3.1 Hilbert Space

In the previous two chapters we have stumbled on a number of interesting properties of simple quantum systems. Some of these are "accidental" features of specific potentials (the even spacing of energy levels for the harmonic oscillator, for example), but others seem to be more general, and it would be nice to prove them once and for all (the uncertainty principle, for instance, and the orthogonality of stationary states). The purpose of this chapter is to recast the theory in more powerful form, with that in mind. There is not much here that is genuinely *new*; the idea, rather, is to make coherent sense of what we have already discovered in particular cases.

Quantum theory is based on two constructs: wave functions and operators. The **state** of a system is represented by its wave function, **observables** are represented by operators. Mathematically, wave functions satisfy the defining conditions for abstract **vectors**, and operators act on them as **linear transformations**. So the natural language of quantum mechanics is **linear algebra**.[1]

But it is not, I suspect, a form of linear algebra with which you may be familiar. In an N-dimensional space it is simplest to represent a vector, $|\alpha\rangle$, by the N-tuple of its components, $\{a_n\}$, with respect to a specified orthonormal basis:

$$|\alpha\rangle \to \mathsf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}; \tag{3.1}$$

the **inner product**, $\langle\alpha|\beta\rangle$, of two vectors (generalizing the dot product in three dimensions) is a complex number,

$$\langle\alpha|\beta\rangle = a_1^* b_1 + a_2^* b_2 + \cdots + a_N^* b_N; \tag{3.2}$$

linear transformations, T, are represented by **matrices** (with respect to the specified basis), which act on vectors (to produce new vectors) by the ordinary rules of matrix multiplication:

$$|\beta\rangle = \hat{T}|\alpha\rangle \to \mathsf{b} = \mathsf{T}\mathsf{a} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1N} \\ t_{21} & t_{22} & \cdots & t_{2N} \\ \vdots & \vdots & & \vdots \\ t_{N1} & t_{N2} & \cdots & t_{NN} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}. \tag{3.3}$$

But the "vectors" we encounter in quantum mechanics are (for the most part) *functions*, and they live in *infinite*-dimensional spaces. For them the N-tuple/matrix notation is awkward, at best, and manipulations that are well behaved in the finite-dimensional case can be problematic. (The underlying reason is that whereas the *finite* sum in Equation 3.2 always exists, an *infinite* sum—or an integral—may not converge, in which case the inner product does not exist, and any argument involving inner products is immediately suspect.) So even though most of the terminology and notation should be familiar, it pays to approach this subject with caution.

The collection of *all* functions of x constitutes a vector space, but for our purposes it is much too large. To represent a possible physical state, the wave function Ψ must be *normalized*:

$$\int |\Psi|^2 \, dx = 1.$$

The set of all **square-integrable** functions, on a specified interval,[2]

$$f(x) \quad \text{such that} \quad \int_a^b |f(x)|^2 \, dx < \infty, \tag{3.4}$$

constitutes a (much smaller) vector space (see Problem 3.1(a)). Mathematicians call it $L^2(a, b)$; physicists call it **Hilbert space**.[3] In quantum mechanics, then:

Wave functions live in Hilbert space. (3.5)

We define the **inner product of two functions**, $f(x)$ and $g(x)$, as follows:

$$\langle f|g\rangle \equiv \int_a^b f(x)^* g(x) \, dx. \tag{3.6}$$

If f and g are both square-integrable (that is, if they are both in Hilbert space), their inner product is guaranteed to exist (the integral in Equation 3.6 converges to a finite number).[4] This follows from the integral **Schwarz inequality**:[5]

$$\left| \int_a^b f(x)^* g(x) \, dx \right| \leq \sqrt{\int_a^b |f(x)|^2 \, dx \int_a^b |g(x)|^2 \, dx}. \tag{3.7}$$

You can check for yourself that definition (Equation 3.6) satisfies all the conditions for an inner product (Problem 3.1(b)). Notice in particular that

$$\langle g|f\rangle = \langle f|g\rangle^*. \tag{3.8}$$

Moreover, the inner product of $f(x)$ with *itself*,

$$\langle f|f\rangle = \int_a^b |f(x)|^2 dx, \tag{3.9}$$

is *real* and non-negative; it's *zero* only when $f(x) = 0$.[6]

A function is said to be **normalized** if its inner product with itself is 1; two functions are **orthogonal** if their inner product is 0; and a *set* of functions, $\{f_n\}$, is **orthonormal** if they are normalized and mutually orthogonal:

$$\langle f_m|f_n\rangle = \delta_{mn}. \tag{3.10}$$

Finally, a set of functions is **complete** if any *other* function (in Hilbert space) can be expressed as a linear combination of them:

$$f(x) = \sum_{n=1}^{\infty} c_n f_n(x). \tag{3.11}$$

If the functions $\{f_n(x)\}$ are orthonormal, the coefficients are given by Fourier's trick:

$$c_n = \langle f_n|f\rangle,$$

as you can check for yourself. I anticipated this terminology, of course, back in Chapter 2. (The (3.12) stationary states of the infinite square well (Equation 2.31) constitute a complete orthonormal set on the interval $(0, a)$; the stationary states for the harmonic oscillator (Equation 2.68 or 2.86) are a complete orthonormal set on the interval $(-\infty, \infty)$.)

Problem 3.1

(a) Show that the set of all square-integrable functions is a vector space (refer to Section A.1 for the definition). *Hint:* The main point is to show that the sum of two square-integrable functions is itself square-integrable. Use Equation 3.7. Is the set of all *normalized* functions a vector space?

(b) Show that the integral in Equation 3.6 satisfies the conditions for an inner product (Section A.2).

* **Problem 3.2**

(a) For what range of ν is the function $f(x) = x^{\nu}$ in Hilbert space, on the interval $(0, 1)$? Assume ν is real, but not necessarily positive.

(b) For the specific case $\nu = 1/2$, is $f(x)$ in this Hilbert space? What about $xf(x)$? How about $(d/dx)\, f(x)$?

3.2 Observables

3.2.1 Hermitian Operators

The expectation value of an observable $Q(x, p)$ can be expressed very neatly in inner-product notation:[7]

$$\langle Q \rangle = \int \Psi^* \hat{Q} \Psi dx = \langle \Psi | \hat{Q} \Psi \rangle. \tag{3.13}$$

Now, the outcome of a measurement has got to be *real*, and so, *a fortiori*, is the *average* of many measurements:

$$\langle Q \rangle = \langle Q \rangle^*. \tag{3.14}$$

But the complex conjugate of an inner product reverses the order (Equation 3.8), so

$$\langle \Psi | \hat{Q} \Psi \rangle = \langle \hat{Q} \Psi | \Psi \rangle, \tag{3.15}$$

and this must hold true for any wave function Ψ. Thus operators representing *observables* have the very special property that

$$\langle f | \hat{Q} f \rangle = \langle \hat{Q} f | f \rangle \quad \text{for all } f(x). \tag{3.16}$$

We call such operators **hermitian.**[8]

Actually, most books require an ostensibly stronger condition:

$$\langle f | \hat{Q} g \rangle = \langle \hat{Q} f | g \rangle \quad \text{for all } f(x) \text{ and all } g(x). \tag{3.17}$$

But it turns out, in spite of appearances, that this is perfectly equivalent to my definition (Equation 3.16), as you will prove in Problem 3.3. So use whichever you like. The essential point is that a hermitian operator can be applied either to the first member of an inner product or to the second, with the same result, and hermitian operators naturally arise in quantum mechanics because their expectation values are real:

Observables are represented by hermitian operators. (3.18)

Well, let's *check* this. Is the momentum operator, for example, hermitian?

$$\langle f | \hat{p} g \rangle = \int_{-\infty}^{\infty} f^* (-i\hbar) \frac{dg}{dx} dx = -i\hbar f^* g \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \left(-i\hbar \frac{df}{dx} \right)^* g dx = \langle \hat{p} f | g \rangle. \tag{3.19}$$

I used integration by parts, of course, and threw away the boundary term for the usual reason: If $f(x)$ and $g(x)$ are square integrable, they must go to zero at $\pm\infty$.[9] Notice how the complex conjugation of i compensates for the minus sign picked up from integration by parts—the operator d/dx (without the i) is *not* hermitian, and it does not represent a possible observable.

The **hermitian conjugate** (or **adjoint**) of an operator $\hat{Q}$ is the operator $\hat{Q}^\dagger$ such that

$$\langle f | \hat{Q} g \rangle = \langle \hat{Q}^\dagger f | g \rangle \quad (\text{for all } f \text{ and } g). \tag{3.20}$$

A hermitian operator, then, is equal to its hermitian conjugate: $\hat{Q} = \hat{Q}^\dagger$.

* **Problem 3.3** Show that if $\langle h|\hat{Q}h\rangle = \langle \hat{Q}h|h\rangle$ for all h (in Hilbert space), then $\langle f|\hat{Q}g\rangle = \langle \hat{Q}f|g\rangle$ for all f and g (i.e. the two definitions of "hermitian" — Equations 3.16 and 3.17—are equivalent). *Hint:* First let $h = f + g$, and then let $h = f + ig$.

Problem 3.4

(a) Show that the *sum* of two hermitian operators is hermitian.

(b) Suppose $\hat{Q}$ is hermitian, and $\mathtt{a}$ is a complex number. Under what condition (on $\mathtt{a}$) is $\alpha\hat{Q}$ hermitian?

(c) When is the *product* of two hermitian operators hermitian?

(d) Show that the position operator $(\hat{x})$ and the Hamiltonian operator $(\hat{H} = -(\hbar^2/2m)d^2/dx^2 + V(x))$ are hermitian.

* **Problem 3.5**

(a) Find the hermitian conjugates of x, i, and d/dx.

(b) Show that $(\hat{Q}\hat{R})^\dagger = \hat{R}^\dagger\hat{Q}^\dagger$ (note the reversed order), $(\hat{Q} + \hat{R}) = \hat{Q}^\dagger + \hat{R}^\dagger$, and $(c\hat{Q})^\dagger = c^*\hat{Q}^\dagger$ for a complex number c.

(c) Construct the hermitian conjugate of a_+ (Equation 2.48).

