(b) as the sum of a real matrix **R** and an imaginary matrix **I**;

(c) as the sum of a Hermitian matrix **H** and a skew-Hermitian matrix **K**.

***Problem 3.12** Prove Equations 3.52, 3.53, and 3.58. Show that the product of two unitary matrices is unitary. Under what conditions is the product of two Hermitian matrices Hermitian? Is the sum of two unitary matrices unitary? Is the sum of two Hermitian matrices Hermitian?

Problem 3.13 In the usual basis $(\hat{\imath}, \hat{\jmath}, \hat{k})$, construct the matrix $\mathbf{T}_x$ representing a rotation through angle θ about the x-axis, and the matrix $\mathbf{T}_y$ representing a rotation through angle θ about the y-axis. Suppose now we change bases, to $\hat{\imath}' = \hat{\jmath}$, $\hat{\jmath}' = -\hat{\imath}$, $\hat{k}' = \hat{k}$. Construct the matrix **S** that effects this change of basis, and check that $\mathbf{S}\mathbf{T}_x\mathbf{S}^{-1}$ and $\mathbf{S}\mathbf{T}_y\mathbf{S}^{-1}$ are what you would expect.

Problem 3.14 Show that similarity preserves matrix multiplication (that is if $\mathbf{A}^e\mathbf{B}^e = \mathbf{C}^e$, then $\mathbf{A}^f\mathbf{B}^f = \mathbf{C}^f$). Similarity does *not*, in general, preserve symmetry, reality, or Hermiticity; show, however, that if **S** is *unitary*, and $\mathbf{H}^e$ is Hermitian, then $\mathbf{H}^f$ is Hermitian. Show that **S** carries an orthonormal basis into another orthonormal basis if and only if it is unitary.

***Problem 3.15** Prove that $\mathrm{Tr}(\mathbf{T}_1\mathbf{T}_2) = \mathrm{Tr}(\mathbf{T}_2\mathbf{T}_1)$. It follows immediately that $\mathrm{Tr}(\mathbf{T}_1\mathbf{T}_2\mathbf{T}_3) = \mathrm{Tr}(\mathbf{T}_2\mathbf{T}_3\mathbf{T}_1)$, but is it the case that $\mathrm{Tr}(\mathbf{T}_1\mathbf{T}_2\mathbf{T}_3) = \mathrm{Tr}(\mathbf{T}_2\mathbf{T}_1\mathbf{T}_3)$, in general? Prove it, or disprove it. *Hint*: The best disproof is always a counterexample—and the simpler the better!

Problem 3.16 Show that the rows and columns of a unitary matrix constitute orthonormal sets.

3.1.4 Eigenvectors and Eigenvalues

Consider the linear transformation in three-space consisting of a rotation, about some specified axis, by an angle θ. Most vectors will change in a rather complicated way (they ride around on a cone about the axis), but vectors that happen to lie *along* the axis have very simple behavior: They don't change at all ($\hat{T}|\alpha\rangle = |\alpha\rangle$). If θ is 180°, then vectors which lie in the the "equatorial" plane reverse signs ($\hat{T}|\alpha\rangle = -|\alpha\rangle$). In a complex vector space,[11] *every* linear transformation has "special" vectors like these, which are transformed into simple multiples of themselves:

$$\hat{T}|\alpha\rangle = \lambda|\alpha\rangle; \qquad [3.68]$$

---

[11] This is *not* always true in a *real* vector space (where the scalars are restricted to real values). See Problem 3.17.

they are called **eigenvectors** of the transformation, and the (complex) number λ is their **eigenvalue**. (The *null* vector doesn't count, even though, in a trivial sense, it obeys Equation 3.68 for *any* $\hat{T}$ and *any* λ; technically, an eigenvector is any *nonzero* vector satisfying Equation 3.68.) Notice that any (nonzero) multiple of an eigenvector is still an eigenvector with the same eigenvalue.

With respect to a particular basis, the eigenvector equation assumes the matrix form

$$\mathbf{Ta} = \lambda \mathbf{a} \qquad [3.69]$$

(for nonzero $\mathbf{a}$), or

$$(\mathbf{T} - \lambda \mathbf{1})\mathbf{a} = \mathbf{0}. \qquad [3.70]$$

(Here $\mathbf{0}$ is the **zero matrix**, whose elements are all zero.) Now, if the matrix $(\mathbf{T} - \lambda \mathbf{1})$ had an *inverse*, we could multiply both sides of Equation 3.70 by $(\mathbf{T} - \lambda \mathbf{1})^{-1}$, and conclude that $\mathbf{a} = \mathbf{0}$. But by assumption $\mathbf{a}$ is *not* zero, so the matrix $(\mathbf{T} - \lambda \mathbf{1})$ must in fact be singular, which means that its determinant vanishes:

$$\det(\mathbf{T} - \lambda \mathbf{1}) = \begin{vmatrix} (T_{11} - \lambda) & T_{12} & \cdots & T_{1n} \\ T_{21} & (T_{22} - \lambda) & \cdots & T_{2n} \\ \vdots & \vdots & & \vdots \\ T_{n1} & T_{n2} & \cdots & (T_{nn} - \lambda) \end{vmatrix} = 0. \qquad [3.71]$$

Expansion of the determinant yields an algebraic equation for λ:

$$C_n \lambda^n + C_{n-1}\lambda^{n-1} + \cdots + C_1 \lambda + C_0 = 0, \qquad [3.72]$$

where the coefficients C_i depend on the elements of $\mathbf{T}$ (see Problem 3.19). This is called the **characteristic equation** for the matrix; its solutions determine the eigenvalues. Notice that it's an nth-order equation, so it has n (complex) roots.[12] However, some of these may be duplicates, so all we can say for certain is that an $n \times n$ matrix has *at least one* and *at most* n distinct eigenvalues. To construct the corresponding eigenvectors it is generally easiest simply to plug each λ back into Equation 3.69 and solve "by hand" for the components of $\mathbf{a}$. I'll show you how it goes by working out an example.

Example. Find the eigenvalues and eigenvectors of the following matrix:

$$\mathbf{M} = \begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix}. \qquad [3.73]$$

The characteristic equation is

$$\begin{vmatrix} (2 - \lambda) & 0 & -2 \\ -2i & (i - \lambda) & 2i \\ 1 & 0 & (-1 - \lambda) \end{vmatrix} = -\lambda^3 + (1 + i)\lambda^2 - i\lambda = 0, \qquad [3.74]$$

---

[12] It is here that the case of *real* vector spaces becomes more awkward, because the characteristic equation need not have any (real) solutions at all. See footnote 11 and Problem 3.17.

and its roots are 0, 1, and i. Call the components of the first eigenvector (a_1, a_2, a_3); then

$$\begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = 0 \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

which yields three equations:

$$\begin{aligned} 2a_1 - 2a_3 &= 0, \\ -2ia_1 + ia_2 + 2ia_3 &= 0, \\ a_1 - a_3 &= 0. \end{aligned}$$

The first determines a_3 (in terms of a_1): $a_3 = a_1$; the second determines a_2: $a_2 = 0$; and the third is redundant. We may as well pick $a_1 = 1$ (since any multiple of an eigenvector is still an eigenvector):

$$\mathbf{a}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \text{ for } \lambda_1 = 0. \qquad [3.75]$$

For the second eigenvector (recycling the same notation for the components) we have

$$\begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = 1 \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix},$$

which leads to the equations

$$\begin{aligned} 2a_1 - 2a_3 &= a_1, \\ -2ia_1 + ia_2 + 2ia_3 &= a_2, \\ a_1 - a_3 &= a_3, \end{aligned}$$

with the solution $a_3 = (1/2)a_1$, $a_2 = [(1 - i)/2]a_1$; this time we'll pick $a_1 = 2$, so that

$$\mathbf{a}^{(2)} = \begin{pmatrix} 2 \\ (1 - i) \\ 1 \end{pmatrix}, \text{ for } \lambda_2 = 1. \qquad [3.76]$$

Finally, for the third eigenvector,

$$\begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = i \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} ia_1 \\ ia_2 \\ ia_3 \end{pmatrix},$$

which gives the equations

$$\begin{aligned} 2a_1 - 2a_3 &= ia_1, \\ -2ia_1 + ia_2 + 2ia_3 &= ia_2, \\ a_1 - a_3 &= ia_3, \end{aligned}$$

whose solution is $a_3 = a_1 = 0$, with a_2 undetermined. Choosing $a_2 = 1$, we conclude

$$\mathbf{a}^{(3)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \text{ for } \lambda_3 = i. \qquad [3.77]$$

If the eigenvectors span the space (as they do in the preceding example), we are free to use *them* as a basis:

$$\begin{aligned} \hat{T}|f_1\rangle &= \lambda_1|f_1\rangle, \\ \hat{T}|f_2\rangle &= \lambda_2|f_2\rangle, \\ &\cdots \\ \hat{T}|f_n\rangle &= \lambda_n|f_n\rangle. \end{aligned}$$

The matrix representing $\hat{T}$ takes on a very simple form in this basis, with the eigenvalues strung out along the main diagonal and all other elements zero:

$$\mathbf{T} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}. \qquad [3.78]$$

The (normalized) eigenvectors are equally simple:

$$\mathbf{a}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{a}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots \quad, \mathbf{a}^{(n)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \qquad [3.79]$$

A matrix that can be brought to **diagonal form** (Equation 3.78) by a change of basis is said to be **diagonalizable**. The similarity matrix that accomplishes the transformation can be constructed by using the eigenvectors (in the old basis) as the columns of $\mathbf{S}^{-1}$:

$$(\mathbf{S}^{-1})_{ij} = (\mathbf{a}^{(j)})_i. \qquad [3.80]$$

Example (cont'd). In the example,

$$\mathbf{S}^{-1} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & (1-i) & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

so (using Equation 3.57)

$$\mathbf{S} = \begin{pmatrix} -1 & 0 & 2 \\ 1 & 0 & -1 \\ (i-1) & 1 & (1-i) \end{pmatrix},$$

and you can check for yourself that

$$\mathbf{SMS}^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & i \end{pmatrix}.$$

There is a great advantage in bringing a matrix to diagonal form—it's much easier to work with. Unfortunately, not every matrix *can* be diagonalized—the eigenvectors have to span the space. For an example of a matrix that *cannot* be diagonalized, see Problem 3.18.

∗**Problem 3.17** The 2×2 matrix representing a rotation of the xy-plane is

$$\mathbf{T} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Show that (except for certain special angles—what are they?) this matrix has no real eigenvalues. (This reflects the *geometrical* fact that no vector in the plane is carried into itself under such a rotation; contrast rotations in *three* dimensions.) This matrix *does*, however, have *complex* eigenvalues and eigenvectors. Find them. Construct a matrix $\mathbf{S}$ which diagonalizes $\mathbf{T}$. Perform the similarity transformation $(\mathbf{STS}^{-1})$ explicitly, and show that it reduces $\mathbf{T}$ to diagonal form.

Problem 3.18 Find the eigenvalues and eigenvectors of the following matrix:

$$\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Can this matrix be diagonalized?

Problem 3.19 Show that the first, second, and last coefficients in the characteristic equation (Equation 3.72) are

$$C_n = (-1)^n, \quad C_{n-1} = (-1)^{n-1}\mathrm{Tr}(\mathbf{T}), \quad \text{and } C_0 = \det(\mathbf{T}). \qquad [3.81]$$

For a 3×3 matrix with elements T_{ij}, what is C_1?

Problem 3.20 It is pretty obvious that the trace of a *diagonal* matrix is the sum of its eigenvalues, and its determinant is their product (see Equation 3.78). It follows (from Equations 3.64 and 3.67) that the same holds for any *diagonalizable* matrix. Prove that

$$\det(\mathbf{T}) = \lambda_1\lambda_2\cdots\lambda_n, \quad \mathrm{Tr}(\mathbf{T}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n \qquad [3.82]$$

for *any* matrix. (The λ's are the n solutions to the characteristic equation—in the case of multiple roots, there may be fewer linearly independent eigen*vectors* than there

are solutions, but we still count each λ as many times as it occurs.) *Hint*: Write the characteristic equation in the form

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda) = 0,$$

and use the result of Problem 3.19.

3.1.5 Hermitian Transformations

In Equation 3.48 I defined the Hermitian conjugate (or "adjoint") of a *matrix* as its transpose conjugate: $\mathbf{T}^\dagger = \tilde{\mathbf{T}}^*$. Now I want to give you a more fundamental definition for the Hermitian conjugate of a *linear transformation*: It is that transformation $\hat{T}^\dagger$ which, when applied to the *first* member of an inner product, gives the same result as if $\hat{T}$ itself had been applied to the *second* vector:

$$\langle \hat{T}^\dagger \alpha | \beta \rangle = \langle \alpha | \hat{T} \beta \rangle \qquad [3.83]$$

(for all vectors $|\alpha\rangle$ and $|\beta\rangle$).[13] [I have to warn you that although everybody uses it, this is lousy notation. For α and β are not *vectors* (the *vectors* are $|\alpha\rangle$ and $|\beta\rangle$), they are *labels*—serial numbers ("F43A-9GT"), or names ("Charlie"), or bar codes—anything you care to use to identify the different vectors. In particular, they are endowed with no mathematical properties at all, and the expression "$\hat{T}\beta$" is literally *nonsense*: linear transformations act on *vectors*, not *labels*. But it's pretty clear what the notation *means*: $|\hat{T}\beta\rangle$ means $\hat{T}|\beta\rangle$, and $\langle \hat{T}^\dagger \alpha | \beta \rangle$ means the inner product of the vector $\hat{T}^\dagger|\alpha\rangle$ with the vector $|\beta\rangle$. Notice in particular that

$$\langle \alpha | c\beta \rangle = c \langle \alpha | \beta \rangle, \qquad [3.84]$$

but

$$\langle c\alpha | \beta \rangle = c^* \langle \alpha | \beta \rangle \qquad [3.85]$$

for any scalar c.] If you're working in an orthonormal basis (as we always shall), the Hermitian conjugate of a linear transformation is represented by the Hermitian conjugate of the corresponding matrix (so the terminology is consistent); for (using Equations 3.50 and 3.53),

$$\langle \alpha | \hat{T} \beta \rangle = \mathbf{a}^\dagger \mathbf{T} \mathbf{b} = (\mathbf{T}^\dagger \mathbf{a})^\dagger \mathbf{b} = \langle \hat{T}^\dagger \alpha | \beta \rangle. \qquad [3.86]$$

In quantum mechanics, a fundamental role is played by Hermitian transformations ($\hat{T}^\dagger = \hat{T}$). The eigenvectors and eigenvalues of a Hermitian transformation have three crucial properties:

---

[13] If you're wondering whether such a transformation necessarily *exists*, you should have been a math major. Still, it's a good question, and the answer is yes. See, for instance, Halmos, (footnote 1), Section 44.

1. The eigenvalues of a Hermitian transformation are real.

Proof: Let λ be an eigenvalue of $\hat{T}$: $\hat{T}|\alpha\rangle = \lambda|\alpha\rangle$, with $|\alpha\rangle \neq |0\rangle$. Then

$$\langle\alpha|\hat{T}\alpha\rangle = \langle\alpha|\lambda\alpha\rangle = \lambda\langle\alpha|\alpha\rangle.$$

Meanwhile, if $\hat{T}$ is Hermitian, then

$$\langle\alpha|\hat{T}\alpha\rangle = \langle\hat{T}\alpha|\alpha\rangle = \langle\lambda\alpha|\alpha\rangle = \lambda^*\langle\alpha|\alpha\rangle.$$

But $\langle\alpha|\alpha\rangle \neq 0$ (Equation 3.20), so $\lambda = \lambda^*$, and hence λ is real. QED

2. The eigenvectors of a Hermitian transformation belonging to distinct eigenvalues are orthogonal.

Proof: Suppose $\hat{T}|\alpha\rangle = \lambda|\alpha\rangle$ and $\hat{T}|\beta\rangle = \mu|\beta\rangle$, with $\lambda \neq \mu$. Then

$$\langle\alpha|\hat{T}\beta\rangle = \langle\alpha|\mu\beta\rangle = \mu\langle\alpha|\beta\rangle,$$

and if $\hat{T}$ is Hermitian,

$$\langle\alpha|\hat{T}\beta\rangle = \langle\hat{T}\alpha|\beta\rangle = \langle\lambda\alpha|\beta\rangle = \lambda^*\langle\alpha|\beta\rangle.$$

But $\lambda = \lambda^*$ (from property 1), and $\lambda \neq \mu$, by assumption, so $\langle\alpha|\beta\rangle = 0$. QED

3. The eigenvectors of a Hermitian transformation span the space.

Comment: If all the n roots of the characteristic equation are distinct, then (by property 2) we have n mutually orthogonal eigenvectors, so they *obviously* span the space. But what if there are duplicate roots (or, as they are called, in this context, **degenerate** eigenvalues)? Suppose λ is m-fold degenerate; any *linear combination* of two eigenvectors belonging to the same eigenvalue is still an eigenvector (with the same eigenvalue)—what we must show is that there are m *linearly independent* eigenvectors with eigenvalue λ. The proof is given in most books on linear algebra,[14] and I shall not repeat it here. These eigenvectors can be orthogonalized by the Gram-Schmidt procedure (see Problem 3.4), so in fact the eigenvectors of a Hermitian transformation can always be taken to constitute an orthonormal basis. It follows, in particular, that *any Hermitian matrix can be diagonalized by a similarity transformation, with* $\mathbf{S}$ *unitary*. This rather technical result is, in a sense, the mathematical support on which much of quantum mechanics leans. As we shall see, it turns out to be a thinner reed than one might have hoped.

[14] I like the treatment in F. W. Byron, Jr., and R. W. Fuller, *Mathematics of Classical and Quantum Physics* (Reading, MA: Addison-Wesley, 1969), Vol. I, Section 4.7.

Problem 3.21 A Hermitian linear transformation must satisfy $\langle\alpha|\hat{T}\beta\rangle = \langle\hat{T}\alpha|\beta\rangle$ for all vectors $|\alpha\rangle$ and $|\beta\rangle$. Prove that it is (surprisingly) sufficient that $\langle\gamma|\hat{T}\gamma\rangle = \langle\hat{T}\gamma|\gamma\rangle$ for all vectors $|\gamma\rangle$. Suppose you could show that $\langle e_n|\hat{T}e_n\rangle = \langle\hat{T}e_n|e_n\rangle$ for every member of an orthonormal basis. Does it necessarily follow that $\hat{T}$ is Hermitian? *Hint*: First let $|\gamma\rangle = |\alpha\rangle + |\beta\rangle$, and then let $|\gamma\rangle = |\alpha\rangle + i|\beta\rangle$.

*__Problem 3.22__ Let

$$\mathbf{T} = \begin{pmatrix} 1 & 1-i \\ 1+i & 0 \end{pmatrix}.$$

(a) Verify that **T** is Hermitian.

(b) Find its eigenvalues (note that they are real).

(c) Find and normalize the eigenvectors (note that they are orthogonal).

(d) Construct the unitary diagonalizing matrix **S**, and check explicitly that it diagonalizes **T**.

(e) Check that det(**T**) and Tr(**T**) are the same for **T** as they are for its diagonalized form.

**__Problem 3.23__ Consider the following Hermitian matrix:

$$\mathbf{T} = \begin{pmatrix} 2 & i & 1 \\ -i & 2 & i \\ 1 & -i & 2 \end{pmatrix}.$$

(a) Calculate det(**T**) and Tr(**T**).

(b) Find the eigenvalues of **T**. Check that their sum and product are consistent with (a), in the sense of Equation 3.82. Write down the diagonalized version of **T**.

(c) Find the eigenvectors of **T**. Within the degenerate sector, construct two linearly independent eigenvectors (it is this step that is always possible for a *Hermitian* matrix, but not for an *arbitrary* matrix—contrast Problem 3.18). Orthogonalize them, and check that both are orthogonal to the third. Normalize all three eigenvectors.

(d) Construct the unitary matrix **S** that diagonalizes **T**, and show explicitly that the similarity transformation using **S** reduces **T** to the appropriate diagonal form.

Problem 3.24 A *unitary* linear transformation is one for which $\hat{U}^\dagger\hat{U} = 1$.

(a) Show that unitary transformations preserve inner products, in the sense that $\langle\hat{U}\alpha|\hat{U}\beta\rangle = \langle\alpha|\beta\rangle$, for all vectors $|\alpha\rangle$, $|\beta\rangle$.

(b) Show that the eigenvalues of a unitary transformation have modulus 1.

(c) Show that the eigenvectors of a unitary transformation belonging to distinct eigenvalues are orthogonal.

3.2 FUNCTION SPACES

We are ready now to apply the machinery of linear algebra to the interesting and important case of **function spaces**, in which the "vectors" are (complex) functions of x, inner products are integrals, and derivatives appear as linear transformations.

3.2.1 Functions as Vectors

Do functions really behave as vectors? Well, is the sum of two functions a function? Sure. Is addition of functions commutative and associative? Indeed. Is there a "null" function? Yes: $f(x) \equiv 0$. If you multiply a function by a complex number, do you get another function? Of course. Now, the set of *all* functions is a bit unwieldy—we'll be concerned with special *classes* of functions, such as the set of all polynomials of degree $< N$ (Problem 3.2), or the set of all odd functions that go to zero at $x = 1$, or the set of all periodic functions with period π. Of course, when you start imposing conditions like this, you've got to make sure that you still meet the requirements for a vector space. For example, the set of all functions whose maximum value is 3 would *not* constitute a vector space (multiplication by 2 would give you functions with maximum value 6, which are outside the space).

The inner product of two functions [$f(x)$ and $g(x)$] is defined by the integral

$$\boxed{\langle f|g\rangle = \int f(x)^* g(x)\, dx} \qquad [3.87]$$

(the limits will depend on the domain of the functions in question). You can check for yourself that it satisfies the three conditions (Equations 3.19, 3.20, and 3.21) for an inner product. Of course, this integral may not *converge*, so if we want a function space with an inner product, we must restrict the class of functions so as to ensure that $\langle f|g\rangle$ is always well defined. It is clearly *necessary* that every admissible function be **square integrable**:

$$\int |f(x)|^2\, dx < \infty \qquad [3.88]$$

(otherwise the inner product of f with *itself* wouldn't even exist). As it turns out,

Table 3.1: The first few Legendre polynomials, $P_n(x)$.

$P_0 = 1$
$P_1 = x$
$P_2 = \frac{1}{2}(3x^2 - 1)$
$P_3 = \frac{1}{2}(5x^3 - 3x)$
$P_4 = \frac{1}{8}(35x^4 - 30x^2 + 3)$
$P_5 = \frac{1}{8}(63x^5 - 70x^3 + 15x)$

this restriction is also *sufficient*—if f and g are both square integrable, then the integral in Equation 3.87 is necessarily finite.[15]

For example, consider the set $P(N)$ of all *polynomials* of degree $< N$:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{N-1} x^{N-1}, \qquad [3.89]$$

on the interval $-1 \leq x \leq 1$. They are certainly square integrable, so this is a bona fide inner product space. An obvious basis is the set of powers of x:

$$|e_1\rangle = 1, \ |e_2\rangle = x, \ |e_3\rangle = x^2, \ \ldots, \ |e_N\rangle = x^{N-1}; \qquad [3.90]$$

evidently it's an N-dimensinal vector space. This is *not*, however, an *orthonormal* basis, for

$$\langle e_1|e_1\rangle = \int_{-1}^{1} 1\,dx = 2, \quad \langle e_1|e_3\rangle = \int_{-1}^{1} x^2\,dx = 2/3,$$

and so on. If you apply the Gram-Schmidt procedure, to orthonormalize this basis (Problem 3.25), you get the famous **Legendre polynomials**, $P_n(x)$ (except that Legendre, who had other things on his mind, didn't normalize them properly):

$$|e_n'\rangle = \sqrt{n - (1/2)}\,P_{n-1}(x), \quad (n = 1, 2, \ldots, N). \qquad [3.91]$$

In Table 3.1 I have listed the first few Legendre polynomials.

***Problem 3.25** Orthonormalize the powers of x, on the interval $-1 \leq x \leq 1$, to obtain the first four Legendre polynomials (Equation 3.91).

***Problem 3.26** Let $T(N)$ be the set of all trigonometric functions of the form

$$f(x) = \sum_{n=0}^{N-1} [a_n \sin(n\pi x) + b_n \cos(n\pi x)], \qquad [3.92]$$

[15] There is a quick phoney "proof" of this, based on the Schwarz inequality (Equation 3.27). The trouble is, we *assumed* the existence of the inner product in proving the Schwarz inequality (Problem 3.5), so the logic is circular. For a legitimate proof, see F. Riesz and B. Sz.-Nagy, *Functional Analysis* (New York: Unger, 1955), Section 21.

on the interval $-1 \leq x \leq 1$. Show that

$$|e_n\rangle = \frac{1}{\sqrt{2}} e^{in\pi x}, \quad (n = 0, \pm 1, \ldots, \pm(N-1)) \tag{3.93}$$

constitutes an orthonormal basis. What is the dimension of this space?

Problem 3.27 Consider the set of all functions of the form $p(x)e^{-x^2/2}$, where $p(x)$ is again a polynomial of degree $< N$ in x, on the interval $-\infty < x < \infty$. Check that they constitute an inner product space. The "natural" basis is

$$|e_1\rangle = e^{-x^2/2}, \; |e_2\rangle = xe^{-x^2/2}, \; |e_3\rangle = x^2 e^{-x^2/2}, \; \ldots, \; |e_N\rangle = x^{N-1} e^{-x^2/2}.$$

Orthonormalize the first four of these, and comment on the result.

3.2.2 Operators as Linear Transformations

In function spaces *operators* (such as d/dx, d^2/dx^2, or simply x) behave as linear transformations, provided that they carry functions in the space into other functions in the space and satisfy the linearity condition (Equation 3.29). For example, in the polynomial space $P(N)$ the derivative operator ($\hat{D} \equiv d/dx$) is a linear transformation, but the operator $\hat{x}$ (multiplication by x)[16] is *not*, for it takes $(N-1)$th-order polynomials into Nth-order polynomials, which are no longer in the space.

In a function space, the eigenvectors of an operator $\hat{T}$ are called **eigenfunctions**:

$$\hat{T}f(x) = \lambda f(x). \tag{3.94}$$

For example, the eigenfunctions of $\hat{D}$ are

$$f_\lambda(x) = Ae^{\lambda x}. \tag{3.95}$$

Evidently this operator has only *one* eigenfunction (the one with $\lambda = 0$) in the space $P(N)$.

A Hermitian operator is one that satisfies the defining condition (Equation 3.83):

$$\langle f|\hat{T}g\rangle = \langle \hat{T}f|g\rangle, \tag{3.96}$$

for all functions $f(x)$ and $g(x)$ in the space. Is the derivative operator Hermitian? Well, using integration by parts, we get

$$\langle f|\hat{D}g\rangle = \int_a^b f^* \frac{dg}{dx}\, dx = (f^*g)\Big|_a^b - \int_a^b \frac{df^*}{dx} g\, dx = (f^*g)\Big|_a^b - \langle \hat{D}f|g\rangle. \tag{3.97}$$

---

[16]For consistency, I'll put a hat on x when I'm emphasizing its role as an operator, but you're welcome to ignore it if you think I'm being too fastidious.

It's *close*, but the sign is wrong, and there's an unwanted boundary term. The sign is easily disposed of: $\hat{D}$ itself is (except for the boundary term) *skew* Hermitian, so $i\hat{D}$ would be Hermitian—complex conjugation of the i compensates for the minus sign coming from integration by parts. As for the boundary term, it will go away if we restrict ourselves to functions which have the same value at the two ends:

$$f(b) = f(a). \qquad [3.98]$$

In practice, we shall almost always be working on the *infinite* interval ($a = -\infty$, $b = +\infty$), where square integrability (Equation 3.88) guarantees that $f(a) = f(b) = 0$, and hence that $i\hat{D}$ is Hermitian. But $i\hat{D}$ is *not* Hermitian in the polynomial space $P(N)$.

By now you will realize that when dealing with operators you must always keep in mind the function space you're working in—an innocent-looking operator may not be a legitimate linear transformation, because it carries functions out of the space; the eigenfunctions of an operator may not reside in the space; and an operator that's Hermitian in one space may *not* be Hermitian in another. However, these are relatively harmless problems—they can startle you, if you're not expecting them, but they don't bite. A much more dangerous snake is lurking here, but it only inhabits vector spaces of infinite dimension. I noted a moment ago that $\hat{x}$ is not a linear transformation in the space $P(N)$ (multiplication by x increases the order of the polynomial and hence takes functions outside the space). However, it *is* a linear transformation on $P(\infty)$, the space of *all* polynomials on the interval $-1 \leq x \leq 1$. In fact, it's a *Hermitian* transformation, since (obviously)

$$\int_{-1}^{1} [f(x)]^*[xg(x)]\,dx = \int_{-1}^{1} [xf(x)]^*[g(x)]\,dx.$$

But what are its eigenfunctions? Well,

$$x(a_0 + a_1x + a_2x^2 + \cdots) = \lambda(a_0 + a_1x + a_2x^2 + \cdots),$$

for all x, means

$$\begin{aligned} 0 &= \lambda a_0, \\ a_0 &= \lambda a_1, \\ a_1 &= \lambda a_2, \end{aligned}$$

and so on. If $\lambda = 0$, then all the components are zero, and that's not a legal eigenvector; but if $\lambda \neq 0$, the first equation says $a_0 = 0$, so the second gives $a_1 = 0$, and the third says $a_2 = 0$, and so on, and we're back in the same bind. *This* Hermitian operator doesn't have a complete set of eigenfunctions—in fact it doesn't have *any at all!* Not, at any rate, in $P(\infty)$.

