PART I   THEORY

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 particle's **wave function**, $\Psi(x, t)$, and we get it by solving the **Schrödinger equation**:

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

---

[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 ($v \ll c$).
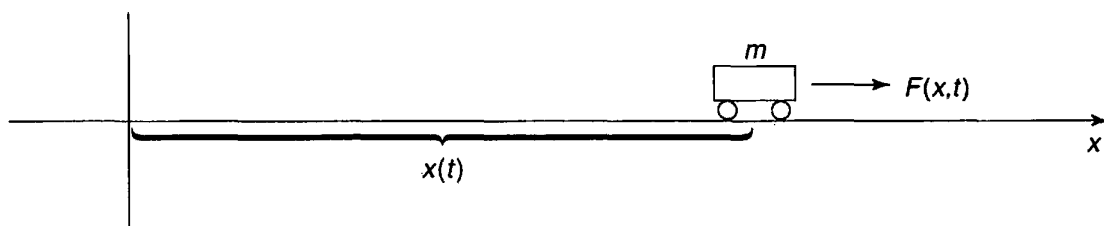
1

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:

$$\hbar = \frac{h}{2\pi} = 1.054572 \times 10^{-34} \text{J s}. \qquad [1.2]$$

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.[2]

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 represent 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,[3]

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

Probability is the *area* under the graph of $|\Psi|^2$. For the wave function in Figure 1.2, you would be quite likely to find the particle in the vicinity of point A, where $|\Psi|^2$ is large, and relatively *un*likely to find it near point B.

---

[2]For a delightful first-hand account of the origins of the Schrödinger equation see the article by Felix Bloch in *Physics Today*, December 1976.

[3]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.
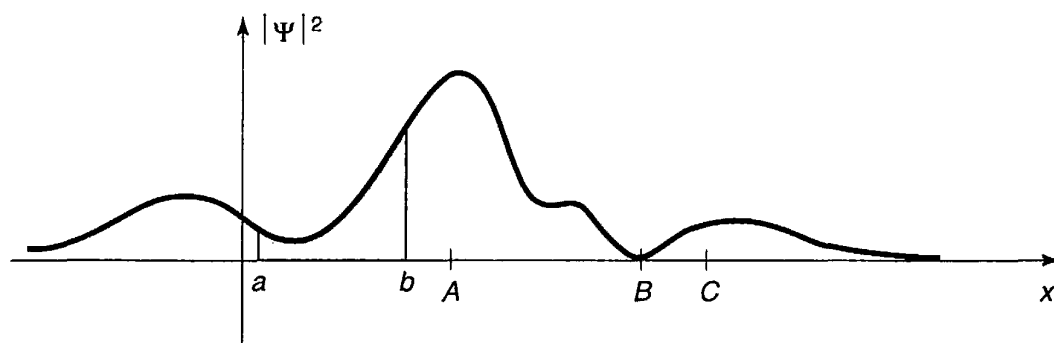
FIGURE 1.2:   A typical wave function. The shaded area represents the probability of finding the particle between *a* and *b*. The particle would be relatively likely to be found near *A*, and unlikely to be found near *B*.

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 particle (to wit: its wave function), still 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, and it is natural to wonder whether it is a fact of nature, or a defect in the theory.

Suppose I *do* measure the position of the particle, and I find it to be at point C.[4] *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."[5] 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

---

[4]Of course, no measuring instrument is perfectly precise: what I *mean* is that the particle was found *in the vicinity* of C, to within the tolerance of the equipment.

[5]Bernard d'Espagnat, "The Quantum Theory and Reality" (Scientific American, November 1979, p. 165).

particle) to assume a definite position."[6] This view (the so-called **Copenhagen interpretation**), is associated with Bohr and his followers. Among physicists it has always been the 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 pejorative 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."[7] For decades this was the "fall-back" position of most physicists: They'd try to sell you the orthodox answer, but if you were persistent they'd retreat to the agnostic response, 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 whether the particle had a precise (though unknown) position prior to the measurement, or not. 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 argument; for now, suffice it to say that the experiments have decisively confirmed the orthodox interpretation:[8] 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.