3.2.2 Determinate States

Ordinarily, when you measure an observable Q on an ensemble of identically prepared systems, all in the same state Ψ, you do *not* get the same result each time—this is the *indeterminacy* of quantum mechanics. *Question:* Would it be possible to prepare a state such that *every* measurement of Q is certain to return the *same* value (call it q)? This would be, if you like, a **determinate state**, for the observable Q. (Actually, we already know one example: Stationary states are determinate states of the Hamiltonian; a measurement of the energy, on a particle in the stationary state Ψ_n, is certain to yield the corresponding "allowed" energy E_n.)

Well, the standard deviation of Q, in a determinate state, would be *zero*, which is to say,

$$\sigma^2 = \left\langle (Q - \langle Q \rangle)^2 \right\rangle = \left\langle \Psi \middle| (\hat{Q} - q)^2 \Psi \right\rangle = \langle (\hat{Q} - q)\Psi \,|\, (\hat{Q} - q)\Psi \rangle = 0. \tag{3.21}$$

(Of course, if every measurement gives q, their average is also q: $\langle Q \rangle = q$. I used the fact that $\hat{Q}$ (and hence also $\hat{Q} - q$) is a *hermitian* operator, to move one factor over to the first term in the inner product.) But the only vector whose inner product with itself vanishes is 0, so

$$\hat{Q}\Psi = q\Psi. \tag{3.22}$$

This is the **eigenvalue equation** for the operator $\hat{Q}$; Ψ is an **eigenfunction** of $\hat{Q}$, and q is the corresponding **eigenvalue**:

Determinate states of Q are eigenfunctions of $\hat{Q}$. (3.23)

Measurement of Q on such a state is certain to yield the eigenvalue, q.[10]

Note that the eigen*value* is a *number* (not an operator or a function). You can multiply any eigenfunction by a constant, and it is still an eigenfunction, with the same eigenvalue. Zero does not count as an eigenfunction (we exclude it by definition—otherwise *every* number would be an eigenvalue, since $\hat{Q}0 = q0 = 0$ for any linear operator $\hat{Q}$ and all q). But there's nothing wrong with zero as an eigen*value*. The collection of all the eigenvalues of an operator is called its **spectrum**. Sometimes two (or more) linearly independent eigenfunctions share the same eigenvalue; in that case the spectrum is said to be **degenerate**. (You encountered this term already, for the case of energy eigenstates, if you worked Problems 2.44 or 2.46.)

For example, determinate states of the total energy are eigenfunctions of the Hamiltonian:

$$\hat{H}\psi = E\psi, \tag{3.24}$$

which is precisely the time-independent Schrödinger equation. In this context we use the letter E for the eigenvalue, and the lower case ψ for the eigenfunction (tack on the wiggle factor $\exp\left(-iEt/\hbar\right)$ to make it Ψ, if you like; it's still an eigenfunction of $\hat{H}$).

Example 3.1

Consider the operator

$$\hat{Q} \equiv i\frac{d}{d\phi}, \tag{3.25}$$

where ϕ is the usual polar coordinate in two dimensions. (This operator might arise in a physical context if we were studying the bead-on-a-ring; see Problem 2.46.) Is $\hat{Q}$ hermitian? Find its eigenfunctions and eigenvalues.

Solution: Here we are working with functions $f(\phi)$ on the *finite* interval $0 \leq \phi \leq 2\pi$, with the property that

$$f(\phi + 2\pi) = f(\phi), \tag{3.26}$$

since ϕ and $\phi + 2\pi$ describe the same physical point. Using integration by parts,

$$\langle f|\hat{Q}g\rangle = \int_0^{2\pi} f^* \left(i\frac{dg}{d\phi}\right) d\phi = if^*g\Big|_0^{2\pi} - \int_0^{2\pi} i\left(\frac{df^*}{d\phi}\right) g\,d\phi = \langle \hat{Q}f|g\rangle,$$

so $\hat{Q}$ *is* hermitian (this time the boundary term disappears by virtue of Equation 3.26).

The eigenvalue equation,

$$i\frac{d}{d\phi}f(\phi) = qf(\phi), \tag{3.27}$$

has the general solution

$$f(\phi) = Ae^{-iq\phi}. \tag{3.28}$$

Equation 3.26 restricts the possible values of the q:

$$e^{-iq2\pi} = 1 \quad \Rightarrow \quad q = 0, \pm 1, \pm 2, \ldots. \tag{3.29}$$

The spectrum of this operator is the set of all integers, and it is nondegenerate.

Problem 3.6 Consider the operator $\hat{Q} = d^2/d\phi^2$, where (as in Example 3.1) ϕ is the azimuthal angle in polar coordinates, and the functions are subject to Equation 3.26. Is $\hat{Q}$ hermitian? Find its eigenfunctions and eigenvalues. What is the spectrum of $\hat{Q}$? Is the spectrum degenerate?

3.3 Eigenfunctions of a Hermitian Operator

Our attention is thus directed to the *eigenfunctions of hermitian operators* (physically: determinate states of observables). These fall into two categories: If the spectrum is **discrete** (i.e. the eigenvalues are separated from one another) then the eigenfunctions lie in Hilbert space and they constitute physically realizable states. If the spectrum is **continuous** (i.e. the eigenvalues fill out an entire range) then the eigenfunctions are not normalizable, and they do not represent possible wave functions (though *linear combinations* of them—involving necessarily a spread in eigenvalues—may be normalizable). Some operators have a discrete spectrum only (for example, the Hamiltonian for the harmonic oscillator), some have only a continuous spectrum (for example, the free particle Hamiltonian), and some have both a discrete part and a continuous part (for example, the Hamiltonian for a finite square well). The discrete case is easier to handle, because the relevant inner products are guaranteed to exist—in fact, it is very similar to the finite-dimensional theory (the eigenvectors of a hermitian *matrix*). I'll treat the discrete case first, and then the continuous one.

3.3.1 Discrete Spectra

Mathematically, the normalizable eigenfunctions of a hermitian operator have two important properties:

Theorem 1 Their eigen*values* are *real*.

Proof: Suppose

$$\hat{Q}f = qf,$$

(i.e. $f(x)$ is an eigenfunction of $\hat{Q}$, with eigenvalue q), and[11]

$$\langle f | \hat{Q}f \rangle = \langle \hat{Q}f | f \rangle$$

($\hat{Q}$ is hermitian). Then

$$q\langle f|f\rangle = q^*\langle f|f\rangle$$

(q is a *number*, so it comes outside the integral, and because the first function in the inner product is complex conjugated (Equation 3.6), so too is the q on the right). But $\langle f|f\rangle$ cannot be zero ($f(x) = 0$ is not a legal eigenfunction), so $q = q^*$, and hence q is real. QED

This is comforting: If you measure an observable on a particle in a determinate state, you will at least get a real number.

Theorem 2 Eigenfunctions belonging to distinct eigenvalues are *orthogonal*.

Proof: Suppose

$$\hat{Q}f = qf, \quad \text{and} \quad \hat{Q}g = q'g,$$

and $\hat{Q}$ is hermitian. Then $\langle f | \hat{Q}g \rangle = \langle \hat{Q}f | g \rangle$, so

$$q'\langle f|g\rangle = q^*\langle f|g\rangle$$

(again, the inner products exist because the eigenfunctions are in Hilbert space). But q is real (from Theorem 1), so if $q' \neq q$ it must be that $\langle f|g\rangle = 0$. QED

That's why the stationary states of the infinite square well, for example, or the harmonic oscillator, are orthogonal—they are eigenfunctions of the Hamiltonian with distinct eigenvalues. But this property is not peculiar to them, or even to the Hamiltonian—the same holds for determinate states of *any* observable.

Unfortunately, Theorem 2 tells us nothing about degenerate states ($q' = q$). However, if two (or more) eigenfunctions share the same eigenvalue, any linear combination of them is itself an eigenfunction, with the same eigenvalue (Problem 3.7), and we can use the **Gram–Schmidt orthogonalization procedure** (Problem A.4) to *construct* orthogonal eigenfunctions within each degenerate subspace. It is almost never necessary to do this explicitly (thank God!), but it can always be done in principle. So *even in the presence of degeneracy* the eigenfunctions can be *chosen* to be orthonormal, and we shall always assume that this has been done. That licenses the use of Fourier's trick, which depends on the orthonormality of the basis functions.

In a *finite*-dimensional vector space the eigenvectors of a hermitian matrix have a third fundamental property: They span the space (every vector can be expressed as a linear combination of them). Unfortunately,

the proof does not generalize to infinite-dimensional spaces. But the property itself is essential to the internal consistency of quantum mechanics, so (following Dirac[12]) we will take it as an *axiom* (or, more precisely, as a restriction on the class of hermitian operators that can represent observables):

> **Axiom:** The eigenfunctions of an observable operator are *complete*: Any function (in Hilbert space) can be expressed as a linear combination of them.[13]

Problem 3.7

(a) Suppose that $f(x)$ and $g(x)$ are two eigenfunctions of an operator $\hat{Q}$, with the same eigenvalue q. Show that any linear combination of f and g is itself an eigenfunction of $\hat{Q}$, with eigenvalue q.

(b) Check that $f(x) = \exp(x)$ and $g(x) = \exp(-x)$ are eigenfunctions of the operator d^2/dx^2, with the same eigenvalue. Construct two linear combinations of f and g that are *orthogonal* eigenfunctions on the interval $(-1, 1)$.

Problem 3.8

(a) Check that the eigenvalues of the hermitian operator in Example 3.1 are real. Show that the eigenfunctions (for distinct eigenvalues) are orthogonal.

(b) Do the same for the operator in Problem 3.6.

3.3.2 Continuous Spectra

If the spectrum of a hermitian operator is *continuous*, the eigenfunctions are not normalizable, and the proofs of Theorems 1 and 2 fail, because the inner products may not exist. Nevertheless, there is a sense in which the three essential properties (reality, orthogonality, and completeness) still hold. I think it's best to approach this case through specific examples.

Example 3.2

Find the eigenfunctions and eigenvalues of the momentum operator (on the interval $-\infty < x < \infty$).

Solution: Let $f_p(x)$ be the eigenfunction and p the eigenvalue:

$$-i\hbar\frac{d}{dx}f_p(x) = pf_p(x). \tag{3.30}$$

The general solution is

$$f_p(x) = Ae^{ipx/\hbar}.$$

This is not square-integrable for *any* (complex) value of p—the momentum operator has *no* eigenfunctions in Hilbert space.