What *would* an eigenfunction of $\hat{x}$ look like? If

$$xg(x) = \lambda g(x), \qquad [3.99]$$

where λ, remember, is a constant, then everywhere *except* at the one point $x = \lambda$ we must have $g(x) = 0$. Evidently the eigenfunctions of $\hat{x}$ are Dirac delta functions:

$$g_\lambda(x) = B\delta(x - \lambda), \qquad [3.100]$$

and since delta functions are certainly not polynomials, it is no wonder that the operator $\hat{x}$ has no eigenfunctions in $P(\infty)$.

The moral of the story is that whereas the first two theorems in Section 3.1.5 are completely general (the eigenvalues of a Hermitian operator are real, and the eigenvectors belonging to different eigenvalues are orthogonal), the third one (completeness of the eigenvectors) is valid (in general) only for finite-dimensional spaces. In infinite-dimensional spaces some Hermitian operators have complete sets of eigenvectors (see Problem 3.32d for an example), some have *in*complete sets, and some (as we just saw) have no eigenvectors (in the space) at all.[17] Unfortunately, the completeness property is absolutely essential in quantum mechanical applications. In Section 3.3 I'll show you how we manage this problem.

Problem 3.28 Show that $\exp(-x^2/2)$ is an eigenfunction of the operator $\hat{Q} = (d^2/dx^2) - x^2$, and find its eigenvalue.

***Problem 3.29**

(a) Construct the matrix **D** representing the derivative operator $\hat{D} = d/dx$ with respect to the (nonorthonormal) basis (Equation 3.90) in $P(N)$.

(b) Construct the matrix representing $\hat{D}$ with respect to the (orthonormal) basis (Equation 3.93) in the space $T(N)$ of Problem 3.26.

(c) Construct the matrix **X** representing the operator $\hat{x} = x$ with respect to the basis (Equation 3.90) in $P(\infty)$. If this is a Hermitian operator (and it *is*), how come the matrix is not equal to its transpose conjugate?

****Problem 3.30** Construct the matrices **D** and **X** in the (orthonormal) basis (Equation 3.91) for $P(\infty)$. You will need to use two recursion formulas for Legendre polynomials:

$$xP_n(x) = \frac{1}{(2n+1)}[(n+1)P_{n+1}(x) + nP_{n-1}(x)]; \qquad [3.101]$$

$$\frac{dP_n}{dx} = \sum_{k=0}(2n - 4k - 1)P_{n-2k-1}(x), \qquad [3.102]$$

---

[17] In an n-dimensional vector space, every linear transformation can be represented (with respect to a particular basis) by an $n \times n$ matrix, and as long as n is finite, the characteristic Equation 3.71 is guaranteed to deliver at least one eigenvalue. But if n is *infinite*, we can't take the determinant, there *is* no characteristic equation, and hence there is no assurance that even a *single* eigenvector exists.

where the sum cuts off at the first term with a negative index. Confirm that **X** is Hermitian but $i\mathbf{D}$ is not.

Problem 3.31 Consider the operator $\hat{D}^2 = d^2/dx^2$. Under what conditions (on the admissable functions) is it a Hermitian operator? Construct the matrix representing $\hat{D}^2$ in $P(N)$ (with respect to the basis Equation 3.90), and confirm that it is the square of the matrix representing $\hat{D}$ (Problem 3.29a).

Problem 3.32

(a) Show that $i\hat{D}$ is Hermitian in the space $T(N)$ of Problem 3.26.

(b) What are its eigenvalues and (normalized) eigenfunctions, in $T(N)$?

(c) Check that your results in (b) satisfy the three theorems in Section 3.1.5.

(d) Confirm that $i\hat{D}$ has a complete set of eigenfunctions in $T(\infty)$ (quote the pertinent theorem from Fourier analysis).

3.2.3 Hilbert Space

To construct the real number system, mathematicians typically begin with the *integers*, and use them to define the *rationals* (ratios of integers). They proceed to show that the rational numbers are "dense," in the sense that between any two of them (no matter how close together they are) you can always find another one (in fact, infinitely many of them). And yet, the set of all rational numbers has "gaps" in it, for you can easily think of infinite *sequences* of rational numbers whose *limit* is *not* a rational number. For example,

$$A_N = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots \pm \frac{1}{N} \qquad [3.103]$$

is a rational number for any finite integer N, but its limit (as $N \to \infty$) is $\ln 2$, which is *not* a rational number. So the final step in constructing the real numbers is to "fill in the gaps", or "complete" the set, by including the limits of all convergent sequences of rational numbers. (Of course, some sequences don't *have* limits, and those we do not include. For example, if you change the minus signs in Equation 3.103 to plus signs, the sequence does not converge, and it doesn't correspond to *any* real number.)

The same thing happens with function spaces. For example, the set of all polynomials, $P(\infty)$, includes functions of the form

$$f_N(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots + \frac{x^N}{N!} \qquad [3.104]$$

(for finite N), but it does *not* include the limit as $N \to \infty$:

$$1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x. \qquad [3.105]$$

For e^x is not itself a polynomial, although it *is* the limit of a *sequence* of polynomials. To *complete* the space, we would like to include all such functions. Of course, some sequences of polynomials don't have limits, or have them only for a restricted range of x. For example, the series

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1 - x}$$

converges only for $|x| < 1$. And even if the sequence does have a limit, the limit function may not be square integrable, so we can't include it in an inner product space. To complete the space, then, we throw in all square-integrable convergent sequences of functions in the space. Notice that completing a space does not involve the introduction of any new basis vectors; it is just that we now allow linear combinations involving an infinite number of terms,

$$|\alpha\rangle = \sum_{j=1}^{\infty} a_j |e_j\rangle, \qquad [3.106]$$

provided $\langle\alpha|\alpha\rangle$ is finite—which is to say (if the basis is orthonormal), provided

$$\sum_{j=1}^{\infty} |a_j|^2 < \infty. \qquad [3.107]$$

A **complete**[18] inner product space is called a **Hilbert space**.[19] The completion of $P(\infty)$ is easy to characterize: It is nothing less than the set of *all* square-integrable functions on the interval $-1 < x < +1$; we call it $L_2(-1, +1)$. More generally, the set of all square-integrable functions on the interval $a < x < b$ is $L_2(a, b)$. We shall be concerned primarily with the Hilbert space $L_2(-\infty, +\infty)$ (or L_2, for short), because this is where quantum mechanical wave functions live. Indeed, to physicists L_2 is practically synonymous with "Hilbert space".

The eigenfunctions of the Hermitian operators $i\hat{D} = id/dx$ and $\hat{x} = x$ are of particular importance. As we have already found (Equations 3.95 and 3.100), they take the form

$$f_\lambda(x) = A_\lambda e^{-i\lambda x}, \quad \text{and} \quad g_\lambda(x) = B_\lambda \delta(x - \lambda),$$

respectively. Note that there is no restriction on the eigenvalues—*every* real number is an eigenvalue of $i\hat{D}$, and *every* real number is an eigenvalue of $\hat{x}$. The set of all eigenvalues of a given operator is called its **spectrum**; $i\hat{D}$ and $\hat{x}$ are operators with **continuous spectra**, in contrast to the **discrete spectra** we have encountered

[18] Note the two entirely different uses of the word "complete": a *set of vectors* is complete if it spans the space; an *inner product space* is complete if it has no "holes" in it (i.e., it includes all its limits).

[19] Every finite-dimensional inner product space is trivially complete, so they're all technically Hilbert spaces, but the term is usually reserved for *infinite*-dimensional spaces.

heretofore. Unfortunately, these eigenfunctions *do not lie in Hilbert space*, and hence, in the strictest sense, do not count as vectors at all. For neither of them is square-integrable:

$$\int_{-\infty}^{\infty} f_\lambda(x)^* f_\lambda(x)\,dx = |A_\lambda|^2 \int_{-\infty}^{\infty} e^{i\lambda x} e^{-i\lambda x}\,dx = |A_\lambda|^2 \int_{-\infty}^{\infty} 1\,dx \to \infty,$$

and

$$\int_{-\infty}^{\infty} g_\lambda(x)^* g_\lambda(x)\,dx = |B_\lambda|^2 \int_{-\infty}^{\infty} \delta(x-\lambda)\delta(x-\lambda)\,dx = |B_\lambda|^2 \delta(\lambda-\lambda) \to \infty.$$

Nevertheless, they do satisfy a kind of orthogonality condition:

$$\int_{-\infty}^{\infty} f_\lambda(x)^* f_\mu(x)\,dx = A_\lambda^* A_\mu \int_{-\infty}^{\infty} e^{i\lambda x} e^{-i\mu x}\,dx = |A_\lambda|^2 2\pi\delta(\lambda-\mu)$$

(see Equation 2.126), and

$$\int_{-\infty}^{\infty} g_\lambda(x)^* g_\mu(x)\,dx = B_\lambda^* B_\mu \int_{-\infty}^{\infty} \delta(x-\lambda)\delta(x-\mu)\,dx = |B_\lambda|^2 \delta(\lambda-\mu).$$

It is customary to "normalize" these (unnormalizable) functions by picking the constant so as to leave an unadorned Dirac delta function on the right side (replacing the *Kronecker* delta in the usual orthonormality condition; Equation 3.23).[20] Thus

$$f_\lambda(x) = \frac{1}{\sqrt{2\pi}} e^{-i\lambda x}, \text{ with } \langle f_\lambda | f_\mu \rangle = \delta(\lambda-\mu), \qquad [3.108]$$

are the "normalized" eigenfunctions of $i\hat{D}$, and

$$g_\lambda(x) = \delta(x-\lambda), \text{ with } \langle g_\lambda | g_\mu \rangle = \delta(\lambda-\mu), \qquad [3.109]$$

are the "normalized" eigenfunctions of $\hat{x}$.[21]

What if we use the "normalized" eigenfunctions of $i\hat{D}$ and $\hat{x}$ as bases for L_2?[22] Because the spectrum is continuous, the linear combination becomes an integral:

$$|f\rangle = \int_{-\infty}^{\infty} a_\lambda |f_\lambda\rangle\,d\lambda; \quad |f\rangle = \int_{-\infty}^{\infty} b_\lambda |g_\lambda\rangle\,d\lambda. \qquad [3.110]$$

---

[20] I'll call this "normalization" (in quotes) so you won't confuse it with the real thing.

[21] We are engaged here in a dangerous stretching of the rules, pioneered by Dirac (who had a kind of inspired confidence that he could get away with it) and disparaged by von Neumann (who was more sensitive to mathematical niceties), in their rival classics (P. A. M. Dirac, *The Principles of Quantum Mechanics*, first published in 1930, 4th ed., Oxford (Clarendon Press) 1958, and J. von Neumann, *The Mathematical Foundations of Quantum Mechanics*, first published in 1932, revised by Princeton Univ. Press, 1955). **Dirac notation** invites us to apply the language and methods of linear algebra to functions that lie in the "almost normalizable" suburbs of Hilbert space. It turns out to be powerful and effective beyond any reasonable expectation.

[22] That's right: We're going to use, as bases, sets of functions none of which is actually in the space! They may not be normalizable, but they *are* complete, and that's all we need.

Taking the inner product with $|f_\mu\rangle$, and exploiting the "orthonormality" of the basis (Equation 3.108), we obtain the "components" a_λ:

$$\langle f_\mu | f \rangle = \int_{-\infty}^{\infty} a_\lambda \langle f_\mu | f_\lambda \rangle \, d\lambda = \int_{-\infty}^{\infty} a_\lambda \delta(\mu - \lambda) \, d\lambda = a_\mu.$$

So

$$a_\lambda = \langle f_\lambda | f \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\lambda x} f(x) \, dx = F(-\lambda); \qquad [3.111]$$

evidently the $-\lambda$ "component" of the vector $|f\rangle$, in the basis of eigenfunctions of $i\hat{D}$, is the *Fourier transform* (Equation 2.85) of the function $f(x)$. Likewise,

$$b_\lambda = \langle g_\lambda | f \rangle = \int_{-\infty}^{\infty} \delta(x - \lambda) f(x) \, dx = f(\lambda), \qquad [3.112]$$

so the λ "component" of the vector $|f\rangle$ in the *position* basis is $f(\lambda)$ itself. [If that sounds like double-talk, remember that $|f\rangle$ is an abstract vector, which can be expressed with respect to any basis you like; in this sense the *function* $f(x)$ is merely the collection of its "components" in the particular basis consisting of eigenvectors of the position operator.] Meanwhile, we can no longer represent operators by matrices because the basis vectors are labeled by a nondenumerable index. Nevertheless, we are still interested in quantities of the form

$$\langle f_\lambda | \hat{T} | f_\mu \rangle,$$

which, by force of habit, we shall call the λ, μ **matrix element** of the operator $\hat{T}$.

****Problem 3.33**

(a) Show that any linear combination of two functions in $L_2(a, b)$ is still in $L_2(a, b)$. If this *weren't* true, of course, $L_2(a, b)$ wouldn't be a vector space at all.

(b) For what range of (real) ν is the function $f(x) = |x|^\nu$ in $L_2(-1, +1)$?

(c) For what range of a is the function $f(x) = 1 - x + x^2 - x^3 + \cdots$ in $L_2(-a, +a)$?

(d) Show that the function $f(x) = e^{-|x|}$ is in L_2, and find its "components" in the basis (Equation 3.108).

(e) Find the matrix elements of the operator $\hat{D}^2$ with respect to the basis (Equation 3.108) of L_2.

Problem 3.34 $L_2(-1, +1)$ includes *discontinuous* functions [such as the step function, $\theta(x)$, Equation [2.125], which are not differentiable. But functions expressible as Taylor series ($f(x) = a_0 + a_1 x + a_2 x^2 + \cdots$) must be *infinitely* differentiable. How, then, can $\theta(x)$ be the limit of a sequence of polynomials? *Note*: This is not a difficult problem, once you see the light, but it is very subtle, so don't waste a lot of time on it if you're not getting anywhere.

3.3 THE GENERALIZED STATISTICAL INTERPRETATION

My next project is to recast the fundamental principles of quantum mechanics (as developed in Chapters 1 and 2) into the more elegant language of linear algebra. Remember that the state of a particle is represented by its wave function, $\Psi(x, t)$, whose absolute square is the probability density for finding the particle at point x, at time t. It follows that Ψ must be *normalized*, which is possible (by dividing off a constant) if and only if it is square integrable. Thus

1. The state of a particle is represented by a normalized vector ($|\Psi\rangle$) in the Hilbert space L_2.

Classical dynamical quantities (such as position, velocity, momentum and kinetic energy) can be expressed as functions of the "canonical" variables x and p (and, in rare cases, t): $Q(x, p, t)$. To each such classical observable we associate a quantum-mechanical *operator*, $\hat{Q}$, obtained from Q by the substitution

$$p \to \frac{\hbar}{i}\frac{\partial}{\partial x}. \qquad [3.113]$$

The expectation value of Q, in the state Ψ, is

$$\langle Q \rangle = \int \Psi(x,t)^* \hat{Q} \Psi(x,t)\, dx,$$

which we now write as an inner product:[23]

$$\langle Q \rangle = \langle \Psi | \hat{Q} \Psi \rangle. \qquad [3.114]$$

Now, the expectation value of an observable quantity has got to be a *real* number (after all, it corresponds to actual measurements in the laboratory, using rulers and clocks and meters), so

$$\langle \Psi | \hat{Q} \Psi \rangle = \langle \Psi | \hat{Q} \Psi \rangle^* = \langle \hat{Q} \Psi | \Psi \rangle, \qquad [3.115]$$

for all vectors $|\Psi\rangle$. It follows (see Problem 3.21) that $\hat{Q}$ must be a *Hermitian* operator. Thus

2. Observable quantities, $Q(x, p, t)$, are represented by Hermitian operators, $\hat{Q}(x, \frac{\hbar}{i}\frac{\partial}{\partial x}, t)$; the expectation value of Q, in the state $|\Psi\rangle$, is $\langle \Psi | \hat{Q} \Psi \rangle$.

[23] The "lousy notation" I warned you about on page 92 is not so bad in this context, for we are using the function Ψ *itself* to label the vector $|\Psi\rangle$, and the expression $\hat{Q}\Psi$ is perfectly self-explanatory.

In general, identical measurements on identically prepared systems (all in the same state Ψ) do not yield reproducible results; however, *some* states are **determinate**, for a particular observable, in the sense that they always give the same result. [A competent measurement of the total energy of a particle in the ground state of the harmonic oscillator, for example, will always return the value $(1/2)\hbar\omega$.] For a determinate state of observable Q, the standard deviation is zero:

$$\begin{aligned} 0 = \sigma_Q^2 &= \langle (\hat{Q} - \langle Q \rangle)^2 \rangle = \langle \Psi | (\hat{Q} - \langle Q \rangle)^2 \Psi \rangle \\ &= \langle (\hat{Q} - \langle Q \rangle)\Psi | (\hat{Q} - \langle Q \rangle)\Psi \rangle = \| (\hat{Q} - \langle Q \rangle)|\Psi\rangle \|^2. \end{aligned} \quad [3.116]$$

[I used the fact that the operator $(\hat{Q} - \langle Q \rangle)$ is Hermitian to peel it off the second member of the inner product and attach it to the first member.] But the only vector with norm zero is the *null* vector (Equation 3.20), so $(\hat{Q} - \langle Q \rangle)|\Psi\rangle = 0$, or

$$\hat{Q}|\Psi\rangle = \langle Q \rangle |\Psi\rangle. \quad [3.117]$$

Evidently determinate states are *eigenvectors* of $\hat{Q}$. Thus

3. A measurement of the observable Q on a particle in the state $|\Psi\rangle$ is *certain* to return the value λ if and only if $|\Psi\rangle$ is an eigenvector of $\hat{Q}$, with eigenvalue λ.

For example, the time-independent Schrödinger equation (Equation 2.11),

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

is nothing but an eigenvalue equation for the Hamiltonian operator, and the solutions are states of determinate energy (as we noted long ago).

Up to this point I have added nothing *new* to the statistical interpretation; I have merely explored its implications in the language of linear algebra. But there is a missing part to the story: Although we can calculate the *average* result of any measurement, we still cannot say what the probability of getting a *particular* result would be if we were to measure a given observable Q on a particle in an arbitrary state $|\Psi\rangle$ (except for the special case of *position* for which the original statistical interpretation supplies the answer). To finish the job, we need the following **generalized statistical interpretation**, which is inspired by postulate 3 above, and subsumes it as a special case:

3′. If you measure an observable Q on a particle in the state $|\Psi\rangle$, you are *certain* to get *one of the eigenvalues* of $\hat{Q}$. The probability of getting the particular eigenvalue λ is equal to the absolute square of the λ component of $|\Psi\rangle$, when expressed in the orthonormal basis of eigenvectors.[24]

[24]Notice that we could calculate from this the expectation value of Q, and it is important to check that the result is consistent with postulate 2 above. See Problem 3.35(c).

To sustain this postulate, it is essential that the eigenfunctions of $\hat{Q}$ span the space. As we have seen, in the *finite*-dimensional case the eigenvectors of a Hermitian operator *always* span the space. But this theorem fails in the infinite-dimensional case—we have encountered examples of Hermitian operators that have no eigenfunctions at all, or for which the eigenfunctions lie outside the Hilbert space. *We shall take it as a restriction on the subset of Hermitian operators that are* **observable**, *that their eigenfunctions constitute a complete set* (though they need not fall inside L_2).[25]

Now, there are two kinds of eigenvectors, which we need to treat separately. If the spectrum is *discrete* (with the distinct eigenvalues separated by finite gaps), we can label the eigenvectors with an integer n:

$$\hat{Q}|e_n\rangle = \lambda_n|e_n\rangle, \quad \text{with } n = 1, 2, 3, \ldots; \tag{3.118}$$

the eigenvectors are orthonormal (or rather, they can always be *chosen* so):

$$\langle e_n|e_m\rangle = \delta_{nm}; \tag{3.119}$$

the completeness relation takes the form of a *sum*:

$$|\Psi\rangle = \sum_{n=1}^{\infty} c_n|e_n\rangle; \tag{3.120}$$

the components are given by "Fourier's trick":

$$c_n = \langle e_n|\Psi\rangle, \tag{3.121}$$

and the probability of getting the particular eigenvalue λ_n is

$$\boxed{|c_n|^2 = |\langle e_n|\Psi\rangle|^2.} \tag{3.122}$$

On the other hand, if the spectrum is *continuous*, the eigenvectors are labeled by a continuous variable (k):

$$\hat{Q}|e_k\rangle = \lambda_k|e_k\rangle, \quad \text{with } -\infty < k < \infty; \tag{3.123}$$

the eigenfunctions are *not* normalizable (so they do not lie in L_2, and do not themselves represent possible particle states), but they satisfy a sort of "orthonormality" condition

$$\langle e_k|e_l\rangle = \delta(k - l) \tag{3.124}$$

[25] Some authors, following Dirac, take this to be an *axiom* of quantum mechanics, but it seems to me peculiar to use that term for something that is *provable* in many particular instances; I prefer to regard it as a part of what it *means* to be "observable".

(or rather, they can always be *chosen* so); the completeness relation takes the form of an *integral*:

$$|\Psi\rangle = \int_{-\infty}^{\infty} c_k |e_k\rangle \, dk; \tag{3.125}$$

the "components" are given by "Fourier's trick":

$$c_k = \langle e_k | \Psi \rangle, \tag{3.126}$$

and the probability of getting an eigenvalue in the *range* dk about λ_k is

$$\boxed{|c_k|^2 \, dk = |\langle e_k | \Psi \rangle|^2 \, dk.} \tag{3.127}$$

The generalized statistical interpretation makes no reference to the observable x; it treats *all* observables on an equal footing. But it includes the "original" form (Equation 1.3) as a special case. The "orthonormal" eigenfunctions of the position operator are

$$e_{x'}(x) = \delta(x - x'), \tag{3.128}$$

and the eigenvalue (x') can take on any value between $-\infty$ and $+\infty$. The x' "component" of $|\Psi\rangle$ is

$$c_{x'} = \langle e_{x'} | \Psi \rangle = \int_{-\infty}^{\infty} \delta(x - x') \Psi(x, t) \, dx = \Psi(x', t), \tag{3.129}$$

so the probability of finding the particle in the range dx' about x' is

$$|c_{x'}|^2 \, dx' = |\Psi(x', t)|^2 \, dx', \tag{3.130}$$

which is the original statistical interpretation of Ψ.

A more illuminating example is provided by the momentum operator. Its "orthonormal" eigenfunctions are (see Problem 3.37)

$$e_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}, \tag{3.131}$$

and the eigenvalue (p) can take on any value in the range $-\infty < p < \infty$. The p "component" of $|\Psi\rangle$ is

$$\boxed{c_p = \langle e_p | \Psi \rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \Psi(x, t) \, dx \equiv \Phi(p, t).} \tag{3.132}$$

We call $\Phi(p, t)$ the **momentum-space wave function**—it is (apart from the factors of $\hbar$) the *Fourier transform* of the "position-space" wave function $\Psi(x, t)$. Evidently the probability of getting a momentum in the range dp is

$$|\Phi(p, t)|^2 \, dp. \tag{3.133}$$

***Problem 3.35**

(a) Show that $\sum |c_n|^2 = 1$, in Equation 3.120.

(b) Show that $\int |c_k|^2\, dk = 1$, in Equation 3.125.

(c) From postulate 3′ it follows that

$$\langle Q \rangle = \sum \lambda_n |c_n|^2, \text{ or } \langle Q \rangle = \int \lambda_k |c_k|^2\, dk, \qquad [3.134]$$

for discrete and continuous spectra, respectively. Show that this is consistent with postulate 2: $\langle Q \rangle = \langle \Psi | \hat{Q} \Psi \rangle$.

***Problem 3.36**

(a) Refer to Problem 2.6. If you measured the energy of this particle, what values might you get, and what is the probability of each? Use the answer to calculate the expectation value of H, and compare the answer you got before.

(b) Do the same for Problem 2.8. *Hint*: To sum the series, look in a math table under "Sums of Reciprocal Powers" or "Riemann Zeta Function."

Problem 3.37 Confirm that $e_p(x)$ (in Equation 3.131) is the "orthonormal" eigenfunction of the momentum operator, with eigenvalue p.

Problem 3.38 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.

3.4 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 shall prove a more general version of the uncertainty principle and explore some of its implications. The argument is beautiful, but rather abstract, so watch closely.

3.4.1 Proof of the Generalized Uncertainty Principle

For any observable A, we have (quoting Equation 3.116)

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

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

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

Therefore (invoking the Schwarz inequality, Equation 3.27),

$$\sigma_A^2\sigma_B^2 = \langle f|f\rangle\langle g|g\rangle \geq |\langle f|g\rangle|^2. \qquad [3.135]$$

Now, for any complex number z,

$$|z|^2 = (\mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2 \geq (\mathrm{Im}(z))^2 = [\frac{1}{2i}(z - z^*)]^2. \qquad [3.136]$$

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

$$\sigma_A^2\sigma_B^2 \geq \left(\frac{1}{2i}[\langle f|g\rangle - \langle g|f\rangle]\right)^2. \qquad [3.137]$$

But

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

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

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \qquad [3.138]$$

is the commutator of the two operators. *Conclusion*:

$$\boxed{\sigma_A^2\sigma_B^2 \geq \left(\frac{1}{2i}\langle[\hat{A}, \hat{B}]\rangle\right)^2.} \qquad [3.139]$$

This is the uncertainty principle in its most general form. (You might think the i makes it trivial—isn't the right side *negative*? No, for the commutator carries its own factor of i, and the two cancel out.)

For example, suppose the first observable is position ($\hat{A} = x$), and the second is momentum ($\hat{B} = (\hbar/i)d/dx$). To determine the commutator, we use an arbitrary

"test function", $f(x)$:

$$[\hat{x}, \hat{p}]f(x) = x\frac{\hbar}{i}\frac{d}{dx}(f) - \frac{\hbar}{i}\frac{d}{dx}(xf) = \frac{\hbar}{i}\left[x\frac{df}{dx} - (f + x\frac{df}{dx})\right] = i\hbar f,$$

so

$$\boxed{[\hat{x}, \hat{p}] = i\hbar.} \qquad [3.140]$$

Accordingly,

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

or, since standard deviations are by their nature positive,

$$\sigma_x\sigma_p \geq \frac{\hbar}{2}. \qquad [3.141]$$

That proves the original Heisenberg uncertainty principle, but we now see that it is just one application of a far more general theorem: There will be an "uncertainty principle" for *any pair of observables whose corresponding operators do not commute.* We call them **incompatible observables**. Evidently, incompatible observables do not have shared eigenvectors—at least, they cannot have a *complete set* of common eigenvectors. Matrices representing incompatible observables cannot be simultaneously diagonalized (that is, they cannot both be brought to diagonal form by the *same* similarity transformation). On the other hand, *compatible* observables (whose operators *do* commute) share a complete set of eigenvectors, and the corresponding matrices *can* be simultaneously diagonalized (see Problem 3.40).

***Problem 3.39** Prove the famous "(your name) uncertainty principle," relating the uncertainty in position ($A = x$) to the uncertainty in energy ($B = p^2/2m + V$):

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

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

Problem 3.40 Prove the following:

(a) If two matrices commute ($[\mathbf{A}, \mathbf{B}] = 0$), and you apply the same similarity transformation to both of them ($\mathbf{A}' = \mathbf{SAS}^{-1}$, $\mathbf{B}' = \mathbf{SBS}^{-1}$), the resulting matrices also commute ($[\mathbf{A}', \mathbf{B}'] = 0$).

(b) Diagonal matrices always commute. (It follows from this that simultaneously diagonalizable matrices must commute. Conversely, if two Hermitian matrices commute, then they are simulatneously diagonalizable—i.e., they have a complete set of common eigenvectors. This is not so easy to prove[26] unless you happen to know that the spectrum of one of them is nondegenerate.)

---

[26] See Byron and Fuller (footnote 14), Theorem 4.22.

(c) If matrices **A** and **B** commute, and $|\alpha\rangle$ is an eigenvector of **A**, and the spectrum of **A** is nondegenerate, then $|\alpha\rangle$ is also an eigenvector of **B**. (In that case the matrix **S** that diagonalizes **A** *also* diagonalizes **B**.)

***Problem 3.41**

(a) Prove the following commutator identity:

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}. \qquad [3.142]$$

(b) Using Equations 3.140 and 3.142, show that

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

(c) For any function $f(x)$ that can be expanded in a power series, show that

$$[f(\hat{x}), \hat{p}] = i\hbar f'(\hat{x}),$$

where the prime denotes differentiation.

3.4.2 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 2.14) and the Gaussian wave packet for the free particle (Problem 2.22). 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 3.135 and Equation 3.136. 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 the angle between the two vectors (Equation 3.28) is zero—that is, when one is a multiple of the other: $|g\rangle = c|f\rangle$, for some scalar c. (Study the proof of the Schwarz inequality in Problem 3.5 if you're not convinced.) Meanwhile, in Equation 3.136 I threw away the real part of z; equality results if $\mathrm{Re}(z) = 0$, which is to say, if $\mathrm{Re}\langle f|g\rangle = \mathrm{Re}(c\langle f|f\rangle) = 0$. Now $\langle f|f\rangle$ is certainly real, so this means the constant c must be purely imaginary—let's call it ia. The necessary and sufficient condition for minimum uncertainty, then, is

$$|g\rangle = ia|f\rangle, \quad \text{where } a \text{ is real.} \qquad [3.143]$$

In particular, for the position-momentum uncertainty principle this criterion becomes

$$\left(\frac{\hbar}{i}\frac{d}{dx} - \langle p\rangle\right)\Psi = ia(x - \langle x\rangle)\Psi, \qquad [3.144]$$

which is a differential equation for Ψ as a function of x, with the general solution (see Problem 3.42)

$$\Psi(x) = Ae^{-a(x-\langle x\rangle)^2/2\hbar}e^{i\langle p\rangle x/\hbar}. \quad [3.145]$$

Evidently the minimum-uncertainty wave packet is a *Gaussian*—and sure enough, the two examples we encountered earlier *were* Gaussians.[27]

Problem 3.42 Solve Equation 3.144 for $\Psi(x)$. (Note that $\langle x\rangle$ and $\langle p\rangle$ are *constants*, as far as x is concerned.)