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

---

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

[7]Quoted by Mermin (footnote 6), p. 40.

[8]This statement is a little too strong: There remain a few theoretical and experimental loopholes, some of which I shall discuss in the Afterword. There exist viable nonlocal hidden variable theories (notably David Bohm's), and 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.
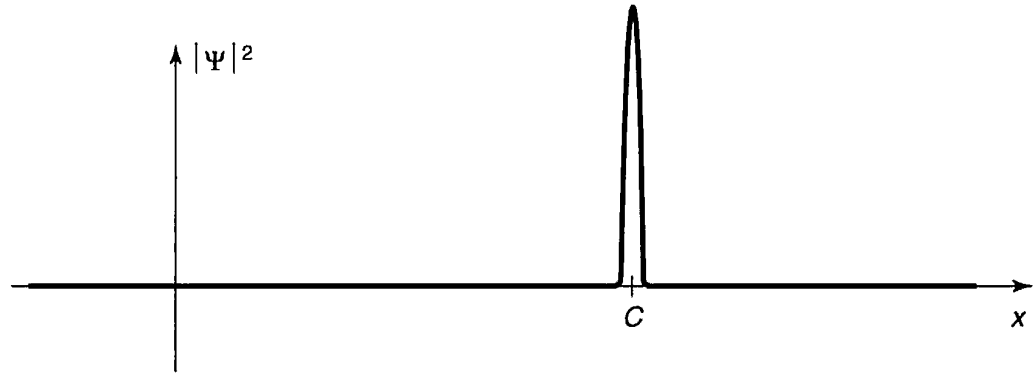
FIGURE 1.3: Collapse of the wave function: graph of $|\Psi|^2$ immediately *after* a measurement has found the particle at point C.

interpretation account for the fact that the second measurement is bound to yield 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 (it soon spreads out again, in accordance with the Schrödinger equation, so the second measurement must be made quickly). There 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.[9]

1.3 PROBABILITY

1.3.1 Discrete Variables

Because of the statistical interpretation, probability plays a central role in quantum mechanics, so I digress now for a brief discussion of probability theory. 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 fourteen people, whose ages are as follows:

one person aged 14,

one person aged 15,

three people aged 16,

---

[9]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.

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

$$N(14) = 1,$$
$$N(15) = 1,$$
$$N(16) = 3,$$
$$N(22) = 2,$$
$$N(24) = 2,$$
$$N(25) = 5,$$

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

$$N = \sum_{j=0}^{\infty} N(j). \tag{1.4}$$

(In the example, 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 that 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,
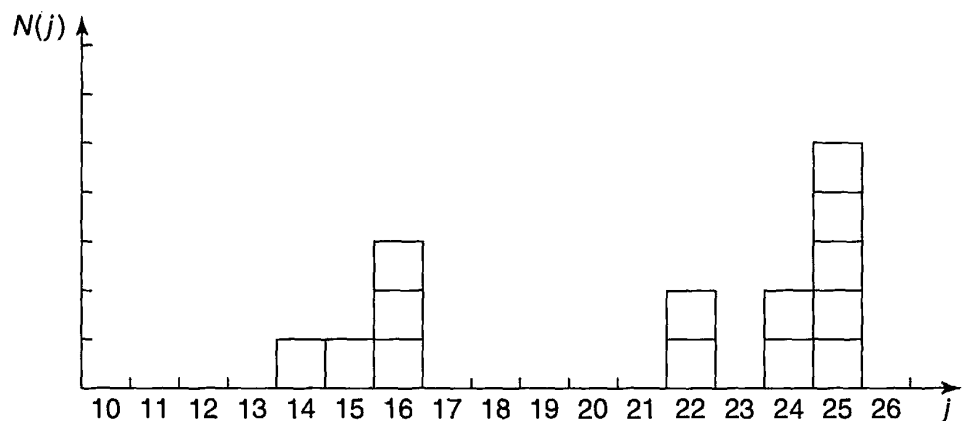
$$P(j) = \frac{N(j)}{N}. \tag{1.5}$$

FIGURE 1.4: Histogram showing the number of people, N(j), with age j, for the distribution in Section 1.3.1.

