

position: $p = mdx/dt$. Once you know how the position and momentum operators behave under a translation, you know how *any* operator does, since

$$\hat{Q}'(\hat{x}, \hat{p}) = \hat{T}^\dagger \hat{Q}(\hat{x}, \hat{p}) \hat{T} = \hat{Q}(\hat{x}', \hat{p}') = \hat{Q}(\hat{x} + a, \hat{p}). \quad (6.8)$$

Problem [6.4](#) will walk you through the proof.

Problem 6.3 Show that the operator \hat{p}' obtained by applying a translation to the operator \hat{p} is $\hat{p}' = \hat{T}^\dagger \hat{p} \hat{T} = \hat{p}$.

Problem 6.4 Prove Equation [6.8](#). You may assume that $\hat{Q}(\hat{x}, \hat{p})$ can be written in a power series

$$\hat{Q}(\hat{x}, \hat{p}) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{mn} \hat{x}^m \hat{p}^n$$

for some constants a_{mn} .

6.2.2 Translational Symmetry

So far we have seen how a *function* behaves under a translation and how an *operator* behaves under a translation. I am now in a position to make precise the notion of a symmetry that I mentioned in the introduction. A system is **translationally invariant** (equivalent to saying it has translational symmetry) if the Hamiltonian is unchanged by the transformation:

$$\hat{H}' = \hat{T}^\dagger \hat{H} \hat{T} = \hat{H}.$$

Because \hat{T} is unitary (Equation 6.4) we can multiply both sides of this equation by \hat{T} to get

$$\hat{H} \hat{T} = \hat{T} \hat{H}.$$

Therefore, a system has translational symmetry if the Hamiltonian commutes with the translation operator:

$$[\hat{H}, \hat{T}] = 0. \quad (6.9)$$

For a particle of mass m moving in a one-dimensional potential, the Hamiltonian is

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

According to Equation 6.8, the transformed Hamiltonian is

$$\hat{H}' = \frac{\hat{p}^2}{2m} + V(x+a)$$

so translational symmetry implies that

$$V(x+a) = V(x). \quad (6.10)$$

Now, there are two very different physical settings where Equation 6.10 might arise. The first is a constant potential, where Equation 6.10 holds for *every* value of a ; such a system is said to have **continuous translational symmetry**. The second is a periodic potential, such as an electron might encounter in a crystal, where Equation 6.10 holds only for a discrete set of a s; such a system is said to have **discrete translational symmetry**. The two cases are illustrated in Figure 6.6.

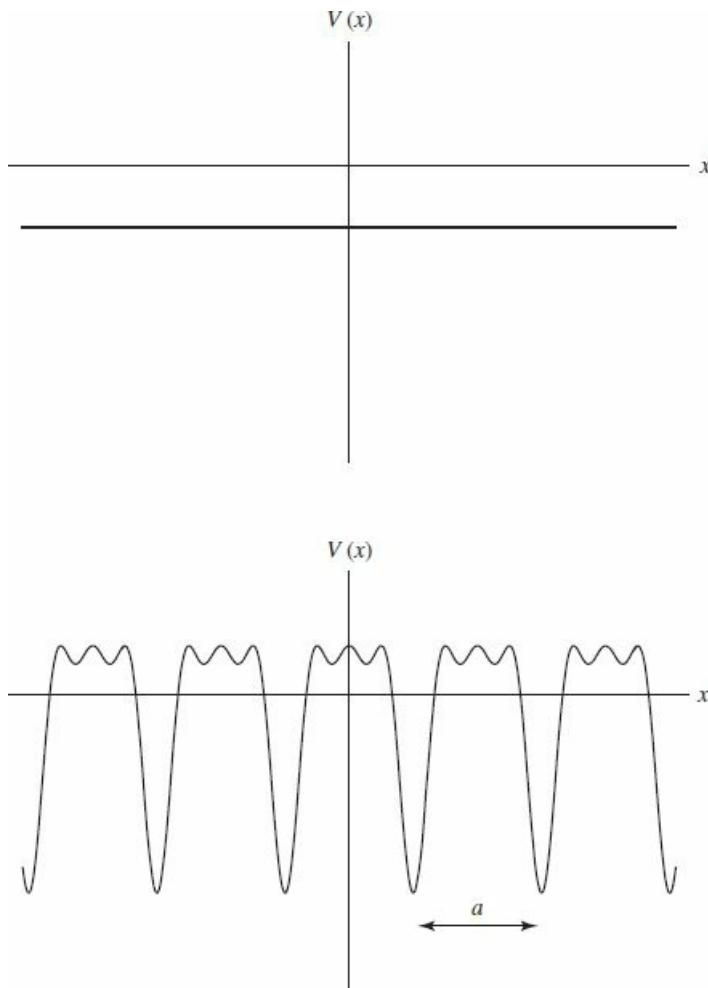


Figure 6.6: Potentials for a system with continuous (top) and discrete (bottom) translational symmetry. In the former case the potential is the same when shifted right or left by *any* amount; in the latter case the potential is the same when shifted right or left by an integer multiple of a .

Discrete Translational Symmetry and Bloch's Theorem

What are the implications of translational symmetry? For a system with a discrete translational symmetry, the most important consequence is Bloch's theorem; the theorem specifies the form taken by the stationary states. We used this theorem in Section 5.3.2; I will now prove it.

In Section A.5 it is shown that if two operators commute, then they have a complete set of simultaneous eigenstates. This means that if the Hamiltonian is translationally invariant (which is to say, if it commutes with the translation operator), then the eigenstates $\psi(x)$ of the Hamiltonian can be chosen to be simultaneously eigenstates of \hat{T} :

$$\hat{H}\psi(x) = E\psi(x), \quad \hat{T}(a)\psi(x) = \lambda\psi(x),$$

where λ is the eigenvalue associated with \hat{T} . Since \hat{T} is unitary, its eigenvalues have magnitude 1 (see Problem A.30), which means that λ can be written as $\lambda = e^{i\phi}$ for some real number ϕ . By convention we write $\phi = -qa$ where $\hbar q$ is called the **crystal momentum**. Therefore, the stationary states of a particle of mass m moving in a periodic potential have the property

$$\psi(x-a) = e^{-iqa}\psi(x). \tag{6.11}$$

There is a more illuminating way to write Equation 6.11:⁷

$$\psi(x) = e^{iqx}u(x)$$

(6.12)

where $u(x)$ is a periodic function of x : $u(x+a) = u(x)$ and e^{iqx} is a traveling wave (recall that a traveling wave by itself describes a *free* particle—Section 2.4) with wavelength $2\pi/q$. Equation 6.12 is **Bloch's theorem** and it says that the stationary states of a particle in a periodic potential are periodic functions multiplying traveling waves. Note that just because the Hamiltonian is translationally invariant, that doesn't mean the stationary states *themselves* are translationally invariant, it simply means that they can be chosen to be eigenstates of the translation operator.

Bloch's theorem is truly remarkable. It tells us that the stationary states of a particle in a periodic potential (such as an electron in a crystal) are, apart from a periodic modulation, traveling waves. As such, they have a nonzero velocity.⁸ This means that an electron could travel through a perfect crystal without scattering! That has dramatic implications for electronic conduction in solids.

Continuous Translational Symmetry and Momentum Conservation

If a system has continuous translation symmetry then the Hamiltonian commutes with $\hat{T}(a)$ for any choice of a . In this case it is useful to consider an **infinitesimal translation**

$$\hat{T}(\delta) = e^{-i\delta \hat{p}/\hbar} \approx 1 - i\frac{\delta}{\hbar} \hat{p},$$

where δ is an infinitesimal length.¹⁰

If the Hamiltonian has continuous translational symmetry, then it must be unchanged under any translation, including an infinitesimal one; equivalently it commutes with the translation operator, and hence

$$[\hat{H}, \hat{T}(\delta)] = [\hat{H}, 1 - i\frac{\delta}{\hbar} \hat{p}] = 0 \Rightarrow [\hat{H}, \hat{p}] = 0.$$

So if the Hamiltonian has continuous translational symmetry, it must commute with the momentum operator. And if the Hamiltonian commutes with momentum, then according to the “generalized Ehrenfest’s theorem” (Equation 3.73)

$$\frac{d}{dt} \langle p \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{p}] \rangle = 0. \quad (6.13)$$

This is a statement of **momentum conservation** and we have now shown that continuous translational symmetry implies that momentum is conserved. This is our first example of a powerful general principle: *symmetries imply conservation laws.*¹⁰

Of course, if we’re talking about a single particle of mass m moving in a potential $V(x)$, the only potential that has continuous translational symmetry is the constant potential, which is equivalent to the free particle. And it is pretty obvious that momentum is conserved in that case. But the analysis here readily extends to a system of interacting particles (see Problem 6.7). The fact that momentum is conserved in that case as well (so long as the Hamiltonian is translationally invariant) is a highly nontrivial result. In any event, the point to remember is that *conservation of momentum is a consequence of translational symmetry*.

Problem 6.5 Show that Equation 6.12 follows from Equation 6.11. *Hint:* First write $\psi(x) = e^{iqx} u(x)$, which is certainly true for *some* $u(x)$, and then show that $u(x)$ is necessarily a periodic function of x .

*** 

Problem 6.6 Consider a particle of mass m moving in a potential $V(x)$ with period a . We know from Bloch’s theorem that the wave function can be written in the form of Equation 6.12. *Note:* It is conventional to label the states with quantum numbers n and q as $\psi_{nq}(x) = e^{-iqx} u_{nq}(x)$ where E_{nq} is the n th energy for a given value of q .

(a) Show that u satisfies the equation

$$-\frac{\hbar^2}{2m} \frac{d^2 u_{nq}}{dx^2} - \frac{i\hbar^2 q}{m} \frac{du_{nq}}{dx} + V(x) u_{nq} = \left(E_{nq} - \frac{\hbar^2 q^2}{2m} \right) u_{nq}.$$

- (b) Use the technique from Problem [2.61](#) to solve the differential equation for u_{nq} . You need to use a two-sided difference for the first derivative so that you have a Hermitian matrix to diagonalize: $\frac{d\psi}{dx} \approx \frac{\psi_{j+1} - \psi_{j-1}}{2\Delta x}$. For the potential in the interval 0 to a let

$$V(x) = \begin{cases} -V_0 & a/4 < x < 3a/4 \\ 0 & \text{else} \end{cases}$$

with $V_0 = 20\hbar^2/2ma^2$. (You will need to modify the technique slightly to account for the fact that the function u_{nq} is periodic.) Find the lowest two energies for the following values of the crystal momentum: $qa = -\pi, -\pi/2, 0, \pi/2, \pi$. Note that q and $q + 2\pi/a$ describe the same wave function (Equation [6.12](#)), so there is no reason to consider values of qa outside of the interval from $-\pi$ to π . In solid state physics, the values of q inside this range constitute the **first Brillouin zone**.

- (c) Make a plot of the energies E_{1q} and E_{2q} for values of q between $-\pi/a$ and π/a . If you've automated the code that you used in part (b), you should be able to show a large number of q values in this range. If not, simply plot the values that you computed in (b).

Problem 6.7 Consider two particles of mass m_1 and m_2 (in one dimension) that interact via a potential that depends only on the distance between the particles $V(|x_1 - x_2|)$, so that the Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m_1} \frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m_2} \frac{\partial^2}{\partial x_2^2} + V(|x_1 - x_2|).$$

Acting on a two-particle wave function the translation operator would be

$$\hat{T}(a) \psi(x_1, x_2) = \psi(x_1 - a, x_2 - a).$$

- (a) Show that the translation operator can be written

$$\hat{T}(a) = e^{-\frac{ia}{\hbar} \hat{P}},$$

where $\hat{P} = \hat{p}_1 + \hat{p}_2$ is the total momentum.

- (b) Show that the total momentum is conserved for this system.

6.3 Conservation Laws

In classical mechanics the meaning of a conservation law is straightforward: the quantity in question is the same before and after some event. Drop a rock, and potential energy is converted into kinetic energy, but the *total* is the same just before it hits the ground as when it was released; collide two billiard balls and momentum is transferred from one to the other, but the total remains unchanged. But in quantum mechanics a system does not in general *have* a definite energy (or momentum) before the process begins (or afterward). What does it *mean*, in that case, to say that the observable Q is (or is not) conserved? Here are two possibilities:

- **First definition:** The *expectation value* $\langle Q \rangle$ is independent of time.
- **Second definition:** The probability of getting any particular value is independent of time.

Under what conditions does each of these conservation laws hold?

Let us stipulate that the observable in question does not depend explicitly on time: $\partial Q / \partial t = 0$. In that case the generalized Ehrenfest theorem (Equation 3.73) tells us that the expectation value of Q is independent of time if *The operator \hat{Q} commutes with the Hamiltonian*. It so happens that the same criterion guarantees conservation by the second definition.

I will now prove this result. Recall that the probability of getting the result q_n in a measurement of Q at time t is (Equation 3.43)

$$P(q_n) = |\langle f_n | \Psi(t) \rangle|^2, \quad (6.15)$$

where f_n is the corresponding eigenvector: $\hat{Q} |f_n\rangle = q_n |f_n\rangle$.¹¹ We know that the time evolution of the wave function is (Equation 2.17)

$$|\Psi(t)\rangle = \sum_m e^{-iE_m t/\hbar} c_m |\psi_m\rangle,$$

where the $|\psi_n\rangle$ are the eigenstates of \hat{H} , and therefore

$$P(q_n) = \left| \sum_m e^{-iE_m t/\hbar} c_m \langle f_n | \psi_m \rangle \right|^2.$$

Now the key point: since \hat{Q} and \hat{H} commute we can find a complete set of simultaneous eigenstates for them (see Section A.5); without loss of generality then $|f_n\rangle = |\psi_n\rangle$. Using the orthonormality of the $|\psi_n\rangle$,

$$P(q_n) = \left| \sum_m e^{-iE_m t/\hbar} c_m \langle \psi_n | \psi_m \rangle \right|^2 = |c_n|^2,$$

which is clearly independent of time.

6.4 Parity

6.4.1 Parity in One Dimension

A spatial inversion is implemented by the parity operator $\hat{\Pi}$; in one dimension,

$$\hat{\Pi}\psi(x) = \psi'(x) = \psi(-x).$$

Evidently, the parity operator is its own inverse: $\hat{\Pi}^{-1} = \hat{\Pi}$; in Problem 6.8 you will show that it is Hermitian: $\hat{\Pi}^\dagger = \hat{\Pi}$. Putting this together, the parity operator is unitary as well:

$$\hat{\Pi}^{-1} = \hat{\Pi} = \hat{\Pi}^\dagger. \quad (6.16)$$

Operators transform under a spatial inversion as

$$\hat{Q}' = \hat{\Pi}^\dagger \hat{Q} \hat{\Pi}. \quad (6.17)$$

I won't repeat the argument leading up to Equation 6.17, since it is identical to the one by which we arrived at Equation 6.6 in the case of translations. The position and momentum operators are "odd under parity" (Problem 6.10):

$$\hat{x}' = \hat{\Pi}^\dagger \hat{x} \hat{\Pi} = -\hat{x}, \quad (6.18)$$

$$\hat{p}' = \hat{\Pi}^\dagger \hat{p} \hat{\Pi} = -\hat{p}, \quad (6.19)$$

and this tells us how *any* operator transforms (see Problem 6.4):

$$\hat{Q}'(\hat{x}, \hat{p}) = \hat{\Pi}^\dagger \hat{Q}(\hat{x}, \hat{p}) \hat{\Pi} = \hat{Q}(-\hat{x}, -\hat{p}).$$

A system has **inversion symmetry** if the Hamiltonian is unchanged by a parity transformation:

$$\hat{H}' = \hat{\Pi}^\dagger \hat{H} \hat{\Pi} = \hat{H},$$

or, using the unitarity of the parity operator,

$$[\hat{H}, \hat{\Pi}] = 0. \quad (6.20)$$

If our Hamiltonian describes a particle of mass m in a one-dimensional potential $V(x)$, then inversion symmetry simply means that the potential is an even function of position:

$$V(x) = V(-x).$$

The implications of inversion symmetry are two: First, we can find a complete set of simultaneous eigenstates of $\hat{\Pi}$ and \hat{H} . Let such an eigenstate be written ψ_n ; it satisfies

$$\hat{\Pi}\psi_n(x) = \psi_n(-x) = \pm\psi_n(x),$$

since the eigenvalues of the parity operator are restricted to ± 1 (Problem 6.8). So the stationary states of a potential that is an even function of position are themselves even or odd functions (or can be chosen as such, in the case of degeneracy).¹² This property is familiar from the simple harmonic oscillator, the infinite square well (if the origin is placed at the center of the well), and the Dirac delta function potential, and you proved it in general in Problem 2.1.

Second, according to Ehrenfest's theorem, if the Hamiltonian has an inversion symmetry then

$$\frac{d}{dt} \langle \Pi \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{\Pi}] \rangle = 0$$

so *parity is conserved* for a particle moving in a symmetric potential. And not just the expectation value, but the *probability* of any particular outcome in a measurement, in accord with the theorem of Section 6.3. Parity conservation means, for example, that if the wave function of a particle in a harmonic oscillator potential is even at $t = 0$ then it will be even at any later time t ; see Figure 6.7.

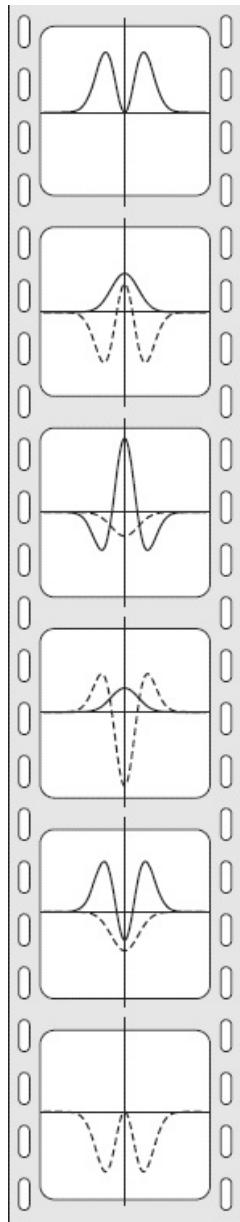


Figure 6.7: This filmstrip shows the time evolution of a particular wave function ($\Psi_0(\xi) = A\xi^2 e^{-\xi^2/2}$) for a particle in the harmonic oscillator potential. The solid and dashed curves are the real and imaginary parts of the wave function respectively, and time increases from top to bottom. Since parity is conserved, a wave function which is initially an even function of position (as this one is) remains an even function at all later times.

* **Problem 6.8**

- (a) Show that the parity operator $\hat{\Pi}$ is Hermitian.
- (b) Show that the eigenvalues of the parity operator are ± 1 .

6.4.2 Parity in Three Dimensions

The spatial inversion generated by the parity operator in three dimensions is

$$\hat{\Pi}\psi(\mathbf{r}) = \psi'(\mathbf{r}) = \psi(-\mathbf{r}).$$

The operators $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ transform as

$$\hat{\mathbf{r}}' = \hat{\Pi}^\dagger \hat{\mathbf{r}} \hat{\Pi} = -\hat{\mathbf{r}}, \quad (6.21)$$

$$\hat{\mathbf{p}}' = \hat{\Pi}^\dagger \hat{\mathbf{p}} \hat{\Pi} = -\hat{\mathbf{p}}. \quad (6.22)$$

Any other operator transforms as

$$\hat{Q}'(\hat{\mathbf{r}}, \hat{\mathbf{p}}) = \hat{\Pi}^\dagger \hat{Q}(\hat{\mathbf{r}}, \hat{\mathbf{p}}) \hat{\Pi} = \hat{Q}(-\hat{\mathbf{r}}, -\hat{\mathbf{p}}). \quad (6.23)$$

Example 6.2

Find the parity-transformed angular momentum operator $\hat{\mathbf{L}}' = \hat{\Pi}^\dagger \hat{\mathbf{L}} \hat{\Pi}$, in terms of $\hat{\mathbf{L}}$.

Solution: Since $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, Equation 6.23 tells us that

$$\hat{\mathbf{L}}' = \hat{\Pi}^\dagger \hat{\mathbf{L}} \hat{\Pi} = \hat{\mathbf{r}}' \times \hat{\mathbf{p}}' = (-\hat{\mathbf{r}}) \times (-\hat{\mathbf{p}}) = \hat{\mathbf{r}} \times \hat{\mathbf{p}} = \hat{\mathbf{L}}. \quad (6.24)$$

We have a special name for vectors like $\hat{\mathbf{L}}$, that are *even* under parity. We call them **pseudovectors**, since they don't change sign under parity the way "true" vectors, such as $\hat{\mathbf{r}}$ or $\hat{\mathbf{p}}$, do. Similarly, scalars that are *odd* under parity are called **pseudoscalars**, since they do not behave under parity the way that "true" scalars (such as $\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}$ which is even under parity) do. See Problem 6.9. *Note:* The labels **scalar** and **vector** describe how the operators behave under *rotations*; we will define these terms carefully in the next section. "True" vectors and pseudovectors behave the same way under a rotation—they are both vectors.

In three dimensions, the Hamiltonian for a particle of mass m moving in a potential $V(\mathbf{r})$ will have inversion symmetry if $V(-\mathbf{r}) = V(\mathbf{r})$. Importantly, any central potential satisfies this condition. As in the one-dimensional case, parity is conserved for such systems, and the eigenstates of the Hamiltonian may be chosen to be simultaneously eigenstates of parity. In Problem 6.1 you proved that the eigenstates of a particle in a central potential, written $\psi_{n\ell m}(r, \theta, \phi) = R_{n\ell}(r) Y_\ell^m(\theta, \phi)$, are eigenstates of parity:¹³

$$\hat{\Pi}\psi_{n\ell m}(r, \theta, \phi) = (-1)^\ell \psi_{n\ell m}(r, \theta, \phi).$$

* Problem 6.9

- (a) Under parity, a "true" scalar operator does not change:

$$\hat{\Pi}^\dagger \hat{f} \hat{\Pi} = \hat{f},$$

whereas a pseudoscalar changes sign. Show therefore that $[\hat{\Pi}, \hat{f}] = 0$ for a "true" scalar, whereas $\{\hat{\Pi}, \hat{f}\} = 0$ for a pseudoscalar. *Note:* the

anti-commutator of two operators \hat{A} and \hat{B} is defined as
 $\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$.

- (b) Similarly, a “true” vector changes sign

$$\hat{\Pi}^\dagger \hat{\mathbf{V}} \hat{\Pi} = -\hat{\mathbf{V}},$$

whereas a pseudovector is unchanged. Show therefore that $\{\hat{\Pi}, \hat{\mathbf{V}}\} = \mathbf{0}$ for a “true” vector and $[\hat{\Pi}, \hat{\mathbf{V}}] = \mathbf{0}$ for a pseudovector.

6.4.3 Parity Selection Rules

Selection rules tell you when a matrix element is zero based on the symmetry of the situation. Recall that a matrix element is any object of the form $\langle b | \hat{Q} | a \rangle$; an expectation value is a special case of a matrix element with $a = b = \psi$. One operator whose selection rules are physically important is the electric dipole moment operator

$$\hat{\mathbf{p}}_e = q\hat{\mathbf{r}}.$$

The selection rules for this operator—the operator itself is nothing more than the charge of the particle times its position—determine which atomic transitions are allowed and which are forbidden (see Chapter 11). It is odd under parity since the position vector $\hat{\mathbf{r}}$ is odd:

$$\hat{\Pi}^\dagger \hat{\mathbf{p}}_e \hat{\Pi} = -\hat{\mathbf{p}}_e. \quad (6.25)$$

Now consider the matrix elements of the electric dipole operator between two states $\psi_{n\ell m}$ and $\psi_{n'\ell' m'}$ (we label the corresponding kets $|n\ell m\rangle$ and $|n'\ell' m'\rangle$). Using Equation 6.25 we have

$$\begin{aligned} \langle n'\ell'm' | \hat{\mathbf{p}}_e | n\ell m \rangle &= -\langle n'\ell'm' | \hat{\Pi}^\dagger \hat{\mathbf{p}}_e \hat{\Pi} | n\ell m \rangle \\ &= -\langle n'\ell'm' | (-1)^{\ell'} \hat{\mathbf{p}}_e (-1)^\ell | n\ell m \rangle \\ &= (-1)^{\ell+\ell'+1} \langle n'\ell'm' | \hat{\mathbf{p}}_e | n\ell m \rangle. \end{aligned}$$

From this we see immediately that

$$\langle n'\ell'm' | \hat{\mathbf{p}}_e | n\ell m \rangle = 0 \text{ if } \ell + \ell' \text{ is even.} \quad (6.26)$$

This is called **Laporte's rule**; it says that matrix elements of the dipole moment operator vanish between states with the same parity. The reasoning by which we obtained Equation 6.26 can be generalized to derive selection rules for any operator, as long as you know how that operator transforms under parity. In particular, *Laporte's rule applies to any operator that is odd under parity*. The selection rule for an operator that is *even* under parity, such as $\hat{\mathbf{L}}$, is derived in Problem 6.11.

Problem 6.10 Show that the position and momentum operators are odd under parity. That is, prove Equations 6.18, 6.19, and, by extension, 6.21 and 6.22.

- * **Problem 6.11** Consider the matrix elements of $\hat{\mathbf{L}}$ between two definite-parity states: $\langle n'\ell'm' | \hat{\mathbf{L}} | n\ell m \rangle$. Under what conditions is this matrix element guaranteed to vanish? Note that the same selection rule would apply to any pseudovector operator, or any “true” scalar operator.

Problem 6.12 Spin angular momentum, $\hat{\mathbf{S}}$, is even under parity, just like orbital angular momentum $\hat{\mathbf{L}}$:

$$\hat{\Pi}^\dagger \hat{\mathbf{S}} \hat{\Pi} = \hat{\mathbf{S}} \quad \text{or} \quad [\hat{\Pi}, \hat{\mathbf{S}}] = 0. \quad (6.27)$$

Acting on a spinor written in the standard basis (Equation 4.139), the parity operator becomes a 2×2 matrix. Show that, due to Equation 6.27, this matrix must be a constant times the identity matrix. As such, the parity of a spinor isn't very interesting since both spin states are parity eigenstates with the same eigenvalue. We can arbitrarily choose that parity to be +1, so the parity operator has no effect on the spin portion of the wave function.¹⁴

*

Problem 6.13 Consider an electron in a hydrogen atom.

- (a) Show that if the electron is in the ground state, then necessarily $\langle \mathbf{p}_e \rangle = 0$.
No calculation allowed.
- (b) Show that if the electron is in an $n = 2$ state, then $\langle \mathbf{p}_e \rangle$ need not vanish.
Give an example of a wave function for the energy level $n = 2$ that has a non-vanishing $\langle \mathbf{p}_e \rangle$ and compute $\langle \mathbf{p}_e \rangle$ for this state.

6.5 Rotational Symmetry

6.5.1 Rotations About the z Axis

The operator that rotates a function about the z axis by an angle φ (Equation 6.2)

$$\hat{R}_z(\varphi) \psi(r, \theta, \phi) = \psi'(r, \theta, \phi) = \psi(r, \theta, \phi - \varphi) \quad (6.28)$$

is closely related to the z component of angular momentum (Equation 4.129). By the same reasoning that led to Equation 6.3,

$$\hat{R}_z(\varphi) = \exp\left[-\frac{i\varphi}{\hbar}\hat{L}_z\right], \quad (6.29)$$

and we say that \hat{L}_z is the **generator of rotations** about the z axis (compare Equation 6.3).

How do the operators $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ transform under rotations? To answer this question we use the infinitesimal form of the operator:

$$\hat{R}_z(\delta) \approx 1 - \frac{i\delta}{\hbar}\hat{L}_z.$$

Then the operator \hat{x} transforms as

$$\begin{aligned} \hat{x}' &= \hat{R}^\dagger \hat{x} \hat{R} \approx \left(1 + \frac{i\delta}{\hbar}\hat{L}_z\right) \hat{x} \left(1 - \frac{i\delta}{\hbar}\hat{L}_z\right) \\ &\approx \hat{x} + \frac{i\delta}{\hbar} [\hat{L}_z, \hat{x}] \approx \hat{x} - \delta \hat{y} \end{aligned}$$

(I used Equation 4.122 for the commutator). Similar calculations show that $\hat{y}' = \hat{y} + \delta \hat{x}$ and $\hat{z}' = \hat{z}$. We can combine these results into a matrix equation

$$\begin{pmatrix} \hat{x}' \\ \hat{y}' \\ \hat{z}' \end{pmatrix} = \begin{pmatrix} 1 & -\delta & 0 \\ \delta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix}. \quad (6.30)$$

That doesn't look quite right for a rotation. Shouldn't it be

$$\begin{pmatrix} \hat{x}' \\ \hat{y}' \\ \hat{z}' \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix} ? \quad (6.31)$$

Yes, but don't forget, we are assuming $\varphi \rightarrow \delta$ is infinitesimal, so (dropping terms of order δ^2 and higher) $\cos \varphi \rightarrow 1$ and $\sin \varphi \rightarrow \delta$.¹⁵

**

Problem 6.14 In this problem you will establish the correspondence between Equations 6.30 and 6.31.

(a) Diagonalize the matrix¹⁶

$$M = \begin{pmatrix} 1 & -\varphi/N \\ \varphi/N & 1 \end{pmatrix}$$

to obtain the matrix

$$\mathbf{M}' = \mathbf{S} \mathbf{M} \mathbf{S}^{-1},$$

where \mathbf{S}^{-1} is the unitary matrix whose columns are the (normalized) eigenvectors of \mathbf{M} .

- (b) Use the binomial expansion to show that $\lim_{N \rightarrow \infty} (\mathbf{M}')^N$ is a diagonal matrix with entries $e^{-i\varphi}$ and $e^{i\varphi}$ on the diagonal.
- (c) Transform back to the original basis to show that

$$\lim_{N \rightarrow \infty} \mathbf{M}^N = \mathbf{S}^{-1} \left[\lim_{N \rightarrow \infty} (\mathbf{M}')^N \right] \mathbf{S}$$

agrees with the matrix in Equation [6.31](#).

6.5.2 Rotations in Three Dimensions

Equation 6.29 can be generalized in the obvious way to a rotation about an axis along the unit vector \mathbf{n} :

$$\boxed{\hat{R}_{\mathbf{n}}(\varphi) = \exp\left[-\frac{i\varphi}{\hbar}\mathbf{n} \cdot \hat{\mathbf{L}}\right]} \quad (6.32)$$

Just as linear momentum is the generator of translations, angular momentum is the **generator of rotations**.

Any operator (with three components) that transforms the same way as the position operator under rotations is called a **vector operator**. By “transforms the same way” we mean that $\hat{\mathbf{V}}' = \mathbf{D}\hat{\mathbf{V}}$ where \mathbf{D} is the same matrix as appears in $\hat{\mathbf{r}}' = \mathbf{D}\hat{\mathbf{r}}$. In particular for a rotation about the z axis, we would have (Equation 6.31)

$$\begin{pmatrix} \hat{V}'_x \\ \hat{V}'_y \\ \hat{V}'_z \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{V}_x \\ \hat{V}_y \\ \hat{V}_z \end{pmatrix}.$$

This transformation rule follows from the commutation relations¹⁷

$$\boxed{[\hat{L}_i, \hat{V}_j] = i\hbar\epsilon_{ijk}\hat{V}_k,} \quad (6.33)$$

(see Problem 6.16), and we may take Equation 6.33 as the *definition* of a vector operator. So far we have encountered three such operators, $\hat{\mathbf{r}}$, $\hat{\mathbf{p}}$ and $\hat{\mathbf{L}}$:

$$[\hat{L}_i, \hat{r}_j] = i\hbar\epsilon_{ijk}\hat{r}_k, \quad [\hat{L}_i, \hat{p}_j] = i\hbar\epsilon_{ijk}\hat{p}_k, \quad [\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k$$

(see Equations 4.99 and 4.122).

A **scalar operator** is a single quantity that is unchanged by rotations; this is equivalent to saying that the operator commutes with $\hat{\mathbf{L}}$:

$$\boxed{[\hat{L}_i, \hat{f}] = 0.} \quad (6.34)$$

We can now classify operators as either scalars or vectors, based on their commutation relations with $\hat{\mathbf{L}}$ (how they transform under a rotation), and as “true” or pseudo-quantities, based on their commutators with $\hat{\Pi}$ (how they transform under parity). These results are summarized in Table 6.1.¹⁸

Table 6.1: Operators are classified as vectors or scalars based on their commutation relations with $\hat{\mathbf{L}}$, which encode how they transform under a rotation, and as pseudo- or “true” quantities based on their commutation relations with $\hat{\Pi}$, which encode how they transform under a spatial inversion. The curly brackets in the first column denote the anti-commutator, defined in Problem 6.9. To include the spin $\hat{\mathbf{S}}$ in this table, one simply replaces \hat{L}_i everywhere it appears in the third column with $\hat{J}_i = \hat{L}_i + \hat{S}_i$ (Problems 6.12 and 6.32, respectively, discuss the effect of parity and rotations on spinors). $\hat{\mathbf{S}}$, like $\hat{\mathbf{L}}$, is then a pseudovector and $\hat{\mathbf{p}} \cdot \hat{\mathbf{S}}$ is a pseudoscalar.

	parity	rotations	examples
true vector $\hat{\mathbf{V}}$	$\{\hat{\Pi}, \hat{V}_i\} = 0$	$[\hat{L}_i, \hat{V}_j] = i \hbar \epsilon_{ijk} \hat{V}_k$	$\hat{\mathbf{r}}, \hat{\mathbf{p}}$
pseudovector $\hat{\mathbf{V}}$	$[\hat{\Pi}, \hat{V}_i] = 0$	$[\hat{L}_i, \hat{V}_j] = i \hbar \epsilon_{ijk} \hat{V}_k$	$\hat{\mathbf{L}}$
true scalar \hat{f}	$[\hat{\Pi}, \hat{f}] = 0$	$[\hat{L}_i, \hat{f}] = 0$	$\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}$
pseudoscalar \hat{f}	$\{\hat{\Pi}, \hat{f}\} = 0$	$[\hat{L}_i, \hat{f}] = 0$	

continuous rotational symmetry

For a particle of mass m moving in a potential $V(\mathbf{r})$, the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\mathbf{r})$$

is rotationally invariant if $V(\mathbf{r}) = V(r)$ (the central potentials studied in Section 4.1.1). In this case the Hamiltonian commutes with a rotation by any angle about an arbitrary axis

$$[\hat{H}, \hat{R}_{\mathbf{n}}(\varphi)] = 0. \quad (6.35)$$

In particular, Equation 6.35 must hold for an *infinitesimal* rotation

$$R_{\mathbf{n}}(\delta) \approx 1 - i \frac{\delta}{\hbar} \mathbf{n} \cdot \hat{\mathbf{L}},$$

which means that the Hamiltonian commutes with the three components of \mathbf{L} :

$$[\hat{H}, \hat{\mathbf{L}}] = \mathbf{0}. \quad (6.36)$$

What, then, are the consequences of rotational invariance?

From Equation 6.36 and Ehrenfest's theorem

$$\frac{d}{dt} \langle \mathbf{L} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{\mathbf{L}}] \rangle = \mathbf{0} \quad (6.37)$$

for a central potential. Thus, *angular momentum conservation is a consequence of rotational invariance*. And beyond the statement 6.37, angular momentum conservation means that the probability distributions (for each component of the angular momentum) are independent of time as well—see Section 6.3.

Since the Hamiltonian for a central potential commutes with all three components of angular momentum, it also commutes with \hat{L}^2 . The operators \hat{H} , \hat{L}_z , and \hat{L}^2 form a **complete set of compatible observables** for the bound states of a central potential. Compatible means that they commute pairwise

$$\begin{aligned} [\hat{H}, \hat{L}^2] &= 0, \\ [\hat{H}, \hat{L}_z] &= 0, \\ [\hat{L}_z, \hat{L}^2] &= 0, \end{aligned} \quad (6.38)$$

so that the eigenstates of \hat{H} can be chosen to be simultaneous eigenstates of \hat{L}^2 and \hat{L}_z .

$$\begin{aligned} \hat{H}\psi_{n\ell m} &= E_n\psi_{n\ell m}, \\ \hat{L}_z\psi_{n\ell m} &= m\hbar\psi_{n\ell m}, \\ \hat{L}^2\psi_{n\ell m} &= \ell(\ell+1)\hbar^2\psi_{n\ell m}. \end{aligned}$$

Saying they are complete means that the quantum numbers n , ℓ , and m uniquely specify a bound state of the Hamiltonian. This is familiar from our solution to the hydrogen atom, the infinite spherical well, and the three-dimensional harmonic oscillator, but it is true for *any* central potential.¹⁹

- * **Problem 6.15** Show how Equation [6.34](#) guarantees that a scalar is unchanged by a rotation: $\hat{f}' = \hat{R}^\dagger \hat{f} \hat{R} = \hat{f}$.

- * **Problem 6.16** Working from Equation [6.33](#), find how the vector operator $\hat{\mathbf{V}}$ transforms for an infinitesimal rotation by an angle δ about the y axis. That is, find the matrix \mathbf{D} in

$$\hat{\mathbf{V}}' = \mathbf{D}\hat{\mathbf{V}}.$$

Problem 6.17 Consider the action of an infinitesimal rotation about the \mathbf{n} axis of an angular momentum eigenstate $\psi_{n\ell m}$. Show that

$$\hat{R}_{\mathbf{n}}(\delta) \psi_{n\ell m} = \sum_{m'} D_{m'm} \psi_{n\ell m'},$$

and find the complex numbers $D_{m'm}$ (they will depend on δ , \mathbf{n} , and ℓ as well as m and m'). This result makes sense: a rotation doesn't change the magnitude of the angular momentum (specified by ℓ) but does change its projection along the z axis (specified by m).

6.6 Degeneracy

Symmetry is the source of most²⁰ degeneracy in quantum mechanics. We have seen that a symmetry implies the existence of an operator \hat{Q} that commutes with the Hamiltonian

$$[\hat{H}, \hat{Q}] = 0. \quad (6.39)$$

So why does symmetry lead to degeneracy in the energy spectrum? The basic idea is this: if we have a stationary state $|\psi_n\rangle$, then $|\psi'_n\rangle = \hat{Q}|\psi_n\rangle$ is a stationary state with the same energy. The proof is straightforward:

$$\hat{H}|\psi'_n\rangle = \hat{H}(\hat{Q}|\psi_n\rangle) = \hat{Q}\hat{H}|\psi_n\rangle = \hat{Q}E_n|\psi_n\rangle = E_n(\hat{Q}|\psi_n\rangle) = E_n|\psi'_n\rangle.$$

For example, if you have an eigenstate of a spherically-symmetric Hamiltonian and you rotate that state about some axis, you must get back another state of the same energy.

You might think that symmetry would *always* lead to degeneracy, and that continuous symmetries would lead to infinite degeneracy, but that is not the case. The reason is that the two states $|\psi_n\rangle$ and $|\psi'_n\rangle$ might be the same.²¹ As an example, consider the Hamiltonian for the harmonic oscillator in one dimension; it commutes with parity. All of its stationary states are either even or odd, so when you act on one with the parity operator you get back the same state you started with (perhaps multiplied by -1 , but that, physically, *is* the same state). There is therefore no degeneracy associated with inversion symmetry in this case.

In fact, if there is only a single symmetry operator \hat{Q} (or if there are multiple symmetry operators that all commute), you do not get degeneracy in the spectrum. The reason is the same theorem we've now quoted many times: since \hat{Q} and \hat{H} commute, we can find simultaneous eigenstates $|\psi_n\rangle$ of \hat{Q} and \hat{H} and these states transform into themselves under the symmetry operation: $\hat{Q}|\psi_n\rangle = q_n|\psi_n\rangle$.

But what if there are *two* operators that commute with the Hamiltonian (call them \hat{Q} and $\hat{\Lambda}$), but do *not* commute with each other? In this case, degeneracy in the energy spectrum is inevitable. Why?

First, consider a state $|\psi\rangle$ that is an eigenstate of both \hat{H} and \hat{Q} with eigenvalues E_n and q_m respectively. Since \hat{H} and $\hat{\Lambda}$ commute we know that the state $|g\rangle = \hat{\Lambda}|\psi\rangle$ is also an eigenstate of \hat{H} with eigenvalue E_n . Since \hat{Q} and $\hat{\Lambda}$ do *not* commute we know (Section A.5) that there cannot exist a complete set of simultaneous eigenstates of all three operators (\hat{Q} , $\hat{\Lambda}$ and \hat{H}). Therefore, there must be some $|\psi\rangle$ such that $\hat{\Lambda}|\psi\rangle$ is distinct from $|\psi\rangle$ (specifically, it is not an eigenstate of $\hat{\Lambda}$) meaning that the energy level E_n is at least doubly degenerate. *The presence of multiple non-commuting symmetry operators guarantees degeneracy of the energy spectrum.*

This is precisely the situation we have encountered in the case of central potentials. Here the Hamiltonian commutes with rotations about any axis (or equivalently with the generators \hat{L}_x , \hat{L}_y and \hat{L}_z) but those rotations don't commute with each other. So we know that there will be degeneracy in the spectrum of a particle in a central potential. The following example shows exactly how much degeneracy is explained by rotational invariance.

Example 6.3

Consider an eigenstate of a central potential $\psi_{n\ell m}$ with energy E_n . Use the fact that the Hamiltonian

for a central potential commutes with any component of $\hat{\mathbf{L}}$, and therefore also with \hat{L}_+ and \hat{L}_- , to show that $\psi_{n\ell m \pm 1}$ are necessarily also eigenstates with the same energy as $\psi_{n\ell m}$.²²

Solution: Since the Hamiltonian commutes with \hat{L}_\pm we have

$$(\hat{H}\hat{L}_\pm - \hat{L}_\pm\hat{H})\psi_{n\ell m} = 0,$$

so

$$\hat{H}\hat{L}_\pm\psi_{n\ell m} = \hat{L}_\pm\hat{H}\psi_{n\ell m} = E_n\hat{L}_\pm\psi_{n\ell m}$$

or

$$\hat{H}\psi_{n\ell m \pm 1} = E_n\psi_{n\ell m \pm 1}$$

(I canceled the constant $\hbar\sqrt{\ell(\ell+1) - m(m \pm 1)}$ from both sides in the last expression). This argument could obviously be repeated to show that $\psi_{n\ell m \pm 2}$ has the same energy as $\psi_{n\ell m \pm 1}$, and so on until you've exhausted the ladder of states. Therefore, *rotational invariance explains why states which differ only in the quantum number m have the same energy*, and since there are $2\ell + 1$ different values of m , $2\ell + 1$ is the "normal" degeneracy for energies in a central potential.

Of course, the degeneracy of hydrogen (neglecting spin) is $n^2 (= 1, 4, 9, \dots)$ (Equation 4.85) which is greater than $2\ell + 1 (= 1, 3, 5, \dots)$.²³ Evidently hydrogen has more degeneracy than is explained by rotational invariance alone. The source of the extra degeneracy is an additional symmetry that is unique to the $1/r$ potential; this is explored in Problem 6.34.²⁴

In this section we have focused on continuous rotational symmetry, but *discrete* rotational symmetry, as experienced (for instance) by an electron in a crystal, can also be of interest. Problem 6.33 explores one such system.

Problem 6.18 Consider the free particle in one dimension: $\hat{H} = \hat{p}^2/2m$. This Hamiltonian has both translational symmetry and inversion symmetry.

- (a) Show that translations and inversion don't commute.
- (b) Because of the translational symmetry we know that the eigenstates of \hat{H} can be chosen to be simultaneous eigenstates of momentum, namely $f_p(x)$ (Equation 3.32). Show that the parity operator turns $f_p(x)$ into $f_{-p}(x)$; these two states must therefore have the same energy.
- (c) Alternatively, because of the inversion symmetry we know that the eigenstates of \hat{H} can be chosen to be simultaneous eigenstates of parity, namely

$$\frac{1}{\sqrt{\pi\hbar}} \cos\left(\frac{px}{\hbar}\right) \text{ and } \frac{1}{\sqrt{\pi\hbar}} \sin\left(\frac{px}{\hbar}\right).$$

Show that the translation operator mixes these two states together; they therefore must be degenerate.

Note: Both parity *and* translational invariance are required to explain the degeneracy in the free-particle spectrum. Without parity, there is no reason for $f_p(x)$ and $f_{-p}(x)$ to have the same energy (I mean no reason based on symmetries discussed thus far ...obviously you can plug them in to the time-independent Schrödinger equation and show it's true).

Problem 6.19 For any vector operator $\hat{\mathbf{V}}$ one can define raising and lowering operators as

$$\hat{V}_\pm = \hat{V}_x \pm i\hat{V}_y.$$

- (a) Using Equation 6.33, show that

$$[\hat{L}_z, \hat{V}_\pm] = \pm \hbar \hat{V}_\pm.$$

$$[\hat{L}^2, \hat{V}_\pm] = 2\hbar^2 \hat{V}_\pm \pm 2\hbar \hat{V}_\pm \hat{L}_z \mp 2\hbar \hat{V}_z \hat{L}_\pm.$$

- (b) Show that, if ψ is an eigenstate of \hat{L}^2 and \hat{L}_z with eigenvalues $\ell(\ell+1)\hbar^2$ and $\ell\hbar$ respectively, then either $\hat{V}_+\psi$ is zero or $\hat{V}_+\psi$ is also an eigenstate of \hat{L}^2 and \hat{L}_z with eigenvalues $(\ell+1)(\ell+2)\hbar^2$ and $(\ell+1)\hbar$ respectively. This means that, acting on a state with maximal $m_\ell = \ell$, the operator \hat{V}_+ either “raises” both the ℓ and m values by 1 or destroys the state.

6.7 Rotational Selection Rules

The most general statement of the rotational selection rules is the Wigner–Eckart Theorem; as a practical matter, it is arguably the most important theorem in all of quantum mechanics. Rather than prove the theorem in full generality I will work out the selection rules for the two classes of operators one encounters most often: scalar operators (in Section 6.7.1) and vector operators (in Section 6.7.2). In deriving these selection rules we consider only how the operators behave under a rotation; therefore, the results of this section apply equally well to “true” scalars and pseudoscalars, and those of the next section apply equally well to “true” vectors and pseudeovectors. These selection rules can be combined with the parity selection rules of Section 6.4.3 to obtain a larger set of selection rules for the operator.

6.7.1 Selection Rules for Scalar Operators

The commutation relations for a scalar operator \hat{f} with the three components of angular momentum (Equation 6.34) can be rewritten in terms of the raising and lowering operators as

$$[\hat{L}_z, \hat{f}] = 0 \quad (6.40)$$

$$[\hat{L}_{\pm}, \hat{f}] = 0 \quad (6.41)$$

$$[\hat{L}^2, \hat{f}] = 0. \quad (6.42)$$

We derive selection rules for \hat{f} by sandwiching these commutators between two states of definite angular momentum, which we will write as $|n\ell m\rangle$ and $|n'\ell' m'\rangle$. These might be hydrogenic orbitals, but they need not be (in fact they need not even be eigenstates of any Hamiltonian but I'll leave the quantum number n there so they look familiar); we require only that $|n\ell m\rangle$ is an eigenstate of \hat{L}^2 and \hat{L}_z with quantum numbers ℓ and m respectively.²⁵

Sandwiching Equation 6.40 between two such states gives

$$\langle n'\ell' m' | [\hat{L}_z, \hat{f}] | n\ell m \rangle = 0$$

or

$$\langle n'\ell' m' | \hat{L}_z \hat{f} | n\ell m \rangle - \langle n'\ell' m' | \hat{f} \hat{L}_z | n\ell m \rangle = 0$$

and therefore

$$(m' - m) \langle n'\ell' m' | \hat{f} | n\ell m \rangle = 0 \quad (6.43)$$

(using the hermiticity of \hat{L}_z). Equation 6.43 says that the matrix elements of a scalar operator vanish unless $m' - m \equiv \Delta m = 0$. Repeating this procedure with Equation 6.42 we get

$$\begin{aligned} \langle n'\ell' m' | [\hat{L}^2, \hat{f}] | n\ell m \rangle &= 0 \\ \langle n'\ell' m' | \hat{L}^2 \hat{f} | n\ell m \rangle - \langle n'\ell' m' | \hat{f} \hat{L}^2 | n\ell m \rangle &= 0 \\ [\ell'(\ell' + 1) - \ell(\ell + 1)] \langle n'\ell' m' | \hat{f} | n\ell m \rangle &= 0. \end{aligned} \quad (6.44)$$

This tells us that the matrix elements of a scalar operator vanish unless $\ell' - \ell \equiv \Delta \ell = 0$.²⁶ These, then, are the selection rules for a scalar operator: $\Delta \ell = 0$ and $\Delta m = 0$.

However, we can get even more information about the matrix elements from the remaining commutators: (I'll just do the + case and leave the - case for Problem 6.20)

$$\begin{aligned} \langle n'\ell' m' | [\hat{L}_+, \hat{f}] | n\ell m \rangle &= 0 \\ \langle n'\ell' m' | \hat{L}_+ \hat{f} | n\ell m \rangle - \langle n'\ell' m' | \hat{f} \hat{L}_+ | n\ell m \rangle &= 0 \\ B_{\ell'}^{m'} \langle n'\ell' (m' - 1) | \hat{f} | n\ell m \rangle - A_{\ell}^m \langle n'\ell' m' | \hat{f} | n\ell (m + 1) \rangle &= 0, \end{aligned} \quad (6.45)$$

where (from Problem 4.21)

$$A_\ell^m = \hbar\sqrt{\ell(\ell+1) - m(m+1)}, \quad \text{and} \quad B_\ell^m = \hbar\sqrt{\ell(\ell+1) - m(m-1)}.$$

(I also used the fact that \hat{L}_\pm is the Hermitian conjugate of \hat{L}_\mp : $\langle \psi | \hat{L}_\pm = \langle L_\mp \psi |$)²⁷ Both terms in Equation 6.45 are zero unless $m' = m + 1$ and $\ell' = \ell$, as we proved in Equations 6.43 and 6.44. When these conditions are satisfied, the two coefficients are equal ($B_\ell^{m+1} = A_\ell^m$) and Equation 6.45 reduces to

$$\langle n'\ell'm' | \hat{f} | n\ell m \rangle = \langle n'\ell(m+1) | \hat{f} | n\ell(m+1) \rangle. \quad (6.46)$$

Evidently the matrix elements of a scalar operator are independent of m .

The results of this section can be summarized as follows:

$$\langle n'\ell'm' | \hat{f} | n\ell m \rangle = \delta_{\ell\ell'} \delta_{mm'} \langle n'\ell \| f \| n\ell \rangle. \quad (6.47)$$

The funny-looking matrix element on the right, with two bars, is called a **reduced matrix element** and is just shorthand for “a constant that depends on n , ℓ , and n' , but not m .”

Example 6.4

- (a) Find $\langle r^2 \rangle$ for all four of the degenerate $n = 2$ states of a hydrogen atom.

Solution: From Equation 6.47 we have, for the states with $\ell = 1$, the following equality:

$$\langle 211 | r^2 | 211 \rangle = \langle 210 | r^2 | 210 \rangle = \langle 21-1 | r^2 | 21-1 \rangle \equiv \langle 21 \| r^2 \| 21 \rangle.$$

To calculate the reduced matrix element we simply pick any one of these expectation values:

$$\begin{aligned} \langle 21 \| r^2 \| 21 \rangle &= \langle 210 | r^2 | 210 \rangle \\ &= \int r^2 |\psi_{210}(r)|^2 d^3\mathbf{r} \\ &= \int_0^\infty r^4 |R_{21}(r)|^2 dr \int |Y_1^0(\theta, \phi)|^2 d\Omega. \end{aligned}$$

The spherical harmonics are normalized (Equation 4.31), so the angular integral is 1, and the radial functions $R_{n\ell}$ are listed in Table 4.7, giving

$$\langle 21 \| r^2 \| 21 \rangle = \int_0^\infty r^4 \frac{1}{24a^3} \frac{r^2}{a^2} e^{-r/a} dr = 30a^2.$$

That determines three of the expectation values. The final expectation value is

$$\begin{aligned}
 \langle 20 | r^2 | 20 \rangle &= \langle 200 | r^2 | 200 \rangle \\
 &= \int r^2 |\psi_{200}(r)|^2 d^3\mathbf{r} \\
 &= \int_0^\infty r^4 |R_{20}(r)|^2 dr \int |Y_0^0(\theta, \phi)|^2 d\Omega \\
 &= \int_0^\infty r^4 \frac{1}{2a^3} \left(1 - \frac{1}{2} \frac{r}{a}\right)^2 e^{-r/a} dr \\
 &= 42a^2.
 \end{aligned}$$

Summarizing:

$$\begin{aligned}
 \langle 200 | r^2 | 200 \rangle &= 42a^2, & \left. \begin{aligned}
 \langle 211 | r^2 | 211 \rangle \\
 \langle 210 | r^2 | 210 \rangle \\
 \langle 21-1 | r^2 | 21-1 \rangle
 \end{aligned} \right\} &= 30a^2.
 \end{aligned} \tag{6.48}$$

(b) Find the expectation value of r^2 for an electron in the superposition state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|200\rangle - i|211\rangle).$$

Solution: We can expand the expectation value as

$$\begin{aligned}
 \langle \psi | r^2 | \psi \rangle &= \frac{1}{2} (\langle 200 | + i \langle 211 |) r^2 (\langle 200 | - i \langle 211 |) \\
 &= \frac{1}{2} \left(\langle 200 | r^2 | 200 \rangle + i \langle 211 | r^2 | 200 \rangle - i \langle 200 | r^2 | 211 \rangle \right. \\
 &\quad \left. + \langle 211 | r^2 | 211 \rangle \right).
 \end{aligned}$$

From Equation 6.47 we see that two of these matrix elements vanish, and

$$\langle \psi | r^2 | \psi \rangle = \frac{1}{2} \left(\langle 20 | r^2 | 20 \rangle + \langle 21 | r^2 | 21 \rangle \right) = 36a^2. \tag{6.49}$$

Problem 6.20 Show that the commutator $[\hat{L}_-, \hat{f}] = 0$ leads to the same rule, Equation 6.46, as does the commutator $[\hat{L}_+, \hat{f}] = 0$.

* **Problem 6.21** For an electron in the hydrogen state

$$\psi = \frac{1}{\sqrt{2}} (\psi_{211} + \psi_{21-1}),$$

find $\langle r \rangle$ after first expressing it in terms of a single reduced matrix element.

6.7.2 Selection Rules for Vector Operators

We now move on to the selection rules for a vector operator $\hat{\mathbf{V}}$. This is significantly more work than the scalar case, but the result is central to understanding atomic transitions (Chapter 11). We begin by defining, by analogy with the angular momentum raising and lowering operators, the operators²⁸

$$\hat{V}_\pm \equiv \hat{V}_x \pm i\hat{V}_y.$$

Written in terms of these operators, Equation 6.33 becomes

$$[\hat{L}_z, \hat{V}_z] = 0 \quad (6.50)$$

$$[\hat{L}_z, \hat{V}_\pm] = \pm \hbar \hat{V}_\pm \quad (6.51)$$

$$[\hat{L}_\pm, \hat{V}_\pm] = 0 \quad (6.52)$$

$$[\hat{L}_\pm, \hat{V}_z] = \mp \hbar \hat{V}_\pm \quad (6.53)$$

$$[\hat{L}_\pm, \hat{V}_\mp] = \pm 2\hbar \hat{V}_z \quad (6.54)$$

as you will show in Problem 6.22(a).²⁹ Just as for the scalar operator in Section 6.7.1, we sandwich each of these commutators between two states of definite angular momentum to derive (a) conditions under which the matrix elements are guaranteed to vanish and (b) relations between matrix elements with differing values of m or different components of $\hat{\mathbf{V}}$.

From Equation 6.51,

$$\langle n' \ell' m' | \hat{L}_z \hat{V}_\pm | n \ell m \rangle - \langle n' \ell' m' | \hat{V}_\pm \hat{L}_z | n \ell m \rangle = \pm \hbar \langle n' \ell' m' | \hat{V}_\pm | n \ell m \rangle,$$

and since our states are eigenstates of \hat{L}_z , this simplifies to

$$[m' - (m \pm 1)] \langle n' \ell' m' | \hat{V}_\pm | n \ell m \rangle = 0. \quad (6.55)$$

Equation 6.55 says that either $m' = m \pm 1$, or else the matrix element of \hat{V}_\pm must vanish. Equation 6.50 works out similarly (see Problem 6.22) and this first set of commutators gives us the selection rules for m :

$$\langle n' \ell' m' | \hat{V}_+ | n \ell m \rangle = 0 \quad \text{unless } m' = m + 1 \quad (6.56)$$

$$\langle n' \ell' m' | \hat{V}_z | n \ell m \rangle = 0 \quad \text{unless } m' = m \quad (6.57)$$

$$\langle n' \ell' m' | \hat{V}_- | n \ell m \rangle = 0 \quad \text{unless } m' = m - 1. \quad (6.58)$$

Note that, if desired, these expressions can be turned back into selection rules for the x - and y -components of our operator, since

$$\begin{aligned}\langle n' \ell' m' | \hat{V}_x | n \ell m \rangle &= \frac{1}{2} [\langle n' \ell' m' | \hat{V}_- | n \ell m \rangle + \langle n' \ell' m' | \hat{V}_+ | n \ell m \rangle] \\ \langle n' \ell' m' | \hat{V}_y | n \ell m \rangle &= \frac{i}{2} [\langle n' \ell' m' | \hat{V}_- | n \ell m \rangle - \langle n' \ell' m' | \hat{V}_+ | n \ell m \rangle].\end{aligned}$$

The remaining commutators, Equations 6.52–6.54, yield a selection rule on ℓ and relations among the nonzero matrix elements. As shown in Problem 6.24, the results may be summarized as³⁰

$$\langle n' \ell' m' | \hat{V}_+ | n \ell m \rangle = -\sqrt{2} C_{m1m'}^{\ell 1 \ell'} \langle n \ell' \| V \| n \ell \rangle \quad (6.59)$$

$$\langle n' \ell' m' | \hat{V}_- | n \ell m \rangle = \sqrt{2} C_{m-1m'}^{\ell 1 \ell'} \langle n \ell' \| V \| n \ell \rangle \quad (6.60)$$

$$\langle n' \ell' m' | \hat{V}_z | n \ell m \rangle = C_{m0m'}^{\ell 1 \ell'} \langle n \ell' \| V \| n \ell \rangle. \quad (6.61)$$

The constants $C_{m_1 m_2 M}^{j_1 j_2 J}$ in these expressions are precisely the Clebsch–Gordan coefficients that appeared in the addition of angular momenta (Section 4.4.3). The Clebsch–Gordan coefficient $C_{m_1 m_2 M}^{j_1 j_2 J}$ vanishes unless $M = m_1 + m_2$ (since the z -components of angular momentum add) and unless $J = j_1 + j_2, j_1 + j_2 - 1, \dots, |j_1 - j_2|$ (Equation 4.182). In particular, the matrix elements of any component of a vector operator, $\langle n' \ell' m' | \hat{V}_i | n \ell m \rangle$, are nonzero only if

$\Delta \ell = 0, \pm 1,$ and $\Delta m = 0, \pm 1.$

(6.62)

Example 6.5

Find all of the matrix elements of $\hat{\mathbf{r}}$ between the states with $n = 2, \ell = 1$ and $n' = 3, \ell' = 2$:

$$\langle 32m' | r_i | 21m \rangle$$

where $m = -1, 0, 1, m' = -2, -1, 0, 1, 2$, and $r_i = x, y, z$.

Solution: With the vector operator $\hat{\mathbf{V}} = \hat{\mathbf{r}}$, our components are $V_z = z$, $V_+ = x + iy$, and $V_- = x - iy$. We start by calculating *one* of the matrix elements,

$$\begin{aligned}\langle 320 | z | 210 \rangle &= \int \psi_{320}(\mathbf{r}) r \cos \theta \psi_{210}(\mathbf{r}) d^3 \mathbf{r} \\ &= \int R_{32}(r)^* r R_{21}(r) r^2 dr \int Y_2^0(\theta, \phi)^* \cos \theta Y_1^0(\theta, \phi) d\Omega \\ &= \frac{2^{12} 3^3 \sqrt{3}}{5^7} a.\end{aligned}$$

From Equation 6.61 we can then determine the reduced matrix element

$$\begin{aligned}\langle 320 | z | 210 \rangle &= C_{000}^{112} \langle 32 \| V \| 21 \rangle \\ \frac{2^{12} 3^3 \sqrt{3}}{5^7} a &= \sqrt{\frac{2}{3}} \langle 32 \| V \| 21 \rangle.\end{aligned}$$

Therefore

$$\langle 32 \| V \| 21 \rangle = \frac{2^{12} 3^4}{5^7 \sqrt{2}} a. \quad (6.63)$$

We can now find all of the remaining matrix elements from Equations 6.59–6.60 with the help of the Clebsch–Gordan table. The relevant coefficients are shown in Figure 6.8. The nonzero matrix elements are

$$\begin{aligned}
 \langle 322 | \hat{V}_+ | 211 \rangle &= -\sqrt{2} C_{112}^{112} \langle 32 \| V \| 21 \rangle = -\sqrt{2} \langle 32 \| V \| 21 \rangle \\
 \langle 321 | \hat{V}_+ | 210 \rangle &= -\sqrt{2} C_{011}^{112} \langle 32 \| V \| 21 \rangle = -\langle 32 \| V \| 21 \rangle \\
 \langle 320 | \hat{V}_+ | 21-1 \rangle &= -\sqrt{2} C_{-110}^{-112} \langle 32 \| V \| 21 \rangle = -\frac{1}{\sqrt{3}} \langle 32 \| V \| 21 \rangle \\
 \langle 320 | \hat{V}_- | 211 \rangle &= \sqrt{2} C_{1-10}^1 \langle 32 \| V \| 21 \rangle = \frac{1}{\sqrt{3}} \langle 32 \| V \| 21 \rangle \\
 \langle 32-1 | \hat{V}_- | 210 \rangle &= \sqrt{2} C_{0-1-1}^1 \langle 32 \| V \| 21 \rangle = \langle 32 \| V \| 21 \rangle \\
 \langle 32-2 | \hat{V}_- | 21-1 \rangle &= \sqrt{2} C_{-1-1-2}^1 \langle 32 \| V \| 21 \rangle = \sqrt{2} \langle 32 \| V \| 21 \rangle \\
 \langle 321 | \hat{V}_z | 211 \rangle &= C_{101}^{112} \langle 32 \| V \| 21 \rangle = \frac{1}{\sqrt{2}} \langle 32 \| V \| 21 \rangle \\
 \langle 320 | \hat{V}_z | 210 \rangle &= C_{000}^{112} \langle 32 \| V \| 21 \rangle = \sqrt{\frac{2}{3}} \langle 32 \| V \| 21 \rangle \\
 \langle 32-1 | \hat{V}_z | 21-1 \rangle &= C_{-10-1}^{-112} \langle 32 \| V \| 21 \rangle = \frac{1}{\sqrt{2}} \langle 32 \| V \| 21 \rangle,
 \end{aligned}$$

with the reduced matrix element given by Equation 6.63. The other thirty-six matrix elements vanish due to the selection rules (Equations 6.56–6.58 and 6.62). We have determined all forty-five matrix elements and have only needed to evaluate a *single* integral. I've left the matrix elements in terms of V_+ and V_- but it's straightforward to write them in terms of x and y using the expressions on page 259.

1x1		2		2		1			
+1	+1	+2		+1	+1				
+1	0	1/2	1/2	2	1	0			
0	+1	1/2	-1/2	0	0	0			
				+1	-1	1/6	1/2	1/3	
				0	0	2/3	0	-1/3	
				-1	+1	1/6	-1/2	1/3	
						0	-1	1/2	2
						-1	0	1/2	-2
								-1	-1
									1

Figure 6.8: The Clebsch–Gordan coefficients for $1 \otimes 1$.

It is no coincidence that the Clebsch–Gordan coefficients appear in Equations 6.59–6.61. States have angular momentum, but *operators* also carry angular momentum. A scalar operator (Equation 6.34) has $\ell = 0$ —it is unchanged by a rotation—just as a state of angular momentum 0 is unchanged. A vector operator (Equation 6.33) has $\ell = 1$; its three components transform into each other under a rotation in the same way

the triplet of states with angular momentum $\ell = 1$ transform into each other.³¹ When we act on a state with an operator, we add together the angular momentum of the state and the operator to obtain the angular momentum of the resultant state; this addition of angular momenta is the source of the Clebsch–Gordan coefficients in Equations 6.59–6.61.³²

* **Problem 6.22**

- (a) Show that the commutation relations, Equations 6.50–6.54, follow from the definition of a vector operator, Equation 6.33. If you did Problem 6.19 you already derived one of these.
- (b) Derive Equation 6.57.

Problem 6.23 The Clebsch–Gordan coefficients are defined by Equation 4.183.

Adding together two states with angular momentum j_1 and j_2 produces a state with total angular momentum J according to

$$|JM\rangle = \sum_{m_1, m_2} C_{m_1 m_2 M}^{j_1 j_2 J} |j_1 j_2 m_1 m_2\rangle. \quad (6.64)$$

- (a) From Equation 6.64, show that the Clebsch–Gordan coefficients satisfy

$$C_{m_1 m_2 M}^{j_1 j_2 J} = \langle j_1 j_2 m_1 m_2 | JM \rangle. \quad (6.65)$$

- (b) Apply $\hat{J}_\pm = \hat{J}_\pm^{(1)} + \hat{J}_\pm^{(2)}$ to Equation 6.64 to derive the recursion relations for Clebsch–Gordan coefficients:

$$\begin{aligned} A_J^M C_{m_1 m_2 M+1}^{j_1 j_2 J} &= B_{j_1}^{m_1} C_{m_1-1 m_2 M}^{j_1 j_2 J} + B_{j_2}^{m_2} C_{m_1 m_2-1 M}^{j_1 j_2 J} \\ B_J^M C_{m_1 m_2 M-1}^{j_1 j_2 J} &= A_{j_1}^{m_1} C_{m_1+1 m_2 M}^{j_1 j_2 J} + A_{j_2}^{m_2} C_{m_1 m_2+1 M}^{j_1 j_2 J} \end{aligned} \quad (6.66)$$

**

Problem 6.24

- (a) Sandwich each of the six commutation relations in Equations 6.52–6.54 between $\langle n' \ell' m' |$ and $|n \ell m\rangle$ to obtain relations between matrix elements of $\hat{\mathbf{V}}$. As an example, Equation 6.52 with the upper signs gives

$$B_{\ell'}^{m'} \langle n' \ell' (m' - 1) | V_+ | n \ell m \rangle = A_\ell^m \langle n' \ell' m' | V_+ | n \ell (m + 1) \rangle.$$

- (b) Using the results in Problem 6.23, show that the six expressions you wrote down in part (a) are satisfied by Equations 6.59–6.61.

*

Problem 6.25 Express the expectation value of the dipole moment \mathbf{p}_e for an electron in the hydrogen state

$$\psi = \frac{1}{\sqrt{2}} (\psi_{211} + \psi_{210})$$

in terms of a single reduced matrix element, and evaluate the expectation value.

Note: this is the expectation value of a vector so you need to compute all three components. Don't forget Laporte's rule!

6.8 Translations in Time

In this section we study time-translation invariance. Consider a solution $\Psi(x, t)$ to the time-dependent Schrödinger equation

$$\hat{H}\Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t).$$

We can define the operator that propagates the wave function forward in time, $\hat{U}(t)$ by

$$\hat{U}(t)\Psi(x, 0) = \Psi(x, t); \quad (6.67)$$

$\hat{U}(t)$ can be expressed in terms of the Hamiltonian, and doing so is straightforward if the Hamiltonian is not itself a function of time. In that case, expanding the right-hand side of Equation 6.67 in a Taylor series gives³³

$$\hat{U}(t)\Psi(x, 0) = \Psi(x, t) = \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{\partial^n}{\partial t^n} \Psi(x, t) \right|_{t=0} t^n \quad (6.68)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar} \hat{H} t \right)^n \Psi(x, 0). \quad (6.69)$$

Therefore, in the case of a time-independent Hamiltonian, the time-evolution operator is³⁴

$$\boxed{\hat{U}(t) = \exp \left[-\frac{i t}{\hbar} \hat{H} \right].} \quad (6.71)$$

We say that the Hamiltonian is the **generator of translations in time**. Note that $\hat{U}(t)$ is a unitary operator (see Problem 6.2).

The time-evolution operator offers a compact way to state the procedure for solving the time-dependent Schrödinger equation. To see the correspondence, write out the wave function at time $t = 0$ as a superposition of stationary states ($\hat{H}\psi_n = E_n \psi_n$):

$$\Psi(x, 0) = \sum_n c_n \psi_n(x).$$

Then

$$\begin{aligned} \Psi(x, t) &= \hat{U}(t)\Psi(x, 0) = \sum_n c_n \hat{U}(t) \psi_n(x) \\ &= \sum_n c_n e^{-i \hat{H} t / \hbar} \psi_n(x) = \sum_n c_n e^{-i E_n t / \hbar} \psi_n(x). \end{aligned}$$

In this sense Equation 6.71 is shorthand for the process of expanding the initial wave function in terms of stationary states and then tacking on the “wiggle factors” to obtain the wave function at a later time (Section 2.1).

6.8.1 The Heisenberg Picture

Just as for the other transformations studied in this chapter, we can examine the effect of applying time translation to *operators*, as well as to wave functions. The transformed operators are called **Heisenberg-picture** operators and we follow the convention of giving them a subscript H rather than a prime:

$$\hat{Q}_H(t) = \hat{U}^\dagger(t) \hat{Q} \hat{U}(t). \quad (6.72)$$

Example 6.6

A particle of mass m moves in one dimension in a potential $V(x)$:

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

Find the position operator in the Heisenberg picture for an infinitesimal time translation δ .

Solution: From Equation 6.71,

$$\hat{U}(\delta) \approx 1 - i\frac{\delta}{\hbar}\hat{H}.$$

Applying Equation 6.72, we have

$$\begin{aligned} \hat{x}_H(\delta) &\approx \left(1 + i\frac{\delta}{\hbar}\hat{H}^\dagger\right)\hat{x}\left(1 - i\frac{\delta}{\hbar}\hat{H}\right) \\ &\approx \hat{x} - i\frac{\delta}{\hbar}[\hat{x}, \hat{H}] \approx \hat{x} - i\frac{\delta}{\hbar}i\hbar\frac{\hat{p}}{m} \end{aligned}$$

so

$$\hat{x}_H(\delta) \approx \hat{x}_H(0) + \frac{1}{m}\hat{p}_H(0)\delta$$

(making use of the fact that the Heisenberg-picture operators at time 0 are just the untransformed operators). This looks exactly like classical mechanics: $x(\delta) \approx x(0) + v(0)\delta$. The Heisenberg picture illuminates the connection between classical and quantum mechanics: the quantum operators obey the classical equations of motion (see Problem 6.29).

Example 6.7

A particle of mass m moves in one dimension in a harmonic-oscillator potential:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2x^2.$$

Find the position operator in the Heisenberg picture at time t .

Solution: Consider the action of \hat{x}_H on a stationary state ψ_n . (Introducing ψ_n allows us to replace the operator $e^{-i\hat{H}t/\hbar}$ with the number $e^{-iE_nt/\hbar}$, since $e^{-i\hat{H}t/\hbar}\psi_n = e^{-iE_nt/\hbar}\psi_n$.) Writing \hat{x} in terms of raising and lowering operators we have (using Equations 2.62, 2.67, and 2.70)

$$\begin{aligned}
 \hat{x}_H(t) \psi_n(x) &= \hat{U}^\dagger(t) \hat{x} \hat{U}(t) \psi_n(x) \\
 &= e^{i\hat{H}t/\hbar} \sqrt{\frac{\hbar}{2m\omega}} (\hat{a}_+ + \hat{a}_-) e^{-i\hat{H}t/\hbar} \psi_n(x) \\
 &= \sqrt{\frac{\hbar}{2m\omega}} e^{-iE_n t/\hbar} e^{i\hat{H}t/\hbar} (\hat{a}_+ + \hat{a}_-) \psi_n(x) \\
 &= \sqrt{\frac{\hbar}{2m\omega}} e^{-iE_n t/\hbar} e^{i\hat{H}t/\hbar} [\sqrt{n+1} \psi_{n+1}(x) + \sqrt{n} \psi_{n-1}(x)] \\
 &= \sqrt{\frac{\hbar}{2m\omega}} e^{-iE_n t/\hbar} [\sqrt{n+1} e^{iE_{n+1}t/\hbar} \psi_{n+1}(x) \\
 &\quad + \sqrt{n} e^{iE_{n-1}t/\hbar} \psi_{n-1}(x)] \\
 &= \sqrt{\frac{\hbar}{2m\omega}} [\sqrt{n+1} e^{i\omega t} \psi_{n+1}(x) + \sqrt{n} e^{-i\omega t} \psi_{n-1}(x)].
 \end{aligned} \tag{6.73}$$

Thus³⁵

$$\hat{x}_H(t) = \sqrt{\frac{\hbar}{2m\omega}} [e^{i\omega t} \hat{a}_+ + e^{-i\omega t} \hat{a}_-].$$

Or, using Equation 2.48 to express \hat{a}_\pm in terms of \hat{x} and \hat{p} ,

$$\hat{x}_H(t) = \hat{x}_H(0) \cos(\omega t) + \frac{1}{m\omega} \hat{p}_H(0) \sin(\omega t). \tag{6.74}$$

As in Example 6.6 we see that the Heisenberg-picture operator satisfies the *classical* equation of motion for a mass on a spring.

In this book we have been working in the **Schrödinger picture**, so-named by Dirac because it was the picture that Schrödinger himself had in mind. In the Schrödinger picture, the wave function evolves in time according to the Schrödinger equation

$$\hat{H}\Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t).$$

The operators $\hat{x} = x$ and $\hat{p} = -i\hbar\partial_x$ have no time dependence of their own, and the time dependence of expectation values (or, more generally, matrix elements) comes from the time dependence of the wave function:³⁶

$$\langle \hat{Q} \rangle = \langle \Psi(t) | \hat{Q} | \Psi(t) \rangle.$$

In the Heisenberg picture, the wave function is constant in time, $\Psi_H(x) = \Psi(x, 0)$, and the *operators* evolve in time according to Equation 6.72. In the Heisenberg picture, the time dependence of expectation values (or matrix elements) is carried by the *operators*.

$$\langle \hat{Q} \rangle = \langle \Psi_H | \hat{Q}_H(t) | \Psi_H \rangle.$$

Of course, the two pictures are entirely equivalent since

$$\langle \Psi(t) | \hat{Q} | \Psi(t) \rangle = \langle \Psi(0) | \hat{U}^\dagger \hat{Q} \hat{U} | \Psi(0) \rangle = \langle \Psi_H | \hat{Q}_H(t) | \Psi_H \rangle.$$

A nice analogy for the two pictures runs as follows. On an ordinary clock, the hands move in a clockwise direction while the numbers stay fixed. But one could equally well design a clock where the hands are stationary and the numbers move in the counter-clockwise direction. The correspondence between these two clocks is roughly the correspondence between the Schrödinger and Heisenberg pictures, the hands representing the wave function and the numbers representing the operator. Other pictures could be introduced as well, in which both the hands of the clock and the numbers on the dial move at intermediate rates such that the clock still tells the correct time.³⁷

- * **Problem 6.26** Work out $\hat{p}_H(t)$ for the system in Example 6.7 and comment on the correspondence with the classical equation of motion.

- ** **Problem 6.27** Consider a free particle of mass m . Show that the position and momentum operators in the Heisenberg picture are given by

$$\begin{aligned}\hat{x}_H(t) &= \hat{x}_H(0) + \frac{1}{m} \hat{p}_H(0) t \\ \hat{p}_H(t) &= \hat{p}_H(0).\end{aligned}$$

Comment on the relationship between these equations and the classical equations of motion. *Hint:* you will first need to evaluate the commutator $[\hat{x}, \hat{H}^n]$; this will allow you to evaluate the commutator $[\hat{x}, \hat{U}]$.

6.8.2 Time-Translation Invariance

If the Hamiltonian is time-dependent one can still write the formal solution to the Schrödinger equation in terms of the time-translation operator, \hat{U} :

$$\Psi(x, t) = \hat{U}(t, t_0) \Psi(x, t_0), \quad (6.75)$$

but \hat{U} no longer takes the simple form 6.71.38 (See Problem 11.23 for the general case.) For an infinitesimal time interval δ (see Problem 6.28)

$$\hat{U}(t_0 + \delta, t_0) \approx 1 - \frac{i}{\hbar} \hat{H}(t_0) \delta. \quad (6.76)$$

Time-translation invariance means that the time evolution is independent of which time interval we are considering. In other words

$$\hat{U}(t_1 + \delta, t_1) = \hat{U}(t_2 + \delta, t_2) \quad (6.77)$$

for any choice of t_1 and t_2 . This ensures that if the system starts in state $|\alpha\rangle$ at time t_1 and evolves for a time δ then it will end up in the same state $|\beta\rangle$ as if the system started in the same state $|\alpha\rangle$ at time t_2 and evolved for the same amount of time δ ; i.e. the experiment proceeds the same on Thursday as it did on Tuesday, assuming identical conditions. Plugging Equation 6.76 into Equation 6.77 we see that the requirement for this to be true is $\hat{H}(t_1) = \hat{H}(t_2)$, and since this must hold true for all t_1 and t_2 , it must be that the Hamiltonian is in fact time-independent after all (for time-translation invariance to hold):

$$\frac{\partial \hat{H}}{\partial t} = 0.$$

In that case the generalized Ehrenfest theorem says

$$\frac{d}{dt} \langle \hat{H} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{H}] \rangle + \left\langle \frac{\partial \hat{H}}{\partial t} \right\rangle = 0.$$

Therefore, **energy conservation is a consequence of time-translation invariance**.

We have now recovered all the classical conservation laws: conservation of momentum, energy, and angular momentum, and seen that they are each related to a continuous symmetry of the Hamiltonian (spatial translation, time translation, and rotation, respectively). And in quantum mechanics, discrete symmetries (such as parity) can also lead to conservation laws.

Problem 6.28 Show that Equations 6.75 and 6.76 are the solution to the Schrödinger equation for an infinitesimal time δ . *Hint:* expand $\Psi(x, t)$ in a Taylor series.

* **Problem 6.29** Differentiate Equation 6.72 to obtain the **Heisenberg equations of motion**

$$i\hbar \frac{d}{dt} \hat{Q}_H(t) = [\hat{Q}_H(t), \hat{H}] \quad (6.78)$$

(for \hat{Q} and \hat{H} independent of time).³⁹ Plug in $\hat{Q} = \hat{x}$ and $\hat{Q} = \hat{p}$ to obtain the differential equations for \hat{x}_H and \hat{p}_H in the Heisenberg picture for a single particle of mass m moving in a potential $V(x)$.

Problem 6.30 Consider a time-independent Hamiltonian for a particle moving in one dimension that has stationary states $\psi_n(x)$ with energies E_n .

- (a) Show that the solution to the time-dependent Schrödinger equation can be written

$$\Psi(x, t) = \hat{U}(t) \Psi(x, 0) = \int K(x, x', t) \Psi(x', 0) dx',$$

where $K(x, x', t)$, known as the **propagator**, is

$$K(x, x', t) = \sum_n \psi_n^*(x') e^{-i E_n t / \hbar} \psi_n(x). \quad (6.79)$$

Here $|K(x, x', t)|^2$ is the probability for a quantum mechanical particle to travel from position x' to position x in time t .

- (b) Find K for a particle of mass m in a simple harmonic oscillator potential of frequency ω . You will need the identity

$$\frac{1}{\sqrt{1-z^2}} \exp\left[-\frac{\xi^2 + \eta^2 - 2\xi\eta z}{1-z^2}\right] = e^{-\xi^2} e^{-\eta^2} \sum_{n=0}^{\infty} \frac{z^n}{2^n n!} H_n(\xi) H_n(\eta).$$

- (c) Find $\Psi(x, t)$ if the particle from part (a) is initially in the state⁴⁰

$$\Psi(x, 0) = \left(\frac{2a}{\pi}\right)^{1/4} e^{-a(x-x_0)^2}.$$

Compare your answer with Problem 2.49. Note: Problem 2.49 is a special case with $a = m\omega/2\hbar$.

- (d) Find K for a free particle of mass m . In this case the stationary states are continuous, not discrete, and one must make the replacement

$$\sum_n \rightarrow \int_{-\infty}^{\infty} dp$$

in Equation 6.79.

- (e) Find $\Psi(x, t)$ for a free particle that starts out in the state

$$\Psi(x, 0) = \left(\frac{2a}{\pi}\right)^{1/4} e^{-ax^2}.$$

Compare your answer with Problem 2.21.

Further Problems on Chapter 6

Problem 6.31 In deriving Equation 6.3 we assumed that our function had a Taylor series. The result holds more generally if we define the exponential of an operator by its spectral decomposition,

$$\hat{T}(a) = \int e^{-iap/\hbar} |p\rangle \langle p| dp, \quad (6.80)$$

rather than its power series. Here I've given the operator in Dirac notation; acting on a position-space function (see the discussion on page 123) this means

$$\hat{T}(a) \psi(x) = \int_{-\infty}^{\infty} e^{-iap/\hbar} f_p(x) \Phi(p) dp, \quad (6.81)$$

where $\Phi(p)$ is the momentum space wave function corresponding to $\psi(x)$ and $f_p(x)$ is defined in Equation 3.32. Show that the operator $\hat{T}(a)$, as given by Equation 6.81, applied to the function

$$\psi(x) = \sqrt{\lambda} e^{-\lambda|x|}$$

(whose first derivative is undefined at $x = 0$) gives the correct result.

- ** **Problem 6.32** Rotations on spin states are given by an expression identical to Equation 6.32, with the spin angular momentum replacing the orbital angular momentum:

$$R_{\mathbf{n}}(\varphi) = \exp \left[-i \frac{\varphi}{\hbar} \mathbf{n} \cdot \mathbf{S} \right].$$

In this problem we will consider rotations of a spin-1/2 state.

- (a) Show that

$$(\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) = \mathbf{a} \cdot \mathbf{b} + i (\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma},$$

where the σ_i are the Pauli spin matrices and \mathbf{a} and \mathbf{b} are ordinary vectors.

Use the result of Problem 4.29.

- (b) Use your result from part (a) to show that

$$\exp \left[-i \frac{\varphi}{\hbar} \mathbf{n} \cdot \mathbf{S} \right] = \cos \left(\frac{\varphi}{2} \right) - i \sin \left(\frac{\varphi}{2} \right) \mathbf{n} \cdot \boldsymbol{\sigma}.$$

Recall that $\mathbf{S} = (\hbar/2) \boldsymbol{\sigma}$.

- (c) Show that your result from part (b) becomes, in the standard basis of spin up and spin down along the z axis, the matrix

$$R_{\mathbf{n}} = \cos \left(\frac{\varphi}{2} \right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - i \sin \left(\frac{\varphi}{2} \right) \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix}$$

where θ and ϕ are the polar coordinates of the unit vector \mathbf{n} that describes the axis of rotation.

- (d) Verify that the matrix $\mathbf{R}_{\mathbf{n}}$ in part (c) is unitary.
- (e) Compute explicitly the matrix $\mathbf{S}'_x = \mathbf{R}^\dagger \mathbf{S}_x \mathbf{R}$ where \mathbf{R} is a rotation by an angle φ about the z axis and verify that it returns the expected result.
Hint: rewrite your result for \mathbf{S}'_x in terms of \mathbf{S}_x and \mathbf{S}_y .
- (f) Construct the matrix for a π rotation about the x axis and verify that it turns an up spin into a down spin.
- (g) Find the matrix describing a 2π rotation about the z axis. Why is this answer surprising?⁴¹



Problem 6.33 Consider a particle of mass m in a two-dimensional infinite square well with sides of length L . With the origin placed at the center of the well, the stationary states can be written as

$$\psi_{n_x n_y}(x, y) = \frac{2}{L} \sin\left[\frac{n_x \pi}{L}\left(x - \frac{L}{2}\right)\right] \sin\left[\frac{n_y \pi}{L}\left(y - \frac{L}{2}\right)\right],$$

with energies

$$E_{n_x n_y} = \frac{\pi^2 \hbar^2}{2m L^2} (n_x^2 + n_y^2),$$

for positive integers n_x and n_y .

- (a) The two states ψ_{ab} and ψ_{ba} for $a \neq b$ are clearly degenerate. Show that a rotation by 90° counterclockwise about the center of the square carries one into the other,

$$\hat{R} \psi_{ab} \propto \psi_{ba},$$

and determine the constant of proportionality. *Hint:* write ψ_{ab} in polar coordinates.

- (b) Suppose that instead of ψ_{ab} and ψ_{ba} we choose the basis ψ_+ and ψ_- for our two degenerate states:

$$\psi_{\pm} = \frac{\psi_{ab} \pm \psi_{ba}}{\sqrt{2}}.$$

Show that if a and b are both even or both odd, then ψ_+ and ψ_- are eigenstates of the rotation operator.

- (c) Make a contour plot of the state ψ_- for $a = 5$ and $b = 7$ and verify (visually) that it is an eigenstate of every symmetry operation of the square (rotation by an integer multiple of $\pi/2$, reflection across a diagonal, or reflection along a line bisecting two sides). The fact that ψ_+ and ψ_- are not connected to each other by any symmetry of the square means that there must be additional symmetry explaining the degeneracy of these two states.⁴²

Problem 6.34 The Coulomb potential has *more symmetry* than simply rotational invariance. This additional symmetry is manifest in an additional conserved quantity, the Laplace–Runge–Lenz vector

$$\hat{\mathbf{M}} = \frac{\hat{\mathbf{p}} \times \hat{\mathbf{L}} - \hat{\mathbf{L}} \times \hat{\mathbf{p}}}{2m} + V(r) \mathbf{r},$$

where $V(\mathbf{r})$ is the potential energy, $V(r) = -e^2/4\pi\epsilon_0 r$.⁴³ The complete set of commutators for the conserved quantities in the hydrogen atom is

- (i) $[\hat{H}, \hat{M}_i] = 0$
- (ii) $[\hat{H}, \hat{L}_i] = 0$
- (iii) $[\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k$
- (iv) $[\hat{L}_i, \hat{M}_j] = i\hbar\epsilon_{ijk}\hat{M}_k$
- (v) $[\hat{M}_i, \hat{M}_j] = \frac{\hbar}{i}\epsilon_{ijk}L_k \frac{2}{m}\hat{H}.$

The *physical* content of these equations is that (i) $\hat{\mathbf{M}}$ is a conserved quantity, (ii) $\hat{\mathbf{L}}$ is a conserved quantity, (iii) $\hat{\mathbf{L}}$ is a vector, and (iv) $\hat{\mathbf{M}}$ is a vector ((v) has no obvious interpretation). There are two additional relations between the quantities $\hat{\mathbf{L}}, \hat{\mathbf{M}}$, and \hat{H} . They are

- (vi) $\hat{M}^2 = \left(\frac{e^2}{4\pi\epsilon_0}\right)^2 + \frac{2}{m}\hat{H}(\hat{L}^2 + \hbar^2)$
- (vii) $\hat{\mathbf{M}} \cdot \hat{\mathbf{L}} = 0.$

- (a) From the result of Problem 6.19, and the fact that $\hat{\mathbf{M}}$ is a conserved quantity, we know that $\hat{M}_+ \psi_{n\ell\ell} = c_{n\ell} \psi_{n(\ell+1)(\ell+1)}$ for some constants $c_{n\ell}$. Apply (vii) to the state $\psi_{n\ell\ell}$ to show that

$$\hat{M}_z \psi_{n\ell\ell} = -\frac{1}{\sqrt{2}} \frac{1}{\sqrt{\ell+1}} c_{n\ell} \psi_{n(\ell+1)\ell}.$$

- (b) Use (vi) to show that

$$\hat{M}_- \hat{M}_+ \psi_{n\ell\ell} = \left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \left[1 - \left(\frac{\ell+1}{n}\right)^2\right] \psi_{n\ell\ell} - \hat{M}_z^2 \psi_{n\ell\ell}.$$

- (c) From your results to parts (a) and (b), obtain the constants $c_{n\ell}$. You should find that $c_{n\ell}$ is nonzero unless $\ell = n - 1$. Hint: Consider $\int |\mathbf{M}_+ \psi_{n\ell m}|^2 d^3\mathbf{r}$ and use the fact that \mathbf{M}_\pm are Hermitian conjugates. Figure 6.9 shows how the degenerate states of hydrogen are related by the generators $\hat{\mathbf{L}}$ and $\hat{\mathbf{M}}$.

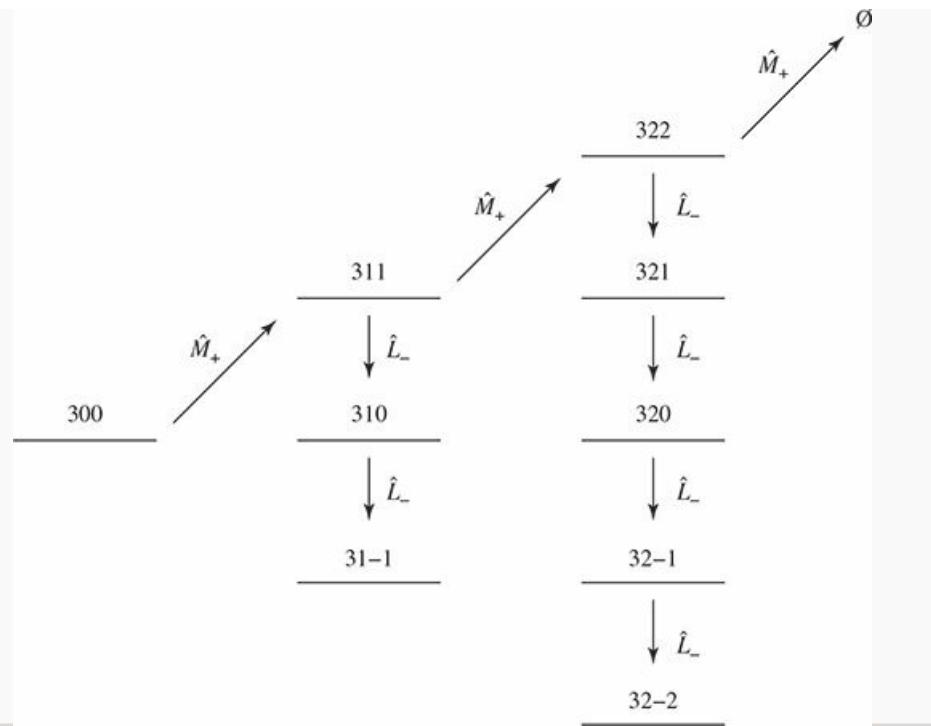


Figure 6.9: The degenerate $n = 3$ states of the hydrogen atom, and the symmetry operations that connect them.

** **Problem 6.35** A Galilean transformation performs a boost from a reference frame \mathcal{S} to a reference frame \mathcal{S}' moving with velocity $-v$ with respect to \mathcal{S} (the origins of the two frames coincide at $t = 0$). The unitary operator that carries out a Galilean transformation at time t is

$$\hat{\Gamma}(v, t) = \exp\left[-\frac{i}{\hbar}v(t\hat{p} - m\hat{x})\right].$$

- (a) Find $\hat{x}' = \hat{\Gamma}^\dagger \hat{x} \hat{\Gamma}$ and $\hat{p}' = \hat{\Gamma}^\dagger \hat{p} \hat{\Gamma}$ for an infinitesimal transformation with velocity δ . What is the physical meaning of your result?
- (b) Show that

$$\begin{aligned}\hat{\Gamma}(v, t) &= \exp\left[\frac{i}{\hbar}\left(mxv - \frac{1}{2}mv^2t\right)\right]\hat{T}(vt) \\ &= \hat{T}(vt)\exp\left[\frac{i}{\hbar}\left(mxv + \frac{1}{2}mv^2t\right)\right].\end{aligned}$$

where \hat{T} is the spatial translation operator (Equation 6.3). You will need to use the Baker–Campbell–Hausdorff formula (Problem 3.29).

- (c) Show that if Ψ is a solution to the time-dependent Schrödinger equation with Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x)$$

then the boosted wave function $\Psi' = \hat{\Gamma}(v, t)\Psi$ is a solution to the time-dependent Schrödinger equation with the potential $V(x)$ in motion:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x - vt).$$

Note: $(d/dt) e^{\hat{A}} = e^{\hat{A}}(d\hat{A}/dt)$ only if $[\hat{A}, (d\hat{A}/dt)] = 0$.

