$\begin{aligned} P(b_2) &= \frac{|\langle b_2|\phi\rangle|^2}{\langle\phi|\phi\rangle} + \frac{|\langle b_3|\phi\rangle|^2}{\langle\phi|\phi\rangle} \\ &= \frac{1}{2}\left|\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}\right|^2 + \frac{1}{2}\left|\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}\right|^2 \\ &= 1. \end{aligned} \tag{3.205}$$

The reason $P(b_2) = 1$ is because the new state $|\phi\rangle$ is an eigenstate of B; in fact $|\phi\rangle = \sqrt{2}/3|b_2\rangle$.

In sum, when measuring A then B, the probability of finding a value of 0 for A and 1 for B is given by the product of the probabilities (3.202) and (3.205):

$$P(a_1, b_2) = P(a_1)P(b_2) = \frac{8}{17}. \tag{3.206}$$

(b) Next we measure B first then A. Since the system is in the state $|\psi(t)\rangle$ and since the value $b_2 = b_3 = 1$ is doubly degenerate, the probability of measuring 1 for B is given by

$$\begin{aligned} P'(b_2) &= \frac{|\langle b_2|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} + \frac{|\langle b_3|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle} \\ &= \frac{36}{17}\frac{1}{36}\left[\left|\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}\right|^2 + \left|\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}\right|^2\right] \\ &= \frac{9}{17}. \end{aligned} \tag{3.207}$$

We now proceed to the measurement of A. The state of the system immediately after measuring B (with a value $b_2 = b_3 = 1$) is given by a projection of $|\psi(t)\rangle$ onto $|b_2\rangle$, and $|b_3\rangle$

$$\begin{aligned} |\chi\rangle &= |b_2\rangle\langle b_2|\psi(t)\rangle + |b_3\rangle\langle b_3|\psi(t)\rangle \\ &= \frac{1}{12}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} \\ &= \frac{1}{6}\begin{pmatrix} 1 \\ -2i \\ 2i \end{pmatrix}. \end{aligned} \tag{3.208}$$

So the probability of finding a value of $a_1 = 0$ when measuring A is given by

$$P'(a_1) = \frac{|\langle a_1|\chi\rangle|^2}{\langle\chi|\chi\rangle} = \frac{36}{9}\left|\frac{1}{6\sqrt{2}}\begin{pmatrix} 0 & i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -2i \\ 2i \end{pmatrix}\right|^2 = \frac{8}{9}, \tag{3.209}$$

since $\langle\chi|\chi\rangle = \frac{9}{36}$.

Therefore, when measuring B then A, the probability of finding a value of 1 for B and 0 for A is given by the product of the probabilities (3.207) and (3.209):

$$P(b_2, a_3) = P'(b_2)P'(a_1) = \frac{9}{17}\frac{8}{9} = \frac{8}{17}. \tag{3.210}$$

(c) The probabilities $P(a_1, b_2)$ and $P(b_2, a_1)$, as shown in (3.206) and (3.210), are equal. This is expected since A and B do commute. The result of the successive measurements of A and B does not depend on the order in which they are carried out.

(d) Neither $\{\hat{A}\}$ nor $\{\hat{B}\}$ forms a CSCO since their eigenvalues are degenerate. The set $\{\hat{A}, \hat{B}\}$, however, does form a CSCO since the opertators $\{\hat{A}\}$ and $\{\hat{B}\}$ commute. The set of eigenstates that are common to $\{\hat{A}, \hat{B}\}$ are given by

$$|a_2,\ b_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}, \qquad |a_1,\ b_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix}, \qquad |a_3,\ b_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3.211}$$

Problem 3.10

Consider a physical system which has a number of observables that are represented by the following matrices:

$$A = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix},\ B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 3 \\ 0 & 3 & 0 \end{pmatrix},\ C = \begin{pmatrix} 0 & 3 & 0 \\ 3 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix},\ D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}.$$

(a) Find the results of the measurements of these observables.

(b) Which among these observables are compatible? Give a basis of eigenvectors common to these observables.

(c) Which among the sets of operators $\{\hat{A}\}$, $\{\hat{B}\}$, $\{\hat{C}\}$, $\{\hat{D}\}$ and their various combinations, such as $\{\hat{A}, \hat{B}\}$, $\{\hat{A}, \hat{C}\}$, $\{\hat{B}, \hat{C}\}$, $\{\hat{A}, \hat{D}\}$, $\{\hat{A}, \hat{B}, \hat{C}\}$, form a complete set of commuting operators (CSCO)?

Solution

(a) The measurements of A, B, C and D yield $a_1 = -1$, $a_2 = 3$, $a_3 = 5$, $b_1 = -3$, $b_2 = 1$, $b_3 = 3$, $c_1 = -1/\sqrt{2}$, $c_2 = 0$, $c_3 = 1/\sqrt{2}$, $d_1 = -1$, $d_2 = d_3 = 1$; the respective eigenvectors of A, B, C and D are

$$|a_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \qquad |a_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad |a_3\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{3.212}$$

$$|b_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \qquad |b_2\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |b_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \tag{3.213}$$

$$|c_1\rangle = \frac{1}{\sqrt{26}}\begin{pmatrix} 3 \\ -\sqrt{13} \\ 2 \end{pmatrix}, \quad |c_2\rangle = \frac{1}{\sqrt{13}}\begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}, \quad |c_3\rangle = \frac{1}{\sqrt{26}}\begin{pmatrix} 3 \\ \sqrt{13} \\ 2 \end{pmatrix}, \tag{3.214}$$

$$|d_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}, \qquad |d_2\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |d_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix}. \tag{3.215}$$

(b) We can verify that, among the observables A, B, C, and D, only A and B are compatible, since the matrices A and B commute; the rest do not commute with one another (neither A nor B commutes with C or D; C and D do not commute).

From (3.212) and (3.213) we see that the three states $|a_1, b_1\rangle$, $|a_2, b_3\rangle$, $|a_3, b_2\rangle$,

$$|a_1, b_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \qquad |a_2, b_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad |a_3, b_2\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{3.216}$$

form a common, complete basis for A and B, since $\hat{A}|a_n, b_m\rangle = a_n|a_n, b_m\rangle$ and $\hat{B}|a_n, b_m\rangle = b_m|a_n, b_m\rangle$.

(c) First, since the eigenvalues of the operators $\{\hat{A}\}$, $\{\hat{B}\}$, and $\{\hat{C}\}$ are all nondegenerate, each one of $\{\hat{A}\}$, $\{\hat{B}\}$, and $\{\hat{C}\}$ forms separately a CSCO. Additionally, since two eigenvalues of $\{\hat{D}\}$ are degenerate ($d_2 = d_3 = 1$), the operator $\{\hat{D}\}$ does not form a CSCO.

Now, among the various combinations $\{\hat{A}, \hat{B}\}$, $\{\hat{A}, \hat{C}\}$, $\{\hat{B}, \hat{C}\}$, $\{\hat{A}, \hat{D}\}$, and $\{\hat{A}, \hat{B}, \hat{C}\}$, only $\{\hat{A}, \hat{B}\}$ forms a CSCO, because $\{\hat{A}\}$ and $\{\hat{B}\}$ are the only operators that commute; the set of their joint eigenvectors are given by $|a_1, b_1\rangle$, $|a_2, b_3\rangle$, $|a_3, b_2\rangle$.

Problem 3.11

Consider a system whose initial state $|\psi(0)\rangle$ and Hamiltonian are given by

$$|\psi(0)\rangle = \frac{1}{5}\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}, \qquad H = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 5 & 0 \end{pmatrix}.$$

(a) If a measurement of the energy is carried out, what values would we obtain and with what probabilities?

(b) Find the state of the system at a later time t; you may need to expand $|\psi(0)\rangle$ in terms of the eigenvectors of H.

(c) Find the total energy of the system at time $t = 0$ and any later time t; are these values different?

(d) Does $\{\hat{H}\}$ form a complete set of commuting operators?

Solution

(a) A measurement of the energy yields the values $E_1 = -5$, $E_2 = 3$, $E_3 = 5$; the respective (orthonormal) eigenvectors of these values are

$$|\phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \qquad |\phi_2\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |\phi_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \tag{3.217}$$

The probabilities of finding the values $E_1 = -5$, $E_2 = 3$, $E_3 = 5$ are given by

$$P(E_1) \;=\; |\langle \phi_1|\psi(0)\rangle|^2 = \left| \frac{1}{5\sqrt{2}} \begin{pmatrix} 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} \right|^2 = \frac{8}{25}, \tag{3.218}$$

$$P(E_2) \;=\; |\langle \phi_2|\psi(0)\rangle|^2 = \left| \frac{1}{5} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} \right|^2 = \frac{9}{25}, \tag{3.219}$$

$$P(E_3) \;=\; |\langle \phi_3|\psi(0)\rangle|^2 = \left| \frac{1}{5\sqrt{2}} \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} \right|^2 = \frac{8}{25}. \tag{3.220}$$

(b) To find $|\psi(t)\rangle$ we need to expand $|\psi(0)\rangle$ in terms of the eigenvectors (3.217):

$$|\psi(0)\rangle = \frac{1}{5}\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} = \frac{2\sqrt{2}}{5}|\phi_1\rangle + \frac{3}{5}|\phi_2\rangle + \frac{2\sqrt{2}}{5}|\phi_3\rangle; \tag{3.221}$$

hence

$$|\psi(t)\rangle = \frac{2\sqrt{2}}{5}e^{-iE_1t}|\phi_1\rangle + \frac{3}{5}e^{-iE_2t}|\phi_2\rangle + \frac{2\sqrt{2}}{5}e^{-iE_3t}|\phi_3\rangle = \frac{1}{5}\begin{pmatrix} 3e^{-3it} \\ -4i\sin 5t \\ 4\cos 5t \end{pmatrix}. \tag{3.222}$$

(c) We can calculate the energy at time $t = 0$ in three quite different ways. The first method uses the bra-ket notation. Since $\langle \psi(0)|\psi(0)\rangle = 1$, $\langle \phi_n)|\phi_m\rangle = \delta_{nm}$ and since $\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$, we have

$$\begin{aligned} E(0) = \langle \psi(0)|\hat{H}|\psi(0)\rangle &= \frac{8}{25}\langle \phi_1|\hat{H}|\phi_1\rangle + \frac{9}{25}\langle \phi_2|\hat{H}|\phi_2\rangle + \frac{8}{25}\langle \phi_3|\hat{H}|\phi_3\rangle \\ &= \frac{8}{25}(-5) + \frac{9}{25}(3) + \frac{8}{25}(5) = \frac{27}{25}. \end{aligned} \tag{3.223}$$

The second method uses matrix algebra:

$$E(0) = \langle \psi(0)|\hat{H}|\psi(0)\rangle = \frac{1}{25}\begin{pmatrix} 3 & 0 & 4 \end{pmatrix}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 5 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} = \frac{27}{25}. \tag{3.224}$$

The third method uses the probabilities:

$$E(0) = \sum_{n=1}^{2} P(E_n)E_n = \frac{8}{25}(-5) + \frac{9}{25}(3) + \frac{8}{25}(5) = \frac{27}{25}. \tag{3.225}$$

The energy at a time t is

$$\begin{aligned} E(t) = \langle \psi(t)|\hat{H}|\psi(t)\rangle &= \frac{8}{25}e^{iE_1t}e^{-iE_1t}\langle \phi_1|\hat{H}|\phi_1\rangle + \frac{9}{25}e^{iE_2t}e^{-iE_2t}\langle \phi_2|\hat{H}|\phi_2\rangle \\ &+ \frac{8}{25}e^{iE_3t}e^{-iE_3t}\langle \phi_3|\hat{H}|\phi_3\rangle = \frac{8}{25}(-5) + \frac{9}{25}(3) + \frac{8}{25}(5) = \frac{27}{25} = E(0). \end{aligned} \tag{3.226}$$

As expected, $E(t) = E(0)$ since $d\langle \hat{H} \rangle/dt = 0$.

(d) Since none of the eigenvalues of $\hat{H}$ is degenerate, the eigenvectors $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$ form a compete (orthonormal) basis. Thus $\{\hat{H}\}$ forms a complete set of commuting operators.

Problem 3.12

(a) Calculate the Poisson bracket between the x and y components of the classical orbital angular momentum.

(b) Calculate the commutator between the x and y components of the orbital angular momentum operator.

(c) Compare the results obtained in (a) and (b).

Solution

(a) Using the definition (3.113) we can write the Poisson bracket $\{l_x, l_y\}$ as

$$\{l_x, l_y\} = \sum_{j=1}^{3} \left(\frac{\partial l_x}{\partial q_j} \frac{\partial l_y}{\partial p_j} - \frac{\partial l_x}{\partial p_j} \frac{\partial l_y}{\partial q_j} \right), \tag{3.227}$$

where $q_1 = x$, $q_2 = y$, $q_3 = z$, $p_1 = p_x$, $p_2 = p_y$, and $p_3 = p_z$. Since $l_x = yp_z - zp_y$, $l_y = zp_x - xp_z$, $l_z = xp_y - yp_x$, the only partial derivatives that survive are $\partial l_x/\partial z = -p_y$, $\partial l_y/\partial p_z = -x$, $\partial l_x/\partial p_z = y$, and $\partial l_y/\partial z = p_x$. Thus, we have

$$\{l_x, l_y\} = \frac{\partial l_x}{\partial z} \frac{\partial l_y}{\partial p_z} - \frac{\partial l_x}{\partial p_z} \frac{\partial l_y}{\partial z} = xp_y - yp_x = l_z. \tag{3.228}$$

(b) The components of $\hat{\vec{L}}$ are listed in (3.26) to (3.28): $\hat{L}_x = \hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y$, $\hat{L}_y = \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z$, and $\hat{L}_Z = \hat{X}\hat{P}_y - \hat{Y}\hat{P}_x$. Since $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ mutually commute and so do $\hat{P}_x$, $\hat{P}_y$, and $\hat{P}_z$, we have

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= [\hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y,\ \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z] \\ &= [\hat{Y}\hat{P}_z, \hat{Z}\hat{P}_x] - [\hat{Y}\hat{P}_z, \hat{X}\hat{P}_z] - [\hat{Z}\hat{P}_y, \hat{Z}\hat{P}_x] + [\hat{Z}\hat{P}_y, \hat{X}\hat{P}_z] \\ &= \hat{Y}[\hat{P}_z, \hat{Z}]\hat{P}_x + \hat{X}[\hat{Z}, \hat{P}_z]\hat{P}_y = i\hbar(\hat{X}\hat{P}_y - \hat{Y}\hat{P}_x) \\ &= i\hbar\hat{L}_z. \end{aligned} \tag{3.229}$$

(c) A comparison of (3.228) and (3.229) shows that

$$\{l_x, l_y\} = l_z \longrightarrow [\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z. \tag{3.230}$$

Problem 3.13

Consider a charged oscillator, of positive charge q and mass m, which is subject to an oscillating electric field $E_0 \cos \omega t$; the particle's Hamiltonian is $\hat{H} = P^2/(2m) + k\hat{X}^2/2 + qE_0\hat{X} \cos \omega t$.