3.4.3 The Energy-Time Uncertainty Principle

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

$$\Delta x\,\Delta p \geq \frac{\hbar}{2}. \quad [3.146]$$

Δx (the "uncertainty" in x) is sloppy notation (and sloppy language) for the standard deviation in the results of repeated measurements on identically prepared systems. Equation 3.146 is often paired with the **energy-time uncertainty principle**,

$$\Delta t\,\Delta E \geq \frac{\hbar}{2}. \quad [3.147]$$

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 3.147 would be a necessary concomitant to Equation 3.146. 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 3.147 is emphatically *not* implied by Equation 3.146. 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 similarity in appearance to the position-momentum uncertainty principle is 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

---

[27] Note that it is only the dependence of Ψ on x that is at issue here—the "constants" A, a, $\langle x\rangle$, and $\langle p\rangle$ may all be functions of time, and as time goes on Ψ may evolve away from the minimal form. All I'm asserting is that if, at some instant, the wave function is Gaussian in x, then (at that instant) the uncertainty product is minimal.

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, $Q(x, p, t)$:

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

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

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

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

$$\boxed{\frac{d}{dt}\langle Q\rangle = \frac{i}{\hbar}\langle[\hat{H}, \hat{Q}]\rangle + \langle\frac{\partial\hat{Q}}{\partial t}\rangle.} \qquad [3.148]$$

This is an interesting and useful result in its own right (see Problems 3.43 and 3.53). In the typical case, where the operator does not depend explicitly on t,[28] 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.

Suppose we pick $A = H$ and $B = Q$, in the generalized uncertainty principle (Equation 3.139), 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,

$$\sigma_H\sigma_Q \geq \frac{\hbar}{2}\left|\frac{d\langle Q\rangle}{dt}\right|. \qquad [3.149]$$

Let's define $\Delta E \equiv \sigma_H$ (with Δ as the usual sloppy notation for standard deviation), and

$$\Delta t \equiv \frac{\sigma_Q}{|d\langle Q\rangle/dt|}. \qquad [3.150]$$

---

[28] As an example of *explicit* time dependence, think of the potential energy of a harmonic oscillator whose spring constant is changing (perhaps the temperature is rising, so the spring becomes more flexible): $Q = (1/2)m[\omega(t)]^2x^2$.

Then

$$\Delta E\,\Delta t \geq \frac{\hbar}{2}, \qquad [3.151]$$

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*. 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. and conversely, if *any* observable changes rapidly, the "uncertainty" in the energy must be large.

Example 1. In the extreme case of a stationary state, for which the energy is uniquely determined, all expectation values are constant in time ($\Delta t = \infty$)—as, in fact, we noticed some time ago (see Equation 2.8). To make something *happen*, you must take a linear combination of at least two stationary states—for example,

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

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, then, $\Delta E = E_2 - E_1$ and $\Delta t = \tau$ (for the *exact* calculation, see Problem 3.44), so

$$\Delta E\,\Delta t = 2\pi\hbar > \frac{\hbar}{2}.$$

Example 2. How long does it take a free particle wave packet to pass by a particular point (Figure 3.1)? Qualitatively (an exact version is explored in Problem 3.45), $\Delta t = \Delta x/v = m\Delta x/p$, but $E = p^2/2m$, so $\Delta E = p\Delta p/m$. Therefore,

$$\Delta E\,\Delta t = \frac{p\Delta p}{m}\frac{m\Delta x}{p} = \Delta x\Delta p \geq \frac{\hbar}{2}.$$

Example 3. The Δ particle lasts about 10^{-23} seconds 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 115

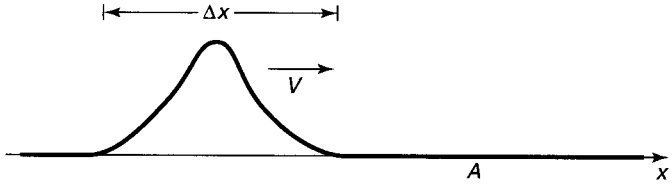


Figure 3.1: A free particle wave packet approaches the point A (Example 2).

MeV/c^2. Why does the rest energy (mc^2) sometimes come out higher than 1232, and sometimes lower? Is this experimental error? No, for

$$\Delta E\,\Delta t = (\frac{115}{2}\ \text{MeV})(10^{-23}\ \text{sec}) = 6 \times 10^{-22}\ \text{MeV sec},$$

whereas $\hbar/2 = 3 \times 10^{-22}$ MeV sec. 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.[29]

Notice the variety of specific meanings attaching to the term Δt in these examples: In Example 1 it's a period of oscillation; in Example 2 it's the time it takes a particle to pass a point; in Example 3 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 that 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. 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 3.151. 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 3.43** Apply Equation 3.148 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 2.35.

****Problem 3.44** Test the energy-time uncertainty principle for the wave function in Problem 2.6 and the observable x by calculating σ_H, σ_x, and $d\langle x\rangle/dt$ exactly.

---

[29] Actually, Example 3 is a bit of a fraud. You can't measure 10^{-23} sec on a stop-watch, and in practice the lifetime of such a short-lived particle is *inferred* from the width of the mass plot, using the uncertainty principle as *input*. However, the point is valid even if the numbers are suspect. Moreover, if you assume the Δ is about the same size as a proton ($\sim 10^{-15}$ m), then 10^{-23} sec is roughly the time it takes light to cross the particle, and it's hard to imagine that the lifetime could be much *less* than that.

Problem 3.45 Test the energy-time uncertainty principle for the free particle wave packet in Problem 2.40 and the observable x by calculating σ_H, σ_x, and $d\langle x\rangle/dt$ exactly.

Problem 3.46 Show that the energy-time uncertainty principle reduces to the "your name" uncertainty principle (Problem 3.39) when the observable in question is x.

FURTHER PROBLEMS FOR CHAPTER 3

Problem 3.47 *Functions* of matrices are defined by their Taylor series expansions; for example,

$$e^M \equiv \mathbf{1} + \mathbf{M} + \frac{1}{2}\mathbf{M}^2 + \frac{1}{3!}\mathbf{M}^3 + \cdots. \tag{3.152}$$

(a) Find $\exp(\mathbf{M})$, if

$$\text{(i) } \mathbf{M} = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}; \quad \text{(ii) } \mathbf{M} = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}.$$

(b) Show that if $\mathbf{M}$ is diagonalizable, then

$$\det\left(e^{\mathbf{M}}\right) = e^{\mathrm{Tr}(\mathbf{M})}. \tag{3.153}$$

(This is actually *true* even if $\mathbf{M}$ is *not* diagonalizable, but it's harder to prove in the general case.)

(c) Show that if the matrices $\mathbf{M}$ and $\mathbf{N}$ *commute*, then

$$e^{\mathbf{M}+\mathbf{N}} = e^{\mathbf{M}}e^{\mathbf{N}}. \tag{3.154}$$

Prove (with the simplest counterexample you can think up) that Equation 3.154 is *not* true, in general, for *non*commuting matrices.

(d) If $\mathbf{H}$ is Hermitian, show that $e^{i\mathbf{H}}$ is unitary.

Problem 3.48 A particle of mass m is in the ground state of the infinite square well (Equation 2.15). Suddenly the well expands to twice its original size—the right wall moving from a to $2a$—leaving the wave function (momentarily) undisturbed. The energy of the particle is now measured.

(a) What is the most probable result? What is the probability of getting that result?

(b) What is the *next* most probable result, and what is its probability?

(c) What is the *expectation value* of the energy? (If you find yourself confronted with an infinite series, try another method.)

Problem 3.49 A harmonic oscillator is in a state such that a measurement of the energy would yield either $(1/2)\hbar\omega$ or $(3/2)\hbar\omega$, with equal probability. What is the largest possible value of $\langle x\rangle$ in such a state? If it assumes this maximal value at time $t = 0$, what is $\Psi(x, t)$?

****Problem 3.50** Find the matrix elements $\langle n|x|n'\rangle$ and $\langle n|p|n'\rangle$ in the (orthonormal) basis consisting of stationary states for the harmonic oscillator (here $|n\rangle$ refers to the state ψ_n, Eq. 2.50). [You already calculated the diagonal elements ($n = n'$) in Problem 2.37; use the same technique for the general case.] Construct the corresponding (infinite) matrices, $\mathbf{X}$ and $\mathbf{P}$. Show that $(1/2m)\mathbf{P}^2 + (m\omega^2/2)\mathbf{X}^2 = \mathbf{H}$ is *diagonal*, in this basis. Are its diagonal elements what you would expect? *Partial answer*:

$$\langle n|x|n'\rangle = \sqrt{\frac{\hbar}{2m\omega}}\left(\sqrt{n'}\delta_{n,n'-1} + \sqrt{n}\delta_{n',n-1}\right). \qquad [3.155]$$

*****Problem 3.51** Show that

$$\langle x\rangle = \int \Phi^*\left(-\frac{\hbar}{i}\frac{\partial}{\partial p}\right)\Phi\, dp, \qquad [3.156]$$

where $\Phi(p, t)$ is the momentum-space wave function. In general,

$$\langle Q(x, p, t)\rangle = \begin{cases} \int \Psi^*\hat{Q}\left(x, \frac{\hbar}{i}\frac{\partial}{\partial x}, t\right)\Psi\, dx, & \text{in position space;} \\ \int \Phi^*\hat{Q}\left(-\frac{\hbar}{i}\frac{\partial}{\partial p}, p, t\right)\Phi\, dp, & \text{in momentum space.} \end{cases} \qquad [3.157]$$

Hint: Notice that $x\exp(ipx/\hbar) = -i\hbar(d/dp)\exp(ipx/\hbar)$.

****Problem 3.52** Find the momentum-space wave function $\Phi_n(p, t)$ for the nth stationary state of the infinite square well. Construct $|\Phi_n|^2$ (it's simplest to write separate formulas for odd and even n). Show that $|\Phi_n|^2$ is finite at $p = \pm n\pi\hbar/a$.

***Problem 3.53** Use Equation 3.148 to show that

$$\frac{d}{dt}\langle xp\rangle = 2\langle T\rangle - \langle x\frac{dV}{dx}\rangle, \qquad [3.158]$$

where T is the kinetic energy ($H = T + V$). In a *stationary* state the left side is zero (why?), so

$$2\langle T\rangle = \langle x\frac{dV}{dx}\rangle. \qquad [3.159]$$

This is called the **virial theorem**. Use it to prove that $\langle T\rangle = \langle V\rangle$ for stationary states of the harmonic oscillator, and check that this is consistent with the results you got in Problems 2.14 and 2.37.

Problem 3.54 What would it mean for an observable Q to be *conserved*, in quantum mechanics? At a minimum, the expectation value of Q should be constant in time, for any state Ψ. The criterion for this (assuming Q has no *explicit* time dependence) is that $\hat{Q}$ commute with the Hamiltonian (Equation 3.148). But we'd like something more: The probability $|c_n|^2$ of getting any particular eigenvalue (λ_n) of $\hat{Q}$ should be independent of t. Show that this, too, is guaranteed by the condition $[\hat{H}, \hat{Q}] = 0$. (Assume that the potential energy is independent of t, but do *not* assume Ψ is a stationary state.) *Hint*: $\hat{Q}$ and $\hat{H}$ are compatible observables, so they have a complete set of simultaneous eigenvalues.

****Problem 3.55**

(a) For a function $f(x)$ that can be expanded in a Taylor series, show that

$$f(x + x_0) = e^{i\hat{p}x_0/\hbar} f(x)$$

(where x_0 is any constant distance). For this reason, $\hat{p}/\hbar$ is called the **generator of translations in space**. (See Problem 3.47 for the meaning of an operator in the exponent.)

(b) If $\Psi(x, t)$ satisfies the (time-dependent) Schrödinger equation, show that

$$\Psi(x, t + t_0) = e^{-i\hat{H}t_0/\hbar} \Psi(x, t)$$

(where t_0 is any constant time); $-\hat{H}/\hbar$ is called the **generator of translations in time**.

(c) Show that the expectation value of a dynamical variable $Q(x, p, t)$, at time $t + t_0$, can be written

$$\langle Q \rangle_{t+t_0} = \langle \Psi(x, t) | e^{i\hat{H}t_0/\hbar} \hat{Q}(\hat{x}, \hat{p}, t + t_0) e^{-i\hat{H}t_0/\hbar} | \Psi(x, t) \rangle.$$

Use this to recover Equation 3.148. *Hint*: Let $t_0 = dt$, and expand to first order in dt.

Problem 3.56 In an interesting version of the energy-time uncertainty principle[30] $\Delta t = \tau/\pi$, where τ is the time it takes $\Psi(x, t)$ to evolve into a state orthogonal to $\Psi(x, 0)$. Test this out, using a wave function that is an equal admixture of two (orthonormal) stationary states of some (arbitrary) potential: $\Psi(x, 0) = (1/\sqrt{2})[\psi_1(x) + \psi_2(x)]$.

****Problem 3.57** Dirac proposed to peel apart the bracket notation for an inner product, $\langle \alpha | \beta \rangle$, into two pieces, which he called **bra** ($\langle \alpha |$) and **ket** ($| \beta \rangle$). 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

[30] See Lev Vaidman, *Am. J. Phys.* **60**, 182 (1992) for a proof.

product.[31] (When an *operator* hits a vector, it delivers another vector; when a *bra* hits a vector, it delivers a number.) Actually, 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 (though I shall not exploit it further in this book). For example, if $|\alpha\rangle$ is a normalized vector, the operator

$$\hat{P} \equiv |\alpha\rangle\langle\alpha| \qquad [3.160]$$

picks out the component 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 $|\alpha\rangle$.

(a) Show that $\hat{P}^2 = \hat{P}$. Determine the eigenvalues of $\hat{P}$, and characterize its eigenvectors.

(b) Suppose $|e_j\rangle$ is an orthonormal basis for an n-dimensional vector space. Show that

$$\sum_{j=1}^{n} |e_j\rangle\langle e_j| = \mathbf{1}. \qquad [3.161]$$

This is the tidiest statement of completeness.

(c) Let $\hat{Q}$ be an operator with a complete set of orthonormal eigenvectors:

$$\hat{Q}|e_j\rangle = \lambda_j|e_j\rangle \quad (j = 1, 2, 3, \ldots n).$$

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

$$\hat{Q} = \sum_{j=1}^{n} \lambda_j |e_j\rangle\langle e_j|. \qquad [3.162]$$

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

$$\hat{Q}|\alpha\rangle = \left\{\sum_{j=1}^{n} \lambda_j |e_j\rangle\langle e_j|\right\} |\alpha\rangle,$$

for any vector $|\alpha\rangle$.

---

[31] In a function space, the bra can be thought of as an instruction to integrate

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

with the "hole" $[\cdots]$ waiting to be filled by whatever function the bra encounters next.

***Problem 3.58** Imagine a system in which there are just *two* linearly independent states:

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

The most general state is a normalized linear combination:

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

Suppose the Hamiltonian matrix is

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

where g and h are real constants. The (time-dependent) Schrödinger equation says

$$\mathbf{H}|\Psi\rangle = i\hbar \frac{d}{dt}|\Psi\rangle.$$

(a) Find the eigenvalues and (normalized) eigenvectors of this Hamiltonian.

(b) Suppose the system starts out (at $t = 0$) in state $|1\rangle$. What is the state at time t? *Answer*:

$$|\Psi(t)\rangle = e^{-iht/\hbar} \begin{pmatrix} \cos(gt/\hbar) \\ -i \sin(gt/\hbar) \end{pmatrix}.$$

Note: This is about the simplest nontrivial quantum system conceivable. It is a crude model for (among other things) **neutrino oscillations**. In that case $|1\rangle$ represents the electron neutrino, and $|2\rangle$ the muon neutrino; if the Hamiltonian has a nonvanishing off-diagonal term g, then in the course of time the electron neutrino will turn into a muon neutrino, and back again. At present this is highly speculative—there is no experimental evidence for neutrino oscillations; however, a very similar phenomenon does occur in the case of neutral K-mesons (K^0 and $\bar{K}^0$).

CHAPTER 4

QUANTUM MECHANICS IN THREE DIMENSIONS

4.1 SCHRÖDINGER EQUATION IN SPHERICAL COORDINATES

The generalization to three dimensions is straightforward. Schrödinger's equation says

$$i\hbar\frac{\partial\Psi}{\partial t} = H\Psi; \qquad [4.1]$$

the Hamiltonian operator[1] H is obtained from the classical energy

$$\frac{1}{2}mv^2 + V = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V$$

by the standard prescription (applied now to y and z, as well as x):

$$p_x \to \frac{\hbar}{i}\frac{\partial}{\partial x}, \quad p_y \to \frac{\hbar}{i}\frac{\partial}{\partial y}, \quad p_z \to \frac{\hbar}{i}\frac{\partial}{\partial z}, \qquad [4.2]$$

or

$$\boxed{\mathbf{p} \to \frac{\hbar}{i}\nabla,} \qquad [4.3]$$

[1] Where confusion might otherwise occur, I have been putting "hats" on operators to distinguish them from the corresponding classical observables. I don't think there will be much occasion for ambiguity in this chapter, and the hats get to be cumbersome, so I am going to leave them off from now on.

for short. Thus

$$\boxed{i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi,} \quad [4.4]$$

where

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad [4.5]$$

is the **Laplacian**, in Cartesian coordinates.

The potential energy V and the wave function Ψ are now functions of $\mathbf{r} = (x, y, z)$ and t. The probability of finding the particle in the infinitesimal volume $d^3\mathbf{r} = dx\,dy\,dx$ is $|\Psi(\mathbf{r}, t)|^2\,d^3\mathbf{r}$, and the normalization condition reads

$$\int |\Psi|^2\,d^3\mathbf{r} = 1, \quad [4.6]$$

with the integral taken over all space. If the potential is independent of time, there will be a complete set of stationary states,

$$\Psi_n(\mathbf{r}, t) = \psi_n(\mathbf{r})e^{-iE_nt/\hbar}, \quad [4.7]$$

where the spatial wave function ψ_n satisfies the time-*independent* Schrödinger equation:

$$-\frac{\hbar^2}{2m}\nabla^2\psi_n + V\psi_n = E_n\psi_n. \quad [4.8]$$

The general solution to the (time-*dependent*) Schrödinger equation is

$$\Psi(\mathbf{r}, t) = \sum c_n\psi_n(\mathbf{r})e^{-iE_nt/\hbar}, \quad [4.9]$$

with the constants c_n determined by the initial wave function, $\Psi(\mathbf{r}, 0)$, in the usual way. (If the potential admits continuum states, then the sum in Equation 4.9 becomes an integral.)

---

*Problem 4.1

(a) Work out all of the **canonical commutation relations** for components of the operators $\mathbf{r}$ and $\mathbf{p}$: $[x, y]$, $[x, p_y]$, $[x, p_x]$, $[p_y, p_z]$, and so on. *Answer*:

$$[r_i, p_j] = -[p_i, r_j] = i\hbar\delta_{ij}, \quad [r_i, r_j] = [p_i, p_j] = 0. \quad [4.10]$$

(b) Show that

$$\frac{d}{dt}\langle\mathbf{r}\rangle = \frac{1}{m}\langle\mathbf{p}\rangle, \quad \text{and} \quad \frac{d}{dt}\langle\mathbf{p}\rangle = \langle -\nabla V\rangle. \quad [4.11]$$

(Each of these, of course, stands for *three* equations—one for each component.) *Hint*: Note that Equation 3.148 is valid in three dimensions.

(c) Formulate Heisenberg's uncertainty principle in three dimensions. *Answer*:

$$\sigma_x \sigma_{p_x} \geq \hbar/2, \quad \sigma_y \sigma_{p_y} \geq \hbar/2, \quad \sigma_z \sigma_{p_z} \geq \hbar/2, \tag{4.12}$$

but there is no restriction on, say, $\sigma_x \sigma_{p_y}$.

4.1.1 Separation of Variables

Typically, the potential is a function only of the distance from the origin. In that case it is natural to adopt **spherical coordinates**, (r, θ, ϕ) (see Figure 4.1). In spherical coordinates the Laplacian takes the form[2]

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\left(\frac{\partial^2}{\partial \phi^2}\right). \tag{4.13}$$

In spherical coordinates, then, the time-independent Schrödinger equation reads

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \psi}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \psi}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\left(\frac{\partial^2 \psi}{\partial \phi^2}\right)\right] + V\psi = E\psi. \tag{4.14}$$

We begin by looking for solutions that are separable into products:

$$\psi(r, \theta, \phi) = R(r)Y(\theta, \phi). \tag{4.15}$$

Putting this into Equation 4.14, we have

$$-\frac{\hbar^2}{2m}\left[\frac{Y}{r^2}\frac{d}{dr}\left(r^2 \frac{dR}{dr}\right) + \frac{R}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial Y}{\partial \theta}\right) + \frac{R}{r^2 \sin^2\theta}\frac{\partial^2 Y}{\partial \phi^2}\right] + VRY = ERY.$$

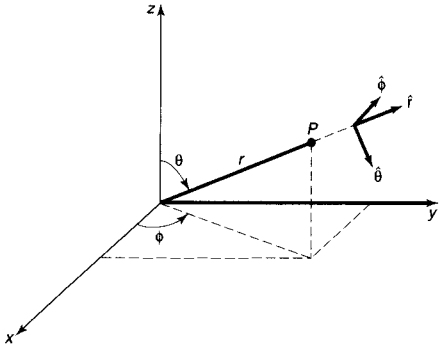


Figure 4.1: Spherical coordinates: radius r, polar angle θ, and azimuthal angle ϕ.

[2] In principle, this can be obtained by change of variables from the Cartesian expression (Equation 4.5). However, there are much more efficient ways of getting it; see, for instance, M. Boas, *Mathematical Methods in the Physical Sciences*, 2nd ed. (New York: John Wiley and Sons, Inc., 1983) Chapter 10, Section 9.

Dividing by YR and multiplying by $-2mr^2/\hbar^2$:

$$\left\{\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - \frac{2mr^2}{\hbar^2}[V(r) - E]\right\}$$

$$+\frac{1}{Y}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right\} = 0.$$

The term in the first curly bracket depends only on r, whereas the remainder depends only on θ and ϕ; accordingly, each must be a constant. For reasons that will appear in due course, I will write this "separation constant" in the form $l(l+1)$:[3]

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - \frac{2mr^2}{\hbar^2}[V(r) - E] = l(l+1); \qquad [4.16]$$

$$\frac{1}{Y}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right\} = -l(l+1). \qquad [4.17]$$

***Problem 4.2** Use separation of variables in Cartesian coordinates to solve the infinite cubical well (or "particle in a box"):

$$V(x, y, z) = \begin{cases} 0, & \text{if } x, y, z \text{ are all between 0 and } a; \\ \infty, & \text{otherwise.} \end{cases}$$

(a) Find the stationary state wave functions and the corresponding energies.

(b) Call the distinct energies $E_1, E_2, E_3, \ldots$, in order of increasing energy. Find E_1, E_2, E_3, E_4, E_5, and E_6. Determine the degeneracy of each of these energies (that is, the number of different states that share the same energy). Recall (Problem 2.42) that degenerate bound states do not occur in *one* dimension, but they are common in three dimensions.

(c) What is the degeneracy of E_{14}, and why is this case interesting?

4.1.2 The Angular Equation

Equation 4.17 determines the dependence of ψ on θ and ϕ; multiplying by $Y\sin^2\theta$, it becomes

$$\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{\partial^2 Y}{\partial\phi^2} = -l(l+1)\sin^2\theta Y. \qquad [4.18]$$

---

[3]Note that there is no loss of generality here—at this stage l could be any complex number. Later on we'll discover that l must in fact be an *integer*, and it is in anticipation of that result that I express the separation constant in a way that looks peculiar now.

You may have encountered this equation already—it occurs in the solution to Laplace's equation in classical electrodynamics. As always, we try separation of variables:

$$Y(\theta, \phi) = \Theta(\theta)\Phi(\phi). \qquad [4.19]$$

Plugging this in, and dividing by $\Theta\Phi$, we find

$$\left\{\frac{1}{\Theta}\left[\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)\right] + l(l+1)\sin^2\theta\right\} + \frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2} = 0.$$

The first term is a function only of θ, and the second is a function only of ϕ, so each must be a constant. This time I'll call the separation constant m^2:[4]

$$\frac{1}{\Theta}\left[\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)\right] + l(l+1)\sin^2\theta = m^2; \qquad [4.20]$$

$$\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2} = -m^2. \qquad [4.21]$$

The ϕ equation is easy:

$$\frac{d^2\Phi}{d\phi^2} = -m^2\Phi \;\Rightarrow\; \Phi(\phi) = e^{im\phi}. \qquad [4.22]$$

[Actually, there are *two* solutions: $\exp(im\phi)$ and $\exp(-im\phi)$, but we'll cover the latter by allowing m to run negative. There could also be a constant factor in front, but we might as well absorb that into Θ. Incidentally, in electrodynamics we would write the azimuthal function (Φ) in terms of sines and cosines, instead of exponentials, because electric potentials must be *real*. In quantum mechanics there is no such constraint, and the exponentials are a lot easier to work with.] Now, when ϕ advances by 2π, we return to the same point in space (see Figure 4.1), so it is natural to require that[5]

$$\Phi(\phi + 2\pi) = \Phi(\phi). \qquad [4.23]$$

In other words, $\exp[im(\phi + 2\pi)] = \exp(im\phi)$, or $\exp(2\pi im) = 1$. From this it follows that m must be an *integer*:

$$m = 0, \pm 1, \pm 2, \ldots. \qquad [4.24]$$

---

[4]Again, there is no loss of generality here since at this stage m could be any complex number; in a moment, though, we will discover that m must in fact be an *integer*. *Beware:* The letter m is now doing double duty, as *mass* and as the so-called **magnetic quantum number**. There is no graceful way to avoid this since both uses are standard. Some authors now switch to M or μ for mass, but I hate to change notation in midstream, and I don't think confusion will arise as long as you are aware of the problem.

[5]This is a more subtle point than it looks. After all, the *probability* density ($|\Phi|^2$) is single valued *regardless* of m. In Section 4.3 we'll obtain the condition on m by an entirely different—and more compelling—argument.

The θ equation,

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+[l(l+1)\sin^2\theta-m^2]\Theta=0, \tag{4.25}$$

may not be so familiar. The solution is

$$\Theta(\theta)=AP_l^m(\cos\theta), \tag{4.26}$$

where P_l^m is the **associated Legendre function**, defined by[6]

$$P_l^m(x)\equiv(1-x^2)^{|m|/2}\left(\frac{d}{dx}\right)^{|m|}P_l(x), \tag{4.27}$$

and $P_l(x)$ is the lth Legendre polynomial. We encountered the latter (Equation 3.91) as orthogonal polynomials on the interval $(-1,+1)$; for our present purposes it is more convenient to define them by the **Rodrigues formula**:

$$P_l(x)\equiv\frac{1}{2^l l!}\left(\frac{d}{dx}\right)^l(x^2-1)^l. \tag{4.28}$$

For example,

$$P_0(x)=1,\ P_1(x)=\frac{1}{2}\frac{d}{dx}(x^2-1)=x,$$

$$P_2(x)=\frac{1}{4\cdot 2}\left(\frac{d}{dx}\right)^2(x^2-1)^2=\frac{1}{2}(3x^2-1),$$

and so on. The first few Legendre polynomials were listed in Table 3.1. As the name suggests, $P_l(x)$ is a polynomial (of degree l) in x, and is even or odd according to the parity of l. But $P_l^m(x)$ is not, in general, a polynomial—if m is odd it carries a factor of $\sqrt{1-x^2}$:

$$P_2^0(x)=\frac{1}{2}(3x^2-1),\ P_2^1(x)=(1-x^2)^{1/2}\frac{d}{dx}\left[\frac{1}{2}(3x^2-1)\right]=3x\sqrt{1-x^2},$$

$$P_2^2(x)=(1-x^2)\left(\frac{d}{dx}\right)\left[\frac{1}{2}(3x^2-1)\right]=3(1-x^2),$$

etc. [On the other hand, what *we* need is $P_l^m(\cos\theta)$, and $\sqrt{1-\cos^2\theta}=\sin\theta$, so $P_l^m(\cos\theta)$ is always a polynomial in $\cos\theta$, multiplied—if m is odd—by $\sin\theta$. Some associated Legendre functions of $\cos\theta$ are listed in Table 4.1.]

[6]Notice that $P_l^{-m}=P_l^m$. Some authors adopt a different sign convention for negative values of m; see Boas (footnote 2) p. 505.

Table 4.1: Some associated Legendre functions, $P_l^m(\cos\theta)$.

$P_1^1 = \sin\theta$	
	$P_3^3 = 15\sin\theta(1-\cos^2\theta)$
$P_1^0 = \cos\theta$	
	$P_3^2 = 15\sin^2\theta\cos\theta$
$P_2^2 = 3\sin^2\theta$	
	$P_3^1 = \frac{3}{2}\sin\theta(5\cos^2\theta - 1)$
$P_2^1 = 3\sin\theta\cos\theta$	
	$P_3^0 = \frac{1}{2}(5\cos^3\theta - 3\cos\theta)$
$P_2^0 = \frac{1}{2}(3\cos^2\theta - 1)$	

Notice that l must be a nonnegative integer for the Rodrigues formula to make any sense; moreover, if $|m| > l$, then Equation 4.27 says $P_l^m = 0$. For any given l, then, there are $(2l+1)$ possible values of m:

$$l = 0, 1, 2, \ldots; \quad m = -l, -l+1, \ldots, -1, 0, 1, \ldots, l-1, l. \qquad [4.29]$$

But wait! Equation 4.25 is a second-order differential equation: It should have *two* linearly independent solutions, for *any old* values of l and m. Where are all the *other* solutions? *Answer*: They *exist*, of course, as mathematical solutions to the equation, but they are *physically* unacceptable because they blow up at $\theta = 0$ and/or $\theta = \pi$, and do not yield normalizable wave functions (see Problem 4.4).