- (d) Show that the result of Problem 2.50(a) is an example of this result.

Problem 6.36 A ball thrown through the air leaves your hand at position \mathbf{r}_0 with a velocity of \mathbf{v}_0 and arrives a time t later at position \mathbf{r}_1 traveling with a velocity \mathbf{v}_1 (Figure 6.10). Suppose we could instantaneously reverse the ball's velocity when it reaches \mathbf{r}_1 . Neglecting air resistance, it would retrace the path that took it from \mathbf{r}_0 to \mathbf{r}_1 and arrive back at \mathbf{r}_0 after another time t had passed, traveling with a velocity $-\mathbf{v}_0$. This is an example of **time-reversal invariance**—reverse the motion of a particle at any point along its trajectory and it will retrace its path with an equal and opposite velocity at all positions.

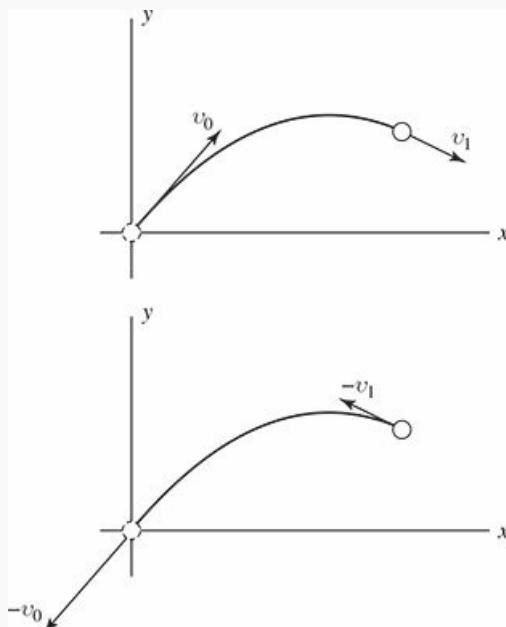


Figure 6.10: A ball thrown through the air (ignore air resistance) is an example of a system with time-reversal symmetry. If we flip the velocity of the particle at any point along its trajectory, it will retrace its path.

Why is this called time reversal? After all, it was the velocity that was reversed, *not* time. Well, if we showed you a movie of the ball traveling from \mathbf{r}_1 to \mathbf{r}_0 , there would be no way to tell if you were watching a movie of the ball after the reversal playing forward, or a movie of the ball before the reversal playing backward. In a time-reversal invariant system, playing the movie backwards represents another possible motion.

A familiar example of a system that does *not* exhibit time-reversal symmetry is a charged particle moving in an external magnetic field.⁴⁴ In that case, when you reverse the velocity of the particle, the Lorentz force will also change sign and the particle will not retrace its path; this is illustrated in Figure 6.11.

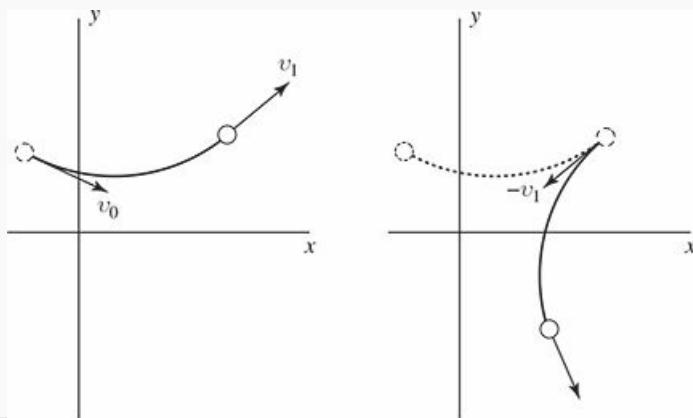


Figure 6.11: An external magnetic field breaks time-reversal symmetry. Shown is the trajectory of a particle of charge $+q$ traveling in a uniform magnetic field pointing into the page. If we flip the particle's velocity from \mathbf{v}_1 to $-\mathbf{v}_1$ at the point shown, the particle does not retrace its path, but instead moves onto a new circular orbit.

The **time-reversal operator** $\hat{\Theta}$ is the operator that reverses the momentum of the particle ($\mathbf{p} \rightarrow -\mathbf{p}$), leaving its position unchanged. A better name would really be the “reversal of the direction of motion” operator.⁴⁵ For a spinless particle, the time-reversal operator $\hat{\Theta}$ simply complex conjugates the position-space wave function⁴⁶

$$\hat{\Theta}\Psi(x, t) = \Psi^*(x, t). \quad (6.82)$$

- (a) Show that the operators \hat{x} and \hat{p} transform under time reversal as

$$\begin{aligned}\hat{x}' &= \hat{\Theta}^{-1}\hat{x}\hat{\Theta} = \hat{x} \\ \hat{p}' &= \hat{\Theta}^{-1}\hat{p}\hat{\Theta} = -\hat{p}.\end{aligned}$$

Hint: Do this by calculating the action of \hat{x}' and \hat{p}' on an arbitrary test function $f(x)$.

- (b) We can write down a mathematical statement of time-reversal invariance from our discussion above. We take a system, evolve it for a time t , reverse its momentum, and evolve it for time t again. If the system is time-reversal invariant it will be back where it started, albeit with its momentum reversed (Figure 6.10). As an operator statement this says

$$\hat{U}(t)\hat{\Theta}\hat{U}(t) = \hat{\Theta}.$$

If this is to hold for any time interval, it must hold in particular for an infinitesimal time interval δ . Show that time-reversal invariance requires

$$[\hat{\Theta}, \hat{H}] = 0. \quad (6.83)$$

- (c) Show that, for a time-reversal invariant Hamiltonian, if $\psi_n(x)$ is a stationary state with energy E_n , then $\psi_n^*(x)$ is also a stationary state with the same energy E_n . If the energy is nondegenerate, this means that the stationary state can be chosen as real.
- (d) What do you get by time-reversing a momentum eigenfunction $f_p(x)$ (Equation 3.32)? How about a hydrogen wave function $\psi_{n\ell m}(r, \theta, \phi)$? Comment on each state's relation to the untransformed state and verify that the transformed and untransformed states share the same energy, as guaranteed by (c).

Problem 6.37 As an angular momentum, a particle's spin must flip under time reversal (Problem 6.36). The action of time-reversal on a spinor (Section 4.4.1) is in fact

$$\hat{\Theta} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b^* \\ a^* \end{pmatrix} \quad (6.84)$$

so that, in addition to the complex conjugation, the up and down components are interchanged.⁴⁷

- (a) Show that $\hat{\Theta}^2 = -1$ for a spin-1/2 particle.
- (b) Consider an eigenstate $|\psi_n\rangle$ of a time-reversal invariant Hamiltonian (Equation 6.83) with energy E_n . We know that $|\psi'_n\rangle = \hat{\Theta}|\psi_n\rangle$ is also an eigenstate of \hat{H} with the same energy E_n . There two possibilities: either $|\psi'_n\rangle$ and $|\psi_n\rangle$ are the *same* state (meaning $|\psi'_n\rangle = c|\psi_n\rangle$ for some complex constant c) or they are distinct states. Show that the first case leads to a contradiction in the case of a spin-1/2 particle, meaning the energy level must be (at least) two-fold degenerate in that case.

Comment: What you have proved is a special case of **Kramer's degeneracy**: for an odd number of spin-1/2 particles (or any half-integer spin for that matter), every energy level (of a time-reversal-invariant Hamiltonian) is at least two-fold degenerate. This is because—as you just showed—for half-integer spin a state and its time-reversed state are necessarily distinct.⁴⁸

¹ A square of course has other symmetries as well, namely mirror symmetries about axes along a diagonal or bisecting two sides. The set of all transformations that leave the square unchanged is called D_4 , the “dihedral group” of degree 4.

² The parity operation in three dimensions can be realized as a mirror reflection followed by a rotation (see Problem 6.1). In two dimensions, the transformation $\psi'(x, y) = \psi(-x, -y)$ is no different from a 180° rotation. We will use the term parity exclusively for spatial inversion, $\hat{\Pi}\psi(\mathbf{r}) = \psi(-\mathbf{r})$, in one or three dimensions.

³ I'm assuming that our function has a Taylor series expansion, but the final result applies more generally. See Problem 6.31 for the details.

⁴ See Section 3.6.2 for the definition of the exponential of an operator.

- ⁵ The term comes from the study of **Lie groups** (the group of translations is an example). If you're interested, an introduction to Lie groups (written for physicists) can be found in George B. Arfken, Hans J. Weber, and Frank E. Harris, *Mathematical Methods for Physicists*, 7th edn, Academic Press, New York (2013), Section 17.7.
- ⁶ Unitary operators are discussed in Problem A.30. A unitary operator is one whose adjoint is also its inverse: $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \mathbf{I}$.
- ⁷ It is clear that Equation 6.12 satisfies Equation 6.11. In Problem 6.5 you'll prove that they are in fact equivalent statements.
- ⁸ For a delightful proof using perturbation theory, see Neil Ashcroft and N. David Mermin, *Solid State Physics*, Cengage, Belmont, 1976 (p. 765), after you have completed Problem 6.6 and studied Chapter 7.
- ⁹ For the case of continuous symmetries, it is often much easier to work with the infinitesimal form of the transformation; any finite transformation can then be built up as a product of infinitesimal transformations. In particular, the finite translation by a is a sequence of N infinitesimal translations with $\delta = a/N$ in the limit that $N \rightarrow \infty$:

$$\lim_{N \rightarrow \infty} \left(1 - i \frac{a}{N} \frac{1}{\hbar} \hat{p} \right)^N = \exp \left[- \frac{ia}{\hbar} \hat{p} \right].$$

For a proof see R. Shankar, *Basic Training in Mathematics: A Fitness Program for Science Students*, Plenum Press, New York, 1995 (p.11).

- ¹⁰ In the case of a discrete translational symmetry, momentum is not conserved, but there is a conserved quantity closely related to the discrete translational symmetry, which is the crystal momentum. For a discussion of crystal momentum see Steven H. Simon, *The Oxford Solid State Basics*, Oxford, 2013, p.84.
- ¹¹ If the spectrum of \hat{Q} is degenerate (there are distinct eigenvectors with the same eigenvalue q_n : $\hat{Q}|f_n^{(i)}\rangle = q_n|f_n^{(i)}\rangle$ for $i = 1, 2, \dots$), then we need to sum over those states:

$$P(q_n) = \sum_i \left| \langle f_n^{(i)} | \Psi(t) \rangle \right|^2.$$

Except for the sum over i the proof proceeds unchanged.

- ¹² For bound (normalizable) states in one dimension, there is no degeneracy and every bound state of a symmetric potential is automatically an eigenstate of parity. (However, see Problem 2.46.) For scattering states, degeneracy does occur.
- ¹³ Note that Equation 6.24 could equivalently be written as $[\hat{\Pi}, \hat{\mathbf{L}}] = \mathbf{0}$. The fact that parity commutes with every component of the angular momentum (and therefore also \hat{L}^2) is the reason you can find simultaneous eigenstates of \hat{L}^2 , \hat{L}_z and $\hat{\Pi}$.
- ¹⁴ However, it turns out that *antiparticles* of spin $1/2$ have opposite parity. Thus the electron is conventionally assigned parity $+1$, but the positron then has parity -1 .
- ¹⁵ To go the other way, from infinitesimal to finite, see Problem 6.14.
- ¹⁶ See Section A.5.
- ¹⁷ The Levi-Civita symbol ϵ_{ijk} is defined in Problem 4.29.
- ¹⁸ Of course, not every operator will fit into one of these categories. Scalar and vector operators are simply the first two instances in a hierarchy of tensor operators. Next come second-rank tensors (the inertia tensor from classical mechanics or the quadrupole tensor from electrodynamics are examples), third-rank tensors, and so forth.
- ¹⁹ This follows from the fact that the radial Schrödinger equation (Equation 4.35) has at most a single normalizable solution so that, once you have specified ℓ and m , the energy uniquely specifies the state. The principal quantum number n indexes those energy values that lead to normalizable solutions.
- ²⁰ When we can't identify the symmetry responsible for a particular degeneracy, we call it an **accidental degeneracy**. In most such cases, the degeneracy turns out to be no accident at all, but instead due to symmetry that is more difficult to identify than, say, rotational invariance. The canonical example is the larger symmetry group of the hydrogen atom (Problem 6.34).
- ²¹ This is highly non-classical. In classical mechanics, if you take a Keplerian orbit there will always be some axis about which you can rotate it to get a different Keplerian orbit (of the same energy) and in fact there will be an infinite number of such orbits with different orientations. In quantum mechanics, if you rotate the ground state of hydrogen you get back exactly the same state regardless of which axis you choose, and if you rotate one of the states with $n = 2$ and $\ell = 1$, you get back a linear combination of the three orthogonal states with these quantum numbers.
- ²² Of course, we already know the energies are equal since the radial equation, Equation 4.35, does not depend on m . This example demonstrates that rotational invariance is behind the degeneracy.
- ²³ I don't mean that they necessarily occur in this order. Look back at the infinite spherical well (Figure 4.3): starting with the ground state the degeneracies are $= 1, 3, 5, 1, 7, 3, 9, 5, \dots$. These are precisely the degrees of degeneracy we expect for rotational invariance ($2\ell + 1$ for integer ℓ) but the symmetry considerations don't tell us where in the spectrum each degeneracy will occur.
- ²⁴ For the three-dimensional harmonic oscillator the degeneracy is $n(n+1)/2 = 1, 3, 6, 10, \dots$ (Problem 4.46) which again is greater than $2\ell + 1$. For a discussion of the additional symmetry in the oscillator problem see D. M. Fradkin, *Am. J. Phys.* 33, 207 (1965).
- ²⁵ Importantly, they satisfy Equations 4.118 and 4.120.
- ²⁶ The other root of the quadratic $\ell'(\ell'+1) - \ell(\ell+1) = 0$ is $\ell' = -(\ell+1)$; since ℓ and ℓ' are non-negative integers this isn't possible.

²⁷ Since \hat{L}_x and \hat{L}_y are Hermitian,

$$\hat{L}_{\pm}^{\dagger} = (\hat{L}_x \pm i\hat{L}_y)^{\dagger} = \hat{L}_x^{\dagger} \pm (-i)\hat{L}_y^{\dagger} = \hat{L}_x \mp i\hat{L}_y = \hat{L}_{\mp}.$$

²⁸ The operators \hat{V}_{\pm} are, up to constants, components of what are known as **spherical tensor** operators of rank 1, written $\hat{T}_q^{(k)}$ where k is the rank and q the component of the operator:

$$\hat{T}_{\pm 1}^{(1)} = \mp \frac{1}{\sqrt{2}} \hat{V}_{\pm} \quad \hat{T}_0^{(1)} = \hat{V}_z.$$

Similarly, the scalar operator f treated in Section 6.7.1 is a rank-0 spherical tensor operator:

$$\hat{T}_0^{(0)} = \hat{f}.$$

²⁹ Equations 6.51–6.54 each stand for *two* equations: read the upper signs all the way across, or the lower signs.

³⁰ A warning about notation: In the selection rules for the scalar operator r ,

$$\langle n' \ell' m' | r | n \ell m \rangle = \delta_{\ell \ell'} \delta_{m m'} \langle n' \ell' \| r \| n \ell \rangle,$$

and for a component (say z) of the vector operator \mathbf{r} ,

$$\langle n' \ell' m' | z | n \ell m \rangle = C_{m 0 m'}^{\ell 1 \ell'} \langle n' \ell' \| r \| n \ell \rangle,$$

the two reduced matrix elements *are not the same*. One is the reduced matrix element for r and one is the reduced matrix element for \mathbf{r} , and these are *different operators* that share the same name. You could tack on a subscript ($\langle n' \ell' \| r \| n \ell \rangle_s$ and $\langle n' \ell' \| r \| n \ell \rangle_v$) to distinguish between the two if that helps keep them straight.

³¹ In the case of the position operator $\hat{\mathbf{r}}$, this correspondence is particularly evident when we rewrite the operator with the help of Table 4.3:

$$\begin{aligned} x \pm iy &= r \sin \theta e^{\pm i\phi} = \mp r \sqrt{\frac{8\pi}{3}} Y_1^{\pm 1}(\theta, \phi) \\ z &= r \cos \theta = r \sqrt{\frac{4\pi}{3}} Y_1^0(\theta, \phi). \end{aligned}$$

³² Since $C_{m 0 m'}^{\ell 0 \ell'} = \delta_{m m'} \delta_{\ell \ell'}$, one could rewrite the selection rules for a scalar operator (Equation 6.47) as

$$\langle n' \ell' m' | f | n \ell m \rangle = C_{m 0 m'}^{\ell 0 \ell'} \langle n' \ell' \| f \| n \ell \rangle.$$

³³ Why is this analysis limited to the case where \hat{H} is independent of time? Whether or not \hat{H} depends on time, Schrödinger's equation says $i\hbar\dot{\Psi} = \hat{H}\Psi$. However, if \hat{H} is time dependent then the second derivative of Ψ is given by

$$\frac{\partial^2}{\partial t^2} \Psi = \frac{\partial}{\partial t} \left(\frac{1}{i\hbar} \hat{H} \Psi \right) = \frac{1}{i\hbar} \frac{\partial \hat{H}}{\partial t} \Psi - \frac{1}{\hbar^2} \hat{H}^2 \Psi$$

and higher derivatives will be even more complicated. Therefore, Equation 6.69 only follows from Equation 6.68 when \hat{H} has no time dependence. See also Problem 11.23.

³⁴ This derivation assumes that the actual solution to Schrodinger's equation, $\Psi(x, t)$, can be expanded as a Taylor series in t , and nothing guarantees that. B. R. Holstein and A. R. Swift, *A. J. Phys.* **40**, 829 (1989) give an innocent-seeming example where such an expansion does not exist. Nonetheless, Equation 6.71 still holds in such cases as long as we define the exponential function through its spectral decomposition (Equation 3.103):

$$\hat{U}(t) \equiv \sum_n e^{-iE_n t/\hbar} |\psi_n\rangle \langle \psi_n|. \quad (6.70)$$

See also M. Amaku et al., *Am. J. Phys.* **85**, 692 (2017).

³⁵ Since Equation 6.73 holds for *any* stationary state $|\psi_n\rangle$ and since the $|\psi_n\rangle$ constitute a complete set of states, the operators must in fact be identical.

³⁶ I am assuming that \hat{Q} , like \hat{x} or \hat{p} , has no *explicit* time dependence.

³⁷ Of these other possible pictures the most important is the **interaction picture** (or Dirac picture) which is often employed in time-dependent perturbation theory.

³⁸ And is a function of *both* the initial time t_0 and the final time t , not simply the amount of time for which the wave function has evolved.

³⁹ For time-dependent \hat{Q} and \hat{H} the generalization is

$$i\hbar \frac{d}{dt} \hat{Q}_H(t) = [\hat{Q}_H(t), \hat{H}_H(t)] + \hat{U}^\dagger \frac{\partial \hat{Q}}{\partial t} \hat{U}.$$

⁴⁰ The integrals in (c)–(e) can *all* be done with the following identity:

$$\int_{-\infty}^{\infty} e^{-ax^2+bx} dx = \sqrt{\frac{\pi}{a}} e^{b^2/4a},$$

which was derived in Problem [2.21](#).

⁴¹ For a discussion of how this sign change is actually *measured*, see S. A. Werner et al., *Phys. Rev. Lett.* **35**, 1053 (1975).

⁴² See F. Leyvraz, et al., *Am. J. Phys.* **65**, 1087 (1997) for a discussion of this “accidental” degeneracy.

⁴³ The full symmetry of the Coulomb Hamiltonian is not just the obvious three-dimensional rotation group (known to mathematicians as SO(3)), but the four-dimensional rotation group (SO(4)), which has six generators (**L** and **M**). (If the four axes are w , x , y , and z , the generators correspond to rotations in each of the six orthogonal planes, wx , wy , wz (that’s **M**) and yz , zx , xy (that’s **L**)).

⁴⁴ By external magnetic field, I mean that we only reverse the velocity of our charge q , and not the velocities of the charges producing the magnetic field. If we reversed those velocities as well, the magnetic field would also switch directions, the Lorentz force on the charge q would be unchanged by the reversal, and the system would in fact be time-reversal invariant.

⁴⁵ See Eugene P. Wigner, *Group Theory and its Applications to Quantum Mechanics and Atomic Spectra* (Academic Press, New York, 1959), p. 325.

⁴⁶ Time reversal is an **anti-unitary operator**. An anti-unitary operator satisfies

$$\langle \Theta f | \Theta g \rangle = \langle f | g \rangle^*$$

$$\hat{\Theta}(a|\alpha\rangle + b|\beta\rangle) = a^* \hat{\Theta}|\alpha\rangle + b^* \hat{\Theta}|\beta\rangle$$

whereas a unitary operator satisfies the same two equations without the complex conjugates. I won’t define the adjoint of an anti-unitary operator; instead I use $\hat{\Theta}^{-1}$ for an anti-unitary operator where we might have used \hat{U}^\dagger or \hat{U}^{-1} interchangeably for a unitary operator.

⁴⁷ For arbitrary spin,

$$\hat{\Theta} = e^{-i\pi \hat{S}_y/\hbar} \hat{K} \tag{6.85}$$

where the first term is a rotation by π about the y axis and \hat{K} is the operator that performs the complex conjugation.

⁴⁸ What about in the case of a spin-0 particle—does time-reversal symmetry tell us anything interesting? Actually it does. For one thing, the stationary states can be chosen as real; you proved this back in Problem [2.2](#) but we now see that it is a consequence of time-reversal symmetry. Another example is the degeneracy of the energy levels in a periodical potential (Section [5.3.2](#) and Problem [6.6](#)) for states with crystal momentum q and $-q$. This can be ascribed to inversion symmetry if the potential is symmetric, but the degeneracy persists even when inversion symmetry is absent (try it out!); that is a result of time-reversal symmetry.

Part II Applications



7

Time-Independent Perturbation Theory

◆

7.1 Nondegenerate Perturbation Theory

7.1.1 General Formulation

Suppose we have solved the (time-independent) Schrödinger equation for some potential (say, the one-dimensional infinite square well):

$$H^0 \psi_n^0 = E_n^0 \psi_n^0, \quad (7.1)$$

obtaining a complete set of orthonormal eigenfunctions, ψ_n^0 ,

$$\langle \psi_n^0 | \psi_m^0 \rangle = \delta_{nm}, \quad (7.2)$$

and the corresponding eigenvalues E_n^0 . Now we perturb the potential slightly (say, by putting a little bump in the bottom of the well—Figure 7.1). We'd like to find the new eigenfunctions and eigenvalues:

$$H \psi_n = E_n \psi_n, \quad (7.3)$$

but unless we are very lucky, we're not going to be able to solve the Schrödinger equation exactly, for this more complicated potential. **Perturbation theory** is a systematic procedure for obtaining *approximate* solutions to the perturbed problem, by building on the known exact solutions to the *unperturbed* case.

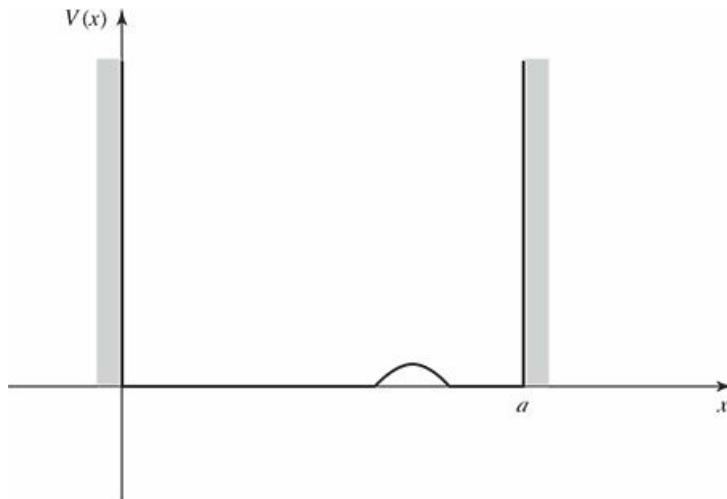


Figure 7.1: Infinite square well with small perturbation.

To begin with we write the new Hamiltonian as the sum of two terms:

$$H = H^0 + \lambda H', \quad (7.4)$$

where H' is the perturbation (the superscript 0 always identifies the *unperturbed* quantity). For the moment we'll take λ to be a small number; later we'll crank it up to 1, and H will be the true Hamiltonian. Next we write ψ_n and E_n as power series in λ :

$$\psi_n = \psi_n^0 + \lambda \psi_n^1 + \lambda^2 \psi_n^2 + \dots; \quad (7.5)$$

$$E_n = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \dots. \quad (7.6)$$

Here E_n^1 is the **first-order correction** to the n th eigenvalue, and ψ_n^1 is the first-order correction to the n th eigenfunction; E_n^2 and ψ_n^2 are the **second-order corrections**, and so on. Plugging Equations 7.5 and 7.6 into

Equation 7.3, we have:

$$\begin{aligned} & (H^0 + \lambda H') [\psi_n^0 + \lambda \psi_n^1 + \lambda^2 \psi_n^2 + \dots] \\ &= (E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \dots) [\psi_n^0 + \lambda \psi_n^1 + \lambda^2 \psi_n^2 + \dots], \end{aligned}$$

or (collecting like powers of λ):

$$\begin{aligned} & H^0 \psi_n^0 + \lambda (H^0 \psi_n^1 + H' \psi_n^0) + \lambda^2 (H^0 \psi_n^2 + H' \psi_n^1) + \dots \\ &= E_n^0 \psi_n^0 + \lambda (E_n^0 \psi_n^1 + E_n^1 \psi_n^0) + \lambda^2 (E_n^0 \psi_n^2 + E_n^1 \psi_n^1 + E_n^2 \psi_n^0) + \dots \end{aligned}$$

To lowest order¹ (λ^0) this yields $H^0 \psi_n^0 = E_n^0 \psi_n^0$, which is nothing new (Equation 7.1). To first order (λ^1),

$$H^0 \psi_n^1 + H' \psi_n^0 = E_n^0 \psi_n^1 + E_n^1 \psi_n^0. \quad (7.7)$$

To second order (λ^2),

$$H^0 \psi_n^2 + H' \psi_n^1 = E_n^0 \psi_n^2 + E_n^1 \psi_n^1 + E_n^2 \psi_n^0, \quad (7.8)$$

and so on. (I'm done with λ , now—it was just a device to keep track of the different orders—so crank it up to 1.)

7.1.2 First-Order Theory

Taking the inner product of Equation 7.7 with ψ_n^0 (that is, multiplying by $(\psi_n^0)^*$ and integrating),

$$\langle \psi_n^0 | H^0 \psi_n^1 \rangle + \langle \psi_n^0 | H' \psi_n^0 \rangle = E_n^0 \langle \psi_n^0 | \psi_n^1 \rangle + E_n^1 \langle \psi_n^0 | \psi_n^0 \rangle.$$

But H^0 is hermitian, so

$$\langle \psi_n^0 | H^0 \psi_n^1 \rangle = \langle H^0 \psi_n^0 | \psi_n^1 \rangle = \langle E_n^0 \psi_n^0 | \psi_n^1 \rangle = E_n^0 \langle \psi_n^0 | \psi_n^1 \rangle,$$

and this cancels the first term on the right. Moreover, $\langle \psi_n^0 | \psi_n^0 \rangle = 1$, so²

$E_n^1 = \langle \psi_n^0 | H' | \psi_n^0 \rangle.$

(7.9)

This is the fundamental result of first-order perturbation theory; as a *practical* matter, it may well be the most frequently used equation in quantum mechanics. It says that the first-order correction to the energy is the *expectation value* of the perturbation, in the *unperturbed* state.

Example 7.1

The unperturbed wave functions for the infinite square well are (Equation 2.31)

$$\psi_n^0(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right).$$

Suppose we perturb the system by simply raising the “floor” of the well a constant amount V_0 (Figure 7.2). Find the first-order correction to the energies.

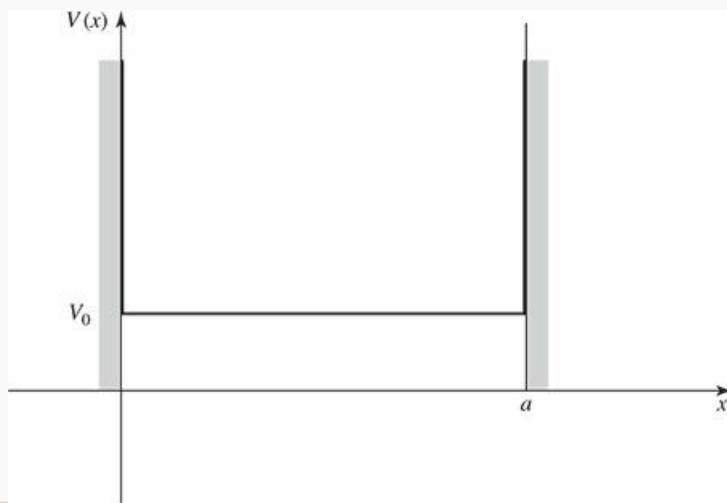


Figure 7.2: Constant perturbation over the whole well.

Solution: In this case $H' = V_0$, and the first-order correction to the energy of the n th state is

$$E_n^1 = \langle \psi_n^0 | V_0 | \psi_n^0 \rangle = V_0 \langle \psi_n^0 | \psi_n^0 \rangle = V_0.$$

The corrected energy levels, then, are $E_n \approx E_n^0 + V_0$; they are simply lifted by the amount V_0 . Of course! The only surprising thing is that in this case the first-order theory yields the *exact* answer. Evidently for a *constant* perturbation all the higher corrections vanish.³ On the other hand, if the perturbation extends only half-way across the well (Figure 7.3), then

$$E_n^1 = \frac{2V_0}{a} \int_0^{a/2} \sin^2 \left(\frac{n\pi}{a} x \right) dx = \frac{V_0}{2}.$$

In this case every energy level is lifted by $V_0/2$. That's not the *exact* result, presumably, but it does seem reasonable, as a first-order approximation.

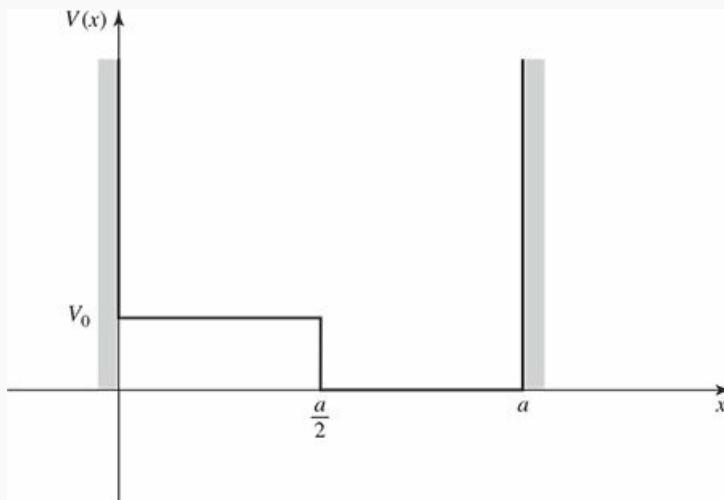


Figure 7.3: Constant perturbation over half the well.

Equation 7.9 is the first-order correction to the *energy*; to find the first-order correction to the *wave function* we rewrite Equation 7.7:

$$(H^0 - E_n^0) \psi_n^1 = - (H' - E_n^1) \psi_n^0. \quad (7.10)$$

The right side is a known function, so this amounts to an inhomogeneous differential equation for ψ_n^1 . Now, the unperturbed wave functions constitute a complete set, so ψ_n^1 (like any other function) can be expressed as a linear combination of them:

$$\psi_n^1 = \sum_{m \neq n} c_m^{(n)} \psi_m^0. \quad (7.11)$$

(There is no need to include $m = n$ in the sum, for if ψ_n^1 satisfies Equation 7.10, so too does $(\psi_n^1 + \alpha \psi_n^0)$, for any constant α , and we can use this freedom to subtract off the ψ_n^0 term.⁴) If we could determine the coefficients $c_m^{(n)}$, we'd be done.

Well, putting Equation 7.11 into Equation 7.10, and using the fact that ψ_m^0 satisfies the unperturbed Schrödinger equation (Equation 7.1), we have

$$\sum_{m \neq n} (E_m^0 - E_n^0) c_m^{(n)} \psi_m^0 = - (H' - E_n^1) \psi_n^0.$$

Taking the inner product with ψ_l^0 ,

$$\sum_{m \neq n} (E_m^0 - E_n^0) c_m^{(n)} \langle \psi_l^0 | \psi_m^0 \rangle = - \langle \psi_l^0 | H' | \psi_n^0 \rangle + E_n^1 \langle \psi_l^0 | \psi_n^0 \rangle.$$

If $l = n$, the left side is zero, and we recover Equation 7.9; if $l \neq n$, we get

$$(E_l^0 - E_n^0) c_l^{(n)} = - \langle \psi_l^0 | H' | \psi_n^0 \rangle,$$

or

$$c_l^{(n)} = \frac{\langle \psi_m^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0}, \quad (7.12)$$

so

$$\boxed{\psi_n^1 = \sum_{m \neq n} \frac{\langle \psi_m^0 | H' | \psi_n^0 \rangle}{(E_n^0 - E_m^0)} \psi_m^0.} \quad (7.13)$$

Notice that the denominator is safe (since there is no coefficient with $m = n$) *as long as the unperturbed energy spectrum is nondegenerate*. But if two different unperturbed states share the same energy, we're in serious trouble (we divided by zero to get Equation 7.12); in that case we need **degenerate perturbation theory**, which I'll come to in Section 7.2.

That completes first-order perturbation theory: The first-order correction to the energy, E_n^1 , is given by Equation 7.9, and the first-order correction to the wave function, ψ_n^1 , is given by Equation 7.13.

- * **Problem 7.1** Suppose we put a delta-function bump in the center of the infinite square well:

$$H' = \alpha \delta(x - a/2),$$

where α is a constant.

- (a) Find the first-order correction to the allowed energies. Explain why the energies are not perturbed for even n .
- (b) Find the first three nonzero terms in the expansion (Equation 7.13) of the correction to the ground state, ψ_1^1 .

- * **Problem 7.2** For the harmonic oscillator [$V(x) = (1/2)kx^2$], the allowed energies are

$$E_n = (n + 1/2)\hbar\omega, \quad (n = 0, 1, 2, \dots),$$

where $\omega = \sqrt{k/m}$ is the classical frequency. Now suppose the spring constant increases slightly: $k \rightarrow (1 + \epsilon)k$. (Perhaps we cool the spring, so it becomes less

flexible.)

- (a) Find the *exact* new energies (trivial, in this case). Expand your formula as a power series in ϵ , up to second order.
- (b) Now calculate the first-order perturbation in the energy, using Equation 7.9. What is H' here? Compare your result with part (a). *Hint:* It is not necessary—in fact, it is not *permitted*—to calculate a single integral in doing this problem.

Problem 7.3 Two identical spin-zero bosons are placed in an infinite square well (Equation 2.22). They interact weakly with one another, via the potential

$$V(x_1, x_2) = -a V_0 \delta(x_1 - x_2)$$

(where V_0 is a constant with the dimensions of energy, and a is the width of the well).

- (a) First, ignoring the interaction between the particles, find the ground state and the first excited state—both the wave functions and the associated energies.
- (b) Use first-order perturbation theory to estimate the effect of the particle-particle interaction on the energies of the ground state and the first excited state.

7.1.3 Second-Order Energies

Proceeding as before, we take the inner product of the *second*-order equation (Equation 7.8) with ψ_n^0 :

$$\langle \psi_n^0 | H^0 \psi_n^2 \rangle + \langle \psi_n^0 | H' \psi_n^1 \rangle = E_n^0 \langle \psi_n^0 | \psi_n^2 \rangle + E_n^1 \langle \psi_n^0 | \psi_n^1 \rangle + E_n^2 \langle \psi_n^0 | \psi_n^0 \rangle.$$

Again, we exploit the hermiticity of H^0 :

$$\langle \psi_n^0 | H^0 \psi_n^2 \rangle = \langle H^0 \psi_n^0 | \psi_n^2 \rangle = E_n^0 \langle \psi_n^0 | \psi_n^2 \rangle,$$

so the first term on the left cancels the first term on the right. Meanwhile, $\langle \psi_n^0 | \psi_n^0 \rangle = 1$, and we are left with a formula for E_n^2 :

$$E_n^2 = \langle \psi_n^0 | H' | \psi_n^1 \rangle - E_n^1 \langle \psi_n^0 | \psi_n^1 \rangle. \quad (7.14)$$

But

$$\langle \psi_n^0 | \psi_n^1 \rangle = \sum_{m \neq n} c_m^{(n)} \langle \psi_n^0 | \psi_m^0 \rangle = 0$$

(because the sum excludes $m = n$, and all the others are orthogonal), so

$$E_n^2 = \langle \psi_n^0 | H' | \psi_n^1 \rangle = \sum_{m \neq n} c_m^{(n)} \langle \psi_n^0 | H' | \psi_m^0 \rangle = \sum_{m \neq n} \frac{\langle \psi_m^0 | H' | \psi_n^0 \rangle \langle \psi_n^0 | H' | \psi_m^0 \rangle}{E_n^0 - E_m^0},$$

or, finally,

$$E_n^2 = \sum_{m \neq n} \frac{|\langle \psi_m^0 | H' | \psi_n^0 \rangle|^2}{E_n^0 - E_m^0}. \quad (7.15)$$

This is the fundamental result of second-order perturbation theory.

We could go on to calculate the second-order correction to the wave function (ψ_n^2), the third-order correction to the energy, and so on, but in practice Equation 7.15 is ordinarily as far as it is useful to pursue this method.⁵

**

Problem 7.4 Apply perturbation theory to the most general two-level system. The unperturbed Hamiltonian is

$$H^0 = \begin{pmatrix} E_a^0 & 0 \\ 0 & E_b^0 \end{pmatrix}$$

and the perturbation is

$$H' = \lambda \begin{pmatrix} V_{aa} & V_{ab} \\ V_{ba} & V_{bb} \end{pmatrix}$$

⁵with $V_{ab} = V_{ba}^*$ and V_{ab} real, so that H' is hermitian. As in Section 7.1.1, λ is

with $V_{ba} = V_{ab}^*$, V_{aa} and V_{bb} real, so that H is hermitian. As in Section 7.1.1, λ is a constant that will later be set to 1.

- (a) Find the exact energies for this two-level system.
- (b) Expand your result from (a) to second order in λ (and then set λ to 1). Verify that the terms in the series agree with the results from perturbation theory in Sections 7.1.2 and 7.1.3. Assume that $E_b > E_a$.
- (c) Setting $V_{aa} = V_{bb} = 0$, show that the series in (b) only converges if

$$\left| \frac{V_{ab}}{E_b^0 - E_a^0} \right| < \frac{1}{2}.$$

Comment: In general, perturbation theory is only valid if the matrix elements of the perturbation are small compared to the energy level spacings. Otherwise, the first few terms (which are all we ever calculate) will give a poor approximation to the quantity of interest and, as shown here, the series may fail to converge at all, in which case the first few terms tell us nothing.

* **Problem 7.5**

- (a) Find the second-order correction to the energies (E_n^2) for the potential in Problem 7.1. *Comment:* You can sum the series explicitly, obtaining $-2m(\alpha/\pi\hbar n)^2$ for odd n .
- (b) Calculate the second-order correction to the ground state energy (E_0^2) for the potential in Problem 7.2. Check that your result is consistent with the exact solution.

**

Problem 7.6 Consider a charged particle in the one-dimensional harmonic oscillator potential. Suppose we turn on a weak electric field (E), so that the potential energy is shifted by an amount $H' = -qEx$.

- (a) Show that there is no first-order change in the energy levels, and calculate the second-order correction. *Hint:* See Problem 3.39.
- (b) The Schrödinger equation can be solved directly in this case, by a change of variables: $x' \equiv x - (qE/m\omega^2)$. Find the exact energies, and show that they are consistent with the perturbation theory approximation.



Problem 7.7 Consider a particle in the potential shown in Figure 7.3.

- (a) Find the first-order correction to the ground-state wave function. The first three nonzero terms in the sum will suffice.
- (b) Using the method of Problem 2.61 find (numerically) the ground-state wave function and energy. Use $V_0 = 4\hbar^2/ma^2$ and $N = 100$. Compare

the energy obtained numerically to the result from first-order perturbation theory (see Example 7.1).

- (c) Make a single plot showing (i) the unperturbed ground-state wave function, (ii) the numerical ground-state wave function, and (iii) the first-order approximation to the ground-state wave function. *Note:* Make sure you've properly normalized your numerical result,

$$1 = \int |\psi(x)|^2 dx \approx \sum_{i=1}^N |\psi_i|^2 \Delta x.$$

7.2 Degenerate Perturbation Theory

If the unperturbed states are degenerate—that is, if two (or more) distinct states (ψ_a^0 and ψ_b^0) share the same energy—then ordinary perturbation theory fails: $c_a^{(b)}$ (Equation 7.12) and E_a^2 (Equation 7.15) blow up (unless, perhaps, the numerator vanishes, $\langle \psi_a^0 | H' | \psi_b^0 \rangle = 0$ —a loophole that will be important to us later on). In the degenerate case, therefore, there is no reason to trust even the *first*-order correction to the energy (Equation 7.9), and we must look for some other way to handle the problem. Note this is *not* a minor problem; almost all applications of perturbation theory involve degeneracy.

7.2.1 Two-Fold Degeneracy

Suppose that

$$H^0 \psi_a^0 = E^0 \psi_a^0, \quad H^0 \psi_b^0 = E^0 \psi_b^0, \quad \langle \psi_a^0 | \psi_b^0 \rangle = 0, \quad (7.16)$$

with ψ_a^0 and ψ_b^0 both normalized. Note that any linear combination of these states,

$$\psi^0 = \alpha \psi_a^0 + \beta \psi_b^0, \quad (7.17)$$

is still an eigenstate of H^0 , with the same eigenvalue E^0 :

$$H^0 \psi^0 = E^0 \psi^0. \quad (7.18)$$

Typically, the perturbation (H') will “break” (or “lift”) the degeneracy: As we increase λ (from 0 to 1), the common unperturbed energy E^0 splits into two (Figure 7.4). Going the other direction, when we turn *off* the perturbation, the “upper” state reduces down to *one* linear combination of ψ_a^0 and ψ_b^0 , and the “lower” state reduces to some (orthogonal) linear combination, but we don’t know *a priori* what these “good” linear combinations will be. For this reason we can’t even calculate the *first*-order energy (Equation 7.9)—we don’t know what unperturbed states to use.

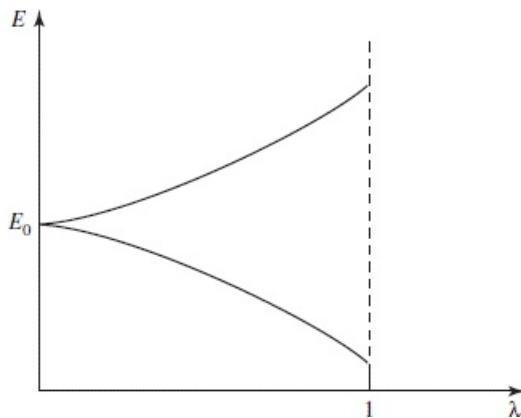


Figure 7.4: “Lifting” of a degeneracy by a perturbation.

The “good” states are *defined* as the limit of the true eigenstates as the perturbation is switched off ($\lambda \rightarrow 0$) but that isn’t how you find them in realistic situations (if you *knew* the exact eigenstates you wouldn’t need perturbation theory). Before I show you the practical techniques for calculating them, we’ll look at an example where we can take the $\lambda \rightarrow 0$ limit of the exact eigenstates.

Example 7.2

Consider a particle of mass m in a two-dimensional oscillator potential

$$H^0 = \frac{p^2}{2m} + \frac{1}{2}m\omega^2(x^2 + y^2)$$

to which is added a perturbation

$$H' = \epsilon m \omega^2 x y.$$

The unperturbed first-excited state (with $E^0 = 2\hbar\omega$) is two-fold degenerate, and *one* basis for those two degenerate states is

$$\begin{aligned}\psi_a^0 &= \psi_0(x) \psi_1(y) = \sqrt{\frac{2}{\pi}} \frac{m\omega}{\hbar} y e^{-\frac{m\omega}{2\hbar}(x^2+y^2)} \\ \psi_b^0 &= \psi_1(x) \psi_0(y) = \sqrt{\frac{2}{\pi}} \frac{m\omega}{\hbar} x e^{-\frac{m\omega}{2\hbar}(x^2+y^2)},\end{aligned}\tag{7.19}$$

where ψ_0 and ψ_1 refer to the one-dimensional harmonic oscillator states (Equation 2.86). To find the “good” linear combinations, solve for the exact eigenstates of $H = H^0 + H'$ and take their limit as $\epsilon \rightarrow 0$. *Hint:* The problem can be solved by rotating coordinates

$$x' = \frac{x+y}{\sqrt{2}} \quad y' = \frac{x-y}{\sqrt{2}}.\tag{7.20}$$

Solution: In terms of the rotated coordinates, the Hamiltonian is

$$H = \frac{1}{2m} \left(\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2} \right) + \frac{1}{2}m(1+\epsilon)\omega^2 x'^2 + \frac{1}{2}m(1-\epsilon)\omega^2 y'^2.$$

This amounts to two independent one-dimensional oscillators. The exact solutions are

$$\psi_{mn} = \psi_m^+(x') \psi_n^-(y'),$$

where ψ_m^\pm are one-dimensional oscillator states with frequencies $\omega_\pm = \sqrt{1 \pm \epsilon}\omega$ respectively. The first few exact energies,

$$E_{mn} = \left(m + \frac{1}{2}\right)\hbar\omega_+ + \left(n + \frac{1}{2}\right)\hbar\omega_-,\tag{7.21}$$

are shown in Figure 7.5.

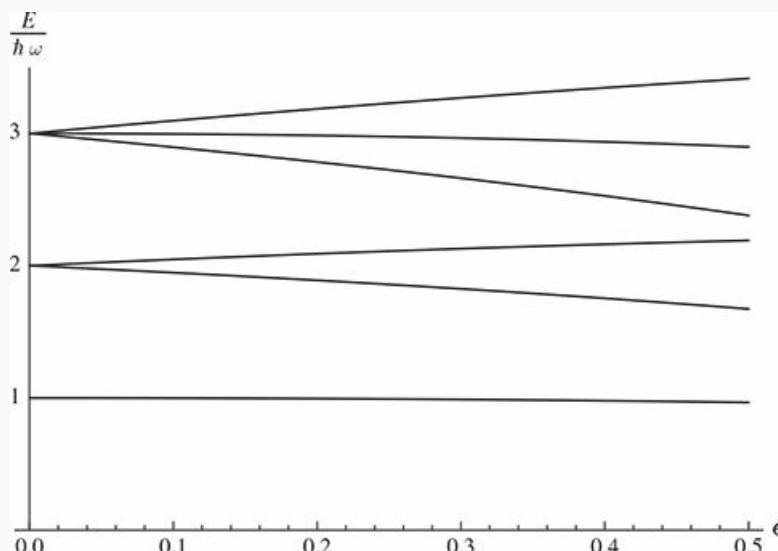


Figure 7.5: Exact energy levels as a function of ϵ for Example 7.2.

The two states which grow out of the degenerate first-excited states as ϵ is increased have $m = 0$, $n = 1$ (lower state) and $m = 1, n = 0$ (upper state). If we track these states back to $\epsilon = 0$ (in that limit $\omega_+ = \omega_- = \omega$) we get

$$\begin{aligned}\lim_{\epsilon \rightarrow 0} \psi_{01}(x) &= \lim_{\epsilon \rightarrow 0} \psi_0^+(x') \psi_1^-(y') = \psi_0\left(\frac{x+y}{\sqrt{2}}\right) \psi_1\left(\frac{x-y}{\sqrt{2}}\right) \\ &= \sqrt{\frac{2}{\pi}} \frac{m\omega}{\hbar} \frac{x-y}{\sqrt{2}} e^{-\frac{m\omega}{2\hbar}(x^2+y^2)} = \frac{-\psi_a^0 + \psi_b^0}{\sqrt{2}}. \\ \lim_{\epsilon \rightarrow 0} \psi_{10}(x) &= \frac{\psi_a^0 + \psi_b^0}{\sqrt{2}}.\end{aligned}\tag{7.22}$$

Therefore the “good” states for this problem are

$$\psi_\pm^0 \equiv \frac{1}{\sqrt{2}} (\psi_b^0 \pm \psi_a^0).\tag{7.23}$$

In this example we were able to find the exact eigenstates of H and then turn off the perturbation to see what states they evolve from. But how do we find the “good” states when we *can’t* solve the system exactly?

For the moment let’s just write the “good” unperturbed states in generic form (Equation 7.17), keeping α and β adjustable. We want to solve the Schrödinger equation,

$$H\psi = E\psi,\tag{7.24}$$

with $H = H^0 + \lambda H'$ and

$$E = E^0 + \lambda E^1 + \lambda^2 E^2 + \dots, \quad \psi = \psi^0 + \lambda \psi^1 + \lambda^2 \psi^2 + \dots\tag{7.25}$$

Plugging these into Equation 7.24, and collecting like powers of λ (as before) we find

$$H^0 \psi^0 + \lambda (H' \psi^0 + H^0 \psi^1) + \dots = E^0 \psi^0 + \lambda (E^1 \psi^0 + E^0 \psi^1) + \dots$$

But $H^0 \psi^0 = E^0 \psi^0$ (Equation 7.18), so the first terms cancel; at order λ^1 we have

$$H^0 \psi^1 + H' \psi^0 = E^0 \psi^1 + E^1 \psi^0.\tag{7.26}$$

Taking the inner product with ψ_a^0 :

$$\langle \psi_a^0 | H^0 \psi^1 \rangle + \langle \psi_a^0 | H' \psi^0 \rangle = E^0 \langle \psi_a^0 | \psi^1 \rangle + E^1 \langle \psi_a^0 | \psi^0 \rangle.$$

Because H^0 is hermitian, the first term on the left cancels the first term on the right. Putting in Equation 7.17, and exploiting the orthonormality condition (Equation 7.16), we obtain

$$\alpha \langle \psi_a^0 | H' | \psi_a^0 \rangle + \beta \langle \psi_a^0 | H' | \psi_b^0 \rangle = \alpha E^1,$$

or, more compactly,

$$\alpha W_{aa} + \beta W_{ab} = \alpha E^1, \quad (7.27)$$

where

$$W_{ij} \equiv \langle \psi_i^0 | H' | \psi_j^0 \rangle. \quad (i, j = a, b). \quad (7.28)$$

Similarly, the inner product with ψ_b^0 yields

$$\alpha W_{ba} + \beta W_{bb} = \beta E^1. \quad (7.29)$$

Notice that the W s are (in principle) *known*—they are just the “matrix elements” of H' , with respect to the unperturbed wave functions ψ_a^0 and ψ_b^0 . Written in matrix form, Equations 7.27 and 7.29 are

$$\underbrace{\begin{pmatrix} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{pmatrix}}_W \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E^1 \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \quad (7.30)$$

The eigenvalues of the matrix W give the first-order corrections to the energy (E^1) and the corresponding eigenvectors tell us the coefficients α and β that determine the “good” states.⁶

The Appendix (Section A.5) shows how to obtain the eigenvalues of a matrix; I’ll reproduce those steps here to find a general solution for E^1 . First, move all the terms in Equation 7.30 to the left-hand side.

$$\begin{pmatrix} W_{aa} - E^1 & W_{ab} \\ W_{ba} & W_{bb} - E^1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0 \quad (7.31)$$

This equation only has non-trivial solutions if the matrix on the left is non-invertible—that is to say, if its determinant vanishes:

$$\begin{vmatrix} W_{aa} - E^1 & W_{ab} \\ W_{ba} & W_{bb} - E^1 \end{vmatrix} = (W_{aa} - E^1)(W_{bb} - E^1) - |W_{ab}|^2 = 0, \quad (7.32)$$

where we used the fact that $W_{ba} = W_{ab}^*$. Solving the quadratic,

$$E_{\pm}^1 = \frac{1}{2} \left[W_{aa} + W_{bb} \pm \sqrt{(W_{aa} - W_{bb})^2 + 4|W_{ab}|^2} \right]. \quad (7.33)$$

This is the fundamental result of degenerate perturbation theory; the two roots correspond to the two perturbed energies.

Example 7.3

Returning to Example 7.2, show that diagonalizing the matrix W gives the same “good” states we found by solving the problem exactly.

Solution: We need to calculate the matrix elements of W . First,

$$\begin{aligned} W_{aa} &= \int \int \psi_a^0(x, y) H' \psi_a^0(x, y) dx dy \\ &= \epsilon m \omega^2 \int |\psi_0(x)|^2 x dx \int |\psi_0(y)|^2 y dy = 0 \end{aligned}$$

(the integrands are both odd functions). Similarly, $W_{bb} = 0$, and we need only compute

$$\begin{aligned} W_{ab} &= \int \int \psi_a^0(x, y) H' \psi_b^0(x, y) dx dy \\ &= \epsilon m \omega^2 \int \psi_0(x) x \psi_1(x) dx \int \psi_1(y) y \psi_0(y) dy. \end{aligned}$$

These two integrals are equal, and recalling (Equation 2.70)

$$x = \sqrt{\frac{\hbar}{2m\omega}} (a_+ + a_-)$$

we have

$$\begin{aligned} W_{ab} &= \epsilon m \omega^2 \left[\int \psi_0(x) \sqrt{\frac{\hbar}{2m\omega}} (a_+ + a_-) \psi_1(x) dx \right]^2 \\ &= \epsilon \frac{\hbar\omega}{2} \left[\int \psi_0(x) \psi_0(x) dx \right]^2 = \epsilon \frac{\hbar\omega}{2}. \end{aligned}$$

Therefore, the matrix W is

$$W = \epsilon \frac{\hbar\omega}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The (normalized) eigenvectors of this matrix are

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

These eigenvectors tell us which linear combination of ψ_a^0 and ψ_b^0 are the good states:

$$\psi_{\pm}^0 = \frac{1}{\sqrt{2}} (\psi_b^0 \pm \psi_a^0),$$

just as in Equation 7.23. The eigenvalues of the matrix W ,

$$E^1 = \pm \epsilon \frac{\hbar\omega}{2}, \tag{7.34}$$

give the first-order corrections to the energy (compare 7.33).

If it happens that $W_{ab} = 0$ in Equation 7.30 then the two eigenvectors are

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and the energies,

$$E_+^1 = W_{aa} = \langle \psi_a^0 | H' | \psi_a^0 \rangle, \quad E_-^1 = W_{bb} = \langle \psi_b^0 | H' | \psi_b^0 \rangle, \tag{7.35}$$

are precisely what we would have obtained using *nondegenerate* perturbation theory (Equation 7.9). We have simply been *lucky*: The states ψ_a^0 and ψ_b^0 were *already* the “good” linear combinations. Obviously, it would be greatly to our advantage if we could somehow *guess* the “good” states right from the start—then we could go ahead and use *nondegenerate* perturbation theory. As it turns out, we can very often do this by exploiting the theorem in the following section.

7.2.2 “Good” States

Theorem: Let A be a hermitian operator that commutes with H^0 and H' . If ψ_a^0 and ψ_b^0 (the degenerate eigenfunctions of H^0) are also eigenfunctions of A , with distinct eigenvalues,

$$A\psi_a^0 = \mu\psi_a^0, \quad A\psi_b^0 = \nu\psi_b^0, \quad \text{and } \mu \neq \nu,$$

then ψ_a^0 and ψ_b^0 are the “good” states to use in perturbation theory.

Proof: Since $H(\lambda) = H^0 + \lambda H'$ and A commute, there exist simultaneous eigenstates $\psi_\gamma(\lambda)$ where

$$H(\lambda)\psi_\gamma(\lambda) = E(\lambda)\psi_\gamma(\lambda) \text{ and } A\psi_\gamma(\lambda) = \gamma\psi_\gamma(\lambda). \quad (7.36)$$

The fact that A is hermitian means

$$\langle \psi_a^0 | A\psi_\gamma(\lambda) \rangle = \langle A\psi_a^0 | \psi_\gamma(\lambda) \rangle \quad (7.37)$$

$$\gamma \langle \psi_a^0 | \psi_\gamma(\lambda) \rangle = \mu^* \langle \psi_a^0 | \psi_\gamma(\lambda) \rangle$$

$$(\gamma - \mu) \langle \psi_a^0 | \psi_\gamma(\lambda) \rangle = 0, \quad (7.38)$$

(making use of the fact that μ is real). This holds true for any value of λ and taking the limit as $\lambda \rightarrow 0$ we have

$$\langle \psi_a^0 | \psi_\gamma(0) \rangle = 0 \text{ unless } \gamma = \mu,$$

and similarly

$$\langle \psi_b^0 | \psi_\gamma(0) \rangle = 0 \text{ unless } \gamma = \nu.$$

Now the good states are linear combinations of ψ_a^0 and ψ_b^0 : $\psi_\gamma(0) = \alpha\psi_a^0 + \beta\psi_b^0$. From above it follows that either $\gamma = \mu$, in which case $\beta = \langle \psi_b^0 | \psi_\gamma(0) \rangle = 0$ and the good state is simply ψ_a^0 , or $\gamma = \nu$ and the good state is ψ_b^0 . QED

Once we identify the “good” states, either by solving Equation 7.30 or by applying this theorem, we can use these “good” states as our unperturbed states and apply ordinary non-degenerate perturbation theory.⁷ In most cases, the operator A will be suggested by symmetry; as you saw in Chapter 6, symmetries are associated with operators that commute with H —precisely what are required to identify the good states.

Example 7.4

Find an operator A that satisfies the requirements of the preceding theorem to construct the “good” states in Examples 7.2 and 7.3.

Solution: The perturbation H' has less symmetry than H^0 . H^0 had continuous rotational symmetry, but $H = H^0 + H'$ is only invariant under rotations by integer multiples of π . For A , take the operator $R(\pi)$ that rotates a function counterclockwise by an angle π . Acting on our states ψ_a and ψ_b

operator $R(\pi)$ that rotates a function counterclockwise by an angle π . Acting on our states ψ_a and ψ_b we have

$$\begin{aligned} R(\pi)\psi_a^0(x, y) &= \psi_a^0(-x, -y) = -\psi_a^0(x, y), \\ R(\pi)\psi_b^0(x, y) &= \psi_b^0(-x, -y) = -\psi_b^0(x, y). \end{aligned}$$

That's no good; we need an operator with distinct eigenvalues. How about the operator that interchanges x and y ? This is a reflection about a 45° diagonal of the well. Call this operator D . D commutes with both H^0 and H' , since they are unchanged when you switch x and y . Now,

$$\begin{aligned} D\psi_a^0(x, y) &= \psi_a^0(y, x) = \psi_b^0(x, y), \\ D\psi_b^0(x, y) &= \psi_b^0(y, x) = \psi_a^0(x, y). \end{aligned}$$

So our degenerate eigenstates are not eigenstates of D . But we can construct linear combinations that are:

$$\psi_{\pm}^0 \equiv \pm\psi_a^0 + \psi_b^0. \quad (7.39)$$

Then

$$D(\pm\psi_a^0 + \psi_b^0) = \pm D\psi_a^0 + D\psi_b^0 = \pm\psi_b^0 + \psi_a^0 = \pm(\pm\psi_a^0 + \psi_b^0).$$

These are “good” states, since they are eigenstates of an operator D with distinct eigenvalues (± 1), and D commutes with both H^0 and H' .

Moral: If you’re faced with degenerate states, look around for some hermitian operator A that commutes with H^0 and H' ; pick as your unperturbed states ones that are simultaneously eigenfunctions of H^0 and A (with *distinct* eigenvalues). Then use *ordinary* first-order perturbation theory. If you can’t find such an operator, you’ll have to resort to Equation 7.33, but in practice this is seldom necessary.

Problem 7.8 Let the two “good” unperturbed states be

$$\psi_{\pm}^0 = \alpha_{\pm}\psi_a^0 + \beta_{\pm}\psi_b^0,$$

where α_{\pm} and β_{\pm} are determined (up to normalization) by Equation 7.27 (or Equation 7.29). Show explicitly that

- (a) ψ_{\pm}^0 are orthogonal ($\langle\psi_+^0|\psi_-^0\rangle = 0$);
- (b) $\langle\psi_+^0|H'|\psi_-^0\rangle = 0$;
- (c) $\langle\psi_{\pm}^0|H'|\psi_{\pm}^0\rangle = E_{\pm}^1$, with E_{\pm}^1 given by Equation 7.33.

Problem 7.9 Consider a particle of mass m that is free to move in a one-dimensional region of length L that closes on itself (for instance, a bead that slides frictionlessly on a circular wire of circumference L , as in Problem 2.46).

- (a) Show that the stationary states can be written in the form

$$\psi_n(x) = \frac{1}{\sqrt{L}} e^{2\pi i n x / L}, \quad (-L/2 < x < L/2),$$

where $n = 0, \pm 1, \pm 2, \dots$, and the allowed energies are

$$E_n = \frac{2}{m} \left(\frac{n\pi\hbar}{L} \right)^2.$$

Notice that—with the exception of the ground state ($n = 0$)—these are all doubly degenerate.

- (b) Now suppose we introduce the perturbation

$$H' = -V_0 e^{-x^2/a^2},$$

where $a \ll L$. (This puts a little “dimple” in the potential at $x = 0$, as though we bent the wire slightly to make a “trap”.) Find the first-order correction to E_n , using Equation 7.33. *Hint:* To evaluate the integrals, exploit the fact that $a \ll L$ to extend the limits from $\pm L/2$ to $\pm \infty$; after all, H' is essentially zero outside $-a < x < a$.

- (c) What are the “good” linear combinations of ψ_n and ψ_{-n} for this problem? (*Hint:* use Eq. 7.27.) Show that with these states you get the first-order correction using Equation 7.9.
- (d) Find a hermitian operator A that fits the requirements of the theorem, and show that the simultaneous eigenstates of H^0 and A are precisely the ones you used in (c).

7.2.3 Higher-Order Degeneracy

In the previous section I assumed the degeneracy was two-fold, but it is easy to see how the method generalizes. In the case of n -fold degeneracy, we look for the eigenvalues of the $n \times n$ matrix

$$W_{ij} = \langle \psi_i^0 | H' | \psi_j^0 \rangle. \quad (7.40)$$

For three-fold degeneracy (with degenerate states ψ_a^0 , ψ_b^0 , and ψ_c^0) the first-order corrections to the energies E^1 are the eigenvalues of \mathbf{W} , determined by solving

$$\begin{pmatrix} W_{aa} & W_{ab} & W_{ac} \\ W_{ba} & W_{bb} & W_{bc} \\ W_{ca} & W_{cb} & W_{cc} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = E^1 \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}, \quad (7.41)$$

and the “good” states are the corresponding eigenvectors:⁸

$$\psi^0 = \alpha \psi_a^0 + \beta \psi_b^0 + \gamma \psi_c^0. \quad (7.42)$$

Once again, if you can think of an operator A that commutes with H^0 and H' , and use the simultaneous eigenfunctions of A and H^0 , then the W matrix will automatically be diagonal, and you won’t have to fuss with calculating the off-diagonal elements of \mathbf{W} or solving the characteristic equation.⁹ (If you’re nervous about my generalization from two-fold degeneracy to n -fold degeneracy, work Problem 7.13.)

Problem 7.10 Show that the first-order energy corrections computed in Example 7.3 (Equation 7.34) agree with an expansion of the exact solution (Equation 7.21) to first order in ϵ .

Problem 7.11 Suppose we perturb the infinite cubical well (Problem 4.2) by putting a delta function “bump” at the point $(a/4, a/2, 3a/4)$:

$$H' = a^3 V_0 \delta(x - a/4) \delta(y - a/2) \delta(z - 3a/4).$$

Find the first-order corrections to the energy of the ground state and the (triply degenerate) first excited states.

* **Problem 7.12** Consider a quantum system with just three linearly independent states. Suppose the Hamiltonian, in matrix form, is

$$H = V_0 \begin{pmatrix} (1 - \epsilon) & 0 & 0 \\ 0 & 1 & \epsilon \\ 0 & \epsilon & 2 \end{pmatrix},$$

where V_0 is a constant, and ϵ is some small number ($\epsilon \ll 1$).

- (a) Write down the eigenvectors and eigenvalues of the *unperturbed* Hamiltonian ($\epsilon = 0$).

- (b) Solve for the *exact* eigenvalues of \mathbf{H} . Expand each of them as a power series in ϵ , up to second order.
- (c) Use first- and second-order *non-degenerate* perturbation theory to find the approximate eigenvalue for the state that grows out of the nondegenerate eigenvector of \mathbf{H}^0 . Compare the exact result, from (b).
- (d) Use *degenerate* perturbation theory to find the first-order correction to the two initially degenerate eigenvalues. Compare the exact results.

Problem 7.13 In the text I asserted that the first-order corrections to an n -fold degenerate energy are the eigenvalues of the W matrix, and I justified this claim as the “natural” generalization of the case $n = 2$. *Prove* it, by reproducing the steps in Section 7.2.1, starting with

$$\psi^0 = \sum_{j=1}^n \alpha_j \psi_j^0$$

(generalizing Equation 7.17), and ending by showing that the analog to Equation 7.27 can be interpreted as the eigenvalue equation for the matrix W .

7.3 The Fine Structure of Hydrogen

In our study of the hydrogen atom (Section 4.2) we took the Hamiltonian—called the Bohr Hamiltonian—to be

$$H_{\text{Bohr}} = -\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{4\pi\epsilon_0 r} \frac{1}{r} \quad (7.43)$$

(electron kinetic energy plus Coulombic potential energy). But this is not quite the whole story. We have already learned how to correct for the motion of the nucleus: Just replace m by the reduced mass (Problem 5.1). More significant is the so-called **fine structure**, which is actually due to two distinct mechanisms: a **relativistic correction**, and **spin-orbit coupling**. Compared to the Bohr energies (Equation 4.70), fine structure is a tiny perturbation—smaller by a factor of α^2 , where

$$\alpha \equiv \frac{e^2}{4\pi\epsilon_0\hbar c} \approx \frac{1}{137.036} \quad (7.44)$$

is the famous **fine structure constant**. Smaller still (by another factor of α) is the **Lamb shift**, associated with the quantization of the electric field, and smaller by yet another order of magnitude is the **hyperfine structure**, which is due to the interaction between the magnetic dipole moments of the electron and the proton. This hierarchy is summarized in Table 7.1. In the present section we will analyze the fine structure of hydrogen, as an application of time-independent perturbation theory.

Table 7.1: Hierarchy of corrections to the Bohr energies of hydrogen.

Bohr energies:	of order	$\alpha^2 mc^2$
Fine structure:	of order	$\alpha^4 mc^2$
Lamb shift:	of order	$\alpha^5 mc^2$
Hyperfine splitting:	of order	$(m/m_p)\alpha^4 mc^2$

Problem 7.14

- (a) Express the Bohr energies in terms of the fine structure constant and the rest energy (mc^2) of the electron.
- (b) Calculate the fine structure constant from first principles (i.e., without recourse to the empirical values of ϵ_0 , e , \hbar , and c). *Comment:* The fine structure constant is undoubtedly the most fundamental pure (dimensionless) number in all of physics. It relates the basic constants of electromagnetism (the charge of the electron), relativity (the speed of light), and quantum mechanics (Planck's constant). If you can solve part (b), you have the most certain Nobel Prize in history waiting for you. But I wouldn't recommend spending a lot of time on it right now; many smart people have tried, and all (so far) have failed.

7.3.1 The Relativistic Correction

The first term in the Hamiltonian is supposed to represent kinetic energy:

$$T = \frac{1}{2}mv^2 = \frac{p^2}{2m}, \quad (7.45)$$

and the canonical substitution $\mathbf{p} \rightarrow -i\hbar\nabla$ yields the operator

$$T = -\frac{\hbar^2}{2m}\nabla^2. \quad (7.46)$$

But Equation 7.45 is the *classical* expression for kinetic energy; the *relativistic* formula is

$$T = \frac{mc^2}{\sqrt{1-(v/c)^2}} - mc^2. \quad (7.47)$$

The first term is the *total* relativistic energy (not counting *potential* energy, which we aren't concerned with at the moment), and the second term is the *rest* energy—the *difference* is the energy attributable to motion.

We need to express T in terms of the (relativistic) momentum,

$$p = \frac{mv}{\sqrt{1-(v/c)^2}}, \quad (7.48)$$

instead of velocity. Notice that

$$p^2c^2 + m^2c^4 = \frac{m^2v^2c^2 + m^2c^4[1-(v/c)^2]}{1-(v/c)^2} = \frac{m^2c^4}{1-(v/c)^2} = (T + mc^2)^2,$$

so

$$T = \sqrt{p^2c^2 + m^2c^4} - mc^2. \quad (7.49)$$

This relativistic equation for kinetic energy reduces (of course) to the classical result (Equation 7.45), in the nonrelativistic limit $p \ll mc$; expanding in powers of the small number (p/mc) , we have

$$\begin{aligned} T &= mc^2 \left[\sqrt{1 + \left(\frac{p}{mc} \right)^2} - 1 \right] = mc^2 \left[1 + \frac{1}{2} \left(\frac{p}{mc} \right)^2 - \frac{1}{8} \left(\frac{p}{mc} \right)^4 \dots - 1 \right] \\ &= \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \dots \end{aligned} \quad (7.50)$$

The lowest-order¹⁰ relativistic correction to the Hamiltonian is therefore

$$H'_r = -\frac{p^4}{8m^3c^2}. \quad (7.51)$$

In first-order perturbation theory, the correction to E_n is given by the expectation value of H' in the unperturbed state (Equation 7.9):

$$E_r^1 = \langle H_r' \rangle = -\frac{1}{8m^3c^2} \langle \psi | p^4 \psi \rangle = -\frac{1}{8m^3c^2} \langle p^2 \psi | p^2 \psi \rangle. \quad (7.52)$$

Now, the Schrödinger equation (for the unperturbed states) says

$$p^2 \psi = 2m(E - V)\psi, \quad (7.53)$$

and hence¹¹

$$E_r^1 = -\frac{1}{2mc^2} \langle (E - V)^2 \rangle = -\frac{1}{2mc^2} [E^2 - 2E \langle V \rangle + \langle V^2 \rangle]. \quad (7.54)$$

So far this is entirely general; but we're interested in hydrogen, for which $V(r) = -(1/4\pi\epsilon_0)e^2/r$:

$$E_r^1 = -\frac{1}{2mc^2} \left[E_n^2 + 2E_n \left(\frac{e^2}{4\pi\epsilon_0} \right) \left\langle \frac{1}{r} \right\rangle + \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \left\langle \frac{1}{r^2} \right\rangle \right], \quad (7.55)$$

where E_n is the Bohr energy of the state in question.

To complete the job, we need the expectation values of $1/r$ and $1/r^2$, in the (unperturbed) state $\psi_{n\ell m}$ (Equation 4.89). The first is easy (see Problem 7.15):

$$\left\langle \frac{1}{r} \right\rangle = \frac{1}{n^2 a}, \quad (7.56)$$

where a is the Bohr radius (Equation 4.72). The second is not so simple to derive (see Problem 7.42), but the answer is¹²

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{1}{(\ell + 1/2)n^3 a^2}. \quad (7.57)$$

It follows that

$$E_r^1 = -\frac{1}{2mc^2} \left[E_n^2 + 2E_n \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{1}{n^2 a} + \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{1}{(\ell + 1/2)n^3 a^2} \right],$$

or, eliminating a (using Equation 4.72) and expressing everything in terms of E_n (using Equation 4.70):

$$E_r^1 = -\frac{(E_n)^2}{2mc^2} \left[\frac{4n}{\ell + 1/2} - 3 \right]. \quad (7.58)$$

Evidently the relativistic correction is smaller than E_n , by a factor of about $E_n/mc^2 = 2 \times 10^{-5}$.

You might have noticed that I used *non-degenerate* perturbation theory in this calculation (Equation 7.52), in spite of the fact that the hydrogen atom is highly degenerate. But the perturbation is spherically symmetric, so it commutes with L^2 and L_z . Moreover, the eigenfunctions of these operators (taken together) have distinct eigenvalues for the n^2 states with a given E_n . Luckily, then, the wave functions $\psi_{n\ell m}$ are the “good” states for this problem (or, as we say, n , ℓ , and m are the **good quantum numbers**), so as it happens the use of nondegenerate perturbation theory was legitimate (see the “Moral” to Section 7.2.1).

From Equation 7.58 we see that *some* of the degeneracy of the n th energy level has lifted. The $(2\ell + 1)$ -fold degeneracy in m remains; as we saw in Example 6.3 it is due to rotational symmetry, a symmetry that remains intact with this perturbation. On the other hand, the “accidental” degeneracy in ℓ has disappeared;

since its source is an additional symmetry unique to the $1/r$ potential (see Problem 6.34), we expect that degeneracy to be broken by practically *any* perturbation.

- * **Problem 7.15** Use the virial theorem (Problem 4.48) to prove Equation 7.56.

Problem 7.16 In Problem 4.52 you calculated the expectation value of r^s in the state ψ_{321} . Check your answer for the special cases $s = 0$ (trivial), $s = -1$ (Equation 7.56), $s = -2$ (Equation 7.57), and $s = -3$ (Equation 7.66). Comment on the case $s = -7$.

- ** **Problem 7.17** Find the (lowest-order) relativistic correction to the energy levels of the one-dimensional harmonic oscillator. *Hint:* Use the technique of Problem 2.12.

- *** **Problem 7.18** Show that p^2 is hermitian, for hydrogen states with $\ell = 0$. *Hint:* For such states ψ is independent of Θ and ϕ , so

$$p^2 = -\frac{\hbar^2}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right)$$

(Equation 4.13). Using integration by parts, show that

$$\langle f | p^2 g \rangle = -4\pi\hbar^2 \left(r^2 f \frac{dg}{dr} - r^2 g \frac{df}{dr} \right) \Big|_0^\infty + \langle p^2 f | g \rangle.$$

Check that the boundary term vanishes for ψ_{n00} , which goes like

$$\psi_{n00} \sim \frac{1}{\sqrt{\pi} (na)^{3/2}} \exp(-r/na)$$

near the origin.

The case of p^4 is more subtle. The Laplacian of $1/r$ picks up a delta function (see, for example, D. J. Griffiths, *Introduction to Electrodynamics*, 4th edn, Eq. 1.102). Show that

$$\nabla^4 [e^{-kr}] = \left(-\frac{4k^3}{r} + k^4 \right) e^{-kr} + 8\pi k \delta^3(\mathbf{r}),$$

and confirm that p^4 is hermitian.¹³

7.3.2 Spin-Orbit Coupling

Imagine the electron in orbit around the nucleus; from the *electron's* point of view, the proton is circling around *it* (Figure 7.6). This orbiting positive charge sets up a magnetic field \mathbf{B} , in the electron frame, which exerts a torque on the spinning electron, tending to align its magnetic moment (μ) along the direction of the field. The Hamiltonian (Equation 4.157) is

$$H = -\mu \cdot \mathbf{B}. \quad (7.59)$$

To begin with, we need to figure out the magnetic field of the proton (\mathbf{B}) and the dipole moment of the electron (μ).

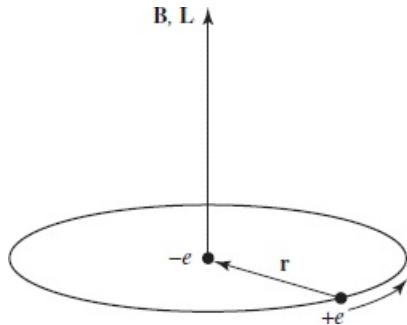


Figure 7.6: Hydrogen atom, from the electron's perspective.

The Magnetic Field of the Proton. If we picture the proton (from the electron's perspective) as a continuous current loop (Figure 7.6), its magnetic field can be calculated from the Biot–Savart law:

$$B = \frac{\mu_0 I}{2r},$$

with an effective current $I = e/T$, where e is the charge of the proton and T is the period of the orbit. On the other hand, the orbital angular momentum of the *electron* (in the rest frame of the *nucleus*) is $L = rmv = 2\pi mr^2/T$. Moreover, \mathbf{B} and \mathbf{L} point in the same direction (up, in Figure 7.6), so

$$\mathbf{B} = \frac{1}{4\pi\epsilon_0} \frac{e}{mc^2 r^3} \mathbf{L}. \quad (7.60)$$

(I used $c = 1/\sqrt{\epsilon_0\mu_0}$ to eliminate μ_0 in favor of ϵ_0 .)

The Magnetic Dipole Moment of the Electron. The magnetic dipole moment of a spinning charge is related to its (spin) angular momentum; the proportionality factor is the gyromagnetic ratio (which we already encountered in Section 4.4.2). Let's derive it, this time, using classical electrodynamics. Consider first a charge q smeared out around a ring of radius r , which rotates about the axis with period T (Figure 7.7). The magnetic dipole moment of the ring is defined as the current (q/T) times the area (πr^2):

$$\mu = \frac{q\pi r^2}{T}.$$

If the mass of the ring is m , its angular momentum is the moment of inertia (mr^2) times the angular velocity ($2\pi/T$):

$$S = \frac{2\pi mr^2}{T}.$$

The gyromagnetic ratio for this configuration is evidently $\mu/S = q/2m$. Notice that it is independent of r (and T). If I had some more complicated object, such as a sphere (all I require is that it be a figure of revolution, rotating about its axis), I could calculate μ and S by chopping it into little rings, and adding up their contributions. As long as the mass and the charge are distributed in the same manner (so that the charge-to-mass ratio is uniform), the gyromagnetic ratio will be the same for each ring, and hence also for the object as a whole. Moreover, the directions of μ and S are the same (or opposite, if the charge is negative), so

$$\mu = \left(\frac{q}{2m}\right) S. \quad (7.61)$$

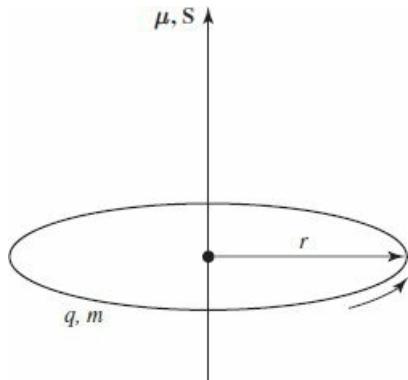


Figure 7.7: A ring of charge, rotating about its axis.

That was a purely *classical* calculation, however; as it turns out the electron's magnetic moment is *twice* the classical value:

$$\mu_e = -\frac{e}{m} \mathbf{S}. \quad (7.62)$$

The “extra” factor of 2 was explained by Dirac, in his relativistic theory of the electron.¹⁴

Putting all this together, we have

$$H = \left(\frac{e^2}{4\pi\epsilon_0}\right) \frac{1}{m^2 c^2 r^3} \mathbf{S} \cdot \mathbf{L}.$$

But there is a serious fraud in this calculation: I did the analysis in the rest frame of the electron, but that's *not an inertial system*—it *accelerates*, as the electron orbits around the nucleus. You can get away with this if you make an appropriate kinematic correction, known as the **Thomas precession**.¹⁵ In this context it throws in a factor of 1/2:¹⁶

$$H'_{\text{so}} = \left(\frac{e^2}{8\pi\epsilon_0}\right) \frac{1}{m^2 c^2 r^3} \mathbf{S} \cdot \mathbf{L}. \quad (7.63)$$

This is the **spin-orbit interaction**; apart from two corrections (the modified gyromagnetic ratio for the electron and the Thomas precession factor—which, coincidentally, exactly cancel one another) it is just what you would expect on the basis of a naive classical model. Physically, it is due to the torque exerted on the

magnetic dipole moment of the spinning electron, by the magnetic field of the proton, in the electron's instantaneous rest frame.

Now the quantum mechanics. In the presence of spin-orbit coupling, the Hamiltonian no longer commutes with \mathbf{L} and \mathbf{S} , so the spin and orbital angular momenta are not separately conserved (see Problem 7.19). However, H'_{so} does commute with L^2, S^2 and the *total* angular momentum

$$\mathbf{J} \equiv \mathbf{L} + \mathbf{S}, \quad (7.64)$$

and hence these quantities *are* conserved (Equation 3.73). To put it another way, the eigenstates of L_z and S_z are not "good" states to use in perturbation theory, but the eigenstates of L^2, S^2, J^2 , and J_z *are*. Now

$$J^2 = (\mathbf{L} + \mathbf{S}) \cdot (\mathbf{L} + \mathbf{S}) = L^2 + S^2 + 2\mathbf{L} \cdot \mathbf{S},$$

so

$$\mathbf{L} \cdot \mathbf{S} = \frac{1}{2} (J^2 - L^2 - S^2), \quad (7.65)$$

and therefore the eigenvalues of $\mathbf{L} \cdot \mathbf{S}$ are

$$\frac{\hbar^2}{2} [j(j+1) - \ell(\ell+1) - s(s+1)].$$

In this case, of course, $s = 1/2$. Meanwhile, the expectation value of $1/r^3$ (see Problem 7.43)¹⁷ is

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{1}{\ell(\ell+1/2)(\ell+1)n^3a^3}, \quad (7.66)$$

and we conclude that

$$E_{\text{so}}^1 = \langle H'_{\text{so}} \rangle = \frac{e^2}{8\pi\epsilon_0 m^2 c^2} \frac{(\hbar^2/2)[j(j+1) - \ell(\ell+1) - 3/4]}{\ell(\ell+1/2)(\ell+1)n^3a^3},$$

or, expressing it all in terms of E_n :¹⁸

$$E_{\text{so}}^1 = \frac{(E_n)^2}{mc^2} \left\{ \frac{n[j(j+1) - \ell(\ell+1) - 3/4]}{\ell(\ell+1/2)(\ell+1)} \right\}. \quad (7.67)$$

It is remarkable, considering the totally different physical mechanisms involved, that the relativistic correction and the spin-orbit coupling are of the same order (E_n^2/mc^2). Adding them together, we get the complete fine-structure formula (see Problem 7.20):

$$E_{\text{fs}}^1 = \frac{(E_n)^2}{2mc^2} \left(3 - \frac{4n}{j+1/2} \right). \quad (7.68)$$

Combining this with the Bohr formula, we obtain the grand result for the energy levels of hydrogen, including fine structure:

$$E_{nj} = -\frac{13.6 \text{ eV}}{n^2} \left[1 + \frac{\alpha^2}{n^2} \left(\frac{n}{j+1/2} - \frac{3}{4} \right) \right]. \quad (7.69)$$

Fine structure breaks the degeneracy in ℓ (that is, for a given n , the different allowed values of ℓ do not all carry the same energy), but it still preserves degeneracy in j (see Figure 7.8). The z -component eigenvalues for orbital and spin angular momentum (m_ℓ and m_s) are no longer “good” quantum numbers—the stationary states are linear combinations of states with different values of these quantities; the “good” quantum numbers are n , ℓ , s , j , and m_j .¹⁹

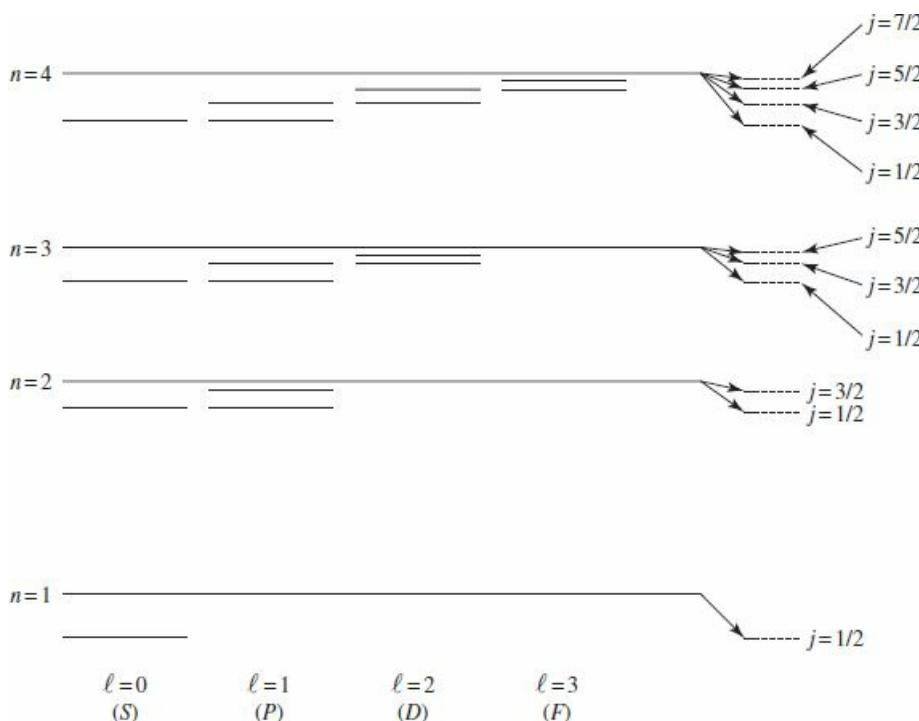


Figure 7.8: Energy levels of hydrogen, including fine structure (not to scale).

Problem 7.19 Evaluate the following commutators: (a) $[\mathbf{L} \cdot \mathbf{S}, \mathbf{L}]$, (b) $[\mathbf{L} \cdot \mathbf{S}, \mathbf{S}]$, (c) $[\mathbf{L} \cdot \mathbf{S}, \mathbf{J}]$, (d) $[\mathbf{L} \cdot \mathbf{S}, L^2]$, (e) $[\mathbf{L} \cdot \mathbf{S}, S^2]$, (f) $[\mathbf{L} \cdot \mathbf{S}, J^2]$. Hint: \mathbf{L} and \mathbf{S} satisfy the fundamental commutation relations for angular momentum (Equations 4.99 and 4.134), but they commute with each other.

* **Problem 7.20** Derive the fine structure formula (Equation 7.68) from the relativistic correction (Equation 7.58) and the spin-orbit coupling (Equation 7.67). Hint: Note that $j = \ell \pm 1/2$ (except for $\ell = 0$, where only the plus sign occurs); treat the plus sign and the minus sign separately, and you’ll find that you get the same final answer either way.

** **Problem 7.21** The most prominent feature of the hydrogen spectrum in the visible region is the red Balmer line, coming from the transition $n = 3$ to $n = 2$. First of all, determine the wavelength and frequency of this line according to the Bohr theory. Fine structure splits this line into several closely-spaced lines; the question

is: *How many, and what is their spacing?* Hint: First determine how many sublevels the $n = 2$ level splits into, and find E_{fs}^1 for each of these, in eV. Then do the same for $n = 3$. Draw an energy level diagram showing all possible transitions from $n = 3$ to $n = 2$. The energy released (in the form of a photon) is $(E_3 - E_2) + \Delta E$, the first part being common to all of them, and ΔE (due to fine structure) varying from one transition to the next. Find ΔE (in eV) for each transition. Finally, convert to photon frequency, and determine the spacing between adjacent spectral lines (in Hz)—*not* the frequency interval between each line and the *unperturbed* line (which is, of course, unobservable), but the frequency interval between each line and the *next* one. Your final answer should take the form: “The red Balmer line splits into (???) lines. In order of increasing frequency, they come from the transitions (1) $j = (???)$ to $j = (???)$, (2) $j = (???)$ to $j = (???)$, The frequency spacing between line (1) and line (2) is (???) Hz, the spacing between line (2) and line (3) is (???) Hz,”

Problem 7.22 The *exact* fine-structure formula for hydrogen (obtained from the Dirac equation without recourse to perturbation theory) is²⁰

$$E_{nj} = mc^2 \left\{ \left[1 + \left(\frac{\alpha}{n - (j + 1/2) + \sqrt{(j + 1/2)^2 - \alpha^2}} \right)^2 \right]^{-1/2} - 1 \right\}.$$

Expand to order α^4 (noting that $\alpha \ll 1$), and show that you recover Equation [7.69](#).

7.4 The Zeeman Effect

When an atom is placed in a uniform external magnetic field \mathbf{B}_{ext} , the energy levels are shifted. This phenomenon is known as the **Zeeman effect**. For a single electron, the perturbation is²¹

$$H'_Z = -(\boldsymbol{\mu}_l + \boldsymbol{\mu}_s) \cdot \mathbf{B}_{\text{ext}}, \quad (7.70)$$

where

$$\boldsymbol{\mu}_s = -\frac{e}{m} \mathbf{S} \quad (7.71)$$

(Equation 7.62) is the magnetic dipole moment associated with electron spin, and

$$\boldsymbol{\mu}_l = -\frac{e}{2m} \mathbf{L} \quad (7.72)$$

(Equation 7.61) is the dipole moment associated with orbital motion.²² Thus

$$H'_Z = \frac{e}{2m} (\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B}_{\text{ext}}. \quad (7.73)$$

The nature of the Zeeman splitting depends critically on the strength of the external field in comparison with the *internal* field (Equation 7.60) that gives rise to spin-orbit coupling. If $B_{\text{ext}} \ll B_{\text{int}}$, then fine structure dominates, and H'_Z can be treated as a small perturbation, whereas if $B_{\text{ext}} \gg B_{\text{int}}$, then the Zeeman effect dominates, and fine structure becomes the perturbation. In the intermediate zone, where the two fields are comparable, we need the full machinery of degenerate perturbation theory, and it is necessary to diagonalize the relevant portion of the Hamiltonian “by hand.” In the following sections we shall explore each of these regimes briefly, for the case of hydrogen.

Problem 7.23 Use Equation 7.60 to estimate the internal field in hydrogen, and characterize quantitatively a “strong” and “weak” Zeeman field.

7.4.1 Weak-Field Zeeman Effect

If $B_{\text{ext}} \ll B_{\text{int}}$, fine structure dominates; we treat $H_{\text{Bohr}} + H'_{fs}$ as the “unperturbed” Hamiltonian and H'_Z as the perturbation. Our “unperturbed” eigenstates are then those appropriate to fine structure: $|n\ell jm_j\rangle$ and the “unperturbed” energies are E_{nj} (Equation 7.69). Even though fine structure has lifted some of the degeneracy in the Bohr model, these states are still degenerate, since the energy does not depend on m_j or ℓ . Luckily the states $|n\ell jm_j\rangle$ are the “good” states for treating the perturbation H'_Z (meaning we don’t have to write down the W matrix for H'_Z —it’s already diagonal) since H'_Z commutes with J_z (so long as we align \mathbf{B}_{ext} with the z axis) and L^2 , and each of the degenerate states is *uniquely* labeled by the two quantum numbers m_j and ℓ .

In first-order perturbation theory, the Zeeman correction to the energy is

$$E_Z^1 = \langle n\ell jm_j | H'_Z | n\ell jm_j \rangle = \frac{e}{2m} B_{\text{ext}} \hat{k} \cdot (\mathbf{L} + 2\mathbf{S}), \quad (7.74)$$

where, as mentioned above, we align \mathbf{B}_{ext} with the z axis to eliminate the off-diagonal elements of W . Now $\mathbf{L} + 2\mathbf{S} = \mathbf{J} + \mathbf{S}$. Unfortunately, we do not immediately know the expectation value of \mathbf{S} . But we can figure it out, as follows: The total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is constant (Figure 7.9); \mathbf{L} and \mathbf{S} precess rapidly about this fixed vector. In particular, the (time) *average* value of \mathbf{S} is just its projection along \mathbf{J} :

$$\mathbf{S}_{\text{ave}} = \frac{(\mathbf{S} \cdot \mathbf{J})}{J^2} \mathbf{J}. \quad (7.75)$$

But $\mathbf{L} = \mathbf{J} - \mathbf{S}$, so $L^2 = J^2 + S^2 - 2\mathbf{J} \cdot \mathbf{S}$, and hence

$$\mathbf{S} \cdot \mathbf{J} = \frac{1}{2} (J^2 + S^2 - L^2) = \frac{\hbar^2}{2} [j(j+1) + s(s+1) - \ell(\ell+1)], \quad (7.76)$$

from which it follows that²³

$$\langle \mathbf{L} + 2\mathbf{S} \rangle = \left\langle \left(1 + \frac{\mathbf{S} \cdot \mathbf{J}}{J^2} \right) \mathbf{J} \right\rangle = \left[1 + \frac{j(j+1) - \ell(\ell+1) + s(s+1)}{2j(j+1)} \right] \langle \mathbf{J} \rangle. \quad (7.78)$$

The term in square brackets is known as the **Landé g-factor**, g_J .²⁴

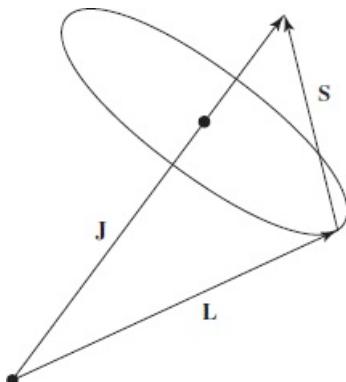


Figure 7.9: In the presence of spin-orbit coupling, \mathbf{L} and \mathbf{S} are not separately conserved; they precess about the fixed total angular momentum, \mathbf{J} .