(a) Calculate $d\langle \hat{X} \rangle/dt$, $d\langle \hat{P} \rangle/dt$, $d\langle \hat{H} \rangle/dt$.

(b) Solve the equation for $d\langle \hat{X} \rangle/dt$ and obtain $\langle \hat{X} \rangle(t)$ such that $\langle \hat{X} \rangle(0) = x_0$.

Solution

(a) Since the position operator $\hat{X}$ does not depend explicitly on time (i.e., $\partial \hat{X}/\partial t = 0$), equation (3.88) yields

$$\frac{d}{dt}\langle \hat{X}\rangle = \frac{1}{i\hbar}\langle[\hat{X}, \hat{H}]\rangle = \frac{1}{i\hbar}\langle\left[\hat{X}, \frac{P^2}{2m}\right]\rangle = \frac{\langle \hat{P}\rangle}{m}. \tag{3.231}$$

Now, since $[\hat{P}, \hat{X}] = -i\hbar$, $[\hat{P}, \hat{X}^2] = -2i\hbar\hat{X}$ and $\partial \hat{P}/\partial t = 0$, we have

$$\frac{d}{dt}\langle \hat{P}\rangle = \frac{1}{i\hbar}\langle[\hat{P}, \hat{H}]\rangle = \frac{1}{i\hbar}\langle\left[\hat{P}, \frac{1}{2}k\hat{X}^2 + qE_0\hat{X}\cos\omega t\right]\rangle = -k\langle \hat{X}\rangle - qE_0\cos\omega t, \tag{3.232}$$

$$\frac{d}{dt}\langle \hat{H}\rangle = \frac{1}{i\hbar}\langle[\hat{H}, \hat{H}]\rangle + \langle\frac{\partial \hat{H}}{\partial t}\rangle = \langle\frac{\partial \hat{H}}{\partial t}\rangle = -qE_0\omega\langle \hat{X}\rangle \sin\omega t. \tag{3.233}$$

(b) To find $\langle \hat{X}\rangle$ we need to take a time derivative of (3.231) and then make use of (3.232):

$$\frac{d^2}{dt^2}\langle \hat{X}\rangle = \frac{1}{m}\frac{d}{dt}\langle \hat{P}\rangle = -\frac{k}{m}\langle \hat{X}\rangle - \frac{qE_0}{m}\cos\omega t. \tag{3.234}$$

The solution of this equation is

$$\langle \hat{X}\rangle(t) = \langle \hat{X}\rangle(0)\cos\left(\sqrt{\frac{k}{m}}t\right) - \frac{qE_0}{m\omega}\sin\omega t + A, \tag{3.235}$$

where A is a constant which can be determined from the initial conditions; since $\langle \hat{X}\rangle(0) = x_0$ we have $A = 0$, and hence

$$\langle \hat{X}\rangle(t) = x_0\cos\left(\sqrt{\frac{k}{m}}t\right) - \frac{qE_0}{m\omega}\sin\omega t. \tag{3.236}$$

Problem 3.14

Consider a one-dimensional free particle of mass m whose position and momentum at time $t = 0$ are given by x_0 and p_0, respectively.

(a) Calculate $\langle \hat{P}\rangle(t)$ and show that $\langle \hat{X}\rangle(t) = p_0 t^2/m + x_0$.

(b) Show that $d\langle \hat{X}^2\rangle/dt = 2\langle \hat{P}\hat{X}\rangle/m + i\hbar/m$ and $d\langle \hat{P}^2\rangle/dt = 0$.

(c) Show that the position and momentum fluctuations are related by $d^2(\Delta x)^2/dt^2 = 2(\Delta p)^2/m^2$ and that the solution to this equation is given by $(\Delta x)^2 = (\Delta p)_0^2 t^2/m^2 + (\Delta x)_0^2$ where $(\Delta x)_0$ and $(\Delta p)_0$ are the initial fluctuations.

Solution

(a) From the Ehrenfest equations $d\langle \hat{P}\rangle/dt = \langle[\hat{P}, \hat{V}(x,t)]\rangle/i\hbar$ as shown in (3.134), and since for a free particle $\hat{V}(x,t) = 0$, we see that $d\langle \hat{P}\rangle/dt = 0$. As expected this leads to $\langle \hat{P}\rangle(t) = p_0$, since the linear momentum of a free particle is conserved. Inserting $\langle \hat{P}\rangle = p_0$ into Ehrenfest's other equation $d\langle \hat{X}\rangle/dt = \langle \hat{P}\rangle/m$ (see (3.132)), we obtain

$$\frac{d\langle \hat{X}\rangle}{dt} = \frac{1}{m}p_0. \tag{3.237}$$

The solution of this equation with the initial condition $\langle \hat{X} \rangle(0) = x_0$ is

$$\langle \hat{X} \rangle(t) = \frac{p_0}{m} t + x_0. \tag{3.238}$$

(b) First, the proof of $d\langle \hat{P}^2 \rangle / dt = 0$ is straightforward. Since $[\hat{P}^2, \hat{H}] = [\hat{P}^2, \hat{P}^2/2m] = 0$ and $\partial \hat{P}^2/\partial t = 0$ (the momentum operator does not depend on time), (3.124) yields

$$\frac{d}{dt}\langle \hat{P}^2 \rangle = \frac{1}{i\hbar}\langle [\hat{P}^2, \hat{H}] \rangle + \langle \frac{\partial \hat{P}^2}{\partial t} \rangle = 0. \tag{3.239}$$

For $d\langle \hat{X}^2 \rangle / dt$ we have

$$\frac{d}{dt}\langle \hat{X}^2 \rangle = \frac{1}{i\hbar}\langle [\hat{X}^2, \hat{H}] \rangle = \frac{1}{2im\hbar}\langle [\hat{X}^2, \hat{P}^2] \rangle, \tag{3.240}$$

since $\partial \hat{X}^2/\partial t = 0$. Using $[\hat{X}, \hat{P}] = i\hbar$, we obtain

$$\begin{aligned} [\hat{X}^2, \hat{P}^2] &= \hat{P}[\hat{X}^2, \hat{P}] + [\hat{X}^2, \hat{P}]\hat{P} \\ &= \hat{P}\hat{X}[\hat{X}, \hat{P}] + \hat{P}[\hat{X}, \hat{P}]\hat{X} + \hat{X}[\hat{X}, \hat{P}]\hat{P} + [\hat{X}, \hat{P}]\hat{X}\hat{P} \\ &= 2i\hbar(\hat{P}\hat{X} + \hat{X}\hat{P}) = 2i\hbar(2\hat{P}\hat{X} + i\hbar); \end{aligned} \tag{3.241}$$

hence

$$\frac{d}{dt}\langle \hat{X}^2 \rangle = \frac{2}{m}\langle \hat{P}\hat{X} \rangle + \frac{i\hbar}{m}. \tag{3.242}$$

(c) As the position fluctuation is given by $(\Delta x)^2 = \langle \hat{X}^2 \rangle - \langle \hat{X} \rangle^2$, we have

$$\frac{d(\Delta x)^2}{dt} = \frac{d\langle \hat{X}^2 \rangle}{dt} - 2\langle \hat{X} \rangle \frac{d\langle \hat{X} \rangle}{dt} = \frac{2}{m}\langle \hat{P}\hat{X} \rangle + \frac{i\hbar}{m} - \frac{2}{m}\langle \hat{X} \rangle \langle \hat{P} \rangle. \tag{3.243}$$

In deriving this expression we have used (3.242) and $d\langle \hat{X} \rangle / dt = \langle \hat{P} \rangle / m$. Now, since $d(\langle \hat{X} \rangle \langle \hat{P} \rangle)/dt = \langle \hat{P} \rangle d\langle \hat{X} \rangle / dt = \langle \hat{P} \rangle^2 / m$ and

$$\frac{d\langle \hat{P}\hat{X} \rangle}{dt} = \frac{1}{i\hbar}\langle [\hat{P}\hat{X},\ \hat{H}] \rangle = \frac{1}{2im\hbar}\langle [\hat{P}\hat{X},\ \hat{P}^2] \rangle = \frac{1}{m}\langle \hat{P}^2 \rangle, \tag{3.244}$$

we can write the second time derivative of (3.243) as follows:

$$\frac{d^2(\Delta x)^2}{dt^2} = \frac{2}{m}\left(\frac{d\langle \hat{P}\hat{X} \rangle}{dt} - \frac{d\langle \hat{X} \rangle \langle \hat{P} \rangle}{dt} \right) = \frac{2}{m^2}\left(\langle \hat{P}^2 \rangle - \langle \hat{P} \rangle^2 \right) = \frac{2}{m^2}(\Delta p)_0^2, \tag{3.245}$$

where $(\Delta p)_0^2 = \langle \hat{P}^2 \rangle - \langle \hat{P} \rangle^2 = \langle \hat{P}^2 \rangle_0 - \langle \hat{P} \rangle_0^2$; the momentum of the free particle is a constant of the motion. We can verify that the solution of the differential equation (3.245) is given by

$$(\Delta x)^2 = \frac{1}{m^2}(\Delta p)_0^2 t^2 + (\Delta x)_0^2. \tag{3.246}$$

This fluctuation is similar to the spreading of a Gaussian wave packet we derived in Chapter 1.

UNIT 6 SCHRÖDINGER EQUATION

Strueture

6.1 INTRODUCTION

In Unit 4, you have seen that a microscopic particle is essentially represented by a matter wave with its wavelength given by the de Broglie relation. How do we describe the motion of such a particle or a system of such particles? Clearly, we cannot make use of Newton's laws of motion for this purpose. So a new theoretical description is needed for the motion of quantum mechanical particles. The new theory should be consistent with the wave nature of particles. It should also reduce to Newtonian mechanics for macroscopic particles. Recall that this condition is similar to the special theory of relativity which reduces to Newtonian mechanics at velocities much smaller than the velocity of light.

In this connection, we would like to recount a story. At the end of a seminar, in **1926,** on de Broglie waves, the physicist Peter Debye said to another physicist that if matter is a wave, there should be a wave equation to describe a matter wave. Debye promptly forgot about it but the other physicist, Erwin Schrodinger proceeded to discover the wave equation for matter waves. This equation is named after him as the Schrödinger equation.

In this unit you will study the one-dimensional Schrödinger equation, and learn about its solutions, We shall also discuss the physical meaning of these solutions. These solutions are acceptable only under certain conditions about which you will study towards the end of the unit. An appendix has been given at the end of the Unit to explain the basic complex algebra used in the text. In the next unit we shall introduce another way of describing quantum mechanical systems, given by Heisenberg and Dirac which makes use of operators and observables.

Objectives

After studying this unit you should be able to

- write the onedimensional time dependent Schrödinger equation and derive the time independent Schrödinger equation from it,
- give a statistical interpretation of the wave function,
- derive the continuity equation for the probability current density,
- normalise a given wave function,
- apply the boundary conditions to a given wave function.

6.2 ONE-DIMENSIONAL SCHRÖDINGER EQUATION

You have already learnt that wave nature is an inherent property of every particle. We now need a wave equation which suitably describes the time evolution of matter waves representing the particle. In Unit 5, you have learnt that one way of **localising** a particle is by constructing a wave packet. However, by solving the **terminal** question 4 you have **also** seen that a wave packet spreads with time. This means that, a wave packet cannot represent a **particle.**

Hence, in quantum mechanics it is postulated that

Postulate 1: Description of the system

> Every particle (or a system of particles) is represented by a **"wave function",** which is a function of space coordinates and time. The wave function determines **all** that can be known about the system it represents.

For one-dimensional motion of a particle, the wave function may be represented by $\psi(x, t)$. Now you may **ask:** What is the form of $\psi(x, t)$ in terms of x and t? To answer this question, consider a classical (macroscopic) particle moving under the influence of a force. Its dynamical behaviour is described by Newton's second law, which is a differential equation. Similarly, Maxwell's equations of classical electromagnetism are also differential equations. Since all objects exhibit particle as well as wave nature, it is natural to expect that the **quantum mechanical** wave function will also be a solution of a certain differential equation, **involving** derivatives of x and t.

The credit for discovering such a differential equation goes to Erwin **Schrödinger** (Fig. 6.1). How did he arrive at his equation? The answer is, he created it intuitively breaking all traditions of such wave equations. The **Schrödinger** equation is one of the most successful equations of quantum mechanics because it predicts results which can be verified experimentally. We will now give you some idea of how he visualised the particular form of his equation.

Let us first state certain preconditions for establishing this equation. Firstly, for a particle of mass m, energy E and momentum p, the equation should be consistent with

(i) the de Broglie relation $\lambda = \frac{h}{p}$, and **(6.1)**

(ii) Planck formula $\nu = \frac{E}{h}$ **(6.2)**

It should also satisfy the relation

(iii) $$E = \frac{p^2}{2m} + V(x, t)$$ **(6.3)**

Fig. 6.1 : Erwin **Schrödinger,** 1887-1961, **Austrian** theoretical physicist. Another **of the** founders **of** the new **quantum** mechanics, **he received** the Nobel Prize in 1933.

for all x and t, where $V(x, t)$ is the potential energy of the particle,

(iv) Finally, the SchrGdinger equation must be linear in x and t. That is, if $\psi_1(x, t)$ and $\psi_2(x, t)$ are two **solutions** of the **Schrödinger** equation for a given potential energy $V(x, t)$ then any linear combination of ψ_1 and ψ_2, say, $C_1 \psi_1 + C_2 \psi_2$, with C_1 and C_2 as arbitrary constants, must also be the solution of the same SchrGdinger equation. This linearity is required so that two waves may be added to produce interference. If the linearity property is to be satisfied, the SchrGdinger (differential) equation must be of degree 1, **i.e.,** the wave functions and its derivatives appearing in it should be only of the first power.