Now, the volume element in spherical coordinates[7] is

$$d^3\mathbf{r} = r^2\sin\theta\, dr\, d\theta\, d\phi, \qquad [4.30]$$

so the normalization condition (Equation 4.6) becomes

$$\int |\psi|^2 r^2 \sin\theta\, dr\, d\theta\, d\phi = \int |R|^2 r^2\, dr \int |Y|^2 \sin\theta\, d\theta\, d\phi = 1.$$

It is convenient to normalize R and Y individually:

$$\int_0^\infty |R|^2 r^2\, dr = 1 \quad \text{and} \quad \int_0^{2\pi}\int_0^\pi |Y|^2 \sin\theta\, d\theta\, d\phi = 1. \qquad [4.31]$$

The normalized angular wave functions[8] are called **spherical harmonics**:

[7] See, for instance, Boas, (footnote 2), Chapter 5, Section 4.

[8] The normalization factor is derived in Problem 4.47. The ϵ factor is chosen for consistency with the notation we will be using in the theory of angular momentum; it is reasonably standard, though some older books use other conventions. Notice that

$$Y_l^{-m} = (-1)^m Y_l^m.$$

Table 4.2: The first few spherical harmonics, $Y_l^m(\theta, \phi)$.

$Y_0^0 = \left(\frac{1}{4\pi}\right)^{1/2}$	$Y_2^{\pm 2} = \left(\frac{15}{32\pi}\right)^{1/2} \sin^2\theta e^{\pm 2i\phi}$
$Y_1^0 = \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta$	$Y_3^0 = \left(\frac{7}{16\pi}\right)^{1/2} (5\cos^3\theta - 3\cos\theta)$
$Y_1^{\pm 1} = \mp\left(\frac{3}{8\pi}\right)^{1/2} \sin\theta e^{\pm i\phi}$	$Y_3^{\pm 1} = \mp\left(\frac{21}{64\pi}\right)^{1/2} \sin\theta(5\cos^2\theta - 1)e^{\pm i\phi}$
$Y_2^0 = \left(\frac{5}{16\pi}\right)^{1/2} (3\cos^2\theta - 1)$	$Y_3^{\pm 2} = \left(\frac{105}{32\pi}\right)^{1/2} \sin^2\theta \cos\theta e^{\pm 2i\phi}$
$Y_2^{\pm 1} = \mp\left(\frac{15}{8\pi}\right)^{1/2} \sin\theta \cos\theta e^{\pm i\phi}$	$Y_3^{\pm 3} = \mp\left(\frac{35}{64\pi}\right)^{1/2} \sin^3\theta e^{\pm 3i\phi}$

$$Y_l^m(\theta, \phi) = \epsilon \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-|m|)!}{(l+|m|)!}} e^{im\phi} P_l^m(\cos\theta), \qquad [4.32]$$

where $\epsilon = (-1)^m$ for $m \geq 0$ and $\epsilon = 1$ for $m \leq 0$. As we shall prove later on, they are automatically orthogonal, so

$$\int_0^{2\pi} \int_0^{\pi} [Y_l^m(\theta, \phi)]^* [Y_{l'}^{m'}(\theta, \phi)] \sin\theta \, d\theta \, d\phi = \delta_{ll'} \delta_{mm'}. \qquad [4.33]$$

In Table 4.2 I have listed the first few spherical harmonics.

∗**Problem 4.3** Use Equations 4.27, 4.28, and 4.32 to construct Y_0^0 and Y_2^1. Check that they are normalized and orthogonal.

Problem 4.4 Show that

$$\Theta(\theta) = A \ln[\tan(\theta/2)]$$

satisfies the θ equation (Equation 4.25) for $l = m = 0$. This is the unacceptable "second solution"—what's *wrong* with it?

∗**Problem 4.5** Using Equation 4.32, find $Y_l^l(\theta, \phi)$ and $Y_3^2(\theta, \phi)$. Check that they satisfy the angular equation (Equation 4.18), for the appropriate values of the parameters l and m.

∗∗**Problem 4.6** Starting from the Rodrigues formula, derive the orthonormality condition for Legendre polynomials:

$$\int_{-1}^{1} P_l(x) P_{l'}(x) \, dx = \left(\frac{2}{2l+1}\right) \delta_{ll'}. \qquad [4.34]$$

Hint: Use integration by parts.

4.1.3 The Radial Equation

Notice that the angular part of the wave function, $Y(\theta, \phi)$, is the same for all spherically symmetric potentials; the actual *shape* of the potential, $V(r)$, affects only the *radial* part of the wave function, $R(r)$, which is determined by Equation 4.16:

$$\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - \frac{2mr^2}{\hbar^2}[V(r) - E]R = l(l+1)R. \qquad [4.35]$$

This equation simplifies if we change variables: Let

$$u(r) \equiv rR(r), \qquad [4.36]$$

so that $R = u/r$, $dR/dr = [r(du/dr) - u]/r^2$, $(d/dr)[r^2(dR/dr)] = rd^2u/dr^2$, and hence

$$\boxed{-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2} + \left[V + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right]u = Eu.} \qquad [4.37]$$

This is called the **radial equation**[9]; it is *identical in form* to the one-dimensional Schrödinger equation (Equation 2.4), except that the effective potential,

$$V_{\text{eff}} = V + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}, \qquad [4.38]$$

contains an extra piece, the so-called **centrifugal term**, $(\hbar^2/2m)[l(l+1)/r^2]$. It tends to throw the particle outward (away from the origin), just like the centrifugal (pseudo-) force in classical mechanics. Meanwhile, the normalization condition (Equation 4.31) becomes

$$\int_0^\infty |u|^2\, dr = 1. \qquad [4.39]$$

We cannot proceed further until a specific potential is provided.

Example. Consider the **infinite spherical well**,

$$V(r) = \begin{cases} 0, & \text{if } r < a; \\ \infty, & \text{if } r > a. \end{cases} \qquad [4.40]$$

Outside the well the wave function is zero; inside the well the radial equation says

$$\frac{d^2u}{dr^2} = \left[\frac{l(l+1)}{r^2} - k^2\right]u, \qquad [4.41]$$

---

[9] Those m's are *masses*, of course—the radial equation makes no reference to the quantum number m.

where

$$k \equiv \frac{\sqrt{2mE}}{\hbar}, \qquad [4.42]$$

as usual. Our problem is to solve this equation, subject to the boundary condition $u(a) = 0$. The case $l = 0$ is easy:

$$\frac{d^2u}{dr^2} = -k^2u \quad \Rightarrow u(r) = A\sin(kr) + B\cos(kr).$$

But remember, the actual radial wave function is $R(r) = u(r)/r$, and $[\cos(kr)]/r$ blows up as $r \to 0$. So[10] we must choose $B = 0$. The boundary condition then requires $\sin(ka) = 0$, and hence $ka = n\pi$, for some integer n. The allowed energies are evidently

$$E_{n0} = \frac{n^2\pi^2\hbar^2}{2ma^2}, \quad (n = 1, 2, 3, \ldots), \qquad [4.43]$$

the same as for the one-dimensional infinite square well (Equation 2.23). Normalizing $u(r)$ yields $A = \sqrt{2/a}$; inclusion of the angular part (constant, in this instance, since $Y_0^0(\theta, \phi) = 1/\sqrt{4\pi}$), we conclude that

$$\psi_{n00} = \frac{1}{\sqrt{2\pi a}} \frac{\sin(n\pi r/a)}{r}. \qquad [4.44]$$

[Notice that the stationary states are labeled by *three* **quantum numbers**, n, l, and m: $\psi_{nlm}(r, \theta, \phi)$. The *energy*, however, depends only on n and l: E_{nl}.]

The general solution to Equation 4.41 (for an arbitrary integer l) is not so familiar:

$$u(r) = Arj_l(kr) + Brn_l(kr), \qquad [4.45]$$

where $j_l(x)$ is the **spherical Bessel function** of order l, and $n_l(x)$ is the **spherical Neumann function** of order l. They are defined as follows:

$$j_l(x) \equiv (-x)^l \left(\frac{1}{x}\frac{d}{dx}\right)^l \frac{\sin x}{x}; \quad n_l(x) \equiv -(-x)^l \left(\frac{1}{x}\frac{d}{dx}\right)^l \frac{\cos x}{x}. \qquad [4.46]$$

For example,

$$j_0(x) = \frac{\sin x}{x}; \quad n_0(x) = -\frac{\cos x}{x};$$

$$j_1(x) = (-x)\frac{1}{x}\frac{d}{dx}\left(\frac{\sin x}{x}\right) = \frac{\sin x}{x^2} - \frac{\cos x}{x};$$

[10] Actually, all we require is that the wave function be *normalizable*, not that it be *finite*: $R(r) \sim 1/r$ at the origin *would* be normalizable (because of the r^2 in Equation 4.31). For a more compelling proof that $B = 0$, see R. Shankar, *Principles of Quantum Mechanics* (New York: Plenum, 1980), p. 351.

$$n_1(x) = -(-x)\frac{1}{x}\frac{d}{dx}\left(\frac{\cos x}{x}\right) = -\frac{\cos x}{x^2} - \frac{\sin x}{x};$$

and so on. The first few spherical Bessel and Neumann functions are listed in Table 4.3. Notice that for small x (where $\sin x \approx x - x^3/3! + x^5/5! - \cdots$ and $\cos x \approx 1 - x^2/2 + x^4/4! - \cdots$),

$$j_0(x) \approx 1; \quad n_0(x) \approx \frac{1}{x}; \quad j_1(x) \approx \frac{x}{3}; \quad n_1(x) \approx -\frac{1}{x^2};$$

etc. The point is that the Bessel functions are finite at the origin, but the Neumann functions blow up at the origin. Accordingly, we must have $B_l = 0$, and hence

$$R(r) = A j_l(kr). \qquad [4.47]$$

There remains the boundary condition, $R(a) = 0$. Evidently k must be chosen such that

$$j_l(ka) = 0; \qquad [4.48]$$

that is, (ka) is a zero of the l^{th}-order spherical Bessel function. Now the Bessel functions are oscillatory (see Figure 4.2); each one has an infinite number of zeros. But (unfortunately, for us) they are not located at nice sensible points (such as n, or $n\pi$, or something); they have to be computed numerically.[11] At any rate, the boundary condition requires that

$$k = \frac{1}{a}\beta_{nl}, \qquad [4.49]$$

Table 4.3: The first few spherical Bessel and Neumann functions, $j_l(x)$ and $n_l(x)$.

$j_0 = \dfrac{\sin x}{x}$	$n_0 = -\dfrac{\cos x}{x}$
$j_1 = \dfrac{\sin x}{x^2} - \dfrac{\cos x}{x}$	$n_1 = -\dfrac{\cos x}{x^2} - \dfrac{\sin x}{x}$
$j_2 = \left(\dfrac{3}{x^3} - \dfrac{1}{x}\right)\sin x - \dfrac{3}{x^2}\cos x$	$n_2 = -\left(\dfrac{3}{x^3} - \dfrac{1}{x}\right)\cos x - \dfrac{3}{x^2}\sin x$
$j_l \to \dfrac{x^l}{(2l+1)!!}$,	$n_l \to -\dfrac{(2l-1)!!}{x^{l+1}}$, for $x \ll 1$.

[11] Abramowitz and Stegun, eds., *Handbook of Mathematical Functions* (New York: Dover, 1965), Chapter 10, provides an extensive listing.

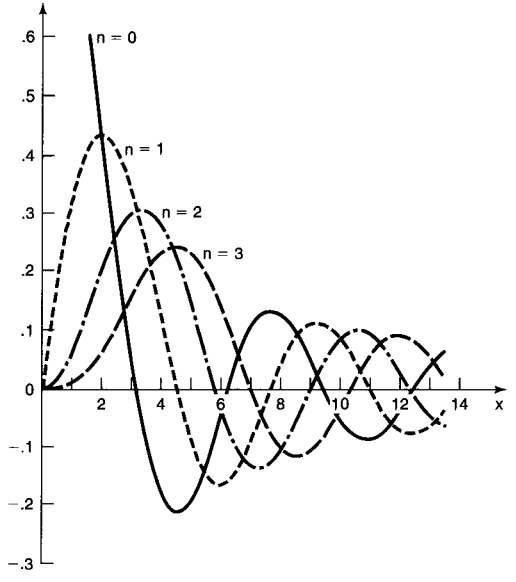


Figure 4.2: Graphs of the first four spherical Bessel functions.

where β_{nl} is the n^{th} zero of the l^{th} spherical Bessel function. The allowed energies, then, are given by

$$E_{nl} = \frac{\hbar^2}{2ma^2}\beta_{nl}^2, \qquad [4.50]$$

and the wave functions are

$$\psi_{nlm}(r, \theta, \phi) = A_{nl} j_l(\beta_{nl} r/a) Y_l^m(\theta, \phi), \qquad [4.51]$$

with the constant A_{nl} to be determined by normalization. Each energy level is $(2l+1)$-fold degenerate, since there are $(2l+1)$ different values of m for each value of l (see Equation 4.29).

Problem 4.7

(a) From the definitions (Equation 4.46), construct $j_2(x)$ and $n_2(x)$.

(b) Expand the sines and cosines to obtain approximate formulas for $j_2(x)$ and $n_2(x)$, valid when $x \ll 1$. Confirm that $j_2(x)$ is finite at the origin but $n_2(x)$ blows up.

Problem 4.8

(a) Check that $Arj_1(kr)$ satisfies the radial equation (Equation 4.37) with $V(r) = 0$ and $l = 1$.

(b) Determine graphically the allowed energies for the infinite spherical well when $l = 1$. Show that for large n, $E_{n1} \approx (\hbar^2\pi^2/2ma^2)(n + 1/2)^2$.

****Problem 4.9** A particle of mass m is placed in a *finite* spherical well:

$$V(r) = \begin{cases} 0, & \text{if } r \leq a; \\ V_0, & \text{if } r > a. \end{cases}$$

Find the ground state by solving the radial equation with $l = 0$. Show that there is no bound state at all if $V_0 a^2 < \pi^2\hbar^2/8m$.

4.2 THE HYDROGEN ATOM

The hydrogen atom consists of a heavy, essentially motionless proton (we may as well put it at the origin) of charge e, together with a much lighter electron (charge $-e$) that circles around it, held in orbit by the mutual attraction of opposite charges (see Figure 4.3). From Coulomb's law, the potential energy (in SI units) is

$$V(r) = -\frac{e^2}{4\pi\epsilon_0}\frac{1}{r}, \qquad [4.52]$$

and the radial equation (Equation 4.37) says

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2} + \left[-\frac{e^2}{4\pi\epsilon_0}\frac{1}{r} + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right]u = Eu. \qquad [4.53]$$

Our problem is to solve this equation for $u(r)$ and determine the allowed electron energies E. The hydrogen atom is such an important case that I'm not going to hand you the solutions this time—we'll work them out in detail by the method we used in the analytical solution to the harmonic oscillator. (If any step in this process is unclear, you may wish to refer back to Section 2.3.2 for a more complete explanation.) Incidentally, the Coulomb potential (Equation 4.52) admits *continuum* states (with $E > 0$), describing electron-proton scattering, as well as discrete *bound* states, representing the hydrogen atom, but we shall confine our attention to the latter.

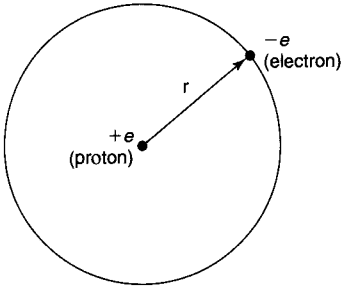


Figure 4.3: The hydrogen atom.

4.2.1 The Radial Wave Function

Our first task is to tidy up the notation. Let

$$\kappa \equiv \frac{\sqrt{-2mE}}{\hbar}. \qquad [4.54]$$

(For bound states, $E < 0$, so κ is real.) Dividing Equation 4.53 by E, we have

$$\frac{1}{\kappa^2}\frac{d^2u}{dr^2} = \left[1 - \frac{me^2}{2\pi\epsilon_0\hbar^2\kappa}\frac{1}{(\kappa r)} + \frac{l(l+1)}{(\kappa r)^2}\right]u.$$

This suggests that we let

$$\rho \equiv \kappa r, \quad \text{and} \quad \rho_0 \equiv \frac{me^2}{2\pi\epsilon_0\hbar^2\kappa}, \qquad [4.55]$$

so that

$$\frac{d^2u}{d\rho^2} = \left[1 - \frac{\rho_0}{\rho} + \frac{l(l+1)}{\rho^2}\right]u. \qquad [4.56]$$

Next we examine the asymptotic form of the solutions. As $\rho \to \infty$, the constant term in the brackets dominates, so (approximately)

$$\frac{d^2u}{d\rho^2} = u.$$

The general solution is

$$u(\rho) = Ae^{-\rho} + Be^{\rho}, \qquad [4.57]$$

but e^{ρ} blows up (as $\rho \to \infty$), so $B = 0$. Evidently,

$$u(\rho) \sim Ae^{-\rho} \qquad [4.58]$$

for large ρ. On the other hand, as $\rho \to 0$ the centrifugal term dominates[12]; approximately, then,

$$\frac{d^2u}{d\rho^2} = \frac{l(l+1)}{\rho^2}u.$$

The general solution (check it!) is

$$u(\rho) = C\rho^{l+1} + D\rho^{-l},$$

but ρ^{-l} blows up (as $\rho \to 0$), so $D = 0$. Thus

$$u(\rho) \sim C\rho^{l+1} \qquad [4.59]$$

[12]This argument does not apply when $l = 0$ (although the conclusion, Equation 4.59, is in fact valid for that case too). But never mind: All I am trying to do is provide some *motivation* for a change of variables (Equation 4.60.)

for small ρ.

The next step is to peel off the asymptotic behavior, introducing the new function $v(\rho)$:

$$u(\rho) = \rho^{l+1} e^{-\rho} v(\rho), \qquad [4.60]$$

in the hope that $v(\rho)$ will turn out to be simpler than $u(\rho)$. The first indications are not auspicious:

$$\frac{du}{d\rho} = \rho^l e^{-\rho} \left[(l + 1 - \rho)v + \rho \frac{dv}{d\rho} \right],$$

and

$$\frac{d^2u}{d\rho^2} = \rho^l e^{-\rho} \left\{ \left[-2l - 2 + \rho + \frac{l(l+1)}{\rho} \right] v + 2(l + 1 - \rho)\frac{dv}{d\rho} + \rho \frac{d^2v}{d\rho^2} \right\}.$$

In terms of $v(\rho)$, then, the radial equation (Equation 4.56) reads

$$\rho \frac{d^2v}{d\rho^2} + 2(l + 1 - \rho)\frac{dv}{d\rho} + [\rho_0 - 2(l + 1)]v = 0. \qquad [4.61]$$

Finally, we assume the solution, $v(\rho)$, can be expressed as a power series in ρ:

$$v(\rho) = \sum_{j=0}^{\infty} a_j \rho^j. \qquad [4.62]$$

Our problem is to determine the coefficients $(a_0, a_1, a_2, \ldots)$. Differentiating term by term,

$$\frac{dv}{d\rho} = \sum_{j=0}^{\infty} j a_j \rho^{j-1} = \sum_{j=0}^{\infty} (j + 1) a_{j+1} \rho^j.$$

[In the second summation I have renamed the "dummy index": $j \to j + 1$. If this troubles you, write out the first few terms explicitly, and *check* it. You might say that the sum should now begin at $j = -1$, but the factor $(j + 1)$ kills that term anyway, so we might as well start at zero.] Differentiating again,

$$\frac{d^2v}{d\rho^2} = \sum_{j=0}^{\infty} j(j + 1) a_{j+1} \rho^{j-1}.$$

Inserting these into Equation 4.61, we have

$$\sum_{j=0}^{\infty} j(j + 1) a_{j+1} \rho^j + 2(l + 1) \sum_{j=0}^{\infty} (j + 1) a_{j+1} \rho^j$$

$$-2 \sum_{j=0}^{\infty} j a_j \rho^j + [\rho_0 - 2(l + 1)] \sum_{j=0}^{\infty} a_j \rho^j = 0.$$

Equating the coefficients of like powers yields

$$j(j+1)a_{j+1} + 2(l+1)(j+1)a_{j+1} - 2ja_j + [\rho_0 - 2(l+1)]a_j = 0,$$

or

$$a_{j+1} = \left\{ \frac{2(j+l+1) - \rho_0}{(j+1)(j+2l+2)} \right\} a_j. \tag{4.63}$$

This recursion formula determines the coefficients, and hence the function $v(\rho)$: We start with $a_0 = A$ (this becomes an overall constant, to be fixed eventually by normalization), and Equation 4.63 gives us a_1; putting this back in, we obtain a_2, and so on.[13]

Now let's see what the coefficients look like for large j (this corresponds to large ρ, where the higher powers dominate). In this regime the recursion formula says

$$a_{j+1} \cong \frac{2j}{j(j+1)} a_j = \frac{2}{j+1} a_j,$$

so

$$a_j \cong \frac{2^j}{j!} A. \tag{4.64}$$

Suppose for a moment that this were the *exact* result. Then

$$v(\rho) = A \sum_{j=0}^{\infty} \frac{2^j}{j!} \rho^j = Ae^{2\rho},$$

and hence

$$u(\rho) = A\rho^{l+1} e^{\rho}, \tag{4.65}$$

which blows up at large ρ. The positive exponential is precisely the asymptotic behavior we *didn't* want in Equation 4.57. (It's no accident that it reappears here; after all, it *does* represent the asymptotic form of *some* solutions to the radial equation—they just don't happen to be the ones we're interested in, because they aren't normalizable.) There is only one way out of this dilemma: *The series must terminate.* There must occur some maximal integer, $j_{\max}$, such that

$$a_{j_{\max}+1} = 0 \tag{4.66}$$

(and beyond which all coefficients vanish automatically). Evidently (Equation 4.63)

$$2(j_{\max} + l + 1) - \rho_0 = 0.$$

[13] You might wonder why I didn't use the series method directly on $u(\rho)$—why factor out the asymptotic behavior before applying this procedure? The reason for peeling off ρ^{l+1} is largely aesthetic: Without this, the sequence would begin with a long string of zeroes (the first nonzero coefficient being a_{l+1}); by factoring out ρ^{l+1} we obtain a series that starts out with ρ^0. The $e^{-\rho}$ factor is more critical—if you *don't* pull that out, you get a three-term recursion formula involving a_{j+2}, a_{j+1}, and a_j (*try* it!), and that is enormously more difficult to work with.

Defining

$$n \equiv j_{\max} + l + 1 \tag{4.67}$$

(the so-called **principal quantum number**), we have

$$\rho_0 = 2n. \tag{4.68}$$

But ρ_0 determines E (Equations 4.54 and 4.55):

$$E = -\frac{\hbar^2\kappa^2}{2m} = -\frac{me^4}{8\pi^2\epsilon_0^2\hbar^2\rho_0^2}, \tag{4.69}$$

so the allowed energies are

$$\boxed{E_n = -\left[\frac{m}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\right]\frac{1}{n^2} = \frac{E_1}{n^2}, \quad n = 1, 2, 3, \ldots.} \tag{4.70}$$

This is the famous **Bohr formula**—by any measure the most important result in all of quantum mechanics. Bohr obtained it in 1913 by a serendipitous mixture of inapplicable classical physics and premature quantum theory (the Schrödinger equation did not come until 1924).

Combining Equations 4.55 and 4.68, we find that

$$\kappa = \left(\frac{me^2}{4\pi\epsilon_0\hbar^2}\right)\frac{1}{n} = \frac{1}{an}, \tag{4.71}$$

where

$$\boxed{a \equiv \frac{4\pi\epsilon_0\hbar^2}{me^2} = 0.529 \times 10^{-10}\ \text{m}} \tag{4.72}$$

is the so-called **Bohr radius**. It follows (again, from Equation 4.55) that

$$\rho = \frac{r}{an}. \tag{4.73}$$

Evidently the spatial wave functions for hydrogen are labeled by three quantum numbers (n, l, and m):

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r)\, Y_l^m(\theta, \phi), \tag{4.74}$$

where (referring back to Equations 4.36 and 4.60)

$$R_{nl}(r) = \frac{1}{r}\rho^{l+1}e^{-\rho}v(\rho), \tag{4.75}$$

and $v(\rho)$ is a polynomial of degree $j_{\max} = n - l - 1$ in ρ, whose coefficients are determined (up to an overall normalization factor) by the recursion formula

$$a_{j+1} = \frac{2(j + l + 1 - n)}{(j + 1)(j + 2l + 2)} a_j. \qquad [4.76]$$

The **ground state** (that is, the state of lowest energy) is the case $n = 1$; putting in the accepted values for the physical constants, we get

$$E_1 = -\left[\frac{m}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\right] = -13.6 \text{ eV}. \qquad [4.77]$$

Evidently the **binding energy** of hydrogen (the amount of energy you would have to impart to the electron in order to ionize the atom) is 13.6 eV. Equation 4.67 forces $l = 0$, whence also $m = 0$ (see Equation 4.29), so

$$\psi_{100}(r, \theta, \phi) = R_{10}(r) Y_0^0(\theta, \phi). \qquad [4.78]$$

The recursion formula truncates after the first term (Equation 4.76 with $j = 0$ yields $a_1 = 0$), so $v(\rho)$ is a constant (a_0) and

$$R_{10}(r) = \frac{a_0}{a} e^{-r/a}. \qquad [4.79]$$

Normalizing it, in accordance with Equation 4.31,

$$\int_0^\infty |R_{10}|^2 r^2 \, dr = \frac{|a_0|^2}{a^2} \int_0^\infty e^{-2r/a} r^2 \, dr = |a_0|^2 \frac{a}{4} = 1,$$

so $a_0 = 2/\sqrt{a}$. Meanwhile, $Y_0^0 = 1/\sqrt{4\pi}$, so

$$\boxed{\psi_{100}(r, \theta, \phi) = \frac{1}{\sqrt{\pi a^3}} e^{-r/a}.} \qquad [4.80]$$

If $n = 2$ the energy is

$$E_2 = \frac{-13.6 \text{ eV}}{4} = -3.4 \text{ eV}; \qquad [4.81]$$

this is the first excited state—or rather, *states*, since we can have either $l = 0$ (in which case $m = 0$) or $l = 1$ (with $m = -1$, 0, or +1), so there are actually four different states that share this energy. If $l = 0$, the recursion relation (Equation 4.76) gives

$$a_1 = -a_0 \text{ (using } j = 0), \quad \text{and } a_2 = 0 \text{ (using } j = 1),$$

so $v(\rho) = a_0(1 - \rho)$, and hence

$$R_{20}(r) = \frac{a_0}{2a}\left(1 - \frac{r}{2a}\right)e^{-r/2a}. \quad [4.82]$$

If $l = 1$ the recursion formula terminates the series after a single term, so $v(\rho)$ is a constant, and we find

$$R_{21}(r) = \frac{a_0}{4a^2} r e^{-r/2a}. \quad [4.83]$$

(In each case the constant a_0 is to be determined by normalization—see Problem 4.11.)

For arbitrary n, the possible values of l (consistent with Equation 4.67) are

$$l = 0, 1, 2, \ldots, n - 1. \quad [4.84]$$

For each l, there are $(2l + 1)$ possible values of m (Equation 4.29), so the total degeneracy of the energy level E_n is

$$d(n) = \sum_{l=0}^{n-1}(2l + 1) = n^2. \quad [4.85]$$

The polynomial $v(\rho)$ (defined by the recursion formula, Equation 4.76) is a function well known to applied mathematicians; apart from normalization, it can be written as

$$v(\rho) = L_{n-l-1}^{2l+1}(2\rho), \quad [4.86]$$

where

$$L_{q-p}^{p}(x) \equiv (-1)^p \left(\frac{d}{dx}\right)^p L_q(x) \quad [4.87]$$

is an **associated Laguerre polynomial**, and

$$L_q(x) \equiv e^x \left(\frac{d}{dx}\right)^q \left(e^{-x}x^q\right) \quad [4.88]$$

is the qth **Laguerre polynomial**.[14] (The first few Laguerre polynomials are listed in Table 4.4; some associated Laguerre polynomials are given in Table 4.5. The first few radial wave functions are listed in Table 4.6 and plotted in Figure 4.4.) The normalized hydrogen wave functions are[15]

$$\boxed{\psi_{nlm} = \sqrt{\left(\frac{2}{na}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}}\, e^{-r/na}\left(\frac{2r}{na}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2r}{na}\right) Y_l^m(\theta, \phi).} \quad [4.89]$$

[14] As usual, there are rival normalization conventions in the literature; I have adopted the most nearly standard one.

[15] If you want to see how the normalization factor is calculated, study (for example), L. Schiff, *Quantum Mechanics*, 2nd ed. (New York: McGraw-Hill, 1968), page 93.

Table 4.4: The first few Laguerre polynomials, $L_q(x)$.

$L_0 = 1$
$L_1 = -x + 1$
$L_2 = x^2 - 4x + 2$
$L_3 = -x^3 + 9x^2 - 18x + 6$
$L_4 = x^4 - 16x^3 + 72x^2 - 96x + 24$
$L_5 = -x^5 + 25x^4 - 200x^3 + 600x^2 - 600x + 120$
$L_6 = x^6 - 36x^5 + 450x^4 - 2400x^3 + 5400x^2 - 4320x + 720$

Table 4.5: Some associated Laguerre polynomials, $L_{q-p}^{p}(x)$.

$L_0^0 = 1$	$L_0^2 = 2$
$L_1^0 = -x + 1$	$L_1^2 = -6x + 18$
$L_2^0 = x^2 - 4x + 2$	$L_2^2 = 12x^2 - 96x + 144$
$L_0^1 = 1$	$L_0^3 = 6$
$L_1^1 = -2x + 4$	$L_1^3 = -24x + 96$
$L_2^1 = 3x^2 - 18x + 18$	$L_2^3 = 60x^2 - 600x + 1200$

They are not pretty, but don't complain—this is one of the very few realistic systems that can be solved at all, in exact closed form. As we will prove later on, they are mutually orthogonal:

$$\int \psi_{nlm}^* \, \psi_{n'l'm'} \, r^2 \sin\theta \, dr \, d\theta \, d\phi = \delta_{nn'} \delta_{ll'} \delta_{mm'}. \qquad [4.90]$$

***Problem 4.10** Work out the radial wave functions R_{30}, R_{31}, and R_{32}, using the recursion formula (Equation 4.76). Don't bother to normalize them.