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:

$$\sum_{j=0}^{\infty} P(j) = 1. \qquad [1.6]$$

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

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

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

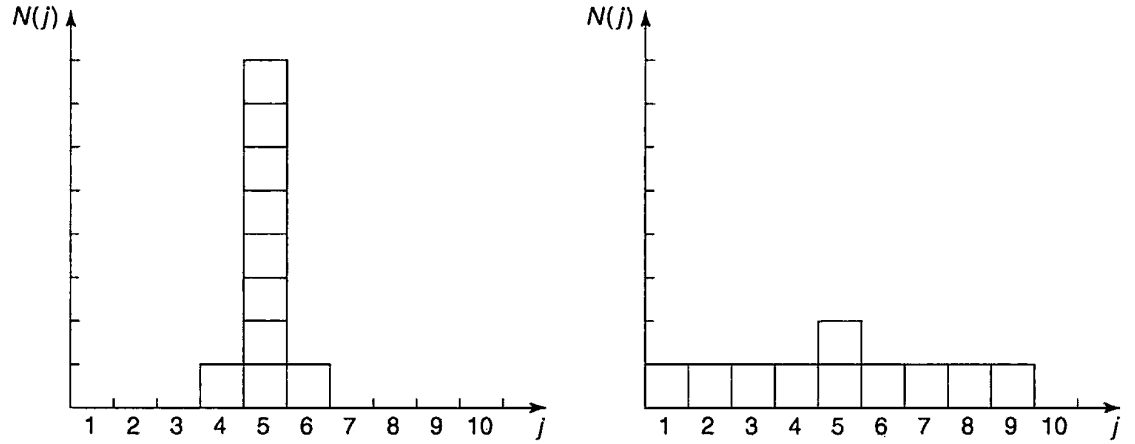$$\langle f(j) \rangle = \sum_{j=0}^{\infty} f(j) P(j). \qquad [1.9]$$

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

(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* equal, in general, 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, the second, perhaps, a rural one-room school-house.) We need a numerical measure of the amount of "spread" in a 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:

$$\langle \Delta j \rangle = \sum (j - \langle j \rangle) P(j) = \sum j P(j) - \langle j \rangle \sum P(j)$$
$$= \langle j \rangle - \langle j \rangle = 0.$$

(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 on variances:

$$\sigma^2 = \langle (\Delta j)^2 \rangle = \sum (\Delta j)^2 P(j) = \sum (j - \langle j \rangle)^2 P(j)$$

$$= \sum (j^2 - 2j\langle j \rangle + \langle j \rangle^2) P(j)$$

$$= \sum j^2 P(j) - 2\langle j \rangle \sum j P(j) + \langle j \rangle^2 \sum P(j)$$

$$= \langle j^2 \rangle - 2\langle j \rangle \langle j \rangle + \langle j \rangle^2 = \langle j^2 \rangle - \langle j \rangle^2.$$

Taking the square root, the standard deviation itself can be written as

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

In practice, this is a much faster way to get σ: Simply calculate $\langle j^2 \rangle$ and $\langle j \rangle^2$, subtract, and take the square root. 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

$$\langle j^2 \rangle \geq \langle j \rangle^2, \qquad [1.13]$$

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).