We now put Eqs. (6.1) and (6.2) in (6.3) to obtain

$$\hbar\omega = \frac{\hbar^2 k^2}{2m} + V(x, t) \qquad (6.4)$$

where $\omega = 2\pi\nu$ and $k = 2\pi/\lambda$. We now consider a simpler situation where the potential energy is constant, say equal to V_0. Under such a situation, if we take the particle to be a photon then it will have a fixed wavelength and frequency (as given by Eqs. (6.1) and (6.2)) and its wavefunction as given by the electromagnetic theory (see Unit 14, PHE-07 entitled Electric and Magnetic Phenomena), will be

$$\psi(x, t) = A\, e^{i(kx - \omega t)} \tag{6.5}$$

Differentiating the above equation once with respect to time you can easily obtain

$$\omega = i \frac{1}{\psi(x, t)} \frac{\partial \psi(x, t)}{\partial t} \tag{6.6}$$

and

$$k^2 = -\frac{1}{\psi(x, t)} \frac{\partial^2 \psi(x, t)}{\partial x^2} \tag{6.7}$$

Putting Eqs. (6.6) and (6.7) in Eq. (6.4) we obtain a differential equation connecting the wave function $\psi(x, t)$ and its derivatives:

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V_0\, \psi(x, t) \tag{6.8}$$

The above equation has been obtained for a special case of constant potential energy V_0. However, Schrödinger made a bold extrapolation and postulated that the form of the Eq. (6.8) does not change even for a particle of mass m moving in a potential which varies with x and t. Thus,

The time-dependent Schrödinger equation for one-dimensional motion of a particle of mass m moving in a potential $V(x, t)$ is given by

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x, t)\, \psi(x, t) \tag{6.9}$$

Postulate 2: Time Evolution of a System

You should note that Eq. (6.9) is consistent with (i) to (iv). The appearance of $\hbar$ in the Schrödinger equation is, of course, crucial. This is how Schrodinger imposed the "quantum condition" on the wave equation of matter.

While reading this discussion, did you wonder what would happen if we had a wave equation with a second order time derivative? We would then have ended up with the relativistic energy-momentum relationship. Actually, to begin with, this is what Schrödinger himself attempted. But very soon he realised that such an equation did not work for electrons — it did not give the correct spectrum for the hydrogen atom. Incidentally, the correct relativistic equation for electrons, discovered by Dirac, does retain the first-order time derivative.

The **Schrödinger** equation is unlike any of the wave equations you have come across so far. Recall that wave equations usually connect a second-order time derivative of the function with its second order spatial derivative. But **Schrödinger's** equation contains only the first-derivative with respect to **time** but the second derivative with respect to space. Hence, time and space coordinates are not treated on an equal footing in this equation. Thus Eq. (6.9) cannot be correct in the relativistic domain. Hence, it is a **non-relativistic** time dependent **Schrödinger** equation.

And, there is a price to pay for having only a first order time derivative in the wave **equation.** The **solutions** of the **Schrödinger** equation are not real physical waves; they we complex functions **with** both a real and an imaginary part. This gives rise to the problem of interpretation of the wave function. What exactly does the wavefunction $\psi(x, t)$ mean physically? We shall discuss the interpretation of ψ given by **Max** Born in the next section.

But before that you might like to verify the linearity property of $\psi(x, t)$. Try the following SAQ.

SAQ 1

If $\psi_1(x, t)$ and $\psi_2(x, t)$ we **two** solutions of the Schrodinger equation **(6.9)**, show that $a\psi_1$ and $a\psi_1 + b\psi_2$ are also solutions of Eq. **(6.9)**, where a and b are arbitrary constants.

6.3 STATISTICAL INTERPRETATION OF THE WAVE FUNCTION

The coefficient of the time derivative of ψ in **Eq.** (6.9) is imaginary. Therefore, it is evident that the wave function ψ, which is the solution of **(6.9)**, will, in general, be complex. Thus, in order to extract any physical information from $\psi(x, t)$, we must establish a quantitative connection between $\psi(x, t)$ and the observables of the **particle.** In 1926, Max Born proposed the following connection:

Postulate 3: Probabilistic interpretation of the wave function

> If, at any instant t, a measurement is made to locate the **object** represented by the wave function $\psi(x, t)$, then the probability $P(x, t)\,dx$ that the object will be found **between** the coordinates x and $x + dx$ is
>
> $$P(x, t)\,dx = \psi^*(x, t)\,\psi(x, t)\,dx = |\psi(x, t)|^2 dx, \tag{6.10}$$
>
> **where** "*" on a **function** represents its complex conjugate.

You can see that $|\psi(x, t)|^2$ **is the** modulus square of the wave function. Here, $P(x, t) = \psi^*(x, t)\,\psi(x, t) = |\psi(x, t)|^2$ is also termed the **probability density.** To put it in words:

The probability of finding a quantum mechanical object in a small interval dx is given ***by*** the product of the modulus square of the wave function representing the object and the interval *itself.*

The probability of finding the particle within some finite length $L = (x_2 - x_1)$ is given **by**

$$P_L(t) = \int_{x_1}^{x_2} P(x, t)\, dx \tag{6.11}$$

Thus, according to Max Born, the **Schrödinger** equation gives probability waves. The wave function just tells us probabilistically where the likelihood of finding the particle will be greater: there the wave will be strong, its amplitude will be larger. If the probability of finding a particle in a region is small, the wave will be weak and its amplitude will be small. It may seem from the above probabilistic interpretation that the phase of the wave function is not important, since it is the modulosquare $|\psi(x, t)|^2$ that we interpret as the probability. However, this is not so. We shall very briefly discuss this aspect.

> To **visualise** this concept, **imagine** you are in a **metropolis** like Delhi and looking for traffic **jams** on Its **mads** from a **helicopter.** If tho vehicles were **described** by **Schrödinger waves, we** would say that the wave **was strong** at **the** location of traffic jam. **Elsewhere,** the wave would be weak.

The phase of the wave function

From the linearity property of **Schrödinger** equation and SAQ 1, you know that if ψ_1 and ψ_2 are solutions of the **Schrödinger** equation, the linear combination

$$\psi(x, t) = a_1\, \psi_1(x, t) + a_2\, \psi_2(x, t) \tag{6.12}$$

is also a solution of the **Schrödinger** equation where a_1 and a_2 are arbitrary complex numbers. This is the **superposition principle.** Now, calculate $|\psi(x, t)|^2$ using Eq. (6.5) with a phase difference, say ϕ, between ψ_1 and ψ_2. You will see that it depends on the relative phase of ψ_1 and ψ_2. Such a superposition is called **coherent superposition** — it gives rise to the interference of matter waves. Thus, the phase of the wave function is important and cannot be ignored.

You should also notice a crucial difference between the use of probability in classical physics and in quantum physics. In classical physics, probabilities add as you have studied in Unit 5 of the physics elective PHE-04 (Mathematical Methods in Physics-I), But in quantum physics, the probability amplitudes add, as in Eq. (6.12) and then we calculate the probabilities from **Eq.** (6.10) giving rise to interference.

Let us come back to the probabilistic interpretation. Since the object must always be somewhere in space, the **total probability of finding it in the whole space is unity.** We obtain this by integrating the probability over all space:

$$\boxed{\int_{-\infty}^{\infty} \psi^*(x, t)\, \psi(x, t)\, dx = 1, \text{ at each instant of time } t} \tag{6.13}$$

The interpretation-of the wave function given by Eqs. (6.10) to (6.13) requires that ψ should be **finite** and **single-valued** everywhere, otherwise the probability of finding an object in a **region** of space will not be finite and unique. Further, Eq. (6.13) requires that we restrict the wave functions used in quantum mechanics to the class of **square integrable functions for** which

$$|\psi(x, t)|^2 \, dx < \infty$$

i.e., $\psi(x, t)$ must approach 0 as $x \to \pm\infty$ at least as fast as $x^{-1/2-\epsilon}$, with $\epsilon > 0$ and arbitrarily small.. In addition, certain continuity conditions need to be imposed on ψ. We will discuss these in Sec. 6.4.1 in detail. Meanwhile, let us further explore the meaning of ψ.

6.3.1 Probability Current Density and the Continuity Equation

Since Eq. (6.13) is true for every *t, the total probability is conserved.* But this can be accomplished only when *the probability,is conserved at each point, and at all times.* Let us examine this aspect in some detail.

Let us consider the concrete example of a fluid moving in the positive x-direction between two points $x = x_1$ and $x = x_2$ with a velocity v which changes with x. Let $\rho(x)$ denote the mass per unit length of the fluid around the point x. The quantity $S_x = v(x)\,\rho(x)$ is then the mass of the fluid crossing a given point x per unit time. What is the net mass accumulated per unit time in between the two points $x = x_1$ and $x = x_2$? Clearly, it is equal to $S_{x=x_1} - S_{x=x_2}$. And if the fluid is to be conserved in this region then this should be equal to the rate of change of mass in this region. Thus

$$\frac{\partial}{\partial t}\int_{x_1}^{x_2} \rho(x)\,dx = S_{x=x_1} - S_{x=x_2} \tag{6.14}$$

We can extend this analogy to the probabilistic interpretation. We say that if *the total probability is to be conserved, the conservation equation should look like* Eq. (6.14) *where* ρ *should be replaced by the probability density* $P(x, t)$ *and* S_x *by a function which we term the-probability flux or the probability current density* $S(x, t)$. Thus, we must have

We can extend this analogy to the probabilistic interpretation. We say that if *the total probability is to be conserved, the conservation equation should look like* Eq. (6.14) *where* ρ *should be replaced by the probability density* $P(x, t)$ *and* S_x *by a function which we term the-probability flux or the probability current density* $S(x, t)$. Thus, we must have

$$\frac{\partial}{\partial t}\int_{x_1}^{x_2} P(x, t)\,dx = S(x_1, t) - S(x_2, t) \tag{6.15}$$

Let us now obtain the definition of the probability flux $S(x, t)$ so that the probability conservation equation (6.15) is obeyed. For this we shall use the **Schrödinger** equation:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi \tag{6.16a}$$

Here and in future we will omit the arguments of ψ and V as long as it does not create any confusion. The complex conjugate of Eq. (6.16a) is

$$-i\hbar \frac{\partial \psi^*}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi^*}{\partial x^2} + V\psi^* \qquad (6.16b)$$

where we have assumed that V is real so that $V^* = V$. You can now multiply Eq. (6.16a) from the left by ψ^* and Eq. (6.16b) by ψ and then subtract (6.16b) from (6.16a) and obtain

$$i\hbar \left(\psi^* \frac{\partial \psi}{\partial t} + \psi \frac{\partial \psi^*}{\partial t}\right) = -\frac{\hbar^2}{2m} \left(\psi^* \frac{\partial^2 \psi}{\partial x^2} - \psi \frac{\partial^2 \psi^*}{\partial x^2}\right)$$

Now carry out a simple algebraic manipulation of the above equation to 'show that

$$\frac{\partial(\psi^*\psi)}{\partial t} = -\frac{\hbar}{2mi} \frac{\partial}{\partial x} \left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x}\right) \qquad (6.17)$$

We now integrate Eq. (6.17) with respect to x from x_1 to x_2 and get

$$\frac{\partial}{\partial t} \int_{x_1}^{x_2} \psi^* \psi \, dx = -\frac{\hbar}{2mi} \left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x}\right) \Bigg|_{x_1}^{x_2} \qquad (6.18)$$

A comparison of Eqs. (6.15) and (6.18) shows that the **probability** density $P(x, t)$ and the *probability* **flux** or **probability** *current* density $S(x, t)$ should **be** defined as

$$\boxed{\begin{aligned} &P(x, t) = \psi^*(x, t)\,\psi(x, t) \\ &\text{and} \\ &S(x, t) = \frac{\hbar}{2mi} \left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x}\right) \end{aligned}} \qquad (6.19)$$

We may rewrite Eq. (6.18) in terms of P and S as

$$\boxed{\frac{\partial P(x, t)}{\partial t} + \frac{\partial S(x, t)}{\partial x} = 0} \qquad (6.20)$$

Doesn't this equation look familiar to you? It has the form of a continuity equation analogous to the continuity equation between the charge density and current density in electrodynamics. This is one reason why $P(x, t)$ is referred to as the probability density and $S(x, t)$ is called the probability current density. In this equation, the first term $\frac{\partial P}{\partial t}$ denotes the rate of change of probability density in a certain fixed length. The second term denotes the net outward flux coming out of the same length. Eq, (6.20) then says that the time rate of change of probability density (which is a negative quantity because as t increases $\partial P/\partial t$ decreases) is numerically equal to the net *outward* flux (an outward flux is always assumed to be positive quantity). This means that within the above length, the particles are neither created nor destroyed (i.e., there are neither sources, nor sinks).

Eq. (6.20) is, therefore, the conservation law expressing the fact that a change in the particle density in a region of space is compensated for by a net change *of flux* from that region. You can now also see why we need to impose continuity conditions on ψ: both ψ and its derivative $\frac{\partial \psi}{\partial x}$ must be finite and continuous for all values of x provided $V(x)$ is finite. That these restrictions on the solutions are necessary may be judged from the following considerations.

The probability density $P(x)$ and the probability flux $S(x, t)$ represent physical quantities and, therefore, have to be well defined. If $\psi(x)$ or its first derivative $\psi'(x)$ were not finite for some values of x then $P(x)$ and/or $S(x, t)$ would not be well defined for all values of x. Further, both $\psi(x)$ and $\psi'(x)$ must be continuous. Otherwise $S(x, t)$ would be singular at some points and these points would act as sources or sinks of probability current. In other words, creation or destruction of matter would take place. This, as you know, is impossible in non-relativistic physics.

Before proceeding further, we would like to point out that Eq. (6.20) is obtained under the condition that V is real. By writing $\psi = \psi_R + i\psi_I$ it is easy to see that both P and S are real and

$$S(x, t) = \frac{\hbar}{m} \, \text{Im} \left(\psi^* \frac{\partial \psi}{\partial x} \right) = \text{Re} \left(\psi^* \frac{\hbar}{im} \frac{\partial \psi}{\partial x} \right) \tag{6.21}$$

where Im(Z) denotes the magnitude of the imaginary part of Z and Re (Z) its real part. You may like to do an exercise to fix the ideas involved.

SAQ 2

(a) Show that for complex potential energy, the continuity equation (6.20) modifies to

$$\frac{\partial P(x, t)}{\partial t} + \frac{\partial S(x, t)}{\partial x} = \frac{2V_I}{\hbar} P(x, t)$$

where V_I is the imaginary part of the potential energy.

(b) The wave function of an object of mass m is given by

$$\psi(x, t) = e^{-(\alpha + i\beta)x} \, e^{-i\omega t}$$

Obtain the values of $P(x, t)$ and $S(x, t)$.

The probabilistic interpretation of ψ also leads us to the concept of normalising a wave function.