The energy corrections are then

$$E_Z^1 = \mu_B g_J B_{\text{ext}} m_j, \quad (7.79)$$

where

$$\mu_B \equiv \frac{e\hbar}{2m} = 5.788 \times 10^{-5} \text{ eV/T} \quad (7.80)$$

is the so-called **Bohr magneton**. Recall (Example 6.3) that degeneracy in the quantum number m is a consequence of rotational invariance.²⁵ The perturbation H'_Z picks out a specific direction in space (the direction of \mathbf{B}) which breaks the rotational symmetry and lifts the degeneracy in m .

The *total* energy is the sum of the fine-structure part (Equation 7.69) and the Zeeman contribution (Equation 7.79). For example, the ground state ($n = 1, \ell = 0, j = 1/2$, and therefore $g_J = 2$) splits into two levels:

$$-13.6 \text{ eV} \left(1 + \alpha^2/4\right) \pm \mu_B B_{\text{ext}}, \quad (7.81)$$

with the plus sign for $m_j = 1/2$, and minus for $m_j = -1/2$. These energies are plotted (as functions of B_{ext}) in Figure 7.10.

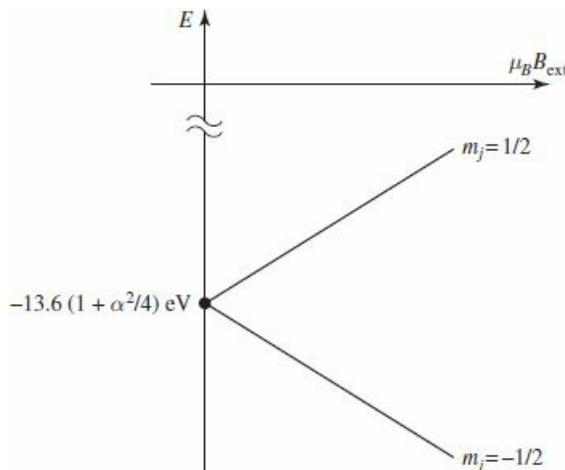


Figure 7.10: Weak-field Zeeman splitting of the ground state of hydrogen; the upper line ($m_j = 1/2$) has slope 1, the lower line ($m_j = -1/2$) has slope -1 .

- * **Problem 7.24** Consider the (eight) $n = 2$ states, $|2\ell jm_j\rangle$. Find the energy of each state, under weak-field Zeeman splitting, and construct a diagram like Figure 7.10 to show how the energies evolve as B_{ext} increases. Label each line clearly, and indicate its slope.

Problem 7.25 Use the Wigner–Eckart theorem (Equations 6.59–6.61) to prove that the matrix elements of any two vector operators, \mathbf{V} and \mathbf{W} , are proportional in a basis of angular-momentum eigenstates:

$$\langle n \ell' m' | \mathbf{V} | n \ell m \rangle = \alpha \langle n \ell' m' | \mathbf{W} | n \ell m \rangle. \quad (7.82)$$

Comment: With ℓ replaced by j (the theorem holds regardless of whether the states are eigenstates of orbital, spin, or total angular momentum), $\mathbf{V} = \mathbf{L} + 2\mathbf{S}$ and $\mathbf{W} = \mathbf{J}$, this proves Equation 7.77.

7.4.2 Strong-Field Zeeman Effect

If $B_{\text{ext}} \gg B_{\text{int}}$, the Zeeman effect dominates²⁶ and we take the “unperturbed” Hamiltonian to be $H_{\text{Bohr}} + H'_Z$ and the perturbation to be H'_{fs} . The Zeeman Hamiltonian is

$$H'_Z = \frac{e}{2m} B_{\text{ext}} (L_z + 2S_z),$$

and it is straightforward to compute the “unperturbed” energies:

$$E_{nm_l m_s} = -\frac{13.6 \text{ eV}}{n^2} + \mu_B B_{\text{ext}} (m_\ell + 2m_s). \quad (7.83)$$

The states we are using here: $|n\ell m_\ell m_s\rangle$ are degenerate, since the energy does not depend on ℓ , and there is an additional degeneracy due to the fact that, for example, $m_\ell = 3$ and $m_s = -1/2$ or $m_\ell = 1$ and $m_s = 1/2$ have the same energy. Again we are lucky; $|n\ell m_\ell m_s\rangle$ are the “good” states for treating the perturbation. The fine structure Hamiltonian H'_{fs} commutes with both L^2 and with J_z (these two operators serve as A in the theorem of Section 7.2.2); the first operator resolves the degeneracy in ℓ and the second resolves the degeneracy from coincidences in $m_\ell + 2m_s = m_j + m_s$.

In first-order perturbation theory the fine structure correction to these levels is

$$E_{\text{fs}}^1 = \langle n\ell m_\ell m_s | (H'_r + H'_{\text{so}}) | n\ell m_\ell m_s \rangle. \quad (7.84)$$

The relativistic contribution is the same as before (Equation 7.58); for the spin-orbit term (Equation 7.63) we need

$$\langle \mathbf{S} \cdot \mathbf{L} \rangle = \langle S_x \rangle \langle L_x \rangle + \langle S_y \rangle \langle L_y \rangle + \langle S_z \rangle \langle L_z \rangle = \hbar^2 m_l m_s \quad (7.85)$$

(note that $\langle S_x \rangle = \langle S_y \rangle = \langle L_x \rangle = \langle L_y \rangle = 0$ for eigenstates of S_z and L_z). Putting all this together (Problem 7.26), we conclude that

$$E_{\text{fs}}^1 = \frac{13.6 \text{ eV}}{n^3} \alpha^2 \left\{ \frac{3}{4n} - \left[\frac{\ell(\ell+1) - m_\ell m_s}{\ell(\ell+1/2)(\ell+1)} \right] \right\}. \quad (7.86)$$

(The term in square brackets is indeterminate for $\ell = 0$; its correct value in this case is 1—see Problem 7.28.) The *total* energy is the sum of the Zeeman part (Equation 7.83) and the fine structure contribution (Equation 7.86).

Problem 7.26 Starting with Equation 7.84, and using Equations 7.58, 7.63, 7.66, and 7.85, derive Equation 7.86.

**

Problem 7.27 Consider the (eight) $n = 2$ states, $|2\ell m_\ell m_s\rangle$. Find the energy of each state, under strong-field Zeeman splitting. Express each answer as the sum of three terms: the Bohr energy, the fine-structure (proportional to α^2), and the Zeeman contribution (proportional to $\mu_B B_{\text{ext}}$). If you ignore fine structure altogether, how many distinct levels are there, and what are their degeneracies?

Problem 7.28 If $\ell = 0$, then $j = s, m_j = m_s$, and the “good” states are the same ($|nm_s\rangle$) for weak *and* strong fields. Determine E_Z^1 (from Equation 7.74) and the fine structure energies (Equation 7.69), and write down the general result for the $\ell = 0$ Zeeman effect—*regardless* of the strength of the field. Show that the strong-field formula (Equation 7.86) reproduces this result, provided that we interpret the indeterminate term in square brackets as 1.

7.4.3 Intermediate-Field Zeeman Effect

In the intermediate regime neither H'_Z nor H'_{fs} dominates, and we must treat the two on an equal footing, as perturbations to the Bohr Hamiltonian (Equation 7.43):

$$H' = H'_Z + H'_{fs}. \quad (7.87)$$

I'll confine my attention here to the case $n = 2$ (you get to do $n = 3$ in Problem 7.30). It's not obvious what the "good" states are, so we'll have to resort to the full machinery of degenerate perturbation theory. I'll choose basis states characterized by ℓ, j , and m_j .²⁷ Using the Clebsch–Gordan coefficients (Problem 4.60 or Table 4.8) to express $|jm_j\rangle$ as linear combinations of $|\ell sm_l m_s\rangle$,²⁸ we have:

$\ell = 0$:

$$\psi_1 \equiv |\frac{1}{2} \frac{1}{2}\rangle = |0 \frac{1}{2} 0 \frac{1}{2}\rangle,$$

$$\psi_2 \equiv |\frac{1}{2} \frac{-1}{2}\rangle = |0 \frac{1}{2} 0 \frac{-1}{2}\rangle,$$

$\ell = 1$:

$$\psi_3 \equiv |\frac{3}{2} \frac{3}{2}\rangle = |1 \frac{1}{2} 1 \frac{1}{2}\rangle,$$

$$\psi_4 \equiv |\frac{3}{2} \frac{-3}{2}\rangle = |1 \frac{1}{2} -1 \frac{-1}{2}\rangle,$$

$$\psi_5 \equiv |\frac{3}{2} \frac{1}{2}\rangle = \sqrt{2/3} |1 \frac{1}{2} 0 \frac{1}{2}\rangle + \sqrt{1/3} |1 \frac{1}{2} 1 \frac{-1}{2}\rangle,$$

$$\psi_6 \equiv |\frac{1}{2} \frac{1}{2}\rangle = -\sqrt{1/3} |1 \frac{1}{2} 0 \frac{1}{2}\rangle + \sqrt{2/3} |1 \frac{1}{2} 1 \frac{-1}{2}\rangle,$$

$$\psi_7 \equiv |\frac{3}{2} \frac{-1}{2}\rangle = \sqrt{1/3} |1 \frac{1}{2} -1 \frac{1}{2}\rangle + \sqrt{2/3} |1 \frac{1}{2} 0 \frac{-1}{2}\rangle,$$

$$\psi_8 \equiv |\frac{1}{2} \frac{-1}{2}\rangle = -\sqrt{2/3} |1 \frac{1}{2} -1 \frac{1}{2}\rangle + \sqrt{1/3} |1 \frac{1}{2} 0 \frac{-1}{2}\rangle.$$

In this basis the nonzero matrix elements of H'_{fs} are all on the diagonal, and given by Equation 7.68; H'_Z has four off-diagonal elements, and the complete matrix – \mathbf{W} is (see Problem 7.29):

$$\begin{pmatrix} 5\gamma - \beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5\gamma + \beta & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma - 2\beta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma + 2\beta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma - \frac{2}{3}\beta & \frac{\sqrt{2}}{3}\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\sqrt{2}}{3}\beta & 5\gamma - \frac{1}{3}\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma + \frac{2}{3}\beta & \frac{\sqrt{2}}{3}\beta \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\sqrt{2}}{3}\beta & 5\gamma + \frac{1}{3}\beta \end{pmatrix}$$

where

$$\gamma \equiv (\alpha/8)^2 13.6 \text{ eV} \quad \text{and} \quad \beta \equiv \mu_B B_{\text{ext}}.$$

The first four eigenvalues are already displayed along the diagonal; it remains only to find the eigenvalues of the two 2×2 blocks. The characteristic equation for the first of these is

$$\lambda^2 - \lambda(6\gamma - \beta) + \left(5\gamma^2 - \frac{11}{3}\gamma\beta\right) = 0,$$

and the quadratic formula gives the eigenvalues:

$$\lambda_{\pm} = 3\gamma - (\beta/2) \pm \sqrt{4\gamma^2 + (2/3)\gamma\beta + (\beta^2/4)}. \quad (7.88)$$

The eigenvalues of the second block are the same, but with the sign of β reversed. The eight energies are listed in Table 7.2, and plotted against B_{ext} in Figure 7.11. In the zero-field limit ($\beta = 0$) they reduce to the fine structure values; for weak fields ($\beta \ll \gamma$) they reproduce what you got in Problem 7.24; for strong fields ($\beta \gg \gamma$) we recover the results of Problem 7.27 (note the convergence to five distinct energy levels, at very high fields, as predicted in Problem 7.27).

Table 7.2: Energy levels for the $n = 2$ states of hydrogen, with fine structure and Zeeman splitting.

ϵ_1	$= E_2 - 5\gamma + \beta$
ϵ_2	$= E_2 - 5\gamma - \beta$
ϵ_3	$= E_2 - \gamma + 2\beta$
ϵ_4	$= E_2 - \gamma - 2\beta$
ϵ_5	$= E_2 - 3\gamma + \beta/2 + \sqrt{4\gamma^2 + (2/3)\gamma\beta + \beta^2/4}$
ϵ_6	$= E_2 - 3\gamma + \beta/2 - \sqrt{4\gamma^2 + (2/3)\gamma\beta + \beta^2/4}$
ϵ_7	$= E_2 - 3\gamma - \beta/2 + \sqrt{4\gamma^2 - (2/3)\gamma\beta + \beta^2/4}$
ϵ_8	$= E_2 - 3\gamma - \beta/2 - \sqrt{4\gamma^2 - (2/3)\gamma\beta + \beta^2/4}$

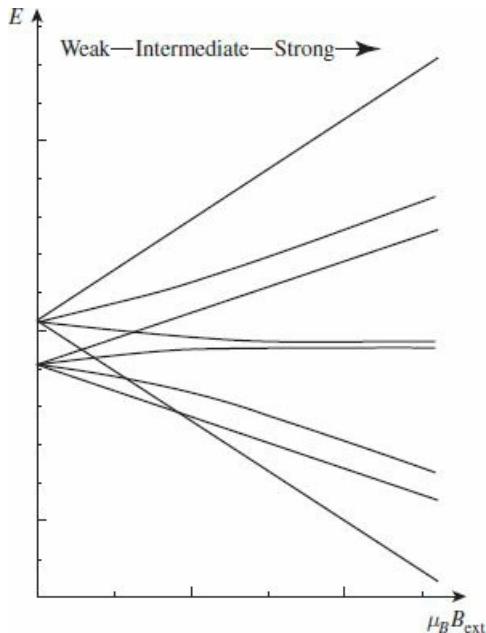


Figure 7.11: Zeeman splitting of the $n = 2$ states of hydrogen, in the weak, intermediate, and strong field regimes.

Problem 7.29 Work out the matrix elements of H'_Z and H'_{fs} , and construct the W matrix given in the text, for $n = 2$.

Problem 7.30 Analyze the Zeeman effect for the $n = 3$ states of hydrogen, in the weak, strong, and intermediate field regimes. Construct a table of energies (analogous to Table 7.2), plot them as functions of the external field (as in Figure 7.11), and check that the intermediate-field results reduce properly in the two limiting cases. *Hint:* The Wigner–Eckart theorem comes in handy here. In Chapter 6 we wrote the theorem in terms of the orbital angular momentum ℓ but it also holds for states of total angular momentum j . In particular,

$$\langle j'm'_j | V^z | jm_j \rangle = C_{m_j 0 m'_j}^{j-1 j} \langle j' \| V \| j \rangle$$

for any vector operator V (and $\mathbf{L} + 2\mathbf{S}$ is a vector operator).

7.5 Hyperfine Splitting in Hydrogen

The proton itself constitutes a magnetic dipole, though its dipole moment is much smaller than the electron's because of the mass in the denominator (Equation 7.62):

$$\boldsymbol{\mu}_p = \frac{g_p e}{2m_p} \mathbf{S}_p, \quad \boldsymbol{\mu}_e = -\frac{e}{m_e} \mathbf{S}_e. \quad (7.89)$$

(The proton is a composite structure, made up of three quarks, and its gyromagnetic ratio is not as simple as the electron's—hence the explicit g -factor (g_p), whose measured value is 5.59 as opposed to 2.00 for the electron.) According to classical electrodynamics, a dipole $\boldsymbol{\mu}$ sets up a magnetic field²⁹

$$\mathbf{B} = \frac{\mu_0}{4\pi r^3} [3(\boldsymbol{\mu} \cdot \hat{r}) \hat{r} - \boldsymbol{\mu}] + \frac{2\mu_0}{3} \boldsymbol{\mu} \delta^3(\mathbf{r}). \quad (7.90)$$

So the Hamiltonian of the electron, in the magnetic field due to the proton's magnetic dipole moment, is (Equation 7.59)

$$H'_{\text{hf}} = \frac{\mu_0 g_p e^2}{8\pi m_p m_e} \frac{[3(\mathbf{S}_p \cdot \hat{r})(\mathbf{S}_e \cdot \hat{r}) - \mathbf{S}_p \cdot \mathbf{S}_e]}{r^3} + \frac{\mu_0 g_p e^2}{3m_p m_e} \mathbf{S}_p \cdot \mathbf{S}_e \delta^3(\mathbf{r}). \quad (7.91)$$

According to perturbation theory, the first-order correction to the energy (Equation 7.9) is the expectation value of the perturbing Hamiltonian:

$$E_{\text{hf}}^1 = \frac{\mu_0 g_p e^2}{8\pi m_p m_e} \left\langle \frac{3(\mathbf{S}_p \cdot \hat{r})(\mathbf{S}_e \cdot \hat{r}) - \mathbf{S}_p \cdot \mathbf{S}_e}{r^3} \right\rangle + \frac{\mu_0 g_p e^2}{3m_p m_e} \langle \mathbf{S}_p \cdot \mathbf{S}_e \rangle |\psi(0)|^2. \quad (7.92)$$

In the ground state (or any other state for which $\ell = 0$) the wave function is spherically symmetric, and the first expectation value vanishes (see Problem 7.31). Meanwhile, from Equation 4.80 we find that $|\psi_{100}(0)|^2 = 1/(\pi a^3)$, so

$$E_{\text{hf}}^1 = \frac{\mu_0 g_p e^2}{3\pi m_p m_e a^3} \langle \mathbf{S}_p \cdot \mathbf{S}_e \rangle, \quad (7.93)$$

in the ground state. This is called **spin-spin coupling**, because it involves the dot product of two spins (contrast spin-orbit coupling, which involves $\mathbf{S} \cdot \mathbf{L}$).

In the presence of spin-spin coupling, the individual spin angular momenta are no longer conserved; the “good” states are eigenvectors of the *total* spin,

$$\mathbf{S} \equiv \mathbf{S}_e + \mathbf{S}_p. \quad (7.94)$$

As before, we square this out to get

$$\mathbf{S}_p \cdot \mathbf{S}_e = \frac{1}{2} (S^2 - S_e^2 - S_p^2). \quad (7.95)$$

But the electron and proton both have spin 1/2, so $S_e^2 = S_p^2 = (3/4) \hbar^2$. In the triplet state (spins “parallel”) the total spin is 1, and hence $S^2 = 2\hbar^2$; in the singlet state the total spin is 0, and $S^2 = 0$. Thus

$$E_{\text{hf}}^1 = \frac{4g_p\hbar^4}{3m_p m_e^2 c^2 a^4} \begin{cases} +1/4, & (\text{triplet}) \\ -3/4, & (\text{singlet}) \end{cases} \quad (7.96)$$

Spin-spin coupling breaks the spin degeneracy of the ground state, lifting the triplet configuration and depressing the singlet (see Figure 7.12). The energy gap is

$$\Delta E = \frac{4g_p\hbar^4}{3m_p m_e^2 c^2 a^4} = 5.88 \times 10^{-6} \text{ eV}. \quad (7.97)$$

The frequency of the photon emitted in a transition from the triplet to the singlet state is

$$\nu = \frac{\Delta E}{h} = 1420 \text{ MHz}, \quad (7.98)$$

and the corresponding wavelength is $c/\nu = 21 \text{ cm}$, which falls in the microwave region. This famous **21-centimeter line** is among the most pervasive forms of radiation in the universe.

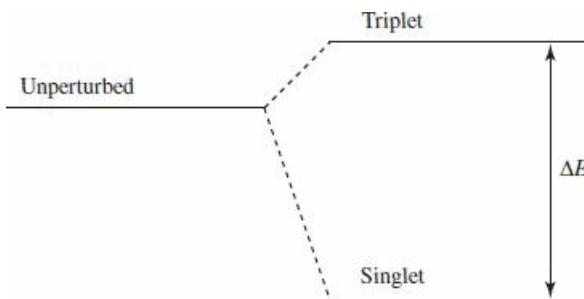


Figure 7.12: Hyperfine splitting in the ground state of hydrogen.

Problem 7.31 Let \mathbf{a} and \mathbf{b} be two constant vectors. Show that

$$\int (\mathbf{a} \cdot \hat{\mathbf{r}}) (\mathbf{b} \cdot \hat{\mathbf{r}}) \sin \theta d\theta d\phi = \frac{4\pi}{3} (\mathbf{a} \cdot \mathbf{b}) \quad (7.99)$$

(the integration is over the usual range: $0 < \theta < \pi$, $0 < \phi < 2\pi$). Use this result to demonstrate that

$$\left\langle \frac{3(\mathbf{S}_p \cdot \hat{\mathbf{r}})(\mathbf{S}_e \cdot \hat{\mathbf{r}}) - \mathbf{S}_p \cdot \mathbf{S}_e}{r^3} \right\rangle = 0,$$

for states with $\ell = 0$. Hint: $\hat{\mathbf{r}} = \sin \theta \cos \phi \hat{i} + \sin \theta \sin \phi \hat{j} + \cos \theta \hat{k}$. Do the angular integrals first.

Problem 7.32 By appropriate modification of the hydrogen formula, determine the hyperfine splitting in the ground state of (a) **muonic hydrogen** (in which a

muon—same charge and g -factor as the electron, but 207 times the mass—substitutes for the electron), (b) **positronium** (in which a positron—same mass and g -factor as the electron, but opposite charge—substitutes for the proton), and (c) **muonium** (in which an anti-muon—same mass and g -factor as a muon, but opposite charge—substitutes for the proton). *Hint:* Don't forget to use the reduced mass (Problem 5.1) in calculating the “Bohr radius” of these exotic “atoms,” but use the *actual* masses in the gyromagnetic ratios. Incidentally, the answer you get for positronium (4.82×10^{-4} eV) is quite far from the experimental value (8.41×10^{-4} eV); the large discrepancy is due to pair annihilation ($e^+ + e^- \rightarrow \gamma + \gamma$), which contributes an extra $(3/4) \Delta E$, and does not occur (of course) in ordinary hydrogen, muonic hydrogen, or muonium.³⁰

Further Problems on Chapter 7

Problem 7.33 Estimate the correction to the ground state energy of hydrogen due to the finite size of the nucleus. Treat the proton as a uniformly charged spherical shell of radius b , so the potential energy of an electron inside the shell is *constant*: $-e^2/(4\pi\epsilon_0 b)$; this isn't very realistic, but it is the simplest model, and it will give us the right order of magnitude. Expand your result in powers of the small parameter (b/a) , where a is the Bohr radius, and keep only the leading term, so your final answer takes the form

$$\frac{\Delta E}{E} = A (b/a)^n.$$

Your business is to determine the constant A and the power n . Finally, put in $b \approx 10^{-15}$ m (roughly the radius of the proton) and work out the actual number. How does it compare with fine structure and hyperfine structure?

Problem 7.34 In this problem you will develop an alternative approach to degenerate perturbation theory. Consider an unperturbed Hamiltonian H^0 with two degenerate states ψ_a^0 and ψ_b^0 (energy E^0), and a perturbation H' . Define the operator that projects³¹ onto the degenerate subspace:

$$P_D = |\psi_a^0\rangle\langle\psi_a^0| + |\psi_b^0\rangle\langle\psi_b^0|. \quad (7.100)$$

The Hamiltonian can be written

$$H = H^0 + H' = \tilde{H}^0 + \tilde{H}' \quad (7.101)$$

where

$$\tilde{H}^0 = H^0 + P_D H' P_D, \quad \tilde{H}' = H' - P_D H' P_D. \quad (7.102)$$

The idea is to treat \tilde{H}^0 as the “unperturbed” Hamiltonian and \tilde{H}' as the perturbation; as you'll soon discover, \tilde{H}^0 is nondegenerate, so we can use ordinary nondegenerate perturbation theory.

(a) First we need to find the eigenstates of \tilde{H}^0 .

- i. Show that any eigenstate ψ_n^0 (other than ψ_a^0 or ψ_b^0) of H^0 is also an eigenstate of \tilde{H}^0 with the same eigenvalue.
- ii. Show that the “good” states $\psi^0 = \alpha\psi_a^0 + \beta\psi_b^0$ (with α and β determined by solving Equation 7.30) are eigenstates of \tilde{H}^0 with energies $E^0 + E_{\pm}^1$.

(b) Assuming that E_+^1 and E_-^1 are distinct, you now have a *nondegenerate* unperturbed Hamiltonian \tilde{H}^0 and you can do nondegenerate

perturbation theory using the perturbation \tilde{H}' . Find an expression for the energy to second order for the states ψ_{\pm}^0 in (ii).

Comment: One advantage of this approach is that it also handles the case where the unperturbed energies are not *exactly* equal, but very close:³² $E_a^0 \approx E_b^0$. In this case one must still use degenerate perturbation theory; an important example of this occurs in the **nearly-free electron approximation** for calculating band structure.³³

Problem 7.35 Here is an application of the technique developed in Problem 7.34.

Consider the Hamiltonian

$$H^0 = \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon' \end{pmatrix}, \quad H' = \begin{pmatrix} 0 & a & b \\ a^* & 0 & c \\ b^* & c^* & 0 \end{pmatrix}.$$

- (a) Find the projection operator P_D (it's a 3×3 matrix) that projects onto the subspace spanned by

$$|\psi_a^0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ and } |\psi_b^0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Then construct the matrices \tilde{H}^0 and \tilde{H}' .

- (b) Solve for the eigenstates of \tilde{H}^0 and verify...
- i. that its spectrum is nondegenerate,
 - ii. that the nondegenerate eigenstate of H^0

$$|\psi_c^0\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

is also an eigenstate of \tilde{H}^0 with the same eigenvalue.

- (c) What are the “good” states, and what are their energies, to first order in the perturbation?

Problem 7.36 Consider the isotropic three-dimensional harmonic oscillator (Problem 4.46). Discuss the effect (in first order) of the perturbation

$$H' = \lambda x^2 y z$$

(for some constant λ) on

- (a) the ground state;
- (b) the (triply degenerate) first excited state. *Hint:* Use the answers to Problems 2.12 and 3.39.

*** **Problem 7.37 Van der Waals interaction.** Consider two atoms a distance R apart.

Because they are electrically neutral you might suppose there would be no force between them, but if they are polarizable there is in fact a weak

attraction. To model this system, picture each atom as an electron (mass m , charge $-e$) attached by a spring (spring constant k) to the nucleus (charge $+e$), as in Figure 7.13. We'll assume the nuclei are heavy, and essentially motionless. The Hamiltonian for the unperturbed system is

$$H^0 = \frac{1}{2m}p_1^2 + \frac{1}{2}kx_1^2 + \frac{1}{2m}p_2^2 + \frac{1}{2}kx_2^2. \quad (7.103)$$

The Coulomb interaction between the atoms is

$$H' = \frac{1}{4\pi\epsilon_0} \left(\frac{e^2}{R} - \frac{e^2}{R-x_1} - \frac{e^2}{R+x_2} + \frac{e^2}{R-x_1+x_2} \right). \quad (7.104)$$

- (a) Explain Equation 7.104. Assuming that $|x_1|$ and $|x_2|$ are both much less than R , show that

$$H' \cong -\frac{e^2 x_1 x_2}{2\pi\epsilon_0 R^3}. \quad (7.105)$$

- (b) Show that the total Hamiltonian (H^0 plus Equation 7.105) separates into two harmonic oscillator Hamiltonians:

$$H = \left[\frac{1}{2m}p_+^2 + \frac{1}{2} \left(k - \frac{e^2}{2\pi\epsilon_0 R^3} \right) x_+^2 \right] + \left[\frac{1}{2m}p_-^2 + \frac{1}{2} \left(k + \frac{e^2}{2\pi\epsilon_0 R^3} \right) x_-^2 \right], \quad (7.106)$$

under the change variables

$$x_{\pm} \equiv \frac{1}{\sqrt{2}}(x_1 \pm x_2), \quad \text{which entails} \quad p_{\pm} = \frac{1}{\sqrt{2}}(p_1 \pm p_2). \quad (7.107)$$

- (c) The ground state energy for this Hamiltonian is evidently

$$E = \frac{1}{2}\hbar(\omega_+ + \omega_-), \quad \text{where} \quad \omega_{\pm} = \sqrt{\frac{k \mp (e^2/2\pi\epsilon_0 R^3)}{m}}. \quad (7.108)$$

Without the Coulomb interaction it would have been $E_0 = \hbar\omega_0$, where $\omega_0 = \sqrt{k/m}$. Assuming that $k \gg (e^2/2\pi\epsilon_0 R^3)$, show that

$$\Delta V \equiv E - E_0 \cong -\frac{\hbar}{8m^2\omega_0^3} \left(\frac{e^2}{2\pi\epsilon_0} \right)^2 \frac{1}{R^6}. \quad (7.109)$$

Conclusion: There is an attractive potential between the atoms, proportional to the inverse sixth power of their separation. This is the **van der Waals interaction** between two neutral atoms.

- (d) Now do the same calculation using second-order perturbation theory.

Hint: The unperturbed states are of the form $\psi_{n_1}(x_1)\psi_{n_2}(x_2)$, where $\psi_n(x)$ is a one-particle oscillator wave function with mass m and spring constant k ; ΔV is the second-order correction to the ground state energy,

for the perturbation in Equation 7.105 (notice that the *first*-order correction is zero).³⁴



Figure 7.13: Two nearby polarizable atoms (Problem 7.37).

- ** Problem 7.38 Suppose the Hamiltonian H , for a particular quantum system, is a function of some parameter λ ; let $E_n(\lambda)$ and $\psi_n(\lambda)$ be the eigenvalues and eigenfunctions of $H(\lambda)$. The Feynman–Hellmann theorem³⁵ states that

$$\frac{\partial E_n}{\partial \lambda} = \left\langle \psi_n \left| \frac{\partial H}{\partial \lambda} \right| \psi_n \right\rangle \quad (7.110)$$

(assuming either that E_n is nondegenerate, or—if degenerate—that the ψ_n s are the “good” linear combinations of the degenerate eigenfunctions).

- (a) Prove the Feynman–Hellmann theorem. *Hint:* Use Equation 7.9.
- (b) Apply it to the one-dimensional harmonic oscillator, (i) using $\lambda = \omega$ (this yields a formula for the expectation value of V), (ii) using $\lambda = \hbar$ (this yields $\langle T \rangle$), and (iii) using $\lambda = m$ (this yields a relation between $\langle T \rangle$ and $\langle V \rangle$). Compare your answers to Problem 2.12, and the virial theorem predictions (Problem 3.37).

Problem 7.39 Consider a three-level system with the unperturbed Hamiltonian

$$H^0 = \begin{pmatrix} \epsilon_a & 0 & 0 \\ 0 & \epsilon_a & 0 \\ 0 & 0 & \epsilon_c \end{pmatrix} \quad (7.111)$$

($\epsilon_a > \epsilon_c$) and the perturbation

$$H' = \begin{pmatrix} 0 & 0 & V \\ 0 & 0 & V \\ V^* & V^* & 0 \end{pmatrix}. \quad (7.112)$$

Since the (2×2) matrix W is diagonal (and in fact identically 0) in the basis of states $(1, 0, 0)$ and $(0, 1, 0)$, you might assume they are the good states, but they’re not. To see this:

- (a) Obtain the exact eigenvalues for the perturbed Hamiltonian $H = H^0 + H'$.
- (b) Expand your results from part (a) as a power series in $|V|$ up to second order.
- (c) What do you obtain by applying *nondegenerate* perturbation theory to find the energies of all three states (up to second order)? This would work if the assumption about the good states above were correct.

Moral: If any of the eigenvalues of \mathbf{W} are equal, the states that diagonalize \mathbf{W} are not unique, and diagonalizing \mathbf{W} does not determine the “good” states. When this happens (and it’s not uncommon), you need to use *second*-order degenerate perturbation theory (see Problem 7.40).

Problem 7.40 If it happens that the square root in Equation 7.33 vanishes, then $E_+^1 = E_-^1$; the degeneracy is not lifted at first order. In this case, diagonalizing the \mathbf{W} matrix puts no restriction on α and β and you still don’t know what the “good” states are. If you need to determine the “good” states—for example to calculate higher-order corrections—you need to use **second-order degenerate perturbation theory**.

- (a) Show that, for the two-fold degeneracy studied in Section 7.2.1, the first-order correction to the wave function in degenerate perturbation theory is

$$\psi^1 = \sum_{m \neq a, b} \frac{\alpha V_{ma} + \beta V_{mb}}{E^0 - E_m^0} \psi_m^0.$$

- (b) Consider the terms of order λ^2 (corresponding to Equation 7.8 in the nondegenerate case) to show that α and β are determined by finding the eigenvectors of the matrix \mathbf{W}^2 (the superscript denotes second order, not \mathbf{W} squared) where

$$[\mathbf{W}^2]_{ij} = \sum_{m \neq a, b} \frac{\langle \psi_i^0 | H' | \psi_m^0 \rangle \langle \psi_m^0 | H' | \psi_j^0 \rangle}{E^0 - E_m^0}$$

and that the eigenvalues of this matrix correspond to the second-order energies E^2 .

- (c) Show that second-order degenerate perturbation theory, developed in (b), gives the correct energies to second order for the three-state Hamiltonian in Problem 7.39.

- ** **Problem 7.41** A free particle of mass m is confined to a ring of circumference L such that $\psi(x + L) = \psi(x)$. The unperturbed Hamiltonian is

$$H^0 = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2},$$

to which we add a perturbation

$$H' = V_0 \cos\left(2\pi \frac{x}{L}\right).$$

- (a) Show that the unperturbed states may be written

$$\psi_n^0(x) = \frac{1}{\sqrt{L}} e^{i2\pi nx/L}$$

for $n = 0, \pm 1, \pm 2$ and that, apart from $n = 0$, all of these states are two-fold degenerate.

- (b) Find a general expression for the matrix elements of the perturbation:

$$H'_{mn} = \langle \psi_m^0 | H' | \psi_n^0 \rangle.$$

- (c) Consider the degenerate pair of states with $n = \pm 1$. Construct the matrix \mathbf{W} and calculate the first-order energy corrections, E^1 . Note that the degeneracy *does not lift at first order*. Therefore, diagonalizing \mathbf{W} does not tell us what the “good” states are.
- (d) Construct the matrix \mathbf{W}^2 (Problem 7.40) for the states $n = \pm 1$, and show that the degeneracy lifts at second order. What are the good linear combinations of the states with $n = \pm 1$?
- (e) What are the energies, accurate to second order, for these states?³⁶

- ** Problem 7.42 The Feynman–Hellmann theorem (Problem 7.38) can be used to determine the expectation values of $1/r$ and $1/r^2$ for hydrogen.³⁷ The effective Hamiltonian for the radial wave functions is (Equation 4.53)

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dr^2} + \frac{\hbar^2 \ell(\ell+1)}{2m r^2} - \frac{e^2}{4\pi\epsilon_0 r},$$

and the eigenvalues (expressed in terms of ℓ)³⁸ are (Equation 4.70)

$$E_n = -\frac{me^4}{32\pi^2\epsilon_0^2\hbar^2(N+\ell)^2}.$$

- (a) Use $\lambda = e$ in the Feynman–Hellmann theorem to obtain $\langle 1/r \rangle$. Check your result against Equation 7.56.
- (b) Use $\lambda = \ell$ to obtain $\langle 1/r^2 \rangle$. Check your answer with Equation 7.57.

- *** Problem 7.43 Prove Kramers’ relation:³⁹

$$\frac{s+1}{n^2} \langle r^s \rangle - (2s+1) a \langle r^{s-1} \rangle + \frac{s}{4} [(2\ell+1)^2 - s^2] a^2 \langle r^{s-2} \rangle = 0, \quad (7.113)$$

which relates the expectation values of r to three different powers (s , $s-1$, and $s-2$), for an electron in the state $\psi_{n\ell m}$ of hydrogen. Hint: Rewrite the radial equation (Equation 4.53) in the form

$$u'' = \left[\frac{\ell(\ell+1)}{r^2} - \frac{2}{ar} + \frac{1}{n^2 a^2} \right] u,$$

and use it to express $\int (ur^s u'') dr$ in terms of $\langle r^s \rangle$, $\langle r^{s-1} \rangle$, and $\langle r^{s-2} \rangle$. Then use integration by parts to reduce the second derivative. Show that $\int (ur^s u') dr = -(s/2) \langle r^{s-1} \rangle$, and $\int (u'r^s u') dr = -[2/(s+1)] \int (u'' r^{s+1} u') dr$. Take it from there.

Problem 7.44

- (a) Plug $s = 0$, $s = 1$, $s = 2$, and $s = 3$ into Kramers' relation (Equation 7.113) to obtain formulas for $\langle r^{-1} \rangle$, $\langle r \rangle$, $\langle r^2 \rangle$, and $\langle r^3 \rangle$. Note that you could continue indefinitely, to find *any* positive power.
- (b) In the *other* direction, however, you hit a snag. Put in $s = -1$, and show that all you get is a relation between $\langle r^{-2} \rangle$ and $\langle r^{-3} \rangle$.
- (c) But if you can get $\langle r^{-2} \rangle$ by some *other* means, you can apply the Kramers relation to obtain the rest of the negative powers. Use Equation 7.57 (which is derived in Problem 7.42) to determine $\langle r^{-3} \rangle$, and check your answer against Equation 7.66.

Problem 7.45 When an atom is placed in a uniform external electric field \mathbf{E}_{ext} , the energy levels are shifted—a phenomenon known as the **Stark effect** (it is the electrical analog to the Zeeman effect). In this problem we analyze the Stark effect for the $n = 1$ and $n = 2$ states of hydrogen. Let the field point in the z direction, so the potential energy of the electron is

$$H'_S = eE_{\text{ext}}z = eE_{\text{ext}}r \cos \theta.$$

Treat this as a perturbation on the Bohr Hamiltonian (Equation 7.43). (Spin is irrelevant to this problem, so ignore it, and neglect the fine structure.)

- (a) Show that the ground state energy is not affected by this perturbation, in first order.
- (b) The first excited state is four-fold degenerate: ψ_{200} , ψ_{211} , ψ_{210} , ψ_{21-1} . Using degenerate perturbation theory, determine the first-order corrections to the energy. Into how many levels does E_2 split?
- (c) What are the “good” wave functions for part (b)? Find the expectation value of the electric dipole moment ($\mathbf{p}_e = -e\mathbf{r}$), in each of these “good” states. Notice that the results are independent of the applied field—evidently hydrogen in its first excited state can carry a *permanent* electric dipole moment.

Hint: There are lots of integrals in this problem, but almost all of them are zero. So study each one carefully, before you do any calculations: If the ϕ integral vanishes, there's not much point in doing the r and θ integrals! You can avoid those integrals altogether if you use the selection rules of Sections 6.4.3 and 6.7.2. *Partial answer:* $W_{13} = W_{31} = -3eaE_{\text{ext}}$; all other elements are zero.

Problem 7.46 Consider the Stark effect (Problem 7.45) for the $n = 3$ states of hydrogen. There are initially nine degenerate states, $\psi_{3\ell m}$ (neglecting spin, as before), and we turn on an electric field in the z direction.

- (a) Construct the 9×9 matrix representing the perturbing Hamiltonian.
Partial answer: $\langle 300 | z | 310 \rangle = -3\sqrt{6}a$, $\langle 310 | z | 320 \rangle = -3\sqrt{3}a$, $\langle 31\pm1 | z | 32\pm1 \rangle = -(9/2)a$.

(b) Find the eigenvalues, and their degeneracies.

Problem 7.47 Calculate the wavelength, in centimeters, of the photon emitted under a hyperfine transition in the ground state ($n = 1$) of deuterium. Deuterium is “heavy” hydrogen, with an extra neutron in the nucleus; the proton and neutron bind together to form a deuteron, with spin 1 and magnetic moment

$$\mu_d = \frac{g_d e}{2m_d} \mathbf{S}_d;$$

the deuteron g -factor is 1.71.

Problem 7.48 In a crystal, the electric field of neighboring ions perturbs the energy levels of an atom. As a crude model, imagine that a hydrogen atom is surrounded by three pairs of point charges, as shown in Figure 7.14. (Spin is irrelevant to this problem, so ignore it.)

(a) Assuming that $r \ll d_1$, $r \ll d_2$, and $r \ll d_3$, show that

$$H' = V_0 + 3 \left(\beta_1 x^2 + \beta_2 y^2 + \beta_3 z^2 \right) - (\beta_1 + \beta_2 + \beta_3) r^2,$$

where

$$\beta_i \equiv -\frac{e}{4\pi\epsilon_0} \frac{q_i}{d_i^3}, \quad \text{and } V_0 = 2 \left(\beta_1 d_1^2 + \beta_2 d_2^2 + \beta_3 d_3^2 \right).$$

(b) Find the lowest-order correction to the ground state energy.

(c) Calculate the first-order corrections to the energy of the first excited states ($n = 2$). Into how many levels does this four-fold degenerate system split, (i) in the case of **cubic symmetry**, $\beta_1 = \beta_2 = \beta_3$; (ii) in the case of **tetragonal symmetry**, $\beta_1 = \beta_2 \neq \beta_3$; (iii) in the general case of **orthorhombic symmetry** (all three different)? *Note:* you might recognize the “good” states from Problem 4.71.

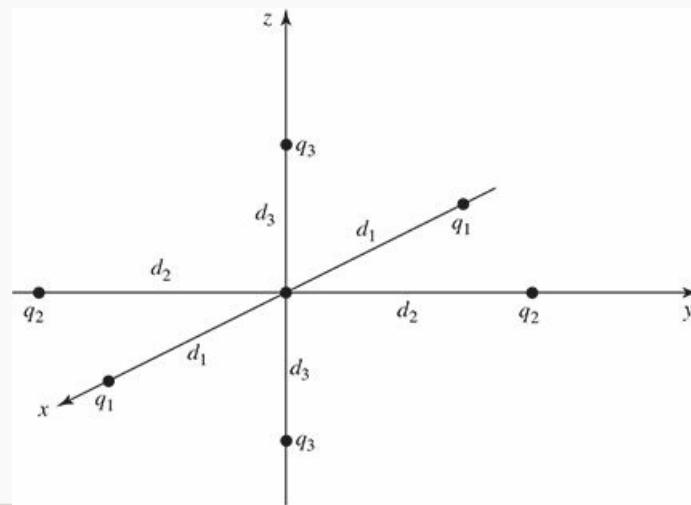


Figure 7.14: Hydrogen atom surrounded by six point charges (crude model)

for a crystal lattice); Problem 7.48.

Problem 7.49 A hydrogen atom is placed in a uniform magnetic field $\mathbf{B}_0 = B_0 \hat{z}$ (the Hamiltonian can be written as in Equation 4.230). Use the Feynman–Hellman theorem (Problem 7.38) to show that

$$\frac{\partial E_n}{\partial B_0} = -\langle \psi_n | \mu_z | \psi_n \rangle \quad (7.114)$$

where the electron's magnetic dipole moment⁴⁰ (orbital plus spin) is

$$\boldsymbol{\mu} = \gamma_o \mathbf{L}_{\text{mechanical}} + \gamma \mathbf{S}.$$

The mechanical angular momentum is defined in Equation 4.231.

Note: From Equation 7.114 it follows that the magnetic susceptibility of N atoms in a volume V and at 0 K (when they're all in the ground state) is⁴¹

$$\chi = \mu_0 \frac{\partial M}{\partial B_0} = -\frac{N}{V} \mu_0 \frac{\partial^2 E_0}{\partial B_0^2}, \quad (7.115)$$

where E_0 is the ground-state energy. Although we derived Equation 7.114 for a hydrogen atom, the expression applies to multi-electron atoms as well—even when electron–electron interactions are included.

Problem 7.50 For an atom in a uniform magnetic field $\mathbf{B}_0 = B_0 \hat{z}$, Equation 4.230 gives

$$H = H_{\text{atom}} - B_0 (\gamma_o L_z + \gamma S_z) + \frac{e^2}{8m} B_0^2 \sum_{i=1}^Z (x_i^2 + y_i^2),$$

where L_z and S_z refer to the total orbital and spin angular momentum of all the electrons.

- (a) Treating the terms involving B_0 as a perturbation, compute the shift of the ground state energy of a helium atom to second order in B_0 . Assume that the helium ground state is given by

$$\psi_0 = \psi_{100}(r_1) \psi_{100}(r_2) \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}$$

where ψ_{100} refers to the hydrogenic ground state (with $Z = 2$).

- (b) Use the results of Problem 7.49 to calculate the magnetic susceptibility of helium. Given a density of 0.166 kg/m^3 , obtain a numerical value for the susceptibility. *Note:* The experimental result is -1.0×10^{-9} (the negative sign means that helium is a **diamagnet**). The results can be brought closer by taking account of screening, which increases the orbital radius (see Section 8.2).

*** **Problem 7.51** Sometimes it is possible to solve Equation 7.10 directly, without having to expand ψ_n^1 in terms of the unperturbed wave functions (Equation 7.11). Here are two particularly nice examples.

(a) Stark effect in the ground state of hydrogen.

- (i) Find the first-order correction to the ground state of hydrogen in the presence of a uniform external electric field E_{ext} (see Problem 7.45). *Hint:* Try a solution of the form

$$(A + Br + Cr^2) e^{-r/a} \cos \theta;$$

your problem is to find the constants A , B , and C that solve Equation 7.10.

- (ii) Use Equation 7.14 to determine the second-order correction to the ground state energy (the first-order correction is zero, as you found in Problem 7.45(a)). *Answer:* $-m(3a^2eE_{\text{ext}}/2\hbar)^2$.

- (b) If the proton had an *electric* dipole moment p , the potential energy of the electron in hydrogen would be perturbed in the amount

$$H' = -\frac{ep \cos \theta}{4\pi \epsilon_0 r^2}.$$

- (i) Solve Equation 7.10 for the first-order correction to the ground state wave function.
- (ii) Show that the *total* electric dipole moment of the atom is (surprisingly) *zero*, to this order.
- (iii) Use Equation 7.14 to determine the second-order correction to the ground state energy. What is the *first*-order correction?

Problem 7.52 Consider a spinless particle of charge q and mass m constrained to move in the xy plane under the influence of the two-dimensional harmonic oscillator potential

$$V(x, y) = \frac{1}{2}m\omega^2(x^2 + y^2).$$

- (a) Construct the ground state wave function, $\psi_0(x, y)$, and write down its energy. Do the same for the (degenerate) first excited states.
- (b) Now imagine that we turn on a weak magnetic field of magnitude B_0 pointing in the z -direction, so that (to first order in B_0) the Hamiltonian acquires an extra term

$$H' = -\mu \cdot \mathbf{B} = -\frac{q}{2m}(\mathbf{L} \cdot \mathbf{B}) = -\frac{qB_0}{2m}(x p_y - y p_x).$$

Treating this as a perturbation, find the first-order corrections to the energies of the ground state and first excited states.



Problem 7.53 Imagine an infinite square well (Equation 2.22) into which we introduce a delta-function perturbation,

$$H'(x) = \lambda \delta(x - x_0),$$

where λ is a positive constant, and $0 < x_0 < a$ (to simplify matters, let $x_0 = pa$, where $0 < p < 1$).⁴²

- (a) Find the first-order correction to the n th allowed energy (Equation 2.30), assuming λ is small. (What does “small” mean, in this context?)
- (b) Find the second-order correction to the allowed energies. (Leave your answer as a sum.)
- (c) Now solve the Schrödinger equation exactly, treating separately the regions $0 \leq x < x_0$ and $x_0 < x < a$, and imposing the boundary conditions at x_0 . Derive the transcendental equation for the energies:

$$u_n \sin(u_n) + \Lambda \sin(pu_n) \sin[(1-p)u_n] = 0 \quad (E > 0). \quad (7.116)$$

Here $\Lambda \equiv 2ma\lambda/\hbar^2$, $u_n \equiv k_n a$, and $k_n \equiv \sqrt{2mE_n}/\hbar$.

Check that Equation 7.116 reproduces your result from part (a), in the appropriate limit.

- (d) Everything so far holds just as well if λ is *negative*, but in that case there may be an additional solution with negative energy. Derive the transcendental equation for a negative-energy state:

$$v \sinh(v) + \Lambda \sinh(pv) \sinh[(1-p)v] = 0 \quad (E < 0), \quad (7.117)$$

where $v \equiv \kappa a$ and $\kappa \equiv \sqrt{-2mE}/\hbar$. Specialize to the symmetrical case $p = 1/2$, and show that you recover the energy of the delta-function well (Equation 2.132), in the appropriate regime.

- (e) There is in fact exactly *one* negative-energy solution, provided that $|\Lambda| > 1/[p(1-p)]$. First, prove this (graphically), for the case $p = 1/2$. (Below that critical value there is no negative-energy solution.) Next, by computer, plot the solution v , as a function of p , for $\Lambda = -4.1, -5$, and -10 . Verify that the solution only exists within the predicted range of p .
- (f) For $p = 1/2$, plot the ground state wave function, $\psi_1(x)$, for $\Lambda = 0, -2, -3, -3.5, -4.5, -5$, and -10 , to show how the sinusoidal shape (Figure 2.2) evolves into the exponential shape (Figure 2.13), as the delta function well “deepens.”⁴³

Problem 7.54 Suppose you want to calculate the expectation value of some observable Ω , in the n th energy eigenstate of a system that is perturbed by H' :

$$\langle \Omega \rangle = \left\langle \psi_n \left| \hat{\Omega} \right| \psi_n \right\rangle.$$

Replacing ψ_n by its perturbation expansion, Equation 7.5,⁴⁴

$$\langle \Omega \rangle = \langle \psi_n^0 | \hat{\Omega} | \psi_n^0 \rangle + \lambda \left[\langle \psi_n^1 | \hat{\Omega} | \psi_n^0 \rangle + \langle \psi_n^0 | \hat{\Omega} | \psi_n^1 \rangle \right] + \lambda^2 (\dots) + \dots$$

The first-order correction to $\langle \Omega \rangle$ is therefore

$$\langle \Omega \rangle^1 = 2\text{Re} \left[\langle \psi_n^0 | \hat{\Omega} | \psi_n^1 \rangle \right],$$

or, using Equation 7.13,

$$\langle \Omega \rangle^1 = 2 \text{Re} \sum_{m \neq n} \frac{\langle \psi_n^0 | \hat{\Omega} | \psi_m^0 \rangle \langle \psi_m^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0} \quad (7.118)$$

(assuming the unperturbed energies are nondegenerate, or that we are using the “good” basis states).

- (a) Suppose $\Omega = H'$ (the perturbation itself). What does Equation 7.118 tell us in this case? Explain (carefully) why this is consistent with Equation 7.15.
- (b) Consider a particle of charge q (maybe an electron in a hydrogen atom, or a pith ball connected to a spring), that is placed in a weak electric field E_{ext} pointing in the x direction, so that

$$H' = -q E_{\text{ext}} x.$$

The field will induce an electric dipole moment, $p_e = qx$, in the “atom.” The expectation value of p_e is proportional to the applied field, and the proportionality factor is called the **polarizability**, α . Show that

$$\alpha = -2q^2 \sum_{m \neq n} \frac{|\langle \psi_n^0 | x | \psi_m^0 \rangle|^2}{E_n^0 - E_m^0}. \quad (7.119)$$

Find the polarizability of the ground state of a one-dimensional harmonic oscillator. Compare the classical answer.

- (c) Now imagine a particle of mass m in a one-dimensional harmonic oscillator with a small anharmonic perturbation⁴⁵

$$H' = -\frac{1}{6} \kappa x^3. \quad (7.120)$$

Find $\langle x \rangle$ (to first order), in the n th energy eigenstate. *Answer:* $\left(n + \frac{1}{2}\right) \hbar\omega / (2m^2\omega^3)$. *Comment:* As the temperature increases, higher-energy states are populated, and the particles move farther (on average) from their equilibrium positions; that’s why most solids expand with rising temperature.

Problem 7.55 Crandall’s Puzzle.⁴⁶ Stationary states of the one-dimensional Schrödinger equation ordinarily respect three “rules of thumb”: (1) the

energies are nondegenerate, (2) the ground state has no nodes, the first excited state has one node, the second has two, and so on, and (3) if the potential is an even function of x , the ground state is even, the first excited state is odd, the second is even, and so on. We have already seen that the “bead-on-a-ring” (Problem 2.46) violates the first of these; now suppose we introduce a “nick” in at the origin:

$$H' = -\alpha\delta(x).$$

(If you don’t like the delta function, make it a gaussian, as in Problem 7.9.) This lifts the degeneracy, but what is the sequence of even and odd wave functions, and what is the sequence of node numbers? *Hint:* You don’t really need to do any calculations, here, and you’re welcome to assume that α is small, but by all means solve the Schrödinger equation exactly if you prefer.

*** **Problem 7.56** In this problem we treat the electron–electron repulsion term in the helium Hamiltonian (Equation 5.38) as a perturbation,

$$H' = \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

(This will not be very accurate, because the perturbation is not small, in comparison to the Coulomb attraction of the nucleus ...but it’s a start.)

(a) Find the first-order correction to the ground state,

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2).$$

(You have already done this calculation, if you worked Problem 5.15—only we didn’t call it perturbation theory back then.)

(b) Now treat the first excited state, in which one electron is in the hydrogenic ground state, ψ_{100} , and the other is in the state ψ_{200} . Actually, there are *two* such states, depending on whether the electron spins occupy the singlet configuration (parahelium) or the triplet (orthohelium):⁴⁷

$$\psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} [\psi_{100}(\mathbf{r}_1)\psi_{200}(\mathbf{r}_2) \pm \psi_{200}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2)].$$

Show that

$$E_{\pm}^1 = \frac{1}{2}(K \pm J),$$

where

$$K \equiv 2 \int \psi_{100}(\mathbf{r}_1)\psi_{200}(\mathbf{r}_2) H' \psi_{100}(\mathbf{r}_1)\psi_{200}(\mathbf{r}_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2,$$

$$J \equiv 2 \int \psi_{100}(\mathbf{r}_1)\psi_{200}(\mathbf{r}_2) H' \psi_{200}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2.$$

Evaluate these two integrals, put in the actual numbers, and compare your results with Figure 5.2 (the measured energies are – 59.2 eV and – 58.4 eV).⁴⁸

Problem 7.57 The Hamiltonian for the Bloch functions (Equation 6.12) can be analyzed with perturbation theory by defining H^0 and H' such that

$$H^0 u_{n0} = E_{n0} u_{n0},$$

$$(H^0 + H') u_{nq} = E_{nq} u_{nq}.$$

In this problem, don't assume anything about the form of $V(x)$.

- (a) Determine the operators H^0 and H' (express them in terms of \hat{p}).
- (b) Find E_{nq} to second order in q . That is, find expressions for A_n , B_n , and C_n (in terms of the E_{n0} and matrix elements of \hat{p} in the “unperturbed” states u_{n0})

$$E_{nq} \approx A_n + B_n q + C_n q^2.$$

- (c) Show that the constants B_n are all zero. *Hint:* See Problem 2.1(b) to get started. Remember that $u_{n0}(x)$ is periodic.

Comment: It is conventional to write $C_n = \hbar^2 / 2m_n^*$ where m_n^* is the effective mass of particles in the n th band since then, as you've just shown,

$$E_{nq} \approx \text{constant} + \frac{\hbar^2 q^2}{2m_n^*}$$

just like the free particle (Equation 2.92) with $k \rightarrow q$.

¹ As always (footnote 34, page 49) the uniqueness of power series expansions guarantees that the coefficients of like powers are equal.

² In this context it doesn't matter whether we write $\langle \psi_n^0 | H' | \psi_n^0 \rangle$ or $\langle \psi_n^0 | H' | \psi_n^0 \rangle$ (with the extra vertical bar), because we are using the wave function itself to label the state. But the latter notation is preferable, because it frees us from this convention. For instance, if we used $|n\rangle$ to denote the n th state of the harmonic oscillator (Equation 2.86), $H' |n\rangle$ makes sense, but $|H'n\rangle$ is unintelligible (operators act on vectors/functions, not on numbers).

³ Incidentally, nothing here depends on the specific nature of the infinite square well—the same holds for *any* potential, when the perturbation is a constant.

⁴ Alternatively, a glance at Equation 7.5 reveals that any ψ_n^0 component in ψ_n^1 might as well be pulled out and combined with the first term. In fact, the choice $c_n^{(n)} = 0$ ensures that ψ_n —with 1 as the coefficient of ψ_n^0 in Equation 7.5—is *normalized* (to first order in λ): $\langle \psi_n | \psi_n \rangle = \langle \psi_n^0 | \psi_n^0 \rangle + \lambda (\langle \psi_n^1 | \psi_n^0 \rangle + \langle \psi_n^0 | \psi_n^1 \rangle) + \lambda^2 (\dots) + \dots$, but the orthonormality of the unperturbed states means that the first term is 1 and $\langle \psi_n^1 | \psi_n^0 \rangle = \langle \psi_n^0 | \psi_n^1 \rangle = 0$, as long as ψ_n^1 has no ψ_n^0 component.

⁵ In the short-hand notation $V_{mn} \equiv \langle \psi_m^0 | H' | \psi_n^0 \rangle$, $\Delta_{mn} \equiv E_m^0 - E_n^0$, the first three corrections to the n th energy are

$$E_n^1 = V_{nn}, \quad E_n^2 = \sum_{m \neq n} \frac{|V_{nm}|^2}{\Delta_{nm}}, \quad E_n^3 = \sum_{l,m \neq n} \frac{V_{nl} V_{lm} V_{mn}}{\Delta_{nl} \Delta_{nm}} - V_{nn} \sum_{m \neq n} \frac{|V_{nm}|^2}{\Delta_{nm}^2}.$$

The third-order correction is given in Landau and Lifschitz, *Quantum Mechanics: Non-Relativistic Theory*, 3rd edn, Pergamon, Oxford (1977), page 136; the fourth and fifth orders (together with a powerful general technique for obtaining the higher orders) are developed by Nicholas Wheeler, *Higher-Order Spectral Perturbation* (unpublished Reed College report, 2000). Illuminating alternative formulations of time-independent perturbation theory include the Dalgarno–Lewis method and the closely related “logarithmic” perturbation theory (see, for example, T. Imbo and U. Sukhatme, *Am. J. Phys.* **52**, 140 (1984), for LPT, and H. Mavromatis, *Am. J. Phys.* **59**, 738 (1991), for Delgarno–Lewis).

- ⁶ This assumes that the eigenvalues of \mathbf{W} are *distinct* so that the degeneracy lifts at first order. If not, any choice of α and β satisfies Equation 7.30; you *still* don't know what the good states are. The first-order energies are correctly given by Equation 7.33 when this happens, and in many cases that's all you require. But if you need to know the "good" states—for example to calculate higher-order corrections—you will have to use *second-order* degenerate perturbation theory (see Problems 7.39, 7.40, and 7.41) or employ the theorem of Section 7.2.2.
- ⁷ Note that the theorem is more general than Equation 7.30. In order to identify the good states from Equation 7.30, the energies E_{\pm}^1 need to be different. In some cases they are the same and the energies of the degenerate states split at second, third, or higher order in perturbation theory. But the *theorem* allows you to identify the good states in every case.
- ⁸ If the eigenvalues are degenerate, see footnote 6.
- ⁹ Degenerate perturbation theory amounts to diagonalization of the *degenerate part* of the Hamiltonian; see Problems 7.34 and 7.35.
- ¹⁰ The kinetic energy of the electron in hydrogen is on the order of 10 eV, which is minuscule compared to its rest energy (511,000 eV), so the hydrogen atom is basically nonrelativistic, and we can afford to keep only the lowest-order correction. In Equation 7.50, p is the *relativistic* momentum (Equation 7.48), *not* the classical momentum mv . It is the former that we now associate with the quantum operator $-i\hbar\nabla$, in Equation 7.51.
- ¹¹ An earlier edition of this book claimed that p^4 is not hermitian for states with $\ell = 0$ (calling into question the maneuver leading to Equation 7.54). That was incorrect— p^4 is hermitian, for all ℓ (see Problem 7.18).
- ¹² The general formula for the expectation value of *any* power of r is given in Hans A. Bethe and Edwin E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms*, Plenum, New York (1977), p. 17.
- ¹³ Thanks to Edward Ross and Li Yi-ding for fixing this problem.
- ¹⁴ We have already noted that it can be dangerous to picture the electron as a spinning sphere (see Problem 4.28), and it is not too surprising that the naive classical model gets the gyromagnetic ratio wrong. The *deviation* from the classical expectation is known as the **g-factor**: $\mu = g(q/2m)\mathbf{S}$. Thus the g-factor of the electron, in Dirac's theory, is exactly 2. But quantum electrodynamics reveals tiny corrections to this: ge is actually $2 + (\alpha/\pi) + \dots = 2.002\dots$. The calculation and measurement (which agree to exquisite precision) of the so-called **anomalous magnetic moment** of the electron were among the greatest achievements of twentieth-century physics.
- ¹⁵ One way of thinking of it is that the electron is continually stepping from one inertial system to another; Thomas precession amounts to the cumulative effect of all these Lorentz transformations. We could avoid the whole problem, of course, by staying in the *lab* frame, in which the nucleus is at rest. In that case the field of the proton is purely *electric*, and you might wonder why it exerts any torque on the electron. Well, the fact is that a moving *magnetic* dipole acquires an *electric* dipole moment, and in the lab frame the spin-orbit coupling is due to the interaction of the *electric* field of the nucleus with the *electric* dipole moment of the electron. Because this analysis requires more sophisticated electrodynamics, it seems best to adopt the electron's perspective, where the physical mechanism is more transparent.
- ¹⁶ More precisely, Thomas precession subtracts 1 from the gyromagnetic ratio (see R. R. Haar and L. J. Curtis, *Am. J. Phys.*, 55, 1044 (1987)).
- ¹⁷ In Problem 7.43 the expectation values are calculated using the hydrogen wave functions $\psi(n\ell m)$ —that is, eigenstates of L_z —whereas we now want eigenstates of J_z —which are linear combinations of $m = m_j + \frac{1}{2}$ and $m = m_j - \frac{1}{2}$. But since $\langle r^s \rangle$ is independent of m , it doesn't matter.
- ¹⁸ The case $\ell = 0$ looks problematic, since we are ostensibly dividing by zero. On the other hand, the numerator is *also* zero, since in this case $j = s$; so Equation 7.67 is indeterminate. On physical grounds there shouldn't be any spin-orbit coupling when $\ell = 0$. In any event, the problem disappears when the spin-orbit coupling is added to the relativistic correction, and their *sum* (Equation 7.68) is correct for *all* ℓ . If you're feeling uneasy about this whole calculation, I don't blame you; take comfort in the fact that the *exact* solution can be obtained by using the (relativistic) Dirac equation in place of the (nonrelativistic) Schrödinger equation, and it confirms the results we obtain here by less rigorous means (see Problem 7.22).
- ¹⁹ To write $|jm_j\rangle$ (for given ℓ and s) as a linear combination of $|\ell sm_{\ell}m_s\rangle$ we would use the appropriate Clebsch–Gordan coefficients (Equation 4.183).
- ²⁰ Bethe and Salpeter (footnote 12, page 298), page 238.
- ²¹ This is correct to first order in B . We are ignoring a term of order B^2 in the Hamiltonian (the exact result was calculated in Problem 4.72). In addition, the orbital magnetic moment (Equation 7.72) is proportional to the mechanical angular momentum, not the canonical angular momentum (see Problem 7.49). These neglected terms give corrections of order B^2 , comparable to the second-order corrections from H'_Z . Since we're working to first order, they are safe to ignore in this context.
- ²² The gyromagnetic ratio for *orbital* motion is just the classical value $(q/2m)$ —it is only for *spin* that there is an "extra" factor of 2.
- ²³ While Equation 7.78 was derived by replacing \mathbf{S} by its average value, the result is not an approximation; $\mathbf{L} + 2\mathbf{S}$ and \mathbf{J} are both vector operators and the states are angular-momentum eigenstates. Therefore, the matrix elements can be evaluated by use of the Wigner–Eckart theorem (Equations 6.59–6.61). It follows (Problem 7.25) that the matrix elements are proportional:

$$\langle n\ell jm_j | \mathbf{L} + 2\mathbf{S} | n\ell jm'_j \rangle = g_J \langle n\ell jm_j | \mathbf{J} | n\ell jm'_j \rangle \quad (7.77)$$

and the constant of proportionality g_J is the ratio of reduced matrix elements. All that remains is to evaluate g_J : see Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloë, *Quantum Mechanics*, Wiley, New York (1977), Vol. 2, Chapter X.

²⁴ In the case of a single electron, where $j = \ell \pm \frac{1}{2}$, $g_J = (2j+1)/(2\ell+1)$.

- ²⁵ That example specifically treated orbital angular momentum, but the same argument holds for the total angular momentum.
- ²⁶ In this regime the Zeeman effect is also known as the **Paschen–Back effect**.
- ²⁷ You can use ℓ, m_ℓ, m_s states if you prefer—this makes the matrix elements of H'_Z easier, but those of H'_{fs} more difficult; the W matrix will be more complicated, but its eigenvalues (which are independent of basis) are the same either way.
- ²⁸ Don't confuse the notation $|\ell s m_\ell m_s\rangle$ in the Clebsch–Gordan tables with $|n \ell j m_j\rangle$ (in Section 7.4.1) or $|n \ell m_\ell m_s\rangle$ (in Section 7.4.2); here n is always 2, and s (of course) is always $1/2$.
- ²⁹ If you are unfamiliar with the delta function term in Equation 7.90, you can derive it by treating the dipole as a spinning charged spherical shell, in the limit as the radius goes to zero and the charge goes to infinity (with μ held constant). See D. J. Griffiths, *Am. J. Phys.*, **50**, 698 (1982).
- ³⁰ For details see Griffiths, footnote 29, page 311.
- ³¹ See page 118 for a discussion of projection operators.
- ³² See Problem 7.4 for a discussion of what close means in this context.
- ³³ See, for example, Steven H. Simon, *The Oxford Solid State Basics* (Oxford University Press, 2013), Section 15.1.
- ³⁴ There is an interesting fraud in this well-known problem. If you expand H' to order $1/R^5$, the extra term has a nonzero expectation value in the ground state of H^0 , so there is a nonzero first-order perturbation, and the dominant contribution goes like $1/R^5$, not $1/R^6$. The model gets the power “right” in three dimensions (where the expectation value is zero), but not in one. See A. C. Ipsen and K. Splittorff, *Am. J. Phys.* **83**, 150 (2015).
- ³⁵ Feynman obtained Equation 7.110 while working on his undergraduate thesis at MIT (R. P. Feynman, *Phys. Rev.* **56**, 340, 1939); Hellmann's work was published four years earlier in an obscure Russian journal.
- ³⁶ See D. Kiang, *Am. J. Phys.* **46** (11), 1978 and L.-K. Chen, *Am. J. Phys.* **72** (7), 2004 for further discussion of this problem. It is shown that each degenerate energy level, $E_{\pm n}^0$, splits at order $2n$ in perturbation theory. The exact solution to the problem can also be obtained as the time-independent Schrödinger equation for $H = H^0 + H^1$ reduces to the **Mathieu equation**.
- ³⁷ C. Sánchez del Rio, *Am. J. Phys.*, **50**, 556 (1982); H. S. Valk, *Am. J. Phys.*, **54**, 921 (1986).
- ³⁸ In part (b) we treat ℓ as a continuous variable; n becomes a function of ℓ , according to Equation 4.67, because N , which must be an integer, is fixed. To avoid confusion, I have eliminated n , to reveal the dependence on ℓ explicitly.
- ³⁹ This is also known as the (second) **Pasternack relation**. See H. Beker, *Am. J. Phys.* **65**, 1118 (1997). For a proof based on the Feynman–Hellmann theorem (Problem 7.38) see S. Balasubramanian, *Am. J. Phys.* **68**, 959 (2000).
- ⁴⁰ For most purposes we can take this to be the magnetic moment of the atom as well. The proton's larger mass means that its contribution to the dipole moment is orders of magnitude smaller than the electron's contribution.
- ⁴¹ See Problem 5.33 for the definition of magnetic susceptibility. This formula does not apply when the ground state is degenerate (see Neil W. Ashcroft and N. David Mermin, *Solid State Physics* (Belmont: Cengage, 1976), p. 655); atoms with non-degenerate ground states have $J = 0$ (see Table 5.1).
- ⁴² We adopt the notation of Y. N. Joglekar, *Am. J. Phys.* **77**, 734 (2009), from which this problem is drawn.
- ⁴³ For the corresponding analysis of the delta function *barrier* (positive λ) see Problem 11.34.
- ⁴⁴ In general, Equation 7.5 does not deliver a *normalized* wave function, but the choice $c_n^{(n)} = 0$ in Equation 7.11 guarantees normalization to first order in λ , which is all we require here (see footnote 4, page 282).
- ⁴⁵ This is just a generic tweak to the simple harmonic oscillator potential, $\frac{1}{2}\kappa x^2$; κ is some constant, and the factor of $-1/6$ is for convenience.
- ⁴⁶ Richard Crandall introduced me to this problem.
- ⁴⁷ It seems strange, at first glance, that spin has anything to do with it, since the perturbation itself doesn't involve spin (and I'm not even bothering to include the spin state explicitly). The point, of course, is that an antisymmetric spin state forces a symmetric (position) wave function, and vice versa, and this *does* affect the result.
- ⁴⁸ If you want to pursue this problem further, see R. C. Massé and T. G. Walker, *Am. J. Phys.* **83**, 730 (2015).

8

The Variational Principle

◆

8.1 Theory

Suppose you want to calculate the ground state energy, E_{gs} , for a system described by the Hamiltonian H , but you are unable to solve the (time-independent) Schrödinger equation. The **variational principle** will get you an *upper bound* for E_{gs} , which is sometimes all you need, and often, if you're clever about it, very close to the exact value. Here's how it works: Pick *any normalized function ψ whatsoever*; I claim that

$$E_{\text{gs}} \leq \langle \psi | H | \psi \rangle \equiv \langle H \rangle. \quad (8.1)$$

That is, the expectation value of H , in the (presumably incorrect) state ψ is certain to *overestimate* the ground state energy. Of course, if ψ just happens to be one of the *excited states*, then *obviously* $\langle H \rangle$ exceeds E_{gs} ; the point is that the same holds for any ψ whatsoever.

Proof: Since the (unknown) eigenfunctions of H form a complete set, we can express ψ as a linear combination of them:¹

$$\psi = \sum_n c_n \psi_n, \quad \text{with } H \psi_n = E_n \psi_n.$$

Since ψ is normalized,

$$1 = \langle \psi | \psi \rangle = \left\langle \sum_m c_m \psi_m \left| \sum_n c_n \psi_n \right. \right\rangle = \sum_m \sum_n c_m^* c_n \langle \psi_m | \psi_n \rangle = \sum_n |c_n|^2,$$

(assuming the eigenfunctions themselves have been orthonormalized: $\langle \psi_m | \psi_n \rangle = \delta_{mn}$). Meanwhile,

$$\langle H \rangle = \left\langle \sum_m c_m \psi_m \left| H \sum_n c_n \psi_n \right. \right\rangle = \sum_m \sum_n c_m^* E_n c_n \langle \psi_m | \psi_n \rangle = \sum_n E_n |c_n|^2.$$

But the ground state energy is, by definition, the *smallest* eigenvalue, $E_{\text{gs}} \leq E_n$, and hence

$$\langle H \rangle \geq E_{\text{gs}} \sum_n |c_n|^2 = E_{\text{gs}},$$

which is what we were trying to prove.

This is hardly surprising. After all, ψ might be the *actual* wave function (at, say, $t = 0$). If you measured the particle's energy you'd be certain to get one of the eigenvalues of H , the smallest of which is E_{gs} , so the *average* of multiple measurements ($\langle H \rangle$) cannot be lower than E_{gs} .

Example 8.1

Suppose we want to find the ground state energy for the one-dimensional harmonic oscillator:

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2.$$

Of course, we already know the *exact* answer in this case (Equation 2.62): $E_{\text{gs}} = (1/2) \hbar\omega$; but this makes it a good test of the method. We might pick as our “trial” wave function the gaussian,

$$\psi(x) = Ae^{-bx^2}, \quad (8.2)$$

where b is a constant, and A is determined by normalization:

$$1 = |A|^2 \int_{-\infty}^{\infty} e^{-2bx^2} dx = |A|^2 \sqrt{\frac{\pi}{2b}} \Rightarrow A = \left(\frac{2b}{\pi}\right)^{1/4}. \quad (8.3)$$

Now

$$\langle H \rangle = \langle T \rangle + \langle V \rangle, \quad (8.4)$$

where, in this case,

$$\langle T \rangle = -\frac{\hbar^2}{2m}|A|^2 \int_{-\infty}^{\infty} e^{-bx^2} \frac{d^2}{dx^2} (e^{-bx^2}) dx = \frac{\hbar^2 b}{2m}, \quad (8.5)$$

and

$$\langle V \rangle = \frac{1}{2}m\omega^2|A|^2 \int_{-\infty}^{\infty} e^{-2bx^2} x^2 dx = \frac{m\omega^2}{8b},$$

so

$$\langle H \rangle = \frac{\hbar^2 b}{2m} + \frac{m\omega^2}{8b}. \quad (8.6)$$

According to Equation 8.1, this exceeds E_{gs} for any b ; to get the *tightest* bound, let’s minimize $\langle H \rangle$:

$$\frac{d}{db} \langle H \rangle = \frac{\hbar^2}{2m} - \frac{m\omega^2}{8b^2} = 0 \Rightarrow b = \frac{m\omega}{2\hbar}.$$

Putting this back into $\langle H \rangle$, we find

$$\langle H \rangle_{\min} = \frac{1}{2}\hbar\omega. \quad (8.7)$$

In this case we hit the ground state energy right on the nose—because (obviously) I “just happened” to pick a trial function with precisely the form of the *actual* ground state (Equation 2.60). But the gaussian is very easy to work with, so it’s a popular trial function, even when it bears little resemblance to the true ground state.

Example 8.2

Suppose we’re looking for the ground state energy of the delta function potential:

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} - \alpha\delta(x).$$

Again, we already know the exact answer (Equation 2.132): $E_{\text{gs}} = -m\alpha^2/2\hbar^2$. As before, we'll use a gaussian trial function (Equation 8.2). We've already determined the normalization, and calculated $\langle T \rangle$; all we need is

$$\langle V \rangle = -\alpha|A|^2 \int_{-\infty}^{\infty} e^{-2bx^2} \delta(x) dx = -\alpha \sqrt{\frac{2b}{\pi}}.$$

Evidently

$$\langle H \rangle = \frac{\hbar^2 b}{2m} - \alpha \sqrt{\frac{2b}{\pi}}, \quad (8.8)$$

and we know that this exceeds E_{gs} for all b . Minimizing it,

$$\frac{d}{db} \langle H \rangle = \frac{\hbar^2}{2m} - \frac{\alpha}{\sqrt{2\pi b}} = 0 \Rightarrow b = \frac{2m^2\alpha^2}{\pi\hbar^4}.$$

So

$$\langle H \rangle_{\min} = -\frac{m\alpha^2}{\pi\hbar^2}, \quad (8.9)$$

which is indeed somewhat higher than E_{gs} , since $\pi > 2$.

I said you can use *any* (normalized) trial function ψ whatsoever, and this is true in a sense. However, for *discontinuous* functions it takes some fancy footwork to assign a sensible meaning to the second derivative (which you need, in order to calculate $\langle T \rangle$). Continuous functions with kinks in them are fair game, however, as long as you are careful; the next example shows how to handle them.³

Example 8.3

Find an upper bound on the ground state energy of the one-dimensional infinite square well (Equation 2.22), using the “triangular” trial wave function (Figure 8.1).³

$$\psi(x) = \begin{cases} Ax, & 0 \leq x \leq a/2, \\ A(a-x), & a/2 \leq x \leq a, \\ 0, & \text{otherwise,} \end{cases} \quad (8.10)$$

where A is determined by normalization:

$$1 = |A|^2 \left[\int_0^{a/2} x^2 dx + \int_{a/2}^a (a-x)^2 dx \right] = |A|^2 \frac{a^3}{12} \Rightarrow A = \frac{2}{a} \sqrt{\frac{3}{a}}. \quad (8.11)$$

In this case

$$\frac{d\psi}{dx} = \begin{cases} A, & 0 \leq x \leq a/2, \\ -A, & a/2 \leq x \leq a, \\ 0, & \text{otherwise,} \end{cases}$$

as indicated in Figure 8.2. Now, the derivative of a step function is a delta function (see Problem 2.23(b)):

$$\frac{d^2\psi}{dx^2} = A\delta(x) - 2A\delta(x - a/2) + A\delta(x - a), \quad (8.12)$$

and hence

$$\begin{aligned}\langle H \rangle &= -\frac{\hbar^2 A}{2m} \int [\delta(x) - 2\delta(x - a/2) + \delta(x - a)] \psi(x) dx \\ &= -\frac{\hbar^2 A}{2m} [\psi(0) - 2\psi(a/2) + \psi(a)] = \frac{\hbar^2 A^2 a}{2m} = \frac{12\hbar^2}{2ma^2}.\end{aligned} \quad (8.13)$$

The exact ground state energy is $E_{\text{gs}} = \pi^2 \hbar^2 / 2ma^2$ (Equation 2.30), so the theorem works ($12 > \pi^2$).

Alternatively, you can exploit the hermiticity of \hat{p} :

$$\begin{aligned}\langle H \rangle &= \frac{1}{2m} \langle \hat{p}^2 \rangle = \frac{1}{2m} \langle \hat{p} \psi | \hat{p} \psi \rangle = \frac{1}{2m} \int_0^a \left(-i\hbar \frac{d\psi}{dx} \right)^* \left(-i\hbar \frac{d\psi}{dx} \right) dx \\ &= \frac{\hbar^2}{2m} \left[\int_0^{a/2} (A)^2 dx + \int_{a/2}^a (-A)^2 dx \right] = \frac{\hbar^2}{2m} A^2 a = \frac{12\hbar^2}{2ma^2}.\end{aligned} \quad (8.14)$$

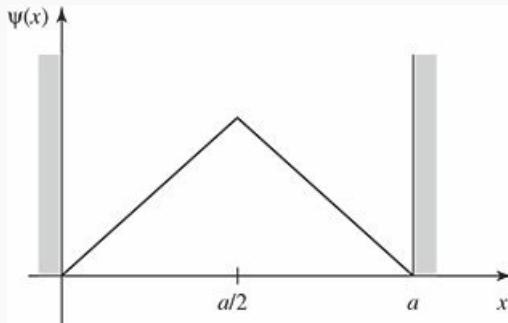


Figure 8.1: Triangular trial wave function for the infinite square well (Equation 8.10).

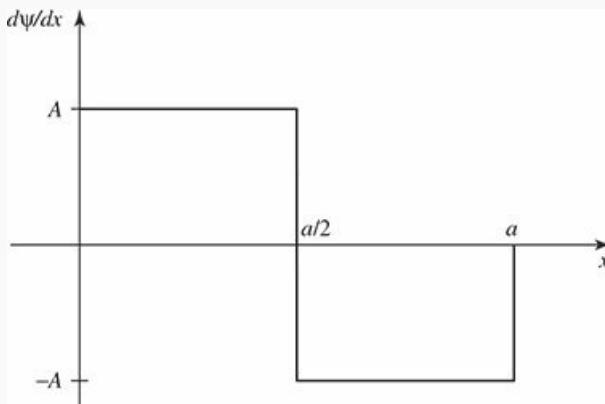


Figure 8.2: Derivative of the wave function in Figure 8.1.

The variational principle is extraordinarily powerful, and embarrassingly easy to use. What a physical chemist does, to find the ground state energy of some complicated molecule, is write down a trial wave function with a large number of adjustable parameters, calculate $\langle H \rangle$, and tweak the parameters to get the lowest possible value. Even if ψ has little resemblance to the true wave function, you often get miraculously accurate values for E_{gs} . Naturally, if you have some way of guessing a *realistic* ψ , so much the better. The only trouble with the method is that you never know for sure how close you are to the target—all you can be *certain* of is that you've got an *upper bound*.⁴ Moreover, as it stands the technique applies only to the ground state (see, however, Problem 8.4).⁵

- * **Problem 8.1** Use a gaussian trial function (Equation 8.2) to obtain the lowest upper bound you can on the ground state energy of (a) the linear potential: $V(x) = \alpha|x|$; (b) the quartic potential: $V(x) = \alpha x^4$.

- ** **Problem 8.2** Find the best bound on E_{gs} for the one-dimensional harmonic oscillator using a trial wave function of the form

$$\psi(x) = \frac{A}{x^2 + b^2},$$

where A is determined by normalization and b is an adjustable parameter.

Problem 8.3 Find the best bound on E_{gs} for the delta function potential $V(x) = -\alpha\delta(x)$, using a triangular trial function (Equation 8.10, only centered at the origin). This time a is an adjustable parameter.

Problem 8.4

- (a) Prove the following corollary to the variational principle: If $\langle \psi | \psi_{\text{gs}} \rangle = 0$, then $\langle H \rangle \geq E_{\text{fe}}$, where E_{fe} is the energy of the first excited state.

Comment: If we can find a trial function that is orthogonal to the exact ground state, we can get an upper bound on the *first excited state*. In general, it's difficult to be sure that ψ is orthogonal to ψ_{gs} since (presumably) we don't *know* the latter. However, if the potential $V(x)$ is an *even* function of x , then the ground state is likewise even, and hence any *odd* trial function will automatically meet the condition for the corollary.⁶

- (b) Find the best bound on the first excited state of the one-dimensional harmonic oscillator using the trial function

$$\psi(x) = Ax e^{-bx^2}.$$

Problem 8.5 Using a trial function of your own devising, obtain an upper bound on the ground state energy for the “bouncing ball” potential (Equation 2.185), and compare it with the exact answer (Problem 2.59):

$$E_{\text{gs}} = 2.33811 \left(mg^2 \hbar^2 / 2 \right)^{1/3}.$$

Problem 8.6

- (a) Use the variational principle to prove that first-order non-degenerate perturbation theory always *overestimates* (or at any rate never *underestimates*) the ground state energy.
- (b) In view of (a), you would expect that the *second*-order correction to the ground state is always negative. Confirm that this is indeed the case, by examining Equation 7.15.

8.2 The Ground State of Helium

The helium atom (Figure 8.3) consists of two electrons in orbit around a nucleus containing two protons (also some neutrons, which are irrelevant to our purpose). The Hamiltonian for this system (ignoring fine structure and smaller corrections) is:

$$H = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - \frac{e^2}{4\pi\epsilon_0} \left(\frac{2}{r_1} + \frac{2}{r_2} - \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right). \quad (8.15)$$

Our problem is to calculate the ground state energy, E_{gs} . Physically, this represents the amount of energy it would take to strip off both electrons. (Given E_{gs} it is easy to figure out the “ionization energy” required to remove a *single* electron—see Problem 8.7.) The ground state energy of helium has been measured to great precision in the laboratory:

$$E_{\text{gs}} = -78.975 \text{ eV} \quad (\text{experimental}). \quad (8.16)$$

This is the number we would like to reproduce theoretically.

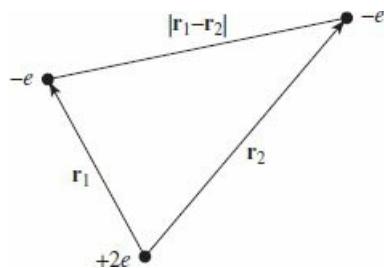


Figure 8.3: The helium atom.

It is curious that such a simple and important problem has no known exact solution.⁷ The trouble comes from the electron-electron repulsion,

$$V_{ee} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (8.17)$$

If we ignore this term altogether, H splits into two independent hydrogen Hamiltonians (only with a nuclear charge of $2e$, instead of e); the exact solution is just the product of hydrogenic wave functions:

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2) \equiv \psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2) = \frac{8}{\pi a^3} e^{-2(r_1+r_2)/a}, \quad (8.18)$$

and the energy is $8E_1 = -109 \text{ eV}$ (Equation 5.42).⁸ This is a long way from -79 eV , but it's a start.

To get a better approximation for E_{gs} we'll apply the variational principle, using ψ_0 as the trial wave function. This is a particularly convenient choice because it's an eigenfunction of *most* of the Hamiltonian:

$$H\psi_0 = (8E_1 + V_{ee})\psi_0. \quad (8.19)$$

Thus

$$\langle H \rangle = 8E_1 + \langle V_{ee} \rangle, \quad (8.20)$$

where⁹

$$\langle V_{ee} \rangle = \left(\frac{e^2}{4\pi\epsilon_0} \right) \left(\frac{8}{\pi a^3} \right)^2 \int \frac{e^{-4(r_1+r_2)/a}}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3\mathbf{r}_1 d^3\mathbf{r}_2. \quad (8.21)$$

I'll do the \mathbf{r}_2 integral first; for this purpose \mathbf{r}_1 is fixed, and we may as well orient the \mathbf{r}_2 coordinate system so that the polar axis lies along \mathbf{r}_1 (see Figure 8.4). By the law of cosines,

$$|\mathbf{r}_1 - \mathbf{r}_2| = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta_2}, \quad (8.22)$$

and hence

$$I_2 \equiv \int \frac{e^{-4r_2/a}}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3\mathbf{r}_2 = \int \frac{e^{-4r_2/a}}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta_2}} r_2^2 \sin \theta_2 dr_2 d\theta_2 d\phi_2. \quad (8.23)$$

The ϕ_2 integral is trivial (2π); the θ_2 integral is

$$\begin{aligned} \int_0^\pi \frac{\sin \theta_2}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta_2}} d\theta_2 &= \frac{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta_2}}{r_1 r_2} \Big|_0^\pi \\ &= \frac{1}{r_1 r_2} \left(\sqrt{r_1^2 + r_2^2 + 2r_1 r_2} - \sqrt{r_1^2 + r_2^2 - 2r_1 r_2} \right) \\ &= \frac{1}{r_1 r_2} [(r_1 + r_2) - |r_1 - r_2|] = \begin{cases} 2/r_1, & r_2 < r_1, \\ 2/r_2, & r_2 > r_1. \end{cases} \end{aligned} \quad (8.24)$$

Thus

$$\begin{aligned} I_2 &= 4\pi \left(\frac{1}{r_1} \int_0^{r_1} e^{-4r_2/a} r_2^2 dr_2 + \int_{r_1}^\infty e^{-4r_2/a} r_2 dr_2 \right) \\ &= \frac{\pi a^3}{8r_1} \left[1 - \left(1 + \frac{2r_1}{a} \right) e^{-4r_1/a} \right]. \end{aligned} \quad (8.25)$$

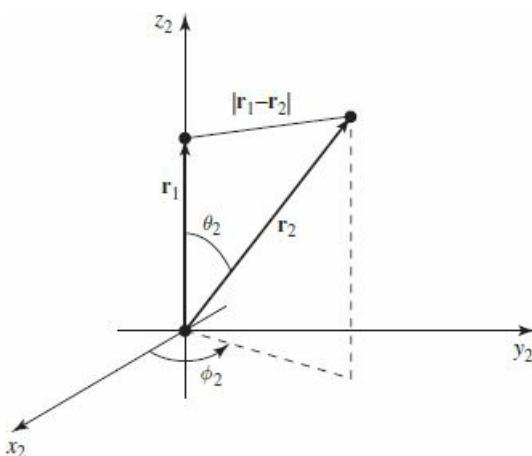


Figure 8.4: Choice of coordinates for the \mathbf{r}_2 -integral (Equation 8.21).

It follows that $\langle V_{ee} \rangle$ is equal to

$$\left(\frac{e^2}{4\pi\epsilon_0}\right)\left(\frac{8}{\pi a^3}\right)\int \left[1 - \left(1 + \frac{2r_1}{a}\right)e^{-4r_1/a}\right] e^{-4r_1/a} r_1 \sin\theta_1 dr_1 d\theta_1 d\phi_1.$$

The angular integrals are easy (4π), and the r_1 integral becomes

$$\int_0^\infty \left[re^{-4r/a} - \left(r + \frac{2r^2}{a}\right)e^{-8r/a}\right] dr = \frac{5a^2}{128}.$$

Finally, then,

$$\langle V_{ee} \rangle = \frac{5}{4a} \left(\frac{e^2}{4\pi\epsilon_0}\right) = -\frac{5}{2} E_1 = 34 \text{ eV}, \quad (8.26)$$

and therefore

$$\langle H \rangle = -109 \text{ eV} + 34 \text{ eV} = -75 \text{ eV}. \quad (8.27)$$

Not bad (remember, the experimental value is -79 eV). But we can do better.

We need to think up a more realistic trial function than ψ_0 (which treats the two electrons as though they did not interact at all). Rather than completely *ignoring* the influence of the other electron, let us say that, on the average, each electron represents a cloud of negative charge which partially *shields* the nucleus, so that the other electron actually sees an *effective* nuclear charge (Z) that is somewhat *less* than 2. This suggests that we use a trial function of the form

$$\psi_1(\mathbf{r}_1, \mathbf{r}_2) \equiv \frac{Z^3}{\pi a^3} e^{-Z(r_1+r_2)/a}. \quad (8.28)$$

We'll treat Z as a variational parameter, picking the value that minimizes $\langle H \rangle$. (Please note that in the variational method we *never touch the Hamiltonian itself* —the Hamiltonian for helium is, and remains, Equation 8.15. But it's fine to *think* about approximating the Hamiltonian *as a way of motivating the choice of the trial wave function*.)

This wave function is an eigenstate of the “unperturbed” Hamiltonian (neglecting electron repulsion), only with Z , instead of 2, in the Coulomb terms. With this in mind, we rewrite H (Equation 8.15) as follows:

$$H = -\frac{\hbar^2}{2m} \left(\nabla_1^2 + \nabla_2^2\right) - \frac{e^2}{4\pi\epsilon_0} \left(\frac{Z}{r_1} + \frac{Z}{r_2}\right) + \frac{e^2}{4\pi\epsilon_0} \left(\frac{(Z-2)}{r_1} + \frac{(Z-2)}{r_2} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\right). \quad (8.29)$$

The expectation value of H is evidently

$$\langle H \rangle = 2Z^2 E_1 + 2(Z-2) \left(\frac{e^2}{4\pi\epsilon_0}\right) \left\langle \frac{1}{r} \right\rangle + \langle V_{ee} \rangle. \quad (8.30)$$

Here $\langle 1/r \rangle$ is the expectation value of $1/r$ in the (one-particle) hydrogenic ground state ψ_{100} (with nuclear charge Z); according to Equation 7.56,

$$\left\langle \frac{1}{r} \right\rangle = \frac{Z}{a}. \quad (8.31)$$

The expectation value of V_{ee} is the same as before (Equation 8.26), except that instead of $Z = 2$ we now want arbitrary Z —so we multiply a by $2/Z$:

$$\langle V_{ee} \rangle = \frac{5Z}{8a} \left(\frac{e^2}{4\pi\epsilon_0} \right) = -\frac{5Z}{4} E_1. \quad (8.32)$$

Putting all this together, we find

$$\langle H \rangle = \left[2Z^2 - 4Z(Z-2) - (5/4)Z \right] E_1 = \left[-2Z^2 + (27/4)Z \right] E_1. \quad (8.33)$$

According to the variational principle, this quantity exceeds E_{gs} for *any* value of Z . The *lowest* upper bound occurs when $\langle H \rangle$ is minimized:

$$\frac{d}{dZ} \langle H \rangle = [-4Z + (27/4)] E_1 = 0,$$

from which it follows that

$$Z = \frac{27}{16} = 1.69. \quad (8.34)$$

This seems reasonable; it tells us that the other electron partially screens the nucleus, reducing its effective charge from 2 down to about 1.69. Putting in this value for Z , we find

$$\langle H \rangle = \frac{1}{2} \left(\frac{3}{2} \right)^6 E_1 = -77.5 \text{ eV}. \quad (8.35)$$

The ground state of helium has been calculated with great precision in this way, using increasingly complicated trial wave functions, with more and more adjustable parameters.¹⁰ But we're within 2% of the correct answer, and, frankly, at this point my own interest in the problem begins to wane.¹¹

Problem 8.7 Using $E_{gs} = -79.0$ eV for the ground state energy of helium, calculate the ionization energy (the energy required to remove just *one* electron).
Hint: First calculate the ground state energy of the helium ion, He^+ , with a single electron orbiting the nucleus; then subtract the two energies.