1.3.2 Continuous Variables

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 3.333 ... seconds is *zero*. The only sensible thing to speak about is the probability that her age lies in some *interval*—say, between 16 and 17. 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 that day—in which case we have simply 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

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

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

$$P_{ab} = \int_a^b \rho(x)\,dx. \qquad [1.15]$$

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

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

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

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

$$\sigma^2 \equiv \langle (\Delta x)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2. \qquad [1.19]$$

Example 1.1   Suppose I drop a rock off a cliff of height h. As it falls, I snap a million photographs, at random intervals. On each picture I measure the distance the rock has fallen. *Question:* What is the *average* of all these distances? That is to say, what is the *time average* of the distance traveled?[10]

Solution: The rock starts out at rest, and picks up speed as it falls; it spends more time near the top, so the average distance must be less than $h/2$. Ignoring air resistance, the distance x at time t is

$$x(t) = \frac{1}{2}gt^2.$$

The velocity is $dx/dt = gt$, and the total flight time is $T = \sqrt{2h/g}$. The probability that the camera flashes in the interval dt is dt/T, so the probability that a given

---

[10]A statistician will complain that I am confusing the average of a *finite sample* (a million, in this case) with the "true" average (over the whole continuum). This can be an awkward problem for the experimentalist, especially when the sample size is small, but here I am only concerned, of course, with the true average, to which the sample average is presumably a good approximation.
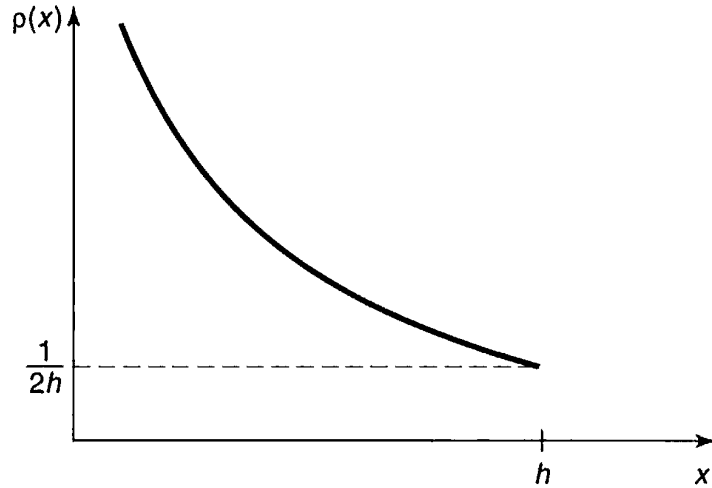
FIGURE 1.6: The probability density in Example 1.1: $p(x) = 1/(2\sqrt{hx})$.

photograph shows a distance in the corresponding range dx is

$$\frac{dt}{T} = \frac{dx}{gt}\sqrt{\frac{g}{2h}} = \frac{1}{2\sqrt{hx}}dx.$$

Evidently the probability *density* (Equation 1.14) is

$$p(x) = \frac{1}{2\sqrt{hx}}, \quad (0 \le x \le h)$$

(outside this range, of course, the probability density is zero).
We can check this result, using Equation 1.16:

$$\int_0^h \frac{1}{2\sqrt{hx}}dx = \frac{1}{2\sqrt{h}}\left(2x^{1/2}\right)\Big|_0^h = 1.$$

The *average* distance (Equation 1.17) is

$$\langle x \rangle = \int_0^h x\frac{1}{2\sqrt{hx}}dx = \frac{1}{2\sqrt{h}}\left(\frac{2}{3}x^{3/2}\right)\Big|_0^h = \frac{h}{3},$$

which is somewhat less than $h/2$, as anticipated.

Figure 1.6 shows the graph of $p(x)$. Notice that a probability *density* can be infinite, though probability itself (the *integral* of p) must of course be finite (indeed, less than or equal to 1).

*Problem 1.1 For the distribution of ages in Section 1.3.1:

(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

(a) Find the standard deviation of the distribution in Example 1.1.

(b) What is the probability that a photograph, selected at random, would show a distance x more than one standard deviation away from the average?

*Problem 1.3 Consider the **gaussian** distribution

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

where A, a, and λ are positive real 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):

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

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 go imposing an extraneous condition on Ψ without checking that the two are consistent. Well, 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.[11]

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 remarkable 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.