And yet, if we restrict ourselves to *real* eigenvalues, we do recover a kind of *ersatz* "orthonormality." Referring to Problems 2.23(a) and 2.26,

$$\int_{-\infty}^{\infty} f_{p'}^*(x)f_p(x)dx = |A|^2\int_{-\infty}^{\infty} e^{i(p-p')x/\hbar}dx = |A|^2\,2\pi\hbar\delta(p-p'). \tag{3.31}$$

If we pick $A = 1/\sqrt{2\pi\hbar}$, so that

$$f_p(x) = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}, \tag{3.32}$$

then

$$\langle f_{p'}|f_p\rangle = \delta(p-p'), \tag{3.33}$$

which is reminiscent of *true* orthonormality (Equation 3.10)—the indices are now continuous variables, and the Kronecker delta has become a Dirac delta, but otherwise it looks just the same. I'll call Equation 3.33 **Dirac orthonormality**.

Most important, the eigenfunctions (with real eigenvalues) are *complete*, with the sum (in Equation 3.11) replaced by an integral: Any (square-integrable) function $f(x)$ can be written in the form

$$f(x) = \int_{-\infty}^{\infty} c(p)f_p(x)dp = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} c(p)e^{ipx/\hbar}dp. \tag{3.34}$$

The "coefficients" (now a *function*, $c(p)$) are obtained, as always, by Fourier's trick:

$$\langle f_{p'}|f\rangle = \int_{-\infty}^{\infty} c(p)\langle f_{p'}|f_p\rangle dp = \int_{-\infty}^{\infty} c(p)\,\delta(p-p')\,dp = c(p'). \tag{3.35}$$

Alternatively, you can get them from Plancherel's theorem (Equation 2.103); indeed, the expansion (Equation 3.34) is nothing but a Fourier transform.

The eigenfunctions of momentum (Equation 3.32) are sinusoidal, with wavelength

$$\lambda = \frac{2\pi\hbar}{p}. \tag{3.36}$$

This is the old de Broglie formula (Equation 1.39), which I promised to justify at the appropriate time. It turns out to be a little more subtle than de Broglie imagined, because we now know that there is actually *no such thing* as a particle with determinate momentum. But we could make a normalizable wave *packet* with a narrow range of momenta, and it is to such an object that the de Broglie relation applies.

What are we to make of Example 3.2? Although none of the eigenfunctions of $\hat{p}$ lives in Hilbert space, a certain family of them (those with real eigenvalues) resides in the nearby "suburbs," with a kind of quasi-normalizability. They do not represent possible physical states, but they are still very useful (as we have already seen, in our study of one-dimensional scattering).[14]

Example 3.3

Find the eigenfunctions and eigenvalues of the position operator.

Solution: Let $g_y(x)$ be the eigenfunction and y the eigenvalue:

$$\hat{x}\,g_y(x) = x g_y(x) = y g_y(x). \tag{3.37}$$

Here y is a fixed number (for any given eigenfunction), but x is a continuous variable. What function of x has the property that multiplying it by x is the same as multiplying it by the constant y? Obviously it's got to be *zero*, except at the one point $x = y$; in fact, it is nothing but the Dirac delta function:

$$g_y(x) = A\delta(x-y).$$

This time the eigenvalue *has* to be real; the eigenfunctions are not square integrable, but again they admit *Dirac* orthonormality:

$$\int_{-\infty}^{\infty} g_{y'}^*(x) g_y(x) dx = |A|^2 \int_{-\infty}^{\infty} \delta(x-y')\delta(x-y)dx = |A|^2\,\delta(y-y'). \tag{3.38}$$

If we pick $A = 1$, so

$$g_y(x) = \delta(x-y), \tag{3.39}$$

then

$$\langle g_{y'}|g_y\rangle = \delta(y-y'). \tag{3.40}$$

These eigenfunctions are also *complete*:

$$f(x) = \int_{-\infty}^{\infty} c(y) g_y(x) dy = \int_{-\infty}^{\infty} c(y) \delta(x - y) dy, \tag{3.41}$$

with

$$c(y) = f(y) \tag{3.42}$$

(trivial, in this case, but you can get it from Fourier's trick if you insist).

If the spectrum of a hermitian operator is *continuous* (so the eigenvalues are labeled by a continuous variable—p or y, in the examples; z, generically, in what follows), the eigenfunctions are not normalizable, they are not in Hilbert space and they do not represent possible physical states; nevertheless, the eigenfunctions with real eigenvalues are *Dirac* orthonormalizable and complete (with the sum now an integral). Luckily, this is all we really require.

Problem 3.9

(a) Cite a Hamiltonian from Chapter 2 (*other* than the harmonic oscillator) that has only a *discrete* spectrum.

(b) Cite a Hamiltonian from Chapter 2 (*other* than the free particle) that has only a *continuous* spectrum.

(c) Cite a Hamiltonian from Chapter 2 (*other* than the finite square well) that has both a discrete and a continuous part to its spectrum.

Problem 3.10 Is the ground state of the infinite square well an eigenfunction of momentum? If so, what is its momentum? If not, *why* not? [For further discussion, see Problem 3.34.]

3.4 Generalized Statistical Interpretation

In Chapter 1 I showed you how to calculate the probability that a particle would be found in a particular location, and how to determine the expectation value of any observable quantity. In Chapter 2 you learned how to find the possible outcomes of an energy measurement, and their probabilities. I am now in a position to state the **generalized statistical interpretation**, which subsumes all of this, and enables you to figure out the possible results of *any* measurement, and their probabilities. Together with the Schrödinger equation (which tells you how the wave function evolves in time) it is the foundation of quantum mechanics.

Generalized statistical interpretation: If you measure an observable $Q(x, p)$ on a particle in the state $\Psi(x, t)$, you are certain to get *one of the eigenvalues* of the hermitian operator $\hat{Q}(x, -i\hbar d/dx)$.[15] If the spectrum of $\hat{Q}$ is discrete, the probability of getting the particular eigenvalue q_n associated with the (orthonormalized) eigenfunction $f_n(x)$ is

$$|c_n|^2, \quad \text{where} \quad c_n = \langle f_n|\Psi\rangle. \tag{3.43}$$

If the spectrum is continuous, with real eigenvalues $q(z)$ and associated (Dirac-orthonormalized) eigenfunctions $f_z(x)$, the probability of getting a result in the range dz is

$$|c(z)|^2\,dz \quad \text{where} \quad c(z) = \langle f_z|\Psi\rangle. \tag{3.44}$$

Upon measurement, the wave function "collapses" to the corresponding eigenstate.[16]

The statistical interpretation is radically different from anything we encounter in classical physics. A somewhat different perspective helps to make it plausible: The eigenfunctions of an observable operator are *complete*, so the wave function can be written as a linear combination of them:

$$\Psi(x,t) = \sum_n c_n(t) f_n(x). \tag{3.45}$$

(For simplicity, I'll assume that the spectrum is discrete; it's easy to generalize this discussion to the continuous case.) Because the eigenfunctions are *orthonormal*, the coefficients are given by Fourier's trick:[17]

$$c_n(t) = \langle f_n|\Psi\rangle = \int f_n(x)^*\,\Psi(x,t)dx. \tag{3.46}$$

Qualitatively, c_n tells you "how much f_n is contained in Ψ," and given that a measurement has to return one of the eigenvalues of $\hat{Q}$, it seems reasonable that the probability of getting the particular eigenvalue q_n would be determined by the "amount of f_n" in Ψ. But because probabilities are determined by the absolute *square* of the wave function, the precise measure is actually $|c_n|^2$. That's the essential message of the generalized statistical interpretation.[18]

Of course, the *total* probability (summed over all possible outcomes) has got to be *one*:

$$\sum_n |c_n|^2 = 1, \tag{3.47}$$

and sure enough, this follows from the normalization of the wave function:

$$1 = \langle \Psi | \Psi \rangle = \left\langle \left(\sum_{n'} c_{n'} f_{n'} \right) \middle| \left(\sum_{n} c_n f_n \right) \right\rangle = \sum_{n'} \sum_{n} c_{n'}^* c_n \langle f_{n'} | f_n \rangle$$
$$= \sum_{n'} \sum_{n} c_{n'}^* c_n \delta_{n'n} = \sum_{n} c_n^* c_n = \sum_{n} |c_n|^2 . \tag{3.48}$$

Similarly, the expectation value of Q should be the sum over all possible outcomes of the eigenvalue times the probability of getting that eigenvalue:

$$\langle Q \rangle = \sum_{n} q_n |c_n|^2 . \tag{3.49}$$

Indeed,

$$\langle Q \rangle = \langle \Psi | \hat{Q} \Psi \rangle = \left\langle \left(\sum_{n'} c_{n'} f_{n'} \right) \middle| \left(\hat{Q} \sum_{n} c_n f_n \right) \right\rangle , \tag{3.50}$$

but $\hat{Q} f_n = q_n f_n$, so

$$\langle Q \rangle = \sum_{n'} \sum_{n} c_{n'}^* c_n q_n \langle f_{n'} | f_n \rangle = \sum_{n'} \sum_{n} c_{n'}^* c_n q_n \delta_{n'n} = \sum_{n} q_n |c_n|^2 . \tag{3.51}$$

So far, at least, everything looks consistent.

Can we reproduce, in this language, the original statistical interpretation for position measurements? Sure—it's overkill, but worth checking. A measurement of x on a particle in state Ψ must return one of the eigenvalues of the position operator. Well, in Example 3.3 we found that every (real) number y is an eigenvalue of x, and the corresponding (Dirac-orthonormalized) eigenfunction is $g_y(x) = \delta(x - y)$. Evidently

$$c(y) = \langle g_y | \Psi \rangle = \int_{-\infty}^{\infty} \delta(x - y) \Psi(x, t) dx = \Psi(y, t), \tag{3.52}$$

so the probability of getting a result in the range dy is $|\Psi(y, t)|^2 dy$, which is precisely the original statistical interpretation.