* **Problem 8.8** Apply the techniques of this Section to the H^- and Li^+ ions (each has two electrons, like helium, but nuclear charges $Z = 1$ and $Z = 3$, respectively). Find the effective (partially shielded) nuclear charge, and determine the best upper bound on E_{gs} for each case. *Comment:* In the case of H^- you should find that $\langle H \rangle > -13.6$ eV, which would appear to indicate that there is no bound state at all, since it would be energetically favorable for one electron to fly off, leaving behind a neutral hydrogen atom. This is not entirely surprising, since the electrons are less strongly attracted to the nucleus than they are in helium, and the electron repulsion tends to break the atom apart. However, it turns out to be incorrect. With a more sophisticated trial wave function (see

Problem 8.25) it can be shown that $E_{\text{gs}} < -13.6 \text{ eV}$, and hence that a bound state *does* exist. It's only *barely* bound, however, and there are no excited bound states,¹² so H^- has no discrete spectrum (all transitions are to and from the continuum). As a result, it is difficult to study in the laboratory, although it exists in great abundance on the surface of the sun.¹³

8.3 The Hydrogen Molecule Ion

Another classic application of the variational principle is to the hydrogen molecule ion, H_2^+ , consisting of a single electron in the Coulomb field of two protons (Figure 8.5). I shall assume for the moment that the protons are fixed in position, a specified distance R apart, although one of the most interesting byproducts of the calculation is going to be the actual *value* of R . The Hamiltonian is

$$H = -\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{r} + \frac{1}{r'} \right), \quad (8.36)$$

where r and r' are the distances to the electron from the respective protons. As always, our strategy will be to guess a reasonable trial wave function, and invoke the variational principle to get a bound on the ground state energy. (Actually, our main interest is in finding out whether this system bonds at *all*—that is, whether its energy is less than that of a neutral hydrogen atom plus a free proton. If our trial wave function indicates that there *is* a bound state, a *better* trial function can only make the bonding even stronger.)

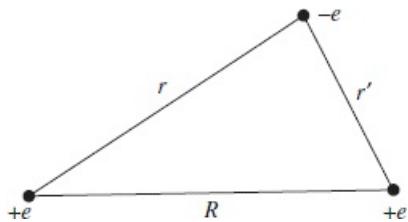


Figure 8.5: The hydrogen molecule ion, H_2^+ .

To construct the trial wave function, imagine that the ion is formed by taking a hydrogen atom in its ground state (Equation 4.80),

$$\psi_0(\mathbf{r}) = \frac{1}{\sqrt{\pi a^3}} e^{-r/a}, \quad (8.37)$$

bringing the second proton in from “infinity,” and nailing it down a distance R away. If R is substantially greater than the Bohr radius, the electron’s wave function probably isn’t changed very much. But we would like to treat the two protons on an equal footing, so that the electron has the same probability of being associated with either one. This suggests that we consider a trial function of the form

$$\psi = A [\psi_0(r) + \psi_0(r')]. \quad (8.38)$$

(Quantum chemists call this the LCAO technique, because we are expressing the *molecular* wave function as a linear combination of atomic orbitals.)

Our first task is to *normalize* the trial function:

$$1 = \int |\psi|^2 d^3\mathbf{r} = |A|^2 \left[\int \psi_0(r)^2 d^3\mathbf{r} + \int \psi_0(r')^2 d^3\mathbf{r} + 2 \int \psi_0(r)\psi_0(r') d^3\mathbf{r} \right]. \quad (8.39)$$

The first two integrals are 1 (since ψ_0 itself is normalized); the third is more tricky. Let

$$I \equiv \langle \psi_0(r) | \psi_0(r') \rangle = \frac{1}{\pi a^3} \int e^{-(r+r')/a} d^3 \mathbf{r}. \quad (8.40)$$

Picking coordinates so that proton 1 is at the origin and proton 2 is on the z axis at the point R (Figure 8.6), we have

$$r' = \sqrt{r^2 + R^2 - 2rR \cos \theta}, \quad (8.41)$$

and therefore

$$I = \frac{1}{\pi a^3} \int e^{-r/a} e^{-\sqrt{r^2+R^2-2rR \cos \theta}/a} r^2 \sin \theta dr d\theta d\phi. \quad (8.42)$$

The ϕ integral is trivial (2π). To do the θ integral, let $y \equiv \sqrt{r^2 + R^2 - 2rR \cos \theta}$, so that $d(y^2) = 2ydy = 2rR \sin \theta d\theta$. Then

$$\begin{aligned} \int_0^\pi e^{-\sqrt{r^2+R^2-2rR \cos \theta}/a} \sin \theta d\theta &= \frac{1}{rR} \int_{|r-R|}^{r+R} e^{-y/a} y dy \\ &= -\frac{a}{rR} \left[e^{-(r+R)/a} (r+R+a) - e^{-|r-R|/a} (|r-R|+a) \right]. \end{aligned}$$

The r integral is now straightforward:

$$\begin{aligned} I &= \frac{2}{a^2 R} \left[-e^{-R/a} \int_0^\infty (r+R+a) e^{-2r/a} r dr + e^{-R/a} \int_0^R (R-r+a) r dr \right. \\ &\quad \left. + e^{R/a} \int_R^\infty (r-R+a) e^{-2r/a} r dr \right]. \end{aligned}$$

Evaluating the integrals, we find (after some algebraic simplification),

$$I = e^{-R/a} \left[1 + \left(\frac{R}{a} \right) + \frac{1}{3} \left(\frac{R}{a} \right)^2 \right]. \quad (8.43)$$

I is called an **overlap** integral; it measures the amount by which $\psi_0(r)$ overlaps $\psi_0(r')$ (notice that it goes to 1 as $R \rightarrow 0$, and to 0 as $R \rightarrow \infty$). In terms of I , the normalization factor (Equation 8.39) is

$$|A|^2 = \frac{1}{2(1+I)}. \quad (8.44)$$

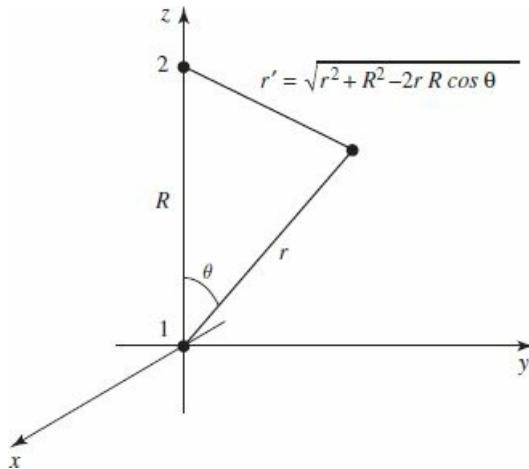


Figure 8.6: Coordinates for the calculation of I (Equation 8.40).

Next we must calculate the expectation value of H in the trial state ψ . Noting that

$$\left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r}\right)\psi_0(r) = E_1\psi_0(r)$$

(where $E_1 = -13.6$ eV is the ground state energy of atomic hydrogen)—and the same with r' in place of r —we have

$$\begin{aligned} H\psi &= A \left[-\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{r} + \frac{1}{r'} \right) \right] [\psi_0(r) + \psi_0(r')] \\ &= E_1\psi - A \left(\frac{e^2}{4\pi\epsilon_0} \right) \left[\frac{1}{r'}\psi_0(r) + \frac{1}{r}\psi_0(r') \right]. \end{aligned}$$

It follows that

$$\langle H \rangle = E_1 - 2|A|^2 \left(\frac{e^2}{4\pi\epsilon_0} \right) \left[\left\langle \psi_0(r) \left| \frac{1}{r'} \right| \psi_0(r) \right\rangle + \left\langle \psi_0(r) \left| \frac{1}{r} \right| \psi_0(r') \right\rangle \right]. \quad (8.45)$$

I'll let you calculate the two remaining quantities, the so-called **direct integral**,

$$D \equiv a \left\langle \psi_0(r) \left| \frac{1}{r'} \right| \psi_0(r) \right\rangle, \quad (8.46)$$

and the **exchange integral**,

$$X \equiv a \left\langle \psi_0(r) \left| \frac{1}{r} \right| \psi_0(r') \right\rangle. \quad (8.47)$$

The results (see Problem 8.9) are

$$D = \frac{a}{R} - \left(1 + \frac{a}{R} \right) e^{-2R/a}, \quad (8.48)$$

and

$$X = \left(1 + \frac{R}{a} \right) e^{-R/a}. \quad (8.49)$$

Putting all this together, and recalling (Equations 4.70 and 4.72) that $E_1 = -(e^2/4\pi\epsilon_0)(1/2a)$, we conclude:

$$\langle H \rangle = \left[1 + 2 \frac{(D+X)}{(1+I)} \right] E_1. \quad (8.50)$$

According to the variational principle, the ground state energy is *less* than $\langle H \rangle$. Of course, this is only the *electron's* energy—there is also potential energy associated with the proton–proton repulsion:

$$V_{pp} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{R} = -\frac{2a}{R} E_1. \quad (8.51)$$

Thus the *total* energy of the system, in units of $-E_1$, and expressed as a function of $x \equiv R/a$, is less than

$$F(x) = -1 + \frac{2}{x} \left\{ \frac{(1 - (2/3)x^2)e^{-x} + (1+x)e^{-2x}}{1 + (1+x + (1/3)x^2)e^{-x}} \right\}. \quad (8.52)$$

This function is plotted in Figure 8.7. Evidently bonding *does* occur, for there exists a region in which the graph goes below -1 , indicating that the energy is less than that of a neutral atom plus a free proton (-13.6 eV). It's a **covalent bond**, with the electron shared equally by the two protons. The equilibrium separation of the protons is about 2.4 Bohr radii, or 1.3 Å (the experimental value is 1.06 Å). The calculated binding energy is 1.8 eV, whereas the experimental value is 2.8 eV (the variational principle, as always, *over* estimates the ground state energy—and hence *under* estimates the strength of the bond—but never mind: The essential point was to see whether binding occurs at all; a better variational function can only make the potential well even deeper).

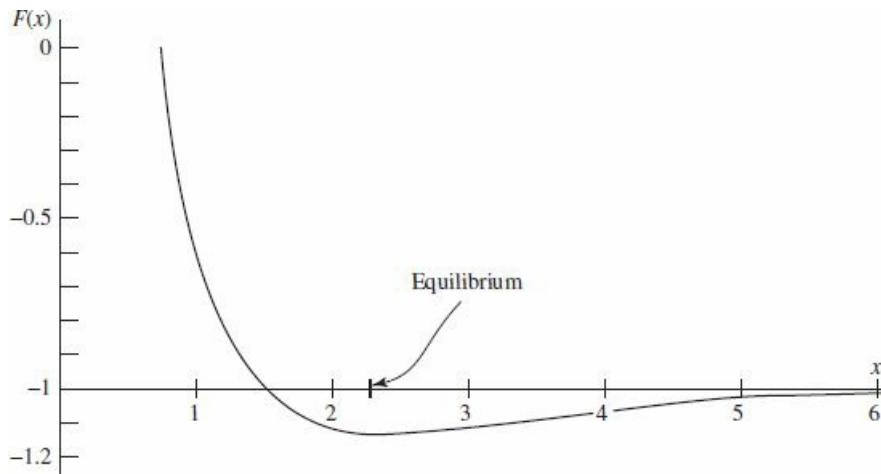


Figure 8.7: Plot of the function $F(x)$, Equation 8.52, showing existence of a bound state.

- * **Problem 8.9** Evaluate D and X (Equations 8.46 and 8.47). Check your answers against Equations 8.48 and 8.49.

**

Problem 8.10 Suppose we used a *minus* sign in our trial wave function (Equation [8.38](#)):

$$\psi = A [\psi_0(r) - \psi_0(r')] . \quad (8.53)$$

Without doing any new integrals, find $F(x)$ (the analog to Equation [8.52](#)) for this case, and construct the graph. Show that there is no evidence of bonding.¹⁴ (Since the variational principle only gives an *upper bound*, this doesn't *prove* that bonding cannot occur for such a state, but it certainly doesn't look promising.)

*** 

Problem 8.11 The second derivative of $F(x)$, at the equilibrium point, can be used to estimate the natural frequency of vibration (ω) of the two protons in the hydrogen molecule ion (see Section [2.3](#)). If the ground state energy ($\hbar\omega/2$) of this oscillator exceeds the binding energy of the system, it will fly apart. Show that in fact the oscillator energy is small enough that this will *not* happen, and estimate how many bound vibrational levels there are. *Note:* You're not going to be able to obtain the position of the minimum—still less the second derivative at that point—analytically. Do it numerically, on a computer.

8.4 The Hydrogen Molecule

Now consider the hydrogen molecule itself, adding a second electron to the hydrogen molecule ion we studied in Section 8.3. Taking the two protons to be at rest, the Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m} \left(\nabla_1^2 + \nabla_2^2 \right) + \frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{r_{12}} + \frac{1}{R} - \frac{1}{r_1} - \frac{1}{r'_1} - \frac{1}{r_2} - \frac{1}{r'_2} \right) \quad (8.54)$$

where r_1 and r'_1 are the distances of electron 1 from each proton and r_2 and r'_2 are the distances of electron 2 from each proton; as shown in Figure 8.8. The six potential energy terms describe the repulsion between the two electrons, the repulsion between the two protons, and the attraction of each electron to each proton.

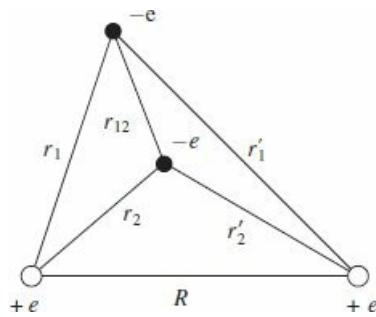


Figure 8.8: Diagram of H₂ showing the distances on which the potential energy depends.

For the variational wave function, associate one electron with each proton, and symmetrize:

$$\psi_+(\mathbf{r}_1, \mathbf{r}_2) = A_+ [\psi_0(r_1)\psi_0(r'_2) + \psi_0(r'_1)\psi_0(r_2)]. \quad (8.55)$$

We'll calculate the normalization A_+ in a moment. Since this spatial wave function is symmetric under interchange, the electrons must occupy the antisymmetric (singlet) spin state. Of course, we could also choose the trial wave function

$$\psi_-(\mathbf{r}_1, \mathbf{r}_2) = A_- [\psi_0(r_1)\psi_0(r'_2) - \psi_0(r'_1)\psi_0(r_2)] \quad (8.56)$$

in which case the electrons would be in a symmetric (triplet) spin state. These two variational wave functions constitute the **Heitler–London approximation**.¹⁵ It is not obvious which of Equations 8.55 or 8.56 would be energetically favored, so let's calculate the energy of each one, and find out.¹⁶

First we need to normalize the wave functions. Note that

$$|\psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2)|^2 = A_{\pm}^2 \left[\psi_0(r_1)^2 \psi_0(r'_2)^2 + \psi_0(r'_1)^2 \psi_0(r_2)^2 \right. \\ \left. \pm 2\psi_0(r_1)\psi_0(r'_2)\psi_0(r'_1)\psi_0(r_2) \right]. \quad (8.58)$$

Normalization requires

$$\begin{aligned}
 1 &= \int \int |\psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2)|^2 d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 \\
 &= A_{\pm}^2 \left[\int \psi_0(r_1)^2 d^3 \mathbf{r}_1 \int \psi_0(r'_2)^2 d^3 \mathbf{r}_2 + \int \psi_0(r'_1)^2 d^3 \mathbf{r}_1 \int \psi_0(r_2)^2 d^3 \mathbf{r}_2 \right. \\
 &\quad \left. \pm 2 \int \psi_0(r_1) \psi_0(r'_1) d^3 \mathbf{r}_1 \int \psi_0(r'_2) \psi_0(r_2) d^3 \mathbf{r}_2 \right].
 \end{aligned} \tag{8.59}$$

The individual orbitals are normalized and the overlap integral was given the symbol I and calculated in Equation 8.43. Thus

$$A_{\pm} = \frac{1}{\sqrt{2(1 \pm I^2)}}. \tag{8.60}$$

To calculate the expectation value of the energy, we will start with the kinetic energy of particle 1. Since ψ_0 is the ground state of the hydrogen Hamiltonian, the same trick that brought us to Equation 8.45 gives

$$\begin{aligned}
 -\frac{\hbar^2}{2m} \nabla_1^2 \psi_{\pm} &= A_{\pm} \left[\left(-\frac{\hbar^2}{2m} \nabla_1^2 \psi_0(r_1) \right) \psi_0(r'_2) \pm \left(-\frac{\hbar^2}{2m} \nabla_1^2 \psi_0(r'_1) \right) \psi_0(r_2) \right] \\
 &= A_{\pm} \left[\left(E_1 + \frac{e^2}{4\pi\epsilon_0 r_1} \right) \psi_0(r_1) \psi_0(r'_2) \right. \\
 &\quad \left. \pm \left(E_1 + \frac{e^2}{4\pi\epsilon_0 r'_1} \right) \psi_0(r'_1) \psi_0(r_2) \right] \\
 &= E_1 \psi_{\pm} + \frac{e^2}{4\pi\epsilon_0 a} A_{\pm} \left(\frac{a}{r_1} \psi_0(r_1) \psi_0(r'_2) \pm \frac{a}{r'_1} \psi_0(r'_1) \psi_0(r_2) \right).
 \end{aligned}$$

Taking the inner product with ψ_{\pm} then gives

$$\begin{aligned}
 \left\langle -\frac{\hbar^2}{2m} \nabla_1^2 \right\rangle &= E_1 + \left(\frac{e^2}{4\pi\epsilon_0 a} \right) A_{\pm}^2 \left[\left\langle \psi_0(r_1) \left| \frac{a}{r_1} \right| \psi_0(r_1) \right\rangle \langle \psi_0(r'_2) | \psi_0(r'_2) \rangle \right. \\
 &\quad \pm \left\langle \psi_0(r'_1) \left| \frac{a}{r_1} \right| \psi_0(r_1) \right\rangle \langle \psi_0(r_2) | \psi_0(r'_2) \rangle \\
 &\quad \pm \left\langle \psi_0(r_1) \left| \frac{a}{r'_1} \right| \psi_0(r'_1) \right\rangle \langle \psi_0(r'_2) | \psi_0(r_2) \rangle \\
 &\quad \left. + \left\langle \psi_0(r'_1) \left| \frac{a}{r'_1} \right| \psi_0(r'_1) \right\rangle \langle \psi_0(r_2) | \psi_0(r_2) \rangle \right].
 \end{aligned} \tag{8.61}$$

These inner products were calculated in Section 8.3 and the kinetic energy of particle 1 is

$$\left\langle -\frac{\hbar^2}{2m} \nabla_1^2 \right\rangle = E_1 + \left(\frac{e^2}{4\pi\epsilon_0 a} \right) \frac{1 \pm IX}{1 \pm I^2}. \tag{8.62}$$

The kinetic energy of particle 2 is of course the same, so the total kinetic energy is simply twice Equation 8.62. The calculation of the electron–proton potential energy is similar; you will show in Problem 8.13 that

$$\left\langle -\frac{e^2}{4\pi\epsilon_0 r_1} \right\rangle = -\frac{1}{2} \left(\frac{e^2}{4\pi\epsilon_0 a} \right) \frac{1 + D \pm 2IX}{1 \pm I^2} \tag{8.63}$$

and the total electron–proton potential energy is four times this amount.

The electron–electron potential energy is given by

$$\begin{aligned}\langle V_{ee} \rangle &= \left(\frac{e^2}{4\pi\epsilon_0 a} \right) \int \int |\psi_{\pm}(\mathbf{r}_1, \mathbf{r}_2)|^2 \frac{a}{r_{12}} d^3\mathbf{r}_1 d^3\mathbf{r}_2 \\ &= \left(\frac{e^2}{4\pi\epsilon_0 a} \right) A_{\pm}^2 \left[\int \int \psi_0(r_1)^2 \frac{a}{r_{12}} \psi_0(r'_2)^2 d^3\mathbf{r}_1 d^3\mathbf{r}_2 \right. \\ &\quad + \int \int \psi_0(r'_1)^2 \frac{a}{r_{12}} \psi_0(r_2)^2 d^3\mathbf{r}_1 d^3\mathbf{r}_2 \\ &\quad \left. \pm 2 \int \int \psi_0(r_1) \psi_0(r'_1) \frac{a}{r_{12}} \psi_0(r_2) \psi_0(r'_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2 \right].\end{aligned}\tag{8.64}$$

The first two integrals in Equation 8.64 are equal, as you can see by interchanging the labels 1 and 2. We will give the two remaining integrals the names

$$D_2 = \int \int |\psi_0(r_1)|^2 \frac{a}{r_{12}} |\psi_0(r'_2)|^2 d^3\mathbf{r}_1 d^3\mathbf{r}_2\tag{8.65}$$

$$X_2 = \int \int \psi_0(r_1) \psi_0(r'_1) \frac{a}{r_{12}} \psi_0(r_2) \psi_0(r'_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2\tag{8.66}$$

so that

$$\langle V_{ee} \rangle = \left(\frac{e^2}{4\pi\epsilon_0 a} \right) \frac{D_2 \pm X_2}{1 \pm I^2}.\tag{8.67}$$

The evaluation of these integrals is discussed in Problem 8.14. Note that the integral D_2 is just the electrostatic potential energy of two charge distributions $\rho_1 = |\psi_0(r_1)|^2$ and $\rho_2 = |\psi_0(r_2)|^2$. The exchange term X_2 has no such classical counterpart.

When we add all of the contributions to the energy—the kinetic energy, the electron–proton potential energy, the electron–electron potential energy, and the proton–proton potential energy (which is a constant, $e^2/4\pi\epsilon_0 R$)—we get

$$\langle H \rangle_{\pm} = 2E_1 \left[1 - \frac{a}{R} + \frac{2D - D_2 \pm (2IX - X_2)}{1 \pm I^2} \right].\tag{8.68}$$

A plot of $\langle H \rangle_+$ and $\langle H \rangle_-$ is shown in Figure 8.9. Recall that the state ψ_+ requires placing the two electrons in the singlet spin configuration, whereas ψ_- means putting them in a triplet spin configuration. According to the figure, bonding only occurs if the two electrons are in a singlet configuration—something that is confirmed experimentally. Again, it's a covalent bond.

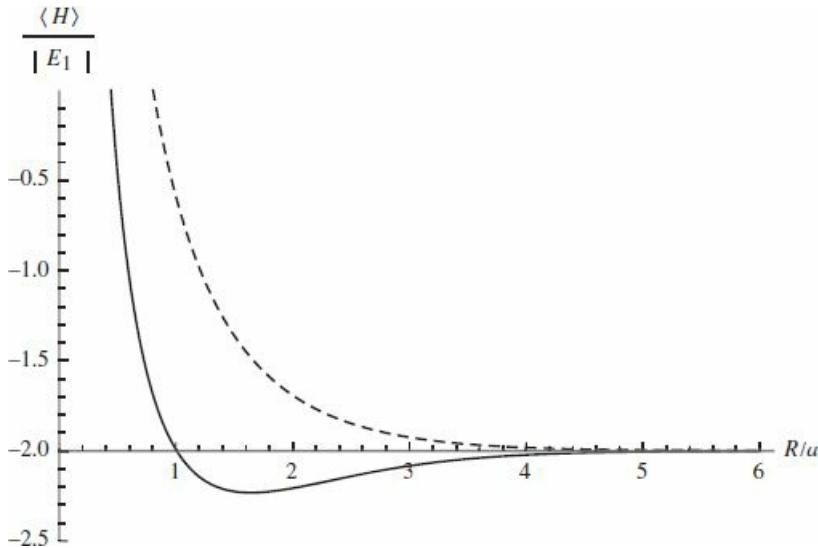


Figure 8.9: The total energy of the singlet (solid curve) and triplet (dashed curve) states for H_2 , as a function of the separation R between the protons. The singlet state has a minimum at around 1.6 Bohr radii, representing a stable bond. The triplet state is unstable and will dissociate, as the energy is minimized for $R \rightarrow \infty$.

Locating the minimum on the plot, our calculation predicts a bond length of 1.64 Bohr radii (the experimental value is 1.40 Bohr radii), and suggests a binding energy of 3.15 eV (whereas the experimental value is 4.75 eV). The trends here follow those of the Hydrogen molecule ion: the calculation overestimates the bond length and underestimates the binding energy, but the agreement is surprisingly good for a variational calculation with no adjustable parameters.

The difference between the singlet and triplet energies is called the **exchange splitting** J . In the Heitler–London approximation it is

$$J = \langle H \rangle_+ - \langle H \rangle_- = 4E_1 \frac{(D_2 - 2D) I^2 - (X_2 - 2IX)}{1 - I^4}, \quad (8.69)$$

which is roughly -10 eV (negative because the singlet is lower in energy) at the equilibrium separation. This means a strong preference for having the electron spins anti-aligned. But in this treatment of H_2 we've left out completely the (magnetic) spin–spin interaction between the electrons—remember that the spin–spin interaction between the *proton* and the electron is what leads to hyperfine splitting (Section 7.5). Were we right to ignore it here? Absolutely: applying Equation 7.92 to two electrons a distance R apart, the energy of the spin–spin interaction is something like 10^{-4} eV in this system, five orders of magnitude smaller than the exchange splitting.

This calculation shows us that different spin configurations can have very different energies, even when the interaction between the spins is negligible. And that helps us understand ferromagnetism (where the spins in a material align) and anti-ferromagnetism (where the spins alternate). As we've just seen, the spin–spin interaction is *way* too weak to account for this—but the exchange splitting isn't. Counterintuitively, it's not a *magnetic* interaction that accounts for ferromagnetism, but an electrostatic one! H_2 is a sort of inchoate anti-ferromagnet where the Hamiltonian, which is independent of the spin, selects a certain *spatial* ground state and the *spin* state comes along for the ride, to satisfy the Fermi statistics.

Problem 8.12 Show that the antisymmetric state (Equation 8.56) can be expressed in terms of the molecular orbitals of Section 8.3—specifically, by placing one electron in the bonding orbital (Equation 8.38) and one in the anti-bonding orbital (Equation 8.53).

Problem 8.13 Verify Equation 8.63 for the electron–proton potential energy.

Problem 8.14 The two-body integrals D_2 and X_2 are defined in Equations 8.65 and 8.66. To evaluate D_2 we write

$$\begin{aligned} D_2 &= \int |\psi_0(r'_2)|^2 \Phi(r_2) d^3\mathbf{r}_2 \\ &= \int \int \int \frac{e^{-2\sqrt{R^2+r_2^2-2Rr_2 \cos \theta_2}/a}}{\pi a^3} \Phi(r_2) r_2^2 dr_2 \sin \theta_2 d\theta_2 d\phi_2 \end{aligned}$$

where θ_2 is the angle between \mathbf{R} and \mathbf{r}_2 (Figure 8.8), and

$$\Phi(r_2) \equiv \int |\psi_0(r_1)|^2 \frac{a}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3\mathbf{r}_1.$$

- (a) Consider first the integral over \mathbf{r}_1 . Align the z axis with \mathbf{r}_2 (which is a constant vector for the purposes of this first integral) so that

$$\Phi(r_2) = \frac{1}{\pi a^3} \int \int \int \frac{ae^{-2r_1/a}}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta_1}} r_1^2 dr_1 \sin \theta_1 d\theta_1 d\phi_1.$$

Do the angular integration first and show that

$$\Phi(r_2) = \frac{a}{r_2} - \left(1 + \frac{a}{r_2}\right) e^{-2r_2/a}.$$

- (b) Plug your result from part (a) back into the relation for D_2 , and show that

$$D_2 = \frac{a}{R} - e^{-2R/a} \left[\frac{1}{6} \left(\frac{R}{a}\right)^2 + \frac{3}{4} \left(\frac{R}{a}\right) + \frac{11}{8} + \frac{a}{R} \right]. \quad (8.70)$$

Again, do the angular integration first.

Comment: The integral X_2 can also be evaluated in closed form, but the procedure is rather involved.¹⁷ We will simply quote the result,

$$\begin{aligned} X_2 &= e^{-2R/a} \left[\frac{5}{8} - \frac{23}{20} \frac{R}{a} - \frac{3}{5} \left(\frac{R}{a}\right)^2 - \frac{1}{15} \left(\frac{R}{a}\right)^3 \right] \\ &\quad + \frac{6}{5} \frac{a}{R} I^2 \left[\gamma + \log \left(\frac{R}{a}\right) + \left(\frac{I}{R}\right)^2 \text{Ei}\left(-\frac{4R}{a}\right) - 2 \frac{I}{R} \text{Ei}\left(-\frac{2R}{a}\right) \right], \end{aligned}$$

where $\gamma = 0.5772 \dots$ is Euler's constant, $Ei(x)$ is the exponential integral (8.71)

$$Ei(x) = - \int_{-x}^{\infty} \frac{e^{-t}}{t} dt,$$

and \tilde{I} is obtained from I by switching the sign of R :

$$\tilde{I} = e^{R/a} \left[1 - \frac{R}{a} + \frac{1}{3} \left(\frac{R}{a} \right)^2 \right]. \quad (8.72)$$



Problem 8.15 Make a plot of the kinetic energy for both the singlet and triplet states of H_2 , as a function of R/a . Do the same for the electron-proton potential energy and for the electron-electron potential energy. You should find that the triplet state has lower potential energy than the singlet state for all values of R . However, the singlet state's *kinetic* energy is *so* much smaller that its total energy comes out lower. *Comment:* In situations where there is not a large kinetic energy cost to aligning the spins, such as two electrons in a partially filled orbital in an atom, the triplet state can come out lower in energy. This is the physics behind Hund's first rule.

Further Problems on Chapter 8

Problem 8.16

- (a) Use the function $\psi(x) = Ax(a-x)$ (for $0 < x < a$, otherwise 0) to get an upper bound on the ground state of the infinite square well.
- (b) Generalize to a function of the form $\psi(x) = A[x(a-x)]^p$, for some real number p . What is the optimal value of p , and what is the best bound on the ground state energy? Compare the exact value. *Answer:* $(5 + 2\sqrt{6}) \hbar^2 / 2ma^2$.

Problem 8.17

- (a) Use a trial wave function of the form

$$\psi(x) = \begin{cases} A \cos(\pi x/a), & -a/2 < x < a/2, \\ 0, & \text{otherwise,} \end{cases}$$

to obtain a bound on the ground state energy of the one-dimensional harmonic oscillator. What is the “best” value of a ? Compare $\langle H \rangle_{\min}$ with the exact energy. *Note:* This trial function has a “kink” in it (a discontinuous derivative) at $\pm a/2$; do you need to take account of this, as I did in Example 8.3?

- (b) Use $\psi(x) = B \sin(\pi x/a)$ on the interval $(-a, a)$ to obtain a bound on the first excited state. Compare the exact answer.

**** Problem 8.18**

- (a) Generalize Problem 8.2, using the trial wave function¹⁸

$$\psi(x) = \frac{A}{(x^2 + b^2)^n},$$

for arbitrary n . *Partial answer:* The best value of b is given by

$$b^2 = \frac{\hbar}{m\omega} \left[\frac{n(4n-1)(4n-3)}{2(2n+1)} \right]^{1/2}.$$

- (b) Find the least upper bound on the first excited state of the harmonic oscillator using a trial function of the form

$$\psi(x) = \frac{Bx}{(x^2 + b^2)^n}.$$

Partial answer: The best value of b is given by

$$b^2 = \frac{\hbar}{m\omega} \left[\frac{n(4n-5)(4n-3)}{2(2n+1)} \right]^{1/2}.$$

- (c) Notice that the bounds approach the exact energies as $n \rightarrow \infty$. Why is that? *Hint:* Plot the trial wave functions for $n = 2$, $n = 3$, and $n = 4$, and compare them with the true wave functions (Equations 2.60 and 2.63). To do it analytically, start with the identity

$$e^z = \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n.$$

Problem 8.19 Find the lowest bound on the ground state of hydrogen you can get using a gaussian trial wave function

$$\psi(\mathbf{r}) = A e^{-br^2},$$

where A is determined by normalization and b is an adjustable parameter. *Answer:* – 11.5 eV.

Problem 8.20 Find an upper bound on the energy of the first excited state of the hydrogen atom. A trial function with $\ell = 1$ will automatically be orthogonal to the ground state (see footnote 6); for the radial part of ψ you can use the same function as in Problem 8.19.

- ** **Problem 8.21** If the photon had a nonzero mass ($m_\gamma \neq 0$), the Coulomb potential would be replaced by the Yukawa potential,

$$V(\mathbf{r}) = -\frac{e^2}{4\pi\epsilon_0} \frac{e^{-\mu r}}{r}, \quad (8.73)$$

where $\mu = m_\gamma c/\hbar$. With a trial wave function of your own devising, estimate the binding energy of a “hydrogen” atom with this potential. Assume $\mu a \ll 1$, and give your answer correct to order $(\mu a)^2$.

Problem 8.22 Suppose you’re given a two-level quantum system whose (time-independent) Hamiltonian H^0 admits just two eigenstates, ψ_a (with energy E_a), and ψ_b (with energy E_b). They are orthogonal, normalized, and nondegenerate (assume E_a is the smaller of the two energies). Now we turn on a perturbation H' , with the following matrix elements:

$$\langle \psi_a | H' | \psi_a \rangle = \langle \psi_b | H' | \psi_b \rangle = 0; \quad \langle \psi_a | H' | \psi_b \rangle = \langle \psi_b | H' | \psi_a \rangle = h, \quad (8.74)$$

where h is some specified constant.

- (a) Find the exact eigenvalues of the perturbed Hamiltonian.
- (b) Estimate the energies of the perturbed system using second-order perturbation theory.
- (c) Estimate the ground state energy of the perturbed system using the variational principle, with a trial function of the form

$$\psi = (\cos \phi) \psi_a + (\sin \phi) \psi_b, \quad (8.75)$$

where ϕ is an adjustable parameter. *Note:* Writing the linear combination in this way is just a neat way to guarantee that ψ is normalized.

- (d) Compare your answers to (a), (b), and (c). Why is the variational principle so accurate, in this case?

Problem 8.23 As an explicit example of the method developed in Problem 8.22, consider an electron at rest in a uniform magnetic field $\mathbf{B} = B_z \hat{k}$, for which the Hamiltonian is (Equation 4.158):

$$H^0 = \frac{eB_z}{m} S_z. \quad (8.76)$$

The eigenspinors, χ_a and χ_b , and the corresponding energies, E_a and E_b , are given in Equation 4.161. Now we turn on a perturbation, in the form of a uniform field in the x direction:

$$H' = \frac{eB_x}{m} S_x. \quad (8.77)$$

- (a) Find the matrix elements of H' , and confirm that they have the structure of Equation 8.74. What is b ?
- (b) Using your result in Problem 8.22(b), find the new ground state energy, in second-order perturbation theory.
- (c) Using your result in Problem 8.22(c), find the variational principle bound on the ground state energy.

Problem 8.24 Although the Schrödinger equation for helium itself cannot be solved exactly, there exist “helium-like” systems that do admit exact solutions. A simple example¹⁹ is “rubber-band helium,” in which the Coulomb forces are replaced by Hooke’s law forces:

$$H = -\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) + \frac{1}{2} m\omega^2 (r_1^2 + r_2^2) - \frac{\lambda}{4} m\omega^2 |\mathbf{r}_1 - \mathbf{r}_2|^2. \quad (8.78)$$

- (a) Show that the change of variables from $\mathbf{r}_1, \mathbf{r}_2$, to

$$\mathbf{u} \equiv \frac{1}{\sqrt{2}} (\mathbf{r}_1 + \mathbf{r}_2), \quad \mathbf{v} \equiv \frac{1}{\sqrt{2}} (\mathbf{r}_1 - \mathbf{r}_2), \quad (8.79)$$

turns the Hamiltonian into two independent three-dimensional harmonic oscillators:

$$H = \left[-\frac{\hbar^2}{2m} \nabla_u^2 + \frac{1}{2} m\omega^2 u^2 \right] + \left[-\frac{\hbar^2}{2m} \nabla_v^2 + \frac{1}{2} (1 - \lambda) m\omega^2 v^2 \right]. \quad (8.80)$$

- (b) What is the *exact* ground state energy for this system?
- (c) If we didn’t know the exact solution, we might be inclined to apply the method of Section 8.2 to the Hamiltonian in its original form (Equation

[8.78](#)). Do so (but don't bother with shielding). How does your result compare with the exact answer? *Answer:* $\langle H \rangle = 3\hbar\omega(1 - \lambda/4)$.

Problem 8.25 In Problem [8.8](#) we found that the trial wave function with shielding (Equation [8.28](#)), which worked well for helium, is inadequate to confirm the existence of a bound state for the negative hydrogen ion. Chandrasekhar²⁰ used a trial wave function of the form

$$\psi(\mathbf{r}_1, \mathbf{r}_2) \equiv A [\psi_1(r_1)\psi_2(r_2) + \psi_2(r_1)\psi_1(r_2)], \quad (8.81)$$

where

$$\psi_1(r) \equiv \sqrt{\frac{Z_1^3}{\pi a^3}} e^{-Z_1 r/a}, \quad \text{and} \quad \psi_2(r) \equiv \sqrt{\frac{Z_2^3}{\pi a^3}} e^{-Z_2 r/a}. \quad (8.82)$$

In effect, he allowed two *different* shielding factors, suggesting that one electron is relatively close to the nucleus, and the other is farther out. (Because electrons are identical particles, the spatial wave function must be symmetrized with respect to interchange. The *spin* state—which is irrelevant to the calculation—is evidently antisymmetric.) Show that by astute choice of the adjustable parameters Z_1 and Z_2 you can get $\langle H \rangle$ less than -13.6 eV. *Answer:*

$$\langle H \rangle = \frac{E_1}{x^6 + y^6} \left(-x^8 + 2x^7 + \frac{1}{2}x^6y^2 - \frac{1}{2}x^5y^2 - \frac{1}{8}x^3y^4 + \frac{11}{8}xy^6 - \frac{1}{2}y^8 \right),$$

where $x \equiv Z_1 + Z_2$ and $y \equiv 2\sqrt{Z_1 Z_2}$. Chandrasekhar used $Z_1 = 1.039$ (since this is larger than 1, the motivating interpretation as an effective nuclear charge cannot be sustained, but never mind—it's still an acceptable trial wave function) and $Z_2 = 0.283$.

Problem 8.26 The fundamental problem in harnessing nuclear fusion is getting the two particles (say, two deuterons) close enough together for the attractive (but short-range) nuclear force to overcome the Coulomb repulsion. The “bulldozer” method is to heat the particles up to fantastic temperatures, and allow the random collisions to bring them together. A more exotic proposal is **muon catalysis**, in which we construct a “hydrogen molecule ion,” only with deuterons in place of protons, and a *muon* in place of the electron. Predict the equilibrium separation distance between the deuterons in such a structure, and explain why muons are superior to electrons for this purpose.²¹

Problem 8.27 Quantum dots. Consider a particle constrained to move in two dimensions in the cross-shaped region shown in Figure [8.10](#). The “arms” of the cross continue out to infinity. The potential is zero within the cross, and infinite in the shaded areas outside. Surprisingly, this configuration admits a positive-energy bound state.²²

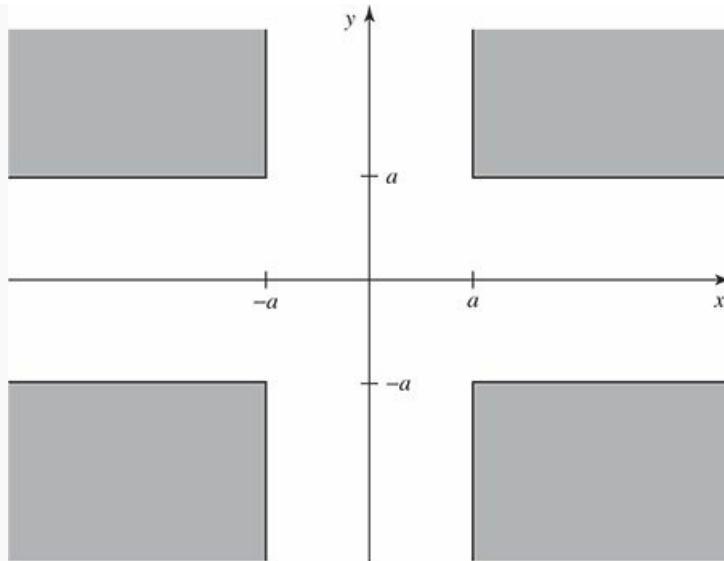


Figure 8.10: The cross-shaped region for Problem 8.27.

- (a) Show that the lowest energy that can propagate off to infinity is

$$E_{\text{threshold}} = \frac{\pi^2 \hbar^2}{8ma^2};$$

any solution with energy *less* than that has to be a bound state. *Hint:* Go way out one arm (say, $x \gg a$), and solve the Schrödinger equation by separation of variables; if the wave function propagates out to infinity, the dependence on x must take the form $\exp(i k_x x)$ with $k_x > 0$.

- (b) Now use the variational principle to show that the ground state has energy less than $E_{\text{threshold}}$. Use the following trial wave function (suggested by Jim McTavish):

$$\psi(x, y) = A \begin{cases} [\cos(\pi x/2a) + \cos(\pi y/2a)] e^{-\alpha}, & |x| \leq a \text{ and } |y| \leq a \\ \cos(\pi x/2a) e^{-\alpha|y|/a}, & |x| \leq a \text{ and } |y| > a \\ \cos(\pi y/2a) e^{-\alpha|x|/a}, & |x| > a \text{ and } |y| \leq a \\ 0, & \text{elsewhere.} \end{cases}$$

Normalize it to determine A , and calculate the expectation value of H .

Answer:

$$\langle H \rangle = \frac{\hbar^2}{ma^2} \left[\frac{\pi^2}{8} - \left(\frac{1 - (\alpha/4)}{1 + (8/\pi^2) + (1/2\alpha)} \right) \right].$$

Now minimize with respect to α , and show that the result is less than $E_{\text{threshold}}$. *Hint:* Take full advantage of the symmetry of the problem—you only need to integrate over 1/8 of the open region, since the other seven integrals will be the same. Note however that whereas the trial wave function is continuous, its *derivatives* are *not*—there are “roof-lines” at the joins, and you will need to exploit the technique of Example 8.3.[23](#)

Problem 8.28 In Yukawa's original theory (1934), which remains a useful approximation in nuclear physics, the "strong" force between protons and neutrons is mediated by the exchange of π -mesons. The potential energy is

$$V(r) = -r_0 V_0 \frac{e^{-r/r_0}}{r}, \quad (8.83)$$

where r is the distance between the nucleons, and the range r_0 is related to the mass of the meson: $r_0 = \hbar/m_\pi c$. *Question:* Does this theory account for the existence of the **deuteron** (a bound state of the proton and the neutron)?

The Schrödinger equation for the proton/neutron system is (see Problem 5.1):

$$-\frac{\hbar^2}{2\mu} \nabla^2 \psi(\mathbf{r}) + V(r) \psi(\mathbf{r}) = E \psi(\mathbf{r}), \quad (8.84)$$

where μ is the reduced mass (the proton and neutron have almost identical masses, so call them both m), and \mathbf{r} is the position of the neutron (say) relative to the proton: $\mathbf{r} = \mathbf{r}_n - \mathbf{r}_p$. Your task is to show that there exists a solution with negative energy (a bound state), using a variational trial wave function of the form

$$\psi_\beta(\mathbf{r}) = A e^{-\beta r/r_0}. \quad (8.85)$$

- (a) Determine A , by normalizing $\psi_\beta(\mathbf{r})$.
- (b) Find the expectation value of the Hamiltonian $\left(H = -\frac{\hbar^2}{2\mu} \nabla^2 + V \right)$ in the state ψ_β . *Answer:*

$$E(\beta) = \frac{\hbar^2}{2\mu r_0^2} \beta^2 \left[1 - \frac{4\gamma\beta}{(1+2\beta)^2} \right], \quad \text{where } \gamma \equiv \frac{2\mu r_0^2}{\hbar^2} V_0. \quad (8.86)$$

- (c) Optimize your trial wave function, by setting $dE(\beta)/d\beta = 0$. This tells you β as a function of γ (and hence—everything else being constant—of V_0), but let's use it instead to eliminate γ in favor of β :

$$E_{\min} = \frac{\hbar^2}{2\mu r_0^2} \frac{\beta^2 (1-2\beta)}{(3+2\beta)}. \quad (8.87)$$

- (d) Setting $\hbar^2/2\mu r_0^2 = 1$, plot E_{\min} as a function of β , for $0 \leq \beta \leq 1$. What does this tell you about the binding of the deuteron? What is the minimum value of V_0 for which you can be confident there is a bound state (look up the necessary masses)? The experimental value is 52 MeV.

Problem 8.29 Existence of Bound States. A potential "well" (in one dimension) is a function $V(x)$ that is never positive ($V(x) \leq 0$ for all x), and goes to zero at infinity ($V(x) \rightarrow 0$ as $x \rightarrow \pm\infty$).²⁴

- (a) Prove the following **Theorem**: If a potential well $V_1(x)$ supports at least one bound state, then any deeper/wider well ($V_2(x) \leq V_1(x)$ for all x) will also support at least one bound state. *Hint*: Use the ground state of V_1 , $\psi_1(x)$, as a variational test function.
- (b) Prove the following **Corollary**: Every potential well in one dimension has a bound state.²⁵ *Hint*: Use a finite square well (Section 2.6) for V_1 .
- (c) Does the Theorem generalize to two and three dimensions? How about the Corollary? *Hint*: You might want to review Problems 4.11 and 4.51.



Problem 8.30 Performing a variational calculation requires finding the minimum of the energy, as a function of the variational parameters. This is, in general, a very hard problem. However, if we choose the form of our trial wave function judiciously, we can develop an efficient algorithm. In particular, suppose we use a *linear* combination of functions $\phi_n(x)$:

$$\psi(x) = \sum_{n=1}^N c_n \phi_n(x), \quad (8.88)$$

where the c_n are the variational parameters. If the ϕ_n are an orthonormal set ($\langle \phi_m | \phi_n \rangle = \delta_{mn}$), but $\psi(x)$ is not necessarily normalized, then $\langle H \rangle$ is

$$\varepsilon = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\sum_{mn} c_m^* H_{mn} c_n}{\sum_n |c_n|^2} \quad (8.89)$$

where $H_{mn} = \langle \phi_m | H | \phi_n \rangle$. Taking the derivative with respect to c_j^* (and setting the result equal to 0) gives²⁶

$$\sum_n H_{jn} c_n = \varepsilon c_j, \quad (8.90)$$

recognizable as the j th row in an eigenvalue problem:

$$\begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1N} \\ H_{21} & H_{22} & \cdots & H_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} & H_{N2} & \cdots & H_{NN} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} = \varepsilon \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix}. \quad (8.91)$$

The smallest eigenvalue of this matrix \mathbf{H} gives a bound on the ground state energy and the corresponding eigenvector determines the best variational wave function of the form 8.88.

- (a) Verify Equation 8.90.
- (b) Now take the derivative of Equation 8.89 with respect to c_j and show that you get a result redundant with Equation 8.90.
- (c) Consider a particle in an infinite square well of width a , with a sloping floor:

$$V(x) = \begin{cases} \infty & x < 0, \\ V_0 x/a & 0 \leq x \leq a, \\ \infty & x > a \end{cases} .$$

Using a linear combination of the first ten stationary states of the infinite square well as the basis functions,

$$\phi_n = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right),$$

determine a bound for the ground state energy in the case $V_0 = 100\hbar^2/ma^2$. Make a plot of the optimized variational wave function. [Note: The exact result is $39.9819\hbar^2/ma^2$.]

- ¹ If the Hamiltonian admits scattering states, as well as bound states, then we'll need an integral as well as a sum, but the argument is unchanged.
- ² For a collection of interesting examples see W. N. Mei, *Int. J. Math. Educ. Sci. Tech.* **30**, 513 (1999).
- ³ There is no point in trying a function (such as the gaussian) that extends outside the well, because you'll get $\langle V \rangle = \infty$, and Equation 8.1 tells you nothing.
- ⁴ In practice this isn't much of a limitation, and there are sometimes ways of estimating the accuracy. The binding energy of helium has been calculated to many significant digits in this way (see for example G. W. Drake *et al.*, *Phys. Rev. A* **65**, 054501 (2002), or Vladimir I. Korobov, *Phys. Rev. A* **66**, 024501 (2002)).
- ⁵ For a systematic extension of the variational principle to the calculation of excited state energies see, for example, Linus Pauling and E. Bright Wilson, *Introduction to Quantum Mechanics, With Applications to Chemistry*, McGraw-Hill, New York (1935, paperback edition 1985), Section 26.
- ⁶ You can extend this trick to other symmetries. Suppose there is a Hermitian operator A such that $[A, H] = 0$. The ground state (assuming it is nondegenerate) must be an eigenstate of A ; call the eigenvalue λ : $A\psi_{\text{gs}} = \lambda\psi_{\text{gs}}$. If you choose a variational function ψ that is an eigenstate of A with a *different* eigenvalue: $A\psi = \nu\psi$ with $\lambda \neq \nu$, you can be certain that ψ and ψ_{gs} are orthogonal (see Section 3.3). For an application see Problem 8.20.
- ⁷ There do exist exactly soluble three-body problems with many of the qualitative features of helium, but using non-Coulombic potentials (see Problem 8.24).
- ⁸ Here a is the ordinary Bohr radius and $E_n = -13.6/n^2$ eV is the n th Bohr energy; recall that for a nucleus with atomic number Z , $E_n \rightarrow Z^2 E_n$ and $a \rightarrow a/Z$ (Problem 4.19). The spin configuration associated with Equation 8.18 will be antisymmetric (the singlet).
- ⁹ You can, if you like, interpret Equation 8.21 as first-order perturbation theory, with $H' = V_{ee}$ (Problem 7.56(a)). However, I regard this as a misuse of the method, since the perturbation is comparable in size to the unperturbed potential. I prefer, therefore, to think of it as a variational calculation, in which we are looking for a rigorous upper bound on E_{gs} .
- ¹⁰ The classic studies are E. A. Hylleraas, *Z. Phys.* **65**, 209 (1930); C. L. Pekeris, *Phys. Rev.* **115**, 1216 (1959). For more recent work, see footnote 4.
- ¹¹ The first excited state of helium can be calculated in much the same way, using a trial wave function orthogonal to the ground state. See Phillip J. E. Peebles, *Quantum Mechanics*, Princeton U.P., Princeton, NJ (1992), Section 40.
- ¹² Robert N. Hill, *J. Math. Phys.* **18**, 2316 (1977).
- ¹³ For further discussion see Hans A. Bethe and Edwin E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms*, Plenum, New York (1977), Section 34.
- ¹⁴ The wave function with the plus sign (Equation 8.38) is called the **bonding orbital**. Bonding is associated with a buildup of electron probability in between the two nuclei. The odd linear combination (Equation 8.53) has a *node* at the center, so it's not surprising that this configuration doesn't lead to bonding; it is called the **anti-bonding orbital**.
- ¹⁵ W. Heitler and F. London, *Z. Phys.* **44**, 455 (1928). For an English translation see Hinne Hettema, *Quantum Chemistry: Classic Scientific Papers*, World Scientific, New Jersey, PA, 2000.
- ¹⁶ Another natural variational wave function consists of placing both electrons in the bonding orbital studied in Section 8.3:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \left(\frac{\psi_0(r_1) + \psi_0(r'_1)}{\sqrt{2(1+I)}} \right) \left(\frac{\psi_0(r_2) + \psi_0(r'_2)}{\sqrt{2(1+I)}} \right), \quad (8.57)$$

also paired with a singlet spin state. If you expand this function you'll see that half the terms—such as $\psi_0(r_1)\psi_0(r_2)$ —involve attaching two electrons to the same proton, which is energetically costly because of the electron–electron repulsion in Equation 8.54. The Heitler–London approximation, Equation 8.55, amounts to dropping the offending terms from Equation 8.57.

¹⁷ The calculation was done by Y. Sugiura, *Z. Phys.* **44**, 455 (1927).

¹⁸ W. N. Mei, *Int. J. Educ. Sci. Tech.* **27**, 285 (1996).

¹⁹ For a more sophisticated model, see R. Crandall, R. Whitnell, and R. Bettega, *Am. J. Phys.* **52**, 438 (1984).

²⁰ S. Chandrasekhar, *Astrophys. J.* **100**, 176 (1944).

²¹ The classic paper on muon-catalyzed fusion is J. D. Jackson, *Phys. Rev.* **106**, 330 (1957); for a more recent popular review, see J. Rafelski and S. Jones, *Scientific American*, November 1987, page 84.

²² This model is taken from R. L. Schult *et al.*, *Phys. Rev. B* **39**, 5476 (1989). For further discussion see J. T. Lonergan and D. P. Murdock, *Am. J. Phys.* **80**, 1085 (2012). In the presence of quantum tunneling a classically bound state can become unbound; this is the reverse: A classically *unbound* state is quantum mechanically *bound*.

²³ W.-N. Mei gets a somewhat better bound (and avoids the roof-lines) using

$$\psi(x, y) = Ae^{-\alpha(x^2+y^2)/a^2} \begin{cases} \left(1 - x^2y^2/a^4\right), \\ \left(1 - x^2/a^2\right), \\ \left(1 - y^2/a^2\right), \end{cases}$$

but the integrals have to be done numerically.

²⁴ To exclude trivial cases, we also assume it has nonzero area ($\int V(x) dx \neq 0$). Notice that for the purposes of this problem neither the infinite square well nor the harmonic oscillator is a “potential well,” though both of them, of course, have bound states.

²⁵ K. R. Brownstein, *Am. J. Phys.* **68**, 160 (2000) proves that any one-dimensional potential satisfying $\int_{-\infty}^{\infty} V(x) dx \leq 0$ admits a bound state (as long as $V(x)$ is not identically zero)—even if it runs positive in some places.

²⁶ Each c_j , being complex, stands for two independent parameters (its real and imaginary parts). One *could* take derivatives with respect to the real and imaginary parts,

$$\frac{\partial}{\partial \operatorname{Re}[c_j]} E = 0 \quad \text{and} \quad \frac{\partial}{\partial \operatorname{Im}[c_j]} E = 0,$$

but it is also legitimate (and simpler) to treat c_j and c_j^* as the independent parameters:

$$\frac{\partial}{\partial c_j} E = 0 \quad \text{and} \quad \frac{\partial}{\partial c_j^*} E = 0.$$

You get the same result either way.

9

The WKB Approximation

◊

The WKB (Wentzel, Kramers, Brillouin)¹ method is a technique for obtaining approximate solutions to the time-independent Schrödinger equation in one dimension (the same basic idea can be applied to many other differential equations, and to the radial part of the Schrödinger equation in three dimensions). It is particularly useful in calculating bound state energies and tunneling rates through potential barriers.

The essential idea is as follows: Imagine a particle of energy E moving through a region where the potential $V(x)$ is *constant*. If $E > V$, the wave function is of the form

$$\psi(x) = Ae^{\pm ikx}, \quad \text{with} \quad k \equiv \sqrt{2m(E - V)/\hbar}.$$

The plus sign indicates that the particle is traveling to the right, and the minus sign means it is going to the left (the general solution, of course, is a linear combination of the two). The wave function is oscillatory, with fixed wavelength ($\lambda = 2\pi/k$) and unchanging amplitude (A). Now suppose that $V(x)$ is *not* constant, but varies rather slowly in comparison to λ , so that over a region containing many full wavelengths the potential is *essentially* constant. Then it is reasonable to suppose that ψ remains *practically* sinusoidal, except that the wavelength and the amplitude change slowly with x . This is the inspiration behind the WKB approximation. In effect, it identifies two different levels of x -dependence: rapid oscillations, *modulated* by gradual variation in amplitude and wavelength.

By the same token, if $E < V$ (and V is constant), then ψ is exponential:

$$\psi(x) = Ae^{\pm \kappa x}, \quad \text{with} \quad \kappa \equiv \sqrt{2m(V - E)/\hbar}.$$

And if $V(x)$ is *not* constant, but varies slowly in comparison with $1/\kappa$, the solution remains *practically* exponential, except that A and κ are now slowly-varying functions of x .

Now, there is one place where this whole program is bound to fail, and that is in the immediate vicinity of a classical **turning point**, where $E \approx V$. For here λ (or $1/\kappa$) goes to infinity, and $V(x)$ can hardly be said to vary “slowly” in comparison. As we shall see, a proper handling of the turning points is the most difficult aspect of the WKB approximation, though the final results are simple to state and easy to implement.

9.1 The “Classical” Region

The Schrödinger equation,

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

can be rewritten in the following way:

$$\frac{d^2\psi}{dx^2} = -\frac{p^2}{\hbar^2}\psi, \quad (9.1)$$

where

$$p(x) \equiv \sqrt{2m [E - V(x)]} \quad (9.2)$$

is the classical formula for the (magnitude of the) momentum of a particle with total energy E and potential energy $V(x)$. For the moment, I'll assume that $E > V(x)$, so that $p(x)$ is *real*; we call this the “classical” region, for obvious reasons—classically the particle is *confined* to this range of x (see Figure 9.1). In general, ψ is some complex function; we can express it in terms of its *amplitude*, $A(x)$, and its *phase*, $\phi(x)$ —both of which are *real*:

$$\psi(x) = A(x)e^{i\phi(x)}. \quad (9.3)$$

Using a prime to denote the derivative with respect to x ,

$$\frac{d\psi}{dx} = (A' + iA\phi')e^{i\phi},$$

and

$$\frac{d^2\psi}{dx^2} = [A'' + 2iA'\phi' + iA\phi'' - A(\phi')^2]e^{i\phi}. \quad (9.4)$$

Putting this into Equation 9.1:

$$A'' + 2iA'\phi' + iA\phi'' - A(\phi')^2 = -\frac{p^2}{\hbar^2}A. \quad (9.5)$$

This is equivalent to two *real* equations, one for the real part and one for the imaginary part:

$$A'' - A(\phi')^2 = -\frac{p^2}{\hbar^2}A, \quad \text{or} \quad A'' = A\left[(\phi')^2 - \frac{p^2}{\hbar^2}\right], \quad (9.6)$$

and

$$2A'\phi' + A\phi'' = 0, \quad \text{or} \quad (A^2\phi')' = 0. \quad (9.7)$$

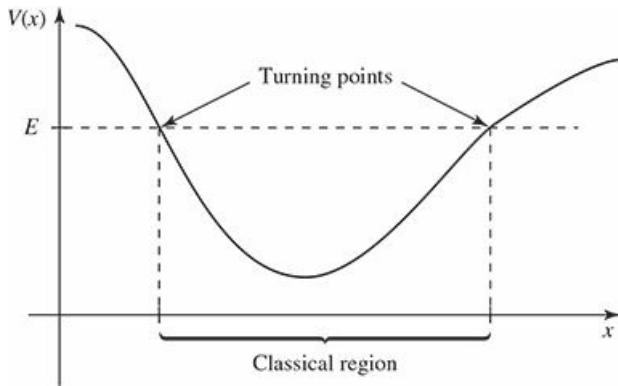


Figure 9.1: Classically, the particle is confined to the region where $E \geq V(x)$.

Equations 9.6 and 9.7 are entirely equivalent to the original Schrödinger equation. The second one is easily solved:

$$A^2 \phi' = C^2, \quad \text{or} \quad A = \frac{C}{\sqrt{|\phi'|}}, \quad (9.8)$$

where C is a (real) constant. The first one (Equation 9.6) cannot be solved in general—so here comes the approximation: *We assume that the amplitude A varies slowly*, so the A'' term is negligible. (More precisely, we assume that A''/A is much less than both $(\phi')^2$ and p^2/\hbar^2 .) In that case we can drop the left side of Equation 9.6, and we are left with

$$(\phi')^2 = \frac{p^2}{\hbar^2}, \quad \text{or} \quad \frac{d\phi}{dx} = \pm \frac{p}{\hbar},$$

and therefore

$$\phi(x) = \pm \frac{1}{\hbar} \int p(x) dx. \quad (9.9)$$

(I'll write this as an *indefinite* integral, for now—any constant of integration can be absorbed into C , which thereby becomes complex. I'll also absorb a factor of $\sqrt{\hbar}$.) Then

$$\psi(x) \approx \frac{C}{\sqrt{p(x)}} e^{\pm \frac{i}{\hbar} \int p(x) dx}.$$

(9.10)

Notice that

$$|\psi(x)|^2 \approx \frac{|C|^2}{p(x)}, \quad (9.11)$$

which says that the probability of finding the particle at point x is inversely proportional to its (classical) momentum (and hence its velocity) at that point. This is exactly what you would expect—the particle doesn't spend long in the places where it is moving rapidly, so the probability of getting caught there is small. In fact, the WKB approximation is sometimes *derived* by starting with this “semi-classical” observation, instead of by dropping the A'' term in the differential equation. The latter approach is cleaner mathematically, but the former offers a more illuminating physical rationale. The general (approximate) solution, of course, will be a linear combination the two solutions in Equation 9.10, one with each sign.

Example 9.1

Potential well with two vertical walls. Suppose we have an infinite square well with a bumpy bottom (Figure 9.2):

$$V(x) = \begin{cases} \text{some specified function,} & (0 < x < a), \\ \infty, & (\text{otherwise}). \end{cases} \quad (9.12)$$

Inside the well (assuming $E > V(x)$ throughout) we have

$$\psi(x) \approx \frac{1}{\sqrt{p(x)}} [C_+ e^{i\phi(x)} + C_- e^{-i\phi(x)}], \quad (9.13)$$

or, more conveniently,

$$\psi(x) \approx \frac{1}{\sqrt{p(x)}} [C_1 \sin \phi(x) + C_2 \cos \phi(x)], \quad (9.14)$$

where (exploiting the freedom noted earlier to impose a convenient lower limit on the integral)²

$$\phi(x) = \frac{1}{\hbar} \int_0^x p(x') dx'. \quad (9.15)$$

Now, $\psi(x)$ must go to zero at $x = 0$, and therefore (since $\phi(0) = 0$) $C_2 = 0$. Also, $\psi(x)$ goes to zero at $x = a$, so

$$\phi(a) = n\pi \quad (n = 1, 2, 3, \dots). \quad (9.16)$$

Conclusion:

$$\int_0^a p(x) dx = n\pi\hbar. \quad (9.17)$$

This quantization condition determines the (approximate) allowed energies.

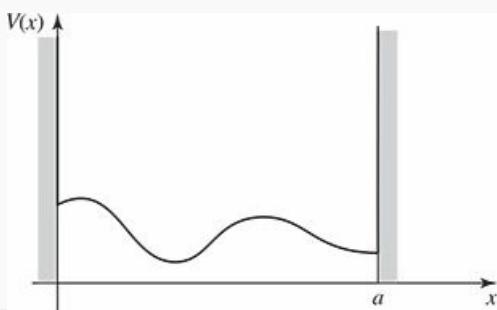


Figure 9.2: Infinite square well with a bumpy bottom.

For instance, if the well has a *flat* bottom ($V(x) = 0$), then $p(x) = \sqrt{2mE}$ (a constant), and Equation 9.17 says $pa = n\pi\hbar$, or

$$E_n = \frac{n^2\pi^2\hbar^2}{2ma^2},$$

which is the old formula for the energy levels of the infinite square well (Equation 2.30). In this case the WKB approximation yields the *exact* answer (the amplitude of the true wave function is *constant*, so dropping A'' cost us nothing).

- * **Problem 9.1** Use the WKB approximation to find the allowed energies (E_n) of an infinite square well with a “shelf,” of height V_0 , extending half-way across (Figure 7.3):

$$V(x) = \begin{cases} V_0, & (0 < x < a/2), \\ 0, & (a/2 < x < a), \\ \infty, & (\text{otherwise}). \end{cases}$$

Express your answer in terms of V_0 and $E_n^0 \equiv (n\pi\hbar)^2 / 2ma^2$ (the n th allowed energy for the infinite square well with *no* shelf). Assume that $E_1^0 > V_0$, but do *not* assume that $E_n \gg V_0$. Compare your result with what we got in Section 7.1.2, using first-order perturbation theory. Note that they are in agreement if either V_0 is very small (the perturbation theory regime) or n is very large (the WKB—semi-classical—regime).

**

- Problem 9.2** An alternative derivation of the WKB formula (Equation 9.10) is based on an expansion in powers of \hbar . Motivated by the free-particle wave function, $\psi = A \exp(\pm ipx/\hbar)$, we write

$$\psi(x) = e^{if(x)/\hbar},$$

where $f(x)$ is some complex function. (Note that there is no loss of generality here—*any* nonzero function can be written in this way.)

- (a) Put this into Schrödinger’s equation (in the form of Equation 9.1), and show that

$$i\hbar f'' - (f')^2 + p^2 = 0.$$

- (b) Write $f(x)$ as a power series in \hbar :

$$f(x) = f_0(x) + \hbar f_1(x) + \hbar^2 f_2(x) + \dots,$$

and, collecting like powers of \hbar , show that

$$(f'_0)^2 = p^2, \quad i f''_0 = 2 f'_0 f'_1, \quad i f''_1 = 2 f'_0 f''_2 + (f'_1)^2, \quad \text{etc.}$$

- (c) Solve for $f_0(x)$ and $f_1(x)$, and show that—to first order in \hbar —you recover Equation 9.10.

Note: The logarithm of a negative number is defined by $\ln(-z) = \ln(z) + in\pi$, where n is an odd integer. If this formula is new to you, try exponentiating both sides, and you’ll see where it comes from.

9.2 Tunneling

So far, I have assumed that $E > V$, so $p(x)$ is real. But we can easily write down the corresponding result in the *non-classical* region ($E < V$)—it's the same as before (Equation 9.10), only now $p(x)$ is *imaginary*:³

$$\psi(x) \approx \frac{C}{\sqrt{|p(x)|}} e^{\pm \frac{1}{\hbar} \int |p(x)| dx}. \quad (9.18)$$

Consider, for example, the problem of scattering from a rectangular barrier with a bumpy top (Figure 9.3). To the left of the barrier ($x < 0$),

$$\psi(x) = A e^{ikx} + B e^{-ikx}, \quad (9.19)$$

where A is the incident amplitude, B is the reflected amplitude, and $k \equiv \sqrt{2mE}/\hbar$ (see Section 2.5). To the right of the barrier ($x > a$),

$$\psi(x) = F e^{ikx}, \quad (9.20)$$

where F is the transmitted amplitude. The transmission probability is

$$T = \frac{|F|^2}{|A|^2}. \quad (9.21)$$

In the tunneling region ($0 \leq x \leq a$), the WKB approximation gives

$$\psi(x) \approx \frac{C}{\sqrt{|p(x)|}} e^{\frac{1}{\hbar} \int_0^x |p(x')| dx'} + \frac{D}{\sqrt{|p(x)|}} e^{-\frac{1}{\hbar} \int_0^x |p(x')| dx'}. \quad (9.22)$$

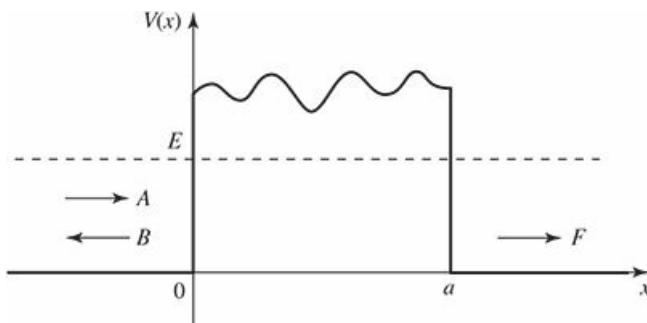


Figure 9.3: Scattering from a rectangular barrier with a bumpy top.

If the barrier is very high and/or very wide (which is to say, if the probability of tunneling is small), then the coefficient of the exponentially *increasing* term (C) must be small (in fact, it would be *zero* if the barrier were *infinitely* broad), and the wave function looks something like⁴ Figure 9.4. The relative amplitudes of the incident and transmitted waves are determined essentially by the total decrease of the exponential over the nonclassical region:

$$\frac{|F|}{|A|} \sim e^{-\frac{1}{\hbar} \int_0^a |p(x')| dx'},$$

so

$$T \approx e^{-2\gamma}, \quad \text{with} \quad \gamma \equiv \frac{1}{\hbar} \int_0^a |p(x)| dx. \quad (9.23)$$

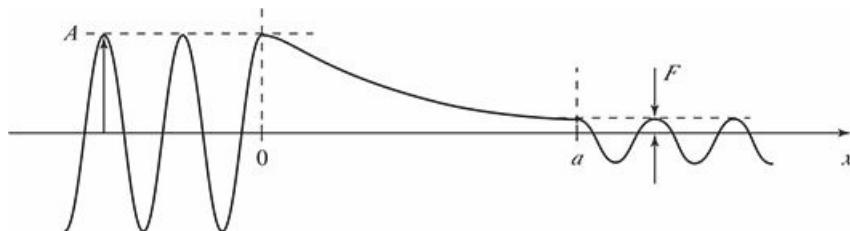


Figure 9.4: Qualitative structure of the wave function, for scattering from a high, broad barrier.

Example 9.2

Gamow's theory of alpha decay.⁵ In 1928, George Gamow (and, independently, Condon and Gurney) used Equation 9.23 to provide the first successful explanation of alpha decay (the spontaneous emission of an alpha particle—two protons and two neutrons—by certain radioactive nuclei).⁶ Since the alpha particle carries a positive charge ($2e$), it will be electrically repelled by the leftover nucleus (charge Ze), as soon as it gets far enough away to escape the nuclear binding force. But first it has to negotiate a potential barrier that was already known (in the case of uranium) to be more than twice the energy of the emitted alpha particle. Gamow approximated the potential energy by a finite square well (representing the attractive nuclear force), extending out to r_1 (the radius of the nucleus), joined to a repulsive Coulombic tail (Figure 9.5), and identified the escape mechanism as quantum tunneling (this was, by the way, the first time that quantum mechanics had been applied to nuclear physics).

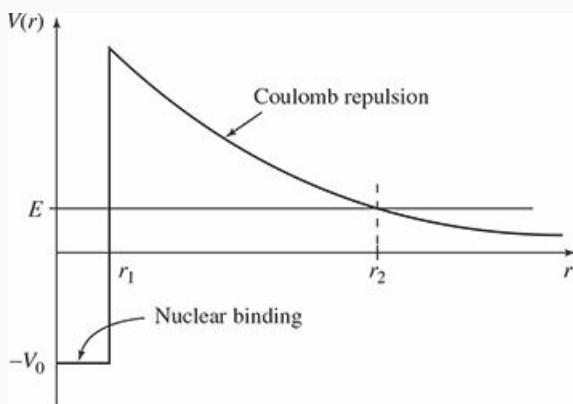


Figure 9.5: Gamow's model for the potential energy of an alpha particle in a radioactive nucleus.

If E is the energy of the emitted alpha particle, the outer turning point (r_2) is determined by

$$\frac{1}{4\pi\epsilon_0} \frac{2Ze^2}{r_2} = E. \quad (9.24)$$

The exponent γ (Equation 9.23) is evidently⁷

$$\gamma = \frac{1}{\hbar} \int_{r_1}^{r_2} \sqrt{2m \left(\frac{1}{4\pi\epsilon_0} \frac{2Ze^2}{r} - E \right)} dr = \frac{\sqrt{2mE}}{\hbar} \int_{r_1}^{r_2} \sqrt{\frac{r_2}{r} - 1} dr.$$

The integral can be done by substitution ($let r \equiv r_2 \sin^2 u$), and the result is

$$\gamma = \frac{\sqrt{2mE}}{\hbar} \left[r_2 \left(\frac{\pi}{2} - \sin^{-1} \sqrt{\frac{r_1}{r_2}} \right) - \sqrt{r_1(r_2 - r_1)} \right]. \quad (9.25)$$

Typically, $r_1 \ll r_2$, and we can simplify this result using the small angle approximation ($\sin \epsilon \approx \epsilon$):

$$\gamma \approx \frac{\sqrt{2mE}}{\hbar} \left[\frac{\pi}{2} r_2 - 2\sqrt{r_1 r_2} \right] = K_1 \frac{Z}{\sqrt{E}} - K_2 \sqrt{Z r_1}, \quad (9.26)$$

where

$$K_1 \equiv \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{\pi \sqrt{2m}}{\hbar} = 1.980 \text{ MeV}^{1/2}, \quad (9.27)$$

and

$$K_2 \equiv \left(\frac{e^2}{4\pi\epsilon_0} \right)^{1/2} \frac{4\sqrt{m}}{\hbar} = 1.485 \text{ fm}^{-1/2}. \quad (9.28)$$

(One fermi (fm) is 10^{-15} m, which is about the size of a typical nucleus.)

If we imagine the alpha particle rattling around inside the nucleus, with an average velocity v , the time between “collisions” with the “wall” is about $2r_1/v$, and hence the *frequency* of collisions is $v/2r_1$. The probability of escape at each collision is $e^{-2\gamma}$, so the probability of emission, per unit time, is $(v/2r_1) e^{-2\gamma}$, and hence the *lifetime* of the parent nucleus is about

$$\tau = \frac{2r_1}{v} e^{2\gamma}. \quad (9.29)$$

Unfortunately, we don’t know v —but it hardly matters, for the exponential factor varies over a *fantastic* range (twenty-five orders of magnitude), as we go from one radioactive nucleus to another; relative to this the variation in v is pretty insignificant. In particular, if you plot the *logarithm* of the experimentally measured lifetime against $1/\sqrt{E}$, the result is a beautiful straight line (Figure 9.6),⁸ just as you would expect from Equations 9.26 and 9.29.

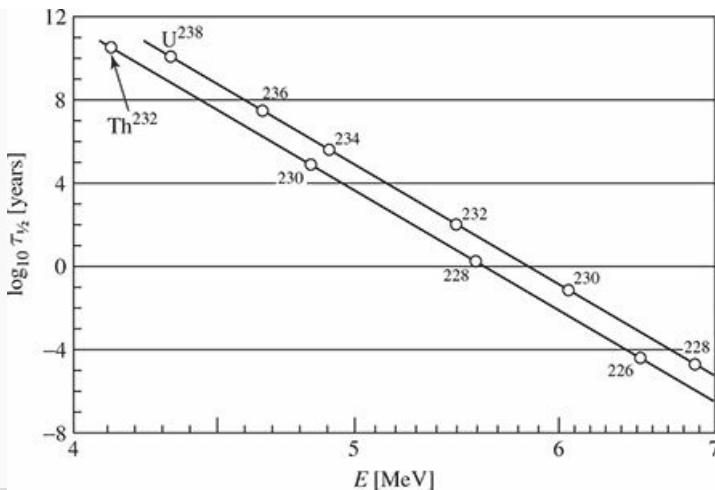


Figure 9.6: Graph of the logarithm of the half-life ($\tau_{1/2} = \tau \ln 2$) versus $1/\sqrt{E}$ (where E is the energy of the emitted alpha particle), for isotopes of uranium and thorium.

* **Problem 9.3** Use Equation 9.23 to calculate the approximate transmission probability for a particle of energy E that encounters a finite square barrier of height $V_0 > E$ and width $2a$. Compare your answer with the exact result (Problem 2.33), to which it should reduce in the WKB regime $T \ll 1$.

** **Problem 9.4** Calculate the lifetimes of U^{238} and Po^{212} , using Equations 9.26 and 9.29. Hint: The density of nuclear matter is relatively constant (i.e. the same for all nuclei), so $(r_1)^3$ is proportional to A (the number of neutrons plus protons). Empirically,

$$r_1 \cong (1.07 \text{ fm}) A^{1/3}. \quad (9.30)$$

The energy of the emitted alpha particle can be deduced by using Einstein's formula ($E = mc^2$):

$$E = m_p c^2 - m_d c^2 - m_\alpha c^2, \quad (9.31)$$

where m_p is the mass of the parent nucleus, m_d is the mass of the daughter nucleus, and m_α is the mass of the alpha particle (which is to say, the He^4 nucleus). To figure out what the daughter nucleus is, note that the alpha particle carries off two protons and two neutrons, so Z decreases by 2 and A by 4. Look up the relevant nuclear masses. To estimate v , use $E = (1/2) m_\alpha v^2$; this ignores the (negative) potential energy inside the nucleus, and surely underestimates v , but it's about the best we can do at this stage. Incidentally, the experimental lifetimes are 6×10^9 yrs and $0.5 \mu\text{s}$, respectively.

Problem 9.5 Zener Tunneling. In a semiconductor, an electric field (if it's large enough) can produce transitions between energy bands—a phenomenon known as Zener tunneling. A uniform electric field $\mathbf{E} = -E_0 \hat{i}$, for which

$$H' = -eE_0x,$$

makes the energy bands position dependent, as shown in Figure 9.7. It is then possible for an electron to tunnel from the valence (lower) band to the conduction (upper) band; this phenomenon is the basis for the **Zener diode**. Treating the gap as a potential barrier through which the electron may tunnel, find the tunneling probability in terms of E_g and E_0 (as well as m , \hbar , e).

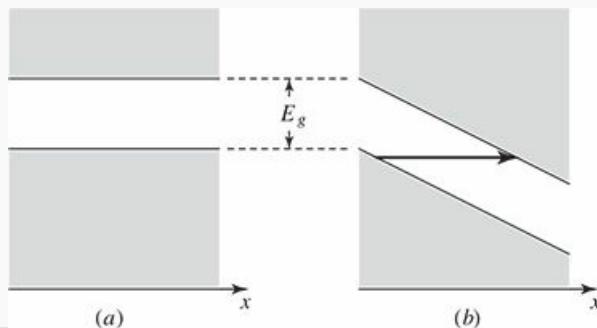


Figure 9.7: (a) The energy bands in the absence of an electric field. (b) In the presence of an electric field an electron can tunnel between the energy bands.

9.3 The Connection Formulas

In the discussion so far I have assumed that the “walls” of the potential well (or the barrier) are *vertical*, so that the “exterior” solution is simple, and the boundary conditions trivial. As it turns out, our main results (Equations 9.17 and 9.23) are reasonably accurate even when the edges are not so abrupt (indeed, in Gamow’s theory they were applied to just such a case). Nevertheless, it is of some interest to study more closely what happens to the wave function at a turning point ($E = V$), where the “classical” region joins the “nonclassical” region, and the WKB approximation itself breaks down. In this section I’ll treat the bound state problem (Figure 9.1); you get to do the scattering problem for yourself (Problem 9.11).⁹

For simplicity, let’s shift the axes over so that the right hand turning point occurs at $x = 0$ (Figure 9.8). In the WKB approximation, we have

$$\psi(x) \approx \begin{cases} \frac{1}{\sqrt{p(x)}} \left[Be^{\frac{i}{\hbar} \int_x^0 p(x') dx'} + Ce^{-\frac{i}{\hbar} \int_x^0 p(x') dx'} \right], & x < 0, \\ \frac{1}{\sqrt{|p(x)|}} De^{-\frac{1}{\hbar} \int_0^x |p(x')| dx'}, & x > 0. \end{cases} \quad (9.32)$$

(Assuming $V(x)$ remains greater than E for *all* $x > 0$, we can exclude the positive exponent in this region, because it blows up as $x \rightarrow \infty$.) Our task is to join the two solutions at the boundary. But there is a serious difficulty here: In the WKB approximation, ψ goes to *infinity* at the turning point (where $p(x) \rightarrow 0$). The *true* wave function, of course, has no such wild behavior—as anticipated, the WKB method simply fails in the vicinity of a turning point. And yet, it is precisely the boundary conditions at the turning points that determine the allowed energies. What we need to do, then, is *splice* the two WKB solutions together, using a “patching” wave function that straddles the turning point.

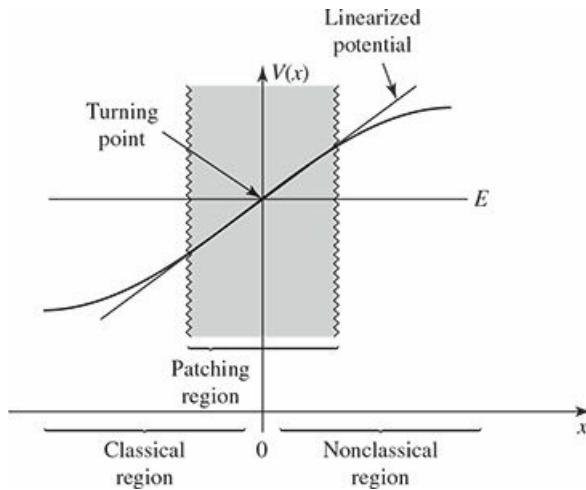


Figure 9.8: Enlarged view of the right-hand turning point.

Since we only need the patching wave function (ψ_p) in the neighborhood of the origin, we’ll *approximate the potential by a straight line*:

$$V(x) \approx E + V'(0)x, \quad (9.33)$$

and solve the Schrödinger equation for this linearized V :

$$-\frac{\hbar^2}{2m} \frac{d^2\psi_p}{dx^2} + (E + V'(0)x)\psi_p = E\psi_p,$$

or

$$\frac{d^2\psi_p}{dx^2} = \alpha^3 x \psi_p, \quad (9.34)$$

where

$$\alpha \equiv \left[\frac{2m}{\hbar^2} V'(0) \right]^{1/3}. \quad (9.35)$$

The α s can be absorbed into the independent variable by defining

$$z \equiv \alpha x, \quad (9.36)$$

so that

$$\frac{d^2\psi_p}{dz^2} = z\psi_p. \quad (9.37)$$

This is **Airy's equation**, and the solutions are called **Airy functions**.¹⁰ Since the Airy equation is a second-order differential equation, there are two linearly independent Airy functions, $\text{Ai}(z)$ and $\text{Bi}(z)$. They are related to Bessel functions of order 1/3; some of their properties are listed in Table 9.1 and they are plotted in Figure 9.9. Evidently the patching wave function is a linear combination of $\text{Ai}(z)$ and $\text{Bi}(z)$:

$$\psi_p(x) = a\text{Ai}(\alpha x) + b\text{Bi}(\alpha x), \quad (9.38)$$

for appropriate constants a and b .

Table 9.1: Some properties of the Airy functions.

<i>Differential Equation:</i>	$\frac{d^2y}{dz^2} = zy.$
<i>Solutions:</i>	Linear combinations of Airy functions, $\text{Ai}(z)$ and $\text{Bi}(z)$.
<i>Integral Representation:</i>	$\text{Ai}(z) = \frac{1}{\pi} \int_0^\infty \cos \left(\frac{s^3}{3} + sz \right) ds,$ $\text{Bi}(z) = \frac{1}{\pi} \int_0^\infty \left[e^{-\frac{s^3}{3} + sz} + \sin \left(\frac{s^3}{3} + sz \right) \right] ds.$
<i>Asymptotic Forms:</i>	$\begin{aligned} \text{Ai}(z) &\sim \frac{1}{2\sqrt{\pi}z^{1/4}} e^{-\frac{2}{3}z^{3/2}} & \text{Ai}(z) &\sim \frac{1}{\sqrt{\pi}(-z)^{1/4}} \sin \left[\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4} \right] \\ \text{Bi}(z) &\sim \frac{1}{\sqrt{\pi}z^{1/4}} e^{\frac{2}{3}z^{3/2}} & \text{Bi}(z) &\sim \frac{1}{\sqrt{\pi}(-z)^{1/4}} \cos \left[\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4} \right] \end{aligned} \quad \left. \begin{array}{l} z \gg 0; \\ z \ll 0. \end{array} \right\}$

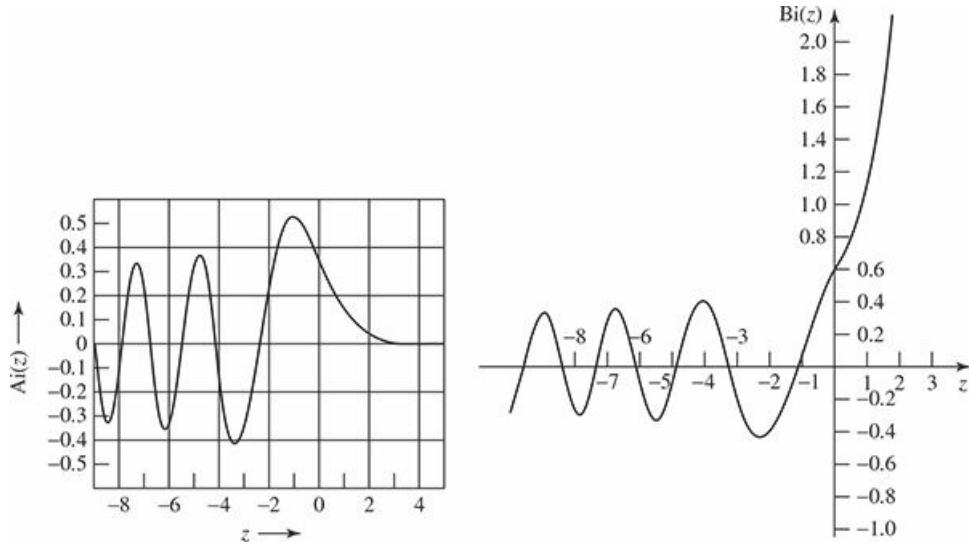


Figure 9.9: Graph of the Airy functions.

Now, ψ_p is the (approximate) wave function in the neighborhood of the origin; our job is to match it to the WKB solutions in the overlap regions on either side (see Figure 9.10). These overlap zones are close enough to the turning point that the linearized potential is reasonably accurate (so that ψ_p is a good approximation to the true wave function), and yet far enough away from the turning point that the WKB approximation is reliable.¹¹ In the overlap regions Equation 9.33 holds, and therefore (in the notation of Equation 9.35)

$$p(x) \approx \sqrt{2m(E - E - V'(0)x)} = \hbar\alpha^{3/2}\sqrt{-x}. \quad (9.39)$$

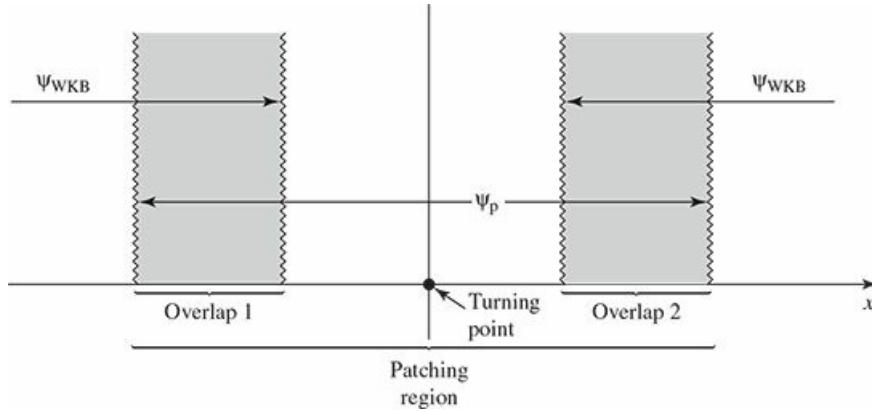


Figure 9.10: Patching region and the two overlap zones.

In particular, in overlap region 2,

$$\int_0^x |p(x')| dx' \approx \hbar\alpha^{3/2} \int_0^x \sqrt{x'} dx' = \frac{2}{3} \hbar(\alpha x)^{3/2},$$

and therefore the WKB wave function (Equation 9.32) can be written as

$$\psi(x) \approx \frac{D}{\sqrt{\hbar\alpha^{3/4}x^{1/4}}} e^{-\frac{2}{3}(\alpha x)^{3/2}}. \quad (9.40)$$

Meanwhile, using the large- z asymptotic forms¹² of the Airy functions (from Table 9.1), the patching wave function (Equation 9.38) in overlap region 2 becomes

$$\psi_p(x) \approx \frac{a}{2\sqrt{\pi}(\alpha x)^{1/4}} e^{-\frac{2}{3}(\alpha x)^{3/2}} + \frac{b}{\sqrt{\pi}(\alpha x)^{1/4}} e^{\frac{2}{3}(\alpha x)^{3/2}}. \quad (9.41)$$

Comparing the two solutions, we see that

$$a = \sqrt{\frac{4\pi}{\alpha\hbar}} D, \quad \text{and} \quad b = 0. \quad (9.42)$$

Now we go back and repeat the procedure for overlap region 1. Once again, $p(x)$ is given by Equation 9.39, but this time x is *negative*, so

$$\int_x^0 p(x') dx' \approx \frac{2}{3}\hbar(-\alpha x)^{3/2} \quad (9.43)$$

and the WKB wave function (Equation 9.32) is

$$\psi(x) \approx \frac{1}{\sqrt{\hbar\alpha^{3/4}(-x)^{1/4}}} \left[B e^{i\frac{2}{3}(-\alpha x)^{3/2}} + C e^{-i\frac{2}{3}(-\alpha x)^{3/2}} \right]. \quad (9.44)$$

Meanwhile, using the asymptotic form of the Airy function for large *negative* z (Table 9.1), the patching function (Equation 9.38, with $b = 0$) reads

$$\begin{aligned} \psi_p(x) &\approx \frac{a}{\sqrt{\pi}(-\alpha x)^{1/4}} \sin \left[\frac{2}{3}(-\alpha x)^{3/2} + \frac{\pi}{4} \right] \\ &= \frac{a}{\sqrt{\pi}(-\alpha x)^{1/4}} \frac{1}{2i} \left[e^{i\pi/4} e^{i\frac{2}{3}(-\alpha x)^{3/2}} - e^{-i\pi/4} e^{-i\frac{2}{3}(-\alpha x)^{3/2}} \right]. \end{aligned} \quad (9.45)$$

Comparing the WKB and patching wave functions in overlap region 1, we find

$$\frac{a}{2i\sqrt{\pi}} e^{i\pi/4} = \frac{B}{\sqrt{\hbar\alpha}} \quad \text{and} \quad \frac{-a}{2i\sqrt{\pi}} e^{-i\pi/4} = \frac{C}{\sqrt{\hbar\alpha}},$$

or, putting in Equation 9.42 for a :

$$B = -ie^{i\pi/4}D, \quad \text{and} \quad C = ie^{-i\pi/4}D. \quad (9.46)$$

These are the so-called **connection formulas**, joining the WKB solutions at either side of the turning point. We're done with the patching wave function now—its only purpose was to bridge the gap. Expressing everything in terms of the one normalization constant D , and shifting the turning point back from the origin to an arbitrary point x_2 , the WKB wave function (Equation 9.32) becomes

$$\psi(x) \approx \begin{cases} \frac{2D}{\sqrt{p(x)}} \sin \left[\frac{1}{\hbar} \int_x^{x_2} p(x') dx' + \frac{\pi}{4} \right], & x < x_2; \\ \frac{D}{\sqrt{|p(x)|}} \exp \left[-\frac{1}{\hbar} \int_{x_2}^x |p(x')| dx' \right], & x > x_2. \end{cases} \quad (9.47)$$

Example 9.3

Potential well with one vertical wall. Imagine a potential well that has one vertical side (at $x = 0$) and

one sloping side (Figure 9.11). In this case $\psi(0) = 0$, so Equation 9.47 says

$$\frac{1}{\hbar} \int_0^{x_2} p(x) dx + \frac{\pi}{4} = n\pi, \quad (n = 1, 2, 3, \dots),$$

or

$$\int_0^{x_2} p(x) dx = \left(n - \frac{1}{4}\right) \pi \hbar. \quad (9.48)$$

For instance, consider the “half-harmonic oscillator”,

$$V(x) = \begin{cases} \frac{1}{2}m\omega^2x^2, & (x > 0), \\ \infty, & (\text{otherwise}). \end{cases} \quad (9.49)$$

In this case

$$p(x) = \sqrt{2m [E - (1/2)m\omega^2x^2]} = m\omega\sqrt{x_2^2 - x^2},$$

where

$$x_2 = \frac{1}{\omega} \sqrt{\frac{2E}{m}}$$

is the turning point. So

$$\int_0^{x_2} p(x) dx = m\omega \int_0^{x_2} \sqrt{x_2^2 - x^2} dx = \frac{\pi}{4} m\omega x_2^2 = \frac{\pi E}{2\omega},$$

and the quantization condition (Equation 9.48) yields

$$E_n = \left(2n - \frac{1}{2}\right) \hbar\omega = \left(\frac{3}{2}, \frac{7}{2}, \frac{11}{2}, \dots\right) \hbar\omega. \quad (9.50)$$

In this particular case the WKB approximation actually delivers the *exact* allowed energies (which are precisely the *odd* energies of the *full* harmonic oscillator—see Problem 2.41).