***Problem 4.11**

(a) Normalize R_{20} (Equation 4.82), and construct the function ψ_{200}.

(b) Normalize R_{21} (Equation 4.83), and construct ψ_{211}, ψ_{210}, and ψ_{21-1}.

****Problem 4.12**

(a) Using Equation 4.88, work out the first four Laguerre polynomials.

(b) Using Equations 4.86, 4.87, and 4.88, find $v(\rho)$ for the case $n = 5$, $l = 2$.

Table 4.6: The first few radial wave functions for hydrogen, $R_{nl}(r)$.

$R_{10} = 2a^{-3/2}\exp(-r/a)$
$R_{20} = \frac{1}{\sqrt{2}}a^{-3/2}\left(1-\frac{1}{2}\frac{r}{a}\right)\exp(-r/2a)$
$R_{21} = \frac{1}{\sqrt{24}}a^{-3/2}\frac{r}{a}\exp(-r/2a)$
$R_{30} = \frac{2}{\sqrt{27}}a^{-3/2}\left(1-\frac{2}{3}\frac{r}{a}+\frac{2}{27}\left(\frac{r}{a}\right)^2\right)\exp(-r/3a)$
$R_{31} = \frac{8}{27\sqrt{6}}a^{-3/2}\left(1-\frac{1}{6}\frac{r}{a}\right)\left(\frac{r}{a}\right)\exp(-r/3a)$
$R_{32} = \frac{4}{81\sqrt{30}}a^{-3/2}\left(\frac{r}{a}\right)^2\exp(-r/3a)$
$R_{40} = \frac{1}{4}a^{-3/2}\left(1-\frac{3}{4}\frac{r}{a}+\frac{1}{8}\left(\frac{r}{a}\right)^2-\frac{1}{192}\left(\frac{r}{a}\right)^3\right)\exp(-r/4a)$
$R_{41} = \frac{\sqrt{5}}{16\sqrt{3}}a^{-3/2}\left(1-\frac{1}{4}\frac{r}{a}+\frac{1}{80}\left(\frac{r}{a}\right)^2\right)\frac{r}{a}\exp(-r/4a)$
$R_{42} = \frac{1}{64\sqrt{5}}a^{-3/2}\left(1-\frac{1}{12}\frac{r}{a}\right)\left(\frac{r}{a}\right)^2\exp(-r/4a)$
$R_{43} = \frac{1}{768\sqrt{35}}a^{-3/2}\left(\frac{r}{a}\right)^3\exp(-r/4a)$

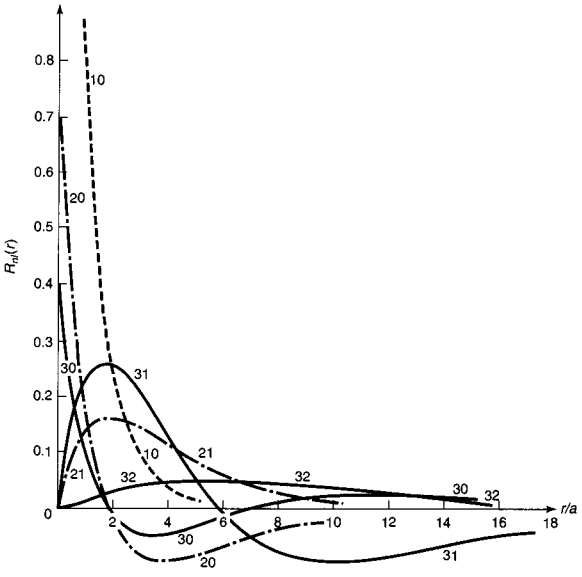


Figure 4.4: Graphs of the first few hydrogen radial wave functions, $R_{nl}(r)$.

(c) Again, find $v(\rho)$ for the case $n = 5, l = 2$, but this time get it from the recursion formula (Equation 4.76).

---

*Problem 4.13

(a) Find $\langle r \rangle$ and $\langle r^2 \rangle$ for an electron in the ground state of hydrogen. Express your answers in terms of the Bohr radius a.

(b) Find $\langle x \rangle$ and $\langle x^2 \rangle$ for an electron in the ground state of hydrogen. *Hint*: This requires no new integration—note that $r^2 = x^2 + y^2 + z^2$, and exploit the symmetry of the ground state.

(c) Find $\langle x^2 \rangle$ in the state $n = 2, l = 1, m = 1$. *Hint*: This state is *not* symmetrical in x, y, z. Use $x = r \sin\theta \cos\phi$.

Problem 4.14 What is the probability that an electron in the ground state of hydrogen will be found *inside the nucleus*?

(a) First calculate the *exact* answer, assuming that the wave function (Equation 4.80) is correct all the way down to $r = 0$. Let b be the radius of the nucleus.

(b) Expand your result as a power series in the small number $\varepsilon \equiv 2b/a$, and show that the lowest-order term is the cubic: $P \approx (4/3)(b/a)^3$. This should be a suitable approximation, provided that $b \ll a$ (which it *is*).

(c) Alternatively, we might assume that $\psi(r)$ is essentially constant over the (tiny) volume of the nucleus, so that $P \approx (4/3)\pi b^3 |\psi(0)|^2$. Check that you get the same answer this way.

(d) Use $b \approx 10^{-15}$ m and $a \approx 0.5 \times 10^{-10}$ m to get a numerical estimate for P. Roughly speaking, this represents the "fraction of its time that the electron spends inside the nucleus".

Problem 4.15

(a) Use the recursion formula (Equation 4.76) to confirm that when $l = n - 1$ the radial wave function takes the form

$$R_{n(n-1)} = N_n r^{n-1} e^{-r/na},$$

and determine the normalization constant N_n by direct integration.

(b) Calculate $\langle r \rangle$ and $\langle r^2 \rangle$ for states of the form $\psi_{n(n-1)m}$.

(c) Show that $\sigma_r = \langle r \rangle / \sqrt{2n+1}$ for such states. Note that the fractional spread in r decreases with increasing n (in this sense the system "begins to look classical" for large n). Sketch the radial wave functions for several values of n to illustrate this point.

4.2.2 The Spectrum of Hydrogen

In principle, if you put a hydrogen atom into some stationary state Ψ_{nlm}, it should stay there forever. However, if you *tickle* it slightly (by collision with another atom, say, or by shining light on it), then the atom may undergo a **transition** to some other stationary state—either by *absorbing* energy and moving up to a higher-energy state, or by *giving off* energy (typically in the form of electromagnetic radiation) and moving down.[16] In practice such perturbations are *always* present; transitions (or, as they are sometimes called, "quantum jumps") are constantly occurring, and the result is that a container of hydrogen gives off light (photons), whose energy corresponds to the *difference* in energy between the initial and final states:

$$E_\gamma = E_i - E_f = -13.6\,\text{eV}\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right). \qquad [4.91]$$

Now, according to the **Planck formula**,[17] the energy of a photon is proportional to its frequency:

$$E_\gamma = h\nu. \qquad [4.92]$$

Meanwhile, the *wavelength* is given by $\lambda = c/\nu$, so

$$\frac{1}{\lambda} = R\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right), \qquad [4.93]$$

where

$$R \equiv \frac{m}{4\pi c\hbar^3}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 = 1.097 \times 10^7\,\text{m}^{-1}. \qquad [4.94]$$

R is known as the **Rydberg constant**, and Equation 4.93 is the **Rydberg formula** for the spectrum of hydrogen. It was discovered empirically in the nineteenth century, and the greatest triumph of Bohr's theory was its ability to account for this result—and to calculate R in terms of the fundamental constants of nature. Transitions to the ground state ($n_f = 1$) lie in the ultraviolet; they are known to spectroscopists as the **Lyman series**. Transitions to the first excited state ($n_f = 2$) fall in the visible region; they constitute the **Balmer series**. Transitions to $n_f = 3$ (the **Paschen series**) are in the infrared, and so on (see Figure 4.5). (At room temperature, most hydrogen atoms are in the ground state; to obtain the emission spectrum, you must first pump them up into the various excited states; typically this is done by passing an electric spark through the gas.)

---

[16] By its nature, this involves a time-*de*pendent interaction, and the details will have to wait for Chapter 9; for our present purposes the actual mechanism involved is immaterial.

[17] The photon is a quantum of electromagnetic radiation; it's a relativistic object if there ever was one, and therefore outside the scope of nonrelativistic quantum mechanics. It will be useful in a few places to speak of photons and to invoke the Planck formula for their energy, but please bear in mind that this is external to the theory we are developing.

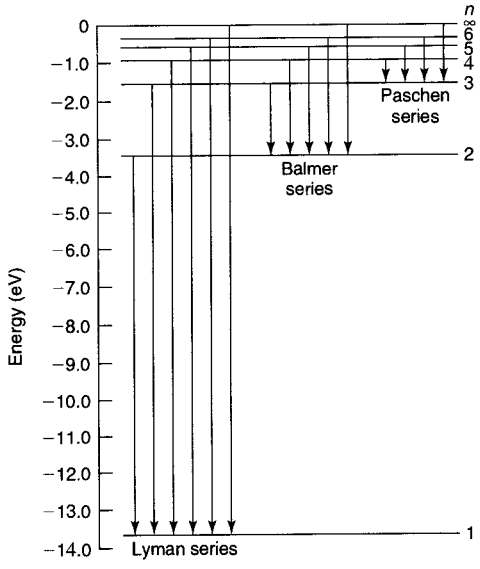


Figure 4.5: Energy levels and transitions in the spectrum of hydrogen.

Problem 4.16 Consider the earth-sun system as a gravitational analog to the hydrogen atom.

(a) What is the potential energy function (replacing Equation 4.52)? (Let m be the mass of the earth and M the mass of the sun.)

(b) What is the "Bohr radius" for this system? Work out the actual numerical value.

(c) Write down the gravitational "Bohr formula", and, by equating E_n to the classical energy of a planet in a circular orbit of radius r_0, show that $n = \sqrt{r_0/a}$. From this, estimate the quantum number n of the earth.

(d) Suppose the earth made a transition to the next lower level $(n - 1)$. How much energy (in Joules) would be released? What would the wavelength of the emitted photon (or, more likely, graviton) be?

***Problem 4.17** A **hydrogenic atom** consists of a single electron orbiting a nucleus with Z protons. ($Z = 1$ would be hydrogen itself, $Z = 2$ is ionized helium, $Z = 3$ is doubly ionized lithium, and so on.) Determine the Bohr energies $E_n(Z)$, the binding energy $E_1(Z)$, the Bohr radius $a(Z)$, and the Rydberg constant $R(Z)$ for a hydrogenic atom. (Express your answers as appropriate multiples of the hydrogen values.) Where in the electromagnetic spectrum would the Lyman series fall, for $Z = 2$ and $Z = 3$?

4.3 ANGULAR MOMENTUM

In classical mechanics, the angular momentum of a particle (with respect to the origin) is given by the formula

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \tag{4.95}$$

which is to say,

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad \text{and} \quad L_z = xp_y - yp_x. \tag{4.96}$$

The corresponding quantum operators are obtained by the standard prescription (Equation 4.2):

$$L_x = \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right); \; L_y = \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right); \\ L_z = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \tag{4.97}$$

In the following sections we will deduce the eigenvalues and eigenfunctions of these operators.

4.3.1 Eigenvalues

L_x and L_y do not commute; in fact [providing a test function, $f(x, y, z)$, for them to act upon]:

$$\begin{aligned}
[L_x, L_y]f &= \left(\frac{\hbar}{i}\right)^2 \left\{\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) f \right. \\
&\quad \left. - \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) f\right\} \\
&= \left(\frac{\hbar}{i}\right)^2 \left\{y\frac{\partial}{\partial z}\left(z\frac{\partial f}{\partial x}\right) - y\frac{\partial}{\partial z}\left(x\frac{\partial f}{\partial z}\right)\right. \\
&\quad -z\frac{\partial}{\partial y}\left(z\frac{\partial f}{\partial x}\right) + z\frac{\partial}{\partial y}\left(x\frac{\partial f}{\partial z}\right) - z\frac{\partial}{\partial x}\left(y\frac{\partial f}{\partial z}\right) \\
&\quad \left. +z\frac{\partial}{\partial x}\left(z\frac{\partial f}{\partial y}\right) + x\frac{\partial}{\partial z}\left(y\frac{\partial f}{\partial z}\right) - x\frac{\partial}{\partial z}\left(z\frac{\partial f}{\partial y}\right)\right\} \\
&= \left(\frac{\hbar}{i}\right)^2 \left(y\frac{\partial f}{\partial x} + yz\frac{\partial^2 f}{\partial z \partial x} - yx\frac{\partial^2 f}{\partial z^2} - z^2\frac{\partial^2 f}{\partial y \partial x} + zx\frac{\partial^2 f}{\partial y \partial z}\right. \\
&\quad \left. -zy\frac{\partial^2 f}{\partial x \partial z} + z^2\frac{\partial^2 f}{\partial x \partial y} + xy\frac{\partial^2 f}{\partial z^2} - x\frac{\partial f}{\partial y} - xz\frac{\partial^2 f}{\partial z \partial y}\right).
\end{aligned}$$

All the terms cancel in pairs (by virtue of the equality of cross-derivatives) except two:

$$[L_x, L_y]f = \left(\frac{\hbar}{i}\right)^2 \left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) f = i\hbar L_z f,$$

and we conclude (dropping the test function)

$$\boxed{[L_x, L_y] = i\hbar L_z.} \tag{4.98}$$

By cyclic permutation of the indices it follows also that

$$[L_y, L_z] = i\hbar L_x \quad \text{and} \quad [L_z, L_x] = i\hbar L_y. \tag{4.99}$$

From these fundamental commutation relations the entire theory of angular momentum can be deduced.

Evidently L_x, L_y, and L_z are *incompatible* observables. According to the generalized uncertainty principle (Equation 3.139),

$$\sigma_{L_x}^2 \sigma_{L_y}^2 \geq \left(\frac{1}{2i}\langle i\hbar L_z\rangle\right)^2 = \frac{\hbar^2}{4}\langle L_z\rangle^2,$$

or

$$\sigma_{L_x}\sigma_{L_y} \geq \frac{\hbar}{2}|\langle L_z\rangle|. \tag{4.100}$$

It would therefore be futile to look for states that are simultaneously eigenfunctions of L_x and of L_y. On the other hand, the *square* of the *total* angular momentum,

$$L^2 \equiv L_x^2 + L_y^2 + L_z^2, \tag{4.101}$$

does commute with L_x:

$$\begin{aligned}
[L^2, L_x] &= [L_x^2, L_x] + [L_y^2, L_x] + [L_z^2, L_x] \\
&= L_y[L_y, L_x] + [L_y, L_x]L_y + L_z[L_z, L_x] + [L_z, L_x]L_z \\
&= L_y(-i\hbar L_z) + (-i\hbar L_z)L_y + L_z(i\hbar L_y) + (i\hbar L_y)L_z \\
&= 0.
\end{aligned}$$

(I used Equation 3.142 and the fact that *any* operator commutes with *itself*.[18]) It follows, of course, that L^2 also commutes with L_y and L_z:

$$[L^2, L_x] = 0, \quad [L^2, L_y] = 0, \quad [L^2, L_z] = 0, \tag{4.102}$$

[18]Note that all the operators we encounter in quantum mechanics (see footnote 8, Chapter 1) are **linear**, in the sense that $\hat{A}(f+g) = \hat{A}f + \hat{A}g$, and therefore *distributive* with respect to addition: $\hat{A}(\hat{B}+\hat{C}) = \hat{A}\hat{B} + \hat{A}\hat{C}$. In particular, $[\hat{A}, \hat{B}+\hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}]$.

or, more compactly,

$$[L^2, \mathbf{L}] = 0. \tag{4.103}$$

So L^2 is compatible with each component of $\mathbf{L}$, and we *can* hope to find simultaneous eigenstates of L^2 and (say) L_z:

$$L^2 f = \lambda f \quad \text{and} \quad L_z f = \mu f. \tag{4.104}$$

We'll use a "ladder operator" technique, very similar to the one we applied to the harmonic oscillator back in Section 2.3.1. Let

$$L_\pm \equiv L_x \pm i L_y. \tag{4.105}$$

Its commutator with L_z is

$$[L_z, L_\pm] = [L_z, L_x] \pm i[L_z, L_y] = i\hbar L_y \pm i(-i\hbar L_x) = \pm\hbar(L_x \pm i L_y),$$

so

$$[L_z, L_\pm] = \pm\hbar L_\pm. \tag{4.106}$$

And, of course,

$$[L^2, L_\pm] = 0. \tag{4.107}$$

I claim that if f is an eigenfunction of L^2 and L_z, so also is $L_\pm f$. For Equation 4.107 says

$$L^2(L_\pm f) = L_\pm(L^2 f) = L_\pm(\lambda f) = \lambda(L_\pm f), \tag{4.108}$$

so $L_\pm f$ is an eigenfunction of L^2, with the same eigenvalue λ, and Equation 4.106 says

$$\begin{aligned} L_z(L_\pm f) &= (L_z L_\pm - L_\pm L_z) f + L_\pm L_z f = \pm\hbar L_\pm f + L_\pm(\mu f) \\ &= (\mu \pm \hbar)(L_\pm f), \end{aligned} \tag{4.109}$$

so $L_\pm f$ is an eigenfunction of L_z with the *new* eigenvalue $\mu \pm \hbar$. L_+ is called the "raising" operator because it *increases* the eigenvalue of L_z by $\hbar$, and L_- is called the "lowering" operator because it *lowers* the eigenvalue by $\hbar$.

For a given value of λ, then, we obtain a "ladder" of states, with each "rung" separated from its neighbors by one unit of $\hbar$ in the eigenvalue of L_z (see Figure 4.6). To ascend the ladder we apply the raising operator, and to descend, the lowering operator. But this process cannot go on forever: Eventually we're going to reach a state for which the z-component exceeds the *total*, and that cannot be (see Problem 4.18). So there must exist a "top rung," f_t, such that[19]

$$L_+ f_t = 0. \tag{4.110}$$

[19] Actually, all we can conclude is that $L_+ f_t$ is *not normalizable*—its norm could be *infinite*, instead of zero. Problem 4.19 eliminates this alternative.

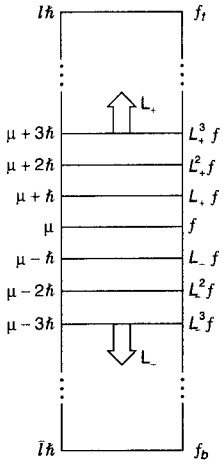


Figure 4.6: The "ladder" of angular momentum states.

Let $\hbar l$ be the eigenvalue of L_z at this top rung (the appropriateness of the letter l—sometimes called the **azimuthal quantum number**—will appear in a moment):

$$L_z f_t = \hbar l f_t; \quad L^2 f_t = \lambda f_t. \tag{4.111}$$

Now

$$\begin{aligned} L_\pm L_\mp &= (L_x \pm iL_y)(L_x \mp iL_y) = L_x^2 + L_y^2 \mp i(L_x L_y - L_y L_x) \\ &= L^2 - L_z^2 \mp i(i\hbar L_z), \end{aligned}$$

or, putting it the other way around,

$$L^2 = L_\pm L_\mp + L_z^2 \mp \hbar L_z. \tag{4.112}$$

It follows that

$$L^2 f_t = (L_- L_+ + L_z^2 + \hbar L_z) f_t = (0 + \hbar^2 l^2 + \hbar^2 l) f_t = \hbar^2 l(l+1) f_t,$$

and hence

$$\lambda = \hbar^2 l(l+1). \tag{4.113}$$

This tells us the eigenvalue of L^2 in terms of the maximum eigenvalue of L_z.

Meanwhile, there is also (for the same reason) a *bottom* rung, f_b, such that

$$L_- f_b = 0. \tag{4.114}$$

Let $\hbar\bar{l}$ be the eigenvalue of L_z at this bottom rung:

$$L_z f_b = \hbar \bar{l} f_b; \quad L^2 f_b = \lambda f_b. \tag{4.115}$$

Using Equation 4.112, we have

$$L^2 f_b = (L_+ L_- + L_z^2 - \hbar L_z) f_b = (0 + \hbar^2 \bar{l}^2 - \hbar^2 \bar{l}) f_b = \hbar^2 \bar{l}(\bar{l} - 1) f_b,$$

and therefore

$$\lambda = \hbar^2 \bar{l}(\bar{l} - 1). \qquad [4.116]$$

Comparing Equations 4.113 and 4.116, we see that $l(l + 1) = \bar{l}(\bar{l} - 1)$, so either $\bar{l} = l + 1$ (which is absurd—the bottom rung is higher than the top rung!), or else

$$\bar{l} = -l. \qquad [4.117]$$

Evidently the eigenvalues of L_z are $m\hbar$, where m (the appropriateness of this letter will also be clear in a moment) goes from $-l$ to $+l$ in N integer steps. In particular, it follows that $l = -l + N$, and hence $l = N/2$, so l must be *an integer or a half-integer*. The eigenfunctions are characterized by the numbers l and m:

$$L^2 f_l^m = \hbar^2 l(l + 1) f_l^m; \quad L_z f_l^m = \hbar m f_l^m, \qquad [4.118]$$

where

$$l = 0, 1/2, 1, 3/2, \ldots; \quad m = -l, -l + 1, \ldots, l - 1, l. \qquad [4.119]$$

For a given value of l, there are $2l + 1$ different values of m (i.e., $2l + 1$ "rungs" on the "ladder").

I hope you're impressed: By *purely algebraic means*, starting with the fundamental commutation relations (Equations 4.98 and 4.99), we have determined the eigenvalues of L^2 and L_z—without ever seeing the eigenfunctions themselves! We turn now to the problem of constructing the eigenfunctions, but I should warn you that this is a much messier business. Just so you know where we're headed, I'll tell you the punch line before we begin: $f_l^m = Y_l^m$—the eigenfunctions of L^2 and L_z are nothing but the old spherical harmonics, which we came upon by a quite different route in Section 4.1.2 (that's why I chose the letters l and m, of course).

Problem 4.18

(a) Prove that if f is simultaneously an eigenfunction of L^2 and of L_z (Equation 4.104), the square of the eigenvalue of L_z cannot exceed the eigenvalue of L^2. *Hint*: Examine the expectation value of L^2.

(b) As it turns out (see Equations 4.118 and 4.119), the square of the eigenvalue of L_z never even *equals* the eigenvalue of L^2 (except in the special case $l = m = 0$). Comment on the implications of this result. Show that it is enforced by the uncertainty principle (Equation 4.100), and explain how the special case gets away with it.

***Problem 4.19** The raising and lowering operators change the value of m by one unit:

$$L_\pm f_l^m = (A_l^m) f_l^{m\pm 1}, \qquad [4.120]$$

where A_l^m is some constant. *Question*: What *is* A_l^m, if the eigenfunctions are to be *normalized*? *Hint*: First show that $L_\mp$ is the Hermitian conjugate of $L_\pm$ (since L_x

and L_y are observables, you may assume they are Hermitian, but prove it if you like), then use Equation 4.112. *Answer*:

$$A_l^m = \hbar\sqrt{l(l+1) - m(m \pm 1)}. \quad [4.121]$$

Note what happens at the top and bottom of the ladder.

***Problem 4.20**

(a) Starting with the canonical commutation relations for position and momentum, Equation 4.10, work out the following commutators:

$$\begin{array}{lll} [L_z, x] = i\hbar y, & [L_z, y] = -i\hbar x, & [L_z, z] = 0 \\ [L_z, p_x] = i\hbar p_y, & [L_z, p_y] = -i\hbar p_x, & [L_z, p_z] = 0. \end{array} \quad [4.122]$$

(b) Use these results to obtain $[L_z, L_x] = i\hbar L_y$ directly from Equation 4.96.

(c) Evaluate the commutators $[L_z, r^2]$ and $[L_z, p^2]$ (where, of course, $r^2 = x^2 + y^2 + z^2$ and $p^2 = p_x^2 + p_y^2 + p_z^2$).

(d) Show that the Hamiltonian $H = (p^2/2m) + V$ commutes with all three components of $\mathbf{L}$, provided that V depends only on r. (Thus H, L^2, and L_z are mutually compatible observables.)

****Problem 4.21**

(a) Prove that for a particle in a potential $V(\mathbf{r})$ the rate of change of the expectation value of the orbital angular momentum $\mathbf{L}$ is equal to the expectation value of the torque:

$$\frac{d}{dt}\langle \mathbf{L} \rangle = \langle \mathbf{N} \rangle,$$

where

$$\mathbf{N} = \mathbf{r} \times (-\nabla V).$$

(This is the rotational analog to Ehrenfest's theorem.)

(b) Show that $d\langle \mathbf{L} \rangle / dt = 0$ for any spherically symmetric potential. (This is one form of the quantum statement of **conservation of angular momentum**.)

4.3.2 Eigenfunctions

First of all we need to rewrite L_x, L_y, and L_z in spherical coordinates. Now $\mathbf{L} = (\hbar/i)(\mathbf{r} \times \nabla)$, and the gradient, in spherical coordinates, is[20]

$$\nabla = \hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\phi}\frac{1}{r \sin\theta}\frac{\partial}{\partial \phi}; \quad [4.123]$$

---

[20] George Arfken, *Mathematical Methods for Physicists*, 3rd ed. (Orlando, FL: Academic Press, 1985), Section 2.5.

meanwhile, $\mathbf{r} = r\hat{r}$, so

$$\mathbf{L} = \frac{\hbar}{i}\left[r(\hat{r}\times\hat{r})\frac{\partial}{\partial r} + (\hat{r}\times\hat{\theta})\frac{\partial}{\partial\theta} + (\hat{r}\times\hat{\phi})\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\right].$$

But $(\hat{r}\times\hat{r}) = 0$, $(\hat{r}\times\hat{\theta}) = \hat{\phi}$, and $(\hat{r}\times\hat{\phi}) = -\hat{\theta}$ (see Figure 4.1), and hence

$$\mathbf{L} = \frac{\hbar}{i}\left(\hat{\phi}\frac{\partial}{\partial\theta} - \hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\right). \qquad [4.124]$$

The unit vectors $\hat{\theta}$ and $\hat{\phi}$ can be resolved into their Cartesian components:

$$\hat{\theta} = (\cos\theta\cos\phi)\hat{\imath} + (\cos\theta\sin\phi)\hat{\jmath} - (\sin\theta)\hat{k}; \qquad [4.125]$$

$$\hat{\phi} = -(\sin\phi)\hat{\imath} + (\cos\phi)\hat{\jmath}. \qquad [4.126]$$

Thus

$$\mathbf{L} = \frac{\hbar}{i}\left[(-\sin\phi\hat{\imath} + \cos\phi\hat{\jmath})\frac{\partial}{\partial\theta} - (\cos\theta\cos\phi\hat{\imath} + \cos\theta\sin\phi\hat{\jmath} - \sin\theta\hat{k})\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\right].$$

Evidently,

$$L_x = \frac{\hbar}{i}\left(-\sin\phi\frac{\partial}{\partial\theta} - \cos\phi\cot\theta\frac{\partial}{\partial\phi}\right), \qquad [4.127]$$

$$L_y = \frac{\hbar}{i}\left(+\cos\phi\frac{\partial}{\partial\theta} - \sin\phi\cot\theta\frac{\partial}{\partial\phi}\right), \qquad [4.128]$$

and

$$L_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi}. \qquad [4.129]$$

We shall also need the raising and lowering operators:

$$L_\pm = L_x \pm iL_y = \frac{\hbar}{i}\left[(-\sin\phi \pm i\cos\phi)\frac{\partial}{\partial\theta} - (\cos\phi \pm i\sin\phi)\cot\theta\frac{\partial}{\partial\phi}\right].$$

But $\cos\phi \pm i\sin\phi = e^{\pm i\phi}$, so

$$L_\pm = \pm\hbar e^{\pm i\phi}\left(\frac{\partial}{\partial\theta} \pm i\cot\theta\frac{\partial}{\partial\phi}\right). \qquad [4.130]$$

We are now in a position to determine $f_l^m(\theta,\phi)$ (I'll drop the subscript and superscript for now). It's an eigenfunction of L_z, with eigenvalue $\hbar m$:

$$L_z f = \frac{\hbar}{i}\frac{\partial f}{\partial\phi} = \hbar m f,$$

so

$$f = g(\theta)e^{im\phi}. \qquad [4.131]$$

[Here $g(\theta)$ is a constant of integration, as far as ϕ is concerned, but it can still depend on θ.] And f is also an eigenfunction of L^2 (which we'll write in terms of $L_\pm$ and L_z, using Equation 4.112), with eigenvalue $\hbar^2 l(l+1)$:

$$\begin{aligned} L^2 f &= (L_+ L_- + L_z^2 - \hbar L_z) f \\ &= \hbar e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right)\left(-\hbar e^{-i\phi}\right)\left(\frac{\partial f}{\partial\theta} - i\cot\theta\frac{\partial f}{\partial\phi}\right) - \hbar^2\frac{\partial^2 f}{\partial\phi^2} - \frac{\hbar^2}{i}\frac{\partial f}{\partial\phi} \\ &= \hbar^2 l(l+1) f. \end{aligned}$$

But in view of Equation 4.131, $\partial f/\partial\theta = e^{im\phi} dg/d\theta$ and $\partial f/\partial\phi = ime^{im\phi} g$, so

$$\begin{aligned} &-e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right)\left(e^{i(m-1)\phi}\right)\left(\frac{dg}{d\theta} + mg\cot\theta\right) + m^2 g e^{im\phi} - mge^{im\phi} \\ &= e^{im\phi}\Big[-\frac{d}{d\theta}\left(\frac{dg}{d\theta} + mg\cot\theta\right) + (m-1)\cot\theta\left(\frac{dg}{d\theta} + mg\cot\theta\right) \\ &\qquad + m(m-1)g\Big] = l(l+1)ge^{im\phi}. \end{aligned}$$

Canceling $e^{im\phi}$,

$$\begin{aligned} &-\frac{d^2 g}{d\theta^2} - m\frac{dg}{d\theta}\cot\theta + mg\csc^2\theta + (m-1)\cot\theta\frac{dg}{d\theta} + m(m-1)(1+\cot^2\theta)g \\ &= -\frac{d^2 g}{d\theta^2} - \cot\theta\frac{dg}{d\theta} + m^2 g\csc^2\theta = l(l+1)g, \end{aligned}$$

or, multiplying through by $-\sin^2\theta$:

$$\sin^2\theta\frac{d^2 g}{d\theta^2} + \sin\theta\cos\theta\frac{dg}{d\theta} - m^2 g = -l(l+1)\sin^2\theta g.$$

This is a differential equation for $g(\theta)$; it can be written in a more familiar form:

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{dg}{d\theta}\right) + [l(l+1)\sin^2\theta - m^2]g = 0. \qquad [4.132]$$

But this is precisely the equation for the θ-dependent part, $\Theta(\theta)$, of $Y_l^m(\theta, \phi)$ (compare Equation 4.25). Meanwhile, the ϕ-dependent part of f (to wit, $e^{im\phi}$) is identical to $\Phi(\phi)$ (Equation 4.22). Conclusion: The *spherical harmonics are precisely the* (normalized) *eigenfunctions of* L^2 *and* L_z.