What about momentum? In Example 3.2 we found the (Dirac-orthonormalized) eigenfunctions of the momentum operator, $f_p(x) = (1/\sqrt{2\pi\hbar}) \exp(ipx/\hbar)$, so

$$c(p) = \langle f_p | \Psi \rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \Psi(x, t) dx. \tag{3.53}$$

This is such an important quantity that we give it a special name and symbol: the **momentum space wave function**, $\Phi(p, t)$. It is essentially the *Fourier transform* of the (**position space**) wave function $\Psi(x, t)$—which, by Plancherel's theorem, is its *inverse* Fourier transform:

$$\Phi(p, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \Psi(x, t) dx; \tag{3.54}$$

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{ipx/\hbar} \Phi(p, t) dp. \tag{3.55}$$

According to the generalized statistical interpretation, the probability that a measurement of momentum would yield a result in the range dp is

$$|\Phi(p,t)|^2\,dp. \tag{3.56}$$

Example 3.4

A particle of mass m is bound in the delta function well $V(x) = -\alpha\delta(x)$. What is the probability that a measurement of its momentum would yield a value greater than $p_0 = m\alpha/\hbar$?

Solution: The (position space) wave function is (Equation 2.132)

$$\Psi(x,t) = \frac{\sqrt{m\alpha}}{\hbar}e^{-m\alpha|x|/\hbar^2}e^{-iEt/\hbar}$$

(where $E = -m\alpha^2/2\hbar^2$). The momentum space wave function is therefore

$$\Phi(p,t) = \frac{1}{\sqrt{2\pi\hbar}}\frac{\sqrt{m\alpha}}{\hbar}e^{-iEt/\hbar}\int_{-\infty}^{\infty}e^{-ipx/\hbar}e^{-m\alpha|x|/\hbar^2}dx = \sqrt{\frac{2}{\pi}}\frac{p_0^{3/2}e^{-iEt/\hbar}}{p^2+p_0^2}$$

(I looked up the integral). The probability, then, is

$$\frac{2}{\pi}p_0^3\int_{p_0}^{\infty}\frac{1}{(p^2+p_0^2)^2}dp = \frac{1}{\pi}\left[\frac{pp_0}{p^2+p_0^2} + \tan^{-1}\left(\frac{p}{p_0}\right)\right]\Bigg|_{p_0}^{\infty} = \frac{1}{4} - \frac{1}{2\pi} = 0.0908$$

(again, I looked up the integral).

Problem 3.11 Find the momentum-space wave function, $\Phi(p,t)$, for a particle in the ground state of the harmonic oscillator. What is the probability (to two significant digits) that a measurement of p on a particle in this state would yield a value outside the classical range (for the same energy)? *Hint:* Look in a math table under "Normal Distribution" or "Error Function" for the numerical part—or use Mathematica.

Problem 3.12 Find $\Phi(p,t)$ for the free particle in terms of the function $\phi(k)$ introduced in Equation 2.101. Show that for the free particle $|\Phi(p,t)|^2$ is independent of time. *Comment*: the time independence of $|\Phi(p,t)|^2$ for the free particle is a manifestation of momentum conservation in this system.

* **Problem 3.13** Show that

$$\langle x\rangle = \int \Phi^*\left(i\hbar\frac{\partial}{\partial p}\right)\Phi dp. \tag{3.57}$$

Hint: Notice that $x \exp(ipx/\hbar) = -i\hbar(\partial/\partial p)\exp(ipx/\hbar)$, and use Equation 2.147. In momentum space, then, the position operator is $i\hbar\,\partial/\partial p$. More generally,

$$\langle Q(x,p,t)\rangle = \begin{cases} \int \Psi^* \hat{Q}\left(x, -i\hbar\frac{\partial}{\partial x}, t\right)\Psi dx, & \text{in position space;} \\ \int \Phi^* \hat{Q}\left(i\hbar\frac{\partial}{\partial p}, p, t\right)\Phi dp, & \text{in momentum space.} \end{cases} \tag{3.58}$$

In principle you can do all calculations in momentum space just as well (though not always as *easily*) as in position space.

3.5 The Uncertainty Principle

I stated the uncertainty principle (in the form $\sigma_x \sigma_p \geq \hbar/2$), back in Section 1.6, and you have checked it several times, in the problems. But we have never actually *proved* it. In this section I will prove a more general version of the uncertainty principle, and explore some of its ramifications. The argument is beautiful, but rather abstract, so watch closely.

3.5.1 Proof of the Generalized Uncertainty Principle

For any observable A, we have (Equation 3.21):

$$\sigma_A^2 = \left\langle \left(\hat{A} - \langle A \rangle\right)\Psi \middle| \left(\hat{A} - \langle A \rangle\right)\Psi \right\rangle = \langle f | f \rangle$$

where $f \equiv \left(\hat{A} - \langle A \rangle\right)\Psi$. Likewise, for any *other* observable, B,

$$\sigma_B^2 = \langle g | g \rangle, \quad \text{where } g \equiv \left(\hat{B} - \langle B \rangle\right)\Psi.$$

Therefore (invoking the Schwarz inequality, Equation 3.7),

$$\sigma_A^2 \sigma_B^2 = \langle f | f \rangle \langle g | g \rangle \geq |\langle f | g \rangle|^2. \tag{3.59}$$

Now, for any complex number z,

$$|z|^2 = [\mathrm{Re}(z)]^2 + [\mathrm{Im}(z)]^2 \geq [\mathrm{Im}(z)]^2 = \left[\frac{1}{2i}(z - z^*)\right]^2. \tag{3.60}$$

Therefore, letting $z = \langle f | g \rangle$,

$$\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i}\left[\langle f | g \rangle - \langle g | f \rangle\right]\right)^2. \tag{3.61}$$

But (exploiting the hermiticity of $\left(\hat{A} - \langle A \rangle\right)$ in the first line)

$$\begin{aligned}
\langle f | g \rangle &= \left\langle \left(\hat{A} - \langle A \rangle\right)\Psi \middle| \left(\hat{B} - \langle B \rangle\right)\Psi \right\rangle = \left\langle \Psi \middle| \left(\hat{A} - \langle A \rangle\right)\left(\hat{B} - \langle B \rangle\right)\Psi \right\rangle \\
&= \left\langle \Psi \middle| \left(\hat{A}\hat{B} - \hat{A}\langle B \rangle - \hat{B}\langle A \rangle + \langle A \rangle \langle B \rangle\right)\Psi \right\rangle \\
&= \left\langle \Psi \middle| \hat{A}\hat{B}\Psi \right\rangle - \langle B \rangle \left\langle \Psi \middle| \hat{A}\Psi \right\rangle - \langle A \rangle \left\langle \Psi \middle| \hat{B}\Psi \right\rangle + \langle A \rangle \langle B \rangle \langle \Psi | \Psi \rangle \\
&= \langle \hat{A}\hat{B} \rangle - \langle B \rangle \langle A \rangle - \langle A \rangle \langle B \rangle + \langle A \rangle \langle B \rangle \\
&= \langle \hat{A}\hat{B} \rangle - \langle A \rangle \langle B \rangle.
\end{aligned}$$

(Remember, $\langle A \rangle$ and $\langle B \rangle$ are *numbers*, not operators, so you can write them in either order.) Similarly,

$$\langle g | f \rangle = \langle \hat{B}\hat{A} \rangle - \langle A \rangle \langle B \rangle,$$

so

$$\langle f | g \rangle - \langle g | f \rangle = \langle \hat{A}\hat{B} \rangle - \langle \hat{B}\hat{A} \rangle = \langle [\hat{A}, \hat{B}] \rangle,$$

where

$$\left[\hat{A}, \hat{B}\right] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$$

is the commutator of the two operators (Equation 2.49). *Conclusion:*

$$\boxed{\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i} \langle [\hat{A}, \hat{B}] \rangle \right)^2.} \tag{3.62}$$

This is the (generalized) **uncertainty principle**. (You might think the i makes it trivial—isn't the right side *negative*? No, for the commutator of two hermitian operators carries its own factor of i, and the two cancel out;[19] the quantity in parentheses is *real*, and its square is *positive*.)

As an example, suppose the first observable is position $\left(\hat{A} = x\right)$, and the second is momentum $\left(\hat{B} = -i\hbar d/dx\right)$. We worked out their commutator back in Chapter 2 (Equation 2.52):

$$[\hat{x}, \hat{p}] = i\hbar.$$

So

$$\sigma_x^2 \sigma_p^2 \geq \left(\frac{1}{2i} i\hbar \right)^2 = \left(\frac{\hbar}{2} \right)^2,$$

or, since standard deviations are by their nature positive,

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}. \tag{3.63}$$

That's the original Heisenberg uncertainty principle, but we now see that it is just one application of a much more general theorem.

There is, in fact, an "uncertainty principle" for *every pair of observables whose operators do not commute*—we call them **incompatible observables**. Incompatible observables do not have shared eigenfunctions—at least, they cannot have a *complete set* of common eigenfunctions (see Problem 3.16). By contrast, *compatible* (commuting) observables *do* admit complete sets of simultaneous eigenfunctions (that is: states that are determinate for both observables).[20] For example, in the hydrogen atom (as we shall see in Chapter 4) the Hamiltonian, the magnitude of the angular momentum, and the z component of angular momentum are mutually compatible observables, and we will construct simultaneous eigenfunctions of all three, labeled by their respective eigenvalues. But there is *no* eigenfunction of position that is also an eigenfunction of momentum; these operators are *in*compatible.

Note that the uncertainty principle is not an *extra* assumption in quantum theory, but rather a *consequence* of the statistical interpretation. You might wonder how it is enforced in the laboratory—*why* can't you determine (say) both the position and the momentum of a particle? You can certainly measure the position of the particle, but the act of measurement collapses the wave function to a narrow spike, which necessarily carries a broad range of wavelengths (hence momenta) in its Fourier decomposition. If you now measure the momentum, the state will collapse to a long sinusoidal wave, with (now) a well-defined wavelength—but the particle no longer has the position you got in the first measurement.[21] The problem, then, is that the second measurement renders the outcome of the first measurement obsolete. Only if the wave function were simultaneously an eigenstate of both observables would it be possible to make the second measurement without disturbing the state of the particle (the second collapse wouldn't change anything, in that case). But this is only possible, in general, if the two observables are compatible.