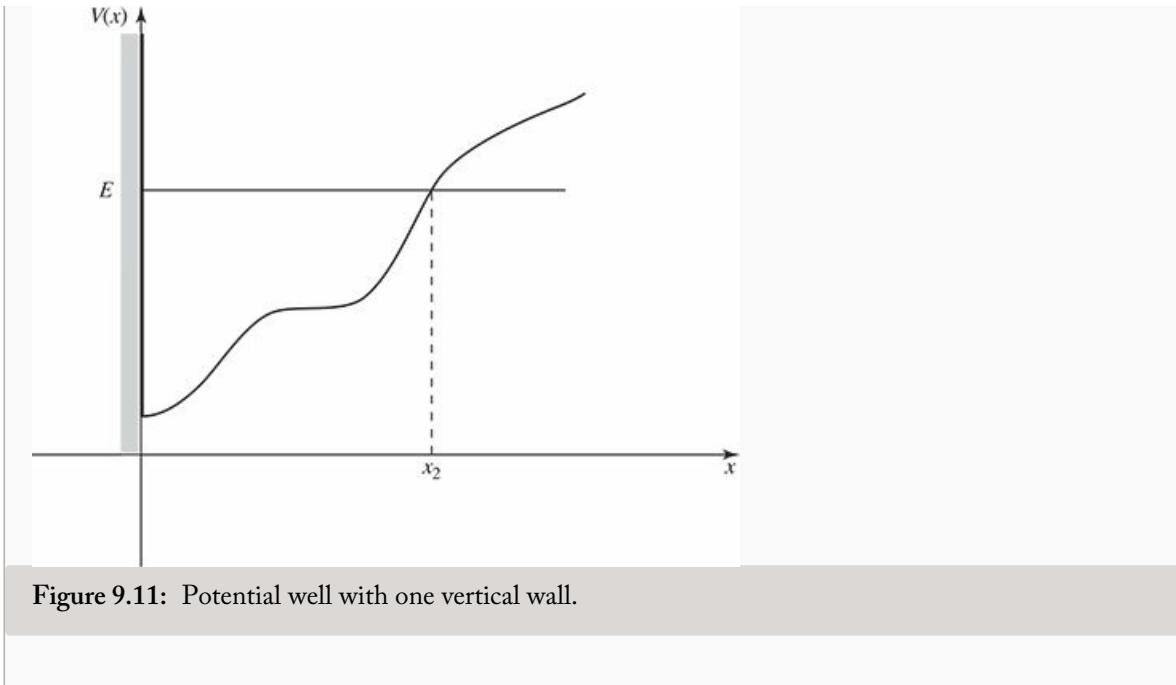


Figure 9.11: Potential well with one vertical wall.

Example 9.4

Potential well with no vertical walls. Equation 9.47 connects the WKB wave functions at a turning point where the potential slopes *upward* (Figure 9.12(a)); the same reasoning, applied to a *downward*-sloping turning point (Figure 9.12(b)), yields (Problem 9.10)

$$\psi(x) \approx \begin{cases} \frac{D'}{\sqrt{|p(x)|}} \exp \left[-\frac{1}{\hbar} \int_x^{x_1} |p(x')| dx' \right], & x < x_1; \\ \frac{2D'}{\sqrt{p(x)}} \sin \left[\frac{1}{\hbar} \int_{x_1}^x p(x') dx' + \frac{\pi}{4} \right], & x > x_1. \end{cases} \quad (9.51)$$

In particular, if we're talking about a potential *well* (Figure 9.12(c)), the wave function in the “interior” region ($x_1 < x < x_2$) can be written *either* as

$$\psi(x) \approx \frac{2D}{\sqrt{p(x)}} \sin \theta_2(x), \quad \text{where } \theta_2(x) \equiv \frac{1}{\hbar} \int_x^{x_2} p(x') dx' + \frac{\pi}{4},$$

(Equation 9.47), or as

$$\psi(x) \approx \frac{-2D'}{\sqrt{p(x)}} \sin \theta_1(x), \quad \text{where } \theta_1(x) \equiv -\frac{1}{\hbar} \int_{x_1}^x p(x') dx' - \frac{\pi}{4},$$

(Equation 9.51). Evidently the arguments of the sine functions must be equal, modulo π :¹³ $\theta_2 = \theta_1 + n\pi$, from which it follows that

$$\int_{x_1}^{x_2} p(x) dx = \left(n - \frac{1}{2} \right) \pi \hbar, \quad \text{with } n = 1, 2, 3, \dots \quad (9.52)$$

This quantization condition determines the allowed energies for the “typical” case of a potential well with two sloping sides. Notice that it differs from the formulas for two vertical walls (Equation 9.17) or one vertical wall (Equation 9.48) only in the number that is subtracted from n (0, 1/4, or 1/2).

Since the WKB approximation works best in the semi-classical (large n) regime, the distinction is more in appearance than in substance. In any event, the result is extraordinarily powerful, for it enables us to calculate (approximate) allowed energies *without ever solving the Schrödinger equation*, by simply evaluating one integral. The wave function itself has dropped out of sight.

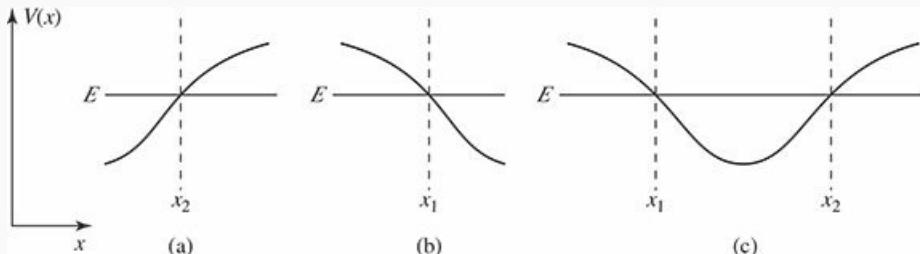


Figure 9.12: Upward-sloping and downward-sloping turning points.

**

Problem 9.6 The “bouncing ball” revisited. Consider the quantum mechanical analog to the classical problem of a ball (mass m) bouncing elastically on the floor.¹⁴

- What is the potential energy, as a function of height x above the floor? (For negative x , the potential is *infinite*—the ball can’t get there at all.)
- Solve the Schrödinger equation for this potential, expressing your answer in terms of the appropriate Airy function (note that $\text{Bi}(z)$ blows up for large z , and must therefore be rejected). Don’t bother to normalize $\psi(x)$.
- Using $g = 9.80 \text{ m/s}^2$ and $m = 0.100 \text{ kg}$, find the first four allowed energies, in joules, correct to three significant digits. *Hint:* See Milton Abramowitz and Irene A. Stegun, *Handbook of Mathematical Functions*, Dover, New York (1970), page 478; the notation is defined on page 450.
- What is the ground state energy, in eV, of an *electron* in this gravitational field? How high off the ground is this electron, on the average? *Hint:* Use the virial theorem to determine $\langle x \rangle$.

*

Problem 9.7 Analyze the bouncing ball (Problem 9.6) using the WKB approximation.

- Find the allowed energies, E_n , in terms of m , g , and \hbar .
- Now put in the particular values given in Problem 9.6(c), and compare the WKB approximation to the first four energies with the “exact” results.
- About how large would the quantum number n have to be to give the ball an average height of, say, 1 meter above the ground?

- * **Problem 9.8** Use the WKB approximation to find the allowed energies of the harmonic oscillator.

Problem 9.9 Consider a particle of mass m in the n th stationary state of the harmonic oscillator (angular frequency ω).

- Find the turning point, x_2 .
- How far (d) could you go *above* the turning point before the error in the linearized potential (Equation 9.33, but with the turning point at x_2) reaches 1%? That is, if

$$\frac{V(x_2 + d) - V_{\text{lin}}(x_2 + d)}{V(x_2)} = 0.01,$$

what is d ?

- The asymptotic form of $\text{Ai}(z)$ is accurate to 1% as long as $z \geq 5$. For the d in part (b), determine the smallest n such that $\alpha d \geq 5$. (For any n larger than this there exists an overlap region in which the linearized potential is good to 1% *and* the large- z form of the Airy function is good to 1%).

**

Problem 9.10 Derive the connection formulas at a downward-sloping turning point, and confirm Equation 9.51.

Problem 9.11 Use appropriate connection formulas to analyze the problem of scattering from a barrier with sloping walls (Figure 9.13). Hint: Begin by writing the WKB wave function in the form

$$\psi(x) \approx \begin{cases} \frac{1}{\sqrt{|p(x)|}} \left[A e^{-\frac{i}{\hbar} \int_{x_1}^x p(x') dx'} + B e^{\frac{i}{\hbar} \int_{x_1}^x p(x') dx'} \right], & x < x_1; \\ \frac{1}{\sqrt{|p(x)|}} \left[C e^{\frac{i}{\hbar} \int_{x_1}^x |p(x')| dx'} + D e^{-\frac{i}{\hbar} \int_{x_1}^x |p(x')| dx'} \right], & x_1 < x < x_2; \\ \frac{1}{\sqrt{|p(x)|}} \left[F e^{\frac{i}{\hbar} \int_{x_2}^x p(x') dx'} \right], & x > x_2. \end{cases} \quad (9.53)$$

Do *not* assume $C = 0$. Calculate the tunneling probability, $T = |F|^2/|A|^2$, and show that your result reduces to Equation 9.23 in the case of a broad, high barrier.

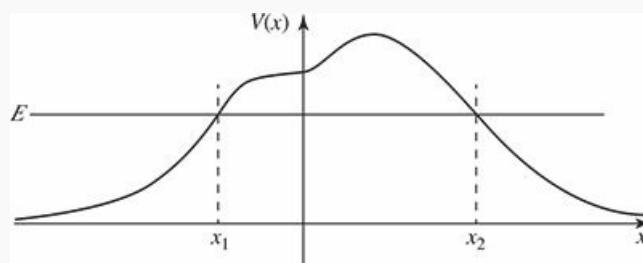


Figure 9.13: Barrier with sloping walls.



Problem 9.12 For the “half-harmonic oscillator” (Example 9.3), make a plot comparing the normalized WKB wave function for $n = 3$ to the exact solution. You’ll have to experiment to determine how wide to make the patching region. *Note:* You *can* do the integrals of $p(x)$ by hand, but feel free to do them numerically. You’ll need to do the integral of $|\psi_{WKB}|^2$ numerically to normalize the wave function.

Further Problems on Chapter 9

- ** **Problem 9.13** Use the WKB approximation to find the allowed energies of the general power-law potential:

$$V(x) = \alpha|x|^v,$$

where v is a positive number. Check your result for the case $v = 2$. *Answer:*¹⁵

$$E_n = \alpha \left[(n - 1/2) \hbar \sqrt{\frac{\pi}{2m\alpha}} \frac{\Gamma\left(\frac{1}{v} + \frac{3}{2}\right)}{\Gamma\left(\frac{1}{v} + 1\right)} \right]^{\left(\frac{2v}{v+2}\right)}. \quad (9.54)$$

- ** **Problem 9.14** Use the WKB approximation to find the bound state energy for the potential in Problem [2.52](#). Compare the exact answer:
 $-\left[(9/8) - (1/\sqrt{2})\right]\hbar^2a^2/m.$

Problem 9.15 For spherically symmetrical potentials we can apply the WKB approximation to the radial part (Equation [4.37](#)). In the case $l = 0$ it is reasonable¹⁶ to use Equation [9.48](#) in the form

$$\int_0^{r_0} p(r)dr = (n - 1/4)\pi\hbar, \quad (9.55)$$

where r_0 is the turning point (in effect, we treat $r = 0$ as an infinite wall). Exploit this formula to estimate the allowed energies of a particle in the logarithmic potential

$$V(r) = V_0 \ln(r/a)$$

(for constants V_0 and a). Treat only the case $l = 0$. Show that the spacing between the levels is independent of mass. *Partial answer:*

$$E_{n+1} - E_n = V_0 \ln\left(\frac{n + 3/4}{n - 1/4}\right).$$

- ** **Problem 9.16** Use the WKB approximation in the form of Equation [9.52](#),

$$\int_{r_1}^{r_2} p(r)dr = (n' - 1/2)\pi\hbar \quad (9.56)$$

to estimate the bound state energies for hydrogen. Don't forget the centrifugal term in the effective potential (Equation [4.38](#)). The following integral may help:

$$\int_a^b \frac{1}{x} \sqrt{(x-a)(b-x)} dx = \frac{\pi}{2} (\sqrt{b} - \sqrt{a})^2. \quad (9.57)$$

Answer:

$$E_{n'\ell} \approx \frac{-13.6 \text{ eV}}{\left[n' - (1/2) + \sqrt{\ell(\ell+1)} \right]^2}. \quad (9.58)$$

I put a prime on n' , because there is no reason to suppose it corresponds to the n in the Bohr formula. Rather, it orders the states *for a given ℓ* , counting the number of nodes in the radial wave function.¹⁷ In the notation of Chapter 4, $n' = N = n - \ell$ (Equation 4.67). Put this in, expand the square root $(\sqrt{1+\epsilon} = 1 + \frac{1}{2}\epsilon - \frac{1}{8}\epsilon^2 + \dots)$, and compare your result to the Bohr formula.

Problem 9.17 Consider the case of a symmetrical double well, such as the one pictured in Figure 9.14. We are interested in bound states with $E < V(0)$.

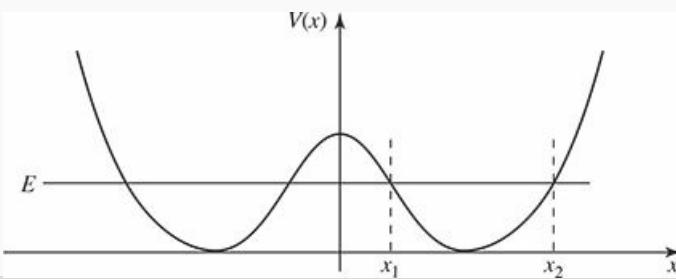


Figure 9.14: Symmetric double well; Problem 9.17.

- (a) Write down the WKB wave functions in regions (i) $x > x_2$, (ii) $x_1 < x < x_2$, and (iii) $0 < x < x_1$. Impose the appropriate connection formulas at x_1 and x_2 (this has already been done, in Equation 9.47, for x_2 ; you will have to work out x_1 for yourself), to show that

$$\psi(x) \approx \begin{cases} \frac{D}{\sqrt{|p(x)|}} \exp \left[-\frac{1}{\hbar} \int_{x_2}^x |p(x')| dx' \right], & \text{(i)} \\ \frac{2D}{\sqrt{|p(x)|}} \sin \left[\frac{1}{\hbar} \int_x^{x_2} p(x') dx' + \frac{\pi}{4} \right], & \text{(ii)} \\ \frac{D}{\sqrt{|p(x)|}} \left[2 \cos \theta e^{\frac{1}{\hbar} \int_x^{x_1} |p(x')| dx'} + \sin \theta e^{-\frac{1}{\hbar} \int_x^{x_1} |p(x')| dx'} \right], & \text{(iii)} \end{cases}$$

where

$$\theta \equiv \frac{1}{\hbar} \int_{x_1}^{x_2} p(x) dx. \quad (9.59)$$

- (b) Because $V(x)$ is symmetric, we need only consider even (+) and odd (-) wave functions. In the former case $\psi'(0) = 0$, and in the latter case $\psi(0) = 0$. Show that this leads to the following quantization condition:

$$\tan \theta = \pm 2e^\phi, \quad (9.60)$$

where

$$\phi \equiv \frac{1}{\hbar} \int_{-x_1}^{x_1} |p(x')| dx'. \quad (9.61)$$

Equation 9.60 determines the (approximate) allowed energies (note that E comes into x_1 and x_2 , so θ and ϕ are both functions of E).

- (c) We are particularly interested in a high and/or broad central barrier, in which case ϕ is large, and e^ϕ is *huge*. Equation 9.60 then tells us that θ must be very close to a half-integer multiple of π . With this in mind, write $\theta = (n + 1/2)\pi + \epsilon$, where $|\epsilon| \ll 1$, and show that the quantization condition becomes

$$\theta \approx \left(n + \frac{1}{2}\right)\pi \mp \frac{1}{2}e^{-\phi}. \quad (9.62)$$

- (d) Suppose each well is a parabola:¹⁸

$$V(x) = \begin{cases} \frac{1}{2}m\omega^2(x+a)^2, & x < 0, \\ \frac{1}{2}m\omega^2(x-a)^2, & x > 0. \end{cases} \quad (9.63)$$

Sketch this potential, find θ (Equation 9.59), and show that

$$E_n^\pm \approx \left(n + \frac{1}{2}\right)\hbar\omega \mp \frac{\hbar\omega}{2\pi}e^{-\phi}. \quad (9.64)$$

Comment: If the central barrier were *impenetrable* ($\phi \rightarrow \infty$), we would simply have two detached harmonic oscillators, and the energies, $E_n = (n + 1/2)\hbar\omega$, would be doubly degenerate, since the particle could be in the left well or in the right one. When the barrier becomes *finite* (putting the two wells into “communication”), the degeneracy is lifted. The even states (ψ_n^+) have slightly *lower* energy, and the odd ones (ψ_n^-) have slightly higher energy.

- (e) Suppose the particle starts out in the *right* well—or, more precisely, in a state of the form

$$\Psi(x, 0) = \frac{1}{\sqrt{2}}(\psi_n^+ + \psi_n^-),$$

which, assuming the phases are picked in the “natural” way, will be concentrated in the right well. Show that it oscillates back and forth between the wells, with a period

$$\tau = \frac{2\pi^2}{\omega}e^\phi. \quad (9.65)$$

- (f) Calculate ϕ , for the specific potential in part (d), and show that for $V(0) \gg E$, $\phi \sim m\omega a^2/\hbar$.

Problem 9.18 Tunneling in the Stark Effect. When you turn on an external electric field, the electron in an atom can, in principle, tunnel out, ionizing the

atom. *Question:* Is this likely to happen in a typical Stark effect experiment? We can estimate the probability using a crude one-dimensional model, as follows. Imagine a particle in a very deep finite square well (Section 2.6).

- What is the energy of the ground state, measured up from the bottom of the well? Assume $V_0 \gg \hbar^2/ma^2$. *Hint:* This is just the ground state energy of the *infinite* square well (of width $2a$).
- Now introduce a perturbation $H' = -\alpha x$ (for an electron in an electric field $\mathbf{E} = -E_{\text{ext}}\hat{i}$ we would have $\alpha = eE_{\text{ext}}$). Assume it is relatively weak ($\alpha a \ll \hbar^2/ma^2$). Sketch the total potential, and note that the particle can now tunnel out, in the direction of positive x .
- Calculate the tunneling factor γ (Equation 9.23), and estimate the time it would take for the particle to escape (Equation 9.29). *Answer:*

$$\gamma = \sqrt{8mV_0^3/3\alpha\hbar}, \tau = (8ma^2/\pi\hbar)e^{2\gamma}.$$
- Put in some reasonable numbers: $V_0 = 20$ eV (typical binding energy for an outer electron), $a = 10^{-10}$ m (typical atomic radius), $E_{\text{ext}} = 7 \times 10^6$ V/m (strong laboratory field), e and m the charge and mass of the electron. Calculate τ , and compare it to the age of the universe.

Problem 9.19 About how long would it take for a (full) can of beer at room temperature to topple over spontaneously, as a result of quantum tunneling?

Hint: Treat it as a uniform cylinder of mass m , radius R , and height b . As the can tips, let x be the height of the center above its equilibrium position ($h/2$). The potential energy is mgx , and it topples when x reaches the critical value $x_0 = \sqrt{R^2 + (h/2)^2} - h/2$. Calculate the tunneling probability (Equation 9.23), for $E = 0$. Use Equation 9.29, with the thermal energy $((1/2)mv^2 = (1/2)k_B T)$ to estimate the velocity. Put in reasonable numbers, and give your final answer in years.¹⁹

Problem 9.20 Equation 9.23 tells us the (approximate) transmission probability for tunneling through a barrier, when $E < V_{\text{max}}$ —a classically forbidden process. In this problem we explore the complementary phenomenon: reflection from a barrier when $E > V_{\text{max}}$ (again, a classically forbidden process). We'll assume that $V(x)$ is an even analytic function, that goes to zero as $x \rightarrow \pm\infty$ (Figure 9.15). *Question:* What is the analog to Equation 9.23?

- Try the obvious approach: assume the potential vanishes for $|x| \geq a$, and use the WKB approximation (Equation 9.13) in the scattering region:

$$\psi(x) \begin{cases} = A e^{ikx} + B e^{-ikx}, & x < a, \\ \approx \frac{1}{\sqrt{p(x)}} [C_+ e^{i\phi(x)} + C_- e^{-i\phi(x)}], & -a < x < a, \\ = C e^{ikx}, & x > a. \end{cases} \quad (9.66)$$

Impose the usual boundary conditions at $\pm a$, and solve for the reflection probability, $R = |B|^2/|A|^2$.

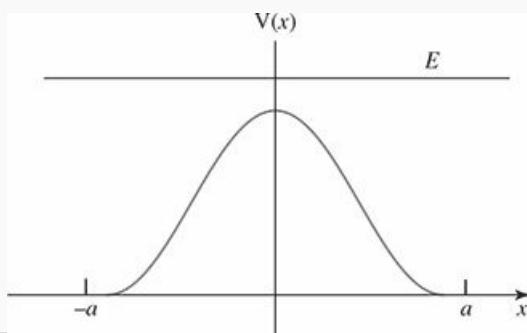


Figure 9.15: Reflection from a barrier (Problem 9.20).

Unfortunately, the result ($R = 0$) is uninformative. It's true that the R is exponentially small (just as the transmission coefficient is, for $E < V_{\max}$), but we've thrown the baby out with the bath water—this approximation is simply too drastic. The correct formula is

$$R = e^{-2\lambda}, \quad \text{where} \quad \lambda \equiv \frac{2}{\hbar} \int_0^{y_0} p(iy) dy \quad (9.67)$$

and y_0 is defined by $p(iy_0) = 0$. Notice that λ (like γ in Equation 9.23) goes like $1/\hbar$; it is in fact the leading term in an expansion in powers of \hbar : $\lambda = c_1/\hbar + c_2 + c_3\hbar + c_4\hbar^2 + \dots$. In the classical limit ($\hbar \rightarrow 0$), λ and γ go to infinity, so R and T go to zero, as expected. It is not easy to derive Equation 9.67,²⁰ but let's look at some examples.

- (b) Suppose $V(x) = V_0 \operatorname{sech}^2(x/a)$, for some positive constants V_0 and a . Plot $V(x)$, plot $p(iy)$ for $0 \leq y \leq y_0$, and show that $\lambda = (\pi a/\hbar) (\sqrt{2mE} - \sqrt{2mV_0})$. Plot R as a function of E , for fixed V_0 .
- (c) Suppose $V(x) = V_0 / [1 + (x/a)^2]$. Plot $V(x)$, and express λ in terms of an elliptic integral. Plot R as a function of E .

¹ In Holland it's KWB, in France it's BWK, and in England it's JWKB (for Jeffreys).

² We might as well take the positive sign, since both are covered by Equation 9.13.

³ In this case the wave function is *real*, and the analogs to Equations 9.6 and 9.7 do not follow *necessarily* from Equation 9.5, although they are still *sufficient*. If this bothers you, study the alternative derivation in Problem 9.2.

⁴ This heuristic argument can be made more rigorous—see Problem 9.11.

⁵ For a more complete discussion, and alternative formulations, see B. R. Holstein, *Am. J. Phys.* **64**, 1061 (1996).

⁶ For an interesting brief history see E. Merzbacher, “The Early History of Quantum Tunneling,” *Physics Today*, August 2002, p. 44.

⁷ In this case the potential does not drop to zero on the left side of the barrier (moreover, this is really a three-dimensional problem), but the essential idea, contained in Equation 9.23, is all we really need.

⁸ This figure is reprinted by permission from David Park, *Introduction to the Quantum Theory*, 3rd edn, Dover Publications, New York (2005); it was adapted from I. Perlman and J. O. Rasmussen, “Alpha Radioactivity,” *Encyclopedia of Physics*, Vol. 42, Springer (1957).

⁹ *Warning:* The following argument is quite technical, and you may wish to skip it on a first reading.

¹⁰ Classically, a linear potential means a constant force, and hence a constant acceleration—the simplest nontrivial motion possible, and the starting point for elementary mechanics. It is ironic that the same potential in quantum mechanics yields stationary states that are unfamiliar transcendental functions, and plays only a peripheral role in the theory. Still, wave packets can be reasonably simple—see Problem 2.51 and especially footnote 61, page 81.

- ¹¹ This is a delicate double constraint, and it is possible to concoct potentials so pathological that no such overlap region exists. However, in practical applications this seldom occurs. See Problem 9.9.
- ¹² At first glance it seems absurd to use a *large-z* approximation in this region, which after all is supposed to be reasonably close to the turning point at $z = 0$ (so that the linear approximation to the potential is valid). But notice that the argument here is αx , and if you study the matter carefully (see Problem 9.9) you will find that there *is* (typically) a region in which αx is large, but at the same time it is reasonable to approximate $V(x)$ by a straight line. Indeed, the asymptotic forms of the Airy functions *are* precisely the WKB solutions to Airy's equation, and since we are already using ψ_{WKB} in the overlap region (Figure 9.10) it is not really a *new* approximation to do the same for ψ_p .
- ¹³ *Not* 2π —an overall minus sign can be absorbed into the normalization factors D and D' .
- ¹⁴ For more on the quantum bouncing ball see Problem 2.59, J. Gea-Banacloche, *Am. J. Phys.* **67**, 776 (1999), and N. Wheeler, “Classical/quantum dynamics in a uniform gravitational field”, unpublished Reed College report (2002). This may sound like an awfully artificial problem, but the experiment has actually been done, using neutrons (V. V. Nesvizhevsky *et al.*, *Nature* **415**, 297 (2002)).
- ¹⁵ As always, the WKB result is most accurate in the semi-classical (large n) regime. In particular, Equation 9.54 is not very good for the ground state ($n = 1$). See W. N. Mei, *Am. J. Phys.* **66**, 541 (1998).
- ¹⁶ Application of the WKB approximation to the radial equation raises some delicate and subtle problems, which I will not go into here. The classic paper on the subject is R. Langer, *Phys. Rev.* **51**, 669 (1937).
- ¹⁷ I thank Ian Gatland and Owen Vajk for pointing this out.
- ¹⁸ Even if $V(x)$ is not strictly parabolic in each well, this calculation of θ , and hence the result (Equation 9.64) will be *approximately* correct, in the sense discussed in Section 2.3, with $\omega \equiv \sqrt{V''(x_0)/m}$, where x_0 is the position of the minimum.
- ¹⁹ R. E. Crandall, *Scientific American*, February, 1997, p. 74.
- ²⁰ L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, Pergamon Press, Oxford (1958), pages 190–191. R. L. Jaffe, *Am. J. Phys.* **78**, 620 (2010) shows that *reflection* (for $E > V_{\max}$) can be regarded as *tunneling* in momentum space, and obtains Equation 9.67 by a clever analog to the argument yielding Equation 9.23.

10
Scattering
◆

10.1 Introduction

10.1.1 Classical Scattering Theory

Imagine a particle incident on some scattering center (say, a marble bouncing off a bowling ball, or a proton fired at a heavy nucleus). It comes in with energy E and **impact parameter** b , and it emerges at some **scattering angle** θ —see Figure 10.1. (I'll assume for simplicity that the target is symmetrical about the z axis, so the trajectory remains in one plane, and that the target is very heavy, so its recoil is negligible.) The essential problem of classical scattering theory is this: *Given the impact parameter, calculate the scattering angle.* Ordinarily, of course, the smaller the impact parameter, the greater the scattering angle.

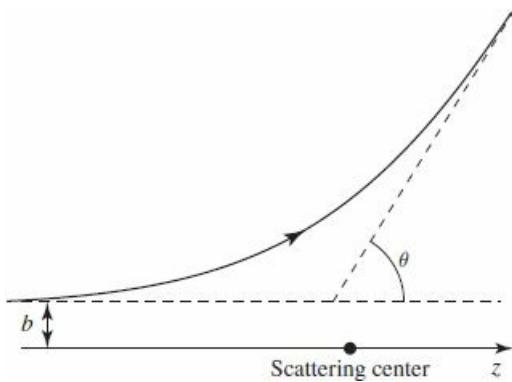


Figure 10.1: The classical scattering problem, showing the impact parameter b and the scattering angle θ .

Example 10.1

Hard-sphere scattering. Suppose the target is a billiard ball, of radius R , and the incident particle is a BB, which bounces off elastically (Figure 10.2). In terms of the angle α , the impact parameter is $b = R \sin \alpha$, and the scattering angle is $\theta = \pi - 2\alpha$, so

$$b = R \sin \left(\frac{\pi}{2} - \frac{\theta}{2} \right) = R \cos \left(\frac{\theta}{2} \right). \quad (10.1)$$

Evidently

$$\theta = \begin{cases} 2 \cos^{-1} (b/R), & (b \leq R), \\ 0, & (b \geq R). \end{cases} \quad (10.2)$$

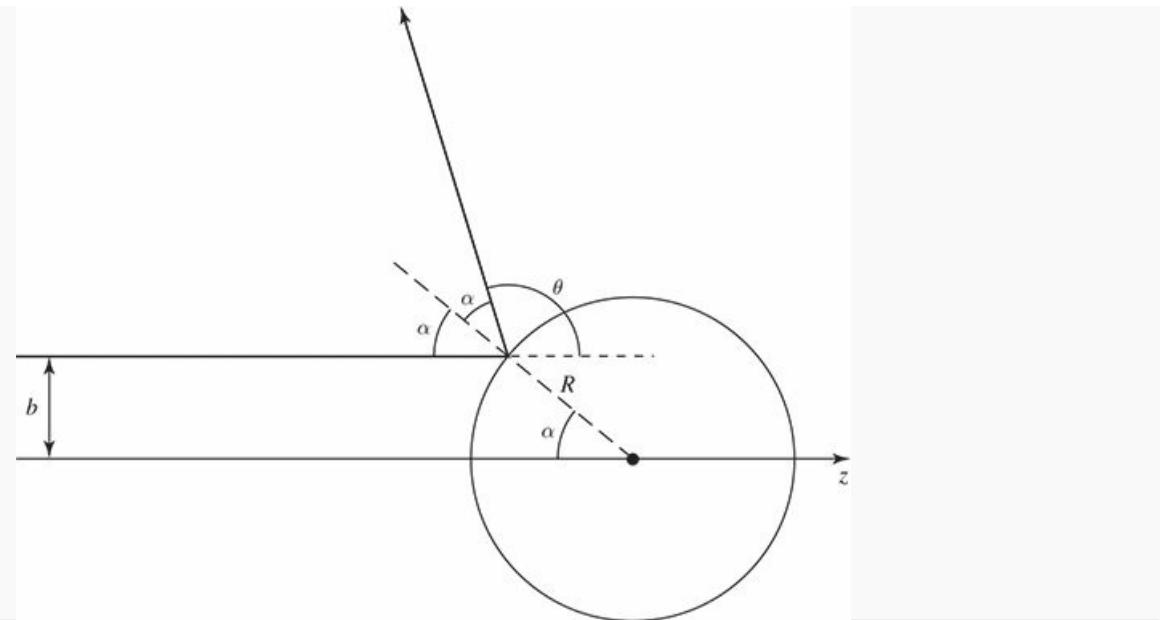


Figure 10.2: Elastic hard-sphere scattering.

More generally, particles incident within an infinitesimal patch of cross-sectional area $d\sigma$ will scatter into a corresponding infinitesimal solid angle $d\Omega$ (Figure 10.3). The larger $d\sigma$ is, the bigger $d\Omega$ will be; the proportionality factor, $D(\theta) \equiv d\sigma/d\Omega$, is called the differential (scattering) cross-section.¹

$$d\sigma = D(\theta) d\Omega. \quad (10.3)$$

In terms of the impact parameter and the azimuthal angle ϕ , $d\sigma = b db d\phi$ and $d\Omega = \sin\theta d\theta d\phi$, so

$$D(\theta) = \frac{b}{\sin\theta} \left| \frac{db}{d\theta} \right|. \quad (10.4)$$

(Since θ is typically a *decreasing* function of b , the derivative is actually negative—hence the absolute value sign.)

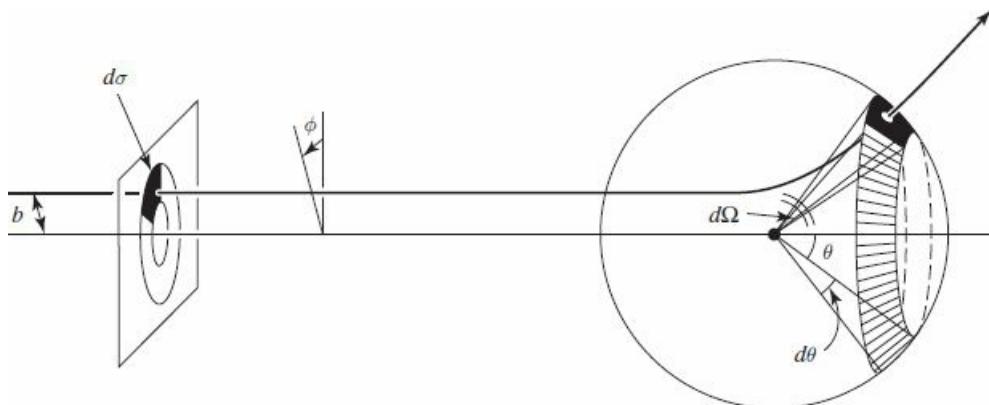


Figure 10.3: Particles incident in the area $d\sigma$ scatter into the solid angle $d\Omega$.

Example 10.2

Hard-sphere scattering (continued). In the case of hard-sphere scattering (Example 10.1)

$$\frac{db}{d\theta} = -\frac{1}{2}R \sin\left(\frac{\theta}{2}\right), \quad (10.5)$$

so

$$D(\theta) = \frac{R \cos(\theta/2)}{\sin \theta} \left(\frac{R \sin(\theta/2)}{2} \right) = \frac{R^2}{4}. \quad (10.6)$$

This example is unusual, in that the differential cross-section is independent of θ .

The **total cross-section** is the *integral* of $D(\theta)$, over all solid angles:

$$\sigma \equiv \int D(\theta) d\Omega; \quad (10.7)$$

roughly speaking, it is the total area of incident beam that is scattered by the target. For example, in the case of hard-sphere scattering,

$$\sigma = (R^2/4) \int d\Omega = \pi R^2, \quad (10.8)$$

which is just what we would expect: It's the cross-sectional area of the sphere; BB's incident within this area will hit the target, and those farther out will miss it completely. But the virtue of the formalism developed here is that it applies just as well to "soft" targets (such as the Coulomb field of a nucleus) that are *not* simply "hit-or-miss".

Finally, suppose we have a *beam* of incident particles, with uniform intensity (or **luminosity**, as particle physicists call it)

$$\mathcal{L} \equiv \text{number of incident particles per unit area, per unit time.} \quad (10.9)$$

The number of particles entering area $d\sigma$ (and hence scattering into solid angle $d\Omega$), per unit time, is $dN = \mathcal{L} d\sigma = \mathcal{L} D(\theta) d\Omega$, so

$$D(\theta) = \frac{1}{\mathcal{L}} \frac{dN}{d\Omega}. \quad (10.10)$$

This is sometimes taken as the *definition* of the differential cross-section, because it makes reference only to quantities easily measured in the laboratory: If the detector subtends a solid angle $d\Omega$, we simply count the *number* recorded per unit time (the **event rate**, dN), divide by $d\Omega$, and normalize to the luminosity of the incident beam.

Problem 10.1 Rutherford scattering. An incident particle of charge q_1 and kinetic energy E scatters off a heavy stationary particle of charge q_2 .

- (a) Derive the formula relating the impact parameter to the scattering angle.² *Answer:* $b = (q_1 q_2 / 8\pi \epsilon_0 E) \cot(\theta/2)$.
- (b) Determine the differential scattering cross-section. *Answer:*

$$D(\theta) = \left[\frac{q_1 q_2}{16\pi \epsilon_0 E \sin^2(\theta/2)} \right]^2. \quad (10.11)$$

- (c) Show that the total cross-section for Rutherford scattering is *infinite*.

10.1.2 Quantum Scattering Theory

In the quantum theory of scattering, we imagine an incident plane wave, $\psi(z) = Ae^{ikz}$, traveling in the z direction, which encounters a scattering potential, producing an outgoing *spherical* wave (Figure 10.4).³ That is, we look for solutions to the Schrödinger equation of the generic form

$$\boxed{\psi(r, \theta) \approx A \left\{ e^{ikz} + f(\theta) \frac{e^{ikr}}{r} \right\}, \quad \text{for large } r.} \quad (10.12)$$

(The spherical wave carries a factor of $1/r$, because this portion of $|\psi|^2$ must go like $1/r^2$ to conserve probability.) The **wave number** k is related to the energy of the incident particles in the usual way:

$$k \equiv \frac{\sqrt{2mE}}{\hbar}. \quad (10.13)$$

(As before, I assume the target is azimuthally symmetrical; in the more general case f would depend on ϕ as well as θ .)

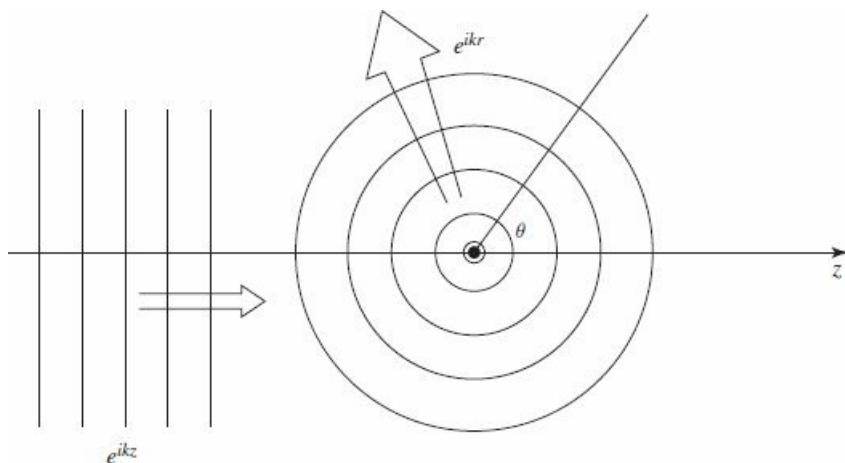


Figure 10.4: Scattering of waves; an incoming plane wave generates an outgoing spherical wave.

The whole problem is to determine the **scattering amplitude** $f(\theta)$; it tells you the *probability of scattering in a given direction* θ , and hence is related to the differential cross-section. Indeed, the probability that the incident particle, traveling at speed v , passes through the infinitesimal area $d\sigma$, in time dt , is (see Figure 10.5)

$$dP = |\psi_{\text{incident}}|^2 dV = |A|^2 (v dt) d\sigma.$$

But this is equal to the probability that the particle scatters into the corresponding solid angle $d\Omega$:

$$dP = |\psi_{\text{scattered}}|^2 dV = \frac{|A|^2 |f|^2}{r^2} (v dt) r^2 d\Omega,$$

from which it follows that $d\sigma = |f|^2 d\Omega$, and hence

$$\boxed{D(\theta) = \frac{d\sigma}{d\Omega} = |f(\theta)|^2.} \quad (10.14)$$

Evidently the differential cross-section (which is the quantity of interest to the experimentalist) is equal to the absolute square of the scattering amplitude (which is obtained by solving the Schrödinger equation). In the following sections we will study two techniques for calculating the scattering amplitude: **partial wave analysis** and the **Born approximation**.

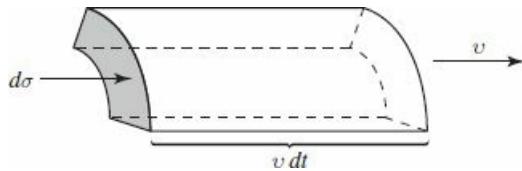


Figure 10.5: The volume dV of incident beam that passes through area $d\sigma$ in time dt .

Problem 10.2 Construct the analogs to Equation [10.12](#) for one-dimensional and two-dimensional scattering.

10.2 Partial Wave Analysis

10.2.1 Formalism

As we found in Chapter 4, the Schrödinger equation for a spherically symmetrical potential $V(r)$ admits the separable solutions

$$\psi(r, \theta, \phi) = R(r)Y_\ell^m(\theta, \phi), \quad (10.15)$$

where Y_ℓ^m is a spherical harmonic (Equation 4.32), and $u(r) = rR(r)$ satisfies the radial equation (Equation 4.37):

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left[V(r) + \frac{\hbar^2 \ell(\ell+1)}{2m r^2} \right] u = Eu. \quad (10.16)$$

At *very* large r the potential goes to zero, and the centrifugal contribution is negligible, so

$$\frac{d^2u}{dr^2} \approx -k^2 u.$$

The general solution is

$$u(r) = C e^{ikr} + D e^{-ikr};$$

the first term represents an *outgoing* spherical wave, and the second an *incoming* one—for the scattered wave we want $D = 0$. At very large r , then,

$$R(r) \sim \frac{e^{ikr}}{r},$$

as we already deduced (on physical grounds) in the previous section (Equation 10.12).

That's for *very* large r (more precisely, for $kr \gg 1$; in optics it would be called the **radiation zone**). As in one-dimensional scattering theory, we assume that the potential is “localized,” in the sense that exterior to some finite scattering region it is essentially zero (Figure 10.6). In the intermediate region (where V can be ignored but the centrifugal term cannot),⁴ the radial equation becomes

$$\frac{d^2u}{dr^2} - \frac{\ell(\ell+1)}{r^2} u = -k^2 u, \quad (10.17)$$

and the general solution (Equation 4.45) is a linear combination of spherical Bessel functions:

$$u(r) = A r j_\ell(kr) + B r n_\ell(kr). \quad (10.18)$$

However, neither j_ℓ (which is somewhat like a sine function) nor n_ℓ (which is a sort of generalized cosine function) represents an outgoing (or an incoming) wave. What we need are the linear combinations analogous to e^{ikr} and e^{-ikr} ; these are known as **spherical Hankel functions**:

$$h_\ell^{(1)}(x) \equiv j_\ell(x) + i n_\ell(x); \quad h_\ell^{(2)}(x) \equiv j_\ell(x) - i n_\ell(x). \quad (10.19)$$

The first few spherical Hankel functions are listed in Table 10.1. At large r , $h_\ell^{(1)}(kr)$ (the **Hankel function of the first kind**) goes like e^{ikr}/r , whereas $h_\ell^{(2)}(kr)$ (the **Hankel function of the second kind**) goes like e^{-ikr}/r ; for outgoing waves, then, we need spherical Hankel functions of the first kind:

$$R(r) \sim h_{\ell}^{(1)}(kr). \quad (10.20)$$

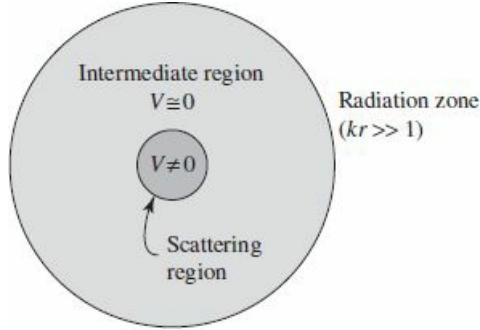


Figure 10.6: Scattering from a localized potential: the scattering region (dark), the intermediate region, where $V = 0$ (shaded), and the radiation zone (where $kr \gg 1$).

Table 10.1: Spherical Hankel functions, $h_{\ell}^{(1)}(x)$ and $h_{\ell}^{(2)}(x)$.

$h_0^{(1)} = -i \frac{e^{ix}}{x}$ $h_1^{(1)} = \left(-\frac{i}{x^2} - \frac{1}{x} \right) e^{ix}$ $h_2^{(1)} = \left(-\frac{3i}{x^3} - \frac{3}{x^2} + \frac{i}{x} \right) e^{ix}$ $h_{\ell}^{(1)} \rightarrow \frac{1}{x} (-i)^{\ell+1} e^{ix}$	$h_0^{(2)} = i \frac{e^{-ix}}{x}$ $h_1^{(2)} = \left(\frac{i}{x^2} - \frac{1}{x} \right) e^{-ix}$ $h_2^{(2)} = \left(\frac{3i}{x^3} - \frac{3}{x^2} - \frac{i}{x} \right) e^{-ix}$ $h_{\ell}^{(2)} \rightarrow \frac{1}{x} (i)^{\ell+1} e^{-ix}$
$\left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \text{for } x \gg 1$	

The exact wave function, in the exterior region (where $V(r) = 0$), is

$$\psi(r, \theta, \phi) = A \left\{ e^{ikz} + \sum_{\ell, m} C_{\ell, m} h_{\ell}^{(1)}(kr) Y_{\ell}^m(\theta, \phi) \right\}. \quad (10.21)$$

The first term is the incident plane wave, and the sum (with expansion coefficients $C_{\ell, m}$) is the scattered wave. But since we are assuming the potential is spherically symmetric, the wave function cannot depend on ϕ .⁵ So only terms with $m = 0$ survive (remember, $Y_{\ell}^m \sim e^{im\phi}$). Now (from Equations 4.27 and 4.32)

$$Y_{\ell}^0(\theta, \phi) = \sqrt{\frac{2\ell+1}{4\pi}} P_{\ell}(\cos \theta), \quad (10.22)$$

where P_{ℓ} is the ℓ th Legendre polynomial. It is customary to redefine the expansion coefficients ($C_{\ell, 0} \equiv i^{\ell+1} k \sqrt{4\pi (2\ell+1)} a_{\ell}$):

$\psi(r, \theta) = A \left\{ e^{ikz} + k \sum_{\ell=0}^{\infty} i^{\ell+1} (2\ell+1) a_{\ell} h_{\ell}^{(1)}(kr) P_{\ell}(\cos \theta) \right\}.$	(10.23)
---	-----------

You'll see in a moment why this peculiar notation is convenient; a_{ℓ} is called the ℓ th partial wave amplitude.

For *very large r*, the Hankel function goes like $(-i)^{\ell+1} e^{ikr}/kr$ (Table 10.1), so

$$\psi(r, \theta) \approx A \left\{ e^{ikz} + f(\theta) \frac{e^{ikr}}{r} \right\}, \quad (10.24)$$

where

$$f(\theta) = \sum_{\ell=0}^{\infty} (2\ell + 1) a_{\ell} P_{\ell}(\cos \theta). \quad (10.25)$$

This confirms more rigorously the general structure postulated in Equation 10.12, and tells us how to compute the scattering amplitude, $f(\theta)$, in terms of the partial wave amplitudes (a_{ℓ}). The differential cross-section is

$$D(\theta) = |f(\theta)|^2 = \sum_{\ell} \sum_{\ell'} (2\ell + 1) (2\ell' + 1) a_{\ell}^* a_{\ell'} P_{\ell}(\cos \theta) P_{\ell'}(\cos \theta), \quad (10.26)$$

and the total cross-section is

$$\sigma = 4\pi \sum_{\ell=0}^{\infty} (2\ell + 1) |a_{\ell}|^2. \quad (10.27)$$

(I used the orthogonality of the Legendre polynomials, Equation 4.34, to do the angular integration.)

10.2.2 Strategy

All that remains is to determine the partial wave amplitudes, a_ℓ , for the potential in question. This is accomplished by solving the Schrödinger equation in the *interior* region (where $V(r)$ is *not* zero), and matching it to the exterior solution (Equation 10.23), using the appropriate boundary conditions. The only problem is that as it stands my notation is hybrid: I used *spherical* coordinates for the scattered wave, but *cartesian* coordinates for the incident wave. We need to rewrite the wave function in a more consistent notation.

Of course, e^{ikz} satisfies the Schrödinger equation with $V = 0$. On the other hand, I just argued that the *general* solution to the Schrödinger equation with $V = 0$ can be written in the form

$$\sum_{\ell,m} [A_{\ell,m} j_\ell(kr) + B_{\ell,m} n_\ell(kr)] Y_\ell^m(\theta, \phi).$$

In particular, then, it must be possible to express e^{ikz} in this way. But e^{ikz} is finite at the origin, so no Neumann functions are allowed in the sum ($n_\ell(kr)$ blows up at $r = 0$), and since $z = r \cos \theta$ has no ϕ dependence, only $m = 0$ terms occur. The resulting expansion of a plane wave in terms of spherical waves is known as Rayleigh's formula:⁶

$$e^{ikz} = \sum_{\ell=0}^{\infty} i^\ell (2\ell + 1) j_\ell(kr) P_\ell(\cos \theta). \quad (10.28)$$

Using this, the wave function in the exterior region (Equation 10.23) can be expressed entirely in terms of r and θ :

$$\psi(r, \theta) = A \sum_{\ell=0}^{\infty} i^\ell (2\ell + 1) [j_\ell(kr) + ik a_\ell h_\ell^{(1)}(kr)] P_\ell(\cos \theta). \quad (10.29)$$

Example 10.3

Quantum hard-sphere scattering. Suppose

$$V(r) = \begin{cases} \infty, & (r \leq a), \\ 0, & (r > a). \end{cases} \quad (10.30)$$

The boundary condition, then, is

$$\psi(a, \theta) = 0, \quad (10.31)$$

so

$$\sum_{\ell=0}^{\infty} i^\ell (2\ell + 1) [j_\ell(ka) + ik a_\ell h_\ell^{(1)}(ka)] P_\ell(\cos \theta) = 0 \quad (10.32)$$

for all θ , from which it follows (Problem 10.3) that

$$a_\ell = i \frac{j_\ell(ka)}{k h_\ell^{(1)}(ka)}. \quad (10.33)$$

In particular, the total cross-section (Equation 10.27) is

$$\sigma = \frac{4\pi}{k^2} \sum_{\ell=0}^{\infty} (2\ell + 1) \left| \frac{j_\ell(ka)}{h_\ell^{(1)}(ka)} \right|^2. \quad (10.34)$$

That's the *exact* answer, but it's not terribly illuminating, so let's consider the limiting case of *low-energy scattering*: $ka \ll 1$. (Since $k = 2\pi/\lambda$, this amounts to saying that the wavelength is much greater than the radius of the sphere.) Referring to Table 4.4, we note that $n_\ell(z)$ is much larger than $j_\ell(z)$, for small z , so

$$\begin{aligned} \frac{j_\ell(z)}{h_\ell^{(1)}(z)} &= \frac{j_\ell(z)}{j_\ell(z) + i n_\ell(z)} \approx -i \frac{j_\ell(z)}{n_\ell(z)} \\ &\approx -i \frac{2^\ell \ell! z^\ell / (2\ell + 1)!}{-(2\ell)! z^{-\ell-1} / 2^\ell \ell!} = \frac{i}{2\ell + 1} \left[\frac{2^\ell \ell!}{(2\ell)!} \right]^2 z^{2\ell+1}, \end{aligned} \quad (10.35)$$

and hence

$$\sigma \approx \frac{4\pi}{k^2} \sum_{\ell=0}^{\infty} \frac{1}{2\ell + 1} \left[\frac{2^\ell \ell!}{(2\ell)!} \right]^4 (ka)^{4\ell+2}.$$

But we're assuming $ka \ll 1$, so the higher powers are negligible—in the low-energy approximation the scattering is dominated by the $\ell = 0$ term. (This means that the differential cross-section is independent of θ , just as it was in the classical case.) Evidently

$$\sigma \approx 4\pi a^2, \quad (10.36)$$

for low energy hard-sphere scattering. Surprisingly, the scattering cross-section is *four times* the geometrical cross-section—in fact, σ is the *total surface area of the sphere*. This “larger effective size” is characteristic of long-wavelength scattering (it would be true in optics, as well); in a sense, these waves “feel” their way around the whole sphere, whereas classical *particles* only see the head-on cross-section (Equation 10.8).

Problem 10.3 Prove Equation 10.33, starting with Equation 10.32. *Hint:* Exploit the orthogonality of the Legendre polynomials to show that the coefficients with different values of ℓ must separately vanish.

**

Problem 10.4 Consider the case of low-energy scattering from a spherical delta-function shell:

$$V(r) = \alpha \delta(r - a),$$

where α and a are constants. Calculate the scattering amplitude, $f(\theta)$, the differential cross-section, $D(\theta)$, and the total cross-section, σ . Assume $ka \ll 1$, so that only the $\ell = 0$ term contributes significantly. (To simplify matters, throw out all $\ell \neq 0$ terms right from the start.) The main problem, of course, is to

determine C_0 . Express your answer in terms of the dimensionless quantity
 $\beta \equiv 2ma\alpha/\hbar^2$. Answer: $\sigma = 4\pi a^2\beta^2/(1 + \beta)^2$.

10.3 Phase Shifts

Consider first the problem of *one*-dimensional scattering from a localized potential $V(x)$ on the half-line $x < 0$ (Figure 10.7). I'll put a “brick wall” at $x = 0$, so a wave incident from the left,

$$\psi_i(x) = Ae^{ikx} \quad (x < -a) \quad (10.37)$$

is entirely reflected

$$\psi_r(x) = Be^{-ikx} \quad (x < -a). \quad (10.38)$$

Whatever happens in the interaction region ($-a < x < 0$), the amplitude of the reflected wave has *got* to be the same as that of the incident wave ($|B| = |A|$), by conservation of probability. But it need not have the same *phase*. If there were no potential at all (just the wall at $x = 0$), then $B = -A$, since the total wave function (incident plus reflected) must vanish at the origin:

$$\psi(x) = A \left(e^{ikx} - e^{-ikx} \right) \quad (V(x) = 0). \quad (10.39)$$

If the potential is *not* zero, the wave function (for $x < -a$) takes the form

$$\psi(x) = A \left(e^{ikx} - e^{i(2\delta - kx)} \right) \quad (V(x) \neq 0). \quad (10.40)$$

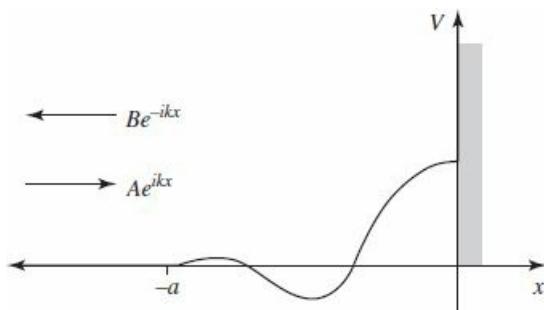


Figure 10.7: One-dimensional scattering from a localized potential bounded on the right by an infinite wall.

The whole theory of scattering reduces to the problem of calculating the **phase shift**⁷ δ (as a function of k , and hence of the energy $E = \hbar^2 k^2 / 2m$), for a specified potential. We do this, of course, by solving the Schrödinger equation in the scattering region ($-a < x < 0$), and imposing appropriate boundary conditions (see Problem 10.5). The advantage of working with the phase shift (as opposed to the complex number B) is that it exploits the physics to simplify the mathematics (trading a complex quantity—two real numbers—for a single real quantity).

Now let's return to the three-dimensional case. The incident plane wave (Ae^{ikz}) carries no angular momentum in the z direction (Rayleigh's formula contains no terms with $m \neq 0$), but it includes all values of the *total* angular momentum ($\ell = 0, 1, 2, \dots$). Because angular momentum is conserved (by a spherically symmetric potential), each **partial wave** (labelled by a particular ℓ) scatters independently, with (again) no change in amplitude⁸—only in phase.

If there is no potential at all, then $\psi = Ae^{ikz}$, and the ℓ th partial wave is (Equation 10.28)

$$\psi^{(\ell)} = A i^\ell (2\ell + 1) j_\ell(kr) P_\ell(\cos \theta) \quad (V(r) = 0). \quad (10.41)$$

But (from Equation 10.19 and Table 10.1)

$$j_\ell(x) = \frac{1}{2} \left[h_\ell^{(1)}(x) + h_\ell^{(2)}(x) \right] \approx \frac{1}{2x} \left[(-i)^{\ell+1} e^{ix} + i^{\ell+1} e^{-ix} \right] \quad (x \gg 1). \quad (10.42)$$

So for large r

$$\psi^{(\ell)} \approx A \frac{(2\ell + 1)}{2ikr} \left[e^{ikr} - (-1)^\ell e^{-ikr} \right] P_\ell(\cos \theta) \quad (V(r) = 0). \quad (10.43)$$

The second term inside the square brackets represents an incoming spherical wave; it comes from the incident plane wave, and is unchanged when we now introduce a potential. The first term is the outgoing wave; it picks up a phase shift (due to the scattering potential):

$$\psi^{(\ell)} \approx A \frac{(2\ell + 1)}{2ikr} \left[e^{i(kr+2\delta_\ell)} - (-1)^\ell e^{-ikr} \right] P_\ell(\cos \theta) \quad (V(r) \neq 0). \quad (10.44)$$

Think of it as a converging spherical wave (the e^{-ikr} term, due exclusively to the $h_\ell^{(2)}$ component in e^{ikz}), which is phase shifted an amount δ_ℓ on the way in, and again δ_ℓ on the way out (hence the 2), emerging as an outgoing spherical wave (the e^{ikr} term, due to the $h_\ell^{(1)}$ part of e^{ikz} plus the scattered wave).

In Section 10.2.1 the whole theory was expressed in terms of the partial wave amplitudes a_ℓ ; now we have formulated it in terms of the phase shifts δ_ℓ . There must be a connection between the two. Indeed, comparing the asymptotic (large r) form of Equation 10.23

$$\psi^{(\ell)} \approx A \left\{ \frac{(2\ell + 1)}{2ikr} \left[e^{ikr} - (-1)^\ell e^{-ikr} \right] + \frac{(2\ell + 1)}{r} a_\ell e^{ikr} \right\} P_\ell(\cos \theta) \quad (10.45)$$

with the generic expression in terms of δ_ℓ (Equation 10.44), we find⁹

$$a_\ell = \frac{1}{2ik} \left(e^{2i\delta_\ell} - 1 \right) = \frac{1}{k} e^{i\delta_\ell} \sin(\delta_\ell). \quad (10.46)$$

It follows in particular (Equation 10.25) that

$$f(\theta) = \frac{1}{k} \sum_{\ell=0}^{\infty} (2\ell + 1) e^{i\delta_\ell} \sin(\delta_\ell) P_\ell(\cos \theta) \quad (10.47)$$

and (Equation 10.27)

$$\sigma = \frac{4\pi}{k^2} \sum_{\ell=0}^{\infty} (2\ell + 1) \sin^2(\delta_\ell). \quad (10.48)$$

Again, the advantage of working with phase shifts (as opposed to partial wave amplitudes) is that they are easier to interpret physically, and simpler mathematically—the phase shift formalism exploits conservation of angular momentum to reduce a complex quantity a_ℓ (two real numbers) to a single real one δ_ℓ .

Problem 10.5 A particle of mass m and energy E is incident from the left on the potential

$$V(x) = \begin{cases} 0, & (x < -a), \\ -V_0, & (-a \leq x \leq 0), \\ \infty, & (x > 0). \end{cases}$$

- (a) If the incoming wave is Ae^{ikx} (where $k = \sqrt{2mE}/\hbar$), find the reflected wave. *Answer:*

$$Ae^{-2ika} \left[\frac{k - ik' \cot(k'a)}{k + ik' \cot(k'a)} \right] e^{-ikx}, \text{ where } k' = \sqrt{2m(E + V_0)}/\hbar.$$

- (b) Confirm that the reflected wave has the same amplitude as the incident wave.
 (c) Find the phase shift δ (Equation 10.40) for a very deep well ($E \ll V_0$).
Answer: $\delta = -ka$.

Problem 10.6 What are the partial wave phase shifts (δ_ℓ) for hard-sphere scattering (Example 10.3)?

Problem 10.7 Find the S -wave ($\ell = 0$) partial wave phase shift $\delta_0(k)$ for scattering from a delta-function shell (Problem 10.4). Assume that the radial wave function $u(r)$ goes to 0 as $r \rightarrow 0$. *Answer:*

$$-\cot^{-1} \left[\cot(ka) + \frac{ka}{\beta \sin^2(ka)} \right], \quad \text{where } \beta \equiv \frac{2m\alpha a}{\hbar^2}.$$

10.4 The Born Approximation

10.4.1 Integral Form of the Schrödinger Equation

The time-independent Schrödinger equation,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi, \quad (10.49)$$

can be written more succinctly as

$$(\nabla^2 + k^2)\psi = Q, \quad (10.50)$$

where

$$k \equiv \frac{\sqrt{2mE}}{\hbar} \quad \text{and} \quad Q \equiv \frac{2m}{\hbar^2}V\psi. \quad (10.51)$$

This has the superficial appearance of the **Helmholtz equation**; note, however, that the “inhomogeneous” term (Q) *itself* depends on ψ . Suppose we could find a function $G(\mathbf{r})$ that solves the Helmholtz equation with a *delta function* “source”:

$$(\nabla^2 + k^2)G(\mathbf{r}) = \delta^3(\mathbf{r}). \quad (10.52)$$

Then we could express ψ as an integral:

$$\psi(\mathbf{r}) = \int G(\mathbf{r} - \mathbf{r}_0)Q(\mathbf{r}_0)d^3\mathbf{r}_0. \quad (10.53)$$

For it is easy to show that this satisfies Schrödinger’s equation, in the form of Equation 10.50:

$$\begin{aligned} (\nabla^2 + k^2)\psi(\mathbf{r}) &= \int [(\nabla^2 + k^2)G(\mathbf{r} - \mathbf{r}_0)]Q(\mathbf{r}_0)d^3\mathbf{r}_0 \\ &= \int \delta^3(\mathbf{r} - \mathbf{r}_0)Q(\mathbf{r}_0)d^3\mathbf{r}_0 = Q(\mathbf{r}). \end{aligned}$$

$G(\mathbf{r})$ is called the **Green’s function** for the Helmholtz equation. (In general, the Green’s function for a linear differential equation represents the “response” to a delta-function source.)

Our first task¹⁰ is to solve Equation 10.52 for $G(\mathbf{r})$. This is most easily accomplished by taking the Fourier transform, which turns the *differential* equation into an *algebraic* equation. Let

$$G(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int e^{i\mathbf{s}\cdot\mathbf{r}}g(\mathbf{s})d^3\mathbf{s}. \quad (10.54)$$

Then

$$(\nabla^2 + k^2)G(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int [(\nabla^2 + k^2)e^{i\mathbf{s}\cdot\mathbf{r}}]g(\mathbf{s})d^3\mathbf{s}.$$

But

$$\nabla^2e^{i\mathbf{s}\cdot\mathbf{r}} = -s^2e^{i\mathbf{s}\cdot\mathbf{r}}, \quad (10.55)$$

and (see Equation 2.147)

$$\delta^3(\mathbf{r}) = \frac{1}{(2\pi)^3} \int e^{i\mathbf{s}\cdot\mathbf{r}} d^3\mathbf{s}, \quad (10.56)$$

so Equation 10.52 says

$$\frac{1}{(2\pi)^{3/2}} \int (-s^2 + k^2) e^{i\mathbf{s}\cdot\mathbf{r}} g(\mathbf{s}) d^3\mathbf{s} = \frac{1}{(2\pi)^3} \int e^{i\mathbf{s}\cdot\mathbf{r}} d^3\mathbf{s}.$$

It follows¹¹ that

$$g(\mathbf{s}) = \frac{1}{(2\pi)^{3/2} (k^2 - s^2)}. \quad (10.57)$$

Putting this back into Equation 10.54, we find:

$$G(\mathbf{r}) = \frac{1}{(2\pi)^3} \int e^{i\mathbf{s}\cdot\mathbf{r}} \frac{1}{(k^2 - s^2)} d^3\mathbf{s}. \quad (10.58)$$

Now, \mathbf{r} is *fixed*, as far as the \mathbf{s} integration is concerned, so we may as well choose spherical coordinates (s, θ, ϕ) with the polar axis along \mathbf{r} (Figure 10.8). Then $\mathbf{s} \cdot \mathbf{r} = sr \cos \theta$, the ϕ integral is trivial (2π), and the θ integral is

$$\int_0^\pi e^{isr \cos \theta} \sin \theta d\theta = -\frac{e^{isr \cos \theta}}{isr} \Big|_0^\pi = \frac{2 \sin(sr)}{sr}. \quad (10.59)$$

Thus

$$G(\mathbf{r}) = \frac{1}{(2\pi)^2 r} \int_0^\infty \frac{s \sin(sr)}{k^2 - s^2} ds = \frac{1}{4\pi^2 r} \int_{-\infty}^\infty \frac{s \sin(sr)}{k^2 - s^2} ds. \quad (10.60)$$

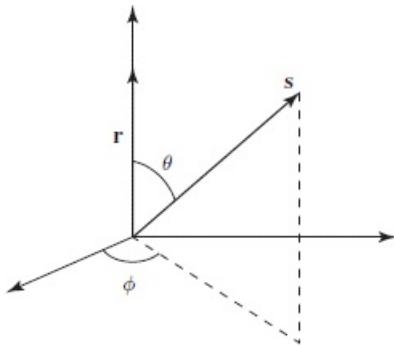


Figure 10.8: Convenient coordinates for the integral in Equation 10.58.

The remaining integral is not so simple. It pays to revert to exponential notation, and factor the denominator:

$$\begin{aligned} G(\mathbf{r}) &= \frac{i}{8\pi^2 r} \left\{ \int_{-\infty}^\infty \frac{se^{isr}}{(s-k)(s+k)} ds - \int_{-\infty}^\infty \frac{se^{-isr}}{(s-k)(s+k)} ds \right\} \\ &= \frac{i}{8\pi^2 r} (I_1 - I_2). \end{aligned} \quad (10.61)$$

These two integrals can be evaluated using Cauchy's integral formula:

$$\oint \frac{f(z)}{(z - z_0)} dz = 2\pi i f(z_0), \quad (10.62)$$

if z_0 lies within the contour (otherwise the integral is zero). In the present case the integration is along the real axis, and it passes *right over* the pole singularities at $\pm k$. We have to decide how to skirt the poles—I'll go *over* the one at $-k$ and *under* the one at $+k$ (Figure 10.9). (You're welcome to choose some *other* convention if you like—even winding seven times around each pole—you'll get a different Green's function, but, as I'll show you in a minute, they're all equally acceptable.)¹²

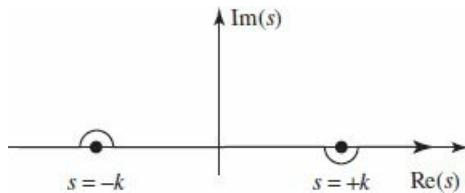


Figure 10.9: Skirting the poles in the contour integral (Equation 10.61).

For each integral in Equation 10.61 I must “close the contour” in such a way that the semicircle at infinity contributes nothing. In the case of I_1 , the factor e^{isr} goes to zero when s has a large *positive* imaginary part; for this one I close *above* (Figure 10.10(a)). The contour encloses only the singularity at $s = +k$, so

$$I_1 = \oint \left[\frac{se^{isr}}{s+k} \right] \frac{1}{s-k} ds = 2\pi i \left[\frac{se^{isr}}{s+k} \right] \Big|_{s=k} = i\pi e^{ikr}. \quad (10.63)$$

In the case of I_2 , the factor e^{-isr} goes to zero when s has a large *negative* imaginary part, so we close *below* (Figure 10.10(b)); this time the contour encloses the singularity at $s = -k$ (and it goes around in the *clockwise* direction, so we pick up a minus sign):

$$I_2 = -\oint \left[\frac{se^{-isr}}{s-k} \right] \frac{1}{s+k} ds = -2\pi i \left[\frac{se^{-isr}}{s-k} \right] \Big|_{s=-k} = -i\pi e^{ikr}. \quad (10.64)$$

Conclusion:

$$G(\mathbf{r}) = \frac{i}{8\pi^2 r} \left[(i\pi e^{ikr}) - (-i\pi e^{ikr}) \right] = -\frac{e^{ikr}}{4\pi r}. \quad (10.65)$$

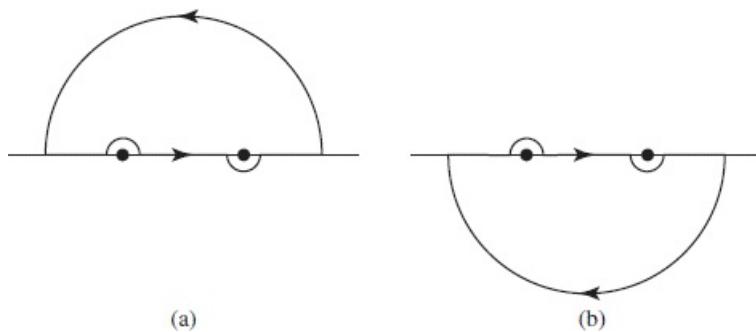


Figure 10.10: Closing the contour in Equations 10.63 and 10.64.

This, finally, is the Green's function for the Helmholtz equation—the solution to Equation 10.52. (If you got lost in all that analysis, you might want to *check* the result by direct differentiation—see Problem

[10.8.](#)) Or rather, it is a Green's function for the Helmholtz equation, for we can add to $G(\mathbf{r})$ any function $G_0(\mathbf{r})$ that satisfies the *homogeneous* Helmholtz equation:

$$\left(\nabla^2 + k^2 \right) G_0(\mathbf{r}) = 0; \quad (10.66)$$

clearly, the result $(G + G_0)$ still satisfies Equation [10.52](#). This ambiguity corresponds precisely to the ambiguity in how to skirt the poles—a different choice amounts to picking a different function $G_0(\mathbf{r})$.

Returning to Equation [10.53](#), the general solution to the Schrödinger equation takes the form

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) - \frac{m}{2\pi\hbar^2} \int \frac{e^{ik|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|} V(\mathbf{r}_0) \psi(\mathbf{r}_0) d^3\mathbf{r}_0, \quad (10.67)$$

where ψ_0 satisfies the *free*-particle Schrödinger equation,

$$\left(\nabla^2 + k^2 \right) \psi_0 = 0. \quad (10.68)$$

Equation [10.67](#) is the **integral form of the Schrödinger equation**; it is entirely equivalent to the more familiar differential form. At first glance it *looks* like an explicit *solution* to the Schrödinger equation (for any potential)—which is too good to be true. Don't be deceived: There's a ψ under the integral sign on the right hand side, so you can't do the integral unless you already know the solution! Nevertheless, the integral form can be very powerful, and it is particularly well suited to scattering problems, as we'll see in the following section.

Problem 10.8 Check that Equation [10.65](#) satisfies Equation [10.52](#), by direct substitution. Hint: $\nabla^2 (1/r) = -4\pi\delta^3(\mathbf{r})$.^{[13](#)}

**

Problem 10.9 Show that the ground state of hydrogen (Equation [4.80](#)) satisfies the integral form of the Schrödinger equation, for the appropriate V and E (note that E is *negative*, so $k = i\kappa$, where $\kappa \equiv \sqrt{-2mE}/\hbar$).

10.4.2 The First Born Approximation

Suppose $V(\mathbf{r}_0)$ is localized about $\mathbf{r}_0 = 0$ —that is, the potential drops to zero outside some finite region (as is typical for a scattering problem), and we want to calculate $\psi(\mathbf{r})$ at points *far away* from the scattering center. Then $|\mathbf{r}| \gg |\mathbf{r}_0|$ for all points that contribute to the integral in Equation 10.67, so

$$|\mathbf{r} - \mathbf{r}_0|^2 = r^2 + r_0^2 - 2\mathbf{r} \cdot \mathbf{r}_0 \approx r^2 \left(1 - 2\frac{\mathbf{r} \cdot \mathbf{r}_0}{r^2}\right), \quad (10.69)$$

and hence

$$|\mathbf{r} - \mathbf{r}_0| \approx r - \hat{\mathbf{r}} \cdot \mathbf{r}_0. \quad (10.70)$$

Let

$$\mathbf{k} \equiv k\hat{\mathbf{r}}; \quad (10.71)$$

then

$$e^{ik|\mathbf{r}-\mathbf{r}_0|} \approx e^{ikr} e^{-i\mathbf{k} \cdot \mathbf{r}_0}, \quad (10.72)$$

and therefore

$$\frac{e^{ik|\mathbf{r} - \mathbf{r}_0|}}{|\mathbf{r} - \mathbf{r}_0|} \approx \frac{e^{ikr}}{r} e^{-i\mathbf{k} \cdot \mathbf{r}_0}. \quad (10.73)$$

(In the *denominator* we can afford to make the more radical approximation $|\mathbf{r} - \mathbf{r}_0| \approx r$; in the *exponent* we need to keep the next term. If this puzzles you, try including the next term in the expansion of the denominator. What we are doing is expanding in powers of the small quantity (r_0/r) , and dropping all but the lowest order.)

In the case of scattering, we want

$$\psi_0(\mathbf{r}) = A e^{ikz}, \quad (10.74)$$

representing an incident plane wave. For large r , then,

$$\psi(\mathbf{r}) \approx A e^{ikz} - \frac{m}{2\pi\hbar^2} \frac{e^{ikr}}{r} \int e^{-i\mathbf{k} \cdot \mathbf{r}_0} V(\mathbf{r}_0) \psi(\mathbf{r}_0) d^3 \mathbf{r}_0. \quad (10.75)$$

This is in the standard form (Equation 10.12), and we can read off the scattering amplitude:

$$f(\theta, \phi) = -\frac{m}{2\pi\hbar^2 A} \int e^{-i\mathbf{k} \cdot \mathbf{r}_0} V(\mathbf{r}_0) \psi(\mathbf{r}_0) d^3 \mathbf{r}_0. \quad (10.76)$$

This is *exact*.¹⁴ Now we invoke the **Born approximation**: Suppose the incoming plane wave is *not substantially altered by the potential*; then it makes sense to use

$$\psi(\mathbf{r}_0) \approx \psi_0(\mathbf{r}_0) = A e^{ikz_0} = A e^{i\mathbf{k}' \cdot \mathbf{r}_0}, \quad (10.77)$$

where

$$\mathbf{k}' \equiv k\hat{z}, \quad (10.78)$$

inside the integral. (This would be the *exact* wave function, if V were zero; it is essentially a *weak potential* approximation.¹⁵) In the Born approximation, then,

$$f(\theta, \phi) \approx -\frac{m}{2\pi\hbar^2} \int e^{i(\mathbf{k}'-\mathbf{k}) \cdot \mathbf{r}_0} V(\mathbf{r}_0) d^3\mathbf{r}_0. \quad (10.79)$$

(In case you have lost track of the definitions of \mathbf{k}' and \mathbf{k} , they both have magnitude k , but the former points in the direction of the incident beam, while the latter points toward the detector—see Figure 10.11; $\hbar(\mathbf{k} - \mathbf{k}')$ is the **momentum transfer** in the process.)

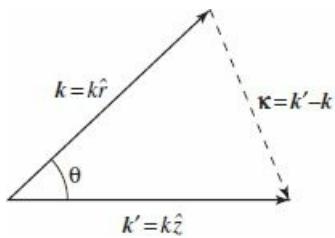


Figure 10.11: Two wave vectors in the Born approximation: \mathbf{k}' points in the *incident* direction, \mathbf{k} in the *scattered* direction.

In particular, for low energy (long wavelength) scattering, the exponential factor is essentially constant over the scattering region, and the Born approximation simplifies to

$$f(\theta, \phi) \approx -\frac{m}{2\pi\hbar^2} \int V(\mathbf{r}) d^3\mathbf{r}, \quad (\text{low energy}). \quad (10.80)$$

(I dropped the subscript on \mathbf{r} , since there is no likelihood of confusion at this point.)

Example 10.4

Low-energy soft-sphere scattering.¹⁶ Suppose

$$V(\mathbf{r}) = \begin{cases} V_0, & (r \leq a), \\ 0, & (r > a). \end{cases} \quad (10.81)$$

In this case the low-energy scattering amplitude is

$$f(\theta, \phi) \approx -\frac{m}{2\pi\hbar^2} V_0 \left(\frac{4}{3}\pi a^3 \right), \quad (10.82)$$

(independent of θ and ϕ), the differential cross-section is

$$\frac{d\sigma}{d\Omega} = |f|^2 \approx \left(\frac{2m V_0 a^3}{3\hbar^2} \right)^2, \quad (10.83)$$

and the total cross-section is

$$\sigma \approx 4\pi \left(\frac{2m V_0 a^3}{3\hbar^2} \right)^2. \quad (10.84)$$

For a spherically symmetrical potential, $V(\mathbf{r}) = V(r)$ —but *not* necessarily at low energy—the Born approximation again reduces to a simpler form. Define

$$\kappa \equiv \mathbf{k}' - \mathbf{k}, \quad (10.85)$$

and let the polar axis for the \mathbf{r}_0 integral lie along κ , so that

$$(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}_0 = \kappa r_0 \cos \theta_0. \quad (10.86)$$

Then

$$f(\theta) \approx -\frac{m}{2\pi\hbar^2} \int e^{i\kappa r_0 \cos \theta_0} V(r_0) r_0^2 \sin \theta_0 dr_0 d\theta_0 d\phi_0. \quad (10.87)$$

The ϕ_0 integral is trivial (2π), and the θ_0 integral is one we have encountered before (see Equation 10.59). Dropping the subscript on r , we are left with

$$f(\theta) \approx -\frac{2m}{\hbar^2 \kappa} \int_0^\infty r V(r) \sin(\kappa r) dr, \text{ (spherical symmetry)}. \quad (10.88)$$

The angular dependence of f is carried by κ ; in Figure 10.11 we see that

$$\kappa = 2k \sin(\theta/2). \quad (10.89)$$

Example 10.5

Yukawa scattering. The **Yukawa potential** (which is a crude model for the binding force in an atomic nucleus) has the form

$$V(r) = \beta \frac{e^{-\mu r}}{r}, \quad (10.90)$$

where β and μ are constants. The Born approximation gives

$$f(\theta) \approx -\frac{2m\beta}{\hbar^2 \kappa} \int_0^\infty e^{-\mu r} \sin(\kappa r) dr = -\frac{2m\beta}{\hbar^2 (\mu^2 + \kappa^2)}. \quad (10.91)$$

(You get to work out the integral for yourself, in Problem 10.11.)

Example 10.6

Rutherford scattering. If we put in $\beta = q_1 q_2 / 4\pi\epsilon_0$, $\mu = 0$, the Yukawa potential reduces to the Coulomb potential, describing the electrical interaction of two point charges. Evidently the scattering amplitude is

$$f(\theta) \approx -\frac{2m q_1 q_2}{4\pi\epsilon_0 \hbar^2 \kappa^2}, \quad (10.92)$$

or (using Equations 10.89 and 10.51):

$$f(\theta) \approx -\frac{q_1 q_2}{16\pi\epsilon_0 E \sin^2(\theta/2)}. \quad (10.93)$$

The differential cross-section is the square of this:

$$\frac{d\sigma}{d\Omega} = \left[\frac{q_1 q_2}{16\pi\epsilon_0 E \sin^2(\theta/2)} \right]^2, \quad (10.94)$$

which is precisely the Rutherford formula (Equation 10.11). It happens that for the Coulomb potential classical mechanics, the Born approximation, and quantum field theory all yield the same result. As they say in the computer business, the Rutherford formula is amazingly “robust.”

- * **Problem 10.10** Find the scattering amplitude, in the Born approximation, for soft-sphere scattering at arbitrary energy. Show that your formula reduces to Equation 10.82 in the low-energy limit.

Problem 10.11 Evaluate the integral in Equation 10.91, to confirm the expression on the right.

- ** **Problem 10.12** Calculate the total cross-section for scattering from a Yukawa potential, in the Born approximation. Express your answer as a function of E .

- * **Problem 10.13** For the potential in Problem 10.4,
 - calculate $f(\theta)$, $D(\theta)$, and σ , in the low-energy Born approximation;
 - calculate $f(\theta)$ for arbitrary energies, in the Born approximation;
 - show that your results are consistent with the answer to Problem 10.4, in the appropriate regime.

10.4.3 The Born Series

The Born approximation is similar in spirit to the **impulse approximation** in classical scattering theory. In the impulse approximation we begin by pretending that the particle keeps going in a straight line (Figure 10.12), and compute the transverse impulse that would be delivered to it in that case:

$$I = \int F_{\perp} dt. \quad (10.95)$$

If the deflection is relatively small, this should be a good approximation to the transverse momentum imparted to the particle, and hence the scattering angle is

$$\theta \approx \tan^{-1}(I/p), \quad (10.96)$$

where p is the incident momentum. This is, if you like, the “first-order” impulse approximation (the *zeroth*-order is what we *started* with: no deflection at all). Likewise, in the zeroth-order Born approximation the incident plane wave passes by with no modification, and what we explored in the previous section is really the first-order correction to this. But the same idea can be iterated to generate a series of higher-order corrections, which presumably converge to the exact answer.

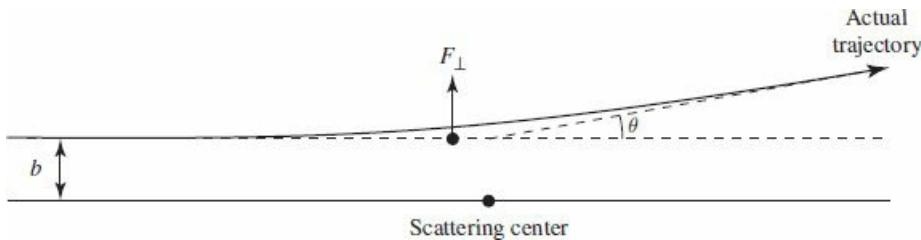


Figure 10.12: The impulse approximation assumes that the particle continues undeflected, and calculates the transverse momentum delivered.

The integral form of the Schrödinger equation reads

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) + \int g(\mathbf{r} - \mathbf{r}_0) V(\mathbf{r}_0) \psi(\mathbf{r}_0) d^3 \mathbf{r}_0, \quad (10.97)$$

where ψ_0 is the incident wave,

$$g(\mathbf{r}) \equiv -\frac{m}{2\pi\hbar^2} \frac{e^{ikr}}{r} \quad (10.98)$$

is the Green’s function (into which I have now incorporated the factor $2m/\hbar^2$, for convenience), and V is the scattering potential. Schematically,

$$\psi = \psi_0 + \int gV\psi. \quad (10.99)$$

Suppose we take this expression for ψ , and plug it in under the integral sign:

$$\psi = \psi_0 + \int gV\psi_0 + \int \int gVgV\psi. \quad (10.100)$$

Iterating this procedure, we obtain a formal series for ψ :

$$\psi = \psi_0 + \int gV\psi_0 + \int \int gVgV\psi_0 + \int \int \int gVgVgV\psi_0 + \dots \quad (10.101)$$

In each integrand only the *incident* wave function (ψ_0) appears, together with more and more powers of gV . The *first* Born approximation truncates the series after the second term, but it is pretty clear how one generates the higher-order corrections.

The Born series can be represented diagrammatically as shown in Figure 10.13. In zeroth order ψ is untouched by the potential; in first order it is “kicked” once, and then “propagates” out in some new direction; in second order it is kicked, propagates to a new location, is kicked again, and then propagates out; and so on. In this context the Green’s function is sometimes called the **propagator**—it tells you how the disturbance propagates between one interaction and the next. The Born series was the inspiration for Feynman’s formulation of relativistic quantum mechanics, which is expressed entirely in terms of **vertex factors** (V) and propagators (g), connected together in **Feynman diagrams**.

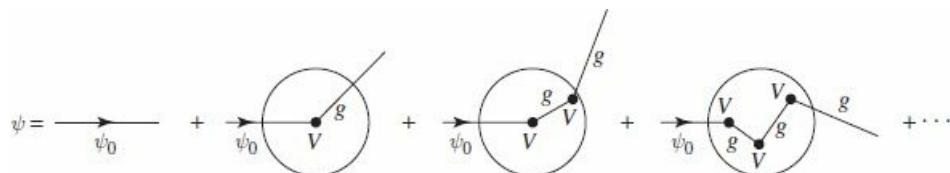


Figure 10.13: Diagrammatic interpretation of the Born series (Equation 10.101).

Problem 10.14 Calculate θ (as a function of the impact parameter) for Rutherford scattering, in the impulse approximation. Show that your result is consistent with the exact expression (Problem 10.1(a)), in the appropriate limit.

Problem 10.15 Find the scattering amplitude for low-energy soft-sphere scattering in the *second* Born approximation. *Answer:*

$$-\left(2mV_0a^3/3\hbar^2\right)\left[1 - \left(4mV_0a^2/5\hbar^2\right)\right].$$

Further Problems on Chapter 10

Problem 10.16 Find the Green's function for the *one*-dimensional Schrödinger equation, and use it to construct the integral form (analogous to Equation 10.66). *Answer:*

$$\psi(x) = \psi_0(x) - \frac{im}{\hbar^2 k} \int_{-\infty}^{\infty} e^{ik|x-x_0|} V(x_0) \psi(x_0) dx_0. \quad (10.102)$$

**

Problem 10.17 Use your result in Problem 10.16 to develop the Born approximation for one-dimensional scattering (on the interval $-\infty < x < \infty$, with no “brick wall” at the origin). That is, choose $\psi_0(x) = Ae^{ikx}$, and assume $\psi(x_0) \approx \psi_0(x_0)$ to evaluate the integral. Show that the reflection coefficient takes the form:

$$R \approx \left(\frac{m}{\hbar^2 k} \right)^2 \left| \int_{-\infty}^{\infty} e^{2ikx} V(x) dx \right|^2. \quad (10.103)$$

Problem 10.18 Use the one-dimensional Born approximation (Problem 10.17) to compute the transmission coefficient ($T = 1 - R$) for scattering from a delta function (Equation 2.117) and from a finite square well (Equation 2.148). Compare your results with the exact answers (Equations 2.144 and 2.172).

Problem 10.19 Prove the **optical theorem**, which relates the total cross-section to the imaginary part of the forward scattering amplitude:

$$\sigma = \frac{4\pi}{k} \text{Im} [f(0)]. \quad (10.104)$$

Hint: Use Equations 10.47 and 10.48.

Problem 10.20 Use the Born approximation to determine the total cross-section for scattering from a gaussian potential

$$V(\mathbf{r}) = V_0 e^{-\mu r^2/a^2}.$$

Express your answer in terms of the constants V_0 , a , and m (the mass of the incident particle), and $k \equiv \sqrt{2mE}/\hbar$, where E is the incident energy.

Problem 10.21 Neutron diffraction. Consider a beam of neutrons scattering from a crystal (Figure 10.14). The interaction between neutrons and the nuclei in the crystal is short ranged, and can be approximated as

$$V(\mathbf{r}) = \frac{2\pi\hbar^2 b}{m} \sum_i \delta^3(\mathbf{r} - \mathbf{r}_i),$$

where the \mathbf{r}_i are the locations of the nuclei and the strength of the potential is expressed in terms of the **nuclear scattering length** b .

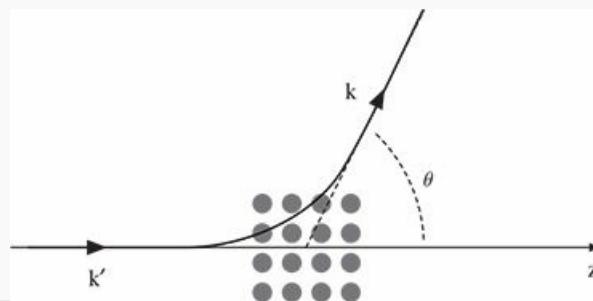


Figure 10.14: Neutron scattering from a crystal.

- (a) In the first Born approximation, show that

$$\frac{d\sigma}{d\Omega} = b^2 \left| \sum_i e^{-i\mathbf{q}\cdot\mathbf{r}_i} \right|^2$$

where $\mathbf{q} \equiv \mathbf{k} - \mathbf{k}'$.

- (b) Now consider the case where the nuclei are arranged on a cubic lattice with spacing a . Take the positions to be

$$\mathbf{r}_i = l a \hat{i} + m a \hat{j} + n a \hat{k}$$

where l , m , and n all range from 0 to $N - 1$, so there are a total of N^3 nuclei.¹⁷ Show that

$$\frac{d\sigma}{d\Omega} = b^2 \frac{\sin^2(N q_x a/2)}{\sin^2(q_x a/2)} \frac{\sin^2(N q_y a/2)}{\sin^2(q_y a/2)} \frac{\sin^2(N q_z a/2)}{\sin^2(q_z a/2)}.$$



- (c) Plot

$$\frac{1}{N} \frac{\sin^2(N q_x a/2)}{\sin^2(q_x a/2)}$$

as a function of $q_x a$ for several values of N ($N = 1, 5, 10$) to show that the function describes a series of peaks that become progressively sharper as N increases.

- (d) In light of (c), in the limit of large N the differential scattering cross section is negligibly small except at one of these peaks:

$$\mathbf{q} = \mathbf{G}_{lmn} = \frac{2\pi}{a} (l \hat{i} + m \hat{j} + n \hat{k})$$

for integer l , m , and n . The vectors \mathbf{G}_{lmn} are called **reciprocal lattice vectors**. Find the scattering angles (θ) at which peaks occur. If the neutron's wavelength is equal to the crystal spacing a , what are the three smallest (nonzero) angles?

Comment: Neutron diffraction is one method used, to determine crystal structures (electrons and x-rays can also be used and the same expression for the locations of the peaks holds). In this problem we looked at a cubic arrangement of atoms, but a different arrangement (hexagonal for example) would produce peaks at a different set of angles. Thus from the scattering data one can infer the underlying crystal structure.

Problem 10.22 Two-dimensional scattering theory. By analogy with Section [10.2](#), develop partial wave analysis for two dimensions.

- (a) In polar coordinates (r, θ) the Laplacian is

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2}. \quad (10.105)$$

Find the separable solutions to the (time-independent) Schrödinger equation, for a potential with azimuthal symmetry ($V(r, \theta) \rightarrow V(r)$).

Answer:

$$\psi(r, \theta) = R(r) e^{ij\theta}, \quad (10.106)$$

where j is an integer, and $u \equiv \sqrt{r} R$ satisfies the radial equation

$$-\frac{\hbar^2}{2m} \frac{d^2 u}{dr^2} + \left[V(r) + \frac{\hbar^2}{2m} \frac{(j^2 - 1/4)}{r^2} \right] u = Eu. \quad (10.107)$$

- (b) By solving the radial equation for very large r (where both $V(r)$ and the centrifugal term go to zero), show that an outgoing radial wave has the asymptotic form

$$R(r) \sim \frac{e^{ikr}}{\sqrt{r}}, \quad (10.108)$$

where $k \equiv \sqrt{2mE}/\hbar$. Check that an incident wave of the form Ae^{ikx} satisfies the Schrödinger equation, for $V(r) = 0$ (this is trivial, if you use cartesian coordinates). Write down the two-dimensional analog to Equation [10.12](#), and compare your result to Problem [10.2](#). *Answer:*

$$\psi(r, \theta) \approx A \left[e^{ikx} + f(\theta) \frac{e^{ikr}}{\sqrt{r}} \right], \quad \text{for large } r. \quad (10.109)$$

- (c) Construct the analog to Equation [10.21](#) (the wave function in the region where $V(r) = 0$ but the centrifugal term *cannot* be ignored). *Answer:*

$$\psi(r, \theta) = A \left\{ e^{ikx} + \sum_{j=-\infty}^{\infty} c_j H_j^{(1)}(kr) e^{ij\theta} \right\}, \quad (10.110)$$

where $H_j^{(1)}$ is the Hankel function (*not* the *spherical* Hankel function!) of order j .¹⁸

(d) For large z ,

$$H_j^{(1)}(z) \sim \sqrt{2/\pi} e^{-i\pi/4} (-i)^j \frac{e^{iz}}{\sqrt{z}}. \quad (10.111)$$

Use this to show that

$$f(\theta) = \sqrt{2/\pi k} e^{-i\pi/4} \sum_{j=-\infty}^{\infty} (-i)^j c_j e^{ij\theta}. \quad (10.112)$$

(e) Adapt the argument of Section 10.1.2 to this two-dimensional geometry.

Instead of the *area* $d\sigma$, we have a *length*, db , and in place of the solid angle $d\Omega$ we have the increment of scattering angle $|d\theta|$; the role of the differential cross-section is played by

$$D(\theta) \equiv \left| \frac{db}{d\theta} \right|, \quad (10.113)$$

and the effective “width” of the target (analogous to the total cross-section) is

$$B \equiv \int_0^{2\pi} D(\theta) d\theta. \quad (10.114)$$

Show that

$$D(\theta) = |f(\theta)|^2, \quad \text{and} \quad B = \frac{4}{k} \sum_{j=-\infty}^{\infty} |c_j|^2. \quad (10.115)$$

(f) Consider the case of scattering from a hard disk (or, in three dimensions, an infinite cylinder¹⁹) of radius a :

$$V(r) = \begin{cases} \infty, & (r \leq a), \\ 0, & (r > a). \end{cases} \quad (10.116)$$

By imposing appropriate boundary conditions at $r = a$, determine B . You’ll need the analog to Rayleigh’s formula:

$$e^{ikx} = \sum_{j=-\infty}^{\infty} (i)^j J_j(kr) e^{ij\theta} \quad (10.117)$$

(where J_j is the Bessel function of order j). Plot B as a function of ka , for $0 < ka < 2$.