This is important, so we'd better pause for a careful proof. To begin with,

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

(Note that the *integral* is a function only of t, so I use a *total* derivative (d/dt) in the first expression, 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

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

---

[11]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.

The integral in Equation 1.21 can now be evaluated explicitly:

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

But $\Psi(x,t)$ must go to zero as x goes to ($\pm$) infinity—otherwise the wave function would not be normalizable.[12] It follows that

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

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

Problem 1.4 At time $t = 0$ a particle is represented by the wave function

$$\Psi(x, 0) = \begin{cases} A\dfrac{x}{a}, & \text{if } 0 \le x \le a, \\[2mm] A\dfrac{(b-x)}{(b-a)}, & \text{if } a \le x \le b, \\[2mm] 0, & \text{otherwise,} \end{cases}$$

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 1.5** Consider the wave function

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

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.

---

[12] A good mathematician can supply you with pathological counterexamples, but they do not arise in physics; for us the wave function *always* goes to zero at infinity.

(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?

1.5 MOMENTUM

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

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

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 have to 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[13]

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

---

[13]To keep things from getting too cluttered. I'll suppress the limits of integration.

This expression can be simplified using integration-by-parts:[14]

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

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

$$\langle v \rangle = \frac{d\langle x \rangle}{dt}.$$ [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:

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

---

[14]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.

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

$$\langle x \rangle = \int \Psi^*(x)\Psi\, dx, \qquad\qquad [1.34]$$

$$\langle p \rangle = \int \Psi^* \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right) \Psi\, dx. \qquad\qquad [1.35]$$

We say that the **operator**[15] 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 quantities? The fact is, *all* classical dynamical variables can be expressed 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 *any* such quantity, $Q(x, p)$, we simply replace every p by $(\hbar/i)(\partial/\partial x)$, insert the resulting operator between Ψ^* and Ψ, and integrate:

$$\langle Q(x, p) \rangle = \int \Psi^* Q \left(x, \frac{\hbar}{i} \frac{\partial}{\partial x} \right) \Psi\, dx. \qquad\qquad [1.36]$$

For example, the expectation value of the kinetic energy is

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

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 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.

---

[15]An "operator" is an instruction to *do something* to the function that follows it. 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.

Problem 1.6 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 1.7 Calculate $d\langle p\rangle/dt$. *Answer*:

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

Equations 1.32 (or the first part of 1.33) and 1.38 are instances of **Ehrenfest's theorem**, which tells us that *expectation values obey classical laws*.

Problem 1.8 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 1.7). 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 1.8), 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? 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.[16] A theorem in Fourier analysis makes all this rigorous, but for the moment I am only concerned with the qualitative argument.

---

[16]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.
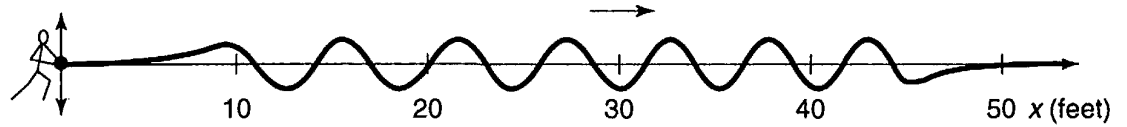
FIGURE 1.7: A wave with a (fairly) well-defined *wavelength*, but an ill-defined *position*.
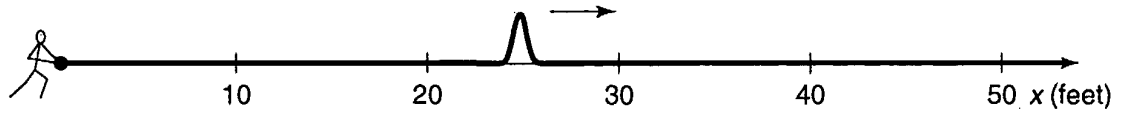

FIGURE 1.8: A wave with a (fairly) well-defined *position*, but an ill-defined *wavelength*.