6.3.2 Normalisation of Wave Functions

Recall Eq. (6.13) which tells us that since the particle must be somewhere in space, the probability integrated over all x (i.e., all space in one-dimension) must equal 1. A wave function satisfying Eq. (6.13) is said to be **normalised.**

Now in quantum mechanics we deal with two types of wave functions. For one of them the value of the integral $\int_{-\infty}^{\infty} \psi^{*\prime}(x, t)\, \psi'(x, t)\, dx$ is finite, say equal to N, where ψ' is the solution of the Schrbdinger equation. Such *functions are* said to be normalisable and N is known as the **norm** of the wave **function.** It also means that ψ' is a square integrable function which vanishes as $|x| \to \infty$.

You also know that Schrodinger equation is linear. And you have shown in SAQ 1 that if ψ' is a solution of Eq. (6.9) then ψ' multiplied by a constant (independent of t and x) is also a solution. Hence we can always choose the constant to be $N^{-1/2}$ and take $\psi = N^{-1/2}\,\psi'$. Then for ψ, Eq. (6.13) is satisfied and the wave function ψ is said to be a **normalised** wave function. Note that N is independent of time (otherwise ψ will not be a solution of Eq. (6.9)). Hence, a wave function which is ***normalised*** at any instant of time stays normalised at all other times.

However, there exists a category of wave functions **for which** the value of the integral in Eq. (6.13) is infinite. Such functions do not represent a physical system in the strictest sense. But, we shall see later that such functions are extensively used in quantum mechanics to describe free particles. In fact, we have already used $e^{i(kx-\omega t)}$ to represent a free particle. It is a wave function whose norm is infinite or, in other words, it is unnormalisable. Wave functions of this form do not go to zero as $x \to \pm\infty$. We shall discuss the normalisation of such functions in detail in Block 3.

Let us now further analyse the Schrodinger Equation.

6.4 TIME INDEPENDENT SCHRÖDINGER EQUATION

The Schrodinger equation (6.9) represents the time development of the wave functions $\psi(x, t)$. We have seen that the position probability density of the particle is related to $\psi(x, t)$. We can also transform $\psi(x, t)$ in such a way that we obtain momentum probability distributions of the particle. (Such transforms are called Fourier transforms). Thus, given the position and momentum distributions of the particle at one time, the same can be obtained at any **later** time with the help of the **Schrödinger** equation. Isn't this situation analogous to the one in classical mechanics? In classical mechanics, if the position and momentum of a particle are known at some initial time then we can use Newton's equation of motion to find the position and momentum of the particle at any later time. However, there is a difference between classical and quantum mechanics. **What** is it? Recall the uncertainty principle which tells us that in quantum mechanics both the position and the momentum of the particle cannot be known precisely at the same instant of time. We can know only their distributions.

In many pmblems of classical mechanics such as Kepler's planetary orbits, Rutherford scattering, we assume that the particle has definite energy **and/or** angular momentum. However, in quantum mechanics if we assume the energy of the particle to be known precisely, **i.e.**, if AE = 0 then, according to the uncertainty relation ΔE At $\geq \hbar$ and At is infinite. This means that an infinite amount of time should be available to make energy measurements. In other words, the probability density ψ^* (x, t) $\psi(x, t)$ should not change

> **Richard Feynman, whose Feynman Lectures on Physics are an essential reading for every physics student, once said, "Electron waves are probability waves in the ocean of uncertainty;" now you know what that means!**

with time. Hence, for a system of constant energy the wave function $\psi(x, t)$ should take the form

$$\psi(x, t) = \psi(x) \exp\{ig(t)\} \tag{6.22}$$

where $g(t)$ is any function of t. For a system represented by such a wave function, the energy of the system does not change with time, i.e., its energy is conserved. Hence, if the particle is initially in a specified energy state represented by Eq. (6.22) then it stays in it indefinitely unless it is disturbed by some external agency. Such energy states are called *stationary* states.

Classically, the energy (which is also termed the I-Iamiltnnian) is a constant of motion if it does not contain time explicitly. Now, the *Hamiltortian is the sum of the kinetic energy and the potential energy*. Thus, for the total energy to be conserved, the potential in which the particle is moving should be independent of time. For such potentials, the Schrödinger equation (6.9) may be separated in x and t as follows. Substituting

$$\psi(x, t) = \psi(x)\, f(t) \tag{6.23}$$

in Eq. *(6.9)* and rearranging the terms, we get

$$-\frac{\hbar^2}{2m}\frac{1}{\psi(x)}\frac{\partial^2 \psi(x)}{\partial x^2} + V(x) = \frac{i\hbar}{f(t)}\frac{\partial f(t)}{\partial t} \tag{6.24}$$

The left side of this equation is constant for fixed x nt all t. Similarly, the right side is constant for Fixed t at all values of x. Hence, Eq. *(6.24)* will hold only when both the sides are equal to a constant C which is independent of x and t. Thus we obtain

$$i\hbar\frac{df}{dt} = C\, f(t) \tag{6.25}$$

and

$$-\frac{\hbar^2}{2m}\psi''(x) + V(x)\,\psi(x) = C\,\psi(x), \tag{6.26}$$

where $\psi''(x) \equiv d^2\psi(x)/dx^2$.

You can solve Eq. (6.25) to obtain

$$f(t) = A\exp\{-iCt/\hbar\} = A\exp\{-i\omega t\} \tag{6.27}$$

where A is the normalisation constant and $\omega = C/\hbar$. From Eqs. *(6.26)* and *(6.27)* it is clear that C should have dimensions of energy and be equal to the total energy E, Eq. *(6.27)* is, therefore, written as

$$\boxed{-\frac{\hbar^2}{2m}\psi''(x) + V(x)\,\psi(x) = E\,\psi(x)} \tag{6.28}$$

The above equation is known as time independent Schrödinger *equation.*

We can now write the general solution or the stationary state solution of the time-dependent Schrödinger equation for a particle of definite energy E as

$$\psi(x, t) = \psi(x)\exp(-iEt/\hbar), \tag{6.29}$$

where $\psi(x)$ satisfies Eq. *(6.28)*. The probability density and the probability flux in such cases are given by

$$P(x) = \psi^*(x)\,\psi(x) \tag{6.30}$$

and

$$S(x) = \frac{\hbar}{2mi} [\psi^* (x) \frac{d}{dx} \psi(x) - \psi(x) \frac{d}{dx} \psi^* (x)] \tag{6.31}$$

It is evident that both $P(x)$ and $S(x)$ are independent of time. You should note that Eq. (6.28) contains no imaginary quantities and hence $\psi(x)$ is not necessarily complex although $\psi(x, t)$ is. The normalisation condition (6.13) for the stationary state function reduces to a form

$$\int_{-\infty}^{\infty} \psi^* (x) \psi(x) \, dx = 1 \tag{6.32}$$

You may now like to apply the concepts discussed so far.

SAQ 3

The wave function for a steady state is given by

$$\psi(x) = N \exp\left(-\frac{x^2}{2}\right)$$

Calculate the value of the normalisation constant N.

In this course we shall confine our study to those problems which require solutions of time independent Schrtidinger equation, i.e., we shall study stationary state problems. Let us now examine the conditions which the wave function $\psi(x)$ has to satisfy in order to be a physically acceptable solution.

Recall that the probabilistic interpretation imposes the following conditions on the wave function $\psi(x)$:

1) $\psi(x)$ should be finite and single-valued everywhere,

2) $\psi(x)$ should be square integrable and

3) both $\psi(x)$ and $\frac{\partial \psi}{\partial x}$ should be continuous everywhere.

We can rewrite the time independent Schrödinger equation (6.28) as

$$\frac{d^2 \psi(x)}{dx^2} = \frac{2m}{\hbar^2} [V(x) - E] \psi(x) \tag{6.33}$$

You know that $\psi(x)$ represents a probabilistic wave satisfying Eq. (6.13). Further, whether $\psi(x)$ is an acceptable solution of Eq. (6.33) or not is also determined by $V(x)$ as well as by the boundary conditions which depend upon the nature of the problem. Let us consider an example of a particle bound in a potential well shown in Fig. 6.2. Here $V(x) > E$ for $x < x_1$ and $x > x_2$.

Classically, if the particle is initially between x_1 and x_2 then for all times to come it will be confined between the same space, i.e., the particle is bounded between x_1 and x_2. Then we say that the particle is in a bound state. Quantum mechanically we expect a large probability of finding the particle between the space $x_1 < x < x_2$. However, there also exists a decreasingly small probability of finding the particle outside this space, which is forbidden classically. This, in turn, demands for the (boundary) condition that the bound state wave functions must vanish at infinity, A very interesting result follows from the consideration of such a boundary condition which you shall study in the next block in detail. Here we just mention it: The (acceptable) solutions of the time independent Schrödinger equation exist only for certain discrete values of the total

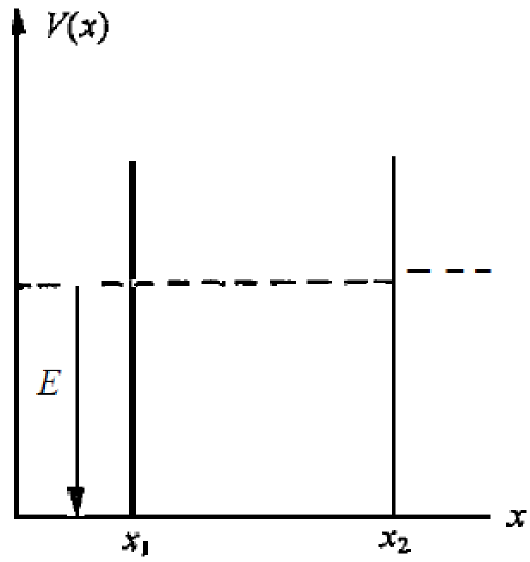

Fig. 6.2 : Particle in a potential well

energy E. Thus the energy quantization is an inherent property of the *Schrödinger* equation for the bound states.

You have just learnt that the probabilistic interpretation puts another **restriction** on an acceptable (or well behaved) solution: the wave function and its first derivative must be finite *and* continuous. Eq. (6.28) shows that if $V(x)$, E and $\psi(x)$ are **finite** then $\psi''(x)$ is also finite. This, in **turn**, means that $\psi'(x)$ is continuous. However, if for certain values of x, $V(x)$ becomes infinite then Eq. (6.28) yields an infinite value for $\psi''(x)$ at those points. Hence at those points $\psi'(x)$ may not be continuous.

We now end this section by summarising the properties and the boundary conditions that a wave function must satisfy to be acceptable:

Properties of the Wave Function

(1) $\psi(x)$ must be **single valued, finite** and **continuous** for all values of x.

(2) $\psi'(x)$ must be **finite** and **continuous** for all values of x, except at those points where $V \to \infty$. At these points $\psi'(x)$ has a finite discontinuity but ψ remains continuous.

(3) For bound states, the probability of finding the particle between x and $x + dx$, i.e., $|\psi|^2 dx$ must vanish as $|x| \to \infty$. Hence $|\psi(x)| \to 0$ as $|x| \to \infty$, i.e., $\psi(x)$ is a **square integrable** wave function.

Let us now summarise what you have studied in this unit.

6.5 SUMMARY

- In this unit we have concentrated on one-dimensional motion of particles. You have learnt **three postulates of quantum mechanics:**

1. Every system can be represented by a **wave function:**

2. The wave function satisfies a differential equation, called the **Schriidinger equation** given by

3. The **probability** $P(x, t)$ of finding a particle at (x, t) in the elementary element dx is given by

$$P(x, t)\, dx = \psi^*(x, t)\, \psi(x, t)\, dx$$

where $P(x, t)$ is the **probability density.**

- **Schrödinger** equation can be used to derive a **continuity equation** which connects the probability density with an associated probability current density $S(x, t)$ as follows:

$$\frac{\partial P(x, t)}{\partial t} + \frac{\partial S(x, t)}{\partial x} = 0$$

where

$$S(x, t) = \frac{\hbar}{2mi}\left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x}\right)$$

The continuity equation tells us that a particle moving under a real potential is neither destroyed nor is another particle created; the change in particle density in a region equals the net change of flux into or away from that region.

- The **total probability** of finding **a** particle in the whole space always remains unity:

$$\int_{-\infty}^{\infty} \psi^*(x)\, \psi(x)\, dx = 1$$

Wave functions which satisfy this condition are **said** to be **normalised.**

- For a conservative system we can write

$$\psi(x, t) = \psi(x)\, e^{-iEt/\hbar}$$

where E is the total energy of the system and $\psi(x)$ is the solution of the time independent Schrodinger equation given by

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V\psi = E\psi$$

- In order that $\psi(x)$ represent a physical system, the wave function $\psi(x)$ must be single valued, finite and continuous at all values of x. Its first derivative $d\psi(x)/dx$ must also be finite and continuous at all values of x, except at those points where $V(x) \to \infty$. At these points the first derivative has a finite discontinuity.

6.6 TERMINAL QUESTIONS

Spend 45 min

1. The wave function of an object of energy E and momentum p is given by

   $$\psi(x, t) = A\, e^{i(px - Et)/\hbar}.$$

   (i) Does ψ represent a bound state?

   (ii) Is the wave function normalisable?

   (iii) Using the above calculate the probability current density $S(x, t)$ in terms of the velocity v of an object and a constant A, which is complex.

2. The unnormalised wave function of a system is given by $x \exp(-x^2/2)$. Obtain the value of its normalisation constant.

3. A certain function is given by

$$\psi(x) = N(1 + ix)\exp(-x) \quad \text{for } x > 1$$
$$= 0 \qquad\qquad \text{for } x < 1$$

Obtain the value of the normalisation constant N. Why can it not represent a physical system?

4. The potential energy of a simple harmonic oscillator of mass m and frequency ν is equal to $2m(\pi\nu x)^2$. Write down its time independent Schrödinger equation and show that it can be rewritten as

$$\frac{d^2\psi}{d\xi^2} + \left(\frac{\alpha}{\beta} - \xi^2\right)\psi = 0$$

where $\alpha = \frac{2m}{\hbar^2}E$, $\beta = 2\pi m\nu/\hbar$ and $\xi = \sqrt{\beta}\,x$, E being the total energy of the oscillator. For what values of α/β are the functions $\psi(\xi) = \exp(-\xi^2/2)$ and $\psi(\xi) = \xi\exp(-\xi^2/2)$ solutions of the above equation?