Problem 10.23 Scattering of identical particles. The results for scattering of a particle from a fixed target also apply to the scattering of two particles in the center of mass frame. With $\psi(\mathbf{R}, \mathbf{r}) = \psi_R(\mathbf{R}) \psi_r(\mathbf{r})$, $\psi_r(\mathbf{r})$ satisfies

$$-\frac{\hbar^2}{2\mu} \nabla^2 \psi_r + V(r) \psi_r = E_r \psi_r \quad (10.118)$$

(see Problem 5.1) where $V(\mathbf{r})$ is the interaction between the particles (assumed here to depend only on their separation distance). This is the *one-particle* Schrödinger equation (with the reduced mass μ in place of m).

(a) Show that if the two particles are identical (spinless) bosons, then $\psi_r(\mathbf{r})$ must be an even function of \mathbf{r} (Figure 10.15).

(b) By symmetrizing Equation 10.12 (why is this allowed?), show that the scattering amplitude in this case is

$$f_B(\theta) = f(\theta) + f(\pi - \theta)$$

where $f(\theta)$ is the scattering amplitude of a single particle of mass μ from a fixed target $V(r)$.

(c) Show that the partial wave amplitudes of f_B vanish for all odd powers of ℓ .

(d) How are the results of (a)–(c) different if the particles are identical fermions (in a triplet spin state).

(e) Show that the scattering amplitude for identical fermions vanishes at $\pi/2$.

(f) Plot the logarithm of the differential scattering cross section for fermions and for bosons in Rutherford scattering (Equation 10.93).²⁰

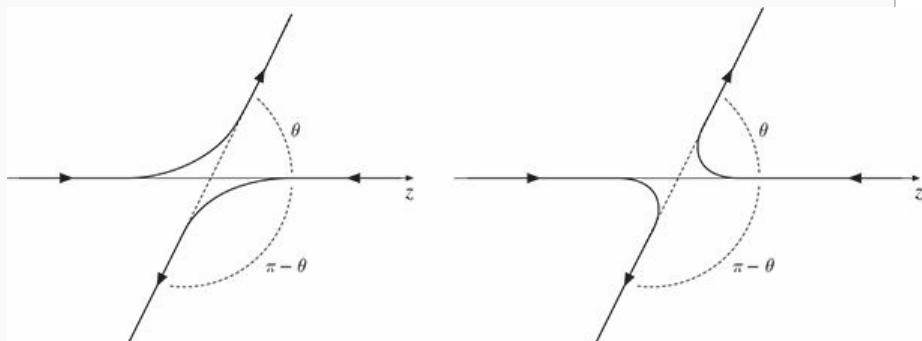


Figure 10.15: Scattering of identical particles.

¹ This is terrible language: D isn't a *differential*, and it isn't a cross-section. To my ear, the words "differential cross-section" would attach more naturally to $d\sigma$. But I'm afraid we're stuck with this terminology. I should also warn you that the notation $D(\theta)$ is nonstandard—most people just call it $d\sigma/d\Omega$ (which makes Equation 10.3 look like a tautology). I think it will be less confusing if we give the differential cross-section its own symbol.

² This isn't easy, and you might want to refer to a book on classical mechanics, such as Jerry B. Marion and Stephen T. Thornton, *Classical Dynamics of Particles and Systems*, 4th edn, Saunders, Fort Worth, TX (1995), Section 9.10.

³ For the moment, there's not much *quantum* mechanics in this; what we're really talking about is the scattering of *waves*, as opposed to *particles*, and you could even think of Figure 10.4 as a picture of water waves encountering a rock, or (better, since we're interested in three-dimensional scattering) sound waves bouncing off a basketball.

⁴ What follows does not apply to the Coulomb potential, since $1/r$ goes to zero more slowly than $1/r^2$, as $r \rightarrow \infty$, and the centrifugal term does *not* dominate in this region. In this sense the Coulomb potential is not localized, and partial wave analysis is inapplicable.

- ⁵ There's nothing wrong with θ dependence, of course, because the incoming plane wave defines a z direction, breaking the spherical symmetry. But the *azimuthal* symmetry remains; the incident plane wave has no ϕ dependence, and there is nothing in the scattering process that could introduce any ϕ dependence in the outgoing wave.
- ⁶ For a guide to the proof, see George Arfken and Hans-Jurgen Weber, *Mathematical Methods for Physicists*, 7th edn, Academic Press, Orlando (2013), Exercises 15.2.24 and 15.2.25.
- ⁷ The 2 in front of δ is conventional. We think of the incident wave as being phase shifted once on the way in, and again on the way out; δ is the “one way” phase shift, and the *total* is 2δ .
- ⁸ One reason this subject can be so confusing is that practically everything is called an “amplitude”: $f(\theta)$ is the “scattering amplitude”, a_ℓ is the “partial wave amplitude”, but the first is a function of θ , and both are complex numbers. I'm *now* talking about “amplitude” in the original sense: the (*real*, of course) height of a sinusoidal wave.
- ⁹ Although I used the asymptotic form of the wave function to draw the connection between a_ℓ and δ_ℓ , there is nothing approximate about the result (Equation 10.46). Both of them are *constants* (independent of r), and δ_ℓ means the phase shift in the asymptotic region (where the Hankel functions have settled down to $e^{\pm ikr}/kr$).
- ¹⁰ *Warning:* You are approaching two pages of heavy analysis, including contour integration; if you wish, skip straight to the answer, Equation 10.65.
- ¹¹ This is clearly *sufficient*, but it is also *necessary*, as you can easily show by combining the two terms into a single integral, and using Plancherel's theorem, Equation 2.103.
- ¹² If you are unfamiliar with this technique you have every right to be suspicious. In truth, the integral in Equation 10.60 is simply ill-defined—it does not converge, and it's something of a miracle that we can make sense of it at all. The root of the problem is that $G(\mathbf{r})$ doesn't really have a legitimate Fourier transform; we're exceeding the speed limit, here, and just hoping we won't get caught.
- ¹³ See, for example, D. Griffiths, *Introduction to Electrodynamics*, 4th edn (Cambridge University Press, Cambridge, UK, 2017), Section 1.5.3.
- ¹⁴ Remember, $f(\theta, \phi)$ is by definition the coefficient of Ae^{ikr}/r at large r .
- ¹⁵ Typically, partial wave analysis is useful when the incident particle has low energy, for then only the first few terms in the series contribute significantly; the Born approximation is more useful at *high* energy, when the deflection is relatively small.
- ¹⁶ You can't apply the Born approximation to *hard-sphere* scattering ($V_0 = \infty$)—the integral blows up. The point is that we assumed the potential is *weak*, and doesn't change the wave function much in the scattering region. But a *hard* sphere changes it *radically*—from Ae^{ikz} to zero.
- ¹⁷ It makes no difference that this crystal isn't “centered” at the origin: shifting the crystal by \mathbf{R} amounts to adding \mathbf{R} to each of the \mathbf{r}_i , and that doesn't affect $d\sigma/d\Omega$. After all, we're assuming an incident plane wave, which extends to $\pm\infty$ in the x and y directions.
- ¹⁸ See Mary Boas, *Mathematical Methods in the Physical Sciences*, 3rd edn (Wiley, New York, 2006), Section 12.17.
- ¹⁹ S. McAlinden and J. Shertzer, *Am. J. Phys.* 84, 764 (2016).
- ²⁰ Equation 10.93 was derived by taking the limit of Yukawa scattering (Example 10.5) and the result for $f(\theta)$ is missing a phase factor (see Albert Messiah, *Quantum Mechanics*, Dover, New York, NY (1999), Section XI.7). That factor drops out of the cross-section for scattering from a fixed potential—giving the correct answer in Example 10.6—but would show up in the cross-section for scattering of identical particles.

11

Quantum Dynamics

◊

So far, practically everything we have done belongs to the subject that might properly be called **quantum statics**, in which the *potential energy function is independent of time*: $V(\mathbf{r}, t) = V(\mathbf{r})$. In that case the (time-dependent) Schrödinger equation,

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi, \quad (11.1)$$

can be solved by separation of variables:

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-iEt/\hbar}, \quad (11.2)$$

where $\psi(\mathbf{r})$ satisfies the time-*independent* Schrödinger equation,

$$\hat{H}\psi = E\psi. \quad (11.3)$$

Because the time dependence of separable solutions is carried by the exponential factor ($e^{-iEt/\hbar}$), which cancels out when we construct the physically relevant quantity $|\Psi|^2$, all probabilities and expectation values (for such states) are constant in time. By forming *linear combinations* of these stationary states we obtain wave functions with more interesting time dependence,

$$\Psi(\mathbf{r}, t) = \sum c_n \psi_n(\mathbf{r})e^{-iE_nt/\hbar}, \quad (11.4)$$

but even then the possible values of the energy (E_n), and their respective probabilities ($|c_n|^2$), are constant.

If we want to allow for **transitions** (**quantum jumps**, as they are sometimes called) between one energy level and another, we must introduce a *time-dependent* potential (**quantum dynamics**). There are precious few exactly solvable problems in quantum dynamics. However, if the time-dependent part of the Hamiltonian is *small* (compared to the time-independent part), it can be treated as a perturbation. The main purpose of this chapter is to develop **time-dependent perturbation theory**, and study its most important application: the emission or absorption of radiation by an atom.

Problem 11.1 Why isn't it *trivial* to solve the time-dependent Schrödinger equation (11.1), in its dependence on t ? After all, it's a first-order differential equation.

- (a) How would you solve the equation

$$\frac{df}{dt} = kf$$

(for $f(t)$), if k were a constant?

- (b) What if k is itself a function of t ? (Here $k(t)$ and $f(t)$ might also depend on other variables, such as \mathbf{r} —it doesn't matter.)
- (c) Why not do the same thing for the Schrödinger equation (with a time-dependent Hamiltonian)? To see that this doesn't work, consider the

simple case

$$\hat{H}(t) = \begin{cases} \hat{H}_1, & (0 < t < \tau), \\ \hat{H}_2, & (t > \tau), \end{cases}$$

where \hat{H}_1 and \hat{H}_2 are themselves time-independent. If the solution in part (b) held for the Schrödinger equation, the wave function at time $t > \tau$ would be

$$\Psi(t) = e^{-i[\hat{H}_1\tau + \hat{H}_2(t-\tau)]/\hbar} \Psi(0),$$

but of course we could also write

$$\Psi(t) = e^{-i\hat{H}_2(t-\tau)/\hbar} \Psi(\tau) = e^{-i\hat{H}_2(t-\tau)/\hbar} e^{-i\hat{H}_1\tau/\hbar} \Psi(0).$$

Why are these generally *not* the same? [This is a subtle matter; if you want to pursue it further, see Problem [11.23](#).]

11.1 Two-Level Systems

To begin with, let us suppose that there are just *two* states of the (unperturbed) system, ψ_a and ψ_b . They are eigenstates of the unperturbed Hamiltonian, \hat{H}^0 :

$$\hat{H}^0\psi_a = E_a\psi_a, \quad \text{and} \quad \hat{H}^0\psi_b = E_b\psi_b, \quad (11.5)$$

and they are orthonormal:

$$\langle \psi_i | \psi_j \rangle = \delta_{ij}, \quad (i, j = a, b). \quad (11.6)$$

Any state can be expressed as a linear combination of them; in particular,

$$\Psi(0) = c_a\psi_a + c_b\psi_b. \quad (11.7)$$

The states ψ_a and ψ_b might be position-space wave functions, or spinors, or something more exotic—it doesn't matter. It is the *time* dependence that concerns us here, so when I write $\Psi(t)$, I simply mean the state of the system at time t . In the absence of any perturbation, each component evolves with its characteristic wiggle factor:

$$\Psi(t) = c_a\psi_a e^{-iE_at/\hbar} + c_b\psi_b e^{-iE_bt/\hbar}. \quad (11.8)$$

Informally, we say that $|c_a|^2$ is the “probability that the particle is in state ψ_a ”—by which we *really* mean the probability that a measurement of the energy would yield the value E_a . Normalization of Ψ requires, of course, that

$$|c_a|^2 + |c_b|^2 = 1. \quad (11.9)$$

11.1.1 The Perturbed System

Now suppose we turn on a time-dependent perturbation, $\hat{H}'(t)$. Since ψ_a and ψ_b constitute a complete set, the wave function $\Psi(t)$ can still be expressed as a linear combination of them. The only difference is that c_a and c_b are now *functions of t*:

$$\Psi(t) = c_a(t)\psi_a e^{-iE_a t/\hbar} + c_b(t)\psi_b e^{-iE_b t/\hbar}. \quad (11.10)$$

(I could absorb the exponential factors into $c_a(t)$ and $c_b(t)$, and some people prefer to do it this way, but I think it is nicer to keep visible the part of the time dependence that would be present even *without* the perturbation.) The whole problem is to determine c_a and c_b , as functions of time. If, for example, the particle started out in the state ψ_a ($c_a(0) = 1$, $c_b(0) = 0$), and at some later time t_1 we find that $c_a(t_1) = 0$, $c_b(t_1) = 1$, we shall report that the system underwent a transition from ψ_a to ψ_b .

We solve for $c_a(t)$ and $c_b(t)$ by demanding that $\Psi(t)$ satisfy the time-dependent Schrödinger equation,

$$\hat{H}\Psi = i\hbar \frac{\partial \Psi}{\partial t}, \quad \text{where} \quad \hat{H} = \hat{H}^0 + \hat{H}'(t). \quad (11.11)$$

From Equations 11.10 and 11.11, we find:

$$\begin{aligned} & c_a \left(\hat{H}^0 \psi_a \right) e^{-iE_a t/\hbar} + c_b \left(\hat{H}^0 \psi_b \right) e^{-iE_b t/\hbar} + c_a \left(\hat{H}' \psi_a \right) e^{-iE_a t/\hbar} \\ & + c_b \left(\hat{H}' \psi_b \right) e^{-iE_b t/\hbar} = i\hbar \left[\dot{c}_a \psi_a e^{-iE_a t/\hbar} + \dot{c}_b \psi_b e^{-iE_b t/\hbar} \right. \\ & \left. + c_a \psi_a \left(-\frac{iE_a}{\hbar} \right) e^{-iE_a t/\hbar} + c_b \psi_b \left(-\frac{iE_b}{\hbar} \right) e^{-iE_b t/\hbar} \right]. \end{aligned}$$

In view of Equation 11.5, the first two terms on the left cancel the last two terms on the right, and hence

$$c_a \left(\hat{H}' \psi_a \right) e^{-iE_a t/\hbar} + c_b \left(\hat{H}' \psi_b \right) e^{-iE_b t/\hbar} = i\hbar \left(\dot{c}_a \psi_a e^{-iE_a t/\hbar} + \dot{c}_b \psi_b e^{-iE_b t/\hbar} \right). \quad (11.12)$$

To isolate \dot{c}_a , we use the standard trick: Take the inner product with ψ_a and exploit the orthogonality of ψ_a and ψ_b (Equation 11.6):

$$c_a \langle \psi_a | \hat{H}' | \psi_a \rangle e^{-iE_a t/\hbar} + c_b \langle \psi_a | \hat{H}' | \psi_b \rangle e^{-iE_b t/\hbar} = i\hbar \dot{c}_a e^{-iE_a t/\hbar}.$$

For short, we define

$$H'_{ij} \equiv \langle \psi_i | \hat{H}' | \psi_j \rangle; \quad (11.13)$$

note that the hermiticity of \hat{H}' entails $H'_{ji} = (H'_{ij})^*$. Multiplying through by $- (i/\hbar) e^{iE_a t/\hbar}$, we conclude that:

$$\dot{c}_a = -\frac{i}{\hbar} \left[c_a H'_{aa} + c_b H'_{ab} e^{-i(E_b - E_a)t/\hbar} \right]. \quad (11.14)$$

Similarly, the inner product with ψ_b picks out \dot{c}_b :

$$c_a \langle \psi_b | \hat{H}' | \psi_a \rangle e^{-iE_a t/\hbar} + c_b \langle \psi_b | \hat{H}' | \psi_b \rangle e^{-iE_b t/\hbar} = i\hbar \dot{c}_b e^{-iE_b t/\hbar},$$

and hence

$$\dot{c}_b = -\frac{i}{\hbar} [c_b H'_{bb} + c_a H'_{ba} e^{i(E_b - E_a)t/\hbar}]. \quad (11.15)$$

Equations 11.14 and 11.15 determine $c_a(t)$ and $c_b(t)$; taken together, they are completely equivalent to the (time-dependent) Schrödinger equation, for a two-level system. Typically, the diagonal matrix elements of \hat{H}' vanish (see Problem 11.5 for the general case):

$$H'_{aa} = H'_{bb} = 0. \quad (11.16)$$

If so, the equations simplify:

$$\dot{c}_a = -\frac{i}{\hbar} H'_{ab} e^{-i\omega_0 t} c_b, \quad \dot{c}_b = -\frac{i}{\hbar} H'_{ba} e^{i\omega_0 t} c_a, \quad (11.17)$$

where

$$\omega_0 \equiv \frac{E_b - E_a}{\hbar}. \quad (11.18)$$

(I'll assume that $E_b \geq E_a$, so $\omega_0 \geq 0$.)

* **Problem 11.2** A hydrogen atom is placed in a (time-dependent) electric field $\mathbf{E} = E(t)\hat{\mathbf{k}}$. Calculate all four matrix elements H'_{ij} of the perturbation $\hat{H}' = eE\hat{\mathbf{z}}$ between the ground state ($n = 1$) and the (quadruply degenerate) first excited states ($n = 2$). Also show that $H'_{ii} = 0$ for all five states. *Note:* There is only one integral to be done here, if you exploit oddness with respect to z ; only one of the $n = 2$ states is “accessible” from the ground state by a perturbation of this form, and therefore the system functions as a two-state configuration—assuming transitions to higher excited states can be ignored.

* **Problem 11.3** Solve Equation 11.17 for the case of a *time-independent* perturbation, assuming that $c_a(0) = 1$ and $c_b(0) = 0$. Check that $|c_a(t)|^2 + |c_b(t)|^2 = 1$. *Comment:* Ostensibly, this system oscillates between “pure ψ_a ” and “some ψ_b .” Doesn't this contradict my general assertion that no transitions occur for time-independent perturbations? No, but the reason is rather subtle: In this case ψ_a and ψ_b are not, and never were, eigenstates of the Hamiltonian—a measurement of the energy *never* yields E_a or E_b . In time-dependent perturbation theory we typically contemplate turning *on* the perturbation for a while, and then turning it *off* again, in order to examine the system. At the beginning, and at the end, ψ_a and ψ_b are eigenstates of the exact Hamiltonian, and only in this context does it make sense to say that the system underwent a transition from one to the other. For the present problem, then,

assume that the perturbation was turned on at time $t = 0$, and off again at time T —this doesn't affect the *calculations*, but it allows for a more sensible interpretation of the result.

** **Problem 11.4** Suppose the perturbation takes the form of a delta function (in time):

$$\hat{H}' = \hat{U}\delta(t);$$

assume that $U_{aa} = U_{bb} = 0$, and let $U_{ab} = U_{ba}^* \equiv \alpha$. If $c_a(-\infty) = 1$ and $c_b(-\infty) = 0$, find $c_a(t)$ and $c_b(t)$, and check that $|c_a(t)|^2 + |c_b(t)|^2 = 1$. What is the net probability ($P_{a \rightarrow b}$ for $t \rightarrow \infty$) that a transition occurs? Hint: You might want to treat the delta function as the limit of a sequence of rectangles. *Answer:* $P_{a \rightarrow b} = \sin^2(|\alpha|/\hbar)$.

11.1.2 Time-Dependent Perturbation Theory

So far, everything is *exact*: We have made no assumption about the *size* of the perturbation. But if \hat{H}' is “small,” we can solve Equation 11.17 by a process of successive approximations, as follows. Suppose the particle starts out in the lower state:

$$c_a(0) = 1, \quad c_b(0) = 0. \quad (11.19)$$

If there were *no perturbation at all*, they would stay this way forever:

Zeroth Order:

$$c_a^{(0)}(t) = 1, \quad c_b^{(0)}(t) = 0. \quad (11.20)$$

(I'll use a superscript in parentheses to indicate the order of the approximation.)

To calculate the first-order approximation, we insert the zeroth-order values on the right side of Equation 11.17:

First Order:

$$\begin{aligned} \frac{dc_a^{(1)}}{dt} &= 0 \Rightarrow c_a^{(1)}(t) = 1; \\ \frac{dc_b^{(1)}}{dt} &= -\frac{i}{\hbar} H'_{ba} e^{i\omega_0 t} \Rightarrow c_b^{(1)} = -\frac{i}{\hbar} \int_0^t H'_{ba}(t') e^{i\omega_0 t'} dt'. \end{aligned} \quad (11.21)$$

Now we insert *these* expressions on the right side of Equation 11.17 to obtain the *second*-order approximation:

Second Order:

$$\begin{aligned} \frac{dc_a^{(2)}}{dt} &= -\frac{i}{\hbar} H'_{ab} e^{-i\omega_0 t} \left(-\frac{i}{\hbar} \right) \int_0^t H'_{ba}(t') e^{i\omega_0 t'} dt' \Rightarrow \\ c_a^{(2)}(t) &= 1 - \frac{1}{\hbar^2} \int_0^t H'_{ab}(t') e^{-i\omega_0 t'} \left[\int_0^{t'} H'_{ba}(t'') e^{i\omega_0 t''} dt'' \right] dt', \end{aligned} \quad (11.22)$$

while c_b is unchanged ($c_b^{(2)}(t) = c_b^{(1)}(t)$). (Notice that $c_a^{(2)}(t)$ *includes* the zeroth-order term; the second-order *correction* would be the integral part alone.)

In principle, we could continue this ritual indefinitely, always inserting the n th-order approximation into the right side of Equation 11.17, and solving for the $(n + 1)$ th order. The zeroth order contains *no* factors of \hat{H}' , the first-order correction contains *one* factor of \hat{H}' , the second-order correction has *two* factors of \hat{H}' , and so on.¹ The error in the first-order approximation is evident in the fact that $|c_a^{(1)}(t)|^2 + |c_b^{(1)}(t)|^2 \neq 1$ (the *exact* coefficients must, of course, obey Equation 11.9). However, $|c_a^{(1)}(t)|^2 + |c_b^{(1)}(t)|^2$ is equal to 1 to *first order in* \hat{H}' , which is all we can expect from a first-order approximation. And the same goes for the higher orders.

Equation 11.21 can be written in the form

$$c_b^{(1)}(t) e^{-iE_bt/\hbar} = -\frac{i}{\hbar} \int_0^t e^{-iE_b(t-t')/\hbar} H'_{ba}(t') e^{-iE_at'/\hbar} dt' \quad (11.23)$$

(where I've restored the exponential we factored out in Equation 11.10). This suggests a nice pictorial interpretation: reading from right to left, the system remains in state a from time 0 to time t' (picking up the “wiggle factor” $e^{-iE_a t'/\hbar}$), makes a transition from state a to state b at time t' , and then remains in state b until time t (picking up the “wiggle factor” $e^{-iE_b(t-t')/\hbar}$). This process is represented in Figure 11.1. (Don't take the picture too literally: there is no sharp transition between these states; in fact, you integrate over all the times t' at which this transition can occur.)

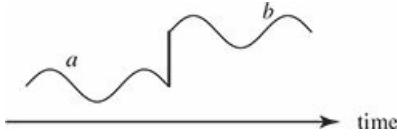


Figure 11.1: Pictorial representation of Equation 11.23.

This interpretation of the perturbation series is especially illuminating at higher orders and for multi-level systems, where the expressions become complicated. Consider Equation 11.22, which can be written

$$c_a^{(2)}(t) e^{-iE_a t/\hbar} = e^{-iE_a t/\hbar} + \left(-\frac{i}{\hbar}\right)^2 \int_0^t \int_0^{t'} e^{-iE_a(t-t')/\hbar} H'_{ab}(t') e^{-iE_b(t'-t'')/\hbar} H'_{ba}(t'') e^{-iE_a t''/\hbar} dt'' dt'. \quad (11.24)$$

The two terms here describe a process where the system remains in state a for the entire time, and a second process where the system transitions from a to b at time t'' and then back to a at time t' . Graphically, this is shown in Figure 11.2.



Figure 11.2: Pictorial representation of Equation 11.24.

With the insight provided by these pictures, it is easy to write down the general result for a multi-level system:²

$$c_n^{(2)}(t) e^{-iE_n t/\hbar} = \delta_{ni} e^{-iE_i t/\hbar} + \left(\frac{-i}{\hbar}\right) \int_0^t e^{-iE_n(t-t')/\hbar} H_{ni}(t') e^{-iE_i t'/\hbar} dt' + \sum_m \left(\frac{-i}{\hbar}\right)^2 \int_0^t \int_0^{t'} e^{-iE_n(t-t')/\hbar} H_{nm}(t') e^{-iE_m(t'-t'')/\hbar} \times H_{mi}(t'') e^{-iE_i t''/\hbar} dt'' dt'. \quad (11.25)$$

For $n \neq i$, this is represented by the diagram in Figure 11.3. The first-order term describes a direct transition from i to n , and the second-order term describes a process where the transition occurs via an intermediate (or “virtual”) state m .

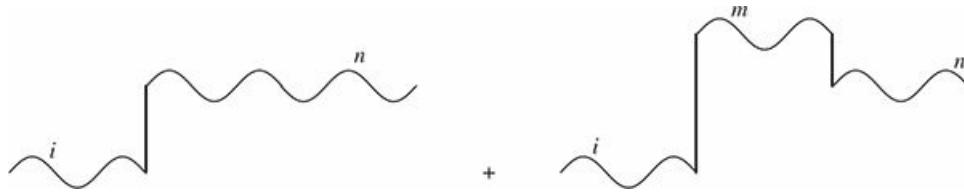


Figure 11.3: Pictorial representation of Equation 11.25 for $n \neq i$.

**

Problem 11.5 Suppose you *don't* assume $H'_{aa} = H'_{bb} = 0$.

- Find $c_a(t)$ and $c_b(t)$ in first-order perturbation theory, for the case $c_a(0) = 1, c_b(0) = 0$. Show that $|c_a^{(1)}(t)|^2 + |c_b^{(1)}(t)|^2 = 1$, to first order in \hat{H}' .
- There is a nicer way to handle this problem. Let

$$d_a \equiv e^{\frac{i}{\hbar} \int_0^t H'_{aa}(t') dt'} c_a, \quad d_b \equiv e^{\frac{i}{\hbar} \int_0^t H'_{bb}(t') dt'} c_b. \quad (11.26)$$

Show that

$$\dot{d}_a = -\frac{i}{\hbar} e^{i\phi} H'_{ab} e^{-i\omega_0 t} d_b; \quad \dot{d}_b = -\frac{i}{\hbar} e^{-i\phi} H'_{ba} e^{i\omega_0 t} d_a, \quad (11.27)$$

where

$$\phi(t) \equiv \frac{1}{\hbar} \int_0^t [H'_{aa}(t') - H'_{bb}(t')] dt'. \quad (11.28)$$

So the equations for d_a and d_b are identical in structure to Equation 11.17 (with an extra factor $e^{i\phi}$ tacked onto \hat{H}').

- Use the method in part (b) to obtain $c_a(t)$ and $c_b(t)$ in first-order perturbation theory, and compare your answer to (a). Comment on any discrepancies.

*

Problem 11.6 Solve Equation 11.17 to second order in perturbation theory, for the general case $c_a(0) = a, c_b(0) = b$.

**

Problem 11.7 Calculate $c_a(t)$ and $c_b(t)$, to second order, for the perturbation in Problem 11.3. Compare your answer with the exact result.

*

Problem 11.8 Consider a perturbation to a two-level system with matrix elements

$$H'_{ab} = H'_{ba} = \frac{\alpha}{\sqrt{\pi\tau}} e^{-(t/\tau)^2}, \quad H'_{aa} = H'_{bb} = 0.$$

where τ and α are positive constants with the appropriate units.

- (a) According to first-order perturbation theory, if the system starts off in the state $c_a = 1, c_b = 0$ at $t = -\infty$, what is the probability that it will be found in the state b at $t = \infty$?
- (b) In the limit that $\tau \rightarrow 0$, $H'_{ab} = \alpha\delta(t)$. Compute the $\tau \rightarrow 0$ limit of your expression from part (a) and compare the result of Problem [11.4](#).
- (c) Now consider the opposite extreme: $\omega_0\tau \gg 1$. What is the limit of your expression from part (a)? *Comment:* This is an example of the adiabatic theorem (Section [11.5.2](#)).

11.1.3 Sinusoidal Perturbations

Suppose the perturbation has sinusoidal time dependence:

$$\hat{H}'(\mathbf{r}, t) = V(\mathbf{r}) \cos(\omega t), \quad (11.29)$$

so that

$$H'_{ab} = V_{ab} \cos(\omega t), \quad (11.30)$$

where

$$V_{ab} \equiv \langle \psi_a | V | \psi_b \rangle. \quad (11.31)$$

(As before, I'll assume the diagonal matrix elements vanish, since this is almost always the case in practice.) To first order (from now on we'll work *exclusively* in first order, and I'll dispense with the superscripts) we have (Equation 11.21):

$$\begin{aligned} c_b(t) &\approx -\frac{i}{\hbar} V_{ba} \int_0^t \cos(\omega t') e^{i\omega_0 t'} dt' = -\frac{i V_{ba}}{2\hbar} \int_0^t [e^{i(\omega_0+\omega)t'} + e^{i(\omega_0-\omega)t'}] dt' \\ &= -\frac{V_{ba}}{2\hbar} \left[\frac{e^{i(\omega_0+\omega)t} - 1}{\omega_0 + \omega} + \frac{e^{i(\omega_0-\omega)t} - 1}{\omega_0 - \omega} \right]. \end{aligned} \quad (11.32)$$

That's the *answer*, but it's a little cumbersome to work with. Things simplify substantially if we restrict our attention to driving frequencies (ω) that are very close to the transition frequency (ω_0), so that the second term in the square brackets dominates; specifically, we assume:

$$\omega_0 + \omega \gg |\omega_0 - \omega|. \quad (11.33)$$

This is not much of a limitation, since perturbations at *other* frequencies have a negligible probability of causing a transition anyway. Dropping the first term, we have

$$\begin{aligned} c_b(t) &\approx -\frac{V_{ba}}{2\hbar} \frac{e^{i(\omega_0-\omega)t/2}}{\omega_0 - \omega} [e^{i(\omega_0-\omega)t/2} - e^{-i(\omega_0-\omega)t/2}] \\ &= -i \frac{V_{ba}}{\hbar} \frac{\sin[(\omega_0 - \omega)t/2]}{\omega_0 - \omega} e^{i(\omega_0-\omega)t/2}. \end{aligned} \quad (11.34)$$

The **transition probability**—the probability that a particle which started out in the state ψ_a will be found, at time t , in the state ψ_b —is

$$P_{a \rightarrow b}(t) = |c_b(t)|^2 \approx \frac{|V_{ab}|^2}{\hbar^2} \frac{\sin^2[(\omega_0 - \omega)t/2]}{(\omega_0 - \omega)^2}. \quad (11.35)$$

The most remarkable feature of this result is that, as a function of time, the transition probability *oscillates* sinusoidally (Figure 11.4). After rising to a maximum of $|V_{ab}|^2 / \hbar^2 (\omega_0 - \omega)^2$ —necessarily much less than 1, else the assumption that the perturbation is “small” would be invalid—it drops back down to zero! At times $t_n = 2n\pi / |\omega_0 - \omega|$, where $n = 1, 2, 3, \dots$, the particle is *certain* to be back in the lower state. If you want to maximize your chances of provoking a transition, you should *not* keep the perturbation on for a long period; you do better to *turn it off* after a time $\pi / |\omega_0 - \omega|$, and hope to “catch” the system in the upper state. In

Problem 11.9 it is shown that this “flopping” is not an artifact of perturbation theory—it occurs also in the exact solution, though the flopping frequency is modified somewhat.

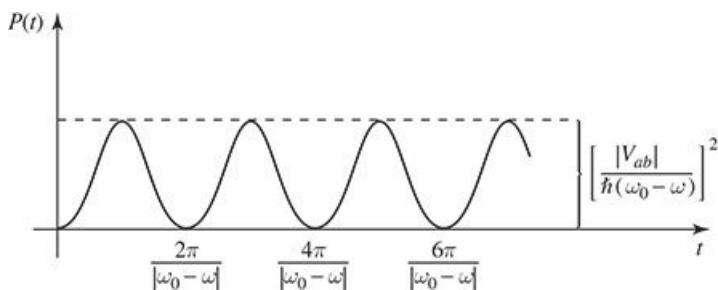


Figure 11.4: Transition probability as a function of time, for a sinusoidal perturbation (Equation 11.35).

As I noted earlier, the probability of a transition is greatest when the driving frequency is close to the “natural” frequency, ω_0 .³ This is illustrated in Figure 11.5, where $P_{a \rightarrow b}$ is plotted as a function of ω . The peak has a height of $(|V_{ab}| t / 2\hbar)^2$ and a width $4\pi/t$; evidently it gets higher and narrower as time goes on. (Ostensibly, the maximum increases without limit. However, the perturbation assumption breaks down before it gets close to 1, so we can believe the result only for relatively small t . In Problem 11.9 you will see that the *exact* result never exceeds 1.)

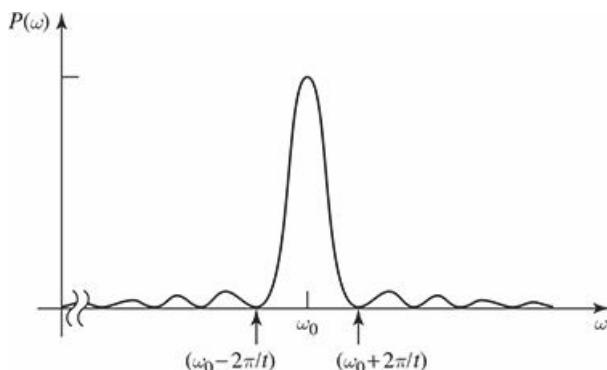


Figure 11.5: Transition probability as a function of driving frequency (Equation 11.35).

**

Problem 11.9 The first term in Equation 11.32 comes from the $e^{i\omega t}/2$ part of $\cos(\omega t)$, and the second from $e^{-i\omega t}/2$. Thus dropping the first term is formally equivalent to writing $\hat{H}' = (V/2)e^{-i\omega t}$, which is to say,

$$H'_{ba} = \frac{V_{ba}}{2} e^{-i\omega t}, \quad H'_{ab} = \frac{V_{ab}}{2} e^{i\omega t}. \quad (11.36)$$

(The latter is required to make the Hamiltonian matrix hermitian—or, if you prefer, to pick out the dominant term in the formula analogous to Equation 11.32 for $c_a(t)$.) Rabi noticed that if you make this so-called **rotating wave approximation** at the *beginning* of the calculation, Equation 11.17 can be solved exactly, with no need for perturbation theory, and no assumption about the strength of the field.

- (a) Solve Equation 11.17 in the rotating wave approximation (Equation 11.36), for the usual initial conditions: $c_a(0) = 1$, $c_b(0) = 0$. Express your results ($c_a(t)$ and $c_b(t)$) in terms of the Rabi flopping frequency,

$$\omega_r \equiv \frac{1}{2} \sqrt{(\omega - \omega_0)^2 + (|V_{ab}|/\hbar)^2}. \quad (11.37)$$

- (b) Determine the transition probability, $P_{a \rightarrow b}(t)$, and show that it never exceeds 1. Confirm that $|c_a(t)|^2 + |c_b(t)|^2 = 1$.
- (c) Check that $P_{a \rightarrow b}(t)$ reduces to the perturbation theory result (Equation 11.35) when the perturbation is “small,” and state precisely what small means in this context, as a constraint on V .
- (d) At what time does the system first return to its initial state?

11.2 Emission and Absorption of Radiation

11.2.1 Electromagnetic Waves

An electromagnetic wave (I'll refer to it as "light", though it could be infrared, ultraviolet, microwave, x-ray, etc.; these differ only in their frequencies) consists of transverse (and mutually perpendicular) oscillating electric and magnetic fields (Figure 11.6). An atom, in the presence of a passing light wave, responds primarily to the electric component. If the wavelength is long (compared to the size of the atom), we can ignore the *spatial* variation in the field;⁴ the atom, then, is exposed to a sinusoidally oscillating electric field

$$\mathbf{E} = E_0 \cos(\omega t) \hat{k} \quad (11.38)$$

(for the moment I'll assume the light is monochromatic, and polarized along the z direction). The perturbing Hamiltonian is⁵

$$H' = -q E_0 z \cos(\omega t), \quad (11.39)$$

where q is the charge of the electron.⁶ Evidently⁷

$$H'_{ba} = -\wp E_0 \cos(\omega t), \quad \text{where } \wp \equiv q \langle \psi_b | z | \psi_a \rangle. \quad (11.40)$$

Typically, ψ is an even or odd function of z ; in either case $z |\psi|^2$ is odd, and integrates to zero (this is Laporte's rule, Section 6.4.3; for some examples see Problem 11.2). This licenses our usual assumption that the diagonal matrix elements of \hat{H}' vanish. Thus the interaction of light with matter is governed by precisely the kind of oscillatory perturbation we studied in Section 11.1.3, with

$$V_{ba} = -\wp E_0. \quad (11.41)$$

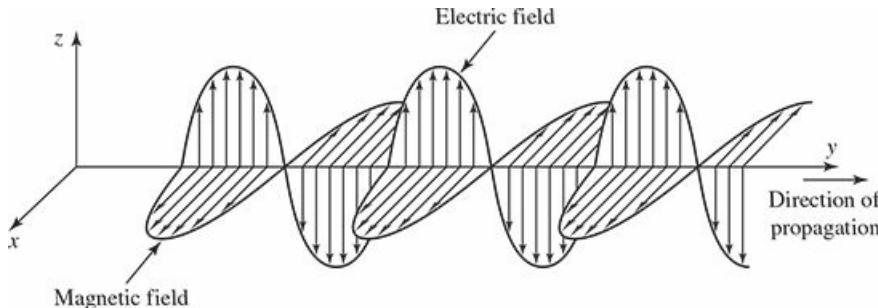


Figure 11.6: An electromagnetic wave.

11.2.2 Absorption, Stimulated Emission, and Spontaneous Emission

If an atom starts out in the “lower” state ψ_a and you shine a polarized monochromatic beam of light on it, the probability of a transition to the “upper” state ψ_b is given by Equation 11.35, which (in view of Equation 11.41) takes the form

$$P_{a \rightarrow b}(t) = \left(\frac{|\phi| E_0}{\hbar} \right)^2 \frac{\sin^2 [(\omega_0 - \omega) t / 2]}{(\omega_0 - \omega)^2}. \quad (11.42)$$

In this process, the atom absorbs energy $E_b - E_a = \hbar\omega_0$ from the electromagnetic field, so it’s called **absorption**. (Informally, we say that the atom has “absorbed a photon” (Figure 11.7(a).) Technically, the word “photon” belongs to **quantum electrodynamics**—the quantum theory of the electromagnetic field—whereas we are treating the field itself *classically*. But this language is convenient, as long as you don’t read too much into it.)

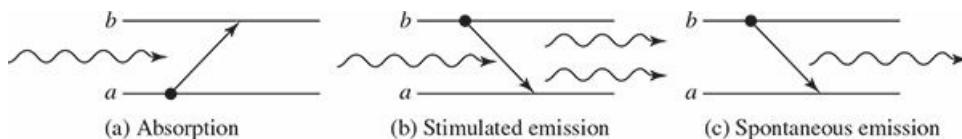


Figure 11.7: Three ways in which light interacts with atoms: (a) absorption, (b) stimulated emission, (c) spontaneous emission.

I could, of course, go back and run the whole derivation for a system that starts off in the *upper* state ($c_a(0) = 0, c_b(0) = 1$). Do it for yourself, if you like; it comes out *exactly the same*—except that this time we’re calculating $P_{b \rightarrow a} = |c_a(t)|^2$, the probability of a transition *down* to the *lower* level:

$$P_{b \rightarrow a}(t) = \left(\frac{|\phi| E_0}{\hbar} \right)^2 \frac{\sin^2 [(\omega_0 - \omega) t / 2]}{(\omega_0 - \omega)^2}. \quad (11.43)$$

(It *has* to come out this way—all we’re doing is switching $a \leftrightarrow b$, which substitutes $-\omega_0$ for ω_0 . When we get to Equation 11.32 we now keep the *first* term, with $-\omega_0 + \omega$ in the denominator, and the rest is the same as before.) But when you stop to think of it, this is an absolutely *astonishing* result: If the particle is in the *upper* state, and you shine light on it, it can make a transition to the *lower* state, and in fact the probability of such a transition is exactly the same as for a transition *upward* from the *lower* state. This process, which was first predicted by Einstein, is called **stimulated emission**.

In the case of stimulated emission the electromagnetic field *gains* energy $\hbar\omega_0$ from the atom; we say that one photon went in and *two* photons came out—the original one that caused the transition plus another one from the transition itself (Figure 11.7(b)). This raises the possibility of *amplification*, for if I had a bottle of atoms, all in the upper state, and triggered it with a single incident photon, a chain reaction would occur, with the first photon producing two, these two producing four, and so on. We’d have an enormous number of photons coming out, all with the same frequency and at virtually the same instant. This is the principle behind the **laser** (Light Amplification by Stimulated Emission of Light). Note that it is essential (for laser action) to get a majority of the atoms into the upper state (a so-called **population inversion**), because **absorption** (which *costs* one photon) competes with stimulated emission (which *creates* one); if you started with an even mixture of the two states, you’d get no amplification at all.

There is a *third* mechanism (in addition to absorption and stimulated emission) by which radiation interacts with matter; it is called **spontaneous emission**. Here an atom in the excited state makes a transition downward, with the release of a photon, but *without* any applied electromagnetic field to initiate the process (Figure 11.7(c)). This is the mechanism that accounts for the normal decay of an atomic excited state. At first sight it is far from clear why spontaneous emission should occur at *all*. If the atom is in a stationary state (albeit an excited one), and there is no external perturbation, it should just sit there forever. And so it *would*, if it were *really* free of all external perturbations. However, in quantum electrodynamics the fields are nonzero *even in the ground state*—just as the harmonic oscillator (for example) has nonzero energy (to wit: $\hbar\omega/2$) in its ground state. You can turn out all the lights, and cool the room down to absolute zero, but there is still some electromagnetic radiation present, and it is this “zero point” radiation that serves to catalyze spontaneous emission. When you come right down to it, there is really no such thing as *truly* spontaneous emission; it’s *all* stimulated emission. The only distinction to be made is whether the field that does the stimulating is one that *you* put there, or one that *God* put there. In this sense it is exactly the reverse of the classical radiative process, in which it’s *all* spontaneous, and there is no such thing as *stimulated* emission.

Quantum electrodynamics is beyond the scope of this book,⁸ but there is a lovely argument, due to Einstein,⁹ which interrelates the three processes (absorption, stimulated emission, and spontaneous emission). Einstein did not identify the *mechanism* responsible for spontaneous emission (perturbation by the ground-state electromagnetic field), but his results nevertheless enable us to calculate the spontaneous emission rate, and from that the natural lifetime of an excited atomic state.¹⁰ Before we turn to that, however, we need to consider the response of an atom to non-monochromatic, unpolarized, incoherent electromagnetic waves coming in from all directions—such as it would encounter, for instance, if it were immersed in thermal radiation.

11.2.3 Incoherent Perturbations

The energy density in an electromagnetic wave is¹¹

$$u = \frac{\epsilon_0}{2} E_0^2, \quad (11.44)$$

where E_0 is (as before) the amplitude of the electric field. So the transition probability (Equation 11.43) is (not surprisingly) proportional to the energy density of the fields:

$$P_{b \rightarrow a}(t) = \frac{2u}{\epsilon_0 \hbar^2} |\phi|^2 \frac{\sin^2 [(\omega_0 - \omega) t/2]}{(\omega_0 - \omega)^2}. \quad (11.45)$$

But this is for a **monochromatic** wave, at a single frequency ω . In many applications the system is exposed to electromagnetic waves at a whole *range* of frequencies; in that case $u \rightarrow \rho(\omega)d\omega$, where $\rho(\omega)d\omega$ is the energy density in the frequency range $d\omega$, and the net transition probability takes the form of an integral:¹²

$$P_{b \rightarrow a}(t) = \frac{2}{\epsilon_0 \hbar^2} |\phi|^2 \int_0^\infty \rho(\omega) \left\{ \frac{\sin^2 [(\omega_0 - \omega) t/2]}{(\omega_0 - \omega)^2} \right\} d\omega. \quad (11.46)$$

The term in curly brackets is sharply peaked about ω_0 (Figure 11.5), whereas $\rho(\omega)$ is ordinarily quite broad, so we may as well replace $\rho(\omega)$ by $\rho(\omega_0)$, and take it outside the integral:

$$P_{b \rightarrow a}(t) \approx \frac{2|\phi|^2}{\epsilon_0 \hbar^2} \rho(\omega_0) \int_0^\infty \frac{\sin^2 [(\omega_0 - \omega) t/2]}{(\omega_0 - \omega)^2} d\omega. \quad (11.47)$$

Changing variables to $x \equiv (\omega_0 - \omega) t/2$, extending the limits of integration to $x = \pm\infty$ (since the integrand is essentially zero out there anyway), and looking up the definite integral

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = \pi, \quad (11.48)$$

we find

$$P_{b \rightarrow a}(t) \approx \frac{\pi |\phi|^2}{\epsilon_0 \hbar^2} \rho(\omega_0) t. \quad (11.49)$$

This time the transition probability is proportional to t . The bizarre “flopping” phenomenon characteristic of a monochromatic perturbation gets “washed out” when we hit the system with an incoherent spread of frequencies. In particular, the **transition rate** ($R \equiv dP/dt$) is now a *constant*:

$$R_{b \rightarrow a} = \frac{\pi}{\epsilon_0 \hbar^2} |\phi|^2 \rho(\omega_0). \quad (11.50)$$

Up to now, we have assumed that the perturbing wave is coming in along the y direction (Figure 11.6), and polarized in the z direction. But we are interested in the case of an atom bathed in radiation coming from *all* directions, and with all possible polarizations; the energy in the fields ($\rho(\omega)$) is shared equally among these different modes. What we need, in place of $|\phi|^2$, is the *average* of $|\phi \cdot \hat{n}|^2$, where

$$\phi \equiv q \langle \psi_b | \mathbf{r} | \psi_a \rangle \quad (11.51)$$

(generalizing Equation 11.40), and the average is over all polarizations and all incident directions.

The averaging can be carried out as follows: Choose spherical coordinates such that the direction of propagation (\hat{k}) is along x , the polarization (\hat{n}) is along z , and the vector ρ defines the spherical angles θ and ϕ (Figure 11.8).¹³ (Actually, ρ is *fixed*, here, and we're averaging over all \hat{k} and \hat{n} consistent with $\hat{k} \perp \hat{n}$ —which is to say, over all θ and ϕ . But we might as well integrate over all directions of ρ , keeping \hat{k} and \hat{n} fixed—it amounts to the same thing.) Then

$$\rho \cdot \hat{n} = \rho \cos \theta, \quad (11.52)$$

and

$$\begin{aligned} |\rho \cdot \hat{n}|_{\text{ave}}^2 &= \frac{1}{4\pi} \int |\rho|^2 \cos^2 \theta \sin \theta d\theta d\phi \\ &= \frac{|\rho|^2}{4\pi} \left(-\frac{\cos^3 \theta}{3} \right) \Big|_0^\pi (2\pi) = \frac{1}{3} |\rho|^2. \end{aligned} \quad (11.53)$$

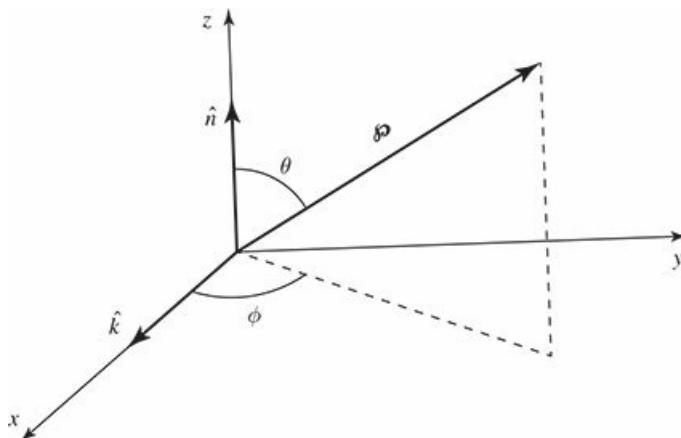


Figure 11.8: Axes for the averaging of $|\rho \cdot \hat{n}|^2$.

Conclusion: The transition rate for stimulated emission from state b to state a , under the influence of incoherent, unpolarized light incident from all directions, is

$$R_{b \rightarrow a} = \frac{\pi}{3\epsilon_0 \hbar^2} |\rho|^2 \rho(\omega_0), \quad (11.54)$$

where ρ is the matrix element of the electric dipole moment between the two states (Equation 11.51), and $\rho(\omega_0)$ is the energy density in the fields, per unit frequency, evaluated at $\omega_0 = (E_b - E_a)/\hbar$.

11.3 Spontaneous Emission

11.3.1 Einstein's *A* and *B* Coefficients

Picture a container of atoms, N_a of them in the lower state (ψ_a), and N_b of them in the upper state (ψ_b). Let A be the spontaneous emission rate,¹⁴ so that the number of particles leaving the upper state by this process, per unit time, is $N_b A$.¹⁵ The transition rate for stimulated emission, as we have seen (Equation 11.54), is proportional to the energy density of the electromagnetic field: $B_{ba}\rho(\omega_0)$, where $B_{ba} = \pi |\mathbf{p}|^2 / 3\epsilon_0\hbar^2$; the number of particles leaving the upper state by this mechanism, per unit time, is $N_b B_{ba}\rho(\omega_0)$. The absorption rate is likewise proportional to $\rho(\omega_0)$ —call it $B_{ab}\rho(\omega_0)$; the number of particles per unit time *joining* the upper level is therefore $N_a B_{ab}\rho(\omega_0)$. All told, then,

$$\frac{dN_b}{dt} = -N_b A - N_b B_{ba}\rho(\omega_0) + N_a B_{ab}\rho(\omega_0). \quad (11.55)$$

Suppose these atoms are in thermal equilibrium with the ambient field, so that the number of particles in each level is *constant*. In that case $dN_b/dt = 0$, and it follows that

$$\rho(\omega_0) = \frac{A}{(N_a/N_b) B_{ab} - B_{ba}}. \quad (11.56)$$

On the other hand, we know from statistical mechanics¹⁶ that the number of particles with energy E , in thermal equilibrium at temperature T , is proportional to the **Boltzmann factor**, $\exp(-E/k_B T)$, so

$$\frac{N_a}{N_b} = \frac{e^{-E_a/k_B T}}{e^{-E_b/k_B T}} = e^{\hbar\omega_0/k_B T}, \quad (11.57)$$

and hence

$$\rho(\omega_0) = \frac{A}{e^{\hbar\omega_0/k_B T} B_{ab} - B_{ba}}. \quad (11.58)$$

But Planck's blackbody formula¹⁷ tells us the energy density of thermal radiation:

$$\rho(\omega) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\hbar\omega/k_B T} - 1}; \quad (11.59)$$

comparing the two expressions, we conclude that

$$B_{ab} = B_{ba} \quad (11.60)$$

and

$$A = \frac{\omega_0^3 \hbar}{\pi^2 c^3} B_{ba}. \quad (11.61)$$

Equation 11.60 confirms what we already knew: the transition rate for stimulated emission is the same as for absorption. But it was an astonishing result in 1917—indeed, Einstein was forced to “invent” stimulated emission in order to reproduce Planck’s formula. Our present attention, however, focuses on Equation 11.61, for this tells us the spontaneous emission rate (A)—which is what we are looking for—in terms of the stimulated emission rate ($B_{ba}\rho(\omega_0)$)—which we already know. From Equation 11.54 we read off

$$B_{ba} = \frac{\pi}{3\epsilon_0\hbar^2} |\mathbf{p}|^2, \quad (11.62)$$

and it follows that the spontaneous emission rate is

$$A = \frac{\omega_0^3 |\rho|^2}{3\pi\epsilon_0 hc^3}. \quad (11.63)$$

Problem 11.10 As a mechanism for downward transitions, spontaneous emission competes with thermally stimulated emission (stimulated emission for which blackbody radiation is the source). Show that at room temperature ($T = 300$ K) thermal stimulation dominates for frequencies well below 5×10^{12} Hz, whereas spontaneous emission dominates for frequencies well above 5×10^{12} Hz. Which mechanism dominates for visible light?

Problem 11.11 You could derive the spontaneous emission rate (Equation 11.63) without the detour through Einstein's A and B coefficients if you knew the ground state energy density of the electromagnetic field, $\rho_0(\omega)$, for then it would simply be a case of stimulated emission (Equation 11.54). To do this honestly would require quantum electrodynamics, but if you are prepared to believe that the ground state consists of *one photon in each classical mode*, then the derivation is fairly simple:

- (a) To obtain the classical modes, consider an empty cubical box, of side l , with one corner at the origin. Electromagnetic fields (in vacuum) satisfy the classical wave equation¹⁸

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) f(x, y, z, t) = 0,$$

where f stands for any component of \mathbf{E} or of \mathbf{B} . Show that separation of variables, and the imposition of the boundary condition $f = 0$ on all six surfaces yields the standing wave patterns

$$f_{n_x, n_y, n_z} = A \cos(\omega t) \sin\left(\frac{n_x \pi}{l} x\right) \sin\left(\frac{n_y \pi}{l} y\right) \sin\left(\frac{n_z \pi}{l} z\right),$$

with

$$\omega = \frac{\pi c}{l} \sqrt{n_x^2 + n_y^2 + n_z^2}.$$

There are two modes for each triplet of positive integers $(n_x, n_y, n_z = 1, 2, 3, \dots)$, corresponding to the two polarization states.

- (b) The energy of a photon is $E = h\nu = \hbar\omega$ (Equation 4.92), so the energy in the mode (n_x, n_y, n_z) is

$$E_{n_x, n_y, n_z} = 2 \frac{\pi \hbar c}{l} \sqrt{n_x^2 + n_y^2 + n_z^2}.$$

What, then, is the *total* energy per unit volume in the frequency range $d\omega$, if each mode gets one photon? Express your answer in the form

$$\frac{1}{l^3} dE = \rho_0(\omega) d\omega$$

and read off $\rho_0(\omega)$. Hint: refer to Figure 5.3.

- (c) Use your result, together with Equation 11.54, to obtain the spontaneous emission rate. Compare Equation 11.63.

11.3.2 The Lifetime of an Excited State

Equation 11.63 is our fundamental result; it gives the transition rate for spontaneous emission. Suppose, now, that you have somehow pumped a large number of atoms into the excited state. As a result of spontaneous emission, this number will decrease as time goes on; specifically, in a time interval dt you will lose a fraction A dt of them:

$$dN_b = -AN_b dt, \quad (11.64)$$

(assuming there is no mechanism to replenish the supply).¹⁹ Solving for $N_b(t)$, we find:

$$N_b(t) = N_b(0)e^{-At}; \quad (11.65)$$

evidently the number remaining in the excited state decreases exponentially, with a time constant

$$\tau = \frac{1}{A}. \quad (11.66)$$

We call this the **lifetime** of the state—technically, it is the time it takes for $N_b(t)$ to reach $1/e \approx 0.368$ of its initial value.

I have assumed all along that there are only *two* states for the system, but this was just for notational simplicity—the spontaneous emission formula (Equation 11.63) gives the transition rate for $\psi_b \rightarrow \psi_a$ regardless of what other states may be accessible (see Problem 11.24). Typically, an excited atom has many different **decay modes** (that is: ψ_b can decay to a large number of different lower-energy states, $\psi_{a_1}, \psi_{a_2}, \psi_{a_3}, \dots$). In that case the transition rates *add*, and the net lifetime is

$$\tau = \frac{1}{A_1 + A_2 + A_3 + \dots}. \quad (11.67)$$

Example 11.1

Suppose a charge q is attached to a spring and constrained to oscillate along the x axis. Say it starts out in the state $|n\rangle$ (Equation 2.68), and decays by spontaneous emission to state $|n'\rangle$. From Equation 11.51 we have

$$\wp = q \langle n | x | n' \rangle \hat{i}.$$

You calculated the matrix elements of x back in Problem 3.39:

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

where ω is the natural frequency of the oscillator (I no longer need this letter for the frequency of the stimulating radiation). But we're talking about *emission*, so n' must be *lower* than n ; for our purposes, then,

$$\wp = q \sqrt{\frac{n\hbar}{2m\omega}} \delta_{n',n-1} \hat{i}. \quad (11.68)$$

Evidently transitions occur only to states one step lower on the “ladder”, and the frequency of the

Evidently transitions occur only to states one step lower on the “ladder”, and the frequency of the photon emitted is

$$\omega_0 = \frac{E_n - E_{n'}}{\hbar} = \frac{(n + 1/2) \hbar\omega - (n' + 1/2) \hbar\omega}{\hbar} = (n - n') \omega = \omega. \quad (11.69)$$

Not surprisingly, the system radiates at the classical oscillator frequency. The transition rate (Equation 11.63) is

$$A = \frac{nq^2\omega^2}{6\pi\epsilon_0mc^3}, \quad (11.70)$$

and the lifetime of the n th stationary state is

$$\tau_n = \frac{6\pi\epsilon_0mc^3}{nq^2\omega^2}. \quad (11.71)$$

Meanwhile, each radiated photon carries an energy $\hbar\omega$, so the *power* radiated is $A\hbar\omega$:

$$P = \frac{q^2\omega^2}{6\pi\epsilon_0mc^3} (n\hbar\omega),$$

or, since the energy of an oscillator in the n th state is $E = (n + 1/2) \hbar\omega$,

$$P = \frac{q^2\omega^2}{6\pi\epsilon_0mc^3} \left(E - \frac{1}{2}\hbar\omega \right). \quad (11.72)$$

This is the average power radiated by a quantum oscillator with (initial) energy E .

For comparison, let's determine the average power radiated by a *classical* oscillator with the same energy. According to classical electrodynamics, the power radiated by an accelerating charge q is given by the **Larmor formula**:²⁰

$$P = \frac{q^2a^2}{6\pi\epsilon_0c^3}. \quad (11.73)$$

For a harmonic oscillator with amplitude x_0 , $x(t) = x_0 \cos(\omega t)$, and the acceleration is $a = -x_0\omega^2 \cos(\omega t)$. Averaging over a full cycle, then,

$$P = \frac{q^2x_0^2\omega^4}{12\pi\epsilon_0c^3}.$$

But the *energy* of the oscillator is $E = (1/2)m\omega^2x_0^2$, so $x_0^2 = 2E/m\omega^2$, and hence

$$P = \frac{q^2\omega^2}{6\pi\epsilon_0mc^3} E. \quad (11.74)$$

This is the average power radiated by a *classical* oscillator with energy E . In the classical limit ($\hbar \rightarrow 0$) the classical and quantum formulas agree;²¹ however, the quantum formula (Equation 11.72) protects the ground state: If $E = (1/2)\hbar\omega$ the oscillator does not radiate.

Problem 11.12 The half-life ($t_{1/2}$) of an excited state is the time it would take for half the atoms in a large sample to make a transition. Find the relation between $t_{1/2}$ and τ (the “lifetime” of the state).

- * **Problem 11.13** Calculate the lifetime (in seconds) for each of the four $n = 2$ states of hydrogen. *Hint:* You’ll need to evaluate matrix elements of the form $\langle \psi_{100} | x | \psi_{200} \rangle$, $\langle \psi_{100} | y | \psi_{211} \rangle$, and so on. Remember that $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$, and $z = r \cos \theta$. Most of these integrals are zero, so inspect them closely before you start calculating. *Answer:* 1.60×10^{-9} seconds for all except ψ_{200} , which is infinite.

11.3.3 Selection Rules

The calculation of spontaneous emission rates has been reduced to a matter of evaluating matrix elements of the form

$$\langle \psi_b | \mathbf{r} | \psi_a \rangle.$$

As you will have discovered if you worked Problem 11.13, (if you *didn't*, go back right now and *do* so!) these quantities are very often *zero*, and it would be helpful to know in advance when this is going to happen, so we don't waste a lot of time evaluating unnecessary integrals. Suppose we are interested in systems like hydrogen, for which the Hamiltonian is spherically symmetrical. In that case we can specify the states with the usual quantum numbers n , ℓ , and m , and the matrix elements are

$$\langle n' \ell' m' | \mathbf{r} | n \ell m \rangle.$$

Now, \mathbf{r} is a vector operator, and we can invoke the results of Chapter 6 to obtain the **selection rules**²²

$$\Delta\ell \equiv \ell' - \ell = \pm 1, \quad \Delta m \equiv m' - m = 0 \text{ or } \pm 1.$$

(11.75)

These conditions follow from symmetry alone. If they are *not* met, then the matrix element is zero, and the transition is said to be **forbidden**. Moreover, it follows from Equations 6.56–6.58 that

$$\begin{cases} \text{if } m' = m, & \text{then } \langle n' \ell' m' | x | n \ell m \rangle = \langle n' \ell' m' | y | n \ell m \rangle = 0, \\ \text{if } m' = m \pm 1, & \text{then } \langle n' \ell' m' | x | n \ell m \rangle = \pm i \langle n' \ell' m' | y | n \ell m \rangle \\ & \text{and } \langle n' \ell' m' | z | n \ell m \rangle = 0. \end{cases} \quad (11.76)$$

So it is never necessary to compute the matrix elements of both x and y ; you can always get one from the other.

Evidently not all transitions to lower-energy states can proceed by electric dipole radiation; most are forbidden by the selection rules. The scheme of allowed transitions for the first four Bohr levels in hydrogen is shown in Figure 11.9. Notice that the $2S$ state (ψ_{200}) is “stuck”: it cannot decay, because there is no lower-energy state with $\ell = 1$. It is called a **metastable** state, and its lifetime is indeed much longer than that of, for example, the $2P$ states (ψ_{211} , ψ_{210} , and ψ_{21-1}). Metastable states do eventually decay, by collisions, or by “forbidden” transitions (Problem 11.31), or by multiphoton emission.

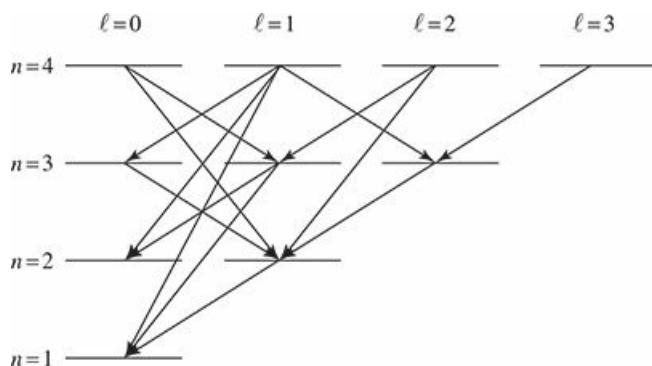


Figure 11.9: Allowed decays for the first four Bohr levels in hydrogen.

Problem 11.14 From the commutators of L_z with x, y , and z (Equation 4.122):

$$[L_z, x] = i\hbar y, \quad [L_z, y] = -i\hbar x, \quad [L_z, z] = 0, \quad (11.77)$$

obtain the selection rule for Δm and Equation 11.76. *Hint:* Sandwich each commutator between $\langle n'\ell'm' |$ and $|n\ell m\rangle$.

**

Problem 11.15 Obtain the selection rule for $\Delta \ell$ as follows:

- (a) Derive the commutation relation

$$[L^2, [L^2, \mathbf{r}]] = 2\hbar^2 (\mathbf{r}L^2 + L^2\mathbf{r}). \quad (11.78)$$

Hint: First show that

$$[L^2, z] = 2i\hbar(xL_y - yL_x - i\hbar z).$$

Use this, and (in the final step) the fact that $\mathbf{r} \cdot \mathbf{L} = \mathbf{r} \cdot (\mathbf{r} \times \mathbf{p}) = \mathbf{0}$, to demonstrate that

$$[L^2, [L^2, z]] = 2\hbar^2 (zL^2 + L^2z).$$

The generalization from z to \mathbf{r} is trivial.

- (b) Sandwich this commutator between $\langle n'\ell'm' |$ and $|n\ell m\rangle$, and work out the implications.

**

Problem 11.16 An electron in the $n = 3, \ell = 0, m = 0$ state of hydrogen decays by a sequence of (electric dipole) transitions to the ground state.

- (a) What decay routes are open to it? Specify them in the following way:

$$|300\rangle \rightarrow |n\ell m\rangle \rightarrow |n'\ell'm'\rangle \rightarrow \dots \rightarrow |100\rangle.$$

- (b) If you had a bottle full of atoms in this state, what fraction of them would decay via each route?
- (c) What is the lifetime of this state? *Hint:* Once it's made the first transition, it's no longer in the state $|300\rangle$, so only the first step in each sequence is relevant in computing the lifetime.

11.4 Fermi's Golden Rule

In the previous sections we considered transitions between two *discrete* energy states, such as two bound states of an atom. We saw that such a transition was most likely when the final energy satisfied the resonance condition: $E_f = E_i + \hbar\omega$, where ω is the frequency associated with the perturbation. I now want to look at the case where E_f falls in a *continuum* of states (Figure 11.10). To stick close to the example of Section 11.2, if the radiation is energetic enough it can ionize the atom—the **photoelectric effect**—exciting the electron from a bound state into the continuum of scattering states.

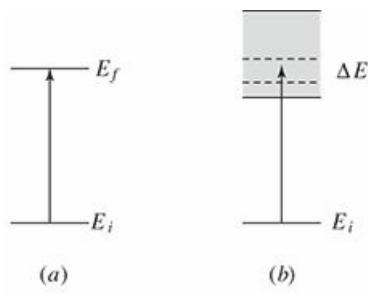


Figure 11.10: A transition (a) between two discrete states and (b) between a discrete state and a continuum of states.

We can't talk about a transition to a *precise* state in that continuum (any more than we can talk about someone being *precisely* 16 years old), but we can compute the probability that the system makes a transition to a state with an energy in some finite range ΔE about E_f . That is given by the integral of Equation 11.35 over all the final states:

$$P = \int_{E_f - \Delta E/2}^{E_f + \Delta E/2} \frac{|V_{in}|^2}{\hbar^2} \left\{ \frac{\sin^2 [(\omega_0 - \omega)t/2]}{(\omega_0 - \omega)^2} \right\} \rho(E_n) dE_n, \quad (11.79)$$

where $\omega_0 = (E_n - E_i)/\hbar$. The quantity $\rho(E)dE$ is the number of states with energy between E and $E + dE$; $\rho(E)$ is called the **density of states**, and I'll show you how it's calculated in Example 11.2.

At short times, Equation 11.79 leads to a transition probability proportional to t^2 , just as for a transition between discrete states. On the other hand, at long times the quantity in curly brackets in Equation 11.79 is sharply peaked: as a function of E_n its maximum occurs at $E_f = E_i + \hbar\omega$ and the central peak has a width of $4\pi\hbar/t$. For sufficiently large t , we can therefore approximate Equation 11.79 as²³

$$P = \frac{|V_{if}|^2}{\hbar^2} \rho(E_f) \int_{-\infty}^{\infty} \frac{\sin^2 [(\omega_0 - \omega)t/2]}{(\omega_0 - \omega)^2} dE_n.$$

The remaining integral was already evaluated in Section 11.2.3:

$$P = \frac{2\pi}{\hbar} \left| \frac{V_{if}}{2} \right|^2 \rho(E_f) t. \quad (11.80)$$

The oscillatory behavior of P has again been “washed out,” giving a constant transition rate:²⁴

$$R = \frac{2\pi}{\hbar} \left| \frac{V_{if}}{2} \right|^2 \rho(E_f). \quad (11.81)$$

Equation 11.81 is known as **Fermi's Golden Rule**.²⁵ Apart from the factor of $2\pi/\hbar$, it says that the transition rate is the square of the matrix element (this encapsulates all the relevant information about the *dynamics* of the process) times the density of states (how many final states are accessible, given the energy supplied by the perturbation—the more roads are open, the faster the traffic will flow). It makes sense.

Example 11.2

Use Fermi's Golden Rule to obtain the differential scattering cross-section for a particle of mass m and incident wave vector \mathbf{k}' scattering from a potential $V(\mathbf{r})$ (Figure 11.11).

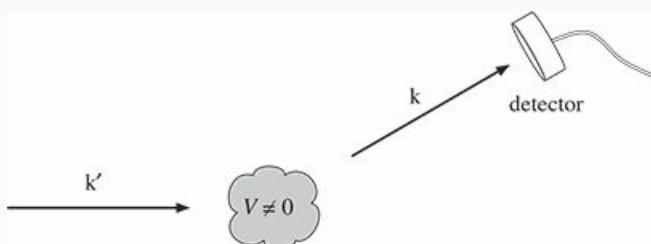


Figure 11.11: A particle with incident wave vector \mathbf{k}' is scattered into a state with wave vector \mathbf{k} .

Solution:

We take our initial and final states to be plane waves:

$$\psi_i = \frac{1}{\sqrt{l^3}} e^{i\mathbf{k}' \cdot \mathbf{r}} \text{ and } \psi_f = \frac{1}{\sqrt{l^3}} e^{i\mathbf{k} \cdot \mathbf{r}}. \quad (11.82)$$

Here I've used a technique called **box normalization**; I place the whole setup inside a box of length l on a side. This makes the free-particle states normalizable and countable. Formally, we want the limit $l \rightarrow \infty$; in practice l will drop out of our final expression. Using periodic boundary conditions,²⁶ the allowed values of \mathbf{k} are

$$\mathbf{k} = \frac{2\pi}{l} (n_x \hat{i} + n_y \hat{j} + n_z \hat{k}) \quad (11.83)$$

for integers n_x , n_y , and n_z . Our perturbation is the scattering potential, $\hat{H}' = V(\mathbf{r})$, and the relevant matrix element is

$$V_{fi} = \int \psi_f^*(\mathbf{r}) V(\mathbf{r}) \psi_i(\mathbf{r}) d^3\mathbf{r} = \frac{1}{l^3} \int e^{i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}} V(\mathbf{r}) d^3\mathbf{r}. \quad (11.84)$$

We need to determine the density of states. In a scattering experiment we measure the number of particles scattered into a solid angle $d\Omega$. We want to count the number of states with energies between E and $E + dE$, with wave vectors \mathbf{k} lying inside $d\Omega$. In k space these states occupy a section of a spherical shell of radius k and thickness dk that subtends a solid angle $d\Omega$; it has a volume

$$k^2 dk d\Omega$$

and contains a number of states²⁷

$$\rho(E) dE = \frac{k^2 dk d\Omega}{(2\pi/l)^3} = \left(\frac{l}{2\pi}\right)^3 k^2 \frac{dk}{dE} dE d\Omega.$$

Since $E = \hbar^2 k^2 / 2m$ this gives

$$\rho(E) = \left(\frac{l}{2\pi}\right)^3 \frac{\sqrt{2m^3 E}}{\hbar^3} d\Omega. \quad (11.85)$$

From Fermi's Golden Rule, the rate at which particles are scattered into the solid angle $d\Omega$ is²⁸

$$R_{i \rightarrow d\Omega} = \frac{2\pi}{\hbar} \frac{1}{l^6} \left| \int e^{i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}} V(\mathbf{r}) d^3 \mathbf{r} \right|^2 \left(\frac{l}{2\pi}\right)^3 \frac{\sqrt{2m^3 E_f}}{\hbar^3} d\Omega.$$

This is closely related to the differential scattering cross section:

$$\frac{d\sigma}{d\Omega} = \frac{R_{i \rightarrow d\Omega}}{J_i d\Omega} \quad (11.86)$$

where J_i is the flux (or probability current) of incident particles. For an incident wave of the form $\psi_i = A e^{i\mathbf{k}' \cdot \mathbf{r}}$, the probability current is (Equation 4.220).

$$J_i = |A|^2 v = \frac{1}{l^3} \frac{\hbar k'}{m} \quad (11.87)$$

and

$$\frac{d\sigma}{d\Omega} = \left| -\frac{m}{2\pi \hbar^2} \int e^{i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}} V(\mathbf{r}) d^3 \mathbf{r} \right|^2. \quad (11.88)$$

This is exactly what we got from the first Born approximation (Equation 10.79).

Problem 11.17 In the photoelectric effect, light can ionize an atom if its energy ($\hbar\omega$) exceeds the binding energy of the electron. Consider the photoelectric effect for the ground state of hydrogen, where the electron is kicked out with momentum $\hbar\mathbf{k}$. The initial state of the electron is $\psi_0(r)$ (Equation 4.80) and its final state is²⁹

$$\psi_f = \frac{1}{\sqrt{l^3}} e^{i\mathbf{k} \cdot \mathbf{r}},$$

as in Example 11.2.

- (a) For light polarized along the z direction, use Fermi's Golden Rule to compute the rate at which electrons are ejected into the solid angle $d\Omega$ in the dipole approximation.³⁰

$$\left[\text{Answer : } R_{i \rightarrow d\Omega} = 256 \frac{\epsilon_0 E_0^2 a^3}{\hbar} \frac{(ka)^3}{[1 + (ka)^2]^6} \cos^2 \theta d\Omega. \right]$$

Hint: To evaluate the matrix element, use the following trick. Write

$$ze^{i\mathbf{k}\cdot\mathbf{r}} = -i \frac{d}{dk_z} e^{i\mathbf{k}\cdot\mathbf{r}},$$

pull d/dk_z outside the integral, and what remains is straightforward to compute.

- (b) The photoelectric cross section is defined as

$$\sigma(k) = \frac{R_{i \rightarrow \text{all}} \hbar \omega}{\frac{1}{2} \epsilon_0 E_0^2 c}$$

where the quantity in the numerator is the rate at which *energy* is absorbed ($\hbar\omega = \frac{\hbar^2 k^2}{2m} - E_1$ per photoelectron) and the quantity in the denominator is the intensity of the incident light. Integrate your result from (a) over all angles to obtain $R_{i \rightarrow \text{all}}$, and compute the photoelectric cross section.

- (c) Obtain a numerical value for the photoelectric cross section for ultraviolet light of wavelength 220\AA (n.b. this is the wavelength of the incident light, not the scattered electron). Express your answer in mega-barns ($\text{Mb} = 10^{-22} \text{m}^2$).

11.5 The Adiabatic Approximation

11.5.1 Adiabatic Processes

Imagine a perfect pendulum, with no friction or air resistance, oscillating back and forth in a vertical plane. If you grab the support and shake it in a jerky manner the bob will swing around chaotically. But if you *very gently* move the support (Figure 11.12), the pendulum will continue to swing in a nice smooth way, in the same plane (or one parallel to it), with the same amplitude. This *gradual change of the external conditions* defines an **adiabatic** process. Notice that there are two characteristic times involved: T_i , the “internal” time, representing the motion of the system itself (in this case the period of the pendulum’s oscillations), and T_e , the “external” time, over which the parameters of the system change appreciably (if the pendulum were mounted on a rotating platform, for example, T_e would be the period of the *platform’s* motion). An adiabatic process is one for which $T_e \gg T_i$ (the pendulum executes many oscillations before the platform has moved appreciably).³¹



Figure 11.12: Adiabatic motion: If the case is transported very gradually, the pendulum inside keeps swinging with the same amplitude, in a plane parallel to the original one.

What if I took this pendulum up to the North Pole, and set it swinging—say, in the direction of Portland (Figure 11.13). For the moment, pretend the earth is not rotating. Very gently (that is, *adiabatically*), I carry it down the longitude line passing through Portland, to the equator. At this point it is swinging north-south. Now I carry it (still swinging north-south) part way around the equator. And finally, I take it back up to the North Pole, along the new longitude line. The pendulum will no longer be swinging in the same plane as it was when I set out—indeed, the new plane makes an angle Θ with the old one, where Θ is the angle between the southbound and the northbound longitude lines. More generally, if you transport the pendulum around a closed loop on the surface of the earth, the angular deviation (between the initial plane of the swing and the final plane) is equal to the solid angle subtended by the path with respect to the center of the sphere, as you can prove for yourself if you are interested.

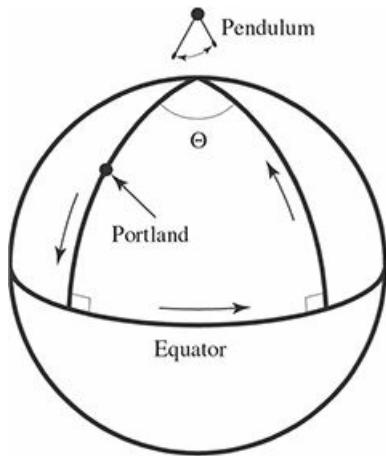


Figure 11.13: Itinerary for adiabatic transport of a pendulum on the surface of the earth.

Incidentally, the **Foucault pendulum** is an example of precisely this sort of adiabatic transport around a closed loop on a sphere—only this time instead of *me* carrying the pendulum around, I let the *rotation of the earth* do the job. The solid angle subtended by a latitude line θ_0 (Figure 11.14) is

$$\Omega = \int \sin \theta d\theta d\phi = 2\pi (-\cos \theta)|_{\theta_0}^{90^\circ} = 2\pi (1 - \cos \theta_0). \quad (11.89)$$

Relative to the earth (which has meanwhile turned through an angle of 2π), the daily precession of the Foucault pendulum is $2\pi \cos \theta_0$ —a result that is ordinarily obtained by appeal to Coriolis forces in the rotating reference frame,³² but is seen in this context to admit a purely *geometrical* interpretation.

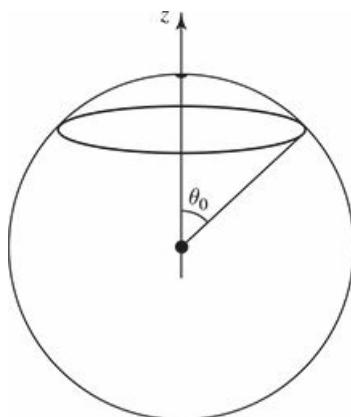


Figure 11.14: Path of a Foucault pendulum, in the course of one day.

The basic strategy for analyzing an adiabatic process is first to solve the problem with the external parameters held *constant*, and only at the *end* of the calculation allow them to vary (slowly) with time. For example, the classical period of a pendulum of (fixed) length L is $2\pi\sqrt{L/g}$; if the length is now gradually *changing*, the period will be $2\pi\sqrt{L(t)/g}$. A more subtle example occurred in our discussion of the hydrogen molecule ion (Section 8.3). We began by assuming that the nuclei were *at rest*, a fixed distance R apart, and we solved for the motion of the electron. Once we had found the ground state energy of the system as a function of R , we located the equilibrium separation and from the curvature of the graph we obtained the frequency of vibration of the nuclei (Problem 8.11). In molecular physics this technique (beginning with

nuclei at rest, calculating electronic wave functions, and using these to obtain information about the positions and—relatively sluggish—motion of the nuclei) is known as the **Born–Oppenheimer approximation**.

11.5.2 The Adiabatic Theorem

In quantum mechanics, the essential content of the **adiabatic approximation** can be cast in the form of a theorem. Suppose the Hamiltonian changes *gradually* from some initial form $\hat{H}(0)$ to some final form $\hat{H}(T)$. The **adiabatic theorem**³³ states that if the particle was initially in the n th eigenstate of $\hat{H}(0)$, it will be carried (under the Schrödinger equation) into the n th eigenstate of $\hat{H}(T)$. (I assume that the spectrum is discrete and nondegenerate throughout the transition, so there is no ambiguity about the ordering of the states; these conditions can be relaxed, given a suitable procedure for “tracking” the eigenfunctions, but I’m not going to pursue that here.)

Example 11.3

Suppose we prepare a particle in the ground state of the infinite square well (Figure 11.15(a)):

$$\psi^i(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi}{a}x\right). \quad (11.90)$$

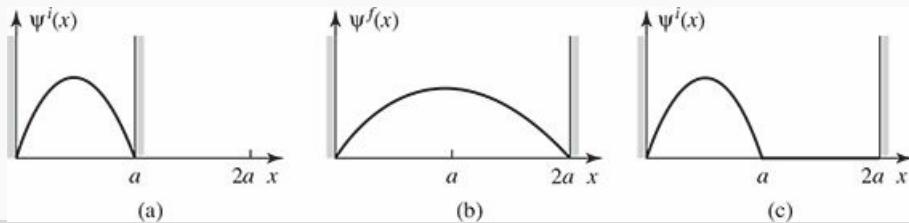


Figure 11.15: (a) Particle starts out in the ground state of the infinite square well. (b) If the wall moves *slowly*, the particle remains in the ground state. (c) If the wall moves *rapidly*, the particle is left (momentarily) in its initial state.

If we now gradually move the right wall out to $2a$, the adiabatic theorem says that the particle will end up in the ground state of the expanded well (Figure 11.15(b)):

$$\psi^f(x) = \sqrt{\frac{1}{a}} \sin\left(\frac{\pi}{2a}x\right), \quad (11.91)$$

(apart from a phase factor, which we’ll discuss in a moment). Notice that we’re not talking about a *small* change in the Hamiltonian (as in perturbation theory)—this one is *huge*. All we require is that it happen *slowly*.

Energy is *not* conserved here—of *course* not: Whoever is moving the wall is extracting energy from the system, just like the piston on a slowly expanding cylinder of gas. By contrast, if the well expands *suddenly*, the resulting state is still $\psi^i(x)$ (Figure 11.15c), which is a complicated linear combination of eigenstates of the new Hamiltonian (Problem 11.18). In this case energy *is* conserved (at least, its *expectation value* is); as in the *free* expansion of a gas (into a vacuum) when the barrier is suddenly removed; no work is done.

According to the adiabatic theorem, a system that starts out in the n th eigenstate of the initial Hamiltonian ($\hat{H}(0)$) will evolve as the n th eigenstate of the instantaneous Hamiltonian ($\hat{H}(t)$), as the

Hamiltonian gradually changes. However, this doesn't tell us what happens to the *phase* of the wave function. For a *constant* Hamiltonian it would pick up the standard "wiggle factor"

$$e^{-iE_n t/\hbar},$$

but the eigenvalue E_n may now itself be a function of time, so the wiggle factor naturally generalizes to

$$e^{i\theta_n(t)}, \quad \text{where} \quad \theta_n(t) \equiv -\frac{1}{\hbar} \int_0^t E_n(t') dt'. \quad (11.92)$$

This is called the **dynamic phase**. But it may not be the end of the story; for all we know there may be an additional phase factor, $\gamma_n(t)$, the so-called **geometric phase**. In the adiabatic limit, then, the wave function at time t takes the form³⁴

$$\Psi_n(t) = e^{i\theta_n(t)} e^{i\gamma_n(t)} \psi_n(t), \quad (11.93)$$

where $\psi_n(t)$ is the n th eigenstate of the instantaneous Hamiltonian,

$$\hat{H}(t)\psi_n(t) = E_n(t)\psi_n(t). \quad (11.94)$$

Equation 11.93 is the formal statement of the adiabatic theorem.

Of course, the phase of $\psi_n(t)$ is itself arbitrary (it's still an eigenfunction, with the same eigenvalue, whatever phase you choose), so the geometric phase itself carries no physical significance. But what if we carry the system around a *closed cycle* (like the pendulum we hauled down to the equator, around, and back to the north pole), so that the Hamiltonian at the end is identical to the Hamiltonian at the beginning? Then the *net* phase change is a measurable quantity. The dynamic phase depends on the elapsed time, but the geometric phase, around an adiabatic closed cycle, depends only on the path taken.³⁵ It is called **Berry's phase**:³⁶

$$\gamma_B \equiv \gamma(T) - \gamma(0). \quad (11.95)$$

Example 11.4

Imagine an electron (charge $-e$, mass m) at rest at the origin, in the presence of a magnetic field whose *magnitude* (B_0) is constant, but whose *direction* sweeps out a cone, of opening angle α , at constant angular velocity ω (Figure 11.16):

$$\mathbf{B}(t) = B_0 \left[\sin \alpha \cos(\omega t) \hat{i} + \sin \alpha \sin(\omega t) \hat{j} + \cos \alpha \hat{k} \right]. \quad (11.96)$$

The Hamiltonian (Equation 4.158) is

$$\begin{aligned} \hat{H}(t) &= \frac{e}{m} \mathbf{B} \cdot \mathbf{S} = \frac{e\hbar B_0}{2m} [\sin \alpha \cos(\omega t) \sigma_x + \sin \alpha \sin(\omega t) \sigma_y + \cos \alpha \sigma_z] \\ &= \frac{\hbar \omega_1}{2} \begin{pmatrix} \cos \alpha & e^{-i\omega t} \sin \alpha \\ e^{i\omega t} \sin \alpha & -\cos \alpha \end{pmatrix}, \end{aligned} \quad (11.97)$$

where

$$\omega_1 \equiv \frac{eB_0}{m}. \quad (11.98)$$

The normalized eigenspinors of $\hat{H}(t)$ are

$$\chi_+(t) = \begin{pmatrix} \cos(\alpha/2) \\ e^{i\omega t} \sin(\alpha/2) \end{pmatrix}, \quad (11.99)$$

and

$$\chi_-(t) = \begin{pmatrix} e^{-i\omega t} \sin(\alpha/2) \\ -\cos(\alpha/2) \end{pmatrix}; \quad (11.100)$$

they represent spin up and spin down, respectively, *along the instantaneous direction of $\mathbf{B}(t)$* (see Problem 4.33). The corresponding eigenvalues are

$$E_{\pm} = \pm \frac{\hbar\omega_1}{2}. \quad (11.101)$$

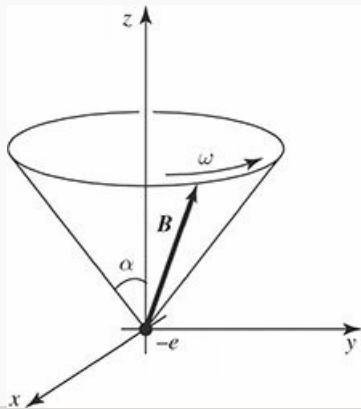


Figure 11.16: The magnetic field sweeps around in a cone, at angular velocity ω (Equation 11.96).

Suppose the electron starts out with spin up, along $\mathbf{B}(0)$:

$$\chi(0) = \begin{pmatrix} \cos(\alpha/2) \\ \sin(\alpha/2) \end{pmatrix}. \quad (11.102)$$

The exact solution to the time-dependent Schrödinger equation is (Problem 11.20):

$$\chi(t) = \begin{pmatrix} \left[\cos(\lambda t/2) - i \frac{(\omega_1 - \omega)}{\lambda} \sin(\lambda t/2) \right] \cos(\alpha/2) e^{-i\omega t/2} \\ \left[\cos(\lambda t/2) - i \frac{(\omega_1 + \omega)}{\lambda} \sin(\lambda t/2) \right] \sin(\alpha/2) e^{+i\omega t/2} \end{pmatrix}, \quad (11.103)$$

where

$$\lambda \equiv \sqrt{\omega^2 + \omega_1^2 - 2\omega\omega_1 \cos \alpha}, \quad (11.104)$$

or, expressing it as a linear combination of χ_+ and χ_- :

$$\begin{aligned} \chi(t) = & \left[\cos\left(\frac{\lambda t}{2}\right) - i \frac{(\omega_1 - \omega \cos \alpha)}{\lambda} \sin\left(\frac{\lambda t}{2}\right) \right] e^{-i\omega t/2} \chi_+(t) \\ & + i \left[\frac{\omega}{\lambda} \sin \alpha \sin\left(\frac{\lambda t}{2}\right) \right] e^{+i\omega t/2} \chi_-(t). \end{aligned} \quad (11.105)$$

Evidently the (exact) probability of a transition to spin down (along the current direction of \mathbf{B}) is

$$|\langle \chi(t) | \chi_-(t) \rangle|^2 = \left[\frac{\omega}{\lambda} \sin \alpha \sin \left(\frac{\lambda t}{2} \right) \right]^2. \quad (11.106)$$

The adiabatic theorem says that this transition probability should vanish in the limit $T_e \gg T_i$, where T_e is the characteristic time for changes in the Hamiltonian (in this case, $1/\omega$) and T_i is the characteristic time for changes in the wave function (in this case, $\hbar/(E_+ - E_-) = 1/\omega_1$). Thus the adiabatic approximation means $\omega \ll \omega_1$: the field rotates slowly, in comparison with the phase of the (unperturbed) wave functions. In the adiabatic regime $\lambda \approx \omega_1$ (Equation 11.104), and therefore

$$|\langle \chi(t) | \chi_-(t) \rangle|^2 \approx \left[\frac{\omega}{\omega_1} \sin \alpha \sin \left(\frac{\lambda t}{2} \right) \right]^2 \rightarrow 0, \quad (11.107)$$

as advertised. The magnetic field leads the electron around by its nose, with the spin always pointing in the direction of \mathbf{B} . By contrast, if $\omega \gg \omega_1$ then $\lambda \approx \omega$, and the system bounces back and forth between spin up and spin down (Figure 11.17).

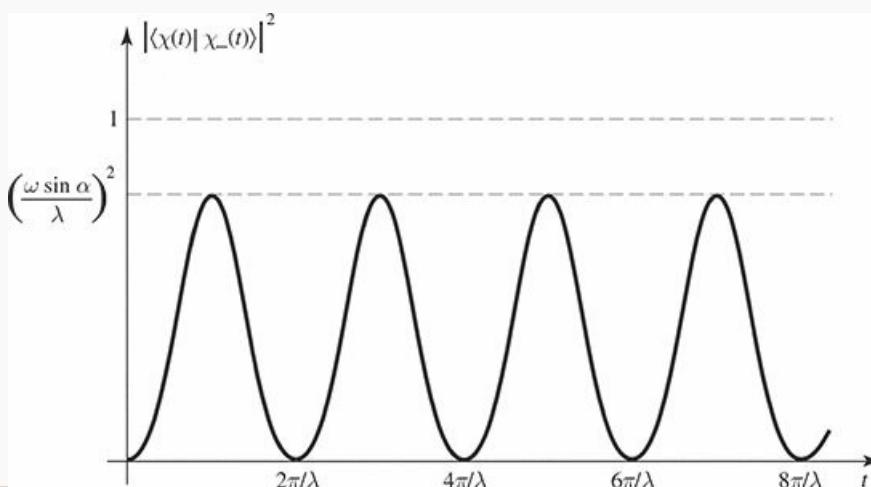


Figure 11.17: Plot of the transition probability, Equation 11.106, in the *non-adiabatic regime* ($\omega \gg \omega_1$).

- * **Problem 11.18** A particle of mass m is in the ground state of the infinite square well (Equation 2.22). 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.
 - What is the most probable result? What is the probability of getting that result?
 - What is the *next* most probable result, and what is its probability? Suppose your measurement returned this value; what would you conclude about conservation of energy?
 - What is the *expectation value* of the energy? Hint: If you find yourself confronted with an infinite series, try another method.

Problem 11.19 A particle is in the ground state of the harmonic oscillator with classical frequency ω , when suddenly the spring constant quadruples, so $\omega' = 2\omega$ without initially changing the wave function (of course, Ψ will now *evolve* differently, because the Hamiltonian has changed). What is the probability that a measurement of the energy would still return the value $\hbar\omega/2$? What is the probability of getting $\hbar\omega$? *Answer:* 0.943.

**

Problem 11.20 Check that Equation 11.103 satisfies the time-dependent Schrödinger equation for the Hamiltonian in Equation 11.97. Also confirm Equation 11.105, and show that the sum of the squares of the coefficients is 1, as required for normalization.

*

Problem 11.21 Find Berry's phase for one cycle of the process in Example 11.4. *Hint:* Use Equation 11.105 to determine the *total* phase change, and subtract off the dynamical part. You'll need to expand λ (Equation 11.104) to first order in ω/ω_1 .

Problem 11.22 The delta function well (Equation 2.117) supports a single bound state (Equation 2.132). Calculate the geometric phase change when a gradually increases from a_1 to a_2 . If the increase occurs at a constant rate ($da/dt = c$), what is the dynamic phase change for this process?³⁷

Further Problems on Chapter 11

*** Problem 11.23 In Problem 11.1 you showed that the solution to

$$\frac{df}{dt} = k(t)f(t)$$

(where $k(t)$ is a function of t) is

$$f(t) = e^{K(t)} f(0), \quad \text{where } K(t) \equiv \int_0^t k(t') dt'.$$

This suggests that the solution to the Schrödinger equation (11.1) might be

$$\Psi(t) = e^{\hat{G}(t)} \Psi(0), \quad \text{where } \hat{G}(t) \equiv -\frac{i}{\hbar} \int_0^t \hat{H}(t') dt'. \quad (11.108)$$

It doesn't work, because $\hat{H}(t)$ is an *operator*, not a function, and $\hat{H}(t_1)$ does not (in general) commute with $\hat{H}(t_2)$.