* **Problem 3.14**

(a) Prove the following commutator identities:

$$\left[\hat{A}+\hat{B},\hat{C}\right]=\left[\hat{A},\hat{C}\right]+\left[\hat{B},\hat{C}\right], \tag{3.64}$$

$$\left[\hat{A}\hat{B},\hat{C}\right]=\hat{A}\left[\hat{B},\hat{C}\right]+\left[\hat{A},\hat{C}\right]\hat{B}. \tag{3.65}$$

(b) Show that

$$\left[x^n,\hat{p}\right]=i\hbar n x^{n-1}.$$

(c) Show more generally that

$$\left[f(x),\hat{p}\right]=i\hbar\frac{df}{dx}, \tag{3.66}$$

for any function $f(x)$ that admits a Taylor series expansion.

(d) Show that for the simple harmonic oscillator

$$\left[\hat{H},\hat{a}_{\pm}\right]=\pm\hbar\omega\hat{a}_{\pm}. \tag{3.67}$$

Hint: Use Equation 2.54.

* **Problem 3.15** Prove the famous "(your name) uncertainty principle," relating the uncertainty in position $(A=x)$ to the uncertainty in energy $\left(B=p^2/2m+V\right)$:

$$\sigma_x\sigma_H\geq\frac{\hbar}{2m}\left|\langle p\rangle\right|.$$

For stationary states this doesn't tell you much—why not?

Problem 3.16 Show that two noncommuting operators cannot have a complete set of common eigenfunctions. *Hint:* Show that *if* $\hat{P}$ and $\hat{Q}$ have a complete set of common eigenfunctions, then $\left[\hat{P},\hat{Q}\right]f=0$ for any function in Hilbert space.

3.5.2 The Minimum-Uncertainty Wave Packet

We have twice encountered wave functions that *hit* the position-momentum uncertainty limit $(\sigma_x \sigma_p = \hbar/2)$: the ground state of the harmonic oscillator (Problem 2.11) and the Gaussian wave packet for the free particle (Problem 2.21). This raises an interesting question: What is the *most general* minimum-uncertainty wave packet? Looking back at the proof of the uncertainty principle, we note that there were two points at which *in*equalities came into the argument: Equation 3.59 and Equation 3.60. Suppose we require that each of these be an *equality*, and see what this tells us about Ψ.

The Schwarz inequality becomes an equality when one function is a multiple of the other: $g(x) = cf(x)$, for some complex number c (see Problem A.5). Meanwhile, in Equation 3.60 I threw away the real part of z; equality results if $\text{Re}(z) = 0$, which is to say, if $\text{Re}\langle f|g\rangle = Re\,(c\,\langle f|f\rangle) = 0$. Now, $\langle f|f\rangle$ is certainly real, so this means the constant c must be pure imaginary—let's call it ia. The necessary and sufficient condition for minimum uncertainty, then, is

$$g(x) = iaf(x), \quad \text{where } a \text{ is real.} \tag{3.68}$$

For the position-momentum uncertainty principle this criterion becomes:

$$\left(-i\hbar\frac{d}{dx} - \langle p\rangle\right)\Psi = ia(x - \langle x\rangle)\Psi, \tag{3.69}$$

which is a differential equation for Ψ as a function of x. Its general solution (see Problem 3.17) is

$$\Psi(x) = Ae^{-a(x-\langle x\rangle)^2/2\hbar}e^{i\langle p\rangle x/\hbar}. \tag{3.70}$$

Evidently the minimum-uncertainty wave packet is a *gaussian*—and, sure enough, the two examples we encountered earlier *were* gaussians.[22]

> **Problem 3.17** Solve Equation 3.69 for $\Psi(x)$. Note that $\langle x\rangle$ and $\langle p\rangle$ are *constants* (independent of x).

3.5.3 The Energy-Time Uncertainty Principle

The position-momentum uncertainty principle is often written in the form

$$\Delta x \Delta p \geq \frac{\hbar}{2}; \tag{3.71}$$

Δx (the "uncertainty" in x) is loose notation (and sloppy language) for the standard deviation of the results of repeated measurements on identically prepared systems.[23] Equation 3.71 is often paired with the **energy-time uncertainty principle,**

$$\Delta t \Delta E \geq \frac{\hbar}{2}. \tag{3.72}$$

Indeed, in the context of special relativity the energy-time form might be thought of as a *consequence* of the position-momentum version, because x and t (or rather, ct) go together in the position-time four-vector, while p and E (or rather, E/c) go together in the energy-momentum four-vector. So in a relativistic theory Equation 3.72 would be a necessary concomitant to Equation 3.71. But we're not doing relativistic quantum mechanics. The Schrödinger equation is explicitly nonrelativistic: It treats t and x on a very unequal footing (as a differential equation it is *first*-order in t, but *second*-order in x), and Equation 3.72 is emphatically *not* implied by Equation 3.71. My purpose now is to *derive* the energy-time uncertainty principle, and in the course of that derivation to persuade you that it is really an altogether different beast, whose superficial resemblance to the position-momentum uncertainty principle is actually quite misleading.

After all, position, momentum, and energy are all dynamical variables—measurable characteristics of the system, at any given time. But time itself is not a dynamical variable (not, at any rate, in a nonrelativistic theory): You don't go out and measure the "time" of a particle, as you might its position or its energy. Time is the *independent* variable, of which the dynamical quantities are *functions*. In particular, the Δt in the energy-time uncertainty principle is not the standard deviation of a collection of time measurements; roughly speaking (I'll make this more precise in a moment) it is the *time it takes the system to change substantially*.

As a measure of how fast the system is changing, let us compute the time derivative of the expectation value of some observable, $Q(x, p, t)$:

$$\frac{d}{dt}\langle Q\rangle = \frac{d}{dt}\left\langle \Psi \middle| \hat{Q}\Psi \right\rangle = \left\langle \frac{\partial \Psi}{\partial t} \middle| \hat{Q}\Psi \right\rangle + \left\langle \Psi \middle| \frac{\partial \hat{Q}}{\partial t}\Psi \right\rangle + \left\langle \Psi \middle| \hat{Q}\frac{\partial \Psi}{\partial t} \right\rangle.$$

Now, the Schrödinger equation says

$$i\hbar\frac{\partial \Psi}{\partial t} = \hat{H}\Psi$$

(where $H = p^2/2m + V$ is the Hamiltonian). So

$$\frac{d}{dt}\langle Q\rangle = -\frac{1}{i\hbar}\left\langle \hat{H}\Psi \middle| \hat{Q}\Psi \right\rangle + \frac{1}{i\hbar}\left\langle \Psi \middle| \hat{Q}\hat{H}\Psi \right\rangle + \left\langle \frac{\partial \hat{Q}}{\partial t} \right\rangle.$$

But $\hat{H}$ is hermitian, so $\langle \hat{H}\Psi | \hat{Q}\Psi\rangle = \langle \Psi | \hat{H}\hat{Q}\Psi\rangle$, and hence

$$\boxed{\frac{d}{dt}\langle Q\rangle = \frac{i}{\hbar}\langle[\hat{H},\hat{Q}]\rangle + \left\langle\frac{\partial\hat{Q}}{\partial t}\right\rangle.} \tag{3.73}$$

This is an interesting and useful result in its own right (see Problems 3.18 and 3.37). It has no name, though it surely deserves one; I'll call it the **generalized Ehrenfest theorem**. In the typical case where the operator does not depend explicitly on time,[24] it tells us that the rate of change of the expectation value is determined by the commutator of the operator with the Hamiltonian. In particular, if $\hat{Q}$ *commutes* with $\hat{H}$, then $\langle Q\rangle$ is constant, and in this sense Q is a *conserved* quantity.

Now, suppose we pick $A = H$ and $B = Q$, in the generalized uncertainty principle (Equation 3.62), and assume that Q does not depend explicitly on t:

$$\sigma_H^2\sigma_Q^2 \geq \left(\frac{1}{2i}\langle[\hat{H},\hat{Q}]\rangle\right)^2 = \left(\frac{1}{2i}\frac{\hbar}{i}\frac{d\langle Q\rangle}{dt}\right)^2 = \left(\frac{\hbar}{2}\right)^2\left(\frac{d\langle Q\rangle}{dt}\right)^2.$$

Or, more simply,

$$\sigma_H\sigma_Q \geq \frac{\hbar}{2}\left|\frac{d\langle Q\rangle}{dt}\right|. \tag{3.74}$$

Let's define $\Delta E \equiv \sigma_H$, and

$$\Delta t \equiv \frac{\sigma_Q}{|d\langle Q\rangle/dt|}. \tag{3.75}$$

Then

$$\Delta E\,\Delta t \geq \frac{\hbar}{2}, \tag{3.76}$$

and that's the energy-time uncertainty principle. But notice what is meant by Δt, here: Since

$$\sigma_Q = \left|\frac{d\langle Q\rangle}{dt}\right|\Delta t,$$

Δt represents the *amount of time it takes the expectation value of Q to change by one standard deviation.*[25] In particular, Δt depends entirely on what observable (Q) you care to look at—the change might be rapid for one observable and slow for another. But if ΔE is small, then the rate of change of *all* observables must be very gradual; or, to put it the other way around, if *any* observable changes rapidly, the "uncertainty" in the energy must be large.

Example 3.5

In the extreme case of a stationary state, for which the energy is uniquely determined, all expectation values are constant in time ($\Delta E = 0 \Rightarrow \Delta t = \infty$)—as in fact we noticed some time ago (see Equation 2.9). To make something *happen* you must take a linear combination of at least two stationary states—say:

$$\Psi(x,t) = a\psi_1(x)e^{-iE_1t/\hbar} + b\psi_2(x)e^{-iE_2t/\hbar}.$$

If a, b, ψ_1, and ψ_2 are real,

$$|\Psi(x,t)|^2 = a^2\,(\psi_1(x))^2 + b^2\,(\psi_2(x))^2 + 2ab\psi_1(x)\psi_2(x)\cos\left(\frac{E_2 - E_1}{\hbar}t\right).$$

The period of oscillation is $\tau = 2\pi\hbar/(E_2 - E_1)$. Roughly speaking, $\Delta E = E_2 - E_1$ and $\Delta t = \tau$ (for the *exact* calculation see Problem 3.20), so

$$\Delta E \Delta t = 2\pi\hbar,$$

which is indeed $\geq \hbar/2$.