6.7 SOLUTIONS AND ANSWERS

Self-Assessment Questions

1. Substituting $\psi = a\psi_1 + b\psi_2$ in Eq. (6.9) we obtain.

$$i\hbar\left(a\frac{\partial\psi_1}{\partial t} + b\frac{\partial\psi_2}{\partial t}\right) = -\frac{\hbar^2}{2m}\left(a\frac{\partial^2\psi_1}{\partial x^2} + b\frac{\partial^2\psi_2}{\partial x^2}\right) + Va\psi_1 + Vb\psi_2$$

or

$$a\left(i\hbar\frac{\partial\psi_1}{\partial t} + \frac{\hbar^2}{2m}\frac{\partial^2\psi_1}{\partial x^2} - V\psi_1\right) + b\left(i\hbar\frac{\partial\psi_2}{\partial t} + \frac{\hbar^2}{2m}\frac{\partial^2\psi_2}{\partial x^2} - V\psi_2\right) = 0$$

Since ψ_1 and ψ_2 satisfy Eq. (6.9), the above equation is identically zero. Hence ψ satisfies Eq. (6.9). Similarly, you can show that $a\psi_1$ is also a solution of Eq. (6.9).

2. (a) Let us express the complex potential energy V as

$$V = V_R + iV_I$$

where V_R is its real part *and* V_I, its imaginary part. The Schrödinger equation (6.16a) and its complex conjugate are then, respectively, written as:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + (V_R + iV_I)\psi$$

$$-i\hbar\frac{\partial\psi^*}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi^*}{\partial x^2} + V_R\psi^* - iV_I\,\psi^*$$

Repealing the process of obtaining Eq. (6.20), we get

$$i\hbar\left(\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}\right) = -\frac{\hbar^2}{2m}\left(\psi^*\frac{\partial^2\psi}{\partial x^2} - \psi\frac{\partial^2\psi^*}{\partial x^2}\right) + 2i\,V_I\,\psi^*\psi$$

or

$$i\hbar\frac{\partial}{\partial t}(\psi^*\psi) = -\frac{\hbar^2}{2m}\frac{\partial}{\partial x}\left(\psi^*\frac{\partial\psi}{\partial x} - \psi\frac{\partial\psi^*}{\partial x}\right) + 2i\,V_I\,\psi^*\psi$$

$$\frac{\partial P(x, t)}{\partial t} + \frac{\partial S(x, t)}{\partial x} = \frac{2V_I}{\hbar} P(x, t)$$

(b) $P(x, t) = \psi^* \psi$

$$= e^{-(\alpha - i\beta)x} e^{i\omega t} e^{-(\alpha + i\beta)x} e^{-i\omega t}$$

$$= e^{-2\alpha x}$$

Using Eq. (6.21) we obtain

$$S(x, t) = \frac{\hbar}{m} \operatorname{Im}\left[\psi^* \frac{\partial \psi}{\partial x}\right]$$

$$= \frac{\hbar}{m} \operatorname{Im}\left[e^{-(\alpha - i\beta)x} e^{i\omega t} (-\alpha - i\beta)\, e^{-(\alpha + i\beta)x} e^{-i\omega t}\right]$$

$$= \frac{\hbar}{m} \operatorname{Im}\left[e^{-2\alpha x} (-\alpha - i\beta)\right] = -\frac{\hbar\beta}{m} e^{-2\alpha x}.$$

3. The normalisation candition is

$$\int_{-\infty}^{\infty} \psi^* \psi \, dx = 1$$

or

$$|N|^2 \int_{-\infty}^{\infty} e^{-x^2} dx = 1$$

or

$$2 |N|^2 \int_0^{\infty} e^{-x^2} dx = 1 \text{ since } e^{-x^2} \text{ is an even function.}$$

By making the substitution $x^2 = t$, the integral takes the form

$|N|^2 \int_0^{\infty} t^{-1/2} e^{-t} dt$. Thus we have

$$|N|^2 \int_0^{\infty} t^{-1/2} e^{-t} dt = 1$$

or $\quad |N|^2 . \Gamma\left(\frac{1}{2}\right) = 1 \qquad \Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} t^{-1/2} e^{-t} dt$

$$= \sqrt{\pi}$$

or $\quad |N|^2 \sqrt{\pi} = 1$

or $\quad |N|^2 = \frac{1}{\sqrt{\pi}}$

and $\quad N = \left(\frac{1}{\pi}\right)^{1/4}$

Terminal Questions

1. (i) No, it does not as $\psi(x)$ does not go to zero as $|x| \to \infty$

(ii) No, its norm is infinite.

(iii) $S(x, t) = \frac{\hbar}{2mi}\left[A^* A e^{-i(px - Et)/\hbar}\left(\frac{ip}{\hbar}\right) e^{i(px - Et)/\hbar}\right.$

$$\left. - AA^* e^{i(px - Et)/\hbar}\left(-\frac{ip}{\hbar}\right) e^{-i(px - Et)/\hbar}\right]$$

$$= \frac{\hbar}{2mi} AA^* \left[\frac{2i\,p}{\hbar}\right]$$

$$= v\, AA^* \quad (\because p = mv)$$

2. Applying the normalisation condition we get

$$|N|^2 \int_{-\infty}^{\infty} x^2 e^{-x^2} dx = 1$$

or

$$2\,|N|^2 \int_0^{\infty} x^2 e^{-x^2} dx = 1 \text{ since the **integrand** is even}$$

Substituting $x^2 = t$, we get

$$|N|^2 \int_0^{\infty} t^{1/2} e^{-t} dt = 1$$

or

$$|N|^2 \frac{\sqrt{\pi}}{2} = 1, \text{ since } \int_0^{\infty} t^{1/2} e^{-t} dt = \frac{\sqrt{\pi}}{2}$$

Since $\Gamma(n + 1) = n\,\Gamma(n)$, $\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$

$$\therefore\ |N|^2 = \frac{2}{\sqrt{\pi}} \text{ and } N = \left(\frac{4}{\pi}\right)^{1/4}.$$

3. The normalisation condition is

$$\int_{-\infty}^{\infty} \psi^* \psi\, dx = 1$$

or

$$|N|^2 \int_1^{\infty} (1 + ix)(1 - ix) \exp(-2x)\, dx = 1$$

since

$$\psi(x) = 0 \text{ for } x < 1,$$

or

$$|N|^2 \int_1^{\infty} (1 + x^2) e^{-2x}\, dx = 1$$

or

$$|N|^2 \int_1^{\infty} e^{-2x} dx + |N|^2 \int_1^{\infty} x^2 e^{-2x} dx = 1$$

or

$$|N|^2 \frac{e^{-2}}{2} + |N|^2 \int_1^{\infty} x^2 e^{-2x} dx = 1$$

Integrating the second term by parts, we get

$$|N|^2\left(\frac{e^{-2}}{2}+\frac{5e^{-2}}{4}\right)=1$$

or

$$|N| = 2e/\sqrt{7}$$

This wave function cannot represent a physical system because it is discontinuous at $x = 1$. Check this out by taking the limit $x \to 1$ from right ($+\infty$) and left ($-\infty$). Both the limits are unequal.

4. The time independent Schrödinger equation for a simple harmonic oscillator is

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+2m\,(\pi\nu x)^2\,\psi = E\psi \tag{1}$$

where we have substituted $V(x) = 2m(\pi\nu x)^2$. Changing the variable to $\xi = \sqrt{\beta}\,x$, where $\beta = \frac{2\pi m\nu}{\hbar}$, we get

$$\frac{d\psi}{dx}=\frac{d\psi}{d\xi}\frac{d\xi}{dx}=\sqrt{\beta}\,\frac{d\psi}{d\xi}$$

$$\frac{d^2\psi}{dx^2}=\beta\,\frac{d^2\psi}{d\xi^2}$$

Thus, (1) becomes

$$-\frac{\hbar^2}{2m}\,\beta\,\frac{d^2\psi}{d\xi^2}+2m\pi^2\nu^2\,\frac{\xi^2}{\beta}\,\psi = E\psi$$

or
$$-\frac{\hbar^2}{2m}\,\frac{2\pi m\nu}{\hbar}\,\frac{d^2\psi}{d\xi^2}+\frac{2m\pi^2\nu^2\hbar}{2\pi m\,\nu}\,\xi^2\,\psi = E\psi$$

or
$$\frac{d^2\psi}{d\xi^2}-\frac{\pi\nu\hbar}{\pi\nu\hbar}\,\xi^2\,\psi = -\frac{2m}{\hbar^2\beta}\,E\,\psi$$

or
$$\frac{d^2\psi}{d\xi^2}+\left(\frac{2mE}{\hbar^2\beta}-\xi^2\right)\psi = 0$$

Defining $\alpha = \frac{2mE}{\hbar^2}$, we can write the equation as

$$\frac{d^2\psi}{d\xi^2}+\left(\frac{\alpha}{\beta}-\xi^2\right)\psi = 0 \qquad \textbf{(A)}$$

Substituting $\psi(\xi) = \exp\left(-\frac{\xi^2}{2}\right)$ in this equation, we get:

$$-\psi+\xi^2\,\psi+\frac{\alpha}{\beta}\,\psi-\xi^2\,\psi = 0 \qquad \left[\because \frac{d^2\psi}{d\xi^2} = -\psi+\xi^2\,\psi\right]$$

or
$$\left(\frac{\alpha}{\beta}-1\right)\psi = 0$$

or
$$\frac{\alpha}{\beta} = 1 \quad \text{since } \psi \neq 0$$

Similarly, substituting $\psi(\xi) = \xi\exp(-\xi^2/2)$ in (A), we get:

$$-3\psi+\xi^2\,\psi+\left(\frac{\alpha}{\beta}-\xi^2\right)\psi = 0$$

or
$$\left(\frac{\alpha}{\beta}-3\right)\psi = 0$$

Power series	$x = r\cos\theta = r\left(1 - \frac{\theta^2}{2!} + \dots\right)$
	$y = r\sin\theta = r\left(\theta - \frac{\theta^3}{3!} + \dots\right)$
	$z = r(\cos 8 + i\sin 8)$
	$= r\left[1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \dots\right]$
Complex number	$z = re^{i\theta}$ (polar form)
Phase factors	$e^{i\theta} = \cos\theta + i\sin 8$ and $e^{-i\theta} = \cos 8 - i\sin\theta$
Real part	$\mathrm{Re}\, e^{i\theta} = \cos\theta = (e^{i\theta} + e^{-i\theta})/2$
Imaginary part	$\mathrm{Im}\, e^{i\theta} = \sin\theta = (e^{i\theta} - e^{-i\theta})/2i$

2022

[This question paper contains 4 printed pages.]

Your Roll No. 2022

C

Sr. No. of Question Paper : 1015

Unique Paper Code : 32221501

Name of the Paper : Quantum Mechanics and Applications

Name of the Course : **B.Sc. (Honors) Physics**

Semester : V

Duration : 3 Hours Maximum Marks : 75

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt **FIVE** questions in all. Question No. **1** is compulsory.
3. All questions carry equal marks.
4. Non programmable calculators are allowed.

1. Attempt any **FIVE** of the following :

   (a) Calculate the commutator $[\widehat{L_x}, \widehat{p_x}]$. (given $[\hat{x}, \widehat{p_x}] = i\hbar$).

   (b) The wave-function of a particle is $\psi(x) = \sqrt{\frac{2}{L}} \sin\frac{n\pi x}{L}$ for $0 \le x \le L$. Determine the probability of finding the particle at x = L/3 for n = 3 state.

P.T.O.

(c) Derive the relation between 'magnetic dipole moment' and 'orbital angular momentum' of an electron revolving around a nucleus.

(d) Write the quantum numbers for the state represented by $4^2F_{5/2}$.

(e) Normalize the wave function e^{-ax^2} in a one-dimensional space.

(f) A free particle of mass m is described by the wave-function $\psi(x) = A\exp(i\mu x)$ where A and μ are constants. Determine the probability current density for this particle.

(g) Determine the uncertainty in position for the normalized wave-function $\psi(x) = \left(\frac{2\alpha}{\pi}\right)^{1/4} e^{-\alpha x^2}$ for $-\infty < x < \infty$. (5×3=15)

2. (a) Explain the concept of expectation values. Give expressions for the expectation values of velocity, momentum and energy in terms of respective operators in three dimensions. Mention the difference between expectation values and eigenvalues of an operator corresponding to a dynamical variable.

(b) The wave-function of a particle of mass m is given by

$$\Psi(x) = \left(\frac{\beta^2}{\pi}\right)^{1/4} e^{-\beta^2 x^2/2} \text{ for } -\infty < x < \infty.$$

Determine the total energy of the particle, if potential energy is $V(x) = \frac{1}{2}m\omega^2x^2$. (7.8)

3. The Gaussian wave packet for a free particle is defined by the wave function

$$\Psi(x,0) = N\,exp\left(-\frac{x^2}{2\sigma^2} + ik_0x\right).$$

Prove that the centre of this Gaussian wave packet travels with a velocity $v = \frac{k_0\hbar}{m}$.

(Use $\int_{-\infty}^{\infty} e^{-x^2/\sigma^2}\,dx = \sigma\sqrt{\pi}$ and $\int_{-\infty}^{\infty} e^{-(ax^2 \pm bx)}\,dx = \sqrt{\frac{\pi}{a}}\,e^{b^2/4a}$).

(15)

4. (a) Solve the Schrodinger equation for a linear Harmonic Oscillator and obtain first two eigenfunctions. (10)

(b) Find ΔX and ΔP for the ground state eigenfunction of linear Harmonic Oscillator and obtain the uncertainty principle. (5)

5. (a) The 'θ' equation obtained after applying separation for variables to the Schrodinger equation for a 3D hydrogen atom in spherical polar coordinates, is given by

$$\frac{1}{sin\theta}\frac{d}{d\theta}\left(sin\theta\frac{d\Theta}{d\theta}\right) + \left(\lambda - \frac{m_l^2}{sin^2\theta}\right)\Theta = 0.$$

Solve the above equation for $m_l = 0$ (or otherwise) to show that

$$\lambda = l(l+1), \quad l = 0,1,2, \dots\dots\dots$$ (12)

P.T.O.

(b) An electron in a hydrogen atom is in a state described by

$$\psi = \frac{1}{\sqrt{6}}[2\psi_{100} + \psi_{211} + \psi_{21-1}]$$

Calculate the expectation value of L_z in this state.

(Given $\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}}$ and

$$\psi_{100} = \frac{1}{\sqrt{\pi}}\left(\frac{1}{a_0}\right)^{\frac{3}{2}} e^{-r/a_0}$$

$$\psi_{211} = \frac{1}{8\sqrt{\pi}a_0^{3/2}}\frac{r}{a_0}e^{-r/2a_0}\sin\theta\, e^{+i\phi}\ )$$ (3)

6. (a) What is Larmor Precession? Draw the relevant diagram and derive the expression for Larmor frequency.