- (a) Try calculating $i\hbar \partial \Psi / \partial t$, using Equation 11.108. Note: as always, the exponentiated operator is to be interpreted as a power series:

$$e^{\hat{G}} \equiv 1 + \hat{G} + \frac{1}{2!} \hat{G}\hat{G} + \frac{1}{3!} \hat{G}\hat{G}\hat{G} + \dots$$

Show that if $[\hat{G}, \hat{H}] = 0$, then Ψ satisfies the Schrödinger equation.

- (b) Check that the correct solution in the general case ($[\hat{G}, \hat{H}] \neq 0$) is

$$\begin{aligned} \Psi(t) = & \left\{ 1 + \left(-\frac{i}{\hbar} \right) \int_0^t \hat{H}(t_1) dt_1 + \left(-\frac{i}{\hbar} \right)^2 \int_0^t \hat{H}(t_1) \left[\int_0^{t_1} \hat{H}(t_2) dt_2 \right] dt_1 \right. \\ & \left. + \left(-\frac{i}{\hbar} \right)^3 \int_0^t \hat{H}(t_1) \left[\int_0^{t_1} \hat{H}(t_2) \left(\int_0^{t_2} \hat{H}(t_3) dt_3 \right) dt_2 \right] dt_1 + \dots \right\} \Psi(0). \end{aligned} \quad (11.109)$$

UGLY! Notice that the operators in each term are “time-ordered,” in the sense that the *latest* \hat{H} appears at the far left, followed by the next latest, and so on ($t \geq t_1 \geq t_2 \geq t_3 \dots$). Dyson introduced the time-ordered product of two operators:

$$\mathbf{T} [\hat{H}(t_i) \hat{H}(t_j)] \equiv \begin{cases} \hat{H}(t_i) \hat{H}(t_j), & t_i \geq t_j \\ \hat{H}(t_j) \hat{H}(t_i), & t_j \geq t_i \end{cases} \quad (11.110)$$

or, more generally,

$$\mathbf{T} [\hat{H}(t_1) \hat{H}(t_2) \dots \hat{H}(t_n)] \equiv \hat{H}(t_{j_1}) \hat{H}(t_{j_2}) \dots \hat{H}(t_{j_n}), \quad (11.111)$$

where $t_{j_1} \geq t_{j_2} \geq \dots \geq t_{j_n}$.

(c) Show that

$$\mathbf{T}\left[\hat{G}\hat{G}\right] = -\frac{2}{\hbar^2} \int_0^t \hat{H}(t_1) \left[\int_0^{t_1} \hat{H}(t_2) dt_2 \right] dt_1,$$

and generalize to higher powers of \hat{G} . In place of \hat{G}^n , in equation 11.108, we really want $\mathbf{T}\left[\hat{G}^n\right]$:

$$\Psi(t) = \mathbf{T}\left[e^{-\frac{i}{\hbar} \int_0^t \hat{H}(t') dt'}\right] \Psi(0). \quad (11.112)$$

This is **Dyson's formula**; it's a compact way of writing Equation 11.109, the formal solution to Schrödinger's equation. Dyson's formula plays a fundamental role in quantum field theory.³⁸

** **Problem 11.24** In this problem we develop time-dependent perturbation theory for a multi-level system, starting with the generalization of Equations 11.5 and 11.6:

$$\hat{H}_0 \psi_n = E_n \psi_n, \quad \langle \psi_n | \psi_m \rangle = \delta_{nm}. \quad (11.113)$$

At time $t = 0$ we turn on a perturbation $H'(t)$, so that the total Hamiltonian is

$$\hat{H} = \hat{H}_0 + \hat{H}'(t). \quad (11.114)$$

(a) Generalize Equation 11.10 to read

$$\Psi(t) = \sum_n c_n(t) \psi_n e^{-i E_n t / \hbar}, \quad (11.115)$$

and show that

$$\dot{c}_m = -\frac{i}{\hbar} \sum_n c_n H'_{mn} e^{i(E_m - E_n)t/\hbar}, \quad (11.116)$$

where

$$H'_{mn} \equiv \left\langle \psi_m \left| \hat{H}' \right| \psi_n \right\rangle. \quad (11.117)$$

(b) If the system starts out in the state ψ_N , show that (in first-order perturbation theory)

$$c_N(t) \approx 1 - \frac{i}{\hbar} \int_0^t H'_{NN}(t') dt'. \quad (11.118)$$

and

$$c_m(t) \approx -\frac{i}{\hbar} \int_0^t H'_{mN}(t') e^{i(E_m - E_N)t'/\hbar} dt', \quad (m \neq N). \quad (11.119)$$

- (c) For example, suppose \hat{H}' is *constant* (except that it was turned on at $t = 0$, and switched off again at some later time T). Find the probability of transition from state N to state M ($M \neq N$), as a function of T . *Answer:*

$$4 |H'_{MN}|^2 \frac{\sin^2 [(E_N - E_M) T / 2\hbar]}{(E_N - E_M)^2}. \quad (11.120)$$

- (d) Now suppose \hat{H}' is a sinusoidal function of time: $\hat{H}' = V \cos(\omega t)$. Making the usual assumptions, show that transitions occur only to states with energy $E_M = E_N \pm \hbar\omega$, and the transition probability is

$$P_{N \rightarrow M} = |V_{MN}|^2 \frac{\sin^2 [(E_N - E_M \pm \hbar\omega) T / 2\hbar]}{(E_N - E_M \pm \hbar\omega)^2}. \quad (11.121)$$

- (e) Suppose a multi-level system is immersed in incoherent electromagnetic radiation. Using Section 11.2.3 as a guide, show that the transition rate for stimulated emission is given by the same formula (Equation 11.54) as for a two-level system.

Problem 11.25 For the examples in Problem 11.24 (c) and (d), calculate $c_m(t)$, to first order. Check the normalization condition:

$$\sum_m |c_m(t)|^2 = 1, \quad (11.122)$$

and comment on any discrepancy. Suppose you wanted to calculate the probability of *remaining* in the original state ψ_N ; would you do better to use $|c_N(t)|^2$, or $1 - \sum_{m \neq N} |c_m(t)|^2$?

Problem 11.26 A particle starts out (at time $t = 0$) in the N th state of the infinite square well. Now the “floor” of the well rises temporarily (maybe water leaks in, and then drains out again), so that the potential inside is uniform but time dependent: $V_0(t)$, with $V_0(0) = V_0(T) = 0$.

- (a) Solve for the *exact* $c_m(t)$, using Equation 11.116, and show that the wave function changes *phase*, but no transitions occur. Find the phase change, $\phi(T)$, in terms of the function $V_0(t)$.
- (b) Analyze the same problem in first-order perturbation theory, and compare your answers.

Comment: The same result holds whenever the perturbation simply adds a constant (constant in x , that is, not in t) to the potential; it has nothing to do with the infinite square well, as such. Compare Problem 1.8.

- * **Problem 11.27** A particle of mass m is initially in the ground state of the (one-dimensional) infinite square well. At time $t = 0$ a “brick” is dropped into the well, so that the potential becomes

$$V(x) = \begin{cases} V_0, & 0 \leq x \leq a/2, \\ 0, & a/2 < x \leq a, \\ \infty, & \text{otherwise,} \end{cases}$$

where $V_0 \ll E_1$. After a time T , the brick is removed, and the energy of the particle is measured. Find the probability (in first-order perturbation theory) that the result is now E_2 .

Problem 11.28 We have encountered stimulated emission, (stimulated) absorption, and spontaneous emission. How come there is no such thing as spontaneous *absorption*?

Problem 11.29 Magnetic resonance. A spin-1/2 particle with gyromagnetic ratio γ , at rest in a static magnetic field $B_0\hat{k}$, precesses at the Larmor frequency $\omega_0 = \gamma B_0$ (Example 4.3). Now we turn on a small transverse radiofrequency (rf) field, $B_{\text{rf}} [\cos(\omega t)\hat{i} - \sin(\omega t)\hat{j}]$, so that the total field is

$$\mathbf{B} = B_{\text{rf}} \cos(\omega t)\hat{i} - B_{\text{rf}} \sin(\omega t)\hat{j} + B_0\hat{k}. \quad (11.123)$$

(a) Construct the 2×2 Hamiltonian matrix (Equation 4.158) for this system.

(b) If $\chi(t) = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}$ is the spin state at time t , show that

$$\dot{a} = \frac{i}{2} (\Omega e^{i\omega t} b + \omega_0 a); \quad \dot{b} = \frac{i}{2} (\Omega e^{-i\omega t} a - \omega_0 b), \quad (11.124)$$

where $\Omega \equiv \gamma B_{\text{rf}}$ is related to the strength of the rf field.

(c) Check that the general solution for $a(t)$ and $b(t)$, in terms of their initial values a_0 and b_0 , is

$$\begin{aligned} a(t) &= \left\{ a_0 \cos(\omega' t/2) + \frac{i}{\omega'} [a_0(\omega_0 - \omega) + b_0 \Omega] \sin(\omega' t/2) \right\} e^{i\omega t/2} \\ b(t) &= \left\{ b_0 \cos(\omega' t/2) + \frac{i}{\omega'} [b_0(\omega - \omega_0) + a_0 \Omega] \sin(\omega' t/2) \right\} e^{-i\omega t/2} \end{aligned}$$

where

$$\omega' \equiv \sqrt{(\omega - \omega_0)^2 + \Omega^2}. \quad (11.125)$$

(d) If the particle starts out with spin up (i.e. $a_0 = 1, b_0 = 0$), find the probability of a transition to spin down, as a function of time. *Answer:* $P(t) = \{\Omega^2 / [(\omega - \omega_0)^2 + \Omega^2]\} \sin^2(\omega' t/2)$.

(e) Sketch the resonance curve,

$$P(\omega) = \frac{\Omega^2}{(\omega - \omega_0)^2 + \Omega^2}, \quad (11.126)$$

as a function of the driving frequency ω (for fixed ω_0 and Ω). Note that the maximum occurs at $\omega = \omega_0$. Find the “full width at half maximum,” $\Delta\omega$.

- (f) Since $\omega_0 = \gamma B_0$, we can use the experimentally observed resonance to determine the magnetic dipole moment of the particle. In a **nuclear magnetic resonance** (nmr) experiment the g -factor of the proton is to be measured, using a static field of 10,000 gauss and an rf field of amplitude 0.01 gauss. What will the resonant frequency be? (See Section 7.5 for the magnetic moment of the proton.) Find the width of the resonance curve. (Give your answers in Hz.)

- ** Problem 11.30 In this problem we will recover the results Section 11.2.1 directly from the Hamiltonian for a charged particle in an electromagnetic field (Equation 4.188). An electromagnetic wave can be described by the potentials

$$\mathbf{A} = \frac{\mathbf{E}_0}{\omega} \sin(\mathbf{k} \cdot \mathbf{r} - \omega t), \quad \varphi = 0$$

where in order to satisfy Maxwell's equations, the wave must be transverse ($\mathbf{E}_0 \cdot \mathbf{k} = 0$) and of course travel at the speed of light ($\omega = c |\mathbf{k}|$).

- (a) Find the electric and magnetic fields for this plane wave.
 (b) The Hamiltonian may be written as $H^0 + H'$ where H^0 is the Hamiltonian in the absence of the electromagnetic wave and H' is the perturbation. Show that the perturbation is given by

$$\hat{H}'(t) = \frac{e}{2im\omega} e^{i\mathbf{k} \cdot \mathbf{r}} \mathbf{E}_0 \cdot \hat{\mathbf{p}} e^{-i\omega t} - \frac{e}{2im\omega} e^{-i\mathbf{k} \cdot \mathbf{r}} \mathbf{E}_0 \cdot \hat{\mathbf{p}} e^{i\omega t}, \quad (11.127)$$

plus a term proportional to E_0^2 that we will ignore. *Note:* the first term corresponds to absorption and the second to emission.

- (c) In the dipole approximation we set $e^{i\mathbf{k} \cdot \mathbf{r}} \approx 1$. With the electromagnetic wave polarized along the z direction, show that the matrix element for absorption is then

$$V_{ba} = -\frac{\omega_0}{\omega} \varphi E_0.$$

Compare Equation 11.41. They're not *exactly* the same; would the difference effect our calculations in Section 11.2.3 or 11.3? Why or why not? *Hint:* To turn the matrix element of \mathbf{p} into a matrix element of \mathbf{r} , you need to prove the following identity: $im [\hat{H}^0, \hat{\mathbf{r}}] = \hbar \hat{\mathbf{p}}$.

- *** Problem 11.31 In Equation 11.38 I assumed that the atom is so small (in comparison to the wavelength of the light) that spatial variations in the field can be ignored. The *true* electric field would be

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \cos(\mathbf{k} \cdot \mathbf{r} - \omega t). \quad (11.128)$$

If the atom is centered at the origin, then $\mathbf{k} \cdot \mathbf{r} \ll 1$ over the relevant volume ($|\mathbf{k}| = 2\pi/\lambda$, so $\mathbf{k} \cdot \mathbf{r} \sim r/\lambda \ll 1$), and that's why we could afford to drop this term. Suppose we keep the first-order correction:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 [\cos(\omega t) + (\mathbf{k} \cdot \mathbf{r}) \sin(\omega t)]. \quad (11.129)$$

The first term gives rise to the allowed (electric dipole) transitions we considered in the text; the second leads to so-called forbidden (magnetic dipole and electric quadrupole) transitions (higher powers of $\mathbf{k} \cdot \mathbf{r}$ lead to even more “forbidden” transitions, associated with higher multipole moments).³⁹

- (a) Obtain the spontaneous emission rate for forbidden transitions (don't bother to average over polarization and propagation directions, though this should really be done to complete the calculation). *Answer:*

$$R_{b \rightarrow a} = \frac{q^2 \omega^5}{\pi \epsilon_0 \hbar c^5} \left| \left\langle a \left| (\hat{n} \cdot \mathbf{r}) (\hat{k} \cdot \mathbf{r}) \right| b \right\rangle \right|^2. \quad (11.130)$$

- (b) Show that for a one-dimensional oscillator the forbidden transitions go from level n to level $n - 2$, and the transition rate (suitably averaged over \hat{n} and \hat{k}) is

$$R = \frac{\hbar q^2 \omega^3 n (n - 1)}{15 \pi \epsilon_0 m^2 c^5}. \quad (11.131)$$

(Note: Here ω is the frequency of the *photon*, not the oscillator.) Find the ratio of the “forbidden” rate to the “allowed” rate, and comment on the terminology.

- (c) Show that the $2S \rightarrow 1S$ transition in hydrogen is not possible even by a “forbidden” transition. (As it turns out, this is true for all the higher multipoles as well; the dominant decay is in fact by two-photon emission, and the lifetime is about a tenth of a second.⁴⁰)

*** **Problem 11.32** Show that the spontaneous emission rate (Equation 11.63) for a transition from n, ℓ to n', ℓ' in hydrogen is

$$\frac{e^2 \omega^3 I^2}{3 \pi \epsilon_0 \hbar c^3} \times \begin{cases} \frac{\ell+1}{2\ell+1}, & \ell' = \ell + 1, \\ \frac{\ell}{2\ell+1}, & \ell' = \ell - 1, \end{cases} \quad (11.132)$$

where

$$I \equiv \int_0^\infty r^3 R_{n\ell}(r) R_{n'\ell'}(r) dr. \quad (11.133)$$

(The atom starts out with a specific value of m , and it goes to *any* of the states m' consistent with the selection rules: $m' = m + 1, m$, or $m - 1$. Notice that the answer is independent of m .) *Hint:* First calculate all the nonzero matrix

elements of x , y , and z between $|n\ell m\rangle$ and $|n'\ell' m'\rangle$ for the case $\ell' = \ell + 1$. From these, determine the quantity

$$|\langle n', \ell + 1, m + 1 | \mathbf{r} | n\ell m \rangle|^2 + |\langle n', \ell + 1, m | \mathbf{r} | n\ell m \rangle|^2 + |\langle n', \ell + 1, m - 1 | \mathbf{r} | n\ell m \rangle|^2.$$

Then do the same for $\ell' = \ell - 1$. You may find useful the following recursion formulas (which hold for $m \geq 0$):⁴¹

$$(2\ell + 1) x P_\ell^m(x) = (\ell + m) P_{\ell-1}^m(x) + (\ell - m + 1) P_{\ell+1}^m(x), \quad (11.134)$$

$$(2\ell + 1) \sqrt{1 - x^2} P_\ell^m(x) = P_{\ell+1}^{m+1}(x) - P_{\ell-1}^{m+1}(x), \quad (11.135)$$

and the orthogonality relation Equation 4.33.

Problem 11.33 The spontaneous emission rate for the 21-cm hyperfine line in hydrogen (Section 7.5) can be obtained from Equation 11.63, except that this is a *magnetic* dipole transition, not an *electric* one:⁴²

$$\boldsymbol{\mu} \rightarrow \frac{1}{c} \mathbf{M} = \frac{1}{c} \langle 1 | (\boldsymbol{\mu}_e + \boldsymbol{\mu}_p) | 0 \rangle,$$

where

$$\boldsymbol{\mu}_e = -\frac{e}{m_e} \mathbf{S}_e, \quad \boldsymbol{\mu}_p = \frac{5.59 e}{2m_p} \mathbf{S}_p$$

are the magnetic moments of the electron and proton (Equation 7.89), and $|0\rangle$, $|1\rangle$ are the singlet and triplet configurations (Equations 4.175 and 4.176). Because $m_p \gg m_e$, the proton contribution is negligible, so

$$A = \frac{\omega_0^3 e^2}{3\pi \epsilon_0 \hbar c^5 m_e^2} |\langle 1 | \mathbf{S}_e | 0 \rangle|^2.$$

Work out $|\langle 1 | \mathbf{S}_e | 0 \rangle|^2$ (use whichever triplet state you like). Put in the actual numbers, to determine the transition rate and the lifetime of the triplet state. *Answer:* 1.1×10^7 years.



Problem 11.34 A particle starts out in the ground state of the infinite square well (on the interval $0 \leq x \leq a$). Now a wall is slowly erected, slightly off-center:⁴³

$$V(x) = f(t) \delta\left(x - \frac{a}{2} - \epsilon\right),$$

where $f(t)$ rises gradually from 0 to ∞ . According to the adiabatic theorem, the particle will remain in the ground state of the evolving Hamiltonian.

- (a) Find (and sketch) the ground state at $t \rightarrow \infty$. *Hint:* This should be the ground state of the infinite square well with an impenetrable barrier at $a/2 + \epsilon$. Note that the particle is confined to the (slightly) larger left “half” of the well.

- (b) Find the (transcendental) equation for the ground state energy at time t .

Answer:

$$z \sin z = T [\cos z - \cos(z\delta)],$$

where $z \equiv ka$, $T \equiv maf(t)/\hbar^2$, $\delta \equiv 2\epsilon/a$, and $k \equiv \sqrt{2mE}/\hbar$.

- (c) Setting $\delta = 0$, solve graphically for z , and show that the smallest z goes from π to 2π as T goes from 0 to ∞ . Explain this result.
- (d) Now set $\delta = 0.01$ and solve numerically for z , using $T = 0, 1, 5, 20, 100$, and 1000.
- (e) Find the probability P_r that the particle is in the right “half” of the well, as a function of z and δ . *Answer:* $P_r = 1/[1 + (I_+/I_-)]$, where $I_{\pm} \equiv [1 \pm \delta - (1/z) \sin(z(1 \pm \delta))] \sin^2[z(1 \mp \delta)/2]$. Evaluate this expression numerically for the T 's and δ in part (d). Comment on your results.
- (e) Plot the ground state wave function for those same values of T and δ . Note how it gets squeezed into the left half of the well, as the barrier grows.⁴⁴

Problem 11.35 The case of an infinite square well whose right wall expands at a *constant velocity* (v) can be solved *exactly*.⁴⁵ A complete set of solutions is

$$\Phi_n(x, t) \equiv \sqrt{\frac{2}{w}} \sin\left(\frac{n\pi}{w}x\right) e^{i(mvx^2 - 2E_n^i at)/2\hbar w}, \quad (11.136)$$

where $w(t) \equiv a + vt$ is the width of the well and $E_n^i \equiv n^2\pi^2\hbar^2/2ma^2$ is the n th allowed energy of the *original* well (width a). The general solution is a linear combination of the Φ 's:

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \Phi_n(x, t), \quad (11.137)$$

whose coefficients c_n are *independent of t*.

- (a) Check that Equation 11.136 satisfies the time-dependent Schrödinger equation, with the appropriate boundary conditions.
- (b) Suppose a particle starts out ($t = 0$) in the ground state of the initial well:

$$\Psi(x, 0) = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi}{a}x\right).$$

Show that the expansion coefficients can be written in the form

$$c_n = \frac{2}{\pi} \int_0^{\pi} e^{-i\alpha z^2} \sin(nz) \sin(z) dz, \quad (11.138)$$

where $\alpha \equiv mva/2\pi^2\hbar$ is a dimensionless measure of the speed with which the well expands. (Unfortunately, this integral cannot be evaluated

in terms of elementary functions.)

- (c) Suppose we allow the well to expand to twice its original width, so the “external” time is given by $w(T_e) = 2a$. The “internal” time is the *period* of the time-dependent exponential factor in the (initial) ground state. Determine T_e and T_i , and show that the adiabatic regime corresponds to $\alpha \ll 1$, so that $\exp(-i\alpha z^2) \approx 1$ over the domain of integration. Use this to determine the expansion coefficients, c_n . Construct $\Psi(x, t)$, and confirm that it is consistent with the adiabatic theorem.

- (d) Show that the phase factor in $\Psi(x, t)$ can be written in the form

$$\theta(t) = -\frac{1}{\hbar} \int_0^t E_1(t') dt', \quad (11.139)$$

where $E_n(t) \equiv n^2 \pi^2 \hbar^2 / 2m w^2$ is the n th *instantaneous* eigenvalue, at time t . Comment on this result. What is the geometric phase? If the well now contracts back to its original size, what is Berry’s phase for the cycle?

Problem 11.36 The driven harmonic oscillator. Suppose the one-dimensional harmonic oscillator (mass m , frequency ω) is subjected to a driving force of the form $F(t) = m\omega^2 f(t)$, where $f(t)$ is some specified function. (I have factored out $m\omega^2$ for notational convenience; $f(t)$ has the dimensions of length.) The Hamiltonian is

$$H(t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} m\omega^2 x^2 - m\omega^2 x f(t). \quad (11.140)$$

Assume that the force was first turned on at time $t = 0$: $f(t) = 0$ for $t \leq 0$. This system can be solved exactly, both in classical mechanics and in quantum mechanics.⁴⁶

- (a) Determine the *classical* position of the oscillator, assuming it started from rest at the origin ($x_c(0) = \dot{x}_c(0) = 0$). *Answer:*

$$x_c(t) = \omega \int_0^t f(t') \sin [\omega(t-t')] dt'. \quad (11.141)$$

- (b) Show that the solution to the (time-dependent) Schrödinger equation for this oscillator, assuming it started out in the n th state of the *undriven* oscillator ($\Psi(x, 0) = \psi_n(x)$ where $\psi_n(x)$ is given by Equation 2.62), can be written as

$$\Psi(x, t) = \psi_n(x - x_c) e^{\frac{i}{\hbar} \left[-\left(n+\frac{1}{2}\right)\hbar\omega t + m\dot{x}_c(x - \frac{x_c}{2}) + \frac{m\omega^2}{2} \int_0^t f(t') x_c(t') dt' \right]}. \quad (11.142)$$

- (c) Show that the eigenfunctions and eigenvalues of $H(t)$ are

$$\psi_n(x, t) = \psi_n(x - f); \quad E_n(t) = \left(n + \frac{1}{2}\right)\hbar\omega - \frac{1}{2}m\omega^2 f^2. \quad (11.143)$$

- (d) Show that in the adiabatic approximation the classical position (Equation 11.141) reduces to $x_c(t) \approx f(t)$. State the precise criterion for adiabaticity, in this context, as a constraint on the time derivative of f . *Hint:* Write $\sin[\omega(t-t')]$ as $(1/\omega)(d/dt')\cos[\omega(t-t')]$ and use integration by parts.
- (e) Confirm the adiabatic theorem for this example, by using the results in (c) and (d) to show that

$$\Psi(x, t) \approx \psi_n(x, t) e^{i\theta_n(t)} e^{i\gamma_n(t)}. \quad (11.144)$$

Check that the dynamic phase has the correct form (Equation 11.92). Is the geometric phase what you would expect?

Problem 11.37 Quantum Zeno Paradox.⁴⁷ Suppose a system starts out in an excited state ψ_b , which has a natural lifetime τ for transition to the ground state ψ_a . Ordinarily, for times substantially less than τ , the probability of a transition is proportional to t (Equation 11.49):

$$P_{b \rightarrow a} = \frac{t}{\tau}. \quad (11.145)$$

If we make a measurement after a time t , then, the probability that the system is still in the *upper* state is

$$P_b(t) = 1 - \frac{t}{\tau}. \quad (11.146)$$

Suppose we *do* find it to be in the upper state. In that case the wave function collapses back to ψ_b , and the process starts all over again. If we make a *second* measurement, at $2t$, the probability that the system is *still* in the upper state is

$$\left(1 - \frac{t}{\tau}\right)^2 \approx 1 - \frac{2t}{\tau}, \quad (11.147)$$

which is the same as it would have been had we never made the first measurement at t (as one would naively expect).

However, for *extremely* short times, the probability of a transition is *not* proportional to t , but rather to t^2 (Equation 11.46):⁴⁸

$$P_{b \rightarrow a} = \alpha t^2. \quad (11.148)$$

- (a) In this case what is the probability that the system is still in the upper state after the two measurements? What *would* it have been (after the same elapsed time) if we had never made the first measurement?
- (b) Suppose we examine the system at n regular (extremely short) intervals, from $t = 0$ out to $t = T$ (that is, we make measurements at $T/n, 2T/n, 3T/n, \dots, T$). What is the probability that the system is still in the upper state at time T ? What is its limit as $n \rightarrow \infty$? *Moral of the story:* Because

of the collapse of the wave function at every measurement, a *continuously* observed system never decays at all!⁴⁹



Problem 11.38 The numerical solution to the time-independent Schrödinger equation in Problem 2.61 can be extended to solve the time-dependent Schrödinger equation. When we discretize the variable x , we obtain the matrix equation

$$\mathbf{H}\Psi = i\hbar \frac{d}{dt}\Psi. \quad (11.149)$$

The solution to this equation can be written

$$\Psi(t + \Delta t) = \mathbf{U}(\Delta t)\Psi(t). \quad (11.150)$$

If \mathbf{H} is time independent, the exact expression for the time-evolution operator is⁵⁰

$$\mathbf{U}(\Delta t) = e^{-i\mathbf{H}\Delta t/\hbar} \quad (11.151)$$

and for Δt small enough, the time-evolution operator can be approximated as

$$\mathbf{U}(\Delta t) \approx 1 - i\mathbf{H}\frac{\Delta t}{\hbar}. \quad (11.152)$$

While Equation 11.152 is the most obvious way to approximate \mathbf{U} , a numerical scheme based on it is unstable, and it is preferable to use Cayley's form for the approximation.⁵¹

$$\mathbf{U}(\Delta t) \approx \frac{1 - \frac{1}{2}i\frac{\Delta t}{\hbar}\mathbf{H}}{1 + \frac{1}{2}i\frac{\Delta t}{\hbar}\mathbf{H}}. \quad (11.153)$$

Combining Equations 11.153 and 11.150 we have

$$\left(1 + \frac{1}{2}i\frac{\Delta t}{\hbar}\mathbf{H}\right)\Psi(t + \Delta t) = \left(1 - \frac{1}{2}i\frac{\Delta t}{\hbar}\mathbf{H}\right)\Psi(t). \quad (11.154)$$

This has the form of a matrix equation $\mathbf{M}\mathbf{x} = \mathbf{b}$ which can be solved for the unknown $\mathbf{x} = \Psi(t + \Delta t)$. Because the matrix $\mathbf{M} = 1 + \frac{1}{2}i\frac{\Delta t}{\hbar}\mathbf{H}$ is tri-diagonal,⁵² efficient algorithms exist for doing so.⁵³

- (a) Show that the approximation in Equation 11.153 is accurate to second order. That is, show that Equations 11.151 and 11.153, expanded as power series in Δt , agree up through terms of order $(\Delta t)^2$. Verify that the matrix in Equation 11.153 is unitary.

As an example, consider a particle of mass m moving in one dimension in a simple harmonic oscillator potential. For the numerical part set $m = 1$, $\omega = 1$, and $\hbar = 1$ (this just defines the units of mass, time, and length).

- (b) Construct the Hamiltonian matrix \mathbf{H} for $N + 1 = 100$ spatial grid points. Set the spatial boundaries where the dimensionless length is $\xi = \pm 10$ (far enough out that we can assume that the wave function vanishes there for low-energy states). By computer, find the lowest two eigenvalues of \mathbf{H} , and compare the exact values. Plot the corresponding eigenfunctions. Are they normalized? If not, normalize them before doing part (c).
- (c) Take $\Psi(0) = (\psi_0 + \psi_1)/\sqrt{2}$ (from part (b)) and use Equation 11.154 to evolve the wave function from time $t = 0$ to $t = 4\pi/\omega$. Create a movie (**Animate**, in Mathematica) showing $\text{Re}(\Psi(t))$, $\text{Im}(\Psi(t))$, and $|\Psi(t)|$, together with the exact result. *Hint:* You need to decide what to use for Δt . In terms of the number of time steps N_t , $N_t \Delta t = 4\pi/\omega$. In order for the approximation of the exponential to hold, we need to have $E \Delta t / \hbar \ll 1$. The energy of our state is of order $\hbar\omega$, and therefore $N_t \gg 4\pi$. So you will need at least (say) 100 time steps.



Problem 11.39 We can use the technique of Problem 11.38 to investigate time evolution when the Hamiltonian *does* depend on time, as long as we choose Δt small enough. Evaluating \mathbf{H} at the midpoint of each time step we simply replace Equation 11.154 with⁵⁴

$$\left[1 + \frac{1}{2}i \frac{\Delta t}{\hbar} \mathbf{H}\left(t + \frac{\Delta t}{2}\right)\right] \Psi(t + \Delta t) = \left[1 - \frac{1}{2}i \frac{\Delta t}{\hbar} \mathbf{H}\left(t + \frac{\Delta t}{2}\right)\right] \Psi(t). \quad (11.155)$$

Consider the driven harmonic oscillator of Problem 11.36 with

$$f(t) = A \sin(\Omega t), \quad (11.156)$$

where A is a constant with the units of length and Ω is the driving frequency. In the following we will set $m = \omega = \hbar = A = 1$ and look at the effect of varying Ω . Use the same parameters for the spatial discretization as in Problem 11.38, but set $N_t = 1000$. For a particle that starts off in the ground state at $t = 0$, create a movie showing the numerical and exact solutions as well as the instantaneous ground state from $t = 0$ to $t = 2\pi/\Omega$ for

- (a) $\Omega = \omega/5$. In line with the adiabatic theorem, you should see that the numerical solution is close (up to a phase) to the instantaneous ground state.
- (b) $\Omega = 5\omega$. In line with what you've learned about sudden perturbations, you should see that the numerical solution is barely affected by the driving force.
- (c) $\Omega = 6\omega/5$.

¹ Notice that c_a is modified in every *even* order, and c_b in every *odd* order; this would not be true if the perturbation included diagonal terms, or if the system started out in a linear combination of the two states.

² Perturbation theory for multi-level systems is treated in Problem 11.24.

- ³ For very small t , $P_{a \rightarrow b}(t)$ is independent of ω ; it takes a couple of cycles for the system to “realize” that the perturbation is periodic.
- ⁴ For visible light $\lambda \sim 5000 \text{ \AA}$, while the diameter of an atom is around 1 \AA , so this approximation is reasonable; but it would *not* be for x-rays. Problem 11.31 explores the effect of spatial variation of the field.
- ⁵ The energy of a charge q in a static field \mathbf{E} is $-q \int \mathbf{E} \cdot d\mathbf{r}$. You may well object to the use of an electrostatic formula for a manifestly time-dependent field. I am implicitly assuming that the period of oscillation is long compared to the time it takes the charge to move around (within the atom).
- ⁶ As usual, we assume the nucleus is heavy and stationary; it is the wave function of the *electron* that concerns us.
- ⁷ The letter ρ is supposed to remind you of **electric dipole moment** (for which, in electrodynamics, the letter p is customarily used—in this context it is rendered as a squiggly ρ to avoid confusion with momentum). Actually, ρ is the off-diagonal matrix element of the z component of the dipole moment operator, $q\mathbf{r}$. Because of its association with electric dipole moments, radiation governed by Equation 11.40 is called **electric dipole radiation**; it is overwhelmingly the dominant kind, at least in the visible region. See Problem 11.31 for generalizations and terminology.
- ⁸ For an accessible treatment see Rodney Loudon, *The Quantum Theory of Light*, 2nd edn (Clarendon Press, Oxford, 1983).
- ⁹ Einstein’s paper was published in 1917, well before the Schrödinger equation. Quantum electrodynamics comes into the argument via the Planck blackbody formula, which dates from 1900.
- ¹⁰ For an alternative derivation using “seat-of-the-pants” quantum electrodynamics, see Problem 11.11.
- ¹¹ David J. Griffiths, *Introduction to Electrodynamics*, 4th edn, (Cambridge University Press, Cambridge, UK, 2017), Section 9.2.3. In general, the energy per unit volume in electromagnetic fields is

$$u = (\epsilon_0/2) E^2 + (1/2\mu_0) B^2.$$

For electromagnetic waves, the electric and magnetic contributions are equal, so

$$u = \epsilon_0 E^2 = \epsilon_0 E_0^2 \cos^2(\omega t),$$

and the average over a full cycle is $(\epsilon_0/2) E_0^2$, since the average of \cos^2 (or \sin^2) is $1/2$.

- ¹² Equation 11.46 assumes that the perturbations at different frequencies are *independent*, so that the total transition probability is a sum of the individual probabilities. If the different components are **coherent** (phase-correlated), then we should add *amplitudes* ($c_b(t)$), not *probabilities* ($|c_b(t)|^2$), and there will be cross-terms. For the applications we will consider the perturbations are always incoherent.
- ¹³ I’ll treat ρ as though it were *real*, even though in general it will be complex. Since

$$|\rho \cdot \hat{n}|^2 = |\operatorname{Re}(\rho) \cdot \hat{n} + i\operatorname{Im}(\rho) \cdot \hat{n}|^2 = |\operatorname{Re}(\rho) \cdot \hat{n}|^2 + |\operatorname{Im}(\rho) \cdot \hat{n}|^2$$

we can do the whole calculation for the real and imaginary parts separately, and simply add the results. In Equation 11.54 the absolute value signs denote *both* the vector magnitude *and* the complex amplitude:

$$|\rho|^2 = |\rho_x|^2 + |\rho_y|^2 + |\rho_z|^2.$$

- ¹⁴ Normally I’d use R for a transition rate, but out of deference to *der Alte* everyone follows Einstein’s notation in this context.
- ¹⁵ Assume that N_a and N_b are very large, so we can treat them as continuous functions of time and ignore statistical fluctuations.
- ¹⁶ See, for example, Daniel Schroeder, *An Introduction to Thermal Physics* (Pearson, Upper Saddle River, NJ, 2000), Section 6.1.
- ¹⁷ Schroeder, footnote 16, Section 7.4.
- ¹⁸ Griffiths, footnote 11, Section 9.2.1.
- ¹⁹ This situation is not to be confused with the case of thermal equilibrium, which we considered in the previous section. We assume here that the atoms have been lifted *out* of equilibrium, and are in the process of cascading back down to their equilibrium levels.
- ²⁰ See, for example, Griffiths, footnote 11, Section 11.2.1.
- ²¹ This is an example of Bohr’s **Correspondence Principle**. In fact, if we express P in terms of the energy *above the ground state*, the two formulas are identical.
- ²² See Equation 6.62 (Equation 6.26 eliminates $\Delta\ell = 0$), or derive them from scratch using Problems 11.14 and 11.15.
- ²³ This is the same set of approximations we made in Equations 11.46–11.48.
- ²⁴ In deriving Equation 11.35, our perturbation was

$$\hat{H}' = V \cos(\omega t) \rightarrow \frac{V}{2} e^{-i\omega t}$$

since we dropped the other (off-resonance) exponential. That is the source of the two inside the absolute value in Equation 11.81. Fermi’s Golden rule can also be applied to a constant perturbation, $\hat{H}' = \hat{V}$, if we set $\omega = 0$ drop the 2:

$$R = \frac{2\pi}{\hbar} |V_{if}|^2 \rho(E_f).$$

- ²⁵ It is actually due to Dirac, but Fermi is the one who gave it the memorable name. See T. Visser, *Am. J. Phys.* **77**, 487 (2009) for the history. Fermi's Golden Rule doesn't just apply to transitions to a continuum of states. For instance, Equation 11.54 can be considered an example. In that case, we integrated over a continuous range of *perturbation frequencies*—not a continuum of final states—but the end result is the same.
- ²⁶ Periodic boundary conditions are discussed in Problem 5.39. In the present context we use periodic boundary conditions—as opposed to impenetrable walls—because they admit traveling-wave solutions.
- ²⁷ Each state in k -space “occupies” a volume of $(2\pi/l)^3$, as shown in Problem 5.39.
- ²⁸ See footnote 24.
- ²⁹ This is an approximation; we really should be using a scattering state of hydrogen. For an extended discussion of the photoelectric effect, including comparison to experiment and the validity of this approximation, see W. Heitler, *The Quantum Theory of Radiation*, 3rd edn, Oxford University Press, London (1954), Section 21.
- ³⁰ The result here is too large by a factor of four; correcting this requires a more careful derivation of the matrix element for radiative transitions (see Problem 11.30). Only the overall factor is affected though; the more interesting features (the dependence on k and θ) are correct.
- ³¹ For an interesting discussion of classical adiabatic processes, see Frank S. Crawford, *Am. J. Phys.* **58**, 337 (1990).
- ³² See, for example, Jerry B. Marion and Stephen T. Thornton, *Classical Dynamics of Particles and Systems*, 4th edn, Saunders, Fort Worth, TX (1995), Example 10.5. Geographers measure latitude (λ) up from the equator, rather than down from the pole, so $\cos\theta_0 = \sin\lambda$.
- ³³ The adiabatic theorem, which is usually attributed to Ehrenfest, is simple to state, and it *sounds* plausible, but it is not easy to prove. The argument will be found in earlier editions of this book, Section 10.1.2.
- ³⁴ I'm suppressing the dependence on other variables; only the time dependence is at issue here.
- ³⁵ As Michael Berry puts it, the dynamic phase answers the question “How long did your trip take?” and the geometric phase, “Where have you been?”
- ³⁶ For more on this subject see Alfred Shapere and Frank Wilczek, eds., *Geometric Phases in Physics*, World Scientific, Singapore (1989); Andrei Bernevig and Taylor Hughes, *Topological Insulators and Topological Superconductors*, Princeton University Press, Princeton, NJ (2013), Chapter 2.
- ³⁷ If $\psi_n(t)$ is *real*, the geometric phase vanishes. You might try to beat the rap by tacking an unnecessary (but perfectly legal) phase factor onto the eigenfunctions: $\psi'_n(t) \equiv e^{i\phi_n} \psi_n(t)$, where ϕ_n is an arbitrary (real) function. Try it. You'll get a nonzero geometric phase, all right, but note what happens when you put it back into Equation 11.93. And for a *closed loop* it gives *zero*.
- ³⁸ The **interaction picture** is intermediate between the Heisenberg and Schrödinger pictures (see Section 6.8.1). In the interaction picture, the wave function satisfies the “Schrödinger equation”

$$i\hbar \frac{d}{dt} |\Psi_I(t)\rangle = \hat{H}'_I(t) |\Psi_I(t)\rangle,$$

where the interaction- and Schrödinger-picture operators are related by

$$\hat{H}'_I(t) = e^{i\hat{H}_0 t/\hbar} \hat{H}'(t) e^{-i\hat{H}_0 t/\hbar}$$

and the wave functions satisfy

$$|\Psi_I(t)\rangle = e^{i\hat{H}_0 t/\hbar} |\Psi(t)\rangle.$$

If you apply the Dyson series to the Schrödinger equation in the interaction picture, you end up with precisely the perturbation series derived in Section 11.1.2. For more details see Ramamurti Shankar, *Principles of Quantum Mechanics*, 2nd edn, Springer, New York (1994), Section 18.3.

- ³⁹ For a systematic treatment (including the role of the magnetic field) see David Park, *Introduction to the Quantum Theory*, 3rd edn (McGraw-Hill, New York, 1992), Chapter 11.
- ⁴⁰ See Masataka Mizushima, *Quantum Mechanics of Atomic Spectra and Atomic Structure*, Benjamin, New York (1970), Section 5.6.
- ⁴¹ George B. Arfken and Hans J. Weber, *Mathematical Methods for Physicists*, 7th edn, Academic Press, San Diego (2013), p. 744.
- ⁴² Electric and magnetic dipole moments have different units—hence the factor of $1/c$ (which you can check by dimensional analysis).
- ⁴³ Julio Gea-Banacloche, *Am. J. Phys.* **70**, 307 (2002) uses a rectangular barrier; the delta-function version was suggested by M. Lakner and J. Peternej, *Am. J. Phys.* **71**, 519 (2003).
- ⁴⁴ Gea-Banacloche (footnote 43) discusses the evolution of the wave function *without* using the adiabatic theorem.
- ⁴⁵ S. W. Doescher and M. H. Rice, *Am. J. Phys.* **37**, 1246 (1969).
- ⁴⁶ See Y. Nogami, *Am. J. Phys.* **59**, 64 (1991), and references therein.

- ⁴⁷ This phenomenon doesn't have much to do with Zeno, but it *is* reminiscent of the old adage, "a watched pot never boils," so it is sometimes called the **watched pot effect**.
- ⁴⁸ In the argument leading to linear time dependence, we assumed that the function $\sin^2(\Omega t/2) / \Omega^2$ in Equation 11.46 was a sharp spike. However, the *width* of the "spike" is of order $\Delta\omega = 4\pi/t$, and for *extremely* short t this assumption fails, and the integral becomes $(t^2/4) \int \rho(\omega) d\omega$.
- ⁴⁹ This argument was introduced by B. Misra and E. C. G. Sudarshan, *J. Math. Phys.* **18**, 756 (1977). The essential result has been confirmed in the laboratory: W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland, *Phys. Rev. A* **41**, 2295 (1990). Unfortunately, the experiment is not as compelling a test of the collapse of the wave function as its designers hoped, for the observed effect can perhaps be accounted for in other ways—see L. E. Ballentine, *Found. Phys.* **20**, 1329 (1990); T. Petrosky, S. Tasaki, and I. Prigogine, *Phys. Lett. A* **151**, 109 (1990).
- ⁵⁰ If you choose Δt small enough, you can actually use this exact form. Routines such as Mathematica's **MatrixExp** can be used to find (numerically) the exponential of a matrix.
- ⁵¹ See A. Goldberg *et al.*, *Am. J. Phys.* **35**, 177 (1967) for further discussion of these approximations.
- ⁵² A tri-diagonal matrix has nonzero entries only along the diagonal and one space to the right or left of the diagonal.
- ⁵³ Use your computing environment's built-in linear equation solver; in Mathematica that would be `x = LinearSolve[M, b]`. To learn how it actually works, see A. Goldberg *et al.*, footnote 51.
- ⁵⁴ C. Lubich, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, edited by J. Grotendorst, D. Marx, and A. Muramatsu (John von Neumann Institute for Computing, Jülich, 2002), Vol. 10, p. 459. Available for download from the Neumann Institute for Computing (NIC) website.

12

Afterword

◊

Now that you have a sound understanding of what quantum mechanics *says*, I would like to return to the question of what it *means*—continuing the story begun in Section 1.2. The source of the problem is the indeterminacy associated with the statistical interpretation of the wave function. For Ψ (or, more generally, the *quantum state*—it could be a spinor, for example) does not uniquely determine the outcome of a measurement; all it tells us is the statistical distribution of possible results. This raises a profound question: Did the physical system “actually have” the attribute in question *prior* to the measurement (the so-called **realist** viewpoint), or did the act of measurement itself “create” the property, limited only by the statistical constraint imposed by the wave function (the **orthodox** position)—or can we duck the issue entirely, on the grounds that it is “metaphysical” (the **agnostic** response)?

According to the realist, quantum mechanics is an *incomplete* theory, for even if you know *everything quantum mechanics has to tell you* about the system (to wit: its wave function), still you cannot determine all of its features. Evidently there is some *other* information, unknown to quantum mechanics, which (together with Ψ) is required for a complete description of physical reality.

The orthodox position raises even more disturbing problems, for if the act of measurement forces the system to “take a stand,” helping to *create* an attribute that was not there previously,¹ then there is something very peculiar about the measurement process. Moreover, in order to account for the fact that an immediately repeated measurement yields the same result, we are forced to assume that the act of measurement **collapses** the wave function, in a manner that is difficult, at best, to reconcile with the normal evolution prescribed by the Schrödinger equation.

In light of this, it is no wonder that generations of physicists retreated to the agnostic position, and advised their students not to waste time worrying about the conceptual foundations of the theory.

12.1 The EPR Paradox

In 1935, Einstein, Podolsky, and Rosen² published the famous **EPR paradox**, which was designed to prove (on purely theoretical grounds) that the realist position is the only tenable one. I'll describe a simplified version of the EPR paradox, due to David Bohm (call it EPRB). Consider the decay of the neutral pi meson into an electron and a positron:

$$\pi^0 \rightarrow e^- + e^+.$$

Assuming the pion was at rest, the electron and positron fly off in opposite directions (Figure 12.1). Now, the pion has spin zero, so conservation of angular momentum requires that the electron and positron occupy the singlet spin configuration:

$$\frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \quad (12.1)$$

If the electron is found to have spin up, the positron must have spin down, and vice versa. Quantum mechanics can't tell you *which* combination you'll get, in any particular pion decay, but it does say that the measurements will be *correlated*, and you'll get each combination half the time (on average). Now suppose we let the electron and positron fly *far* off—10 meters, in a practical experiment, or, in principle, 10 light years—and then you measure the spin of the electron. Say you get spin up. Immediately you know that someone 20 meters (or 20 light years) away will get spin down, if that person examines the positron.

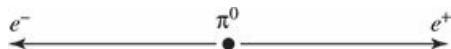


Figure 12.1: Bohm's version of the EPR experiment: A π^0 at rest decays into an electron–positron pair.

To the realist, there's nothing surprising about this—the electron *really had* spin up (and the positron spin down) from the moment they were created ...it's just that quantum mechanics didn't know about it. But the “orthodox” view holds that neither particle had either spin up *or* spin down until the act of measurement intervened: Your measurement of the electron collapsed the wave function, and instantaneously “produced” the spin of the positron 20 meters (or 20 light years) away. Einstein, Podolsky, and Rosen considered such “spooky action-at-a-distance” (Einstein's delightful term) preposterous. They concluded that the orthodox position is untenable; the electron and positron must have had well-defined spins all along, whether quantum mechanics knows it or not.

The fundamental assumption on which the EPR argument rests is that no influence can propagate faster than the speed of light. We call this the principle of **locality**. You might be tempted to propose that the collapse of the wave function is *not* instantaneous, but “travels” at some finite velocity. However, this would lead to violations of angular momentum conservation, for if we measured the spin of the positron before the news of the collapse had reached it, there would be a fifty-fifty probability of finding *both* particles with spin up. Whatever you might think of such a theory in the abstract, the experiments are unequivocal: No such violation occurs—the (anti-)correlation of the spins is perfect. Evidently the collapse of the wave function—whatever its ontological status—is instantaneous.³

Problem 12.1 Entangled states. The singlet spin configuration (Equation 12.1) is the classic example of an *entangled state*—a two-particle state that cannot be expressed as the product of two one-particle states, and for which, therefore, one cannot really speak of “the state” of either particle separately.⁴ You might wonder whether this is somehow an artifact of bad notation—maybe some linear combination of the one-particle states would disentangle the system. Prove the following theorem:

Consider a two-level system, $|\phi_a\rangle$ and $|\phi_b\rangle$, with $\langle\phi_i|\phi_j\rangle = \delta_{ij}$. (For example, $|\phi_a\rangle$ might represent spin up and $|\phi_b\rangle$ spin down.) The two-particle state

$$\alpha|\phi_a(1)\rangle|\phi_b(2)\rangle + \beta|\phi_b(1)\rangle|\phi_a(2)\rangle$$

(with $\alpha \neq 0$ and $\beta \neq 0$) *cannot* be expressed as a product

$$|\psi_r(1)\rangle|\psi_s(2)\rangle,$$

for any one-particle states $|\psi_r\rangle$ and $|\psi_s\rangle$.

Hint: Write $|\psi_r\rangle$ and $|\psi_s\rangle$ as linear combinations of $|\phi_a\rangle$ and $|\phi_b\rangle$.

Problem 12.2 Einstein’s Boxes. In an interesting precursor to the EPR paradox, Einstein proposed the following gedanken experiment.⁵ Imagine a particle confined to a box (make it a one-dimensional infinite square well, if you like). It’s in the ground state, when an impenetrable partition is introduced, dividing the box into separate halves, B_1 and B_2 , in such a way that the particle is equally likely to be found in either one.⁶ Now the two boxes are moved very far apart, and a measurement is made on B_1 to see if the particle is in that box. Suppose the answer is *yes*. Immediately we know that the particle will *not* be found in the (distant) box B_2 .

- (a) What would Einstein say about this?
- (b) How does the Copenhagen interpretation account for it? What is the wave function in B_2 , right after the measurement on B_1 ?

12.2 Bell's Theorem

Einstein, Podolsky, and Rosen did not doubt that quantum mechanics is *correct*, as far as it goes; they only claimed that it is an *incomplete* description of physical reality: The wave function is not the whole story—some *other* quantity, λ , is needed, in addition to Ψ , to characterize the state of a system fully. We call λ the “hidden variable” because, at this stage, we have no idea how to calculate or measure it.⁷ Over the years, a number of hidden variable theories have been proposed, to supplement quantum mechanics;⁸ they tend to be cumbersome and implausible, but never mind—until 1964 the program seemed eminently worth pursuing. But in that year J. S. Bell proved that *any* local hidden variable theory is *incompatible* with quantum mechanics.⁹

Bell suggested a generalization of the EPRB experiment: Instead of orienting the electron and positron detectors along the *same* direction, he allowed them to be rotated independently. The first measures the component of the electron spin in the direction of a unit vector \mathbf{a} , and the second measures the spin of the positron along the direction \mathbf{b} (Figure 12.2). For simplicity, let’s record the spins in units of $\hbar/2$; then each detector registers the value $+1$ (for spin up) or -1 (spin down), along the direction in question. A table of results, for many π^0 decays, might look like this:

electron	positron	product
+1	-1	-1
+1	+1	+1
-1	+1	-1
+1	-1	-1
-1	-1	+1
:	:	:

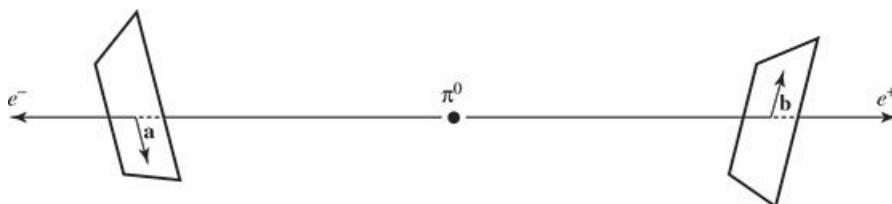


Figure 12.2: Bell’s version of the EPRB experiment: detectors independently oriented in directions \mathbf{a} and \mathbf{b} .

Bell proposed to calculate the *average* value of the *product* of the spins, for a given set of detector orientations. Call this average $P(\mathbf{a}, \mathbf{b})$. If the detectors are parallel ($\mathbf{b} = \mathbf{a}$), we recover the original EPRB configuration; in this case one is spin up and the other spin down, so the product is always -1 , and hence so too is the average:

$$P(\mathbf{a}, \mathbf{a}) = -1. \tag{12.2}$$

By the same token, if they are *anti-parallel* ($\mathbf{b} = -\mathbf{a}$), then every product is $+1$, so

$$P(\mathbf{a}, -\mathbf{a}) = +1. \tag{12.3}$$

For arbitrary orientations, quantum mechanics predicts

$$P(\mathbf{a}, \mathbf{b}) = -\mathbf{a} \cdot \mathbf{b} \quad (12.4)$$

(see Problem 4.59). What Bell discovered is that this result is *incompatible with any local hidden variable theory*.

The argument is stunningly simple. Suppose that the “complete” state of the electron–positron system is characterized by the hidden variable(s) λ (λ varies, in some way that we neither understand nor control, from one pion decay to the next). Suppose further that the outcome of the *electron* measurement is independent of the orientation (\mathbf{b}) of the *positron* detector—which may, after all, be chosen by the experimenter at the positron end just before the electron measurement is made, and hence far too late for any subluminal message to get back to the electron detector. (This is the locality assumption.) Then there exists some function $A(\mathbf{a}, \lambda)$ which determines the result of an electron measurement, and some other function $B(\mathbf{b}, \lambda)$ for the positron measurement. These functions can only take on the values ± 1 :¹⁰

$$A(\mathbf{a}, \lambda) = \pm 1; \quad B(\mathbf{b}, \lambda) = \pm 1. \quad (12.5)$$

When the detectors are aligned, the results are perfectly (anti)-correlated:

$$A(\mathbf{a}, \lambda) = -B(\mathbf{a}, \lambda), \quad (12.6)$$

regardless of the value of λ .

Now, the average of the product of the measurements is

$$P(\mathbf{a}, \mathbf{b}) = \int \rho(\lambda) A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda) d\lambda, \quad (12.7)$$

where $\rho(\lambda)$ is the probability density for the hidden variable. (Like any probability density, it is real, nonnegative, and satisfies the normalization condition $\int \rho(\lambda) d\lambda = 1$, but beyond this we make no assumptions about $\rho(\lambda)$; different hidden variable theories would presumably deliver quite different expressions for ρ .) In view of Equation 12.6, we can eliminate B :

$$P(\mathbf{a}, \mathbf{b}) = - \int \rho(\lambda) A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) d\lambda. \quad (12.8)$$

If \mathbf{c} is any *other* unit vector,

$$P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c}) = - \int \rho(\lambda) [A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) - A(\mathbf{a}, \lambda) A(\mathbf{c}, \lambda)] d\lambda. \quad (12.9)$$

Or, since $[A(\mathbf{b}, \lambda)]^2 = 1$:

$$P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c}) = - \int \rho(\lambda) [1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)] A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) d\lambda. \quad (12.10)$$

But it follows from Equation 12.5 that $|A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda)| = 1$; moreover $\rho(\lambda) [1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)] \geq 0$, so

$$|P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})| \leq \int \rho(\lambda) [1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)] d\lambda, \quad (12.11)$$

or, more simply:

$$|P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})| \leq 1 + P(\mathbf{b}, \mathbf{c}). \quad (12.12)$$

This is the famous **Bell inequality**. It holds for *any* local hidden variable theory (subject only to the minimal requirements of Equations 12.5 and 12.6), for we have made no assumptions whatever as to the nature or number of the hidden variable(s), or their distribution (ρ).

But it is easy to show that the quantum mechanical prediction (Equation 12.4) is incompatible with Bell's inequality. For example, suppose the three vectors lie in a plane, and c makes a 45° angle with a and b (Figure 12.3); in this case quantum mechanics says

$$P(a, b) = 0, \quad P(a, c) = P(b, c) = -0.707,$$

which is patently inconsistent with Bell's inequality:

$$0.707 \not\leq 1 - 0.707 = 0.293.$$

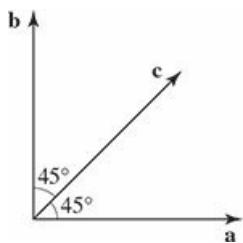


Figure 12.3: An orientation of the detectors that demonstrates quantum violations of Bell's inequality.

With Bell's modification, then, the EPR paradox proves something far more radical than its authors imagined: If they are right, then not only is quantum mechanics *incomplete*, it is downright *wrong*. On the other hand, if quantum mechanics is right, then *no* hidden variable theory is going to rescue us from the nonlocality Einstein considered so preposterous. Moreover, we are provided with a very simple experiment to settle the issue once and for all.¹¹

Many experiments to test Bell's inequality were performed in the 1960s and 1970s, culminating in the work of Aspect, Grangier, and Roger.¹² The details do not concern us here (they actually used two-photon atomic transitions, not pion decays). To exclude the remote possibility that the positron detector might somehow “sense” the orientation of the electron detector, both orientations were set quasi-randomly *after* the photons were already in flight. The results were in excellent agreement with the predictions of quantum mechanics, and inconsistent with Bell's inequality by a wide margin.¹³

Ironically, the experimental confirmation of quantum mechanics came as something of a shock to the scientific community. But not because it spelled the demise of “realism”—most physicists had long since adjusted to this (and for those who could not, there remained the possibility of *nonlocal* hidden variable theories, to which Bell's theorem does not apply).¹⁴ The real shock was the demonstration that *nature itself is fundamentally nonlocal*. Nonlocality, in the form of the instantaneous collapse of the wave function (and for that matter also in the symmetrization requirement for identical particles) had always been a feature of the orthodox interpretation, but before Aspect's experiment it was possible to hope that quantum nonlocality was somehow a nonphysical artifact of the formalism, with no detectable consequences. That hope can no longer be sustained, and we are obliged to reexamine our objection to instantaneous action-at-a-distance.

Why *are* physicists so squeamish about superluminal influences? After all, there are many things that travel faster than light. If a bug flies across the beam of a movie projector, the speed of its shadow is proportional to the distance to the screen; in principle, that distance can be as large as you like, and hence the *shadow* can move at arbitrarily high velocity (Figure 12.4). However, the shadow does not carry any *energy*,

nor can it transmit any *information* from one point on the screen to another. A person at point X cannot *cause anything to happen* at point Y by manipulating the passing shadow.



Figure 12.4: The shadow of the bug moves across the screen at a velocity v' greater than c , provided the screen is far enough away.

On the other hand, a *causal* influence that propagated faster than light would carry unacceptable implications. For according to special relativity there exist inertial frames in which such a signal propagates *backward in time*—the effect preceding the cause—and this leads to inescapable logical anomalies. (You could, for example, arrange to kill your infant grandfather. Think about it ...not a good idea.) The question is, are the superluminal influences predicted by quantum mechanics and detected by Aspect *causal*, in this sense, or are they somehow ethereal enough (like the bug’s shadow) to escape the philosophical objection?

Well, let’s consider Bell’s experiment. Does the measurement of the electron *influence* the outcome of the positron measurement? Assuredly it *does*—otherwise we cannot account for the correlations in the data. But does the measurement of the electron *cause* a particular outcome for the positron? Not in any ordinary sense of the word. There is no way the person manning the electron detector could use his measurement to send a signal to the person at the positron detector, since he does not control the outcome of his own measurement (he cannot *make* a given electron come out spin up, any more than the person at X can affect the passing shadow of the bug). It is true that he can decide *whether to make a measurement at all*, but the positron monitor, having immediate access only to data at his end of the line, cannot tell whether the electron has been measured or not. The lists of data compiled at the two ends, considered separately, are completely random. It is only later, when we *compare* the two lists, that we discover the remarkable correlations. In another reference frame the positron measurements occur *before* the electron measurements, and yet this leads to no logical paradox—the observed correlation is entirely symmetrical in its treatment, and it is a matter of indifference whether we say the observation of the electron influenced the measurement of the positron, or the other way around. This is a wonderfully delicate kind of influence, whose only manifestation is a subtle correlation between two lists of otherwise random data.

We are led, then, to distinguish two types of influence: the “causal” variety, which produce actual changes in some physical property of the receiver, detectable by measurements on that subsystem alone, and an “ethereal” kind, which do not transmit energy or information, and for which the only evidence is a correlation in the data taken on the two separate subsystems—a correlation which by its nature cannot be detected by examining either list alone. Causal influences *cannot* propagate faster than light, but there is no compelling reason why ethereal ones should not. The influences associated with the collapse of the wave function are of the latter type, and the fact that they “travel” faster than light may be surprising, but it is not, after all, catastrophic.¹⁵

**

Problem 12.3 One example¹⁶ of a (local) deterministic (“hidden variable”) theory is ...classical mechanics! Suppose we carried out the Bell experiment with classical objects (baseballs, say) in place of the electron and proton. They are launched (by a kind of double pitching machine) in opposite directions, with equal and opposite spins (angular momenta), \mathbf{S}_a and $\mathbf{S}_b = -\mathbf{S}_a$. Now, these are *classical* objects—their angular momenta can point in any direction, and this direction is set (let’s say randomly) at the moment of launch. Detectors placed 10 meters or so on either side of the launch point measure the spin vectors of their respective baseballs. However, in order to match the conditions for Bell’s theorem, they only *record* the *sign* of the *component* of \mathbf{S} along the directions \mathbf{a} and \mathbf{b} :

$$A \equiv \text{sign}(\mathbf{a} \cdot \mathbf{S}_a), \quad B \equiv \text{sign}(\mathbf{b} \cdot \mathbf{S}_b).$$

Thus each detector records either + 1 or - 1, in any given trial.

In this example the “hidden variable” is the *actual* orientation of \mathbf{S}_a , specified (say) by the polar and azimuthal angles θ and ϕ : $\lambda = (\theta, \phi)$:

- (a) Choosing axes as in the figure, with \mathbf{a} and \mathbf{b} in the x - y plane and \mathbf{a} along the x axis, verify that

$$A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda) = -\text{sign}[\cos(\phi) \cos(\phi - \eta)],$$

where η is the angle between \mathbf{a} and \mathbf{b} (take it to run from $-\pi$ to $+\pi$).

- (b) Assuming the baseballs are launched in such a way that \mathbf{S}_a is equally likely to point in any direction, compute $P(\mathbf{a}, \mathbf{b})$. *Answer:* $(2|\eta|/\pi) - 1$.
- (c) Sketch the graph of $P(\mathbf{a}, \mathbf{b})$, from $\eta = -\pi$ to $+\pi$, and (on the same graph) the quantum formula (Equation 12.4, with $\theta \rightarrow \eta$). For what values of η does this hidden variable theory agree with the quantum-mechanical result?
- (d) Verify that your result satisfies Bell’s inequality, Equation 12.12. *Hint:* The vectors \mathbf{a} , \mathbf{b} , and \mathbf{c} define three points on the surface of a unit sphere; the inequality can be expressed in terms of the distances between those points.

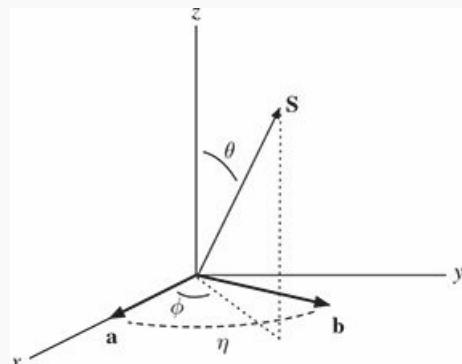


Figure 12.5: Axes for Problem 12.3.

12.3 Mixed States and the Density Matrix

12.3.1 Pure States

In this book we have dealt with particles in **pure states**, $|\Psi\rangle$ —a harmonic oscillator in its n th stationary state, for instance, or in a specific linear combination of stationary states, or a free particle in a gaussian wave packet. The expectation value of some observable A is then

$$\langle A \rangle = \langle \Psi | \hat{A} | \Psi \rangle; \quad (12.13)$$

it's the average of measurements on an ensemble of identically-prepared systems, all of them in the same state $|\Psi\rangle$. We developed the whole theory in terms of $|\Psi\rangle$ (a vector in Hilbert space, or, in the position basis, the wave function).

But there are other ways to formulate the theory, and a particularly useful one starts by defining the **density operator**,¹⁷

$$\hat{\rho} \equiv |\Psi\rangle\langle\Psi|. \quad (12.14)$$

With respect to an orthonormal basis $\{|e_j\rangle\}$ an operator is represented by a *matrix*; the ij element of the matrix \mathbf{A} representing the operator \hat{A} is

$$A_{ij} = \langle e_i | \hat{A} | e_j \rangle. \quad (12.15)$$

In particular, the ij element of the **density matrix** ρ is

$$\rho_{ij} = \langle e_i | \hat{\rho} | e_j \rangle = \langle e_i | \Psi \rangle \langle \Psi | e_j \rangle. \quad (12.16)$$

The density matrix (for pure states) has several interesting properties:

$$\rho^2 = \rho, \quad (\text{it is idempotent}), \quad (12.17)$$

$$\rho^\dagger = \rho, \quad (\text{it is hermitian}), \quad (12.18)$$

$$\text{Tr}(\rho) = \sum_i \rho_{ii} = 1, \quad (\text{its trace is 1}). \quad (12.19)$$

The expectation value of an observable A is

$$\langle A \rangle = \text{Tr}(\rho \mathbf{A}). \quad (12.20)$$

We could do everything using the density matrix, instead of the wave function, to represent the state of a particle.

Example 12.1

In the standard basis

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (12.21)$$

representing spin up and spin down along the z direction (Equation 4.149), construct the density matrix for an electron with spin up along the x direction.

Solution: In this case

$$|\Psi\rangle = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad (12.22)$$

(Equation 4.151). So

$$\begin{aligned}\rho_{11} &= \left[(1 \ 0) \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \right] \left[(1/\sqrt{2} \ 1/\sqrt{2}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} = \frac{1}{2}, \\ \rho_{12} &= \left[(1 \ 0) \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \right] \left[(1/\sqrt{2} \ 1/\sqrt{2}) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} = \frac{1}{2}, \\ \rho_{21} &= \left[(0 \ 1) \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \right] \left[(1/\sqrt{2} \ 1/\sqrt{2}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} = \frac{1}{2}, \\ \rho_{22} &= \left[(0 \ 1) \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \right] \left[(1/\sqrt{2} \ 1/\sqrt{2}) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} = \frac{1}{2},\end{aligned}$$

and hence

$$\rho = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}. \quad (12.23)$$

Or, more efficiently,

$$\rho = |\Psi\rangle\langle\Psi| = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} (1/\sqrt{2} \ 1/\sqrt{2}) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}. \quad (12.24)$$

Note that ρ is hermitian, its trace is 1, and

$$\rho^2 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \rho. \quad (12.25)$$

Problem 12.4

- (a) Prove properties 12.17, 12.18, 12.19, and 12.20.
- (b) Show that the time evolution of the density operator is governed by the equation

$$i\hbar \frac{d\hat{\rho}}{dt} = [\hat{H}, \hat{\rho}]. \quad (12.26)$$

(This is the Schrödinger equation, expressed in terms of $\hat{\rho}$.)

Problem 12.5 Repeat Example 12.1 for an electron with spin down along the y direction.

12.3.2 Mixed States

In practice it is often the case that we simply don't *know* the state of the particle. Suppose, for example, we are interested in an electron emerging from the Stanford Linear Accelerator. It might have spin up (along some chosen direction), or it might have spin down, or it might be in some linear combination of the two—we just don't know.¹⁸ We say that the particle is in a **mixed state**.¹⁹

How should we describe such a particle? I could simply list the *probability*, p_k , that it's in each possible state $|\Psi_k\rangle$. The expectation value of an observable would now be the average of measurements taken over an ensemble of systems that are *not* identically prepared (they are *not* all in the same state); rather, a fraction p_k of them is in each (pure) state $|\Psi_k\rangle$:

$$\langle A \rangle = \sum_k p_k \langle \Psi_k | \hat{A} | \Psi_k \rangle. \quad (12.27)$$

There's a slick way to package this information, by generalizing the density operator:

$$\hat{\rho} \equiv \sum_k p_k |\Psi_k\rangle \langle \Psi_k|. \quad (12.28)$$

Again, it becomes a matrix when referred to a particular basis:

$$\rho_{ij} = \sum_k p_k \langle e_i | \Psi_k \rangle \langle \Psi_k | e_j \rangle. \quad (12.29)$$

The density matrix encodes all the information available to us about the system.

Like any probabilities,

$$0 \leq p_k \leq 1 \quad \text{and} \quad \sum_k p_k = 1. \quad (12.30)$$

The density matrix for mixed states retains most of the properties we identified for pure states:

$$\rho^\dagger = \rho, \quad (12.31)$$

$$\text{Tr}(\rho) = 1, \quad (12.32)$$

$$\langle A \rangle = \text{Tr}(\rho A). \quad (12.33)$$

$$i\hbar \frac{d\hat{\rho}}{dt} = [\hat{H}, \hat{\rho}], \quad (\text{if } \frac{dp_k}{dt} = 0 \text{ for all } k) \quad (12.34)$$

but ρ is idempotent *only* if it represents a pure state:

$$\rho^2 \neq \rho \quad (12.35)$$

(indeed, this is a quick way to test whether the state is pure).

Example 12.2

Construct the density matrix for an electron that is either in the spin-up state or the spin-down state (along z), with equal probability.

Solution: In this case $p_1 = p_2 = 1/2$, so

$$\begin{aligned}\rho &= \sum_k p_k |\Psi_k\rangle\langle\Psi_k| = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \ 0) + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \ 1) \\ &= \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}.\end{aligned}\quad (12.36)$$

Note that ρ is hermitian, and its trace is 1, but

$$\rho^2 = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix} \neq \rho;\quad (12.37)$$

this is *not* a pure state.

Problem 12.6

- (a) Prove properties [12.31](#), [12.32](#), [12.33](#), and [12.34](#).
- (b) Show that $\text{Tr}(\rho^2) \leq 1$, and equal to 1 only if ρ represents a pure state.
- (c) Show that $\rho^2 = \rho$ if and only if ρ represents a pure state.

Problem 12.7

- (a) Construct the density matrix for an electron that is either in the state spin up along x (with probability 1/3) or in the state spin down along y (with probability 2/3).
- (b) Find $\langle S_y \rangle$ for the electron in (a).

Problem 12.8

- (a) Show that the most general density matrix for a spin-1/2 particle can be written in terms of three real numbers (a_1, a_2, a_3) :

$$\rho = \frac{1}{2} \begin{pmatrix} (1 + a_3) & (a_1 - ia_2) \\ (a_1 + ia_2) & (1 - a_3) \end{pmatrix} = \frac{1}{2}(1 + \mathbf{a} \cdot \boldsymbol{\sigma}), \quad (12.38)$$

where $\sigma_1, \sigma_2, \sigma_3$ are the three Pauli matrices. *Hint:* It has to be hermitian, and its trace must be 1.

- (b) In the literature, \mathbf{a} is known as the **Bloch vector**. Show that ρ represents a pure state if and only if $|\mathbf{a}| = 1$, and for a mixed state $|\mathbf{a}| < 1$. *Hint:* Use Problem [12.6\(c\)](#). Thus every density matrix for a spin-1/2 particle corresponds to a point in the **Bloch sphere**, of radius 1. Points on the surface are pure states, and points inside are mixed states.
- (c) What is the probability that a measurement of S_z would return the value $+\hbar/2$, if the tip of the Bloch vector is at (i) the north pole ($\mathbf{a} = (0, 0, 1)$), (ii) the center of the sphere ($\mathbf{a} = (0, 0, 0)$), (iii) the south pole ($\mathbf{a} = (0, 0, -1)$)?

- (d) Find the spinor χ representing the (pure) state of the system, if the Bloch vector lies on the equator, at azimuthal angle ϕ .

12.3.3 Subsystems

There is another context in which one might invoke the density matrix formalism: an entangled state, such as the singlet spin configuration of an electron/positron pair,

$$\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \quad (12.39)$$

Suppose we are interested only in the positron: what is its state? I cannot say ...a measurement could return spin up (fifty-fifty probability) or spin down. This has nothing to do with *ignorance*; I know the state of the system precisely. But the *subsystem* (the positron) by itself does *not* occupy a pure state. If I insist on talking about the positron alone, the best I can do is to tell you its density matrix:

$$\rho = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}, \quad (12.40)$$

representing the 50/50 mixture.

Of course, this is the *same* as the density matrix representing a positron in a specific (but unknown) spin state (Example 12.2). I'll call it a **subsystem** density matrix, to distinguish it from an **ignorance** density matrix. The EPRB paradox illustrates the difference. Before the electron spin was measured, the positron (alone) was represented by the "subsystem" density matrix (Equation 12.40); when the electron is measured the positron is knocked into a definite state ...but we (at the distant positron detector) don't know which. The positron is now represented by the "ignorance" density matrix (Equation 12.36). But the two density matrices are identical! Our description of the state of the positron has not been altered by the measurement of the electron —all that has changed is our reason for using the density matrix formalism.

12.4 The No-Clone Theorem

Quantum measurements are typically **destructive**, in the sense that they alter the state of the system measured. This is how the uncertainty principle is enforced in the laboratory. You might wonder why we don't just make a bunch of identical copies (**clones**) of the original state, and measure *them*, leaving the system itself unscathed. It can't be done. Indeed, if you *could* build a cloning device (a “quantum Xerox machine”), quantum mechanics would be out the window.

For example, it would then be possible to send superluminal messages using the EPRB apparatus.²⁰ Say the message to be transmitted, from the operator of the electron detector (conventionally “Alice”) to the operator of the positron detector (“Bob”), is either “yes” (“drop the bomb”) or “no.” If the message is to be “yes,” Alice measures S_z (of the electron). Never mind what result she gets—all that matters is that she makes the measurement, for this means that the positron is now in the pure state \uparrow or \downarrow (never mind which). If she wants to say “no,” she measures S_x , and that means the positron is now in the definite state \leftarrow or \rightarrow (never mind which). In any case, Bob makes a million clones of the positron, and measures S_z on half of them, and S_x on the other half. If the first group are all in the same state (all \uparrow or all \downarrow), then Alice must have measured S_z , and the message is “yes” (the S_x group should be a 50/50 mixture). If all the S_x measurements yield the same answer (all \leftarrow or all \rightarrow), then Alice must have measured S_x , and the message is “no” (in that case the S_z measurements should be a 50/50 mixture).

It doesn't work, because you *can't make* a quantum Xerox machine, as Wootters, Zurek, and Dieks proved in 1982.²¹ Schematically, we want the machine to take as input a particle in state $|\psi\rangle$ (the one to be copied), plus a second particle in state $|X\rangle$ (the “blank sheet of paper”), and spit out *two* particles in the state $|\psi\rangle$ (original plus copy):

$$|\psi\rangle|X\rangle \rightarrow |\psi\rangle|\psi\rangle. \quad (12.41)$$

Suppose we have made a device that successfully clones the state $|\psi_1\rangle$:

$$|\psi_1\rangle|X\rangle \rightarrow |\psi_1\rangle|\psi_1\rangle, \quad (12.42)$$

and also works for state $|\psi_2\rangle$:

$$|\psi_2\rangle|X\rangle \rightarrow |\psi_2\rangle|\psi_2\rangle \quad (12.43)$$

($|\psi_1\rangle$ and $|\psi_2\rangle$ might be spin up and spin down, for example, if the particle is an electron). So far, so good. But what happens when we feed in a linear combination $|\psi\rangle = \alpha|\psi_1\rangle + \beta|\psi_2\rangle$? Evidently we get²²

$$|\psi\rangle|X\rangle \rightarrow \alpha|\psi_1\rangle|\psi_1\rangle + \beta|\psi_2\rangle|\psi_2\rangle, \quad (12.44)$$

which is not at all what we wanted—what we *wanted* was

$$\begin{aligned} |\psi\rangle|X\rangle &\rightarrow |\psi\rangle|\psi\rangle = [\alpha|\psi_1\rangle + \beta|\psi_2\rangle][\alpha|\psi_1\rangle + \beta|\psi_2\rangle] \\ &= \alpha^2|\psi_1\rangle|\psi_1\rangle + \beta^2|\psi_2\rangle|\psi_2\rangle + \alpha\beta[|\psi_1\rangle|\psi_2\rangle + |\psi_2\rangle|\psi_1\rangle]. \end{aligned} \quad (12.45)$$

You can make a machine to clone spin-up electrons and spin-down electrons, but it will fail for any nontrivial linear combinations (such as eigenstates of S_x). It's as though you bought a Xerox machine that copies vertical lines perfectly, and also horizontal lines, but completely distorts diagonals.

The no-clone theorem turned out to have an importance well beyond “merely” protecting quantum mechanics from superluminal communication (and hence an inescapable conflict with special relativity).²³ In particular, it opened up the field of **quantum cryptography**, which exploits the theorem to detect eavesdropping.²⁴ This time Alice and Bob want to agree on a key for decoding messages, without the cumbersome necessity of actually meeting face-to-face. Alice is to send the key (a string of numbers) to Bob via a stream of carefully prepared photons.²⁵ But they are worried that their nemesis, Eve, might try to intercept this communication, and thereby crack the code, without their knowledge. Alice prepares a string of photons in four different states: linearly polarized (horizontal $|H\rangle$ and vertical $|V\rangle$), and circularly polarized (left $|L\rangle$ and right $|R\rangle$), which she sends to Bob. Eve hopes to capture and clone the photons en route, sending the originals along to Bob, who will be none the wiser. (Later on, she knows, Alice and Bob will compare notes on a sample of the photons, to make sure there has been no tampering—that’s why she has to clone them perfectly, to go undetected.) But the no-clone theorem guarantees that Eve’s Xerox machine will fail;²⁶ Alice and Bob will catch the eavesdropping when they compare the samples. (They will then, presumably, discard that key.)

12.5 Schrödinger's Cat

The measurement process plays a mischievous role in quantum mechanics: It is here that indeterminacy, nonlocality, the collapse of the wave function, and all the attendant conceptual difficulties arise. Absent measurement, the wave function evolves in a leisurely and deterministic way, according to the Schrödinger equation, and quantum mechanics looks like a rather ordinary field theory (much simpler than classical electrodynamics, for example, since there is only *one* field (Ψ), instead of *two* (E and B), and it's a *scalar*). It is the bizarre role of the measurement process that gives quantum mechanics its extraordinary richness and subtlety. But what, exactly, *is* a measurement? What makes it so different from other physical processes?²⁷ And how can we tell when a measurement has occurred?

Schrödinger posed the essential question most starkly, in his famous **cat paradox**:²⁸

A cat is placed in a steel chamber, together with the following hellish contraption.... In a Geiger counter there is a tiny amount of radioactive substance, so tiny that maybe within an hour one of the atoms decays, but equally probably none of them decays. If one decays then the counter triggers and via a relay activates a little hammer which breaks a container of cyanide. If one has left this entire system for an hour, then one would say the cat is living if no atom has decayed. The first decay would have poisoned it. The wave function of the entire system would express this by containing equal parts of the living and dead cat.

At the end of the hour, then, the wave function of the cat has the schematic form

$$\psi = \frac{1}{\sqrt{2}}(\psi_{\text{alive}} + \psi_{\text{dead}}). \quad (12.46)$$

The cat is neither alive nor dead, but rather a linear combination of the two, until a measurement occurs—until, say, you peek in the window to check. At that moment your observation forces the cat to “take a stand”: dead or alive. And if you find him to be dead, then it's really *you* who killed him, by looking in the window.

Schrödinger regarded this as patent nonsense, and I think most people would agree with him. There is something absurd about the very idea of a *macroscopic* object being in a linear combination of two palpably different states. An electron can be in a linear combination of spin up and spin down, but a cat simply cannot *be* in a linear combination of alive and dead. But how are we to reconcile this with quantum mechanics?

The Schrödinger cat paradox forces us to confront the question “What constitutes a ‘measurement,’ in quantum mechanics?” Does the “measurement” really occur when we peek in the keyhole? Or did it happen much earlier, when the atom did (or did not) decay? Or was it when the Geiger counter registered (or did not) the decay, or when the hammer did (or did not) hit the vial of cyanide? Historically, there have been many answers to this question. Wigner held that measurement requires the intervention of human consciousness; Bohr thought it meant the interaction between a microscopic system (subject to the laws of quantum mechanics) and a macroscopic measuring apparatus (described by classical laws); Heisenberg maintained that a measurement occurs when a permanent record is left; others have pointed to the *irreversible* nature of a measurement. The embarrassing fact is that none of these characterizations is entirely satisfactory. Most physicists would say that the measurement occurred (and the cat became either alive or dead) well before we looked in the window, but there is no real consensus as to when or why.

And this still leaves the deeper question of why a macroscopic system cannot occupy a linear combination of two clearly distinct states—a baseball, say, in a linear combination of Seattle and Toronto. Suppose you *could* get a baseball into such a state, what would happen to it? In some ultimate sense the macroscopic system must itself be described by the laws of quantum mechanics. But wave functions, in the first instance, represent

individual elementary particles; the wave function of a macroscopic object would be a monstrously complicated composite structure, built out of the wave functions of its 10^{23} constituent particles. And it is subject to constant bombardment from the environment²⁹ —subject, that is, to continuous “measurement” and the attendant collapse. In this process, presumably, “classical” states are statistically favored, and in practice the linear combination devolves almost instantaneously into one of the ordinary configurations we encounter in everyday life. This phenomenon is called **decoherence**, and although it is still not entirely understood it appears to be the fundamental mechanism by which quantum mechanics reduces to classical mechanics in the macroscopic realm.³⁰

• • • • • •

In this book I have tried to tell a consistent and coherent story: The wave function (Ψ) represents the state of a particle (or system); particles do not in general possess specific dynamical properties (position, momentum, energy, angular momentum, etc.) until an act of measurement intervenes; the probability of getting a particular value in any given experiment is determined by the statistical interpretation of Ψ ; upon measurement the wave function collapses, so that an immediately repeated measurement is certain to yield the same result. There are other possible interpretations—nonlocal hidden variable theories, the “many worlds” picture, “consistent histories,” ensemble models, and others—but I believe this one is conceptually the *simpler*, and certainly it is the one shared by most physicists today.³¹ It has stood the test of time, and emerged unscathed from every experimental challenge. But I cannot believe this is the end of the story; at the very least, we have much to learn about the nature of measurement and the mechanism of collapse. And it is entirely possible that future generations will look back, from the vantage point of a more sophisticated theory, and wonder how we could have been so gullible.

¹ This may be *strange*, but it is not *mystical*, as some popularizers would like to suggest. The so-called **wave–particle duality**, which Niels Bohr elevated to the status of a cosmic principle (**complementarity**), makes electrons sound like unpredictable adolescents, who sometimes behave like adults, and sometimes, for no particular reason, like children. I prefer to avoid such language. When I say that a particle does not have a particular attribute before its measurement, I have in mind, for example, an electron in the spin state $\chi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$; a measurement of the x -component of its angular momentum could return the value $\hbar/2$, or (with equal probability) the value $-\hbar/2$, but until the measurement is made it simply *does not have* a well-defined value of S_x .

² A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).

³ Bohr wrote a famous rebuttal to the EPR paradox (*Phys. Rev.* **48**, 696 (1935)). I doubt many people read it, and certainly very few understood it (Bohr himself later admitted that he had trouble making sense of his own argument), but it was a relief that the great man had solved the problem, and everybody else could go back to business. It was not until the mid-1960s that most physicists began to worry seriously about the EPR paradox.

⁴ Although the term “entanglement” is usually applied to systems of two (or more) particles, the same basic notion can be extended to *single* particle states (Problem 12.2 is an example). For an interesting discussion see D. V. Schroeder, *Am. J. Phys.* **85**, 812 (2017).

⁵ See T. Norsen, *Am. J. Phys.* **73**, 164 (2005).

⁶ The partition is inserted rapidly; if it is done adiabatically the particle may be forced into the (however slightly) larger of the two, as you found in Problem 11.34.

⁷ The hidden variable could be a single number, or it could be a whole *collection* of numbers; perhaps λ is to be calculated in some future theory, or maybe it is for some reason of principle incalculable. It hardly matters. All I am asserting is that there must be *something*—if only a *list* of the outcomes of every possible experiment—associated with the system prior to a measurement.

⁸ D. Bohm, *Phys. Rev.* **85**, 166, 180 (1952).

⁹ Bell’s original paper (*Physics* **1**, 195 (1964), reprinted as Chapter 2 in John S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, UK (1987)) is a gem: brief, accessible, and beautifully written.

¹⁰ This already concedes far more than a *classical* determinist would be prepared to allow, for it abandons any notion that the particles could have well-defined angular momentum vectors with simultaneously determinate components. The point of Bell’s argument is to demonstrate that quantum mechanics is incompatible with *any* local deterministic theory—even one that bends over backwards to be accommodating. Of course, if you reject Equation 12.5, then the theory is *manifestly* incompatible with quantum mechanics.

- ¹¹ It is an embarrassing historical fact that Bell's theorem, which is now universally recognized as one of the most profound discoveries of the twentieth century, was barely noticed at the time, with the exception of an inspired fringe element. For a fascinating account, see David Kaiser, *How the Hippies Saved Physics*, W. W. Norton, New York, 2011.
- ¹² A. Aspect, P. Grangier, and G. Roger, *Phys. Rev. Lett.* **49**, 91 (1982). There were logically possible (if implausible) loopholes in the Aspect experiment, which were gradually closed over the ensuing years; see J. Handsteiner et al., *Phys. Rev. Lett.* **118**, 060401 (2017). It is now possible to test Bell's inequality in the undergraduate laboratory: D. Dehlinger and M. W. Mitchell, *Am. J. Phys.* **70**, 903 (2002).
- ¹³ Bell's theorem involves *averages* and it is conceivable that an apparatus such as Aspect's contains some secret bias which selects out a nonrepresentative sample, thus distorting the average. In 1989, an improved version of Bell's theorem was proposed, in which the contrast between the quantum prediction and that of any local hidden variable theory is even more dramatic. See D. Greenberger, M. Horne, A. Shimony, and A. Zeilinger, *Am. J. Phys.* **58**, 1131 (1990) and N. D. Mermin, *Am. J. Phys.* **58**, 731, (1990). An experiment of this kind suitable for an undergraduate laboratory has been carried out by Mark Beck and his students: *Am. J. Phys.* **74**, 180 (2006).
- ¹⁴ It is a curious twist of fate that the EPR paradox, which *assumed* locality in order to *prove* realism, led finally to the demise of locality and left the issue of realism undecided—the outcome (as Bell put it) Einstein would have liked *least*. Most physicists today consider that if they can't have *local* realism, there's not much point in realism at *all*, and for this reason nonlocal hidden variable theories occupy a rather peripheral niche. Still, some authors—notably Bell himself, in *Speakable and Unspeakable in Quantum Mechanics* (footnote 9 in this chapter)—argue that such theories offer the best hope of bridging the conceptual gap between the measured system and the measuring apparatus, and for supplying an intelligible mechanism for the collapse of the wave function.
- ¹⁵ An enormous amount has been written about Bell's theorem. My favorite is an inspired essay by David Mermin in *Physics Today* (April 1985, page 38). An extensive bibliography will be found in L. E. Ballentine, *Am. J. Phys.* **55**, 785 (1987).
- ¹⁶ This problem is based on George Greenstein and Arthur G. Zajonc, *The Quantum Challenge*, 2nd edn., Jones and Bartlett, Sudbury, MA (2006), Section 5.3.
- ¹⁷ It's actually the “projection operator” onto the state $|\Psi\rangle$ —see Equation 3.91.
- ¹⁸ I'm not talking about any fancy quantum phenomenon (Heisenberg uncertainty or Born indeterminacy, which would apply even if we knew the precise state); I'm talking here about good old-fashioned *ignorance*.
- ¹⁹ Do not confuse a *linear combination* of two pure states, which itself is still a pure state (the sum of two vectors in Hilbert space is another vector in Hilbert space) with a *mixed state*, which is not represented by *any* (single) vector in the Hilbert space.
- ²⁰ Starting around 1975, members of the so-called “Fundamental Fysisks Group” proposed a series of increasingly ingenious schemes for faster-than-light communication—inspiring in turn a succession of increasingly sophisticated rebuttals, culminating in the no-clone theorem, which finally put a stop to the whole misguided enterprise. For a fascinating account, see Chapter 11 of Kaiser's *How the Hippies Saved Physics* (footnote 11, page 451).
- ²¹ W. K. Wootters and W. H. Zurek, *Nature* **299**, 802 (1982); D. Dieks, *Phys. Lett. A* **92**, 271 (1982).
- ²² This assumes that the device acts *linearly* on the state $|\Psi\rangle$, as it must, since the time-dependent Schrödinger equation (which presumably governs the process) is linear.
- ²³ The no-clone theorem is one of the foundations for **quantum information theory**, “**teleportation**,” and **quantum computation**. For a brief history and a comprehensive bibliography, see F. W. Strauch, *Am. J. Phys.* **84**, 495 (2016).
- ²⁴ For a brief summary, see W. K. Wootters and W. H. Zurek, *Physics Today*, February 2009, page 76.
- ²⁵ Electrons would do, but traditionally the story is told using photons. By the way, there is no entanglement involved, and they're not in a hurry—this has nothing to do with EPR or superluminal signals.
- ²⁶ If Alice and Bob were foolish enough to use just *two* orthogonal photon states (say, $|H\rangle$ and $|V\rangle$), then Eve might get lucky, and use a quantum Xerox machine that does faithfully clone those two states. But as long as they include nontrivial linear combinations (such as $|R\rangle$ and $|L\rangle$), the cloning is certain to fail, and the eavesdropping will be detected.
- ²⁷ There is a school of thought that rejects this distinction, holding that the system and the measuring apparatus should be described by one great big wave function which itself evolves according to the Schrödinger equation. In such theories there is no collapse of the wave function, but one must typically abandon any hope of describing individual events—quantum mechanics (in this view) applies only to *ensembles* of identically prepared systems. See, for example, Philip Pearle, *Am. J. Phys.* **35**, 742 (1967), or Leslie E. Ballentine, *Quantum Mechanics: A Modern Development*, 2nd edn, World Scientific, Singapore (1998).
- ²⁸ E. Schrödinger, *Naturwiss.* **48**, 52 (1935); translation by Josef M. Jauch, *Foundations of Quantum Mechanics*, Addison-Wesley, Reading, MA (1968), page 185.
- ²⁹ This is true even if you put it in an almost complete vacuum, cool it down practically to absolute zero, and somehow shield out the cosmic background radiation. It is possible to imagine a single electron avoiding all contact for a significant time, but not a macroscopic object.
- ³⁰ See, for example, M. Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, Springer, (2007), or W. H. Zurek, *Physics Today*, October, 2014, page 44.
- ³¹ See Daniel Styer et al., *Am. J. Phys.* **70**, 288 (2002).

Appendix Linear Algebra

◆

Linear algebra abstracts and generalizes the arithmetic of ordinary vectors, as we encounter them in first-year physics. The generalization is in two directions: (1) we allow the scalars to be *complex* numbers, and (2) we do not restrict ourselves to three dimensions.

A.1 Vectors

A **vector space** consists of a set of **vectors** ($|\alpha\rangle, |\beta\rangle, |\gamma\rangle, \dots$), together with a set of **scalars** (a, b, c, \dots),¹ which is **closed**² under two operations: vector addition and scalar multiplication.

- **Vector Addition**

The “sum” of any two vectors is another vector:

$$|\alpha\rangle + |\beta\rangle = |\gamma\rangle. \quad (\text{A.1})$$

Vector addition is **commutative**:

$$|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle, \quad (\text{A.2})$$

and **associative**:

$$|\alpha\rangle + (|\beta\rangle + |\gamma\rangle) = (|\alpha\rangle + |\beta\rangle) + |\gamma\rangle. \quad (\text{A.3})$$

There exists a **zero** (or **null**) **vector**,³ $|0\rangle$, with the property that

$$|\alpha\rangle + |0\rangle = |\alpha\rangle, \quad (\text{A.4})$$

for every vector $|\alpha\rangle$. And for every vector $|\alpha\rangle$ there is an associated **inverse vector** ($|- \alpha\rangle$),⁴ such that

$$|\alpha\rangle + |- \alpha\rangle = |0\rangle. \quad (\text{A.5})$$

- **Scalar Multiplication**

The “product” of any scalar with any vector is another vector:

$$a|\alpha\rangle = |\gamma\rangle. \quad (\text{A.6})$$

Scalar multiplication is **distributive** with respect to vector addition:

$$a(|\alpha\rangle + |\beta\rangle) = a|\alpha\rangle + a|\beta\rangle, \quad (\text{A.7})$$

and with respect to scalar addition:

$$(a + b)|\alpha\rangle = a|\alpha\rangle + b|\alpha\rangle. \quad (\text{A.8})$$

It is also **associative** with respect to the ordinary multiplication of scalars:

$$a(b|\alpha\rangle) = (ab)|\alpha\rangle. \quad (\text{A.9})$$

Multiplication by the scalars 0 and 1 has the effect you would expect:

$$0|\alpha\rangle = |0\rangle; \quad 1|\alpha\rangle = |\alpha\rangle. \quad (\text{A.10})$$

Evidently $|- \alpha\rangle = (-1)|\alpha\rangle$ (which we write more simply as $-|\alpha\rangle$).