Example 3.6

Let Δt be the time it takes a free-particle wave packet to pass a particular point (Figure 3.1). Qualitatively (an exact version is explored in Problem 3.21), $\Delta t = \Delta x/v = m\Delta x/p$. But $E = p^2/2m$, so $\Delta E = p\Delta p/m$, and therefore,

$$\Delta E \Delta t = \frac{p\Delta p}{m}\frac{m\Delta x}{p} = \Delta x \Delta p,$$

which is $\geq \hbar/2$ by the position-momentum uncertainty principle.

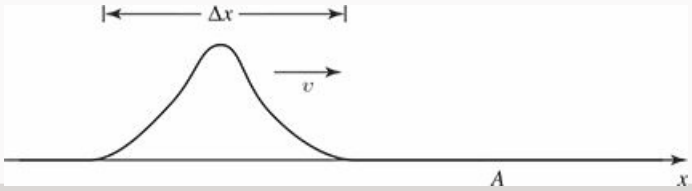

Figure 3.1: A free particle wave packet approaches the point A (Example 3.6).

Example 3.7

The Δ particle lasts about 10^{-23} s, before spontaneously disintegrating. If you make a histogram of all measurements of its mass, you get a kind of bell-shaped curve centered at 1232 MeV/c^2, with a width of about 120 MeV/c^2 (Figure 3.2). Why does the rest energy (mc^2) sometimes come out higher than 1232, and sometimes lower? Is this experimental error? No, for if we take Δt to be the lifetime of the particle (certainly *one* measure of "how long it takes the system to change appreciably"),

$$\Delta E \Delta t = \left(\frac{120}{2}\text{ MeV}\right)\left(10^{-23}\text{ s}\right) = 6 \times 10^{-22}\text{ MeV-s},$$

whereas $\hbar/2 = 3 \times 10^{-22}$ MeV s. So the spread in m is about as small as the uncertainty principle allows—a particle with so short a lifetime just doesn't *have* a very well-defined mass.[26]

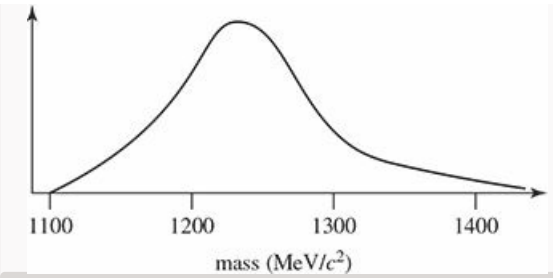

Figure 3.2: Measurements of the Δ mass (Example 3.7).

Notice the variety of specific meanings attaching to the term Δt in these examples: In Example 3.5 it's a period of oscillation; in Example 3.6 it's the time it takes a particle to pass a point; in Example 3.7 it's the lifetime of an unstable particle. In every case, however, Δt is the time it takes for the system to undergo "substantial" change.

It is often said that the uncertainty principle means energy is not strictly conserved in quantum mechanics—that you're allowed to "borrow" energy ΔE, as long as you "pay it back" in a time $\Delta t \approx \hbar/(2\Delta E)$; the greater the violation, the briefer the period over which it can occur. Now, there are many legitimate readings of the energy-time uncertainty principle, but this is not one of them. Nowhere does quantum mechanics license violation of energy conservation, and certainly no such authorization entered into the derivation of Equation 3.76. But the uncertainty principle is extraordinarily robust: It can be misused without leading to seriously incorrect results, and as a consequence physicists are in the habit of applying it rather carelessly.

* **Problem 3.18** Apply Equation 3.73 to the following special cases: (a) $Q = 1$; (b) $Q = H$; (c) $Q = x$; (d) $Q = p$. In each case, comment on the result, with particular reference to Equations 1.27, 1.33, 1.38, and conservation of energy (see remarks following Equation 2.21).

Problem 3.19 Use Equation 3.73 (or Problem 3.18 (c) and (d)) to show that:

(a) For any (normalized) wave packet representing a free particle ($V(x) = 0$), $\langle x \rangle$ moves at constant velocity (this is the quantum analog to Newton's first law). *Note:* You showed this for a *gaussian* wave packet in Problem 2.42, but it is completely general.

(b) For any (normalized) wave packet representing a particle in the harmonic oscillator potential $\left(V(x) = \frac{1}{2}m\omega^2 x^2\right)$, $\langle x \rangle$ oscillates at the classical frequency. *Note:* You showed this for a particular *gaussian* wave packet in Problem 2.49, but it is completely general.

Problem 3.20 Test the energy-time uncertainty principle for the wave function in Problem 2.5 and the observable x, by calculating σ_H, σ_x, and $d\langle x \rangle/dt$ exactly.

Problem 3.21 Test the energy-time uncertainty principle for the free particle wave packet in Problem 2.42 and the observable x, by calculating σ_H, σ_x, and $d\langle x\rangle/dt$ exactly.

Problem 3.22 Show that the energy-time uncertainty principle reduces to the "your name" uncertainty principle (Problem 3.15), when the observable in question is x.

3.6 Vectors and Operators

3.6.1 Bases in Hilbert Space

Imagine an ordinary vector **A** in two dimensions (Fig. 3.3(a)). How would you describe this vector to someone? You might tell them "It's about an inch long, and it points 20° clockwise from straight up, with respect to the page." But that's pretty awkward. A better way would be to introduce cartesian axes, x and y, and specify the *components* of **A**: $A_x = \hat{\imath} \cdot \mathbf{A}$, $A_y = \hat{\jmath} \cdot \mathbf{A}$ (Fig. 3.3(b)). Of course, your sister might draw a different set of axes, x' and y', and she would report different components: $A'_x = \hat{\imath}' \cdot \mathbf{A}$, $A'_y = \hat{\jmath}' \cdot \mathbf{A}$ (Fig. 3.3(c)) …but it's all the same *vector*—we're simply expressing it with respect to two different **bases** ($\{\hat{\imath}, \hat{\jmath}\}$ and $\{\hat{\imath}', \hat{\jmath}'\}$). The vector itself lives "out there in space," independent of anybody's (arbitrary) choice of coordinates.

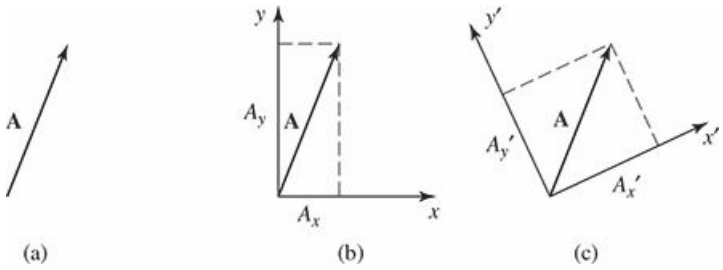

Figure 3.3: (a) Vector **A**. (b) Components of **A** with respect to xy axes. (c) Components of **A** with respect to $x'y'$ axes.

The same is true for the state of a system in quantum mechanics. It is represented by a *vector*, $|\mathcal{S}(t)\rangle$, that lives "out there in Hilbert space," but we can *express* it with respect to any number of different *bases*. The wave function $\Psi(x, t)$ is actually the x "component" in the expansion of $|\mathcal{S}(t)\rangle$ in the basis of position eigenfunctions:

$$\boxed{\Psi(x,t) = \langle x|\mathcal{S}(t)\rangle\,,} \tag{3.77}$$

(the analog to $\hat{\imath} \cdot \mathbf{A}$) with $|x\rangle$ standing for the eigenfunction of $\hat{x}$ with eigenvalue x.[27] The momentum space wave function $\Phi(p, t)$ is the p component in the expansion of $|\mathcal{S}(t)\rangle$ in the basis of momentum eigenfunctions:

$$\Phi(p,t) = \langle p|\mathcal{S}(t)\rangle \tag{3.78}$$

(with $|p\rangle$ standing for the eigenfunction of $\hat{p}$ with eigenvalue p).[28] Or we could expand $|\mathcal{S}(t)\rangle$ in the basis of energy eigenfunctions (supposing for simplicity that the spectrum is discrete):

$$c_n(t) = \langle n|\mathcal{S}(t)\rangle \tag{3.79}$$

(with $|n\rangle$ standing for the nth eigenfunction of $\hat{H}$—Equation 3.46). But it's all the same state; the functions Ψ and Φ, and the collection of coefficients $\{c_n\}$, contain exactly the same information—they are simply three different ways of identifying the same vector:

$$\begin{aligned} |\mathcal{S}(t)\rangle \rightarrow & \int \Psi(y,t)\delta(x-y)dy = \int \Phi(p,t)\frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}dp \\ & = \sum c_n e^{-iE_n t/\hbar}\psi_n(x). \end{aligned} \tag{3.80}$$

Operators (representing observables) are linear transformations on Hilbert space—they "transform" one vector into another:

$$|\beta\rangle = \hat{Q}|\alpha\rangle. \tag{3.81}$$

Just as vectors are represented, with respect to an orthonormal basis $\{|e_n\rangle\}$,[29] by their components,

$$|\alpha\rangle = \sum_n a_n |e_n\rangle, \quad |\beta\rangle = \sum_n b_n |e_n\rangle, \quad a_n = \langle e_n|\alpha\rangle, \quad b_n = \langle e_n|\beta\rangle\,, \tag{3.82}$$

operators are represented (with respect to a particular basis) by their **matrix elements**[30]

$$\left\langle e_m \middle| \hat{Q} \middle| e_n \right\rangle \equiv Q_{mn}. \tag{3.83}$$

In this notation Equation 3.81 says

$$\sum_n b_n |e_n\rangle = \sum_n a_n \hat{Q}|e_n\rangle, \tag{3.84}$$

or, taking the inner product with $|e_m\rangle$,

$$\sum_n b_n \langle e_m|e_n\rangle = \sum_n a_n \left\langle e_m \middle| \hat{Q} \middle| e_n \right\rangle, \tag{3.85}$$

and hence (since $\langle e_m|e_n\rangle = \delta_{mn}$)

$$b_m = \sum_n Q_{mn} a_n. \tag{3.86}$$

Thus the matrix elements of $\hat{Q}$ tell you how the components transform.[31]

Later on we will encounter systems that admit only a finite number N of linearly independent states. In that case $|S(t)\rangle$ lives in an N-dimensional vector space; it can be represented as a column of (N) components (with respect to a given basis), and operators take the form of ordinary $(N \times N)$ matrices. These are the simplest quantum systems—none of the subtleties associated with infinite-dimensional vector spaces arise. Easiest of all is the two-state system, which we explore in the following example.