When we solved the Schrödinger equation by separation of variables, in Section 4.1, we were inadvertantly constructing simultaneous eigenfunctions of the three commuting operators H, L^2, and L_z:

$$H\psi = E\psi, \quad L^2\psi = \hbar^2 l(l+1)\psi, \quad L_z\psi = \hbar m\psi. \qquad [4.133]$$

But there is a curious twist to this story, for the *algebraic* theory of angular momentum permits l (and hence also m) to take on *half*-integer values (Equation 4.119), whereas the *analytic* method yielded eigenfunctions only for *integer* values (Equation 4.29). You might reasonably guess that the half-integer solutions are spurious, but it turns out that they are of profound importance, as we shall see in the following sections.

***Problem 4.22**

(a) What is $L_+ Y_l^l$? (No calculation allowed!)

(b) Use the result of (a), together with the fact that $L_z Y_l^l = \hbar l Y_l^l$, to determine $Y_l^l(\theta, \phi)$, up to a normalization constant.

(c) Determine the normalization constant by direct integration. Compare your final answer to what you got in Problem 4.5.

Problem 4.23 In Problem 4.3 you showed that

$$Y_2^1(\theta, \phi) = -\sqrt{15/8\pi}\, \sin\theta \cos\theta e^{i\phi}.$$

Apply the raising operator to find $Y_2^2(\theta, \phi)$. Use Equation 4.121 to get the normalization.

Problem 4.24

(a) Prove that the spherical harmonics are orthogonal (Equation 4.33). *Hint*: This requires no calculation, if you invoke the appropriate theorem.

(b) Prove the orthogonality of the hydrogen wave functions $\psi_{nlm}(r, \theta, \phi)$ (Equation 4.90).

Problem 4.25 Two particles of mass m are attached to the ends of a massless rigid rod of length a. The system is free to rotate in three dimensions about the center (but the center point itself is fixed).

(a) Show that the allowed energies of this **rigid rotor** are

$$E_n = \frac{\hbar^2 n(n+1)}{ma^2}, \quad \text{for} \quad n = 0, 1, 2, \ldots$$

Hint: First express the (classical) energy in terms of the total angular momentum.

(b) What are the normalized eigenfunctions for this system? What is the degeneracy of the nth energy level?

4.4 SPIN

In *classical* mechanics, a rigid object admits two kinds of angular momentum: **orbital** ($\mathbf{L} = \mathbf{r} \times \mathbf{p}$), associated with the motion *of* the center of mass, and **spin** ($\mathbf{S} = I\boldsymbol{\omega}$). associated with motion *about* the center of mass. For example, the earth has orbital angular momentum attributable to its annual revolution around the sun, and spin angular momentum coming from its daily rotation about the north-south axis. In the classical context this distinction is largely a matter of convenience, for when you come right down to it, **S** is nothing but the sum total of the "orbital" angular momenta of all the rocks and dirt clods that go to make up the earth, as they circle around the axis. But an analogous thing happens in quantum mechanics, and here the distinction is absolutely fundamental. In addition to orbital angular momentum. associated (in the case of hydrogen) with the motion of the electron around the nucleus (and described by the spherical harmonics), the electron also carries *another* form of angular momentum, which has nothing to do with motion in space (and which is not, therefore, described by any function of the position variables r, θ, ϕ) but which is somewhat analogous to classical spin (and for which, therefore, we use the same word). It doesn't pay to press this analogy too far: The electron (as far as we know) is a structureless point particle, and its spin angular momentum cannot be decomposed into orbital angular momenta of constituent parts (see Problem 4.26).[21] Suffice it to say that elementary particles carry **intrinsic** angular momentum (**S**) in addition to their "extrinsic" angular momentum (**L**).

The *algebraic* theory of spin is a carbon copy of the theory of orbital angular momentum, beginning with the fundamental commutation relations[22]:

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y. \tag{4.134}$$

It follows (as before) that the eigenvectors of S^2 and S_z satisfy[23]

$$S^2|s\,m\rangle = \hbar^2 s(s+1)|s\,m\rangle; \quad S_z|s\,m\rangle = \hbar m|s\,m\rangle; \tag{4.135}$$

and

$$S_\pm|s\,m\rangle = \hbar\sqrt{s(s+1) - m(m \pm 1)}\;|s\,(m \pm 1)\rangle, \tag{4.136}$$

---

[21] For a contrary interpretation, see Hans C. Ohanian, "What is Spin?", *Am. J. Phys.* **54**, 500 (1986).

[22] We shall take these as *postulates* for the theory of spin; the analogous formulas for *orbital* angular momentum (Equations 4.98 and 4.99) were *derived* from the known form of the operators (Equation 4.97). In a more sophisticated treatment they can both be obtained from the rotational invariance of the three-dimensional world [see, for example, Leslie E. Ballentine, *Quantum Mechanics* (Englewood Cliffs, NJ: Prentice Hall, 1990), Section 3.3]. Indeed, these fundamental commutation relations apply to *all* forms of angular momentum, whether spin, orbital, or the combined angular momentum of a composite system, which could include some spin and some orbital.

[23] Because the eigenstates of spin are not *functions*, I revert to the "ket" notation for them. (I could have done the same in Section 4.3, writing $|l\,m\rangle$ in place of Y_l^m, but in that context the function notation seems more natural.) By the way, I'm running out of letters, so I'll use m for the eigenvalue of S_z, just as I did for L_z (some authors write m_l and m_s at this stage, just to be absolutely clear).

where $S_\pm \equiv S_x \pm iS_y$. But this time the eigenvectors are *not* spherical harmonics (they're not functions of θ and ϕ at all), and there is no a priori reason to exclude the half-integer values of s and m:

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots; \quad m = -s, -s+1, \ldots, s-1, s. \qquad [4.137]$$

It so happens that every elementary particle has a *specific and immutable* value of s, which we call the **spin** of that particular species: pi mesons have spin 0; electrons have spin 1/2; photons have spin 1; deltas have spin 3/2; gravitons have spin 2; and so on. By contrast, the *orbital* angular momentum quantum number l (for an electron in a hydrogen atom, say) can take on any (integer) value you please, and will change from one to another when the system is perturbed. But s is *fixed*, for any given particle, and this makes the theory of spin comparatively simple.[24]

Problem 4.26 If the electron is a classical solid sphere, with radius

$$r_c = \frac{e^2}{4\pi\epsilon_0 mc^2}, \qquad [4.138]$$

(the so-called **classical electron radius**, obtained by assuming that the electron's mass is attributable to energy stored in its electric field, via the Einstein formula $E = mc^2$), and its angular momentum is $(1/2)\hbar$, then how fast (in m/s) is a point on the "equator" moving? Does this model for spin make sense? (Actually, the radius of the electron is known experimentally to be much less than r_c, but this only makes matters worse.)

4.4.1 Spin 1/2

By far the most important case is $s = 1/2$, for this is the spin of the particles that make up ordinary matter (protons, neutrons, and electrons), as well as all quarks and all leptons. Moreover, once you understand spin 1/2, it is a simple matter to work out the formalism for any higher spin. There are just *two* eigenstates: $|\frac{1}{2}\,\frac{1}{2}\rangle$, which we call **spin up** (informally, $\uparrow$), and $|\frac{1}{2}\,(-\frac{1}{2})\rangle$, which we call **spin down** ($\downarrow$). Using these as basis vectors, the general state of a spin-1/2 particle can be expressed as a two-element column matrix (or **spinor**):

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} = a\chi_+ + b\chi_-, \qquad [4.139]$$

---

[24] Indeed, in a mathematical sense, spin 1/2 is the simplest possible nontrivial quantum system, for it admits just two possible states. In place of an infinite-dimensional Hilbert space, with all its subtleties and complications, we find ourselves working in an ordinary two-dimensional vector space; in place of unfamiliar differential equations and fancy functions, we are confronted with 2×2 matrices and two-component vectors. For this reason, some authors *begin* quantum mechanics with a treatment of the spin-1/2 system. But the price of mathematical simplicity is conceptual abstraction, and I prefer not to do it that way.

with

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{4.140}$$

representing spin up, and

$$\chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.141}$$

for spin down. Meanwhile, the spin operators become 2×2 matrices, which we can work out by noting their effect on χ_+ and χ_-: Equation 4.135 says

$$S^2\chi_+ = \frac{3}{4}\hbar^2\chi_+;\ S^2\chi_- = \frac{3}{4}\hbar^2\chi_-;\ S_z\chi_+ = \frac{\hbar}{2}\chi_+;\ S_z\chi_- = -\frac{\hbar}{2}\chi_-; \tag{4.142}$$

and Equation 4.136 gives

$$S_+\chi_- = \hbar\chi_+; \quad S_-\chi_+ = \hbar\chi_-; \quad S_+\chi_+ = S_-\chi_- = 0. \tag{4.143}$$

Now, $S_\pm = S_x \pm iS_y$, so

$$S_x = \frac{1}{2}(S_+ + S_-) \quad \text{and} \quad S_y = \frac{1}{2i}(S_+ - S_-), \tag{4.144}$$

and it follows that

$$S_x\chi_+ = \frac{\hbar}{2}\chi_-;\ S_x\chi_- = \frac{\hbar}{2}\chi_+;\ S_y\chi_+ = -\frac{\hbar}{2i}\chi_-;\ S_y\chi_- = \frac{\hbar}{2i}\chi_+. \tag{4.145}$$

Thus

$$S^2 = \frac{3}{4}\hbar^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};\ S_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix};\ S_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \tag{4.146}$$

while

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix};\ S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix};\ S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.147}$$

It's a little tidier to divide off the factor of $\hbar/2$: $\mathbf{S} = (\hbar/2)\boldsymbol{\sigma}$, where

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix};\ \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix};\ \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.148}$$

These are the famous **Pauli spin matrices**. Notice that S_x, S_y, S_z, and S^2 are all *Hermitian* (as they should be, since they represent observables). On the other hand, S_+ and S_- are *not* Hermitian—evidently they are not observable.

The eigenspinors of S_z are (of course)

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ (\text{eigenvalue } +\frac{\hbar}{2}); \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ (\text{eigenvalue } -\frac{\hbar}{2}). \tag{4.149}$$

If you measure S_z on a particle in the general state χ (Equation 4.139), you could get $+\hbar/2$, with probability $|a|^2$, or $-\hbar/2$, with probability $|b|^2$. Since these are the *only* possibilities,

$$|a|^2 + |b|^2 = 1 \qquad [4.150]$$

(i.e., the spinor must be normalized).[25]

But what if, instead, you chose to measure S_x? What are the possible results, and what are their respective probabilities? According to the generalized statistical interpretation, we need to know the eigenvalues and eigenspinors of S_x. The characteristic equation is

$$\begin{vmatrix} -\lambda & \hbar/2 \\ \hbar/2 & -\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 = \left(\frac{\hbar}{2}\right)^2 \Rightarrow \lambda = \pm\frac{\hbar}{2}.$$

Not surprisingly, the possible values for S_x are the same as those for S_z. The eigenspinors are obtained in the usual way:

$$\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \pm\frac{\hbar}{2}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \Rightarrow \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \pm\begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

so $\beta = \pm\alpha$. Evidently the (normalized) eigenspinors of S_x are

$$\chi_+^{(x)} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \text{ (eigenvalue } +\frac{\hbar}{2}\text{)}; \quad \chi_-^{(x)} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \text{ (eigenvalue } -\frac{\hbar}{2}\text{)}. \qquad [4.151]$$

As the eigenvectors of a Hermitian matrix, they span the space; the generic spinor χ can be expressed as a linear combination of them:

$$\chi = \left(\frac{a+b}{\sqrt{2}}\right)\chi_+^{(x)} + \left(\frac{a-b}{\sqrt{2}}\right)\chi_-^{(x)}. \qquad [4.152]$$

If you measure S_x, the probability of getting $+\hbar/2$ is $(1/2)|a+b|^2$, and the probability of getting $-\hbar/2$ is $(1/2)|a-b|^2$. (You should check for yourself that these probabilities add up to 1.)

Example. Suppose a spin 1/2 particle is in the state

$$\chi = \frac{1}{\sqrt{6}}\begin{pmatrix} 1+i \\ 2 \end{pmatrix}.$$

If you measure S_z, the probability of getting $+\hbar/2$ is $|(1+i)/\sqrt{6}|^2 = 1/3$, and the probability of getting $-\hbar/2$ is $|2/\sqrt{6}|^2 = 2/3$. If you measure S_x, the probability of

---

[25] People often say that $|a|^2$ is the "probability that the particle is in the spin-up state", but this is sloppy language; the particle is in state χ—*not* χ_+—and what the speaker really *means* is that if you measured S_z, $|a|^2$ is the probability you'd get $\hbar/2$, which is an entirely different assertion.

getting $+\hbar/2$ is $(1/2)|(3+i)/\sqrt{6}|^2 = 5/6$, and the probability of getting $-\hbar/2$ is $(1/2)|(-1+i)/\sqrt{6}|^2 = 1/6$. Evidently the *expectation* value of S_x is

$$\frac{5}{6}\left(+\frac{\hbar}{2}\right) + \frac{1}{6}\left(-\frac{\hbar}{2}\right) = \frac{\hbar}{3},$$

which we could also have obtained more directly:

$$\langle S_x \rangle = \chi^\dagger S_x \chi = \begin{pmatrix} \frac{1-i}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 0 & \hbar/2 \\ \hbar/2 & 0 \end{pmatrix} \begin{pmatrix} (1+i)/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix} = \frac{\hbar}{3}.$$

I'd like now to walk you through an imaginary measurement scenario involving spin 1/2, because it serves to illustrate in very concrete terms some of the abstract ideas we discussed back in Chapter 1. Let's say we start out with a particle in the state χ_+. If someone asks, "What is the z-component of that particle's spin angular momentum?", we could answer unambiguously: $+\hbar/2$. For a measurement of S_z is certain to return that value. But if our interrogator asks instead, "What is the x-component of that particle's spin angular momentum?", we are obliged to equivocate: If you measure S_x, the chances are 50-50 of getting either $\hbar/2$ or $-\hbar/2$. If the questioner is a classical physicist, or a "realist" (in the sense of Section 1.2), he will regard this as an inadequate—not to say impertinent—response: "Are you telling me that you *don't know* the true state of that particle?" On the contrary; I know *precisely* what the state of the particle is: χ_+. "Well, then, how come you can't tell me what the x-component of its spin is?" Because it simply *does not have* a particular x-component of spin. Indeed, it *cannot*, for if both S_x and S_z were well defined, the uncertainty principle would be violated.

At this point our challenger grabs the test tube and *measures* the x-component of its spin; let's say he gets the value $+\hbar/2$. "Aha!" (he shouts in triumph), "You *lied*! This particle has a perfectly well-defined value of S_x: It's $\hbar/2$." Well, sure—it does *now*, but that doesn't prove it *had* that value, prior to your measurement. "You have obviously been reduced to splitting hairs. And anyway, what happened to your uncertainty principle? I now know both S_x *and* S_z." I'm sorry, but you do *not*: In the course of your measurement, you altered the particle's state; it is now in the state $\chi_+^{(x)}$, and whereas you know the value of S_x, you no longer know the value of S_z. "But I was extremely careful not to disturb the particle when I measured S_x."[26] Very well, if you don't believe me, *check it out*: Measure S_z, and see what you get. (Of course, he *may* get $+\hbar/2$, which will be embarrassing to my case—but if we repeat this whole scenario over and over, half the time he will get $-\hbar/2$.)

To the layperson, the philosopher, or the classical physicist, a statement of the form "this particle doesn't have a well-defined position" (or momentum, or

---

[26]Neils Bohr was anxious to track down the *mechanism* by which the measurement of S_x inevitably destroys the value of S_z, in gedanken experiments of this sort. His famous debates with Einstein include many delightful examples, showing in detail how experimental constraints serve to enforce the uncertainty principle.

x-component of spin angular momentum, or whatever) sounds vague, incompetent, or (worst of all) profound. It is none of these. But its precise meaning is, I think, almost impossible to convey to anyone who has not studied quantum mechanics in some depth. If you find your own comprehension slipping, from time to time (if you *don't*, you probably haven't understood the problem), come back to the spin-1/2 system: It is the simplest and cleanest context for thinking through the conceptual paradoxes of quantum mechanics.

Problem 4.27

(a) Check that the spin matrices (Equation 4.147) obey the fundamental commutation relation for angular momentum: $[S_x, S_y] = i\hbar S_z$.

(b) Show that the Pauli spin matrices satisfy

$$\sigma_j \sigma_k = \delta_{jk} + i \sum_l \epsilon_{jkl} \sigma_l, \qquad [4.153]$$

where the indices stand for x, y, or z, and ϵ_{jkl} is the **Levi-Cirita** symbol: $+1$ if $jkl = 123,231$, or 312; -1 if $jkl = 132,213$, or 321; 0 otherwise.

***Problem 4.28** An electron is in the spin state

$$\chi = A \begin{pmatrix} 3i \\ 4 \end{pmatrix}.$$

(a) Determine the normalization constant A.

(b) Find the expectation values of S_x, S_y, and S_z.

(c) Find the "uncertainties" σ_{S_x}, σ_{S_y}, and σ_{S_z}.

(d) Confirm that your results are consistent with all three uncertainty principles (Equation 4.100 and its cyclic permutations—only with S in place of L, of course).

***Problem 4.29** For the most general normalized spinor χ (Equation 4.139), compute $\langle S_x \rangle$, $\langle S_y \rangle$, $\langle S_z \rangle$, $\langle S_x^2 \rangle$, $\langle S_y^2 \rangle$, and $\langle S_z^2 \rangle$. Check that $\langle S_x^2 \rangle + \langle S_y^2 \rangle + \langle S_z^2 \rangle = \langle S^2 \rangle$.

***Problem 4.30**

(a) Find the eigenvalues and eigenspinors of S_y.

(b) If you measured S_y on a particle in the general state χ (Equation 4.139), what values might you get, and what is the probability of each? Check that the probabilities add up to 1.

(c) If you measured S_y^2, what values might you get and with what probability?

Problem 4.31 Construct the matrix S_r representing the component of spin angular momentum along an arbitrary direction $\hat{r}$. Use spherical coordinates, so that

$$\hat{r} = \sin\theta\cos\phi\,\hat{\imath} + \sin\theta\sin\phi\,\hat{\jmath} + \cos\theta\,\hat{k}. \quad [4.154]$$

Find the eigenvalues and (normalized) eigenspinors of S_r. *Answer*:

$$\chi_+^{(r)} = \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\sin(\theta/2) \end{pmatrix}; \quad \chi_-^{(r)} = \begin{pmatrix} \sin(\theta/2) \\ -e^{i\phi}\cos(\theta/2) \end{pmatrix}. \quad [4.155]$$

Problem 4.32 Construct the spin matrices (S_x, S_y, and S_z) for a particle of spin 1. *Hint*: How many eigenstates of S_z are there? Determine the action of S_z, S_+, and S_- on each of these states. Follow the procedure used in the text for spin 1/2.

4.4.2 Electron in a Magnetic Field

A spinning charged particle constitutes a magnetic dipole. Its **magnetic dipole moment** $\boldsymbol{\mu}$ is proportional to its spin angular momentum $\mathbf{S}$:

$$\boldsymbol{\mu} = \gamma\mathbf{S}; \quad [4.156]$$

the proportionality constant γ is called the **gyromagnetic ratio**.[27] When a magnetic dipole is placed in a magnetic field $\mathbf{B}$, it experiences a torque, $\boldsymbol{\mu} \times \mathbf{B}$, which tends to line it up parallel to the field (just like a compass needle). The energy associated with this torque is[28]

$$H = -\boldsymbol{\mu}\cdot\mathbf{B}, \quad [4.157]$$

so the Hamiltonian of a spinning charged particle, at rest[29] in a magnetic field $\mathbf{B}$, becomes

$$H = -\gamma\mathbf{B}\cdot\mathbf{S}, \quad [4.158]$$

where $\mathbf{S}$ is the appropriate spin matrix (Equation 4.147, in the case of spin 1/2).

Example: Larmor precession. Imagine a particle of spin 1/2 at rest in a uniform magnetic field, which points in the z-direction:

$$\mathbf{B} = B_0\hat{k}. \quad [4.159]$$

[27] See, for example, D. Griffiths, *Introduction to Electrodynamics*, 2nd ed. (Englewood Cliffs, NJ: Prentice Hall, 1986), page 239. Classically, the gyromagnetic ratio of a rigid object is $q/2m$, where q is its charge and m is its mass. For reasons that are fully explained only in relativistic quantum theory, the gyromagnetic ratio of the electron is almost exactly *twice* the classical value.

[28] Griffiths, (footnote 27), pages 246 and 268.

[29] If the particle is allowed to *move*, there will also be kinetic energy to consider; moreover, it will be subject to the Lorentz force ($q\mathbf{v}\times\mathbf{B}$), which is not derivable from a potential energy function and hence does not fit the Schrödinger equation as we have formulated it so far. I'll show you later on how to handle this problem, but for the moment let's just assume that the particle is free to *rotate*, but otherwise stationary.

The Hamiltonian matrix is

$$H = -\gamma B_0 S_z = -\frac{\gamma B_0 \hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad [4.160]$$

The eigenstates of H are the same as those of S_z:

$$\begin{cases} \chi_+, & \text{with energy } E_+ = -(\gamma B_0 \hbar)/2, \\ \chi_-, & \text{with energy } E_- = +(\gamma B_0 \hbar)/2. \end{cases} \qquad [4.161]$$

Evidently, the energy is lowest when the dipole moment is parallel to the field—just as it would be classically.

Since the Hamiltonian is time independent, the general solution to the time-dependent Schrödinger equation,

$$i\hbar \frac{\partial \chi}{\partial t} = H\chi, \qquad [4.162]$$

can be expressed in terms of the stationary states:

$$\chi(t) = a\chi_+ e^{-iE_+ t/\hbar} + b\chi_- e^{-iE_- t/\hbar} = \begin{pmatrix} ae^{i\gamma B_0 t/2} \\ be^{-i\gamma B_0 t/2} \end{pmatrix}.$$

The constants a and b are determined by the initial conditions; say

$$\chi(0) = \begin{pmatrix} a \\ b \end{pmatrix},$$

where $|a|^2 + |b|^2 = 1$. With no essential loss of generality[30] I'll write $a = \cos(\alpha/2)$ and $b = \sin(\alpha/2)$, where α is a fixed angle whose physical significance will appear in a moment. Thus

$$\chi(t) = \begin{pmatrix} \cos(\alpha/2)e^{i\gamma B_0 t/2} \\ \sin(\alpha/2)e^{-i\gamma B_0 t/2} \end{pmatrix}. \qquad [4.163]$$

To get a feel for what is happening here, let's calculate the expectation value of the spin $\langle \mathbf{S} \rangle$ as a function of time:

$$\begin{aligned} \langle S_x \rangle &= \chi(t)^\dagger S_x \chi(t) \\ &= \left(\cos(\alpha/2)e^{-i\gamma B_0 t/2} \quad \sin(\alpha/2)e^{i\gamma B_0 t/2}\right) \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \cos(\alpha/2)e^{i\gamma B_0 t/2} \\ \sin(\alpha/2)e^{-i\gamma B_0 t/2} \end{pmatrix} \\ &= \frac{\hbar}{2}\sin\alpha\cos(\gamma B_0 t). \end{aligned} \qquad [4.164]$$

Similarly,

$$\langle S_y \rangle = \chi(t)^\dagger S_y \chi(t) = -\frac{\hbar}{2}\sin\alpha\sin(\gamma B_0 t), \qquad [4.165]$$

---

[30] This does assume that a and b are *real*; you can work out the general case if you like, but all it does is add a constant to t.

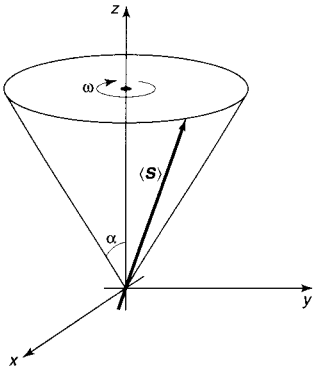


Figure 4.7: Precession of $\langle \mathbf{S} \rangle$ in a uniform magnetic field.

and

$$\langle S_z \rangle = \chi(t)^\dagger S_z \chi(t) = \frac{\hbar}{2} \cos\alpha. \qquad [4.166]$$

Evidently $\langle \mathbf{S} \rangle$ is tilted at a constant angle α to the z-axis, and precesses about the field at the **Larmor frequency**

$$\omega = \gamma B_0, \qquad [4.167]$$

just as it would classically[31] (see Figure 4.7). No *surprise* here—Ehrenfest's theorem (in the form derived in Problem 4.21) guarantees that $\langle \mathbf{S} \rangle$ evolves according to the classical laws. But it's nice to see how this works out in a specific context.

Example: the Stern-Gerlach experiment. In an *inhomogeneous* magnetic field, there is not only a torque, but also a force, on a magnetic dipole[32]:

$$\mathbf{F} = \nabla(\boldsymbol{\mu} \cdot \mathbf{B}). \qquad [4.168]$$

This force can be used to separate out particles with a particular spin orientation. as follows. Imagine a beam of relatively heavy neutral atoms,[33] traveling in the y-direction, which passes through a region of inhomogeneous magnetic field (Figure 4.8)—for example,

$$\mathbf{B}(x, y, z) = -\alpha x \hat{\imath} + (B_0 + \alpha z)\hat{k}, \qquad [4.169]$$

where B_0 is a strong uniform field and the constant α describes a small deviation from homogeneity. (Actually, what we'd *like* is just the z-component of this field, but

---

[31] See, for instance, *The Feynman Lectures on Physics* (Reading, MA: Addison-Wesley, 1964). Volume II, Section 34-3. Of course, in the classical case it is the angular momentum vector itself, not just its expectation value, that precesses around the magnetic field.

[32] Griffiths, (footnote 27), page 247. Note that $\mathbf{F}$ is the negative gradient of the energy (Equation 4.157).

[33] We make them neutral to avoid the large-scale deflection that would otherwise result from the Lorentz force, and heavy so we can construct localized wave packets and treat the motion in terms of classical particle trajectories. In practice, the Stern-Gerlach experiment doesn't work, for example, with a beam of free electrons.

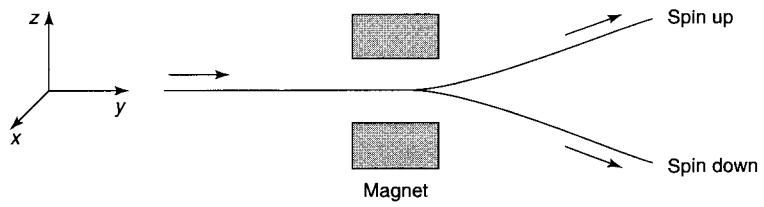


Figure 4.8: The Stern-Gerlach apparatus.

unfortunately that's impossible—it would violate the electromagnetic law $\nabla \cdot \mathbf{B} = 0$; like it or not, the x-component comes along for the ride.) The force on these atoms is

$$\mathbf{F} = \gamma\alpha(-S_x\hat{\imath} + S_z\hat{k}).$$

But because of the Larmor precession about B_0, S_x oscillates rapidly, and averages to zero; the *net* force is in the z-direction:

$$F_z = \gamma\alpha S_z, \qquad [4.170]$$

and the beam is deflected up or down, in proportion to the z-component of the spin angular momentum. Classically we'd expect a smear, but in fact the beam splits into $2s + 1$ individual beams, beautifully demonstrating the quantization of S_z. (If you use silver atoms, for example, all the inner electrons are paired in such a way that their spin and orbital angular momenta cancel. The net spin is simply that of the outermost—unpaired—electron, so in this case $s = 1/2$, and the beam splits in two.)