(b) A beam of silver atoms moving with a velocity 10^7 cm/s passes through a magnetic field of gradient 0.5 Wb/m²/cm for 10 cm. What is the separation between the two components of the beam as it comes out of the magnetic field? (8,7)

7. (a) What is spin orbit coupling? Explain the fine structure splitting in the energy levels due to this. For the 2p level of the hydrogen atom with $E_n = -3.14$ eV, evaluate the fine structure splitting.

(b) Consider a two-electron system with $l_1 = 1$, $l_2 = 1$. Explain the LS coupling scheme in such a case. Write the spectral notation for each state. (10,5)

(2000)

Name of the Department: PHYSICS DEPARTMENT

Name of Course: B.Sc. Hons.–CBCS_Core (NC)

2021

Semester: V- Semester

Name of the Paper: Quantum Mechanics and Applications

Unique Paper Code: 32221501

Time Duration: 3 Hours **Maximum Marks: 75**

Attempt four questions out of six. Each question carries equal marks.

1.

i. A particle is represented at (t=0) by the wave function:

$$\psi(x,0) = \begin{cases} A(a^2 - x^2) & -a \le x \le a \\ 0 & otherwise \end{cases}$$

Find A and expectation value of x, x^2, p and p^2. Find uncertainties in position and momentum.

ii. Show that Divergence of J (probability current density) is zero for stationary states.

iii. Find the Fourier transform of the wave function e^{-ax^2}.

(14.75+2+2)

2.

i. State Heisenberg's Uncertainty principle. What is the origin of concept of uncertainty in position and momentum? Derive $\Delta x \Delta p \geq \hbar/2$.

ii. Verify whether the following operators are linear:

a. $\widehat{f(x)} = \frac{d}{dx} f(x)$

b. $\widehat{f(x)} = \sqrt{f(x)}$

iii. What is uncertainty in the location of a photon of wavelength 5000 Angstrom which is known to an accuracy of one part in 10^7?

(12.75+3+3)

3.

i. Solve Schrodinger's equation for the potential energy $V = (1/2)kx^2$ and show that the energy eigenvalue are $E_n = (n + \frac{1}{2})\hbar\omega$.

ii. Which of the following wave functions

(i) e^{-ax^2} (ii) $\sin(kx)$

are eigenvalues of operator (a) p and (b) p^2.

iii. Find the locations of classical turning points for a One Dimensional Harmonic Oscillator in its ground state.

(12.75+4+2)

4.

i. Describe the Stern-Gerlach Experiment and its theory. Discuss the significance of the experiment. Why is an inhomogeneous magnetic field required?

ii. A beam of silver atoms moving with a velocity of $10^5 cm/s$ passes through a magnetic field of gradient 0.50Wb/m^2/cm for a distance of 10cm. Determine acceleration of Ag atoms, time spent by atoms in the field and displacement of Ag atoms along z-direction as it comes out of the magnetic field (along z-axis).

iii. Show that $\frac{d}{dt}\int_{-\infty}^{\infty} \psi_1^* \psi_2 dx = 0$ for any two (normalizable) solutions to Schrodinger's equation, ψ_1 *and* ψ_2.

(12.75+3+3)

5.

i. Derive an expression for energy difference ΔE between doublets due to Spin-Orbit coupling. How does ΔE depend on quantum numbers n and l?

ii. Show that the angle between angular momentum (L) and z-axis is given by $\theta_{m_l} = cos^{-1}(\frac{m_l}{\sqrt{l(l+1)}})$. Find the values of angle θ_{m_l} for l=2.

iii. Calculate the probability of finding the electron in the region $\frac{a_0}{2} < r < 2a_0$ in a hydrogen atom in ground state given that wave function for the ground state of Hydrogen atom is $\psi(r) = \frac{1}{\sqrt{\pi a_0^3}} e^{\frac{-r}{a_0}}$, where $a_0 = \frac{4\pi\epsilon_0 \hbar^2}{me^2}$.

(12.75+3+3)

6.

i. Solve 1-D time independent Schrodinger's equation for a particle having energy E for a square well of finite depth V_0 ($E<V_0$). Show graphically existence of bound states.

ii. An electron moves in 1-D potential well of width 8 Angstrom and depth 12eV. Find the number of bound states?

iii. Assuming LS coupling scheme, list the possible total angular momentum and spectral terms for three electron having configuration 2p 3p 4d.

(12.75+3+3)

Physical Constants:

Mass of Ag atom = 107.87amu,

Charge of electron = 1.6×10^{-19}C,

h=6.626×10^{-34}Js

Mass of electron = 9.1×10^{-31}Kg.

[This question paper contains 8 printed pages.] 2019

Your Roll No...............

Sr. No. of Question Paper : 7486A J

Unique Paper Code : 32221501

Name of the Paper : Quantum Mechanics and Applications

Name of the Course : B.Sc. (Hons.) Physics

Semester : V

Duration : 3 hours Maximum Marks : 75

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt **five** questions in all **Q.1** and all its parts are compulsory.
3. Attempt any **four** questions from the rest. Also, attempt any **two** parts out of **three** from each question.
4. Non-programmable calculators are allowed.

1. (a) Normalize the wave function, $\Psi(x) = e^{-|x|/a}$. (3)

P.T.O.

(b) Write the Schrodinger equation for a system of two particles of masses m_1 and m_2 carrying charges e_1 and e_2 respectively in what kind of field? (2)

(c) Given that the position and momentum operators are Hermitian, verify whether the operator $\hat{x}^2 + \hat{x}\hat{p}_x$, is Hermitian. (3)

(d) Write the values of quantum numbers n, l, s, j, m_j for the following states :

(i) $2\,^2S_{1/2}$ (ii) $5\,^2F_{5/2}$. (3)

(e) Consider the state, $\psi = \sqrt{\frac{1}{10}}\phi_1 + \sqrt{\frac{3}{5}}\phi_2 + \sqrt{\frac{3}{10}}\phi_3$, where ϕ_n are orthonormal eigenstates of an operator $\hat{A}$. Find the expectation value of the operator $\hat{A}$ in the state ϕ, if it satisfies the eigenvalue equation $\hat{A}\phi_n = (2n^2+1)\phi_n$. (3)

(f) Write down the wave function for a system of (i) two Bosons and (ii) two fermions indistinguishable. (2)

(g) What is the probability that an electron in the state

$$\psi_{210} = \frac{1}{\sqrt{\pi}}\left(\frac{1}{a_0^{5/2}}\right) r e^{-r/2a_0} \cos(\theta)$$ of the hydrogen

atom, exists between a distance of $3a_0$ to $6a_0$ from the nucleus. (3)

2. (a) (i) Set up the time dependent Schrodinger equation and hence derive the time independent Schrodinger equation. (4)

(ii) Starting with the Schrodinger equation in one dimension and using a de Broglie plane wave as a solution, show that when $V = 0$ this leads to the correct nonrelativistic relationship between energy and momentum. (3)

(b) For a Gaussian wave packet

$$\left(\psi(x) = Ae^{\frac{x^2}{4\alpha^2}} e^{-i(k_0 x - w_0 t)}\right)$$

corresponding to a free particle (i) Find the probability current density and (ii) Verify the continuity equation. (4+3)

(c) (i) Explain spreading of a Gaussian wave packet for a flee particle in one dimension. (5)

(ii) Calculate the fractional change in the width of the wave packet in one second if the wave packet corresponds to a particle of mass 6.644×10^{-27} Kg. The initial width being of the order of 10^{-10} m. (2)

3. (a) Write the Schrodinger equation for a linear Harmonic oscillator and solve it to obtain the energy eigen values. (7)

(b) (i) A Harmonic Oscillator has a wave function which is a superposition of the ground state and the second excited state eigenfunctions $\psi(x) = \psi_0(x) + 2\psi_2(x)$.

Find the expectation value of the energy. (3)

(ii) Using the Uncertainty Principle show that the ground state energy for a Harmonic Oscillator is non-zero. (4)

(c) (i) An electron is confined in the ground state of a one-dimensional harmonic oscillator such that $\Delta x = 10^{-10}$m. Assuming that the average

Kinetic energy is equal to the average Potential energy, find the energy in electron volts required to excite it to the first excited state. (4)

(ii) For a linear harmonic oscillator in its ground state, show that the probability of finding it beyond the classical limits is approximately 0.16. (3)

4. (a) (i) Obtain the solution for the Legendre equation

$$\left(1-\xi^2\right)\frac{d^2P(\xi)}{d\xi^2}-2\xi\frac{dP(\xi)}{d\xi}+\lambda P(\xi)=0,$$

What are the conditions that need to be imposed so that the solutions are well behaved? What do the conditions imply. (4)

(ii) Verify whether the function $Y_{1,1}(\theta,\phi) = -\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\phi}$ is an eigenstate of the following angular momentum operator :

$$\hat{L}_x = i\hbar\left(\sin\phi\frac{\partial}{\partial\theta}+\cot\theta\cos\phi\frac{\partial}{\partial\phi}\right) \quad (3)$$

(b) Calculate $\langle V(r)\rangle = -\frac{e^2}{4\pi\epsilon_0}\left\langle\frac{1}{r}\right\rangle$ for the first exci

state of the hydrogen atom with the wave functi

$$\left[\psi_{210} = \frac{1}{\sqrt{\pi}}\left\{\frac{1}{2a_0}\right\}^{5/2} re^{-r/2a_0}\cos\theta\right].$$

(c) The electron in the hydrogen atom is replaced a muon of mass $m_\mu \approx 200m_e$, where m_e is the ma of the electron. Determine the correspondi changes in the following :

(i) The Larmor frequency and hence t Zeeman splitting for the 2p level in t presence of a magnetic field of 1 Tes (Ignore the electron spin)

(ii) The wavelength of the corresponding l line. Will it be in the visible regio (Rydberg constant $R = 1.097 \times 10^7 m^{-1}$ f the hydrogen atom) (3+

5. (a) Consider a particle trapped inside a on dimensional finite square well. Solve tim independent Schrodinger equation for the syste and obtain the bound state eigenfunctions. Discu how the energy levels are obtained graphically. (7

(b) (i) Derive the relationship between magnetic dipole moment and orbital angular momentum of an electron revolving around a nucleus. (3)

(ii) Explain space quantization. Calculate the possible orientation of the total angular momentum vector J corresponding to $j = 3/2$ with respect to a magnetic field along the z-axis. (4)

(c) (i) What is Larmor Precession? Derive the expression for Lannor frequency. (4)

(ii) A beam of electron enters a uniform magnetic field of flux density 1.2 tesla. Calculate the energy difference between electrons whose spins are parallel and antiparallel to the field. (3)

6. (a) (i) Explain Normal Zeeman Effect. (2)

(ii) Write the term diagram for the splitting of the yellow line of sodium $(1s^2, 2s^2, 2p^6)3s^1$ into two components D1 and D2. (2)

(iii) In a Stern-Gerlach experiment, a beam hydrogen atom with velocity 3×10^{3} passes through an inhomogeneous magn field of length 50 cm and having gradien 200 T/m perpendicular to the direction of incident beam. Find out the transve deflection of the atoms at the point wh the beam leaves the field. (Bohr magne 9.24×10^{-24} J/T, $M = 1.67 \times 10^{-27}$Kg).

(b) (i) Write down the normal electron configuration of Carbon atom $(Z = 6)$ a obtain the spectral terms arising fr equivalent electrons.

(ii) The quantum number of two optical electr in a two valence electron atom are

$n_1 = 6,\ l_1 = 3,\ s_1 = 1/2$

$n_2 = 5,\ l_2 = 1,\ s_2 = 1/2$

assuming j-j coupling, find the possible val of J.

(c) (i) What is spin orbit coupling. Calculate change in the energy level due to this.

(ii) Write the term symbol for the ground st of the hydrogen atom in the LS coupl scheme.

2018

This question paper contains 4 printed pages.

Your Roll No. ....

Sl. No. of Ques. Paper: 107 I

Unique Paper Code : 32221501

Name of Paper : Quantum Mechanics and Applications

Name of Course : B.Sc. (Hons.) Physics

Semester : V

Duration : 3 hours

Maximum Marks : 75

(Write your Roll No. on the top immediately on receipt of this question paper.)

Attempt five questions in all.
Q. No. 1 is compulsory.
All questions carry equal marks.

Non-programmable calculators are allowed.

1. Attempt any *five* of the following:

(a) State linearity and superposition principle.

(b) Prove that:

$$[x^n, \hat{p}] = -in\hbar x^{n-1}.$$

(c) What are stationary states? Why are they called so?

(d) What are the conditions for a wavefunction to be physically acceptable?

P. T. O.

(e) What do you mean by space quantization? Explain.

(f) Write the quantum numbers for the state represented by:

$$3^2 D_{3/2}.$$

(g) Define group velocity and phase velocity.

5×3=15

2. (a) Set up the time dependent Schrödinger equation and hence derive the time independent Schrödinger equation.

(b) Derive the expressions for probability density and probability current densities in three dimensions and hence derive the equation of continuity. 7,8

3. (a) Give the theory to explain spreading of a Gaussian wave packet for a free particle in one dimension.

(b) Normalize the following wave function for a particle in one dimension:

$$\begin{cases} A \sin\left(\dfrac{\pi x}{a}\right) & 0<x<a \\ 0 & \text{outside} \end{cases}$$

10,5

4. (a) Solve the Schrödinger equation for a Linear

Harmonic Oscillator to show that the energy eigenvalues are:

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

(b) A Harmonic Oscillator has a wave function which is superposition of its ground state and first excited state normalized eigenfunctions are given by:

$$\Psi(x) = \frac{1}{\sqrt{2}}\ [\psi_0(x) + \psi_1(x)].$$

Find the expectation value of the energy. 10,5

5. Write the Schrödinger equation for a 3D hydrogen atom in spherical polar coordinates. Derive three separate equations for r, θ, φ using the method of separation of variables. Solve the equation for φ to obtain the normalized eigenfunctions and show that they are orthogonal. 15

6. (a) Describe Stern Gerlach experiment with necessary theory. What does it demonstrate?

   (b) Explain Normal Zeeman Effect with examples and energy diagram. 8,7

7. (a) What is spin orbit coupling? Calculate the change in the energy levels due to this.

(b) Show the result of an LS coupling of two non-equivalent p-electrons. 10,5

2017

[This question paper contains 6 printed pages.]

Your Roll No. ....