Example 3.8

Imagine a system in which there are just *two* linearly independent states:[32]

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The most general state is a normalized linear combination:

$$|S\rangle = a|1\rangle + b|2\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \text{with } |a|^2 + |b|^2 = 1.$$

The Hamiltonian can be expressed as a (hermitian) matrix (Equation 3.83); suppose it has the specific form

$$\mathsf{H} = \begin{pmatrix} h & g \\ g & h \end{pmatrix},$$

where g and h are real constants. If the system starts out (at $t = 0$) in state $|1\rangle$, what is its state at time t?

Solution: The (time-dependent) Schrödinger equation[33] says

$$i\hbar \frac{d}{dt}|\mathcal{S}(t)\rangle = \hat{H}|\mathcal{S}(t)\rangle. \tag{3.87}$$

As always, we begin by solving the time-*in*dependent Schrödinger equation:

$$\hat{H}|s\rangle = E|s\rangle; \tag{3.88}$$

that is, we look for the eigenvectors and eigenvalues of $\hat{H}$. The characteristic equation determines the eigenvalues:

$$\det \begin{pmatrix} h-E & g \\ g & h-E \end{pmatrix} = (h-E)^2 - g^2 = 0 \Rightarrow h - E = \mp g \Rightarrow E_\pm = h \pm g.$$

Evidently the allowed energies are $(h + g)$ and $(h - g)$. To determine the eigenvectors, we write

$$\begin{pmatrix} h & g \\ g & h \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = (h \pm g) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \Rightarrow h\alpha + g\beta = (h \pm g)\alpha \Rightarrow \beta = \pm\alpha,$$

so the normalized eigenvectors are

$$|s_\pm\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}.$$

Next we expand the initial state as a linear combination of eigenvectors of the Hamiltonian:

$$|\mathcal{S}(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}(|s_+\rangle + |s_-\rangle).$$

Finally, we tack on the standard time-dependence (the wiggle factor) $\exp(-iE_n t/\hbar)$:

$$\begin{aligned} |S(t)\rangle &= \frac{1}{\sqrt{2}}\left[e^{-i(h+g)t/\hbar}|s_+\rangle + e^{-i(h-g)t/\hbar}|s_-\rangle\right] \\ &= \frac{1}{2}e^{-iht/\hbar}\left[e^{-igt/\hbar}\begin{pmatrix}1\\1\end{pmatrix} + e^{igt/\hbar}\begin{pmatrix}1\\-1\end{pmatrix}\right] \\ &= \frac{1}{2}e^{-iht/\hbar}\begin{pmatrix} e^{-igt/\hbar} + e^{igt/\hbar} \\ e^{-igt/\hbar} - e^{igt/\hbar}\end{pmatrix} = e^{-iht/\hbar}\begin{pmatrix}\cos(gt/\hbar) \\ -i\sin(gt/\hbar)\end{pmatrix}. \end{aligned}$$

If you doubt this result, by all means *check* it: Does it satisfy the time-dependent Schrödinger equation (Equation 3.87)? Does it match the initial state when $t = 0$?[34]

Just as vectors look different when expressed in different bases, so too do operators (or, in the discrete case, the matrices that represent them). We have already encountered a particularly nice example:

$$\hat{x} \text{ (the position operator)} \rightarrow \begin{cases} x & \text{(in position space)}, \\ i\hbar\partial/\partial p & \text{(in momentum space)}; \end{cases}$$

$$\hat{p} \text{ (the momentum operator)} \rightarrow \begin{cases} -i\hbar\partial/\partial x & \text{(in position space)}, \\ p & \text{(in momentum space)}. \end{cases}$$

("Position space" is nothing but the position basis; "momentum space" is the momentum basis.) If someone asked you, "What is the operator, $\hat{x}$, representing position, in quantum mechanics?" you would probably answer "Just x itself." But an equally correct reply would be "$i\hbar\partial/\partial p$," and the best response would be "With respect to what basis?"

I have often said "the state of a system is represented by its wave function, $\Psi(x, t)$," and this is true, in the same sense that an ordinary vector in three dimensions is "represented by" the triplet of its components; but really, I should always add "in the position basis." After all, the state of the system is a vector in Hilbert space, $|\mathcal{S}(t)\rangle$; it makes no reference to any particular basis. Its connection to $\Psi(x, t)$ is given by Equation 3.77: $\Psi(x, t) = \langle x|\mathcal{S}(t)\rangle$. Having said that, for the most part we *do* in fact work in position space, and no serious harm comes from referring to the wave function as "the state of the system."

3.6.2 Dirac Notation

Dirac proposed to chop the bracket notation for the inner product, $\langle\alpha|\beta\rangle$, into two pieces, which he called **bra**, $\langle\alpha|$, and **ket**, $|\beta\rangle$ (I don't know what happened to the c). The latter is a vector, but what exactly is the former? It's a *linear function* of vectors, in the sense that when it hits a vector (to its right) it yields a (complex) number—the inner product. (When an *operator* hits a vector, it delivers another vector; when a *bra* hits a vector, it delivers a number.) In a function space, the bra can be thought of as an instruction to integrate:

$$\langle f| = \int f^*[\cdots]\,dx,$$

with the ellipsis $[\cdots]$ waiting to be filled by whatever function the bra encounters in the ket to its right. In a finite-dimensional vector space, with the kets expressed as columns (of components with respect to some basis),

$$|\alpha\rangle \to \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \tag{3.89}$$

the bras are rows:

$$\langle\beta| \to \begin{pmatrix} b_1^* & b_2^* & \cdots & b_n^* \end{pmatrix}, \tag{3.90}$$

and $\langle\beta|\alpha\rangle = b_1^*a_1 + b_2^*a_2 + \cdots + b_n^*a_n$ is the matrix product. The collection of all bras constitutes another vector space—the so-called **dual space**.

The license to treat bras as separate entities in their own right allows for some powerful and pretty notation. For example, if $|\alpha\rangle$ is a normalized vector, the operator

$$\hat{P} \equiv |\alpha\rangle\langle\alpha| \tag{3.91}$$

picks out the portion of any other vector that "lies along" $|\alpha\rangle$:

$$\hat{P}\,|\beta\rangle = (\langle\alpha|\,\beta\rangle)\,|\alpha\rangle;$$

we call it the **projection operator** onto the one-dimensional subspace spanned by $|\alpha\rangle$. If $\{|e_n\rangle\}$ is a discrete orthonormal basis,

$$\langle e_m|e_n\rangle = \delta_{mn}, \tag{3.92}$$

then

$$\boxed{\sum_n |e_n\rangle\langle e_n| = 1} \tag{3.93}$$

(the identity operator). For if we let this operator act on any vector $|\alpha\rangle$, we recover the expansion of $|\alpha\rangle$ in the $\{|e_n\rangle\}$ basis:

$$\sum_n (\langle e_n\,|\alpha\rangle)|\,e_n\rangle = |\alpha\rangle. \tag{3.94}$$

Similarly, if $\{|e_z\rangle\}$ is a *Dirac* orthonormalized continuous basis,

$$\langle e_z | e_{z'} \rangle = \delta(z - z'), \tag{3.95}$$

then

$$\boxed{\int |e_z\rangle\langle e_z| \, dz = 1.} \tag{3.96}$$

Equations 3.93 and 3.96 are the tidiest ways to express *completeness*.

Technically, the guts of a ket or a bra (the ellipsis in $|\cdots\rangle$ or $\langle\cdots|$) is a *name*—the name of the vector in question: "a," or "n," or for that matter "Alice," or "Bob." It is endowed with no intrinsic mathematical attributes. Of course, it may be helpful to choose an evocative name—for instance, if you're working in the space L^2 of square-integrable functions, it is natural to name each vector after the function it represents: $|f\rangle$. Then, for example, we can write the definition of a hermitian operator as we did in Equation 3.17:

$$\left\langle f \middle| \hat{Q} f \right\rangle = \left\langle \hat{Q} f \middle| f \right\rangle.$$

Strictly speaking, in Dirac notation this is a nonsense expression: f here is a *name*, and operators act on *vectors*, not on *names*. The left side should properly be written as

$$\left\langle f \middle| \hat{Q} \middle| f \right\rangle,$$

but what are we to make of the right side? $\langle \hat{Q} f |$ really means "the bra dual to $\hat{Q}\,|f\rangle$," but what is its name? I suppose we could say

$$\left\langle \left(\text{the name of the vector } \hat{Q}\,|f\rangle \right) \right|,$$

but that's a mouthful. However, since we have chosen to name each vector after the function it represents, and since we *do* know how $\hat{Q}$ acts on the *function* (as opposed to the *name*) f, this in fact becomes[35]

$$\langle \hat{Q} f |,$$

and we are OK after all.[36]

An operator takes one vector in Hilbert space and delivers another:

$$\hat{Q}|\alpha\rangle = |\beta\rangle. \tag{3.97}$$

The sum of two operators is defined in the obvious way,

$$\left(\hat{Q} + \hat{R}\right)|\alpha\rangle = \hat{Q}\,|\alpha\rangle + \hat{R}\,|\alpha\rangle, \tag{3.98}$$

and the product of two operators is

$$\hat{Q}\hat{R}\,|\alpha\rangle = \hat{Q}\left(\hat{R}\,|\alpha\rangle\right) \tag{3.99}$$

(first apply $\hat{R}$ to $|\alpha\rangle$, and then apply $\hat{Q}$ to what you got—being careful, of course, to respect their *ordering*). Occasionally we shall encounter *functions* of operators. They are typically defined by the power series

expansion:

$$e^{\hat{Q}} \equiv 1 + \hat{Q} + \frac{1}{2}\hat{Q}^2 + \frac{1}{3!}\hat{Q}^3 + \cdots \tag{3.100}$$

$$\frac{1}{1-\hat{Q}} \equiv 1 + \hat{Q} + \hat{Q}^2 + \hat{Q}^3 + \hat{Q}^4 + \cdots \tag{3.101}$$

$$\ln\left(1+\hat{Q}\right) \equiv \hat{Q} - \frac{1}{2}\hat{Q}^2 + \frac{1}{3}\hat{Q}^3 - \frac{1}{4}\hat{Q}^4 + \cdots \tag{3.102}$$

and so on. On the right-hand side we have only sums and products, and we *know* how to handle *them*.