That argument was purely *classical*, up to the final step; "force" has no place in a proper quantum calculation, and you might therefore prefer the following approach to the same problem.[34] We examine the process from the perspective of a reference frame that moves along with the beam. In this frame the Hamiltonian starts out zero, turns on for a time T (as the particle passes through the magnet), and then turns off again:

$$H(t) = \begin{cases} 0, & \text{for } t < 0, \\ -\gamma(B_0 + \alpha z)S_z, & \text{for } 0 \le t \le T, \\ 0, & \text{for } t > T. \end{cases} \qquad [4.171]$$

(I ignore the pesky x-component of $\mathbf{B}$, which—for reasons indicated above—is irrelevant to the problem.) Suppose the atom has spin 1/2, and starts out in the state

$$\chi(t) = a\chi_+ + b\chi_-, \quad \text{for } t \le 0.$$

While the Hamiltonian acts, $\chi(t)$ evolves in the usual way:

$$\chi(t) = a\chi_+ e^{-iE_+t/\hbar} + b\chi_- e^{-iE_-t/\hbar}, \text{ for } 0 \le t \le T,$$

[34]This argument follows L. Ballentine, *Quantum Mechanics* (Englewood Cliffs, NJ: Prentice Hall, 1990), page 172.

where (from Equation 4.158)

$$E_{\pm} = \mp\gamma(B_0 + \alpha z)\frac{\hbar}{2}, \qquad [4.172]$$

and hence it emerges in the state

$$\chi(t) = \left(ae^{i\gamma T B_0/2}\chi_+\right)e^{i(\alpha\gamma T/2)z} + \left(be^{-i\gamma T B_0/2}\chi_-\right)e^{-i(\alpha\gamma T/2)z}, \ (t \geq T). \qquad [4.173]$$

The two terms now carry *momentum* in the z-direction (see Equation 3.131); the spin-up component has momentum

$$p_z = \frac{\alpha\gamma T\hbar}{2}, \qquad [4.174]$$

and it moves in the plus-z direction; the spin-down component has the opposite momentum, and it moves in the minus-z direction. Thus the beam splits in two, as before. (Note that Equation 4.174 is consistent with the earlier result, Equation 4.170, for in this case $S_z = \hbar/2$ and $p_z = F_z T$.)

The Stern-Gerlach experiment has played an important role in the philosophy of quantum mechanics, where it serves both as the prototype for the preparation of a quantum state and as an illuminating model for a certain kind of quantum measurement. We casually assume that the initial state of a system is *known* (the Schrödinger equation tells us how it subsequently evolves)—but it is natural to wonder how you get a system into a particular state in the first place. Well, if you want to prepare a beam of atoms in a given spin configuration, you pass an unpolarized beam through a Stern-Gerlach magnet and select the outgoing stream you are interested in (closing off the others with suitable baffles and shutters). Conversely, if you want to *measure* the z-component of an atom's spin, you send it through a Stern-Gerlach apparatus and record which bin it lands in. I do not claim that this is always the most *practical* way to do the job, but it is *conceptually* very clean and hence a useful context in which to explore the problems of state preparation and measurement.

Problem 4.33 In the first example (Larmor precession in a uniform magnetic field):

(a) If you measured the component of spin angular momentum along the x-direction, at time t, what is the probability that you would get $+\hbar/2$?

(b) Same question, but for the y-component.

(c) Same, but for the z-component.

****Problem 4.34** An electron is at rest in an oscillating magnetic field

$$\mathbf{B} = B_0\cos(\omega t)\hat{k},$$

where B_0 and ω are constants.

(a) Construct the Hamiltonian matrix for this system.

(b) The electron starts out (at $t = 0$) in the spin-up state with respect to the x-axis [that is, $\chi(0) = \chi_+^{(x)}$]. Determine $\chi(t)$ at any subsequent time. *Beware:* This is a time-*dependent* Hamiltonian, so you cannot get $\chi(t)$ in the usual way from stationary states. Fortunately, in this case you can solve the time-dependent Schrödinger equation (Equation 4.162) directly.

(c) Find the probability of getting $-\hbar/2$ if you measure S_x. *Answer*:

$$\sin^2\left(\frac{\gamma B_0}{2\omega}\sin(\omega t)\right).$$

(d) What is the minimum field (B_0) required to force a complete flip in S_x?

4.4.3 Addition of Angular Momenta

Suppose now that we have *two* spin-1/2 particles—for example, the electron and the proton in the ground state[35] of hydrogen. Each can have spin up or spin down, so there are four possibilities in all[36]:

$$\uparrow\uparrow,\ \uparrow\downarrow,\ \downarrow\uparrow,\ \downarrow\downarrow, \qquad [4.175]$$

where the first arrow refers to the electron and the second to the proton. *Question*: What is the *total* angular momentum of the atom? Let

$$\mathbf{S} \equiv \mathbf{S}^{(1)} + \mathbf{S}^{(2)}. \qquad [4.176]$$

Each of the four composite states is an eigenstate of S_z—the z-components simply *add*

$$\begin{aligned} S_z\chi_1\chi_2 &= (S_z^{(1)} + S_z^{(2)})\chi_1\chi_2 = (S_z^{(1)}\chi_1)\chi_2 + \chi_1(S_z^{(2)}\chi_2) \\ &= (\hbar m_1\chi_1)\chi_2 + \chi_1(\hbar m_2\chi_2) = \hbar(m_1 + m_2)\chi_1\chi_2, \end{aligned}$$

[note that $\mathbf{S}^{(1)}$ acts only on χ_1, and $\mathbf{S}^{(2)}$ acts only on χ_2]. So m (the quantum number for the composite system) is just $m_1 + m_2$:

$$\begin{aligned} \uparrow\uparrow\text{:}\ m &= 1; \\ \uparrow\downarrow\text{:}\ m &= 0; \\ \downarrow\uparrow\text{:}\ m &= 0; \\ \downarrow\downarrow\text{:}\ m &= -1. \end{aligned}$$

---

[35] I put them in the ground state so there won't be any *orbital* angular momentum to worry about.

[36] More precisely, each particle is in a *linear combination* of spin up and spin down, and the composite system is in a *linear combination* of the four states listed.

At first glance, this doesn't look right: m is supposed to advance in integer steps, from $-s$ to $+s$, so it appears that $s = 1$—but there is an extra state with $m = 0$. One way to untangle this problem is to apply the lowering operator $S_- = S_-^{(1)} + S_-^{(2)}$ to the state $\uparrow\uparrow$, using Equation 4.143:

$$\begin{aligned} S_-(\uparrow\uparrow) &= (S_-^{(1)}\uparrow)\uparrow + \uparrow(S_-^{(2)}\uparrow) \\ &= (\hbar\downarrow)\uparrow + \uparrow(\hbar\downarrow) = \hbar(\downarrow\uparrow + \uparrow\downarrow). \end{aligned}$$

Evidently the three states with $s = 1$ are (in the notation $|s\,m\rangle$):

$$\left\{\begin{array}{lcl} |1\,1\rangle &=& \uparrow\uparrow \\ |1\,0\rangle &=& \frac{1}{\sqrt{2}}(\uparrow\downarrow + \downarrow\uparrow) \\ |1\,{-1}\rangle &=& \downarrow\downarrow \end{array}\right\} \quad s = 1 \text{ (triplet)}. \qquad [4.177]$$

(As a check, try applying the lowering operator to $|1\,0\rangle$; what *should* you get? See Problem 4.35.) This is called the **triplet** combination, for the obvious reason. Meanwhile, the orthogonal state with $m = 0$ carries $s = 0$:

$$\left\{|0\,0\rangle = \frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow)\right\} \quad s = 0 \text{ (singlet)}. \qquad [4.178]$$

(If you apply the raising or lowering operator to this state, you'll get zero. See Problem 4.35.)

I claim, then, that the combination of two spin-1/2 particles can carry a total spin of 1 or 0, depending on whether they occupy the triplet or the singlet configuration. To confirm this, I need to prove that the triplet states are eigenvectors of S^2 with eigenvalue $2\hbar^2$ and the singlet is an eigenvector of S^2 with eigenvalue 0. Now

$$S^2 = (\mathbf{S}^{(1)} + \mathbf{S}^{(2)}) \cdot (\mathbf{S}^{(1)} + \mathbf{S}^{(2)}) = (S^{(1)})^2 + (S^{(2)})^2 + 2\mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)}. \qquad [4.179]$$

Using Equations 4.142 and 4.145, we have

$$\begin{aligned} \mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)}(\uparrow\downarrow) &= (S_x^{(1)}\uparrow)(S_x^{(2)}\downarrow) + (S_y^{(1)}\uparrow)(S_y^{(2)}\downarrow) + (S_z^{(1)}\uparrow)(S_z^{(2)}\downarrow) \\ &= \left(\frac{\hbar}{2}\downarrow\right)\left(\frac{\hbar}{2}\uparrow\right) + \left(\frac{i\hbar}{2}\downarrow\right)\left(\frac{-i\hbar}{2}\uparrow\right) + \left(\frac{\hbar}{2}\uparrow\right)\left(\frac{-\hbar}{2}\downarrow\right) \\ &= \frac{\hbar^2}{4}(2\downarrow\uparrow - \uparrow\downarrow). \end{aligned}$$

Similarly,

$$\mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)}(\downarrow\uparrow) = \frac{\hbar^2}{4}(2\uparrow\downarrow - \downarrow\uparrow).$$

It follows that

$$\mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)}|1\,0\rangle = \frac{\hbar^2}{4}\frac{1}{\sqrt{2}}(2\downarrow\uparrow - \uparrow\downarrow + 2\uparrow\downarrow - \downarrow\uparrow) = \frac{\hbar^2}{4}|1\,0\rangle, \qquad [4.180]$$

and

$$\mathbf{S}^{(1)} \cdot \mathbf{S}^{(2)}|0\,0\rangle = \frac{\hbar^2}{4}\frac{1}{\sqrt{2}}(2\downarrow\uparrow - \uparrow\downarrow - 2\uparrow\downarrow + \downarrow\uparrow) = -\frac{3\hbar^2}{4}|0\,0\rangle. \quad [4.181]$$

Returning to Equation 4.179 (and again using Equation 4.142), we conclude that

$$S^2|1\,0\rangle = \left(\frac{3\hbar^2}{4} + \frac{3\hbar^2}{4} + 2\frac{\hbar^2}{4}\right)|1\,0\rangle = 2\hbar^2|1\,0\rangle, \quad [4.182]$$

so $|1\,0\rangle$ is indeed an eigenstate of S^2 with eigenvalue $2\hbar^2$; and

$$S^2|0\,0\rangle = \left(\frac{3\hbar^2}{4} + \frac{3\hbar^2}{4} - 2\frac{3\hbar^2}{4}\right)|0\,0\rangle = 0, \quad [4.183]$$

so $|0\,0\rangle$ is an eigenstate of S^2 with eigenvalue 0. (I will leave it for you to confirm that $|1\,1\rangle$ and $|1\,-1\rangle$ are eigenstates of S^2, with the appropriate eigenvalue—see Problem 4.35.)

What we have just done (combining spin 1/2 with spin 1/2 to get spin 1 and spin 0) is the simplest example of a larger problem: If you combine spin s_1 with spin s_2, what total spins s can you get?[37] The answer[38] is that you get every spin from $(s_1 + s_2)$ down to $(s_1 - s_2)$—or $(s_2 - s_1)$, if $s_2 > s_1$—in integer steps:

$$s = (s_1 + s_2),\ (s_1 + s_2 - 1),\ (s_1 + s_2 - 2),\ \ldots,\ |s_1 - s_2|. \quad [4.184]$$

(Roughly speaking, the highest total spin occurs when the individual spins are aligned parallel to one another, and the lowest occurs when they are antiparallel.) For example, if you package together a particle of spin 3/2 with a particle of spin 2, you could get a total spin of 7/2, 5/2, 3/2, or 1/2, depending on the configuration. Another example: If a hydrogen atom is in the state ψ_{nlm}, the net angular momentum of the electron (spin plus orbital) is $l + 1/2$ or $l - 1/2$; if you now throw in the spin of the proton, the atom's total angular momentum quantum number is $l + 1$, l, or $l - 1$ (and l can be achieved in two distinct ways, depending on whether the electron alone is in the $l + 1/2$ configuration or the $l - 1/2$ configuration).

The particular state $|s\,m\rangle$ with total spin s and z-component m will be some linear combination of the composite states $|s_1\,m_1\rangle|s_2\,m_2\rangle$:

$$|s\,m\rangle = \sum_{m_1+m_2=m} C^{s_1 s_2 s}_{m_1 m_2 m}|s_1\,m_1\rangle|s_2\,m_2\rangle \quad [4.185]$$

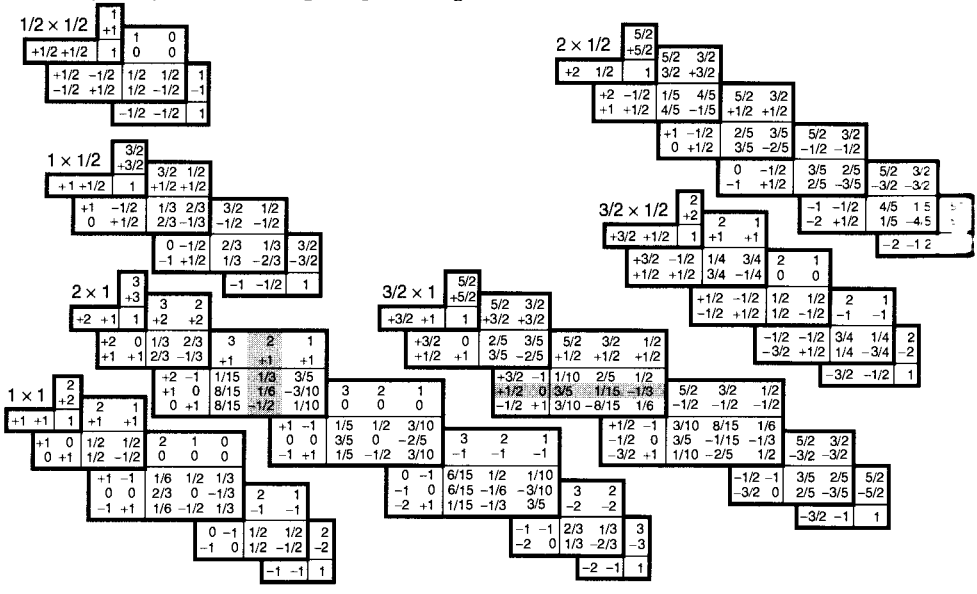
(because the z-components add, the only composite states that contribute are those for which $m_1 + m_2 = m$). Equations 4.177 and 4.178 are special cases of this general

---

[37] I say *spins* for simplicity, but either one (or both) could just as well be *orbital* angular momentum (for which, however, we would use the letter l).

[38] For a proof you must look in a more advanced text; see, for instance, Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloë, *Quantum Mechanics* (New York: John Wiley & Sons, 1977), Vol. 2, Chapter X.

Table 4.7: Clebsch-Gordan coefficients. (A square root sign is understood for every entry; the minus sign, if present, goes *outside* the radical.)



form, with $s_1 = s_2 = 1/2$ (I used the informal notation $\uparrow = |\frac{1}{2}\,\frac{1}{2}\rangle$, $\downarrow = |\frac{1}{2}\,(-\frac{1}{2})\rangle$). The constants $C^{s_1 s_2 s}_{m_1 m_2 m}$ are called **Clebsch-Gordan coefficients**. A few of the simplest cases are listed in Table 4.7.[39] For example, the shaded column of the 2×1 table tells us that

$$|2\,1\rangle = \frac{1}{\sqrt{3}}|2\,2\rangle|1\,{-1}\rangle + \frac{1}{\sqrt{6}}|2\,1\rangle|1\,0\rangle - \frac{1}{\sqrt{2}}|2\,0\rangle|1\,1\rangle.$$

In particular, if two particles (of spin 2 and spin 1) are at rest in a box, and the total spin is 2, and its z-component is 1, then a measurement of $S_z^{(1)}$ could return the value $2\hbar$ (with probability 1/3), or $\hbar$ (with probability 1/6), or 0 (with probability 1/2). Notice that the probabilities add up to 1 (the sum of the squares of any column on the Clebsch-Gordan table is 1).

These tables also work the other way around:

$$|s_1\,m_1\rangle|s_2\,m_2\rangle = \sum_s C^{s_1 s_2 s}_{m_1 m_2 m}|s\,m\rangle. \qquad [4.186]$$

For example, the shaded row in the $3/2 \times 1$ table tells us that

$$|\tfrac{3}{2}\,\tfrac{1}{2}\rangle|1\,0\rangle = \sqrt{\tfrac{3}{5}}\,|\tfrac{5}{2}\,\tfrac{1}{2}\rangle + \sqrt{\tfrac{1}{15}}\,|\tfrac{3}{2}\,\tfrac{1}{2}\rangle - \sqrt{\tfrac{1}{3}}\,|\tfrac{1}{2}\,\tfrac{1}{2}\rangle.$$

[39] The general formula is derived in Arno Bohm, *Quantum Mechanics: Foundations and Applications*, 2nd ed. (New York: Springer-Verlag, 1986), p. 172.

If you put particles of spin 3/2 and spin 1 in the box, and you know that the first has $m_1 = 1/2$ and the second has $m_2 = 0$ (so m is necessarily 1/2), and you measured the total spin s, you could get 5/2 (with probability 3/5), or 3/2 (with probability 1/15), or 1/2 (with probability 1/3). Again, the sum of the probabilities is 1 (the sum of the squares of each row on the Clebsch-Gordan table is 1).

If you think this is starting to sound like mystical numerology, I don't blame you. We will not be using the Clebsch-Gordan tables much in the rest of the book, but I wanted you to know where they fit into the scheme of things, in case you encounter them later on. In a mathematical sense this is all applied **group theory**—what we are talking about is the decomposition of the direct product of two irreducible representations of the rotation group into a direct sum of irreducible representations. (You can quote that to impress your friends.)

***Problem 4.35**

(a) Apply S_- to $|1\,0\rangle$ (Equation 4.177), and confirm that you get $\sqrt{2}\hbar|1\,-1\rangle$.

(b) Apply $S_\pm$ to $|0\,0\rangle$ (Equation 4.178), and confirm that you get zero.

(c) Show that $|1\,1\rangle$ and $|1\,-1\rangle$ (Equation 4.177) are eigenstates of S^2, with the appropriate eigenvalue.

Problem 4.36 **Quarks** carry spin 1/2. Three quarks bind together to make a **baryon** (such as the proton or neutron); two quarks (or more precisely a quark and an antiquark) bind together to make a **meson** (such as the pion or the kaon). Assume the quarks are in the ground state (so the *orbital* angular momentum is zero).

(a) What spins are possible for baryons?

(b) What spins are possible for mesons?

Problem 4.37

(a) A particle of spin 1 and a particle of spin 2 are at rest in a configuration such that the total spin is 3, and its z-component is 1 (that is, the eigenvalue of S_z is $\hbar$). If you measured the z-component of the angular momentum of the spin-2 particle, what values might you get, and what is the probability of each one?

(b) An electron with spin down is in the state ψ_{510} of the hydrogen atom. If you could measure the total angular momentum squared of the electron alone (not including the proton spin), what values might you get, and what is the probability of each?

Problem 4.38 Determine the commutator of S^2 with $S_z^{(1)}$ (where $\mathbf{S} \equiv \mathbf{S}^{(1)} + \mathbf{S}^{(2)}$). Generalize your result to show that

$$[S^2, \mathbf{S}^{(1)}] = 2i\hbar(\mathbf{S}^{(1)} \times \mathbf{S}^{(2)}). \qquad [4.187]$$

Note: Because $S_z^{(1)}$ does not commute with S^2, we cannot hope to find states that are simultaneous eigenvectors of both. To form eigenstates of S^2, we need linear combinations of eigenstates of $S_z^{(1)}$. This is precisely what the Clebsch-Gordan coefficients (in Equation 4.185) do for us. On the other hand, it follows by obvious inference from Equation 4.187 that the sum $\mathbf{S}^{(1)} + \mathbf{S}^{(2)}$ *does* commute with S^2, which only confirms what we already knew (see Equation 4.103).

FURTHER PROBLEMS FOR CHAPTER 4

***Problem 4.39** Consider the **three-dimensional harmonic oscillator**, for which the potential is

$$V(r) = \frac{1}{2}m\omega^2 r^2. \qquad [4.188]$$

(a) Show that separation of variables in Cartesian coordinates turns this into three one-dimensional oscillators, and exploit your knowledge of the latter to determine the allowed energies. *Answer*:

$$E_n = (n + 3/2)\hbar\omega. \qquad [4.189]$$

(b) Determine the degeneracy $d(n)$ of E_n.

***Problem 4.40** Because the three-dimensional harmonic oscillator potential (Equation 4.188) is spherically symmetric, the Schrödinger equation can be handled by separation of variables in *spherical* coordinates as well as Cartesian coordinates. Use the power series method to solve the radial equation. Find the recursion formula for the coefficients, and determine the allowed energies. Check your answer against Equation 4.189.

Problem 4.41

(a) Prove the **three-dimensional virial theorem**

$$2\langle T\rangle = \langle \mathbf{r}\cdot\nabla V\rangle \qquad [4.190]$$

(for stationary states). *Hint*: Refer to Problem 3.53.

(b) Apply the virial theorem to the case of hydrogen, and show that

$$\langle T\rangle = -E_n; \quad \langle V\rangle = 2E_n. \qquad [4.191]$$

(c) Apply the virial theorem to the three-dimensional harmonic oscillator (Problem 4.39), and show that in this case

$$\langle T\rangle = \langle V\rangle = E_n/2. \qquad [4.192]$$

***Problem 4.42** The **momentum-space wave function** in three dimensions is defined by the natural generalization of Equation 3.132:

$$\phi(\mathbf{p}) \equiv \frac{1}{(2\pi\hbar)^{3/2}} \int e^{-i(\mathbf{p}\cdot\mathbf{r})/\hbar} \psi(\mathbf{r})\, d^3\mathbf{r}. \qquad [4.193]$$

(a) Find the momentum-space wave function for the ground state of hydrogen (Equation 4.80). *Hint*: Use spherical coordinates, setting the polar axis along the direction of $\mathbf{p}$. Do the θ integral first. *Answer*:

$$\phi(\mathbf{p}) = \frac{1}{\pi}\left(\frac{2a}{\hbar}\right)^{3/2} \frac{1}{[1 + (ap/\hbar)^2]^2}. \qquad [4.194]$$

(b) Check that $\phi(\mathbf{p})$ is normalized.

(c) Use $\phi(\mathbf{p})$ to calculate $\langle p^2 \rangle$.

(d) What is the expectation value of the kinetic energy in this state? Express your answer as a multiple of E_1, and check that it is consistent with the virial theorem (Equation 4.191).

Problem 4.43

(a) Construct the spatial wave function (ψ) for hydrogen in the state $n = 3$, $l = 2$, $m = 1$. Express your answer as a function of r, θ, ϕ, and a (the Bohr radius) *only*—no other variables (ρ, z, etc.), or functions (Y, v, etc.), or constants (A, a_0, etc.), or derivatives allowed (π is okay, and e, and 2, etc.).

(b) Check that this wave function is properly normalized by carrying out the appropriate integrals over r, θ, and ϕ.

(c) Find the expectation value of r^s in this state. For what range of s is the result finite?

***Problem 4.44** Suppose two spin-1/2 particles are known to be in the singlet configuration (Equation 4.178). Let $S_a^{(1)}$ be the component of the spin angular momentum of particle number 1 in the direction defined by the unit vector $\hat{a}$. Similarly, let $S_b^{(2)}$ be the component of 2's angular momentum in the direction $\hat{b}$. Show that

$$\langle S_a^{(1)} S_b^{(2)} \rangle = -\frac{\hbar^2}{4} \cos\theta, \qquad [4.195]$$

where θ is the angle between $\hat{a}$ and $\hat{b}$.

***Problem 4.45** Work out the Clebsch-Gordan coefficients for the case $s_1 = 1/2$, $s_2 =$ anything. *Hint*: You're looking for the coefficients A and B in

$$|s\, m\rangle = A|\tfrac{1}{2}\, \tfrac{1}{2}\rangle|s_2\, (m - \tfrac{1}{2})\rangle + B|\tfrac{1}{2}\, (-\tfrac{1}{2})\rangle|s_2\, (m + \tfrac{1}{2})\rangle,$$

such that $|s\,m\rangle$ is an eigenstate of S^2. Use the method of Equations 4.179 through 4.182. If you can't figure out what $S_x^{(2)}$ (for instance) does to $|s_2\,m_2\rangle$, refer back to Equations 4.136 and 4.144. Use this general result to construct the $(1/2) \times 1$ table of Clebsch-Gordan coefficients, and check it against Table 4.7. *Answer*:

$$A = \sqrt{\frac{s_2 \pm m + 1/2}{2s_2 + 1}}; \quad B = \pm\sqrt{\frac{s_2 \mp m + 1/2}{2s_2 + 1}},$$

where the signs are determined by $s = s_2 \pm 1/2$.

Problem 4.46 Find the matrix representing S_x for a particle of spin 3/2 (using the basis of eigenstates of S_z). Solve the characteristic equation to determine the eigenvalues of S_x.

***Problem 4.47** Work out the normalization factor for the spherical harmonics, as follows. From Section 4.1.2 we know that

$$Y_l^m = B_l^m e^{im\phi} P_l^m(\cos\theta);$$

the problem is to determine the factor B_l^m (which I quoted, but did not derive, in Equation 4.32). Use Equations 4.120, 4.121, and 4.130 to obtain a recursion relation giving B_l^{m+1} in terms of B_l^m. Solve it by induction on m to get B_l^m up to an overall constant $C(l)$. Finally, use the result of Problem 4.22 to fix the constant. You may find the following formula for the derivative of an associated Legendre function useful:

$$(1 - x^2)\frac{dP_l^m}{dx} = \sqrt{1 - x^2}\,P_l^{m+1} - mx\,P_l^m. \qquad [4.196]$$

Problem 4.48 The electron in a hydrogen atom occupies the combined spin and position state

$$R_{21}\left(\sqrt{\frac{1}{3}}Y_1^0\chi_+ + \sqrt{\frac{2}{3}}Y_1^1\chi_-\right).$$

(a) If you measured the orbital angular momentum squared (L^2), what values might you get, and what is the probability of each?

(b) Same for the z-component of orbital angular momentum (L_z).

(c) Same for the spin angular momentum squared (S^2).

(d) Same for the z-component of spin angular momentum (S_z).

Let $\mathbf{J} \equiv \mathbf{L} + \mathbf{S}$ be the total angular momentum.

(e) If you measured J^2, what values might you get, and what is the probability of each?

(f) Same for J_z.

(g) If you measured the position of the particle, what is the probability density for finding it at r, θ, ϕ?

(h) If you measured both the z-component of the spin *and* the distance from the origin (note that these are compatible observables), what is the probability density for finding the particle with spin up and at radius r?

***Problem 4.49

(a) For a function $f(\phi)$ that can be expanded in a Taylor series, show that

$$f(\phi + \phi_0) = e^{iL_z\phi_0/\hbar} f(\phi)$$

(where ϕ_0 is any constant angle). For this reason, $L_z/\hbar$ is called the **generator of rotations** about the z-axis. *Hint:* Use Equation 4.129, and refer to Problem 3.55.

More generally, $\mathbf{L} \cdot \hat{n}/\hbar$ is the generator of rotations about the direction $\hat{n}$, in the sense that $\exp(i\mathbf{L} \cdot \hat{n}\varphi/\hbar)$ effects a rotation through angle φ (in the right-hand sense) about the axis $\hat{n}$. In the case of spin, the generator of rotations is $\mathbf{S} \cdot \hat{n}/\hbar$. In particular, for spin 1/2

$$\chi' = e^{i(\sigma \cdot \hat{n})\varphi/2}\chi \qquad [4.197]$$

tells us how spinors rotate.

(b) Construct the (2×2) matrix representing rotation by $180°$ about the x-axis, and show that it converts "spin up" (χ_+) into "spin down" (χ_-), as you would expect.

(c) Construct the matrix representing rotation by $90°$ about the y-axis, and check what it does to χ_+.

(d) Construct the matrix representing rotation by $360°$ about the z-axis. If the answer is not quite what you expected, discuss its implications.

(e) Show that

$$e^{i(\sigma \cdot \hat{n})\varphi/2} = \cos(\varphi/2) + i(\hat{n} \cdot \boldsymbol{\sigma}) \sin(\varphi/2). \qquad [4.198]$$

Problem 4.50 The fundamental commutation relations for angular momentum (Equations 4.98 and 4.99) allow for half-integer (as well as integer) eigenvalues. But for *orbital* angular momentum only the integer values occur. There must be some *extra* constraint in the specific form $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ that excludes half-integer values.[40] Let a be some convenient constant with the dimensions of length (the Bohr radius, say, if we're talking about hydrogen), and define the operators

$$q_1 \equiv \frac{1}{\sqrt{2}}\left[x + (a^2/\hbar)p_y\right]; \quad p_1 \equiv \frac{1}{\sqrt{2}}\left[p_x - (\hbar/a^2)y\right];$$
$$q_2 \equiv \frac{1}{\sqrt{2}}\left[x - (a^2/\hbar)p_y\right]; \quad p_2 \equiv \frac{1}{\sqrt{2}}\left[p_x + (\hbar/a^2)y\right].$$

[40] This problem is based on an argument in Ballentine, (footnote 34), page 127.

(a) Verify that $[q_1, q_2] = [p_1, p_2] = 0$; $[q_1, p_1] = [q_2, p_2] = i\hbar$. Thus the q's and the p's satisfy the canonical commutation relations for position and momentum. and those of index 1 are compatible with those of index 2.

(b) Show that

$$L_z = \frac{\hbar}{2a^2}(q_1^2 - q_2^2) + \frac{a^2}{2\hbar}(p_1^2 - p_2^2).$$

(c) Check that $L_z = H_1 - H_2$, where each H is the Hamiltonian for a harmonic oscillator with mass $m = \hbar/a^2$ and frequency $\omega = 1$.

(d) We know that the eigenvalues of the harmonic oscillator Hamiltonian are $(n + 1/2)\hbar\omega$, where $n = 0, 1, 2, \ldots$ (in the algebraic theory of Section 2.3.1, this follows from the form of the Hamiltonian and the canonical commutation relations). Use this to conclude that the eigenvalues of L_z must be integers.