Sr. No. of Question Paper : 5774 **H**

Unique Paper Code : 222502

Name of the Paper : PHHT-516 : Quantum Mechanics

Name of the Course : **B.Sc. (Hons.) Physics**

Semester : V

Duration : 3 Hours .Maximum Marks : 75

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt **five** questions in all.
3. Question No. **1** is compulsory.
4. Use of non-programmable scientific calculator is allowed.
5. Symbols have their usual meaning.

1. Attempt any **five** of the following : (5×3=15)

   (a) What is the minimum uncertainty in the energy state of an atom if an electron remains in this state for 10^{-8} s?

   (b) Using the expression $[x, p_x] = i\hbar$, find $[x, L_y]$.

P.T.O.

(c) A proton is moving non-relativistically having kinetic energy 1 MeV. Find its de-Broglie wavelength.

(d) Determine whether the following wave functions are physically acceptable solutions of Schrodinger wave equation or not :

(i) Ae^{x}

(ii) Ae^{-x}

(iii) Ae^{-x^2}, $-\infty \leq x \leq \infty$

(e) Find the probability that a particle in a box L wide can be found between $x = 0$ and $x = L/2$ when it is in the first excited state.

(f) For hydrogen atom what are the possible values of l and m_l for $n = 2$?

(g) The azimuthal wave function for the hydrogen atom is

$$\Phi(\phi) = Ae^{im_l\phi}, \ 0 \leq \phi \leq 2\pi.$$

Find the normalization constant A.

(h) An electron in H-atom is in the 3p state. Which downward transitions (1s, 2s, 2p) are forbidden by the selection rules?

2. An x-ray photon of wavelength 0.05 nm strikes a free electron at rest and the scattered photon departs at 90° from the initial photon direction.

(a) Determine the momenta of the incident photon, the scattered photon and the scattered electron.

[Given: $\tan^{-1}(0.9542) = 43.66°$, $\sin(43.66°) = 0.69$]

(2,4,6)

(b) Determine the Kinetic energy of the scattered electron. (3)

3. (a) Suppose an electron at rest absorbs the incident photon and moves with the speed v along the direction of incident photon. Using the laws of momentum and energy conservation, determine the value(s) of v. (6)

(b) The photoelectric threshold wavelength for a material is 5000 Å. Find

(i) the work function of this material

(ii) the maximum kinetic energy of the photoelectrons if light of 4000 Å strikes the surface of this material

(iii) the stopping potential for 4000 Å photons

(3,4,2)

4. (a) Determine the phase velocity and group velocity of the wave corresponding to a de Broglie wavelength of $\lambda_s = h/p = h/mv$. (3,5)

(b) A free particle of mass m is described by a wave function

$\psi(x) = e^{ipx/\hbar}$; p is the momentum of the particle

prove that the probability current density is equal to the speed of this particle. (7)

5. (a) A particle of mass m moves inside an infinite potential well :

$$V = \begin{cases} 0\,, 0 < x < L \\ \infty, x < 0 \text{ and } x > L \end{cases}$$

Find the energy eigenvalues and the normalized wave functions of this particle. (10)

(b) Determine the average of squared-momentum, $< p^2 >$ of this particle, when it is in the ground state. (5)

6. (a) A particle of mass m is moving in a harmonic potential well,

$$V(x) = \frac{1}{2} m\,\omega^2\, x^2,\ -\infty \le x \le \infty.$$

If this particle is described by a wave function $\psi(x) = Axe^{-m\omega x^2/2\hbar}$, then find

(i). A

(ii) Energy of this particle in the given state.

(4,6)

(b) Determine the probability of finding this harmonic oscillator in the classically forbidden region, if it is in the ground state. (5)

$$\left[\textit{Given: } \psi_0(x) = \left(\frac{\beta^2}{\pi}\right)^{1/4} \exp\left(\frac{-\beta^2 x^2}{2}\right),\ -\infty \le x \le \infty,\ \beta = \sqrt{\frac{m\omega}{\hbar}}\right.$$

$$\left. \textit{and } \frac{2}{\sqrt{\pi}}\int_0^1 e^{-u^2}\,du = 0.843,\ \frac{2}{\sqrt{\pi}}\int_1^\infty e^{-u^2}\,du = 0.157\right]$$

7. (a) Prove that z-component of angular momentum operator is given by

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$$ (6)

(b) An electron in hydrogen atom is in the state $\psi(\theta, \phi) = A\sin^2\theta\, e^{i2\phi}$, find

(i) A,

(ii) L, the magnitude of angular momentum

(iii) L_z, the magnitude of z-component of angular momentum

$$\left[\textit{Given: } \hat{L}^2 = -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \operatorname{cosec}^2\theta\frac{\partial^2}{\partial\phi^2}\right)\right]$$

(3,4,2)

8. A particle of mass w and energy E moves from a region of potential V_0 towards the region of zero potential, as shown below in Fig. (1).

(a) Explain why energy of this particle should be greater than V_0.

(b) Derive the expressions for the reflection and transmission co-efficients of this particle. (2,13)

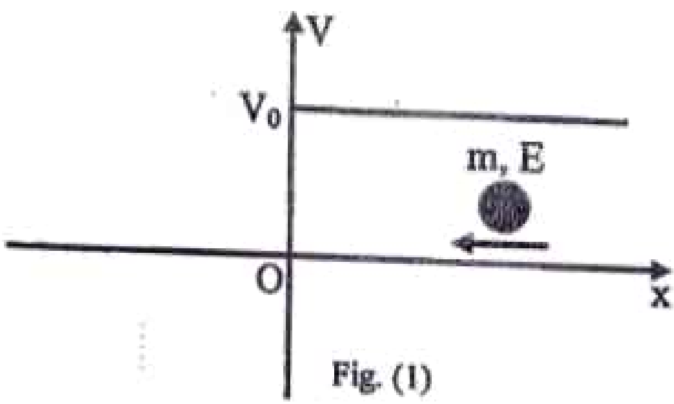

Fig. (1)

Physical Constants:

$h = 6.626 \times 10^{-34}$ J.s $= 4.136 \times 10^{-15}$ eV.s

1 eV $= 1.602 \times 10^{-19}$ J

$m_e = 9.1 \times 10^{-31}$ kg, $m_ec^2 = 0.511$ MeV, $m_pc^2 = 938.3$ MeV

$$\int_0^\infty x^n e^{-ax^m} dx = \frac{1}{m a^{\frac{n+1}{m}}} \Gamma\left(\frac{n+1}{m}\right)$$

This question paper contains 3 printed pages] 2017

Roll No. | | | | | | | | | | | |

S. No. of Question Paper : **6676**

Unique Paper Code : **32221501** **HC**

Name of the Paper : **Quantum Mechanics and Applications**

Name of the Course : **B.Sc. (Honors) Physics**

Semester : **V**

Duration : **3 Hours** Maximum Marks : **75**

(Write your Roll No. on the top immediately on receipt of this question paper.)

Attempt *five* questions in all. Question No.1 is compulsory.

All questions carry equal marks.

Non-programmable calculators are allowed.

1. Attempt any *five* of the following :

(*a*) Give the expressions for Energy, Linear momentum and Angular momentum in operator form.

(*b*) Calculate the commutator $[\widehat{L_x}, \widehat{p_x}]$. (given $[\hat{x}, \widehat{p_x}] = i\hbar$)

(*c*) Explain uncertainty principle with an example.

P.T.O.

(*d*) Derive the relation between 'magnetic dipole moment' and 'orbital angular momentum' of an electron revolving around a nucleus.

(*e*) Write the quantum numbers for the state represented by $4^2F_{5/2}$.

(*f*) What are symmetric and antisymmetric wave functions ?

(*g*) What are free and bound states ? Explain. 5×3=15

2. Consider a particle trapped inside a one-dimensional finite square well. Solve the time independent Schrödinger equation for the system and obtain the bound state eigenfunctions. Discuss how the energy levels are obtained graphically ? 15

3. (*a*) Calculate the expectation value of the momentum for wavefunction.

$$\begin{cases} \sqrt{\frac{2}{a}}\sin\left(\frac{nx}{a}\right) & \text{for } 0<|x|<a \\ 0 & \text{for } |x|>a \end{cases}$$

(*b*) Give the theory to explain spreading of a Gaussian wave packet for a free particle in one-dimension. 5,10

4. (*a*) Solve the Schrödinger equation for a Linear Harmonic Oscillator and hence obtain first three eigenfunctions.

(*b*) A particle of mass 1 mg is attached to a spring of spring constant 0.001 Nm^{-1}. Calculate its zero point energy. 10,5

5. The 'θ' equation obtained after applying separation for variables to the Schrödinger equation for a 3D hydrogen atom in spherical polar coordinates, is given by

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+\left(\lambda-\frac{m_l^2}{\sin 2\theta}\right)\Theta=0.$$

Solve the above equation for $m_l = 0$ (or otherwise) to show that :

$$\lambda = l\,(l+1),\ l = 0,1,2,........$$ 15

6. (*a*) What is Larmor Precession ? Derive the expression for Larmor frequency.

(*b*) Explain Normal Zeeman Effect with examples and energy diagram. 8,7

7. (*a*) What is spin orbit coupling ? Calculate the change in the energy levels due to this.

(*b*) Show the result of a JJ coupling of two nonequivalent p-electrons. 10,5

This question paper contains 2 printed pages]

2016

Roll No. | | | | | | | | | |

S. No. of Question Paper : 91

Unique Paper Code : 222563

G

Name of the Paper : **Physics-V : Quantum Mechanics and Atomic Physics (PHPT-505)**

Name of the Course : **B.Sc. Physical Science / Applied Physical Science**

Semester : **V**

Duration : **3 Hours**

Maximum Marks : **75**

(Write your Roll No. on the top immediately on receipt of this question paper.)

Attempt any *five* questions.

All questions carry equal marks.

1. (*a*) What is a wave packet ? Prove that the de-Broglie wave packet associated with a moving body travels with the same velocity as the body. 2,5

   (*b*) State Heisenberg uncertainty principle and derive it on the basis of wave packets. 2,6

2. (*a*) What is photoelectric effect ? Define threshold frequency and work function. Give an account of Einstein explanation of photoelectric effect on the basis of quantum theory. 3,3,5

   (*b*) Radiations of wavelength 5400 Å fall on a metal plate whose work function is 1.9 eV. Find the kinetic energy of the emitted photoelectrons. 4

3. (*a*) What are admissibe conditions for a wave function ? 3

   (*b*) How do you correlate the operator $\hat{H}$ to its corresponding phyiscal quantity ? 3

   (*c*) Starting with time dependent Schrodinger wave equation, derive an expression for time-independent wave function. 9

P.T.O.

4. Write short notes on any *two* of the following : 15

(*a*) Inadequacies of classical mechanics

(*b*) γ-ray microscope experiment

(*c*) Compton effect

5. (*a*) What is Normal Zeeman effect ? Derive an expression for the frequency shift in Normal Zeeman effect. 3,7

(*b*) Consider the normal Zeeman effect in the $3d \rightarrow 2p$ transition. Draw the energy-level diagram that shows the splitting of $3d$ and $2p$ levels in an external magnetic field. Also indicate all the possible transitions. 5

6. (*a*) What is Bohr Magneton ? Give its unit. 3

(*b*) Discuss symmetric and anti-symmetric wave functions. 5

(*c*) Explain space quantization of L and S with the help of an example. 7

7. (*a*) Discuss LS and *jj* coupling. 6

(*b*) Find the possible values of total angular momentum quantum number J in LS coupling of two atomic electrons having orbital quantum numbers $l_1 = 2$ and $l_2 = 1$. 4

(*c*) What is the physical significance of the magnetic orbital quantum number and magnetic spin quantum number ? 5

8. (*a*) State and explain the Pauli's exclusion principle. 4

(*b*) Obtain an expression for the maximum number of electrons that can be accommodated in a shell. Give electronic configurations for the following elements : 4,2,2

(*i*) Zn (Z = 30) and

(*ii*) La (Z = 57)

(*c*) Find the S, L and J values that correspond to each of the following states : 3

$^2S_{1/2}$, $^3P_2$, $^2D_{3/2}$.

This question paper contains 4+2 printed pages]

Roll No. [ ]

S. No. of Question Paper : 6216

Unique Paper Code : 222502 D

Name of the Paper : **Quantum Mechanics (PHHT-516)**

Name of the Course : **B.Sc. (Hons.) Physics**

Semester : **V**

Duration : **3 Hours** Maximum Marks : **75**

(Write your Roll No. on the top immediately on receipt of this question paper.)

Attempt *Five* questions in all. Question No. 1 is compulsory.

Use of non-programmable scientific calculator is allowed.

Symbols have their usual meaning.

1. Attempt any *five* of the following : $3\times5=15$

(*a*) A metal surface when irradiated with light of wavelength 5896 Å emits electrons for which the stopping potential is 0.12 V. When the same surface is irradiated with 2830 Å, it emits electrons for which the stopping potential is 2.20 V. Calculate the value of Planck's constant.

P.T.O.

(*b*) Compare the de Broglie wavelengths for an electron with a kinetic energy of 1 eV and a ball of mass 300 gm travelling at 100 km/hr.

(*c*) Determine the smallest possible uncertainty in the position of an electron moving with velocity 3×10^7 m/s.

(*d*) Establish time independent form of Schrödinger equation for stationary states.

(*e*) Determine the probability of finding a particle of mass m between $x = 0$ and $x = L/10$, if the particle is described by the normalized wave function :

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}$$

for $0 \le x \le L$ and is in the $n = 3$ state.

(*f*) A radial function in spherical polar coordinates is :

$$R_n(r) = Ce^{-r/2}\, U_n(r),$$

where C is a normalization constant. Discuss the physical acceptability of $R_n(r)$ if $U_n(r)$ behaves as :

(*i*) $1/r^2$ for small values of r and as a polynomial in r otherwise; and

(*ii*) a polynomial in r of degree more than 3 for all values of r.

(g) A one-dimensional harmonic oscillator is in a state described by the wave function :

$$\psi(x, 0) = \frac{1}{2}\psi_0(x) + \frac{i}{\sqrt{2}}\psi_1(x) + \frac{1}{\sqrt{2}}e^{i\pi/3}\psi_2(x)$$

where $\psi_n(x)$ are the usual normalized orthogonal wave functions. Normalize the wave function $\psi(x, 0)$.

(h) Find the classical amplitude of a one-dimensional harmonic oscillator in its ground state with an energy $\frac{1}{2}\hbar\omega$.