Problem 3.23 Show that projection operators are **idempotent**: $\hat{P}^2 = \hat{P}$. Determine the eigenvalues of $\hat{P}$, and characterize its eigenvectors.

Problem 3.24 Show that if an operator $\hat{Q}$ is hermitian, then its matrix elements in any orthonormal basis satisfy $Q_{mn} = Q^*_{nm}$. That is, the corresponding matrix is equal to its transpose conjugate.

Problem 3.25 The Hamiltonian for a certain two-level system is

$$\hat{H} = \epsilon\left(|1\rangle\langle 1| - |2\rangle\langle 2| + |1\rangle\langle 2| + |2\rangle\langle 1|\right),$$

where $|1\rangle, |2\rangle$ is an orthonormal basis and ϵ is a number with the dimensions of energy. Find its eigenvalues and eigenvectors (as linear combinations of $|1\rangle$ and $|2\rangle$). What is the matrix H representing $\hat{H}$ with respect to this basis?

* **Problem 3.26** Consider a three-dimensional vector space spanned by an orthonormal basis $|1\rangle, |2\rangle, |3\rangle$. Kets $|\alpha\rangle$ and $|\beta\rangle$ are given by

$$|\alpha\rangle = i|1\rangle - 2|2\rangle - i|3\rangle, \quad |\beta\rangle = i|1\rangle + 2|3\rangle.$$

(a) Construct $\langle\alpha|$ and $\langle\beta|$ (in terms of the dual basis $\langle 1|, \langle 2|, \langle 3|$).

(b) Find $\langle\alpha|\beta\rangle$ and $\langle\beta|\alpha\rangle$, and confirm that $\langle\beta|\alpha\rangle = \langle\alpha|\beta\rangle^*$.

(c) Find all nine matrix elements of the operator $\hat{A} \equiv |\alpha\rangle\langle\beta|$, in this basis, and construct the matrix A. Is it hermitian?

Problem 3.27 Let $\hat{Q}$ be an operator with a complete set of orthonormal eigenvectors:

$$\hat{Q}|e_n\rangle = q_n|e_n\rangle \quad (n = 1, 2, 3, \ldots).$$

(a) Show that $\hat{Q}$ can be written in terms of its **spectral decomposition**:

$$\hat{Q} = \sum_n q_n \, |e_n\rangle\langle e_n| \, . \tag{3.103}$$

Hint: An operator is characterized by its action on all possible vectors, so what you must show is that

$$\hat{Q}|\alpha\rangle = \left\{ \sum_n q_n |e_n\rangle \langle e_n| \right\} |\alpha\rangle \, ,$$

for any vector $|\alpha\rangle$.

(b) Another way to define a function of $\hat{Q}$ is via the spectral decomposition:

$$f(\hat{Q}) = \sum_n f(q_n) \, |e_n\rangle\langle e_n| \, . \tag{3.104}$$

Show that this is equivalent to Equation 3.100 in the case of $e^{\hat{Q}}$.

Problem 3.28 Let $\hat{D} = d/dx$ (the derivative operator). Find

(a) $\left(\sin \hat{D}\right) x^5$.

(b) $\left(\dfrac{1}{1 - \hat{D}/2}\right) \cos\,(x)$.

** **Problem 3.29** Consider operators $\hat{A}$ and $\hat{B}$ that do not commute with each other $\left(\hat{C} = \left[\hat{A}, \hat{B}\right]\right)$ but do commute with their commutator: $\left[\hat{A}, \hat{C}\right] = \left[\hat{B}, \hat{C}\right] = 0$ (for instance, $\hat{x}$ and $\hat{p}$).

(a) Show that

$$\left[\hat{A}^n, \hat{B}\right] = n\hat{A}^{n-1}\hat{C}.$$

Hint: You can prove this by induction on n, using Equation 3.65.

(b) Show that

$$\left[e^{\lambda\hat{A}}, \hat{B}\right] = \lambda e^{\lambda\hat{A}}\hat{C},$$

where λ is any complex number. *Hint:* Express $e^{\lambda\hat{A}}$ as a power series.

(c) Derive the **Baker–Campbell–Hausdorff formula**:[37]

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}}e^{\hat{B}}e^{-\hat{C}/2}.$$

Hint: Define the functions

$$\hat{f}\,(\lambda) = e^{\lambda\left(\hat{A}+\hat{B}\right)}, \qquad \hat{g}\,(\lambda) = e^{\lambda\hat{A}}e^{\lambda\hat{B}}e^{-\lambda^2\hat{C}/2}.$$

Note that these functions are equal at $\lambda = 0$, and show that they satisfy the same differential equation: $d\hat{f}/d\lambda = \left(\hat{A} + \hat{B}\right)\hat{f}$ and

$d\hat{g}/d\lambda = \left(\hat{A} + \hat{B}\right)\hat{g}$. Therefore, the functions are themselves equal for all λ.[38]

3.6.3 Changing Bases in Dirac Notation

The advantage of Dirac notation is that it frees us from working in any particular basis, and makes transforming between bases seamless. Recall that the identity operator can be written as a projection onto a complete set of states (Equations 3.93 and 3.96); of particular interest are the position eigenstates $|x\rangle$, the momentum eigenstates $|p\rangle$, and the energy eigenstates (we will assume those are discrete) $|n\rangle$:

$$1 = \int dx\, |x\rangle\langle x|\,, \qquad 1 = \int dp\, |p\rangle\langle p|\,, \qquad 1 = \sum |n\rangle\langle n|\,. \tag{3.106}$$

Acting on the state vector $|\mathcal{S}(t)\rangle$ with each of these resolutions of the identity gives

$$|\mathcal{S}(t)\rangle = \int dx\, |x\rangle\langle x|\mathcal{S}(t)\rangle \equiv \int \Psi(x,t)\, |x\rangle\, dx,$$
$$|\mathcal{S}(t)\rangle = \int dp\, |p\rangle\langle p|\mathcal{S}(t)\rangle \equiv \int \Phi(p,t)\, |p\rangle\, dp, \tag{3.107}$$
$$|\mathcal{S}(t)\rangle = \sum_n |n\rangle\langle n|\mathcal{S}(t)\rangle \equiv \sum c_n(t)\, |n\rangle\,.$$

Here we recognize the position-space, momentum-space, and "energy-space" wave functions (Equations 3.77–3.79) as the *components* of the vector $|\mathcal{S}(t)\rangle$ in the respective bases.

Example 3.9

Derive the transformation from the position-space wave function to the momentum-space wave function. (We already know the answer, of course, but I want to show you how this works out in Dirac notation.)

Solution: We want to find $\Phi(p,t) = \langle p|\mathcal{S}(t)\rangle$ given $\Psi(x,t) = \langle x|\mathcal{S}(t)\rangle$. We can relate the two by inserting a resolution of the identity:

$$\begin{aligned}\Phi(p,t) &= \langle p|\mathcal{S}(t)\rangle \\ &= \left\langle p\left|\left(\int dx\, |x\rangle\langle x|\right)\right|\mathcal{S}(t)\right\rangle \\ &= \int \langle p|x\rangle\, \langle x|\mathcal{S}(t)\rangle\, dx \\ &= \int \langle p|x\rangle\, \Psi(x,t)\, dx.\end{aligned} \tag{3.108}$$

Now, $\langle x|p\rangle$ is the momentum eigenstate (with eigenvalue p) in the position basis—what we called $f_p(x)$, in Equation 3.32. So

$$\langle p|x\rangle = \langle x|p\rangle^* = [f_p(x)]^* = \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx/\hbar}.$$

Plugging this into Equation 3.108 gives

$$\Phi(p,t) = \int \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx/\hbar} \Psi(x,t)\, dx,$$

which is precisely Equation 3.54.

Just as the wave function takes different forms in different bases, so do operators. The position operator is given by

$$\hat{x} \to x$$

in the position basis, or

$$\hat{x} \to i\hbar \frac{\partial}{\partial p}$$

in the momentum basis. However, Dirac notation allows us to do away with the arrows and stick to equalities. Operators act on kets (for instance, $\hat{x}\,|\mathcal{S}(t)\rangle$); the outcome of this operation can be expressed in any basis by taking the inner product with an appropriate basis vector. That is,

$$\langle x\,|\hat{x}|\,\mathcal{S}(t)\rangle = \text{action of position operator in } x \text{ basis} = x\,\Psi(x,t)\,, \tag{3.109}$$

or

$$\langle p\,|\hat{x}|\,\mathcal{S}(t)\rangle = \text{action of position operator in } p \text{ basis} = i\hbar \frac{\partial \Phi}{\partial p}\,. \tag{3.110}$$

In this notation it is straightforward to transform operators between bases, as the following example illustrates.

Example 3.10

Obtain the position operator in the momentum basis (Equation 3.110) by inserting a resolution of the identity on the left-hand side.

Solution:

$$\begin{aligned}\langle p\,|\hat{x}|\,\mathcal{S}(t)\rangle &= \left\langle p\left|\hat{x}\int dx\right|x\right\rangle\left\langle x\,\middle|\,\middle|\,\mathcal{S}(t)\right\rangle \\ &= \int \langle p\,|x|\,x\rangle\ \langle x\,|\,S(t)\rangle\ dx\,,\end{aligned}$$

where I've used the fact that $|x\rangle$ is an eigenstate of $\hat{x}\left(\hat{x}\,|x\rangle = x\,|x\rangle\right)$; x can then be pulled out of the inner product (it's just a number) and

$$\begin{aligned}\langle p\,|\hat{x}|\,\mathcal{S}(t)\rangle &= \int x\,\langle p|x\rangle\,\Psi(x,t)\,dx \\ &= \int x\frac{e^{-ipx/\hbar}}{\sqrt{2\pi\hbar}}\Psi(x,t)\,dx \\ &= i\hbar\frac{\partial}{\partial p}\int\frac{e^{-ipx/\hbar}}{\sqrt{2\pi\hbar}}\Psi(x,t)\,dx.\end{aligned}$$

Finally we recognize the integral as $\Phi(p,t)$ (Equation 3.54).

Problem 3.30 Derive the transformation from the position-space wave function to the "energy-space" wave function $(c_n(t))$ using the technique of Example 3.9. Assume that the energy spectrum is discrete, and the potential is time-independent.