***Problem 4.51** In classical electrodynamics the force on a particle of charge q moving with velocity $\mathbf{v}$ through electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ is given by the **Lorentz force law**:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \qquad [4.199]$$

This force cannot be expressed as the gradient of a scalar potential energy function. and therefore the Schrödinger equation in its original form (Equation 1.1) cannot accomodate it. But in the more sophisticated form

$$i\hbar\frac{\partial\Psi}{\partial t} = H\Psi \qquad [4.200]$$

there is no problem; the classical Hamiltonian[41] is

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + q\varphi, \qquad [4.201]$$

where $\mathbf{A}$ is the vector potential ($\mathbf{B} = \nabla \times \mathbf{A}$) and φ is the scalar potential ($\mathbf{E} = -\nabla\varphi - \partial\mathbf{A}/\partial t$), so the Schrödinger equation (making the canonical substitution $\mathbf{p} \to (i/\hbar)\nabla$) becomes

$$i\hbar\frac{\partial\Psi}{\partial t} = \left[\frac{1}{2m}\left(\frac{\hbar}{i}\nabla - q\mathbf{A}\right)^2 + q\varphi\right]\Psi. \qquad [4.202]$$

(a) Show that

$$\frac{d\langle\mathbf{r}\rangle}{dt} = \frac{1}{m}\langle(\mathbf{p} - q\mathbf{A})\rangle. \qquad [4.203]$$

---

[41] See, for example, H. Goldstein, *Classical Mechanics*, 2nd ed., Addison-Wesley, Reading, MA. 1980, page 346.

(b) As always (see Equation 1.32) we identify $d\langle \mathbf{r} \rangle / dt$ with $\langle \mathbf{v} \rangle$. Show that

$$m\frac{d\langle \mathbf{v} \rangle}{dt} = q\langle \mathbf{E} \rangle + \frac{q}{2m}\langle (\mathbf{p} \times \mathbf{B} - \mathbf{B} \times \mathbf{p}) \rangle - \frac{q^2}{m}\langle (\mathbf{A} \times \mathbf{B}) \rangle. \quad [4.204]$$

(c) In particular, if the fields $\mathbf{E}$ and $\mathbf{B}$ are *uniform* over the volume of the wave packet, show that

$$m\frac{d\langle \mathbf{v} \rangle}{dt} = q(\mathbf{E} + \langle \mathbf{v} \rangle \times \mathbf{B}), \quad [4.205]$$

so the *expectation value* of $\langle \mathbf{v} \rangle$ moves according to the Lorentz force law, as we would expect from Ehrenfest's theorem.

***Problem 4.52** [Refer to Problem 4.51 for background.] Suppose

$$\mathbf{A} = \frac{B_0}{2}(x\hat{\jmath} - y\hat{\imath}), \quad \text{and} \quad \varphi = Kz^2,$$

where B_0 and K are constants.

(a) Find the fields $\mathbf{E}$ and $\mathbf{B}$.

(b) Find the allowed energies, for a particle of mass m and charge q, in these fields.

Problem 4.53 [Refer to Problem 4.51 for background.] In classical electrodynamics the potentials $\mathbf{A}$ and φ are not uniquely determined[42]; the *physical* quantities are the *fields*, $\mathbf{E}$ and $\mathbf{B}$.

(a) Show that the potentials

$$\varphi' \equiv \varphi - \frac{\partial \Lambda}{\partial t}, \quad \mathbf{A}' \equiv \mathbf{A} + \nabla\Lambda \quad [4.206]$$

(where Λ is an arbitrary real function of position and time) yield the same fields as φ and $\mathbf{A}$. Equation [4.206] is called a **gauge transformation**, and the theory is said to be **gauge invariant**.

(b) In quantum mechanics the potentials play a more direct role, and it is of interest to know whether the theory remains gauge invariant. Show that

$$\Psi' \equiv e^{iq\Lambda/\hbar}\Psi \quad [4.207]$$

satisfies the Schrödinger equation [4.202] with the gauge-transformed potentials φ' and $\mathbf{A}'$. Since Ψ' differs from Ψ only by a *phase factor*, it represents the

---

[42] See, for example, Griffiths, (footnote 27), section 10.2.4.

same physical state[43], and the theory *is* gauge invariant (see Section 10.2.4 for further discussion).

---

[43]That is to say, $\langle \mathbf{r} \rangle$, $d\langle \mathbf{r} \rangle / dt$, etc. are unchanged. Because Λ depends on position, $\langle \mathbf{p} \rangle$ (with $\mathbf{p}$ represented by the operator $(\hbar/i)\nabla$) *does* change, but as we found in Equation [4.203], $\mathbf{p}$ does not represent the mechanical momentum $(m\mathbf{v})$ in this context (in Lagrangian mechanics it is so-called **canonical momentum**).

CHAPTER 5

IDENTICAL PARTICLES

5.1 TWO-PARTICLE SYSTEMS

For a *single* particle, the wave function $\Psi(\mathbf{r}, t)$ is a function of the spatial coordinates $\mathbf{r}$ and the time t (we'll ignore spin for the moment). The wave function for a *two*-particle system is a function of the coordinates of particle one ($\mathbf{r}_1$), the coordinates of particle two ($\mathbf{r}_2$), and the time:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t). \tag{5.1}$$

Its time evolution is determined (as always) by the Schrödinger equation:

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi, \tag{5.2}$$

where H is the Hamiltonian for the whole system:

$$H = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(\mathbf{r}_1, \mathbf{r}_2, t) \tag{5.3}$$

(the subscript on ∇ indicates differentiation with respect to the coordinates of particle 1 or particle 2, as the case may be). The statistical interpretation carries over in the obvious way:

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 \, d^3\mathbf{r}_1 \, d^3\mathbf{r}_2 \tag{5.4}$$

is the probability of finding particle 1 in the volume $d^3\mathbf{r}_1$ and particle 2 in the volume $d^3\mathbf{r}_2$; evidently Ψ must be normalized in such a way that

$$\int |\Psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 \, d^3\mathbf{r}_1 \, d^3\mathbf{r}_2 = 1. \tag{5.5}$$

For time-independent potentials, we obtain a complete set of solutions by separation of variables:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t) = \psi(\mathbf{r}_1, \mathbf{r}_2)e^{-iEt/\hbar}, \tag{5.6}$$

where the spatial wave function (ψ) satisfies the time-independent Schrödinger equation:

$$-\frac{\hbar^2}{2m_1}\nabla_1^2\psi - \frac{\hbar^2}{2m_2}\nabla_2^2\psi + V\psi = E\psi, \tag{5.7}$$

and E is the total energy of the system.

**Problem 5.1 Typically, the interaction potential depends only on the vector $\mathbf{r} \equiv \mathbf{r}_1 - \mathbf{r}_2$ separating the two particles. In that case the Schrödinger equation separates, if we change variables from $\mathbf{r}_1$, $\mathbf{r}_2$ to $\mathbf{r}$, $\mathbf{R} \equiv (m_1\mathbf{r}_1 + m_2\mathbf{r}_2)/(m_1 + m_2)$ (the center of mass).

(a) Show that $\mathbf{r}_1 = \mathbf{R} + (\mu/m_1)\mathbf{r}$, $\mathbf{r}_2 = \mathbf{R} - (\mu/m_2)\mathbf{r}$, and $\nabla_1 = (\mu/m_2)\nabla_R + \nabla_r$, $\nabla_2 = (\mu/m_1)\nabla_R - \nabla_r$, where

$$\mu \equiv \frac{m_1 m_2}{m_1 + m_2} \tag{5.8}$$

is the **reduced mass** of the system.

(b) Show that the (time-independent) Schrödinger equation becomes

$$-\frac{\hbar^2}{2(m_1 + m_2)}\nabla_R^2\psi - \frac{\hbar^2}{2\mu}\nabla_r^2\psi + V(\mathbf{r})\psi = E\psi.$$

(c) Solve by separation of variables, letting $\psi(\mathbf{R}, \mathbf{r}) = \psi_R(\mathbf{R})\psi_r(\mathbf{r})$. Note that ψ_R satisfies the one-particle Schrödinger equation, with the *total* mass $(m_1 + m_2)$ in place of m, potential zero, and energy E_R, while ψ_r satisfies the one-particle Schrödinger equation with the *reduced* mass in place of m, potential $V(\mathbf{r})$, and energy E_r. The total energy is the sum: $E = E_R + E_r$. *Note*: What this tells us is that the center of mass moves like a free particle, and the *relative* motion (that is, the motion of particle 2 with respect to particle 1) is the same as if we had a *single* particle with the *reduced* mass, subject to the potential V. Exactly the same separation occurs in *classical* mechanics[1]; it reduces the two-body problem to an equivalent one-body problem.

Problem 5.2 In view of Problem 5.1, we can correct for the motion of the nucleus in hydrogen by simply replacing the electron mass with the reduced mass:

(a) Find (to two significant digits) the percent error in the binding energy of hydrogen (Equation 4.77) introduced by our use of m instead of μ.

---

[1] See, for example, Jerry Marion, *Classical Dynamics*, 2nd ed. (New York: Academic Press 1970), Section 8.2.

(b) Find the separation in wavelength between the red Balmer lines ($n = 3 \rightarrow n = 2$) for hydrogen and deuterium.

(c) Find the binding energy of **positronium** (in which the proton is replaced by a positron—positrons have the same mass as electrons but opposite charge).

(d) Suppose you wanted to confirm the existence of **muonic hydrogen**, in which the electron is replaced by a muon (same charge, but 206.77 times heavier). Where (i.e., at what wavelength) would you look for the "Lyman-α" line ($n = 2 \rightarrow n = 1$)?

5.1.1 Bosons and Fermions

Suppose particle 1 is in the (one-particle) state $\psi_a(\mathbf{r})$, and particle 2 is in the state $\psi_b(\mathbf{r})$. In that case $\psi(\mathbf{r}_1, \mathbf{r}_2)$ is a simple *product*:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2). \qquad [5.9]$$

Of course, this assumes that we can tell the particles apart—otherwise it wouldn't make any sense to claim that number 1 is in state ψ_a and number 2 is in state ψ_b; all we could say is that *one* of them is in the state ψ_a and the *other* is in state ψ_b, but we wouldn't know which is which. If we were talking about *classical* mechanics this would be a silly objection: You can *always* tell the particles apart, in principle—just paint one of them red and the other one blue, or stamp identification numbers on them, or hire private detectives to follow them around. But in quantum mechanics the situation is fundamentally different: You can't paint an electron red, or pin a label on it, and a detective's observations will inevitably and unpredictably alter the state, raising doubts as to whether the two had perhaps switched places. The fact is, all electrons are *utterly identical*, in a way that no two classical objects can ever be. It is not merely that *we* don't happen to know which electron is which; *God* doesn't know which is which, because there is *no such thing* as "this" electron, or "that" electron; all we can legitimately speak about is "an" electron.

Quantum mechanics neatly accommodates the existence of particles that are *indistinguishable in principle*: We simply construct a wave function that is *noncommittal* as to which particle is in which state. There are actually *two* ways to do it:

$$\psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2) = A[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \pm \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)]. \qquad [5.10]$$

Thus the theory admits two kinds of identical particles: **bosons**, for which we use the plus sign, and **fermions**, for which we use the minus sign. Photons and mesons are bosons; protons and electrons are fermions. It so happens that

$$\left\{ \begin{array}{l} \text{all particles with } \textit{integer} \text{ spin are bosons, and} \\ \text{all particles with } \textit{half-integer} \text{ spin are fermions.} \end{array} \right. \qquad [5.11]$$

This connection between spin and "statistics" (as we shall see, bosons and fermions have quite different statistical properties) can be *proved* in relativistic quantum mechanics; in the nonrelativistic theory it must be taken as an axiom.

It follows, in particular, that *two identical fermions* (for example, two electrons) *cannot occupy the same state.* For if $\psi_a = \psi_b$, then

$$\psi_-(\mathbf{r}_1, \mathbf{r}_2) = A[\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) - \psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] = 0,$$

and we are left with no wave function at all. This is the famous **Pauli exclusion principle**. It is not (as you may have been led to believe) a bizarre ad hoc assumption applying only to electrons, but rather a consequence of the rules for constructing two-particle wave functions, applying to *all* identical fermions.

I assumed, for the sake of argument, that one particle was in the state ψ_a and the other in state ψ_b, but there is a more general (and more sophisticated) way to formulate the problem. Let us define the **exchange operator** P which interchanges the two particles:

$$Pf(\mathbf{r}_1, \mathbf{r}_2) = f(\mathbf{r}_2, \mathbf{r}_1). \qquad [5.12]$$

Clearly, $P^2 = 1$, and it follows (prove it for yourself) that the eigenvalues of P are ± 1. If the two particles are identical, the Hamiltonian must treat them the same: $m_1 = m_2$ and $V(\mathbf{r}_1, \mathbf{r}_2) = V(\mathbf{r}_2, \mathbf{r}_1)$. It follows that P and H are compatible observables,

$$[P, H] = 0, \qquad [5.13]$$

and hence we can find a complete set of functions that are simultaneous eigenstates of both. That is to say, we can find solutions to the Schrödinger equation that are either symmetric (eigenvalue $+1$) or antisymmetric (eigenvalue -1) under exchange:

$$\boxed{\psi(\mathbf{r}_1, \mathbf{r}_2) = \pm\psi(\mathbf{r}_2, \mathbf{r}_1) \quad (+ \text{ for bosons}, - \text{ for fermions}).} \qquad [5.14]$$

Moreover, if a system starts out in such a state, it will remain in such a state. The *new* law (I'll call it the **symmetrization requirement**) is that for identical particles the wave function is not merely *allowed*, but *required* to satisfy Equation 5.14, with the plus sign for bosons and the minus sign for fermions.[2] This is the *general* statement. of which Equation 5.10 is a special case.

---

[2]It is sometimes suggested that the symmetrization requirement (Equation 5.14) is nothing new—that it is *forced* by the fact that P and H commute. This is false: It is perfectly possible to imagine a system of two *distinguishable* particles (say, an electron and a positron) for which the Hamiltonian is symmetric. and yet there is no requirement that the wave function be symmetric (or antisymmetric). But *identical* particles *have* to occupy symmetric or antisymmetric states, and this is a completely *new fundamental law*—on a par, logically, with Schrödinger's equation and the statistical interpretation. Of course, there didn't *have* to be any such things as identical particles; it could have been that every single particle in nature was clearly distinguishable from every other one. Quantum mechanics allows for the *possibility* of identical particles, and nature (being lazy) seized the opportunity. (But I'm not complaining—this makes matters enormously simpler!)

Example. Suppose we have two noninteracting[3] particles, both of mass m, in the infinite square well (Section 2.2). The one-particle states are

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right), \quad E_n = n^2 K$$

(where $K \equiv \pi^2\hbar^2/2ma^2$). If the particles are *distinguishable*, the composite wave functions are simple products:

$$\psi_{n_1 n_2}(x_1, x_2) = \psi_{n_1}(x_1)\psi_{n_2}(x_2), \quad E_{n_1 n_2} = (n_1^2 + n_2^2)K.$$

For example, the ground state is

$$\psi_{11} = \frac{2}{a}\sin(\pi x_1/a)\sin(\pi x_2/a), \quad E_{11} = 2K;$$

the first excited state is doubly degenerate:

$$\psi_{12} = \frac{2}{a}\sin(\pi x_1/a)\sin(2\pi x_2/a), \quad E_{12} = 5K,$$

$$\psi_{21} = \frac{2}{a}\sin(2\pi x_1/a)\sin(\pi x_2/a), \quad E_{21} = 5K;$$

and so on. If the two particles are identical *bosons*, the ground state is unchanged, but the first excited state is *nondegenerate*:

$$\frac{\sqrt{2}}{a}[\sin(\pi x_1/a)\sin(2\pi x_2/a) + \sin(2\pi x_1/a)\sin(\pi x_2/a)]$$

(still with energy $5K$). And if the particles are identical *fermions*, there is *no* state with energy $2K$; the ground state is

$$\frac{\sqrt{2}}{a}[\sin(\pi x_1/a)\sin(2\pi x_2/a) - \sin(2\pi x_1/a)\sin(\pi x_2/a)],$$

and its energy is $5K$.

*Problem 5.3

(a) If ψ_a and ψ_b are orthogonal, and both normalized, what is the constant A in Equation 5.10?

(b) If $\psi_a = \psi_b$ (and it is normalized), what is A? (This case, of course, occurs only for bosons.)

[3] They pass right through one another—never mind how you would set this up in practice! I'll ignore spin—if this bothers you (after all, a spinless fermion is a contradiction in terms), assume they're in the same spin state.

Problem 5.4

(a) Write down the Hamiltonian for two identical noninteracting particles in the infinite square well. Verify that the fermion ground state given in the example is an eigenfunction of H, with the appropriate eigenvalue.

(b) Find the next two excited states (beyond the ones given in the example)—wave functions and energies—for each of the three cases (distinguishable, identical bosons, identical fermions).

5.1.2 Exchange Forces

To give you some sense of what the symmetrization requirement actually *does*, I'm going to work out a simple one-dimensional example. Suppose one particle is in state $\psi_a(x)$, and the other is in state $\psi_b(x)$, and these two states are orthogonal and normalized. If the two particles are distinguishable, and number 1 is the one in state ψ_a, then the combined wave function is

$$\psi(x_1, x_2) = \psi_a(x_1)\psi_b(x_2); \tag{5.15}$$

if they are identical bosons, the composite wave function is (see Problem 5.3 for the normalization)

$$\psi_+(x_1, x_2) = \frac{1}{\sqrt{2}}[\psi_a(x_1)\psi_b(x_2) + \psi_b(x_1)\psi_a(x_2)]; \tag{5.16}$$

and if they are identical fermions, it is

$$\psi_-(x_1, x_2) = \frac{1}{\sqrt{2}}[\psi_a(x_1)\psi_b(x_2) - \psi_b(x_1)\psi_a(x_2)]. \tag{5.17}$$

Let's calculate the expectation value of the square of the separation distance between the two particles,

$$\langle (x_1 - x_2)^2 \rangle = \langle x_1^2 \rangle + \langle x_2^2 \rangle - 2\langle x_1 x_2 \rangle. \tag{5.18}$$

Case 1: Distinguishable particles. For the wave function in Equation 5.15. we have

$$\langle x_1^2 \rangle = \int x_1^2 |\psi_a(x_1)|^2 \, dx_1 \int |\psi_b(x_2)|^2 \, dx_2 = \langle x^2 \rangle_a$$

(the expectation value of x^2 in the one-particle state ψ_a),

$$\langle x_2^2 \rangle = \int |\psi_a(x_1)|^2 \, dx_1 \int x_2^2 |\psi_b(x_2)|^2 \, dx_2 = \langle x^2 \rangle_b,$$

and

$$\langle x_1 x_2 \rangle = \int x_1 |\psi(x_1)|^2 \, dx_1 \int x_2 |\psi(x_2)|^2 \, dx_2 = \langle x \rangle_a \langle x \rangle_b.$$

In this case, then,

$$\langle (x_1 - x_2)^2 \rangle_d = \langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b. \qquad [5.19]$$

(Incidentally, the answer would—of course—be the same if particle 1 had been in state ψ_b, and particle 2 in state ψ_a.)

Case 2: Identical particles. For the wave functions in Equations 5.16 and 5.17,

$$\begin{aligned}
\langle x_1^2 \rangle &= \frac{1}{2}\Big[\int x_1^2 |\psi_a(x_1)|^2 \, dx_1 \int |\psi_b(x_2)|^2 \, dx_2 \\
&\quad + \int x_1^2 |\psi_b(x_1)|^2 \, dx_1 \int |\psi_a(x_2)|^2 \, dx_2 \\
&\quad \pm \int x_1^2 \psi_a(x_1)^* \psi_b(x_1) \, dx_1 \int \psi_b(x_2)^* \psi_a(x_2) \, dx_2 \\
&\quad \pm \int x_1^2 \psi_b(x_1)^* \psi_a(x_1) \, dx_1 \int \psi_a(x_2)^* \psi_b(x_2) \, dx_2 \Big] \\
&= \frac{1}{2}[\langle x^2 \rangle_a + \langle x^2 \rangle_b \pm 0 \pm 0] = \frac{1}{2}\left(\langle x^2 \rangle_a + \langle x^2 \rangle_b \right).
\end{aligned}$$

Similarly,

$$\langle x_2^2 \rangle = \frac{1}{2}\left(\langle x^2 \rangle_b + \langle x^2 \rangle_a \right).$$

(Naturally, $\langle x_2^2 \rangle = \langle x_1^2 \rangle$, since you can't tell them apart.) But

$$\begin{aligned}
\langle x_1 x_2 \rangle &= \frac{1}{2}\Big[\int x_1 |\psi_a(x_1)|^2 \, dx_1 \int x_2 |\psi_b(x_2)|^2 \, dx_2 \\
&\quad + \int x_1 |\psi_b(x_1)|^2 \, dx_1 \int x_2 |\psi_a(x_2)|^2 \, dx_2 \\
&\quad \pm \int x_1 \psi_a(x_1)^* \psi_b(x_1) \, dx_1 \int x_2 \psi_b(x_2)^* \psi_a(x_2) \, dx_2 \\
&\quad \pm \int x_1 \psi_b(x_1)^* \psi_a(x_1) \, dx_1 \int x_2 \psi_a(x_2)^* \psi_b(x_2) \, dx_2 \Big]
\end{aligned}$$

$$= \frac{1}{2}\left(\langle x\rangle_a\langle x\rangle_b + \langle x\rangle_b\langle x\rangle_a \pm \langle x\rangle_{ab}\langle x\rangle_{ba} \pm \langle x\rangle_{ba}\langle x\rangle_{ab}\right)$$

$$= \langle x\rangle_a\langle x\rangle_b \pm |\langle x\rangle_{ab}|^2,$$

where

$$\langle x\rangle_{ab} \equiv \int x\psi_a(x)^*\psi_b(x)\,dx. \qquad [5.20]$$

Evidently

$$\langle (x_1 - x_2)^2\rangle_\pm = \langle x^2\rangle_a + \langle x^2\rangle_b - 2\langle x\rangle_a\langle x\rangle_b \mp 2|\langle x\rangle_{ab}|^2. \qquad [5.21]$$

Comparing Equations 5.19 and 5.21, we see that the difference resides in the final term:

$$\langle (\Delta x)^2\rangle_\pm = \langle (\Delta x)^2\rangle_d \mp 2|\langle x\rangle_{ab}|^2; \qquad [5.22]$$

identical bosons (the upper signs) tend to be somewhat closer together, and identical fermions (the lower signs) somewhat farther apart, than distinguishable particles in the same two states. Notice that $\langle x\rangle_{ab}$ *vanishes* unless the two wave functions actually *overlap* [if $\psi_a(x)$ is zero wherever $\psi_b(x)$ is *non*zero, the integral in Equation 5.20 is itself zero]. So if ψ_a represents an electron in an atom in Chicago and ψ_b represents an electron in an atom in Seattle, it's not going to make any difference whether you antisymmetrize the wave function or not. As a *practical* matter, therefore, it's okay to pretend that electrons with nonoverlapping wave functions are distinguishable. (Indeed, this is the only thing that allows physicists and chemists to proceed at *all*, for in *principle* every electron in the universe is linked to every other one via the antisymmetrization of their wave functions, and if this really *mattered*, you wouldn't be able to talk about any *one* electron until you were prepared to deal with them *all*!)

The *interesting* case is when there *is* some overlap of the wave functions. The system behaves as though there were a "force of attraction" between identical bosons, pulling them closer together, and a "force of repulsion" between identical fermions, pushing them apart. We call it an **exchange force**, although it's not really a force at all—no physical agency is pushing on the particles; rather, it is a purely *geometrical* consequence of the symmetrization requirement. It is also a strictly quantum mechanical phenomenon, with no classical counterpart. Nevertheless, it has profound consequences. Consider, for example, the hydrogen molecule (H_2). Roughly speaking, the ground state consists of one electron in the atomic ground state (Equation 4.80) centered on nucleus 1, and one electron in the atomic ground state centered at nucleus 2. If electrons were *bosons*, the symmetrization requirement (or, if you like, the "exchange force") would tend to concentrate the electrons toward the middle, between the two protons (Figure 5.1a), and the resulting accumulation of negative charge would attract the protons inward, accounting for the **covalent bond** that holds

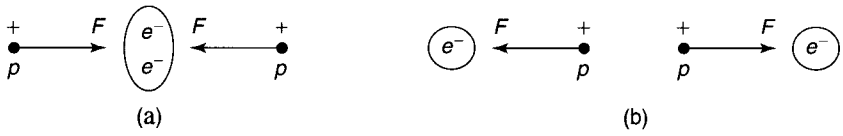


Figure 5.1: Schematic picture of the covalent bond: (a) Symmetric configuration produces attractive force; (b) antisymmetric configuration produces repulsive force.

the molecule together. Unfortunately, electrons *aren't* bosons, they're fermions, and this means that the concentration of negative charge should actually be shifted to the wings (Figure 5.1b), tearing the molecule apart!

But wait. We have been ignoring *spin*. The *complete* state of the electron includes not only its position wave function, but also a spinor, describing the orientation of its spin[4]:

$$\psi(\mathbf{r})\chi(\mathbf{s}). \qquad [5.23]$$

When we put together the two-electron state, it is the *whole works*, not just the spatial part, that has to be antisymmetric with respect to exchange. Now, a glance back at the composite spin states (Equations 4.177 and 4.178) reveals that the singlet combination is antisymmetric (and hence would have to be joined with a *symmetric* spatial function), whereas the three triplet states are all symmetric (and would require an *antisymmetric* spatial function). Evidently, then, the singlet state should lead to *bonding*, and the triplet to *anti*bonding. Sure enough, the chemists tell us that covalent bonding requires the two electrons to occupy the singlet state, with total spin zero.[5]

***Problem 5.5** Imagine two noninteracting particles, each of mass m, in the infinite square well. If one is in the state ψ_n (Equation 2.24) and the other in state ψ_m orthogonal to ψ_n, calculate $\langle(x_1 - x_2)^2\rangle$, assuming that (a) they are distinguishable particles, (b) they are identical bosons, and (c) they are identical fermions.

Problem 5.6 Suppose you had *three* particles, one in state $\psi_a(x)$, one in state $\psi_b(x)$, and one in state $\psi_c(x)$. Assuming that ψ_a, ψ_b, and ψ_c are orthonormal, construct the three-particle states (analogous to Equations 5.15, 5.16, and 5.17) representing (a) distinguishable particles, (b) identical bosons, and (c) identical fermions. Keep in mind that (b) must be completely symmetric under interchange of *any pair* of particles, and (c) must be completely anti-symmetric in the same sense.) *Note*: There's a cute

---

[4] In the absence of coupling between spin and position, we are free to assume that the state is *separable* in its spin and spatial coordinates. This just says that the probability of getting spin up is independent of the *location* of the particle. In the *presence* of coupling, the general state would take the form of a linear combination: $\psi_+(\mathbf{r})\chi_+ + \psi_-(\mathbf{r})\chi_-$.

[5] In casual language, it is often said that the electrons are "oppositely aligned" (one with spin up, and the other with spin down). This is something of an oversimplification, since the same could be said of the $m = 0$ triplet state. The precise statement is that they are in the singlet configuration.

trick for constructing completely antisymmetric wave functions: Form the **Slater determinant**, whose first row is $\psi_a(x_1)$, $\psi_b(x_1)$, $\psi_c(x_1)$, etc., whose second row is $\psi_a(x_2)$, $\psi_b(x_2)$, $\psi_c(x_2)$, etc., and so on (this device works for any number of particles).

5.2 ATOMS

A neutral atom, of atomic number Z, consists of a heavy nucleus, with electric charge Ze, surrounded by Z electrons (mass m and charge $-e$). The Hamiltonian for this system is[6]

$$H = \sum_{j=1}^{Z} \left\{ -\frac{\hbar^2}{2m}\nabla_j^2 - \left(\frac{1}{4\pi\epsilon_0}\right)\frac{Ze^2}{r_j} \right\} + \frac{1}{2}\left(\frac{1}{4\pi\epsilon_0}\right)\sum_{j\neq k}^{Z} \frac{e^2}{|\mathbf{r}_j - \mathbf{r}_k|}. \quad [5.24]$$

The term in curly brackets represents the kinetic plus potential energy of the jth electron in the electric field of the nucleus; the second sum (which runs over all values of j and k except $j = k$) is the potential energy associated with the mutual repulsion of the electrons (the factor of 1/2 in front corrects for the fact that the summation counts each pair twice). The problem is to solve Schrödinger's equation.

$$H\psi = E\psi, \quad [5.25]$$

for the wave function $\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_Z)$. Because electrons are identical fermions, however, not all solutions are acceptable: only those for which the complete state (position and spin),

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_Z)\chi(\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_Z), \quad [5.26]$$

is antisymmetric with respect to interchange of any two electrons. In particular, no two electrons can occupy the *same* state.

Unfortunately, the Schrödinger equation with the Hamiltonian in Equation 5.24 cannot be solved exactly (at any rate, it *hasn't* been) except for the very simplest case, $Z = 1$ (hydrogen). In practice, one must resort to elaborate approximation methods. Some of these we shall explore in Part II; for now I plan only to sketch some of the qualitative features of the solutions, obtained by neglecting the electron repulsion term altogether. In section 5.2.1 we'll study the ground state and excited states of helium, and in section 5.2.2 we'll examine the ground states of higher atoms.

[6] I'm assuming the nucleus is *stationary*. The trick of accounting for nuclear motion by using the reduced mass (Problem 5.1) works only for the *two*-body problem—hydrogen; fortunately, the nucleus is so much more massive than the electrons that the correction is extremely small even in that case (see Problem 5.2a), and it is smaller still for the heavier atoms. There are more interesting effects, due to magnetic interactions associated with electron spin, relativistic corrections, and the finite size of the nucleus. We'll look into these in later chapters, but all of them are minute corrections to the "purely Coulombic" atom described by Equation 5.24.