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**:[17]

$$p = \frac{h}{\lambda} = \frac{2\pi\hbar}{\lambda}.$$ [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 is its momentum. Quantitatively,

$$\boxed{\sigma_x \sigma_p \geq \frac{\hbar}{2},}$$ [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 right away, so you can test it out on the examples in Chapter 2.)

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 identically prepared systems do not yield identical results. You can, if you want, construct a state 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 state with a reproducible momentum (by making

---

[17] 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.

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 create a state for 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 1.9 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?

FURTHER PROBLEMS FOR CHAPTER 1

Problem 1.10 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 1.11 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, $\rho(\theta)$? *Hint:* $\rho(\theta) d\theta$ 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 1.12 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 $\rho(x)$? 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:* $\rho(x)\,dx$ is the probability that the projection lies between x and $(x + dx)$. You know (from Problem 1.11) 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 1.11.

∗∗Problem 1.13 Buffon's needle. 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 1.12.

Problem 1.14 Let $P_{ab}(t)$ be the probability of finding a 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 $J(x,t)$? *Comment:* 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 Problem 1.9. (This is not a very pithy example, I'm afraid; we'll encounter more substantial ones in due course.)

∗∗Problem 1.15 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.

Problem 1.16 Show that

$$\frac{d}{dt} \int_{-\infty}^{\infty} \Psi_1^* \, \Psi_2 \, dx = 0$$

for any two (normalizable) solutions to the Schrödinger equation, Ψ_1 and Ψ_2.

Problem 1.17 A particle is represented (at time $t = 0$) by the wave function

$$\Psi(x, 0) = \begin{cases} A(a^2 - x^2), & \text{if } -a \leq x \leq +a, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Determine the normalization constant A.

(b) What is the expectation value of x (at time $t = 0$)?

(c) What is the expectation value of p (at time $t = 0$)? (Note that you *cannot* get it from $p = md\langle x \rangle/dt$. Why not?)

(d) Find the expectation value of x^2.

(e) Find the expectation value of p^2.

(f) Find the uncertainty in x (σ_x).

(g) Find the uncertainty in p (σ_p).

(h) Check that your results are consistent with the uncertainty principle.

Problem 1.18 In general, quantum mechanics is relevant when the de Broglie wavelength of the particle in question (h/p) is greater than the characteristic size of the system (d). In thermal equilibrium at (Kelvin) temperature T, the average kinetic energy of a particle is

$$\frac{p^2}{2m} = \frac{3}{2}k_B T$$

(where k_B is Boltzmann's constant), so the typical de Broglie wavelength is

$$\lambda = \frac{h}{\sqrt{3mk_B T}}.$$ [1.41]

The purpose of this problem is to anticipate which systems will have to be treated quantum mechanically, and which can safely be described classically.

(a) **Solids.** The lattice spacing in a typical solid is around $d = 0.3$ nm. Find the temperature below which the free[18] *electrons* in a solid are quantum mechanical. Below what temperature are the *nuclei* in a solid quantum mechanical? (Use sodium as a typical case.) *Moral:* The free electrons in a solid are *always* quantum mechanical; the nuclei are almost *never* quantum mechanical. The same goes for liquids (for which the interatomic spacing is roughly the same), with the exception of helium below 4 K.

(b) **Gases.** For what temperatures are the atoms in an ideal gas at pressure P quantum mechanical? *Hint:* Use the ideal gas law $(PV = Nk_B T)$ to deduce the interatomic spacing. *Answer:* $T < (1/k_B)(h^2/3m)^{3/5}P^{2/5}$. Obviously (for the gas to show quantum behavior) we want m to be as *small* as possible, and P as *large* as possible. Put in the numbers for helium at atmospheric pressure. Is hydrogen in outer space (where the interatomic spacing is about 1 cm and the temperature is 3 K) quantum mechanical?

---

[18] In a solid the inner electrons are attached to a particular nucleus, and for them the relevant size would be the radius of the atom. But the outermost electrons are not attached, and for them the relevant distance is the lattice spacing. This problem pertains to the *outer* electrons.