2. (a) In the Compton scattering of a photon of frequency ν by a free electron through an angle □, using the expressions for momentum conservation : 10

$$pc\cos\theta = h\nu - h\nu'\cos\phi \text{ and } pc\sin\theta = h\nu'\sin\phi$$

and the expression for change in wavelength of scattered photon :

$$\lambda' - \lambda = \frac{h}{m_o c}(1 - \cos\phi)$$

prove that :

$$\tan\theta = \frac{1}{[1+\beta]\tan\frac{\phi}{2}}; \text{ where } \beta = \frac{h\nu}{m_o c^2}$$

(b) A metal surface emits electrons with maximum kinetic energies E_1 and E_2 when illuminated with light of wavelengths λ_1 and λ_2 respectively, where $\lambda_1 > \lambda_2$. Prove that the Planck's constant h and work function ϕ of the metal are given by : 5

$$h = \frac{(E_2 - E_1)\lambda_1 \lambda_2}{c(\lambda_1 - \lambda_2)} \text{ and } \phi = \frac{E_2 \lambda_2 - E_1 \lambda_1}{(\lambda_1 - \lambda_2)}.$$

3. (a) Explain de Broglie hypothesis for matter waves. Show that for the de Broglie wave associated with a moving particle the group velocity is equal to the particle velocity. 5

(b) Assume that at time $t = 0$, a single non-interacting electron is located near $x = x_0$ with the probability P dx of finding it between x and $x + dx$ being given by :

$$\psi(x, 0) = Ae^{-(x - x_0)^2/2a^2} e^{ip_0 x/\hbar}.$$

Obtain the expectation values of x and p. Also show that : 10

$$\Delta x \,.\, \Delta p = \hbar/2.$$

4. (a) Describe an experiment to locate the position of a free electron by a microscope using γ ray and hence, obtain an expression for uncertainty principle. 10

(b) Determine the minimum uncertainty in the position of a particle in terms of de Broglie wavelength when the uncertainty in the velocity of a particle in one-tenth of its velocity. 5

5. (a) Obtain the energy eigen values and the normalized wave functions for a free particle of mass m trapped in a one-dimensional box of length L along x-axis in positive direction of x from the origin. 10

(b) The wave function of a particle confined in a box of linear dimension L along x-axis is :

$$\psi(x) = Ae^{ikx};\ 0 \leq x \leq L.$$

Find the probability of finding the particle in the distance $0 \leq x \leq \frac{L}{4}$. 5

6. Solve the time independent Schrödinger equation for the energy levels of a one-dimensional harmonic oscillator. Draw the energy level diagram. Explain the physical significance of zero-point energy. 15

7. A particle of mass m and energy E moves along x-axis from a region of zero potential towards a one-dimensional step potential barrier of height V_0 of infinite extent. Assuming $E > V_0$, derive expressions for the reflection and transmission coefficients. Comment on the wavelengths associated with the incident, reflected and transmitted waves. Also, obtain expressions for probability current densities associated with the incident, reflected and transmitted waves. 15

P.T.O.

ks : 75

3×5=15

or which

330 Å, it

'Planck's

P.T.O.

8. (a) Solve the angular equation of hydrogen atom in spherical polar coordinates, given as :

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial y}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 y}{\partial\phi^2}+\lambda y=0$$

to obtain relation $\lambda = l(l+1)$. 11

(b) Given the wave function of ground state of hydrogen atom is :

$$\psi_{100}(r)=\frac{e^{-r/a_0}}{(\pi a_0^3)^{1/2}},$$

where the symbols have usual meaning and :

$$a_0=\frac{4\pi\epsilon_0\hbar^2}{me^2}.$$

Calculate the most probable distance of electron from nucleus in ground state. 4

$\hbar = 1.054 \times 10^{-34}$ Js

$h = 6.63 \times 10^{-34}$ Js

Rest mass of electron = 9.1×10^{-31} kg

Charge of electron = 1.6×10^{-19} C

Rest mass energy of electron = 512 KeV

Velocity of light in free space = 3×10^8 m/s.

6.1: Time-dependent Schrödinger Equation

When we first introduced quantum mechanics, we saw that the fourth postulate of QM states that: *The evolution of a closed system is unitary (reversible). The evolution is given by the time-dependent Schrödinger equation*

$$i\hbar\frac{\partial|\psi\rangle}{\partial t}=\mathcal{H}|\psi\rangle$$

where $\mathcal{H}$ *is the Hamiltonian of the system (the energy operator) and* $\hbar$ *is the reduced Planck constant* ($\hbar=h/2\pi$ with h the Planck constant, allowing conversion from energy to frequency units).

We will focus mainly on the Schrödinger equation to describe the evolution of a quantum-mechanical system. The statement that the evolution of a closed quantum system is unitary is however more general. It means that the state of a system at a later time t is given by , where $U(t)$ is a unitary operator. An operator is unitary if its adjoint $\mathrm{U}^{\dagger}$ (obtained by taking the transpose and the complex conjugate of the operator, $U^{\dagger}=(U^{*})^{T}$) is equal to its inverse: $U^{\dagger}=U^{-1}$ and $UU^{\dagger}=1$.

Note that the expression $|\psi(t)\rangle=U(t)|\psi(0)\rangle$ is an integral equation relating the state at time zero with the state at time t. For example, classically we could write that $x(t)=x(0)+vt$ (where v is the speed, for constant speed). We can as well write a differential equation that provides the same information: the Schrödinger equation. Classically for example, (in the example above) the equivalent differential equation would be $\frac{dx}{dt}=v$ (more generally we would have Newton's equation linking the acceleration to the force). In QM we have a differential equation that control the evolution of closed systems. This is the Schrödinger equation:

$$\boxed{i\hbar\frac{\partial\psi(x,t)}{\partial t}=\mathcal{H}\psi(x,t)}$$

where $\mathcal{H}$ is the system's Hamiltonian. The solution to this partial differential equation gives the wavefunction $\psi(x,t)$ at any later time, when $\psi(x,0)$ is known.

Solutions to the Schrödinger equation

We first try to find a solution in the case where the Hamiltonian $\mathcal{H}=\frac{\hat{p}^2}{2m}+V(x,t)$ is such that the potential $V(x,t)$ is time independent (we can then write $V(x)$). In this case we can use separation of variables to look for solutions. That is, we look for solutions that are a product of a function of position only and a function of time only:

$$\psi(x,t)=\varphi(x)f(t)$$

Then, when we take the partial derivatives we have that

$$\frac{\partial\psi(x,t)}{\partial t}=\frac{df(t)}{dt}\varphi(x),\quad\frac{\partial\psi(x,t)}{\partial x}=\frac{d\varphi(x)}{dx}f(t)\text{ and }\frac{\partial^2\psi(x,t)}{\partial x^2}=\frac{d^2\varphi(x)}{dx^2}f(t)$$

The Schrödinger equation simplifies to

$$i\hbar\frac{df(t)}{dt}\varphi(x)=-\frac{\hbar^2}{2m}\frac{d^2\varphi(x)}{x^2}f(t)+V(x)\varphi(x)f(t)$$

Dividing by $\psi(x,t)$ we have:

$$i\hbar\frac{df(t)}{dt}\frac{1}{f(t)}=-\frac{\hbar^2}{2m}\frac{d^2\varphi(x)}{x^2}\frac{1}{\varphi(x)}+V(x)$$

Now the LHS is a function of time only, while the RHS is a function of position only. For the equation to hold, both sides have then to be equal to a constant (separation constant):

$$i\hbar\frac{df(t)}{dt}\frac{1}{f(t)}=E,-\frac{\hbar^2}{2m}\frac{d^2\varphi(x)}{x^2}\frac{1}{\varphi(x)}+V(x)=E$$

The two equations we find are a simple equation in the time variable:

$$\frac{df(t)}{dt}=-\frac{i}{\hbar}Ef(t),\ \rightarrow\ f(t)=f(0)e^{-i\frac{Et}{\hbar}}$$

and

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi(x)}{x^2}\frac{1}{\varphi(x)}+V(x)=E$$

that we have already seen as the time-independent Schrödinger equation. We have extensively studied the solutions of the this last equation, as they are the eigenfunctions of the energy-eigenvalue problem, giving the stationary (equilibrium) states of quantum systems. Note that for these stationary solutions $\varphi(x)$ we can still find the corresponding total wavefunction, given as stated above by $\psi(x,t) = \varphi(x)f(t)$, which does describe also the time evolution of the system:

$$\boxed{\psi(x,t) = \varphi(x)e^{-i\frac{Et}{\hbar}}}$$

Does this mean that the states that up to now we called stationary are instead evolving in time?

The answer is yes, but with a caveat. Although the states themselves evolve as stated above, any measurable quantity (such as the probability density $|\psi(x,t)|^2$ or the expectation values of observable, $\langle A\rangle = \int \psi(x,t)^* A[\psi(x,t)]$) are still time-independent. (Check it!)

Thus we were correct in calling these states **stationary** and neglecting in practice their time-evolution when studying the properties of systems they describe.

Notice that the wavefunction built from one energy eigenfunction, $\psi(x,t) = \varphi(x)f(t)$, is only a particular solution of the Schrödinger equation, but many other are possible. These will be complicated functions of space and time, whose shape will depend on the particular form of the potential $V(x)$. How can we describe these general solutions? We know that in general we can write a basis given by the eigenfunction of the Hamiltonian. These are the functions $\{\varphi(x)\}$ (as defined above by the time-independent Schrödinger equation). The eigenstate of the Hamiltonian do not evolve. However we can write any wavefunction as

$$\psi(x,t) = \sum_k c_k(t)\varphi_k(x)$$

This just corresponds to express the wavefunction in the *basis* given by the energy eigenfunctions. As usual, the coefficients $c_k(t)$ can be obtained at any instant in time by taking the inner product: $\langle \varphi_k \mid \psi(x,t)\rangle$.

What is the evolution of such a function? Substituting in the Schrödinger equation we have

$$i\hbar \frac{\partial\left(\sum_k c_k(t)\varphi_k(x)\right)}{\partial t} = \sum_k c_k(t)\mathcal{H}\varphi_k(x)$$

that becomes

$$i\hbar \sum_k \frac{\partial\left(c_k(t)\right)}{\partial t}\varphi_k(x) = \sum_k c_k(t)E_k\varphi_k(x)$$

For each φ_k we then have the equation in the coefficients only

$$i\hbar\frac{dc_k}{dt} = E_k c_k(t) \rightarrow c_k(t) = c_k(0)e^{-i\frac{E_k t}{\hbar}}$$

A general solution of the Schrödinger equation is then

$$\psi(x,t) = \sum_k c_k(0)e^{-i\frac{E_k t}{\hbar}}\varphi_k(x)$$

Obs.

We can define the eigen-frequencies $\hbar\omega_k = E_k$ from the eigen-energies. Thus we see that the wavefunction is a superposition of waves φ_k propagating in time each with a different frequency ω_k.

The behavior of quantum systems –even particles– thus often is similar to the propagation of waves. One example is the diffraction pattern for electrons (and even heavier objects) when scattering from a slit. We saw an example in the electron diffraction video at the beginning of the class.

Obs.

What is the probability of measuring a certain energy E_k at a time t? It is given by the coefficient of the φ_k eigenfunction, $|c_k(t)|^2 = \left|c_k(0)e^{-i\frac{E_k t}{\hbar}}\right|^2 = |c_k(0)|^2$. This means that the probability for the given energy is constant, does not change in time. Energy is then a so-called constant of the motion. This is true only for the energy eigenvalues, not for other observables'.

Example 6.1.1

Consider instead the probability of finding the system at a certain position, $p(x) = |\psi(x,t)|^2$. This of course changes in time. For example, let

$$\psi(x,0) = c_1(0)\varphi_1(x) + c_2(0)\varphi_2(x),$$

with

$$|c_1(0)|^2 + |c_2(0)|^2 = |c_1|^2 + |c_2|^2 = 1$$

and $\varphi_{1,2}$ normalized energy eigenfunctions. Then at a later time we have

$$\psi(x,0) = c_1(0)e^{-i\omega_1 t}\varphi_1(x) + c_2(0)e^{-i\omega_2 t}\varphi_2(x).$$

What is $p(x,t)$?

Solution

$$\begin{aligned}
&\left|c_1(0)e^{-i\omega_1 t}\varphi_1(x) + c_2(0)e^{-i\omega_2 t}\varphi_2(x)\right|^2 \\
&= |c_1(0)|^2|\varphi_1(x)|^2 + |c_2(0)|^2|\varphi_2(x)|^2 + c_1^* c_2 \varphi_1^* \varphi_2 e^{-i(\omega_2-\omega_1)t} + c_1 c_2^* \varphi_1 \varphi_2^* e^{i(\omega_2-\omega_1)t} \\
&= |c_1|^2 + |c_2|^2 + 2\,\mathrm{Re}\left[c_1^* c_2 \varphi_1^* \varphi_2 e^{-i(\omega_2-\omega_1)t}\right]
\end{aligned}$$

The last term describes a wave interference between different components of the initial wavefunction.

Obs.

The expressions found above for the time-dependent wavefunction are only valid if the potential is itself time-independent. If this is not the case, the solutions are even more difficult to obtain.

Unitary Evolution

We saw two equivalent formulation of the quantum mechanical evolution, the Schrödinger equation and the Heisenberg equation. We now present a third possible formulation: following the 4$^{\text{th}}$ postulate we express the evolution of a state in terms of a unitary operator, called the **propagator**:

$$\psi(x,t) = \hat{U}(t)\psi(x,0)$$

with $\hat{U}^\dagger \hat{U} = 1$. (Notice that a priori the unitary operator $\hat{U}$ could also be a function of space). We can show that this is equivalent to the Schrödinger equation, by verifying that $\psi(x,t)$ above is a solution:

$$i\hbar \frac{\partial \hat{U}\psi(x,0)}{\partial t} = \mathcal{H}\hat{U}\psi(x,0) \quad \rightarrow \quad i\hbar \frac{\partial \hat{U}}{\partial t} = \mathcal{H}\hat{U}$$

where in the second step we used the fact that since the equation holds for any wavefunction ψ it must hold for the operator themselves. If the Hamiltonian is time independent, the second equation can be solved easily, obtaining:

$$i\hbar \frac{\partial \hat{U}}{\partial t} = \mathcal{H}\hat{U} \quad \rightarrow \quad \hat{U}(t) = e^{-i\mathcal{H}t/\hbar}$$

where we set $\hat{U}(t=0) = 1$. Notice that as desired $\hat{U}$ is unitary, $\hat{U}^\dagger \hat{U} = e^{i\mathcal{H}t/\hbar} e^{-i\mathcal{H}t/\hbar} = 1$.