There's a lot less here than meets the eye—all I have done is to write down in abstract language the familiar rules for manipulating vectors. The virtue of such abstraction is that we will be able to apply our knowledge and intuition about the behavior of ordinary vectors to other systems that happen to share the same formal properties.

A **linear combination** of the vectors $|\alpha\rangle, |\beta\rangle, |\gamma\rangle, \dots$, is an expression of the form

$$a|\alpha\rangle + b|\beta\rangle + c|\gamma\rangle + \dots \quad (\text{A.11})$$

A vector $|\lambda\rangle$ is said to be **linearly independent** of the set $|\alpha\rangle, |\beta\rangle, |\gamma\rangle, \dots$, if it cannot be written as a linear combination of them. (For example, in three dimensions the unit vector \hat{k} is linearly independent of \hat{i} and \hat{j} , but any vector in the xy plane is linearly *dependent* on \hat{i} and \hat{j} .) By extension, a *set* of vectors is “linearly independent” if each one is linearly independent of all the rest. A collection of vectors is said to **span** the space if *every* vector can be written as a linear combination of the members of this set.⁵ A set of *linearly independent* vectors that spans the space is called a **basis**. The number of vectors in any basis is called the **dimension** of the space. For the moment we shall assume that the dimension (n) is *finite*.

With respect to a prescribed basis

$$|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle, \quad (\text{A.12})$$

any given vector

$$|\alpha\rangle = a_1|e_1\rangle + a_2|e_2\rangle + \dots + a_n|e_n\rangle, \quad (\text{A.13})$$

is uniquely represented by the (ordered) n -tuple of its **components**:

$$|\alpha\rangle \leftrightarrow (a_1, a_2, \dots, a_n). \quad (\text{A.14})$$

It is often easier to work with the components than with the abstract vectors themselves. To add vectors, you add their corresponding components:

$$|\alpha\rangle + |\beta\rangle \leftrightarrow (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n); \quad (\text{A.15})$$

to multiply by a scalar you multiply each component:

$$c|\alpha\rangle \leftrightarrow (ca_1, ca_2, \dots, ca_n); \quad (\text{A.16})$$

the null vector is represented by a string of zeroes:

$$|0\rangle \leftrightarrow (0, 0, \dots, 0); \quad (\text{A.17})$$

and the components of the inverse vector have their signs reversed:

$$|-\alpha\rangle \leftrightarrow (-a_1, -a_2, \dots, -a_n). \quad (\text{A.18})$$

The only *disadvantage* of working with components is that you have to commit yourself to a particular basis, and the same manipulations will look very different to someone using a different basis.

Problem A.1 Consider the ordinary vectors in three dimensions $(a_x\hat{i} + a_y\hat{j} + a_z\hat{k})$, with complex components.

- (a) Does the subset of all vectors with $a_z = 0$ constitute a vector space? If so, what is its dimension; if not, why not?
- (b) What about the subset of all vectors whose z component is 1? *Hint:* Would the sum of two such vectors be in the subset? How about the null vector?
- (c) What about the subset of vectors whose components are all equal?

- * **Problem A.2** Consider the collection of all polynomials (with complex coefficients) of degree $< N$ in x .
- (a) Does this set constitute a vector space (with the polynomials as “vectors”)? If so, suggest a convenient basis, and give the dimension of the space. If not, which of the defining properties does it lack?
 - (b) What if we require that the polynomials be *even* functions?
 - (c) What if we require that the leading coefficient (i.e. the number multiplying x^{N-1}) be 1?
 - (d) What if we require that the polynomials have the value 0 at $x = 1$?
 - (e) What if we require that the polynomials have the value 1 at $x = 0$?

Problem A.3 Prove that the components of a vector with respect to a given basis are *unique*.

A.2 Inner Products

In three dimensions we encounter two kinds of vector products: the dot product and the cross product. The latter does not generalize in any natural way to n -dimensional vector spaces, but the former *does*—in this context it is usually called the **inner product**. The inner product of vectors $|\alpha\rangle$ and $|\beta\rangle$ is a complex number (which we write as $\langle\alpha|\beta\rangle$), with the following properties:

$$\langle\beta|\alpha\rangle = \langle\alpha|\beta\rangle^*, \quad (\text{A.19})$$

$$\langle\alpha|\alpha\rangle \geq 0, \quad \text{and} \quad \langle\alpha|\alpha\rangle = 0 \Leftrightarrow |\alpha\rangle = |0\rangle, \quad (\text{A.20})$$

$$\langle\alpha| (b|\beta\rangle + c|\gamma\rangle) = b\langle\alpha|\beta\rangle + c\langle\alpha|\gamma\rangle. \quad (\text{A.21})$$

Apart from the generalization to complex numbers, these axioms simply codify the familiar behavior of dot products. A vector space with an inner product is called an **inner product space**.

Because the inner product of any vector with itself is a non-negative number (Equation A.20), its square root is *real*—we call this the **norm** of the vector:

$$\|\alpha\| \equiv \sqrt{\langle\alpha|\alpha\rangle}; \quad (\text{A.22})$$

it generalizes the notion of “length.” A **unit vector** (one whose norm is 1) is said to be **normalized** (the word should really be “normal,” but I guess that sounds too judgmental). Two vectors whose inner product is zero are called **orthogonal** (generalizing the notion of “perpendicular”). A collection of mutually orthogonal normalized vectors,

$$\langle\alpha_i|\alpha_j\rangle = \delta_{ij}, \quad (\text{A.23})$$

is called an **orthonormal set**. It is always possible (see Problem A.4), and almost always convenient, to choose an **orthonormal basis**; in that case the inner product of two vectors can be written very neatly in terms of their components:

$$\langle\alpha|\beta\rangle = a_1^*b_1 + a_2^*b_2 + \cdots + a_n^*b_n, \quad (\text{A.24})$$

the norm (squared) becomes

$$\langle\alpha|\alpha\rangle = |a_1|^2 + |a_2|^2 + \cdots + |a_n|^2, \quad (\text{A.25})$$

and the components themselves are

$$a_i = \langle e_i|\alpha\rangle. \quad (\text{A.26})$$

(These results generalize the familiar formulas $\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z$, $|\mathbf{a}|^2 = a_x^2 + a_y^2 + a_z^2$, and $a_x = \hat{i} \cdot \mathbf{a}$, $a_y = \hat{j} \cdot \mathbf{a}$, $a_z = \hat{k} \cdot \mathbf{a}$, for the three-dimensional orthonormal basis \hat{i} , \hat{j} , \hat{k} .) From now on we shall *always* work in orthonormal bases, unless it is explicitly indicated otherwise.

Another geometrical quantity one might wish to generalize is the *angle* between two vectors. In ordinary vector analysis $\cos\theta = (\mathbf{a} \cdot \mathbf{b}) / |\mathbf{a}| |\mathbf{b}|$, but because the inner product is in general a complex number, the analogous formula (in an arbitrary inner product space) does not define a (real) angle θ . Nevertheless, it is still true that the *absolute value* of this quantity is a number no greater than 1,

$$|\langle \alpha | \beta \rangle|^2 \leq \langle \alpha | \alpha \rangle \langle \beta | \beta \rangle. \quad (\text{A.27})$$

This important result is known as the **Schwarz inequality**; the proof is given in Problem A.5. So you can, if you like, define the angle between $|\alpha\rangle$ and $|\beta\rangle$ by the formula

$$\cos \theta = \sqrt{\frac{\langle \alpha | \beta \rangle \langle \beta | \alpha \rangle}{\langle \alpha | \alpha \rangle \langle \beta | \beta \rangle}}. \quad (\text{A.28})$$

- * **Problem A.4** Suppose you start out with a basis $(|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle)$ that is *not* orthonormal. The **Gram–Schmidt procedure** is a systematic ritual for generating from it an orthonormal basis $(|e'_1\rangle, |e'_2\rangle, \dots, |e'_n\rangle)$. It goes like this:

(i) Normalize the first basis vector (divide by its norm):

$$|e'_1\rangle = \frac{|e_1\rangle}{\|e_1\|}.$$

(ii) Find the projection of the second vector along the first, and subtract it off:

$$|e_2\rangle - \langle e'_1 | e_2 \rangle |e'_1\rangle.$$

This vector is orthogonal to $|e'_1\rangle$; normalize it to get $|e'_2\rangle$.

(iii) Subtract from $|e_3\rangle$ its projections along $|e'_1\rangle$ and $|e'_2\rangle$:

$$|e_3\rangle - \langle e'_1 | e_3 \rangle |e'_1\rangle - \langle e'_2 | e_3 \rangle |e'_2\rangle.$$

This is orthogonal to $|e'_1\rangle$ and $|e'_2\rangle$; normalize it to get $|e'_3\rangle$. And so on.

Use the Gram–Schmidt procedure to orthonormalize the 3-space basis $|e_1\rangle = (1+i)\hat{i} + (1)\hat{j} + (i)\hat{k}$, $|e_2\rangle = (i)\hat{i} + (3)\hat{j} + (1)\hat{k}$, $|e_3\rangle = (0)\hat{i} + (28)\hat{j} + (0)\hat{k}$.

Problem A.5 Prove the Schwarz inequality (Equation A.27). *Hint:* Let $|\gamma\rangle = |\beta\rangle - (\langle \alpha | \beta \rangle / \langle \alpha | \alpha \rangle) |\alpha\rangle$, and use $\langle \gamma | \gamma \rangle \geq 0$.

Problem A.6 Find the angle (in the sense of Equation A.28) between the vectors $|\alpha\rangle = (1+i)\hat{i} + (1)\hat{j} + (i)\hat{k}$ and $|\beta\rangle = (4-i)\hat{i} + (0)\hat{j} + (2-2i)\hat{k}$.

Problem A.7 Prove the triangle inequality: $\|(|\alpha\rangle + |\beta\rangle)\| \leq \|\alpha\| + \|\beta\|$.

A.3 Matrices

Suppose you take every vector (in 3-space) and multiply it by 17, or you rotate every vector by 39° about the z axis, or you reflect every vector in the xy plane—these are all examples of **linear transformations**. A linear transformation⁶ (\hat{T}) takes each vector in a vector space and “transforms” it into some other vector ($|\alpha\rangle \rightarrow |\alpha'\rangle = \hat{T}|\alpha\rangle$), subject to the condition that the operation be *linear*:

$$\hat{T}(a|\alpha\rangle + b|\beta\rangle) = a(\hat{T}|\alpha\rangle) + b(\hat{T}|\beta\rangle), \quad (\text{A.29})$$

for any vectors $|\alpha\rangle, |\beta\rangle$ and any scalars a, b .

If you know what a particular linear transformation does to a set of *basis* vectors, you can easily figure out what it does to *any* vector. For suppose that

$$\begin{aligned}\hat{T}|e_1\rangle &= T_{11}|e_1\rangle + T_{21}|e_2\rangle + \cdots + T_{n1}|e_n\rangle, \\ \hat{T}|e_2\rangle &= T_{12}|e_1\rangle + T_{22}|e_2\rangle + \cdots + T_{n2}|e_n\rangle, \\ &\dots \\ \hat{T}|e_n\rangle &= T_{1n}|e_1\rangle + T_{2n}|e_2\rangle + \cdots + T_{nn}|e_n\rangle,\end{aligned}$$

or, more compactly,

$$\hat{T}|e_j\rangle = \sum_{i=1}^n T_{ij}|e_i\rangle, \quad (j = 1, 2, \dots, n). \quad (\text{A.30})$$

If $|\alpha\rangle$ is an arbitrary vector,

$$|\alpha\rangle = a_1|e_1\rangle + a_2|e_2\rangle + \cdots + a_n|e_n\rangle = \sum_{j=1}^n a_j|e_j\rangle, \quad (\text{A.31})$$

then

$$\hat{T}|\alpha\rangle = \sum_{j=1}^n a_j(\hat{T}|e_j\rangle) = \sum_{j=1}^n \sum_{i=1}^n a_j T_{ij}|e_i\rangle = \sum_{i=1}^n \left(\sum_{j=1}^n T_{ij} a_j \right) |e_i\rangle. \quad (\text{A.32})$$

Evidently \hat{T} takes a vector with components a_1, a_2, \dots, a_n into a vector with components⁷

$$a'_i = \sum_{j=1}^n T_{ij} a_j. \quad (\text{A.33})$$

Thus the n^2 elements T_{ij} uniquely characterize the linear transformation \hat{T} (with respect to a given basis), just as the n components a_i uniquely characterize the vector $|\alpha\rangle$ (with respect to that basis):

$$\hat{T} \leftrightarrow (T_{11}, T_{12}, \dots, T_{nn}). \quad (\text{A.34})$$

If the basis is orthonormal, it follows from Equation A.30 that

$$T_{ij} = \langle e_i | \hat{T} | e_j \rangle. \quad (\text{A.35})$$

It is convenient to display these complex numbers in the form of a matrix:⁸

$$\mathbf{T} = \begin{pmatrix} T_{11} & T_{12} & \dots & T_{1n} \\ T_{21} & T_{22} & \dots & T_{2n} \\ \vdots & \vdots & & \vdots \\ T_{n1} & T_{n2} & \dots & T_{nn} \end{pmatrix}. \quad (\text{A.36})$$

The study of linear transformations reduces, then, to the theory of matrices. The *sum* of two linear transformations $(\hat{S} + \hat{T})$ is defined in the natural way:

$$(\hat{S} + \hat{T})|\alpha\rangle = \hat{S}|\alpha\rangle + \hat{T}|\alpha\rangle; \quad (\text{A.37})$$

this matches the usual rule for adding matrices (you add the corresponding elements):

$$\mathbf{U} = \mathbf{S} + \mathbf{T} \Leftrightarrow U_{ij} = S_{ij} + T_{ij}. \quad (\text{A.38})$$

The *product* of two linear transformations $(\hat{S}\hat{T})$ is the net effect of performing them in succession—first \hat{T} , then \hat{S} :

$$|\alpha'\rangle = \hat{T}|\alpha\rangle; \quad |\alpha''\rangle = \hat{S}|\alpha'\rangle = \hat{S}(\hat{T}|\alpha\rangle) = \hat{S}\hat{T}|\alpha\rangle. \quad (\text{A.39})$$

What matrix \mathbf{U} represents the combined transformation $\hat{U} = \hat{S}\hat{T}$? It's not hard to work it out:

$$a''_i = \sum_{j=1}^n S_{ij} a'_j = \sum_{j=1}^n S_{ij} \left(\sum_{k=1}^n T_{jk} a_k \right) = \sum_{k=1}^n \left(\sum_{j=1}^n S_{ij} T_{jk} \right) a_k = \sum_{k=1}^n U_{ik} a_k.$$

Evidently

$$\mathbf{U} = \mathbf{ST} \Leftrightarrow U_{ik} = \sum_{j=1}^n S_{ij} T_{jk} \quad (\text{A.40})$$

—this is the standard rule for matrix multiplication: to find the ik th element of the product, you look at the i th row of \mathbf{S} , and the k th column of \mathbf{T} , multiply corresponding entries, and add. The same prescription allows you to multiply *rectangular* matrices, as long as the number of columns in the first matches the number of rows in the second. In particular, if we write the n -tuple of components of $|\alpha\rangle$ as an $n \times 1$ **column matrix** (or “column vector”):⁹

$$\mathbf{a} \equiv \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad (\text{A.41})$$

the transformation rule (Equation A.33) can be expressed as a matrix product:

$$\mathbf{a}' = \mathbf{T}\mathbf{a}. \quad (\text{A.42})$$

Now some matrix terminology:

- The **transpose** of a matrix (which we shall write with a tilde: $\tilde{\mathbf{T}}$) is the same set of elements, but with

rows and columns interchanged. In particular, the transpose of a *column* matrix is a **row matrix**:

$$\tilde{\mathbf{a}} = (a_1 \quad a_2 \quad \dots \quad a_n). \quad (\text{A.43})$$

For a *square* matrix taking the transpose amounts to reflecting in the **main diagonal** (upper left to lower right):

$$\tilde{\mathbf{T}} = \begin{pmatrix} T_{11} & T_{21} & \dots & T_{n1} \\ T_{12} & T_{22} & \dots & T_{n2} \\ \vdots & \vdots & & \vdots \\ T_{1n} & T_{2n} & \dots & T_{nn} \end{pmatrix}. \quad (\text{A.44})$$

A (square) matrix is **symmetric** if it is equal to its transpose; it is **antisymmetric** if this operation reverses the sign:

$$\text{symmetric : } \tilde{\mathbf{T}} = \mathbf{T}; \quad \text{antisymmetric : } \tilde{\mathbf{T}} = -\mathbf{T}. \quad (\text{A.45})$$

- The (complex) **conjugate** of a matrix (which we denote, as usual, with an asterisk, \mathbf{T}^*), consists of the complex conjugate of every element:

$$\mathbf{T}^* = \begin{pmatrix} T_{11}^* & T_{12}^* & \dots & T_{1n}^* \\ T_{21}^* & T_{22}^* & \dots & T_{2n}^* \\ \vdots & \vdots & & \vdots \\ T_{n1}^* & T_{n2}^* & \dots & T_{nn}^* \end{pmatrix}; \quad \mathbf{a}^* = \begin{pmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_n^* \end{pmatrix}. \quad (\text{A.46})$$

A matrix is **real** if all its elements are real, and **imaginary** if they are all imaginary:

$$\text{real : } \mathbf{T}^* = \mathbf{T}; \quad \text{imaginary : } \mathbf{T}^* = -\mathbf{T}. \quad (\text{A.47})$$

- The **hermitian conjugate** (or **adjoint**) of a matrix (indicated by a dagger, \mathbf{T}^\dagger) is the transpose conjugate:

$$\mathbf{T}^\dagger \equiv \tilde{\mathbf{T}}^* = \begin{pmatrix} T_{11}^* & T_{21}^* & \dots & T_{n1}^* \\ T_{12}^* & T_{22}^* & \dots & T_{n2}^* \\ \vdots & \vdots & & \vdots \\ T_{1n}^* & T_{2n}^* & \dots & T_{nn}^* \end{pmatrix}; \quad \mathbf{a}^\dagger \equiv \tilde{\mathbf{a}}^* = (a_1^* \quad a_2^* \quad \dots \quad a_n^*). \quad (\text{A.48})$$

A square matrix is **hermitian** (or **self-adjoint**) if it is equal to its hermitian conjugate; if hermitian conjugation introduces a minus sign, the matrix is **skew hermitian** (or **anti-hermitian**):

$$\text{hermitian : } \mathbf{T}^\dagger = \mathbf{T}; \quad \text{skew hermitian : } \mathbf{T}^\dagger = -\mathbf{T}. \quad (\text{A.49})$$

In this notation the inner product of two vectors (with respect to an orthonormal basis—Equation A.24), can be written very neatly as a matrix product:

$$\langle \alpha | \beta \rangle = \mathbf{a}^\dagger \mathbf{b}. \quad (\text{A.50})$$

Notice that each of the three operations defined in this paragraph, if applied twice, returns you to the original matrix.

Matrix multiplication is not, in general, commutative ($ST \neq TS$); the *difference* between the two orderings is called the **commutator**:¹⁰

$$[S, T] \equiv ST - TS. \quad (\text{A.51})$$

The transpose of a product is the product of the transposes *in reverse order*:

$$(ST)^\dagger = T^\dagger S^\dagger, \quad (\text{A.52})$$

(see Problem A.11), and the same goes for hermitian conjugates:

$$(ST)^\dagger = T^\dagger S^\dagger. \quad (\text{A.53})$$

The **identity matrix** (representing a linear transformation that carries every vector into itself) consists of ones on the main diagonal, and zeroes everywhere else:

$$I \equiv \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}. \quad (\text{A.54})$$

In other words,

$$I_{ij} = \delta_{ij}. \quad (\text{A.55})$$

The **inverse** of a (square) matrix (written T^{-1}) is defined in the obvious way:¹¹

$$T^{-1}T = TT^{-1} = I. \quad (\text{A.56})$$

A matrix has an inverse if and only if its **determinant**¹² is nonzero; in fact,

$$T^{-1} = \frac{1}{\det T} \tilde{C}, \quad (\text{A.57})$$

where C is the matrix of **cofactors** (the cofactor of element T_{ij} is $(-1)^{i+j}$ times the determinant of the submatrix obtained from T by erasing the i th row and the j th column). A matrix that has no inverse is said to be **singular**. The inverse of a product (assuming it exists) is the product of the inverses *in reverse order*:

$$(ST)^{-1} = T^{-1}S^{-1}. \quad (\text{A.58})$$

A matrix is **unitary** if its inverse is equal to its hermitian conjugate:¹³

$$\text{unitary : } U^\dagger = U^{-1}. \quad (\text{A.59})$$

Assuming the basis is orthonormal, the columns of a unitary matrix constitute an orthonormal set, and so too do its rows (see Problem A.12). Linear transformations represented by unitary matrices preserve inner products, since (Equation A.50)

$$\langle \alpha' | \beta' \rangle = a'^\dagger b' = (Ua)^\dagger (Ub) = a^\dagger U^\dagger Ub = a^\dagger b = \langle \alpha | \beta \rangle. \quad (\text{A.60})$$

*

Problem A.8 Given the following two matrices:

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & i \\ 2 & 0 & 3 \\ 2i & -2i & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 0 & -i \\ 0 & 1 & 0 \\ i & 3 & 2 \end{pmatrix},$$

compute: (a) $\mathbf{A} + \mathbf{B}$, (b) $\mathbf{A} \mathbf{B}$, (c) $[\mathbf{A}, \mathbf{B}]$, (d) $\tilde{\mathbf{A}}$, (e) \mathbf{A}^* , (f) \mathbf{A}^\dagger , (g) $\det(\mathbf{B})$, and (h) \mathbf{B}^{-1} . Check that $\mathbf{B} \mathbf{B}^{-1} = \mathbf{1}$. Does \mathbf{A} have an inverse?

- * **Problem A.9** Using the square matrices in Problem [A.8](#), and the column matrices

$$\mathbf{a} = \begin{pmatrix} i \\ 2i \\ 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ (1-i) \\ 0 \end{pmatrix},$$

find: (a) $\mathbf{A}\mathbf{a}$, (b) $\mathbf{a}^\dagger\mathbf{b}$, (c) $\tilde{\mathbf{a}}\mathbf{B}\mathbf{b}$, (d) $\mathbf{a}\mathbf{b}^\dagger$.

Problem A.10 By explicit construction of the matrices in question, show that any matrix \mathbf{T} can be written

- (a) as the sum of a symmetric matrix \mathbf{S} and an antisymmetric matrix \mathbf{A} ;
- (b) as the sum of a real matrix \mathbf{R} and an imaginary matrix \mathbf{M} ;
- (c) as the sum of a hermitian matrix \mathbf{H} and a skew-hermitian matrix \mathbf{K} .

- * **Problem A.11** Prove Equations [A.52](#), [A.53](#), and [A.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 necessarily unitary? Is the sum of two hermitian matrices always hermitian?

Problem A.12 Show that the rows and columns of a unitary matrix constitute orthonormal sets.

Problem A.13 Noting that $\det(\tilde{\mathbf{T}}) = \det(\mathbf{T})$, show that the determinant of a hermitian matrix is real, the determinant of a unitary matrix has modulus 1 (hence the name), and the determinant of an orthogonal matrix (footnote 13) is either $+1$ or -1 .

A.4 Changing Bases

The components of a vector depend, of course, on your (arbitrary) choice of basis, and so do the elements of the matrix representing a linear transformation. We might inquire how these numbers change when we switch to a different basis. The old basis vectors, $|e_i\rangle$ are—like *all* vectors—linear combinations of the new ones, $|f_i\rangle$:

$$|e_1\rangle = S_{11}|f_1\rangle + S_{21}|f_2\rangle + \cdots + S_{n1}|f_n\rangle,$$

$$|e_2\rangle = S_{12}|f_1\rangle + S_{22}|f_2\rangle + \cdots + S_{n2}|f_n\rangle,$$

...

$$|e_n\rangle = S_{1n}|f_1\rangle + S_{2n}|f_2\rangle + \cdots + S_{nn}|f_n\rangle,$$

(for some set of complex numbers S_{ij}), or, more compactly,

$$|e_j\rangle = \sum_{i=1}^n S_{ij}|f_i\rangle, \quad (j = 1, 2, \dots, n). \quad (\text{A.61})$$

This is *itself* a linear transformation (compare Equation A.30),¹⁴ and we know immediately how the components transform:

$$a_i^f = \sum_{j=1}^n S_{ij}a_j^e, \quad (\text{A.62})$$

(where the superscript indicates the basis). In matrix form

$$\mathbf{a}^f = \mathbf{S}\mathbf{a}^e. \quad (\text{A.63})$$

What about the matrix representing a linear transformation \hat{T} —how is *it* modified by a change of basis? Well, in the old basis we had (Equation A.42)

$$\mathbf{a}'^e = \mathbf{T}^e\mathbf{a}^e,$$

and Equation A.63—multiplying both sides by \mathbf{S}^{-1} —entails¹⁵ $\mathbf{a}^e = \mathbf{S}^{-1}\mathbf{a}^f$, so

$$\mathbf{a}'^f = \mathbf{S}\mathbf{a}'^e = \mathbf{S}(\mathbf{T}^e\mathbf{a}^e) = \mathbf{S}\mathbf{T}^e\mathbf{S}^{-1}\mathbf{a}^f.$$

Evidently

$$\mathbf{T}^f = \mathbf{S}\mathbf{T}^e\mathbf{S}^{-1}. \quad (\text{A.64})$$

In general, two matrices (\mathbf{T}_1 and \mathbf{T}_2) are said to be *similar* if $\mathbf{T}_2 = \mathbf{S}\mathbf{T}_1\mathbf{S}^{-1}$ for some (nonsingular) matrix \mathbf{S} . What we have just found is that *matrices representing the same linear transformation, with respect to different bases, are similar*. Incidentally, if the first basis is orthonormal, the second will also be orthonormal if and only if the matrix \mathbf{S} is *unitary* (see Problem A.16). Since we always work in orthonormal bases, we are interested mainly in *unitary* similarity transformations.

While the *elements* of the matrix representing a given linear transformation may look very different in the new basis, *two* numbers associated with the matrix are unchanged: the determinant and the **trace**. For the determinant of a product is the product of the determinants, and hence

$$\det(\mathbf{T}^f) = \det(\mathbf{S}\mathbf{T}^e\mathbf{S}^{-1}) = \det(\mathbf{S})\det(\mathbf{T}^e)\det(\mathbf{S}^{-1}) = \det(\mathbf{T}^e). \quad (\text{A.65})$$

And the trace, which is the *sum of the diagonal elements*,

$$\text{Tr}(\mathbf{T}) \equiv \sum_{i=1}^m T_{ii}, \quad (\text{A.66})$$

has the property (see Problem A.17) that

$$\text{Tr}(\mathbf{T}_1\mathbf{T}_2) = \text{Tr}(\mathbf{T}_2\mathbf{T}_1), \quad (\text{A.67})$$

(for any two matrices \mathbf{T}_1 and \mathbf{T}_2), so

$$\text{Tr}(\mathbf{T}^f) = \text{Tr}(\mathbf{S}\mathbf{T}^e\mathbf{S}^{-1}) = \text{Tr}(\mathbf{T}^e\mathbf{S}^{-1}\mathbf{S}) = \text{Tr}(\mathbf{T}^e). \quad (\text{A.68})$$

Problem A.14 Using the standard basis $(\hat{i}, \hat{j}, \hat{k})$ for vectors in three dimensions:

- (a) Construct the matrix representing a rotation through angle θ (counterclockwise, looking down the axis toward the origin) about the z axis.
- (b) Construct the matrix representing a rotation by 120° (counterclockwise, looking down the axis) about an axis through the point $(1,1,1)$.
- (c) Construct the matrix representing reflection through the xy plane.
- (d) Check that all these matrices are orthogonal, and calculate their determinants.

Problem A.15 In the usual basis $(\hat{i}, \hat{j}, \hat{k})$, 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 $\hat{i}' = \hat{j}$, $\hat{j}' = -\hat{i}$, $\hat{k}' = \hat{k}$. Construct the matrix \mathbf{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 A.16 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 \mathbf{S} is *unitary*, and \mathbf{H}^e is hermitian, then \mathbf{H}^f is hermitian. Show that \mathbf{S} carries an orthonormal basis into another orthonormal basis if and only if it is unitary.

- * **Problem A.17** Prove that $\text{Tr}(\mathbf{T}_1\mathbf{T}_2) = \text{Tr}(\mathbf{T}_2\mathbf{T}_1)$. It follows immediately that $\text{Tr}(\mathbf{T}_1\mathbf{T}_2\mathbf{T}_3) = \text{Tr}(\mathbf{T}_2\mathbf{T}_3\mathbf{T}_1)$, but is it the case that $\text{Tr}(\mathbf{T}_1\mathbf{T}_2\mathbf{T}_3) = \text{Tr}(\mathbf{T}_2\mathbf{T}_1\mathbf{T}_3)$, in general? Prove it, or disprove it. *Hint:* The best disproof is always a counterexample—the simpler the better!

A.5 Eigenvectors and Eigenvalues

Consider the linear transformation in 3-space consisting of a rotation, about some specified axis, by an angle θ . Most vectors (with tails at the origin) 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 space¹⁶ *every* linear transformation has "special" vectors like these, which are transformed into scalar multiples of themselves:

$$\hat{T}|\alpha\rangle = \lambda|\alpha\rangle; \quad (\text{A.69})$$

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 A.69 for *any* \hat{T} and *any* λ ; technically, an eigenvector is any *nonzero* vector satisfying Equation A.69.) 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

$$\mathbf{T}\mathbf{a} = \lambda\mathbf{a}, \quad (\mathbf{a} \neq \mathbf{0}), \quad (\text{A.70})$$

or

$$(\mathbf{T} - \lambda\mathbf{I})\mathbf{a} = \mathbf{0}. \quad (\text{A.71})$$

(Here $\mathbf{0}$ is the (column) matrix whose elements are all zero.) Now, if the matrix $(\mathbf{T} - \lambda\mathbf{I})$ had an *inverse*, we could multiply both sides of Equation A.71 by $(\mathbf{T} - \lambda\mathbf{I})^{-1}$, and conclude that $\mathbf{a} = \mathbf{0}$. But by assumption \mathbf{a} is *not* zero, so the matrix $(\mathbf{T} - \lambda\mathbf{I})$ must in fact be singular, which means that its determinant is zero:

$$\det(\mathbf{T} - \lambda\mathbf{I}) = \begin{vmatrix} (T_{11} - \lambda) & T_{12} & \dots & T_{1n} \\ T_{21} & (T_{22} - \lambda) & \dots & T_{2n} \\ \vdots & \vdots & & \vdots \\ T_{n1} & T_{n2} & \dots & (T_{nn} - \lambda) \end{vmatrix} = 0. \quad (\text{A.72})$$

Expansion of the determinant yields an algebraic equation for λ :

$$C_n\lambda^n + C_{n-1}\lambda^{n-1} + \dots + C_1\lambda + C_0 = 0, \quad (\text{A.73})$$

where the coefficients C_i depend on the elements of \mathbf{T} (see Problem A.20). This is called the **characteristic equation** for the matrix; its solutions determine the eigenvalues. Notice that it's an n th-order equation, so (by the **fundamental theorem of algebra**) it has n (complex) roots.¹⁷ However, some of these may be multiple roots, so all we can say for certain is that an $n \times n$ matrix has *at least one* and *at most n* distinct eigenvalues. The collection of all the eigenvalues of a matrix is called its **spectrum**; if two (or more) linearly independent eigenvectors share the same eigenvalue, the spectrum is said to be **degenerate**.

To construct the *eigenvectors* it is generally easiest simply to plug each λ back into Equation A.70 and solve "by hand" for the components of \mathbf{a} . I'll show you how it goes by working out an example.

Example 1.1

Find the eigenvalues and eigenvectors of the following matrix:

$$M = \begin{pmatrix} 2 & 0 & -2 \\ -2i & i & 2i \\ 1 & 0 & -1 \end{pmatrix}. \quad (\text{A.74})$$

Solution: The characteristic equation is

$$\begin{vmatrix} (2-\lambda) & 0 & -2 \\ -2i & (i-\lambda) & 2i \\ 1 & 0 & (-1-\lambda) \end{vmatrix} = -\lambda^3 + (1+i)\lambda^2 - i\lambda = 0, \quad (\text{A.75})$$

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):

$$\mathbf{a}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \text{ for } \lambda_1 = 0. \quad (\text{A.76})$$

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 I'll pick $a_1 = 2$, so

$$\mathbf{a}^{(2)} = \begin{pmatrix} 2 \\ 1-i \\ 1 \end{pmatrix}, \text{ for } \lambda_2 = 1. \quad (\text{A.77})$$

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

$$\mathbf{a}^{(3)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \text{ for } \lambda_3 = i. \quad (\text{A.78})$$

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

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

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

$$T = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad (\text{A.79})$$

and the (normalized) eigenvectors are

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \quad (\text{A.80})$$

A matrix that can be brought to **diagonal form** (Equation A.79) by a change of basis is said to be **diagonalizable** (evidently a matrix is diagonalizable if and only if its eigenvectors span the space). The similarity matrix that effects the diagonalization can be constructed by using the eigenvectors (in the old basis) as the columns of S^{-1} :

$$(S^{-1})_{ij} = (a^{(j)})_i. \quad (\text{A.81})$$

Example 1.2

In Example A.1,

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

so (using Equation A.57)

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

you can check for yourself that

$$\mathbf{S}\mathbf{a}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{S}\mathbf{a}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{S}\mathbf{a}^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

and

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

There's an obvious 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. If the characteristic equation has n distinct roots, then the matrix is certainly diagonalizable, but it *may* be diagonalizable even if there are multiple roots. (For an example of a matrix that *cannot* be diagonalized, see Problem A.19.) It would be handy to know in advance (before working out all the eigenvectors) whether a given matrix is diagonalizable. A useful sufficient (though not necessary) condition is the following: A matrix is said to be **normal** if it commutes with its hermitian conjugate:

$$\text{normal : } [\mathbf{N}^\dagger, \mathbf{N}] = 0. \tag{A.82}$$

Every normal matrix is diagonalizable (its eigenvectors span the space). In particular, every *hermitian* matrix is diagonalizable, and so is every *unitary* matrix.

Suppose we have *two* diagonalizable matrices; in quantum applications the question often arises: Can they be **simultaneously diagonalized** (by the *same* similarity matrix \mathbf{S})? That is to say, does there exist a basis all of whose members are eigenvectors of *both* matrices? In this basis, both matrices would be diagonal. The answer is yes *if and only if the two matrices commute*, as we shall now prove. (By the way, if two matrices commute with respect to *one* basis, they commute with respect to *any* basis—see Problem A.23.)

We first show that if a basis of simultaneous eigenvectors exists then the matrices commute. Actually, it's trivial in the (simultaneously) diagonal form:

$$\begin{aligned}
 \mathbf{T}\mathbf{V} &= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_n \end{pmatrix} \\
 &= \begin{pmatrix} \lambda_1 v_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 v_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n v_n \end{pmatrix} = \mathbf{V}\mathbf{T}.
 \end{aligned} \tag{A.83}$$

The converse is trickier. We start with the special case where the spectrum of \mathbf{T} is nondegenerate. Let the basis of eigenvectors of \mathbf{T} be labeled $\mathbf{a}^{(i)}$

$$\mathbf{T}\mathbf{a}^{(i)} = \lambda_i \mathbf{a}^{(i)}. \tag{A.84}$$

We assume $[\mathbf{T}, \mathbf{V}] = 0$ and we want to prove that $\mathbf{a}^{(i)}$ is also an eigenvector of \mathbf{V} .

$$[\mathbf{T}, \mathbf{V}] \mathbf{a}^{(i)} = 0, \quad \text{or} \quad \mathbf{T}\mathbf{V}\mathbf{a}^{(i)} - \mathbf{V}\mathbf{T}\mathbf{a}^{(i)} = 0 \tag{A.85}$$

and from Equation A.84

$$\mathbf{T}(\mathbf{V}\mathbf{a}^{(i)}) = \lambda_i (\mathbf{V}\mathbf{a}^{(i)}). \tag{A.86}$$

Equation A.86 says that the vector $\mathbf{b}^{(i)} \equiv \mathbf{V}\mathbf{a}^{(i)}$ is an eigenvector of \mathbf{T} with eigenvalue λ_i . But by assumption, the spectrum of \mathbf{T} is nondegenerate and that means that $\mathbf{b}^{(i)}$ must be (up to a constant) $\mathbf{a}^{(i)}$ itself. If we call the constant v_i ,

$$\mathbf{b}^{(i)} = \mathbf{V}\mathbf{a}^{(i)} = v_i \mathbf{a}^{(i)}, \tag{A.87}$$

so $\mathbf{a}^{(i)}$ is an eigenvector of \mathbf{V} .

All that remains is to relax the assumption of nondegeneracy. Assume now that \mathbf{T} has at least one degenerate eigenvalue such that both $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ are eigenvectors of \mathbf{T} with the same eigenvalue λ_0 :

$$\mathbf{T}\mathbf{a}^{(1)} = \lambda_0 \mathbf{a}^{(1)}$$

$$\mathbf{T}\mathbf{a}^{(2)} = \lambda_0 \mathbf{a}^{(2)}.$$

We again assume that the matrices \mathbf{T} and \mathbf{V} commute, so

$$[\mathbf{T}, \mathbf{V}] \mathbf{a}^{(1)} = 0$$

$$[\mathbf{T}, \mathbf{V}] \mathbf{a}^{(2)} = 0,$$

which leads to the conclusion (as in the nondegenerate case) that both $\mathbf{b}^{(1)} \equiv \mathbf{V}\mathbf{a}^{(1)}$ and $\mathbf{b}^{(2)} \equiv \mathbf{V}\mathbf{a}^{(2)}$ are eigenvectors of \mathbf{T} with eigenvalue λ_0 . But this time we can't say that $\mathbf{b}^{(1)}$ is a constant times $\mathbf{a}^{(1)}$ since any linear combination of $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ is an eigenvector of \mathbf{T} with eigenvalue λ_0 . All we know is that

$$\mathbf{b}^{(1)} = \mathbf{V}\mathbf{a}^{(1)} = c_{11}\mathbf{a}^{(1)} + c_{21}\mathbf{a}^{(2)}$$

$$\mathbf{b}^{(2)} = \mathbf{V}\mathbf{a}^{(2)} = c_{12}\mathbf{a}^{(1)} + c_{22}\mathbf{a}^{(2)}$$

for some constants c_{ij} . So $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ are not eigenvectors of \mathbf{V} (unless the constants c_{12} and c_{21} just happen to vanish). But suppose we choose a different basis of eigenvectors $\tilde{\mathbf{a}}$

$$\begin{aligned}\tilde{\mathbf{a}}^{(1)} &= d_{11}\mathbf{a}^{(1)} + d_{21}\mathbf{a}^{(2)} \\ \tilde{\mathbf{a}}^{(2)} &= d_{12}\mathbf{a}^{(1)} + d_{22}\mathbf{a}^{(2)},\end{aligned}\tag{A.88}$$

for some constants d_{ij} , such that $\tilde{\mathbf{a}}^{(1)}$ and $\tilde{\mathbf{a}}^{(2)}$ are eigenvectors of V :

$$V\tilde{\mathbf{a}}^{(i)} = v_i \tilde{\mathbf{a}}^{(i)}. \tag{A.89}$$

The $\tilde{\mathbf{a}}$ s are still eigenvectors of T , with the same eigenvalue λ_0 , since any linear combinations of $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ are. But *can* we construct linear combinations (A.88) that are eigenvectors of V —how do we get the appropriate coefficients d_{ij} ? *Answer:* We solve the eigenvalue problem¹⁸

$$\underbrace{\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}}_{C} \begin{pmatrix} d_{1i} \\ d_{2i} \end{pmatrix} = v_i \begin{pmatrix} d_{1i} \\ d_{2i} \end{pmatrix}. \tag{A.90}$$

I'll let you show (Problem A.24) that the eigenvectors $\tilde{\mathbf{a}}^{(i)}$ constructed in this way satisfy Equation A.88, completing the proof.¹⁹ What we have seen is that, when the spectrum contains degeneracy, the eigenvectors of one matrix aren't *automatically* eigenvectors of a second commuting matrix, but we can always *choose* a linear combination of them to form a simultaneous basis of eigenvectors.

- * **Problem A.18** The 2×2 matrix representing a rotation of the xy plane is

$$T = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \tag{A.91}$$

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 S that diagonalizes T . Perform the similarity transformation (STS^{-1}) explicitly, and show that it reduces T to diagonal form.

Problem A.19 Find the eigenvalues and eigenvectors of the following matrix:

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

Can this matrix be diagonalized?

Problem A.20 Show that the first, second, and last coefficients in the characteristic equation (Equation A.73) are:

$$C_n = (-1)^n, \quad C_{n-1} = (-1)^{n-1} \text{Tr}(T), \quad \text{and } C_0 = \det(T). \tag{A.92}$$

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

Problem A.21 It's obvious that the trace of a *diagonal* matrix is the sum of its eigenvalues, and its determinant is their product (just look at Equation A.79). It follows (from Equations A.65 and A.68) that the same holds for any *diagonalizable* matrix. Prove that in fact

$$\det(\mathbf{T}) = \lambda_1 \lambda_2 \cdots \lambda_n, \quad \text{Tr}(\mathbf{T}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \quad (\text{A.93})$$

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 eigenvectors 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 A.20.

Problem A.22 Consider the matrix

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

- (a) Is it normal?
- (b) Is it diagonalizable?

Problem A.23 Show that if two matrices commute in *one* basis, then they commute in *any* basis. That is:

$$[\mathbf{T}_1^e, \mathbf{T}_2^e] = \mathbf{0} \Rightarrow [\mathbf{T}_1^f, \mathbf{T}_2^f] = \mathbf{0}. \quad (\text{A.94})$$

Hint: Use Equation A.64.

Problem A.24 Show that the $\tilde{\mathbf{a}}$ computed from Equations A.88 and A.90 are eigenvectors of \mathbf{V} .

* **Problem A.25** Consider the matrices

$$\mathbf{A} = \begin{pmatrix} 1 & 4 & 1 \\ 4 & -2 & 4 \\ 1 & 4 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & -2 & -1 \\ -2 & 2 & -2 \\ -1 & -2 & 1 \end{pmatrix}.$$

- (a) Verify that they are diagonalizable and that they commute.
- (b) Find the eigenvalues and eigenvectors of \mathbf{A} and verify that its spectrum is nondegenerate.

- (c) Show that the eigenvectors of \mathbf{A} are eigenvectors of \mathbf{B} as well.

Problem A.26 Consider the matrices

$$\mathbf{A} = \begin{pmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & -1 & 2 \\ -1 & 5 & -1 \\ 2 & -1 & 2 \end{pmatrix}.$$

- (a) Verify that they are diagonalizable and that they commute.
- (b) Find the eigenvalues and eigenvectors of \mathbf{A} and verify that its spectrum is degenerate.
- (c) Are the eigenvectors that you found in part (b) also eigenvectors of \mathbf{B} ? If not, find the vectors that are simultaneous eigenvectors of both matrices.

A.6 Hermitian Transformations

In Equation A.48 I defined the hermitian conjugate (or “adjoint”) of a *matrix* as its transpose-conjugate: $T^\dagger = \tilde{T}^*$. 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:

$$\langle \hat{T}^\dagger \alpha | \beta \rangle = \langle \alpha | \hat{T} \beta \rangle, \quad (\text{A.95})$$

(for all vectors $|\alpha\rangle$ and $|\beta\rangle$).²⁰ 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 *names*. 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*: $\hat{T}\beta$ is the name of the vector $\hat{T}|\beta\rangle$, and $\langle \hat{T}^\dagger \alpha | \beta \rangle$ is the inner product of the vector $\hat{T}^\dagger|\alpha\rangle$ with the vector $|\beta\rangle$. Notice in particular that

$$\langle \alpha | c\beta \rangle = c \langle \alpha | \beta \rangle, \quad (\text{A.96})$$

whereas

$$\langle c\alpha | \beta \rangle = c^* \langle \alpha | \beta \rangle, \quad (\text{A.97})$$

for any scalar c .

If you’re working in an orthonormal basis (as we always do), the hermitian conjugate of a linear transformation is represented by the hermitian conjugate of the corresponding matrix; for (using Equations A.50 and A.53),

$$\langle \alpha | \hat{T} \beta \rangle = a^\dagger T b = (T^\dagger a)^\dagger b = \langle \hat{T}^\dagger \alpha | \beta \rangle. \quad (\text{A.98})$$

So the terminology is consistent, and we can speak interchangeably in the language of transformations or of matrices.

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:

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 A.20), so $\lambda = \lambda^*$, and hence λ is real. QED

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

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

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

As we have seen, this is equivalent to the statement that any hermitian matrix can be diagonalized. This rather technical fact is, in a sense, the mathematical support on which much of quantum mechanics leans. It turns out to be a thinner reed than one might have hoped, because the proof does not carry over to infinite-dimensional vector spaces.

Problem A.27 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 $|\gamma\rangle$. *Hint:* First let $|\gamma\rangle = |\alpha\rangle + |\beta\rangle$, and then let $|\gamma\rangle = |\alpha\rangle + i|\beta\rangle$.

* **Problem A.28** Let

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

- (a) Verify that \mathbf{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 \mathbf{S} , and check explicitly that it diagonalizes \mathbf{T} .
- (e) Check that $\det(\mathbf{T})$ and $\text{Tr}(\mathbf{T})$ are the same for \mathbf{T} as they are for its diagonalized form.

** **Problem A.29** Consider the following hermitian matrix:

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

- (a) Calculate $\det(\mathbf{T})$ and $\text{Tr}(\mathbf{T})$.
- (b) Find the eigenvalues of \mathbf{T} . Check that their sum and product are consistent with (a), in the sense of Equation A.93. Write down the diagonalized version of \mathbf{T} .

- (c) Find the eigenvectors of \mathbf{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 A.19). Orthogonalize them, and check that both are orthogonal to the third. Normalize all three eigenvectors.
- (d) Construct the unitary matrix \mathbf{S} that diagonalizes \mathbf{T} , and show explicitly that the similarity transformation using \mathbf{S} reduces \mathbf{T} to the appropriate diagonal form.

Problem A.30 A unitary 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.

Problem A.31 *Functions* of matrices are typically defined by their Taylor series expansions. For example,

$$e^M \equiv I + M + \frac{1}{2}M^2 + \frac{1}{3!}M^3 + \dots \quad (\text{A.99})$$

- (a) Find $\exp(M)$, if

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

- (b) Show that if M is diagonalizable, then

$$\det(e^M) = e^{\text{Tr}(M)}. \quad (\text{A.100})$$

Comment: This is actually true even if M is *not* diagonalizable, but it's harder to prove in the general case.

- (c) Show that if the matrices M and N *commute*, then

$$e^{M+N} = e^M e^N. \quad (\text{A.101})$$

Prove (with the simplest counterexample you can think up) that Equation A.101 is *not* true, in general, for *non-commuting* matrices.²¹

- (d) If H is hermitian, show that e^{iH} is unitary.

¹ For our purposes, the scalars will be ordinary complex numbers. Mathematicians can tell you about vector spaces over more exotic fields, but such objects play no role in quantum mechanics. Note that $\alpha, \beta, \gamma \dots$ are *not* (ordinarily) numbers; they are *names* (labels)—“Charlie,” for instance, or “F43A-9GL,” or whatever you care to use to identify the vector in question.

- ² That is to say, these operations are always well-defined, and will never carry you outside the vector space.
- ³ It is customary, where no confusion can arise, to write the null vector without the adorning bracket: $|0\rangle \rightarrow 0$.
- ⁴ This is funny notation, since $\mathbf{0}$ is not a number. I'm simply adopting the name “–Charlie” for the inverse of the vector whose name is “Charlie.” More natural terminology will suggest itself in a moment.
- ⁵ A set of vectors that spans the space is also called **complete**, though I personally reserve that word for the infinite-dimensional case, where subtle questions of convergence may arise.
- ⁶ In this chapter I'll use a hat (^) to denote linear transformations; this is not inconsistent with my convention in the text (putting hats on operators), for (as we shall see) quantum operators *are* linear transformations.
- ⁷ Notice the reversal of indices between Equations [A.30](#) and [A.33](#). This is not a typographical error. Another way of putting it (switching $i \leftrightarrow j$ in Equation [A.30](#)) is that if the *components* transform with T_{ij} , the *basis* vectors transform with T_{ji} .
- ⁸ I'll use boldface capital letters, sans serif, to denote square matrices.
- ⁹ I'll use a boldface lower-case letters, sans serif, for row and column matrices.
- ¹⁰ The commutator only makes sense for *square* matrices, of course; for rectangular matrices the two orderings wouldn't even be the same size.
- ¹¹ Note that the left inverse is equal to the right inverse, for if $\mathbf{AT} = \mathbf{I}$ and $\mathbf{TB} = \mathbf{I}$, then (multiplying the second on the left by \mathbf{A} and invoking the first) we get $\mathbf{B} = \mathbf{A}$.
- ¹² I assume you know how to evaluate determinants. If not, see Mary L. Boas, *Mathematical Methods in the Physical Sciences*, 3rd edn (John Wiley, New York, 2006), Section [3.3](#).
- ¹³ In a *real* vector space (that is, one in which the scalars are real) the hermitian conjugate is the same as the transpose, and a unitary matrix is **orthogonal**: $\tilde{\mathbf{O}} = \mathbf{O}^{-1}$. For example, rotations in ordinary 3-space are represented by orthogonal matrices.
- ¹⁴ Notice, however, the radically different perspective: In this case we're talking about one and the same *vector*, referred to two completely different *bases*, whereas before we were thinking of a completely *different* vector, referred to the *same* basis.
- ¹⁵ Note that \mathbf{S}^{-1} certainly exists—if \mathbf{S} were singular, the $|f_i\rangle$ s would not span the space, so they wouldn't constitute a basis.
- ¹⁶ This is *not* always true in a *real* vector space (where the scalars are restricted to real values). See Problem [A.18](#).
- ¹⁷ 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 Problem [A.18](#).
- ¹⁸ You might worry that the matrix \mathbf{C} is not diagonalizable, but you need not. The matrix \mathbf{C} is a 2×2 block of the transformation $\hat{\mathbf{V}}$ written in the basis $\mathbf{a}^{(i)}$; it is diagonalizable by virtue of the fact that \mathbf{V} itself is diagonalizable.
- ¹⁹ I've only proved it for a two-fold degeneracy, but the argument extends in the obvious way to a higher-order degeneracy; you simply need to diagonalize a bigger matrix \mathbf{C} .
- ²⁰ If you're wondering whether such a transformation necessarily *exists*, that's a good question, and the answer is “yes.” See, for instance, Paul R. Halmos, *Finite Dimensional Vector Spaces*, 2nd edn, van Nostrand, Princeton (1958), Section 44.
- ²¹ See Problem [3.29](#) for the more general “Baker–Campbell–Hausdorff” formula.

Index

21-centimeter line [313](#)

A

absorption [411–422](#), [436](#)
 active transformation [236–237](#)
 addition of angular momenta [176–180](#)
 adiabatic approximation [426–433](#)
 adiabatic process [426–428](#)
 adiabatic theorem [408](#), [428–433](#)
 adjoint [45](#), [95](#), [471](#)
 agnostic [5](#), [446](#)
 Aharonov–Bohm effect [182–186](#)
 Airy function [364–365](#)
 Airy's equation [364](#)
 allowed energy [28](#)
 bands [220–224](#)
 bouncing ball [369](#)
 delta-function well [66](#)
 finite square well [72–73](#)
 harmonic oscillator [44](#), [51](#), [187](#), [370](#)
 helium atom [210–212](#)
 hydrogen atom [147](#), [303–304](#), [372](#)
 infinite cubical well [132](#), [216–217](#)
 infinite spherical well [139–142](#)
 infinite square well [32](#)
 periodic potential [220–224](#)
 potential well [356–357](#), [367–369](#)
 power-law potential [371](#)
 allowed transitions [246](#), [421](#), [438](#)
 alpha decay [360–361](#)
 alpha particle [360](#)
 angular equation [134–138](#), [164](#)
 angular momentum [157–165](#)
 addition [176–180](#)
 canonical [197](#)
 commutation relations [157](#), [162](#), [166](#), [303](#), [421](#)
 conservation [162](#), [250–251](#)
 eigenfunctions [162–164](#)
 eigenvalues [157–162](#)
 generator of translations [248–249](#)

intrinsic [166](#)
 mechanical [197](#)
 operator [157](#), [163](#)
 orbital [165](#)
 spin [165](#)
 anharmonic oscillator [324](#)
 anomalous magnetic moment [301](#)
 antibonding [340](#)
 anticommutator [246](#)
 antiferromagnetism [193](#), [345](#)
 anti-hermitian matrix [471](#)
 anti-hermitian operator [124](#)
 antisymmetric matrix [471](#)
 antisymmetric state [201](#), [206](#)
 anti-unitary operator [274](#)
 Aspect, A. [451](#)
 associated Laguerre polynomial [150](#)
 associated Legendre function [135](#), [137](#), [192](#), [439](#)
 atomic nomenclature [213](#)–[215](#)
 atoms [209](#)–[216](#)
 average [9](#)–[10](#)
 azimuthal angle [133](#), [135](#)
 azimuthal quantum number [147](#)

B

Baker–Campbell–Hausdorff formula [121](#), [272](#)
 Balmer series [155](#)–[156](#), [304](#)
 band structure [220](#)–[225](#)
 baryon [180](#)
 basis [113](#)–[115](#), [465](#), [473](#)
 bead on a ring [79](#), [97](#), [183](#)–[184](#), [293](#)–[294](#), [317](#)–[318](#), [325](#)
 Bell, J. [5](#)–[6](#), [449](#)
 Bell's inequality [451](#)
 Bell's theorem [449](#)–[454](#)
 Berry, M. [430](#)
 Berry's phase [185](#), [430](#)–[433](#)
 Bessel function [140](#)–[141](#), [381](#)–[382](#)
 spherical [140](#)–[141](#)
 binding energy [148](#)
 Biot–Savart law [300](#)
 blackbody spectrum [416](#)
 Bloch, F. [3](#), [216](#)
 function [240](#), [326](#)
 sphere [458](#)
 theorem [220](#)–[221](#), [238](#)–[240](#)
 vector [458](#)

Bohm, D. [6](#), [447](#)
 Bohr, N. [5](#), [107](#), [462](#)
 energies [147](#), [194](#), [295–296](#)
 formula [147](#)
 Hamiltonian [143](#), [295](#)
 magneton [226](#), [306](#)
 radius [147](#), [298](#)
 Boltzmann constant [23](#), [88](#)
 Boltzmann factor [88](#), [416](#)
 bonding [340](#), [344](#)
 Born, M. [4](#)
 approximation [380](#), [388–397](#)
 –Oppenheimer approximation [428](#)
 series [395–397](#)
 statistical interpretation [3–8](#)
 boson [201](#), [202](#)
 bouncing ball [332](#), [369](#)
 boundary conditions [32](#), [64–65](#)
 delta-function [64–65](#)
 finite square well [71–72](#)
 impenetrable wall [216](#), [230](#)
 periodic [230–231](#), [424](#)
 bound states [61–63](#), [143](#), [352](#)
 degenerate [78](#)
 delta-function [64–66](#)
 finite cylindrical well [70–72](#), [188–189](#)
 finite spherical well [143](#)
 finite square well [72](#)
 variational principle [352](#)
 box normalization [424](#)
 bra [117–118](#)
 Brillouin zone [241](#)
 bulk modulus [220](#)

C

canonical commutation relation [41](#), [132](#)
 canonical momentum [183](#)
 cat paradox [461–462](#)
 Cauchy's integral formula [389](#)
 causal influence [452–453](#)
 Cayley's approximation [443](#)
 central potential [132](#), [199](#)
 centrifugal term [139](#)
 Chandrasekhar, S. [349](#)
 Chandrasechar limit [228](#)
 change of basis [121–123](#)

characteristic equation 476, 481
 classical electron radius 167
 classical region 354–358, 362–363
 classical scattering theory 376–379, 395
 classical velocity 55–56, 58–59
 classical wave equation 417
 Clebsch–Gordan coefficients 178–180, 190, 259–262
 coefficient
 A and B 416–417
 Clebsch–Gordan 178–180, 190, 259–262
 reflection 67, 75, 375
 transmission 67, 359, 370
 cofactor 472
 coherent radiation 414
 coherent state 126–127
 cohesive energy 84–85, 225
 coincident spectral lines 190
 “cold” solid 218, 220
 collapse 6, 102, 170, 443, 446–447, 453
 column matrix 470
 commutator 41, 108
 angular momentum 157, 162, 303, 421
 canonical 41, 132
 matrix 471
 uncertainty principle and 106–107
 commuting operators 40–41
 compatible observables 107–108, 251
 complementarity principle 446
 complete inner product space 92
 completeness
 eigenfunctions 34–36, 99
 set of functions 34–36, 93
 Hilbert space 92
 quantum mechanics 5, 446, 449
 complete set of vectors 465
 component 465
 conductor 223
 conjugate
 complex 471
 hermitian 471
 connection formulas 362–371
 conservation laws 242–243
 angular momentum 162, 250–251
 energy 37, 112–113, 266–267
 momentum 240–241
 parity 244–246

probability [22](#), [187–188](#)
 conservative system [3](#)
 continuity equation [187](#)
 continuous spectrum [99–102](#)
 continuous symmetry [232](#)
 continuous variable [11–14](#)
 continuum state [143](#)
 contour integral [390](#)
 Copenhagen interpretation [5](#)
 correspondence principle [420](#)
 Coulomb barrier [360](#), [349–350](#)
 Coulomb potential [143](#), [209](#), [213](#), [381](#), [394](#)
 Coulomb repulsion [360](#)
 covalent bond [340](#), [344](#)
 Crandall's puzzle [324–325](#)
 cross product [466](#)
 cross-section [377–378](#), [426](#)
 crystal [229–230](#), [320–321](#), [398](#)
 momentum [239–240](#), [275](#)
 cubic symmetry [321](#)
 cyclotron motion [182](#)

D

d (diffuse) [213](#)
 D_4 [232](#)
 de Broglie formula [19](#), [55](#)
 de Broglie wavelength [19](#), [23](#)
 decay modes [418](#)
 decoherence [462](#)
 degeneracy [252–253](#), [269–271](#)
 free particle [254](#)
 higher-order [294–295](#)
 hydrogen [149](#), [253](#), [270–272](#)
 infinite spherical well [253](#)
 Kramers' [275](#)
 lifting [287–288](#)
 pressure [219](#), [227–228](#)
 rotational [253](#)
 three-dimensional oscillator [253](#)
 two-fold [286–294](#)
 degenerate perturbation theory [283](#), [286–295](#), [314–315](#)
 first-order [290](#)
 second-order [317–318](#)
 degenerate spectra [96](#), [476](#)
 degenerate states [78](#)
 delta-function potential [61–70](#)

barrier [68–69](#), [440](#)
 bound state [64–66](#), [329](#), [331](#)
 bump [283](#), [285](#), [294](#), [323](#), [325](#), [440](#)
 Dirac comb [221](#)
 interaction [284](#)
 moving [80](#)
 scattering states [66–69](#)
 shell [385](#), [387](#)
 source [388](#)
 time-dependent [433](#)
 well [64–70](#), [329](#), [331](#)
 density
 matrix [455–459](#)
 of states [228–229](#), [423](#)
 operator [455](#)
 plot [151–152](#)
 d'Espagnat, B. [5](#)
 destructive measurement [459](#)
 determinant [206](#), [472](#)
 determinate state [96–97](#)
 deuterium [320](#)
 deuteron [320](#), [351](#)
 diagonal form [478](#)
 diagonalization [294](#), [478–480](#)
 diamagnetism [196](#), [322](#)
 Dieks, D. [459](#)
 differential scattering cross-section [377](#), [424–425](#)
 dihedral group [232](#)
 dimension [465](#)
 dipole moment
 electric [246–248](#), [302](#), [319](#)
 magnetic [172](#), [299–301](#), [305](#)
 Dirac, P. [301](#)
 comb [220–223](#)
 delta function [61–70](#), [74](#), [78](#), [80](#), [82–83](#), [99–102](#), [330](#)
 equation [304](#)
 notation [117–123](#)
 orthonormality [99–102](#)
 direct integral [339](#)
 Dirichlet's theorem [34](#), [60](#)
 discrete spectrum [98–99](#)
 discrete symmetry [232](#)
 discrete variable [8–11](#)
 dispersion relation [59](#)
 distinguishable particles [203](#)
 distribution [63](#)

domain [130](#)
doping [223–224](#)
dot product [466](#)
double-slit experiment [7–8](#)
double well [70, 79, 372–374](#)
dual space [118](#)
dynamic phase [429–430](#)
Dyson's formula [434](#)

E

effective nuclear charge [335–336](#)
effictive mass [326](#)
effective potential [138](#)
Ehrenfest's theorem [18, 48, 80, 132, 162, 174](#)
eigenfunction [96](#)
angular momentum [162–164](#)
continuous spectra [99–102](#)
determinate states [96](#)
Dirac notation [117–123](#)
discrete spectra [98–99](#)
hermitian operators [97–102](#)
incompatible observables [107](#)
momentum [99–100](#)
position [101](#)
eigenspinor [169](#)
eigenvalue [96, 475–482](#)
angular momentum [157–160](#)
determinate states [96](#)
generalized statistical interpretation [102–105](#)
hermitian transformations [483](#)
eigenvector [475–484](#)
Einstein, A. [5](#)
A and B coefficients [416–417](#)
boxes [448](#)
EPR paradox [447–448, 452](#)
mass-energy formula [362](#)
temperature [89](#)
electric dipole moment [246–248, 302, 319, 411](#)
electric dipole radiation [411](#)
electric dipole transition [438](#)
electric quadrupole transition [438](#)
electromagnetic
field [181–182, 411](#)
Hamiltonian [181](#)
interaction [181–186](#)
wave [411](#)

electron
 configuration [213–215](#)
 gas [216–220](#)
 g-factor [301, 311](#)
 in magnetic field [172–176, 430–432](#)
 interference [7–8](#)
 magnetic dipole moment [300–301](#)
 volt [148](#)

electron-electron repulsion [209–211, 325, 333–335, 343](#)

elements [213–216](#)

energy
 binding [148](#)
 cohesive [84–85](#)
 conservation [37, 112–113, 266–267](#)
 ionization [148, 333, 336](#)
 photon [155](#)
 relativistic [296](#)
 second-order [284–286](#)

energy-time uncertainty principle [109–113](#)

ensemble [16](#)

entangled state [199, 448](#)

EPR paradox [447–448, 451](#)

ethereal influence [453](#)

Euler's formula [26](#)

even function [30, 33, 71–72](#)

event rate [378](#)

exchange force [203–205](#)

exchange integral [339](#)

exchange operator [207](#)

exchange splitting [344](#)

excited state [33](#)
 helium [211–212](#)
 infinite square well [33](#)
 lifetime [418](#)

exclusion principle [202, 213, 218, 223](#)

exotic atom [313](#)

expectation value [10, 16–18](#)
 effect of perturbation [323–324](#)
 generalized Ehrenfest theorem [110](#)
 generalized statistical interpretation [103–104](#)
 Hamiltonian [30](#)
 harmonic oscillator [47–48](#)
 stationary state [27](#)
 time derivative [110](#)

extended uncertainty principle [127–128](#)

F

f (fundamental) [213](#)
 Fermi, E. [423](#)
 energy [218](#), [225](#)
 Golden Rule [422–426](#)
 surface [218](#)
 temperature [219](#)
 fermion [201–208](#), [218](#)
 ferromagnetism [345](#)
 Feynman diagram [396](#)
 Feynman–Hellmann theorem [316](#), [318–319](#), [321](#)
 fine structure [295–304](#)
 constant [295](#)
 exact [304](#)
 hydrogen [295–304](#)
 relativistic correction [295–299](#)
 spin-orbit coupling [295–296](#), [299–304](#)
 finite spherical well [143](#)
 finite square well [70–76](#)
 shallow, narrow [72](#)
 deep, wide [72](#)
 first Born approximation [391–395](#)
 flux quantization [184](#)
 forbidden energies [223–224](#)
 forbidden transition [246](#), [421](#), [438](#)
 Foucault pendulum [426–428](#)
 Fourier series [34](#)
 Fourier transform [56](#), [69–70](#), [104](#)
 inverse [56](#)
 Fourier's trick [34](#), [100–102](#)
 fractional quantum Hall effect [209](#)
 free electron density [221](#)
 free electron gas [216–220](#)
 free particle [55–61](#), [111–112](#), [267–268](#)
 frustration [193](#)
 fundamental theorem of algebra [476](#)
 fusion [349–350](#)

G

Galilean transformation [272](#)
 Gamow, G. [360](#)
 theory of alpha decay [360–361](#)
 gap [223–224](#)
 gauge invariance [182](#)
 gauge transformation [182–183](#)
 gaussian [108–109](#)

function [14](#), [328–329](#), [331–332](#), [347](#)
 integral [61](#)
 wave packet [61](#), [77](#), [108–109](#), [130](#)
 generalized Ehrenfest theorem [110](#)
 generalized function [63](#)
 generalized statistical interpretation [102–105](#)
 generalized symmetrization principle [207](#)
 generalized uncertainty principle [105–108](#)
 generating function [54](#)
 generator
 rotations [248–249](#)
 translations in space [235](#)
 translations in time [263](#)
 geometric phase [429–430](#)
 g-factor [301](#)
 deuteron [320](#)
 electron [301](#), [311](#)
 Landé [306](#)
 muon [313](#)
 positron [313](#)
 proton [311](#)
 Golden Rule [422–426](#)
 “good” quantum number [298](#), [305–308](#)
 “good” state [287–295](#)
 Gram–Schmidt orthogonalization [98](#), [468](#)
 graphical solution [72](#)
 Green’s function [388–391](#), [397](#)
 ground state [33](#), [327](#)
 delta function well [329](#)
 elements [214–215](#)
 harmonic oscillator [328](#)
 helium [332–336](#)
 hydrogen atom [148](#)
 hydrogen ion (H^-) [336](#), [349](#)
 hydrogen molecule [341–346](#)
 hydrogen molecule ion [337–341](#)
 infinite spherical well [139](#)
 infinite square well [33](#), [329–331](#), [346](#)
 lithium atom [212](#)
 upper bound [327](#), [332](#)
 variational principle [327–332](#)
 group theory [180](#)
 group velocity [58–59](#)
 gyromagnetic ratio [172](#), [300](#), [305](#)

H

half harmonic oscillator [77](#), [368](#), [371](#)
 half-integer angular momentum [160](#), [164](#), [201](#)
 half-life [361](#), [420](#)
 Hamiltonian [27–28](#)
 atom [209](#)
 discrete and continuous spectra [102](#)
 electromagnetic [181](#)
 helium [210](#), [333](#)
 hydrogen atom [143](#)
 hydrogen molecule [341](#)
 hydrogen molecule ion [337](#)
 magnetic dipole in magnetic field [172–176](#), [195–196](#), [299](#)
 relativistic correction [296–297](#)
 Hankel function [381–382](#)
 hard-sphere scattering [376–378](#), [384](#)
 harmonic chain [229–230](#)
 harmonic crystal [229–230](#)
 harmonic oscillator [39–54](#), [267–268](#), [315](#)
 algebraic solution [48–54](#)
 allowed energies [44](#)
 analytic solution [39–48](#)
 changing spring constant [432](#)
 coherent states [44](#), [126–127](#)
 driven [441–442](#), [444](#)
 ground state [328](#), [332](#), [346–347](#)
 half [368](#), [371](#)
 perturbed [283–284](#), [286–289](#), [324](#)
 relativistic correction [298](#)
 radiation from [419–420](#)
 stationary states [46](#), [52](#)
 three-dimensional [187](#), [315](#)
 two-dimensional [287–288](#), [322–323](#)
 WKB approximation [370](#)
 heat capacity [89](#)
 Heisenberg, W. [462](#)
 picture [264–267](#), [434](#)
 uncertainty principle [19–20](#), [107](#), [132](#)
 Heitler-London approximation [341–342](#), [344](#)
 helium [210–212](#)
 electron-electron repulsion [210](#), [325](#)
 excited states [211](#), [325](#), [336](#)
 ground state [325](#), [332–336](#)
 ion (H_2^+) [211](#), [336](#)
 ionization energy [336](#)
 ortho- [211–212](#)
 para- [211–212](#)

“rubber band” model 348–349
 helium-3 220
 Helmholtz equation 388, 391
 Hermite polynomial 52–54
 hermitian conjugate 45, 95, 161, 471
 hermitian matrix 471
 hermitian operator 94–95, 297, 299
 continuous spectra 99–102
 discrete spectra 98–99
 eigenfunctions 97–102
 eigenvalues 97–102, 483
 hermitian transformation 482–485
 hidden variable 5–6, 449, 454
 Hilbert space 91–94, 100–101, 113–114
 hole 223–224
 Hooke’s law 39
 Hund’s rules 214–215, 346
 hydrogen atom 143–156
 allowed energies 147, 149, 194
 binding energy 148
 degeneracy 149, 253, 270–272
 fine structure 295–304
 ground state 148, 312, 347
 hyperfine structure 295–296, 311–313
 in infinite spherical well 194
 muonic 200, 313
 potential 143
 radial wave function 144, 152–154
 radius 147
 spectrum 155–156
 Stark effect 319–320, 322, 374
 variational principle 347
 wave functions 151
 Zeeman effect 304–310
 hydrogenic atom 155–156
 hydrogen ion (H^-) 336, 349
 hydrogen molecule 341–346
 hydrogen molecule ion 337–341
 hyperfine splitting 295–296, 311–313, 320, 439

I

idempotent operator 119
 identical particles 198–231
 bosons 201, 205
 fermions 201, 205
 two-particle systems 198–207

identity matrix 472
 identity operator 118, 121–122
 impact parameter 376
 impenetrable walls 216, 230
 impulse approximation 395
 incident wave 66–68, 379
 incoherent perturbation 413–415
 incompatible observables 107–108, 158
 incompleteness 5, 446, 449
 indeterminacy 4, 452
 indistinguishable particles 201
 infinite cubical well 132, 216–217, 294
 infinite spherical well 139–142, 194
 infinite square well 31–39

- moving wall 428–429, 432, 440
- perturbed 279–286, 323
- rising barrier 440
- rising floor 436
- two particles 202–203, 205
- variational principle 329–331, 346

 WKB approximation 356–357
 infinitesimal transformation 240
 inner product 92–93, 466–468
 inner product space 467
 interaction picture 434
 intrinsic angular momentum 166
 insulator 223
 integration by parts 17
 interference 7–8
 inverse beta decay 228
 inverse Fourier transform 56
 inverse matrix 472
 inversion symmetry 243–244
 ionization 148, 336, 422, 425

J

Jordan, P. 5

K

ket 117–118
 kinetic energy 18, 296–297
 Kramers' degeneracy, 275
 Kramers' relation 319
 Kronecker delta 33–34
 Kronig–Penney model 221–222

L

ladder operators 41–47, 158–159, 163, 229–230

Laguerre polynomial 150

Lamb shift 295–296

Landau levels 182

Landé g-factor 306

Laplacian 131, 133

Larmor formula 419

Larmor frequency 173

Larmor precession 172

laser 412–413

Laughlin wave function 208

LCAO representation 337

Legendre function 135, 137

Legendre polynomial 90, 135–136, 138

Levi-Civita symbol 171

Lie group 235

lifetime 23, 112, 361–362, 418

excited state 418–420

lifting degeneracy 287

linear algebra 464–485

changing bases 473–475

eigenvectors and eigenvalues 475–484

inner product 466–468

matrices 468–473

vectors 464–466

linear combination 28, 465

linear independence 465

linear operator 94

linear transformation 91, 94, 468

lithium atom 212

lithium ion 336

locality 447

Lorentz force law 181

lowering operator 41–47, 159, 161–162

luminosity 378

Lyman series 155–156

M

magnetic dipole 172

anomalous moment 301

electron 301, 311

energy 172, 299, 304

force on 174

moment 172, 299, 300–301, 305

proton 311

transition 438

magnetic
 field [172–176](#), [300](#), [430](#)
 flux [183–184](#)
 frustration [193](#)
 quantum number [149](#)
 resonance [436–437](#)
 susceptibility [226](#)
 magnetization [226](#)
 Mandelstam–Tamm uncertainty principle [111](#)
 many worlds interpretation [6](#), [462](#)
 matrix [91](#), [468–473](#)
 adjoint [471](#)
 antisymmetric [471](#)
 characteristic equation [476](#)
 column [476](#)
 complex conjugate [471](#)
 density [455–459](#)
 determinant [472](#)
 diagonal [478](#)
 eigenvectors and eigenvalues [475–482](#)
 element [115](#), [120](#), [126](#), [469](#)
 function of [485](#)
 hermitian [471](#), [483–484](#)
 hermitian conjugate [471](#)
 identity [472](#)
 imaginary [471](#)
 inverse [472](#)
 normal [479](#)
 orthogonal [472](#)
 Pauli [168](#), [171](#)
 real [471](#)
 row [470](#)
 similar [474](#)
 simultaneous diagonalization [479](#)
 singular [472](#)
 skew hermitian [471](#)
 spectrum [476](#)
 spin [168](#), [171–172](#), [191](#)
 symmetric [471](#)
 transpose [470–471](#)
 tri-diagonal [444](#)
 unitary [472](#), [474](#)
 zero [476](#)
 mean [9–11](#)
 measurement [3–8](#), [30](#), [170](#), [462](#)
 cat paradox [461–462](#)

destructive [459](#)
 generalized statistical interpretation [102–105](#)
 indeterminacy [96, 452](#)
 repeated [6, 170](#)
 sequential [124, 194–195](#)
 simultaneous
 uncertainty principle [107–108](#)
 median [9–10](#)
 Mermin, N. [5, 453](#)
 meson [180, 447](#)
 metal [84, 223](#)
 metastable state [421](#)
 minimal coupling [181–182](#)
 minimum-uncertainty [108–109, 193](#)
 mixed state [456–458](#)
 momentum [16–18](#)
 angular [157–165](#)
 canonical [183](#)
 conservation [240–241](#)
 de Broglie formula [19, 23](#)
 eigenfunctions/eigenvalues [99–100](#)
 generator of translations [235](#)
 mechanical [197](#)
 operator [17, 41, 95, 99–100, 131](#)
 relativistic [296–297](#)
 transfer [392](#)
 momentum space [104–105, 121–123, 188](#)
 most probable configuration [9](#)
 Mott insulator [223](#)
 muon [313, 349](#)
 muon catalysis [349–350](#)
 muonic hydrogen [200, 313](#)
 muonium [313](#)

N

nearly-free electron approximation [314](#)
 Neumann function [140–141](#)
 neutrino oscillation [117](#)
 neutron diffraction [397–398](#)
 neutron star [228](#)
 no-clone theorem [459–460](#)
 node [33, 140–142, 146](#)
 nomenclature (atomic) [213–215](#)
 nondegenerate perturbation theory [279–286](#)
 noninteracting particles [198–199](#)
 nonlocality [447, 451–452](#)

non-normalizable function [14](#), [56](#), [66](#)

norm [467](#)

normal matrix [479](#)

normalization [14–16](#), [30](#), [93](#)

box [424](#)

free particle [56](#), [424](#)

harmonic oscillator [45–46](#)

hydrogen [150](#)

spherical coordinates [136](#)

spherical harmonics [191–192](#)

spinor [169](#)

three dimensions [131](#)

two-particle systems [198](#), [203](#)

variational principle [327](#)

vector [467](#)

wave function [14–16](#)

nuclear fusion [349–350](#)

nuclear lifetime [360–362](#)

nuclear magnetic resonance [437](#)

nuclear motion [200–201](#), [209](#)

nuclear scattering length [398](#)

null vector [96](#), [464](#)

O

observable [94–97](#), [99](#)

determinate state [96–97](#)

hermitian operator [94–95](#)

incompatible [107–108](#)

observer [461–462](#)

odd function [33](#)

operator [17](#)

anti-unitary [274](#)

angular mometum [157](#), [163](#)

anti-hermitian [124](#)

commuting [40–41](#)

differentiating [121](#)

Dirac notation [120](#)

exchange [207](#)

Hamiltonian [27–28](#)

hermitian [94–95](#)

identity [118](#), [121–122](#)

incompatible [252–253](#)

ladder [41–47](#), [158–159](#), [161–163](#)

linear [94](#)

lowering [41–47](#), [159](#), [161–163](#)

momentum [17](#), [41](#)

noncommuting 252–253
 parity 233–234, 243–248
 position 17, 101
 product rule 121
 projection 118, 121–122, 314
 raising 41–47, 159, 161–163
 rotation 233–234, 248–251
 scalar 250
 vector 249–250
 optical theorem 397
 orbital 213, 337
 orbital angular momentum 165
 orthodox position 5, 446
 orthogonality 33
 eigenfunctions 98–99
 functions 93
 Gram–Schmidt procedure 98, 468
 hydrogen wave functions 151
 spherical harmonics 137
 vectors 467
 orthogonal matrix 472
 orthohelium 211–212
 orthonormality 33, 46, 467
 Dirac 99–102, 118
 eigenfunctions 100
 functions 93
 vectors 467
 orthorhombic symmetry 321
 overlap integral 339

P

p (principal) 213
 parahelium 211–212
 paramagnetism 196, 226
 parity 233–234, 243–248
 hydrogen states 234
 polar coordinates 234
 spherical harmonics 234
 partial wave 380–387
 Paschen-Back effect 307
 Paschen series 155–156
 passive transformation 236–237
 Pasternack relation 319
 Pauli, W. 5
 exclusion principle 202, 213, 218, 223
 paramagnetism 226

spin matrices [168](#), [171](#)
 periodic boundary conditions
 Periodic Table [213–216](#)
 permanent [206](#)
 perturbation theory [279–326](#)
 constant [281–282](#)
 degenerate [283](#), [286–295](#), [314–315](#), [317](#)
 expectation value [323–324](#)
 first order [279–284](#), [332](#)
 higher order [285](#), [317](#)
 nondegenerate [279–286](#)
 second order [279](#), [284–286](#)
 time-independent [279–326](#)
 time-dependent [402](#), [405–411](#)
 phase
 Berry's [185](#), [430–433](#)
 dynamic [429–430](#)
 gauge transformation [183](#), [185](#)
 geometric [429–430](#)
 wave function [18](#), [32](#), [38](#), [185](#)
 phase shift [385–387](#)
 phase velocity [58–59](#)
 phonon [230](#)
 photoelectric effect [422](#), [425–426](#)
 photon [7](#), [155](#), [412](#), [417](#)
 Plancherel's theorem [56](#), [60](#), [69–70](#)
 Planck formula [155](#), [312](#)
 Planck's blackbody spectrum [416–417](#)
 Planck's constant [3](#)
 polar angle [133](#)
 polarization [411](#), [414](#)
 polarizability [324](#)
 population inversion [413](#)
 position
 eigenfunctions/eigenvalues [101](#)
 generalized statistical interpretation [104–105](#)
 operator [17](#), [265–265](#)
 space
 position-momentum uncertainty principle [19–20](#), [107](#)
 position space wave function [104](#), [121–123](#)
 positronium [200](#), [313](#)
 potential [25](#)
 Coulomb [143](#)
 delta-function [61–70](#)
 Dirac comb [221](#)
 effective [138](#)

finite square well [70–76](#)
 hydrogen [143](#)
 infinite square well [31–39](#)
 Kronig–Penney [221–222](#)
 power law [371](#)
 reflectionless [81](#)
 scalar [181](#)
 sech-squared [81, 371, 375](#)
 spherically symmetrical [132, 371, 393–394](#)
 step [75–76](#)
 super- [129](#)
 vector [181](#)
 Yukawa [347, 351–352](#)

potential well [352, 355–358, 367–369](#)
 power law potential [371](#)
 power series method [40, 49–50, 145](#)
 principal quantum number [140, 147](#)
 probability [8–14](#)
 Born statistical interpretation [3–8](#)
 conservation [22, 188](#)
 continuous variables [11–14](#)
 current [22, 60, 187–188](#)
 density quad [12–13](#)
 discrete variables [8–11](#)
 generalized statistical interpretation [102–105](#)
 reflection [67, 375](#)
 transition [409](#)
 transmission [67, 359](#)
 projection operator [118, 121–122, 314](#)
 propagator [267–268, 396](#)
 proton
 g-factor [311](#)
 magnetic moment [311](#)
 magnetic field [300](#)
 pseudovector [245–247, 250](#)
 pure state [455–456](#)

Q

quantum
 computation [460](#)
 cryptography [460](#)
 dot [350–351](#)
 dynamics [402–445](#)
 electrodynamics [181, 301, 412, 417](#)
 Hall effect [209](#)
 information [460](#)

jump [155](#), [402](#)
 Xerox machine [459–460](#)
 Zeno effect [442–443](#)

quantum number
 angular momentum [160](#), [166](#)
 azimuthal [147](#)
 “good” [298](#), [305–308](#)
 magnetic [147](#)
 principal [140](#), [147](#)

quark [180](#)

R

Rabi flopping [410](#), [414](#)
 radial equation [138–143](#)
 radial wave function [138](#), [144](#), [152](#)
 radiation [411–418](#)
 radiation zone [381](#)
 radius
 Bohr [147](#), [298](#)
 classical electron [167](#)
 raising operator [41–47](#), [159](#), [161–162](#)
 Ramsauer-Townsend effect [74](#)
 Rayleigh’s formula [383](#)
 realist position [5](#), [170](#), [446](#)
 reciprocal lattice [399](#)
 recursion formula [50–52](#), [148](#)
 reduced mass [200](#)
 reduced matrix element [256](#)
 reflected wave [66–68](#)
 reflection coefficient [67](#), [375](#)
 reflectionless potential [81](#)
 relativistic correction [296–299](#)
 harmonic oscillator [298](#)
 hydrogen [295–299](#)
 relativistic energy [296](#)
 relativistic momentum [296](#)
 resonance curve [437](#)
 revival time [76](#)
 Riemann zeta function [36–37](#)
 rigged Hilbert space [100–101](#)
 rigid rotor [165](#)
 Rodrigues formula [54](#), [135](#)
 “roof lines” [351](#)
 rotating wave approximation [410](#)
 rotations [233–234](#), [248–262](#)
 generator [248–249](#)

infinitesimal [248–251](#)
 spinor [269](#)
 row matrix [470](#)
 Runge–Lenz vector [270–272](#)
 Rutherford scattering [379, 394](#)
 Rydberg constant [155](#)
 Rydberg formula [155](#)

S

s (sharp) [213](#)
 scalar [464](#)
 pseudo- [246, 250](#)
 “true” [246, 250](#)
 scalar multiplication [465](#)
 scalar operator [250](#)
 scalar potential [181](#)
 scattering [376–401, 424–425](#)
 amplitude [380](#)
 angle [376](#)
 Born approximation [380, 388–397](#)
 classical [376–379](#)
 cross-section [377–378](#)
 hard-sphere [376–378, 384, 387](#)
 identical particles [400–401](#)
 length [398](#)
 low energy [393](#)
 matrix [81–82](#)
 one dimensional
 partial wave analysis [380–387](#)
 phase shift [385–387](#)
 Rutherford [379, 394](#)
 soft-sphere [393, 395, 397](#)
 two dimensional [399–400](#)
 Yukawa [394–395](#)
 scattering states [61–63, 66–70, 81–82](#)
 delta function [66–70](#)
 finite square well [73–76](#)
 tunneling [68–69, 358–362, 370, 375](#)
 Schrödinger, E. [188](#)
 Schrödinger equation [3, 131–132](#)
 electromagnetic [181](#)
 helium [348](#)
 hydrogen [143](#)
 integral form [388–391](#)
 momentum space [130](#)
 normalization [14–16](#)

radial 138
 spherical coordinates 133
 three-dimensional 131–132
 time dependent 3, 15, 24–31, 131, 402–403
 time independent 25–26, 132
 two-particle systems 198–200
 WKB approximation 354–375
 Schrödinger picture 265, 434
 Schrödinger's cat 461–462
 Schwarz inequality 92, 106, 467–468
 screening 213, 335–336
 sech-squared potential 81, 371, 375
 selection rules 246, 420–422

- parity 246–248
- scalar operator 255–258
- vector operator 258–262

 self-adjoint matrix 471
 self-adjoint operator 130
 semiclassical regime 358
 semiconductor 223
 separable solution 25–31
 separation of variables 25–26, 133–134
 sequential measurements 194–195
 shell 213
 shielding 335–336
 shooting method 194
 similar matrices 474
 simple harmonic oscillator equation 31
 simultaneous diagonalization 479
 singlet configuration 177, 206, 312
 singular matrix 472
 sinusoidal perturbation 408–411
 sinusoidal wave 56
 skew-hermitian matrix 124, 471
 skew-hermitian operator 124
 Slater determinant 206
 S-matrix 81–82
 solenoid 183–184
 solids 23–24, 84, 216–225

- band structure 220–224
- free electron model 216–220
- Kronig–Penney model 221–222

 Sommerfeld, A. 216
 span 465
 spectral decomposition 120
 spectrum 96, 476

blackbody 416–417
 coincident lines 190
 degenerate 96, 476
 hydrogen 155–156
 matrix 476
 spherical Bessel function 140–141, 381–382
 spherical coordinates 132–134
 angular equation 134–138
 radial equation 138–154
 separation of variables 133
 spherical Hankel function 381–382
 spherical harmonic 137–138, 191–192, 234
 spherical Neumann function 140–141
 spherically symmetric potential 132–134, 371, 393–394
 spherical tensor 258
 spin 165–180, 191
 commutation relations 166
 down 167
 entangled states 177, 199, 447
 matrix 168, 191
 one 172
 one-half 167–171
 singlet 177
 statistics 201
 three-halves 191
 triplet 177
 up 167
 spinor 167, 247
 spin-orbit coupling 295, 299–304
 spin-spin coupling 312, 344
 spontaneous emission 412–413, 416–422
 hydrogen 439
 lifetime of excited state 418–420
 selection rules 420–422
 square-integrable function 14, 92–93
 square well
 double 79
 finite 70–76
 infinite 31–39
 standard deviation 11
 Stark effect 286, 319–320, 322, 374
 state
 mixed 456–458
 pure 455–456
 stationary states 25–31, 324–325
 delta-function well 66

free particle 55–56
 harmonic oscillator 44
 infinite square well 31–32
 virial theorem 125
 statistical interpretation 3–8, 102–105
 statistics (spin and) 201
 step function 69, 330
 Stern–Gerlach experiment 174–175, 196
 stimulated emission 412–413
 Stoner criterion 227
 subsystem 459
 superconductor 184
 superluminal influence 452
 superpotential 129
 supersymmetry 40, 129
 symmetric matrix 471
 symmetric state 201, 206
 symmetrization principle 201, 207
 symmetry 232–275
 continuous 232
 cubic 321
 discrete 232
 inversion 243
 orthorhombic 321
 rotational 250–251, 269–270
 spherical 132–134, 371, 393
 tetragonal 321
 translational 238–242

T

Taylor series 39
 Taylor's theorem 49
 tensor operator 250
 tetragonal symmetry 321
 thermal energy 219
 theta function 330
 Thomas precession 302
 three-particle state 225
 time-dependent perturbation theory 402, 405–411, 434–435
 Golden Rule 422–426
 two-level systems 403–411
 time-dependent Schrödinger equation 3, 15, 24–31, 131, 197–198, 402–403, 433–434
 numerical solution 443–445
 time evolution operator 262–268
 time-independent perturbation theory 279–326
 degenerate 283, 286–295, 314–315, 317

nondegenerate 279–286
 time-independent Schrödinger equation 25–26
 bouncing ball 332, 369
 delta-function barrier 440
 delta-function well 63–70
 finite square well 70–74
 free particle 55–61
 harmonic oscillator 39–54
 helium atom 210
 hydrogen atom 143
 hydrogen molecule 341
 hydrogen molecule ion 337
 infinite square well 31–39
 stationary states 25–31
 three dimensions 132
 two-particle 198–200
 time-ordered product 434
 time reversal 272–275
 time translation 266–267
 total cross-section 378, 383–384
 trace 474
 transfer matrix 82–83
 transformation
 active 236–237
 hermitian 482–484
 infinitesimal 240
 linear 91, 468
 of operators 235–237
 unitary 484
 transition 155, 402
 allowed 438
 electric dipole 438
 electric quadrupole 438
 forbidden 421, 438
 magnetic dipole 438
 passive 236–237
 transition probability 409
 transition rate 414–415
 translation 232–233
 generator 235
 infinitesimal 240–241
 operator 232–233, 235–242, 268
 symmetry 238–242
 time 262–268
 transmission coefficient 67, 75, 359, 370
 transmitted wave 66–68

transpose [470–471](#)
triangle inequality [468](#)
tri-diagonal matrix [444](#)
triplet configuration [177, 312](#)
“true” vector [245–247, 250](#)
tunneling [63, 68–69, 358–362, 370](#)
in Stark effect [374](#)
turning point [61–62, 354–355, 363–364, 366](#)
two-fold degeneracy [286–294](#)
two-level system [403–411](#)
two-particle systems [198–200](#)

U

uncertainty principle [19–20, 105–113](#)
angular momentum [132, 158](#)
energy-time [109–113](#)
extended [127–128](#)
generalized [105–108](#)
Heisenberg [132](#)
position-momentum [19–20, 132](#)
minimum-uncertainty wave packet [108–109](#)
unitary matrix [472, 474, 484](#)
unitary transformation [484](#)
unit vector [467](#)
unstable particle [23](#)

V

valence electron [216](#)
van der Waals interaction [315–316](#)
van Hove singularity [229](#)
variables
continuous [11–14](#)
discrete [8–11](#)
hidden [5–6, 449, 454](#)
separation of [25–26, 133](#)
variance [11](#)
variational principle [327–353](#)
bouncing ball [332](#)
delta function well [331](#)
excited states [331–332](#)
harmonic oscillator [328, 331–332](#)
helium [332–336](#)
hydrogen [347](#)
hydrogen ion (H^-) [336](#)
hydrogen molecule [341–346](#)
hydrogen molecule ion [337–341](#)
infinite square well [327–332](#)

quantum dot 350–351

Yukawa potential 347

vector 91, 464–466

addition 464

changing bases 473–475

column 470

Dirac notation

inverse 464

null 464

operator 249–250

pseudo- 245–247, 250

row 470

“true” 245–247, 250

unit 467

zero 464–465

vector potential 181

vector space 464

velocity 17

classical 56, 58–59

group 58–59

phase 58–59

vertex factor 396

virial theorem 125, 187, 298

W

wag the dog method 51, 83–84, 194

watched pot phenomenon 442–443

wave function 3–8, 14–16

collapse 6, 102, 170, 443, 446–447, 453

free particle 55–61

hydrogen 151

infinite square well 32

Laughlin 208–209

momentum space 104–105, 121–123, 188

normalization 14–16

position space 104, 121–123

radial 138, 144, 152

statistical interpretation 3–8

three dimensions 131–132

two-particle 198–200

unstable particle 23

wavelength 19, 155

de Broglie 19, 23

wave number 379

wave packet 56–61

gaussian 61, 77, 108–109, 130

minimum-uncertainty [108–109](#)
wave-particle duality [446](#)
wave vector [217](#)
white dwarf [227](#)
wiggle factor [26, 29](#)
Wigner, E. [462](#)
Wigner-Eckart theorem [255–262, 307](#)
WKB approximation [354–375](#)
“classical” region [355, 363](#)
connection formulas [362–371](#)
double well [372–374](#)
hydrogen [372](#)
non-vertical walls [368–369](#)
one vertical wall [367–368](#)
radial wave function [371](#)
tunneling [358–362, 374–375](#)
two vertical walls [356–357](#)
Wootters, W. [459](#)

Y

Young’s double-slit experiment [7–8](#)
Yukawa potential [347, 351–352, 394](#)
Yukawa scattering [394](#)

Z

Zeeman effect [304–310](#)
intermediate-field [309–310](#)
strong-field [307–308](#)
weak-field [305–307](#)
Zener diode [362](#)
Zener tunneling [362–363](#)
Zeno effect [442–443](#)
zero matrix [476](#)
zero vector [464–465](#)
Zurek, W. [459](#)