FURTHER PROBLEMS FOR CHAPTER 8

**Problem 8.11 Use the WKB approximation to find the allowed energies of the general power-law potential:

$$V(x) = \alpha|x|^{\nu},$$

where ν is a positive number. Check your result for the case $\nu = 2$. *Answer:*

$$E_n = \alpha\left[(n - 1/2)\hbar\sqrt{\frac{\pi}{2m\alpha}}\frac{\Gamma\left(\frac{1}{\nu} + \frac{3}{2}\right)}{\Gamma\left(\frac{1}{\nu} + 1\right)}\right]^{\left(\frac{2\nu}{\nu+2}\right)}. \qquad [8.53]$$

**Problem 8.12 Use the WKB approximation to find the bound-state energy for the potential in Problem 2.48. Compare the exact answer. *Answer:*

$$-[(9/8) - (1/\sqrt{2})]\hbar^2 a^2/m.$$

Problem 8.13 For spherically symmetrical potentials, we can apply the WKB approximation to the radial equation, (Equation 4.37). In the case $l = 0$, it is reasonable[10] to use Equation 8.47 in the form

$$\int_0^{r_0} p(r)\,dr = (n - 1/4)\pi\hbar, \qquad [8.54]$$

where r_0 is the turning point (in effect, we treat $r = 0$ as an infinite wall). Apply 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 8.14 Use the WKB approximation in the form

$$\int_{r_1}^{r_2} p(r)\,dr = (n - 1/2)\pi\hbar \qquad [8.55]$$

[10] 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).

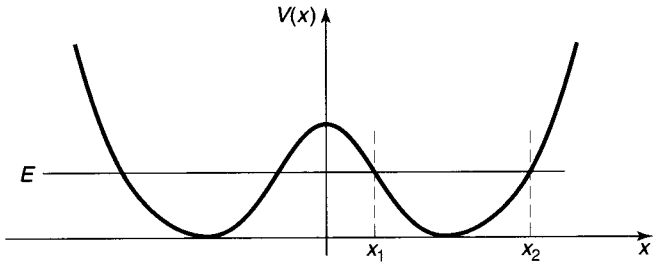

Figure 8.13: Symmetric double well; Problem 8.15.

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:

$$\int_a^b \frac{1}{x}\sqrt{(x-a)(b-x)} = \frac{\pi}{2}(\sqrt{b}-\sqrt{a})^2. \qquad [8.56]$$

Note that you recover the Bohr levels when $n \gg l$ and $n \gg 1/2$. *Answer:*

$$E_{nl} \cong \frac{-13.6\,\text{eV}}{[n-(1/2)+\sqrt{l(l+1)}]^2}. \qquad [8.57]$$

***Problem 8.15** Consider the case of a symmetrical double-well, such as the one pictured in Figure 8.13. We are interested in bound states with $E < V(0)$.

(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 8.46, for x_2; you will have to work out x_1 for yourself), to show that

$$\psi(x) \cong \begin{cases} \dfrac{D}{\sqrt{|p(x)|}} e^{-\frac{1}{\hbar}\int_{x_2}^{x}|p(x')|dx'}, & \text{(i)} \\ \dfrac{2D}{\sqrt{p(x)}} \sin\left[\dfrac{1}{\hbar}\displaystyle\int_x^{x_2} p(x')\,dx' + \dfrac{\pi}{4}\right], & \text{(ii)} \\ \dfrac{D}{\sqrt{|p(x)|}}\left[2\cos\theta\, e^{\frac{1}{\hbar}\int_x^{x_1}|p(x')|dx'} + \sin\theta\, e^{-\frac{1}{\hbar}\int_x^{x_1}|p(x')|dx'}\right], & \text{(iii)} \end{cases}$$

where

$$\theta \equiv \frac{1}{\hbar}\int_{x_1}^{x_2} p(x)\,dx. \qquad [8.58]$$

(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:

$$\tan\theta = \pm 2e^{\phi}, \qquad [8.59]$$

where

$$\phi \equiv \frac{1}{\hbar}\int_{-x_1}^{x_1} |p(x')|\,dx'. \qquad [8.60]$$

Equation 8.59 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 8.59 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

$$\theta \cong \left(n + \frac{1}{2}\right)\pi \mp \frac{1}{2}e^{-\phi}. \qquad [8.61]$$

(d) Suppose each well is a parabola[11]:

$$V(x) = \begin{cases} \frac{1}{2}m\omega^2(x+a)^2, & \text{if } x < 0, \\ \frac{1}{2}m\omega^2(x-a)^2, & \text{if } x > 0. \end{cases} \qquad [8.62]$$

Sketch this potential, find θ (Equation 8.58), and show that

$$E_n^{\pm} \cong \left(n + \frac{1}{2}\right)\hbar\omega \mp \frac{\hbar\omega}{2\pi}e^{-\phi}. \qquad [8.63]$$

Note: If the central barrier were *impenetrable* ($\phi \to \infty$), 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.

[11] Even if $V(x)$ is not strictly parabolic in each well, this calculation of θ, and hence the result (Equation 8.63), 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.

(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

$$\tau = \frac{2\pi^2}{\omega} e^{\phi}. \qquad [8.64]$$

(f) Calculate ϕ, for the specific potential in part (d), and show that for $V(0) \gg E$, $\phi \sim m\omega a^2/\hbar$.

CHAPTER 9

TIME-DEPENDENT PERTURBATION THEORY

Up to this point, 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,

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

can be solved by separation of variables:

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r}) e^{-iEt/\hbar},$$

where $\psi(\mathbf{r})$ satisfies the time-*in*dependent Schrödinger equation,

$$H\psi = E\psi.$$

Because the time dependence of Ψ 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 are constant in time. By forming *linear combinations* of these stationary states we obtain wave functions with more interesting time dependence, but even then the possible values of the energy, and their respective probabilities, are constant.

If we want to allow for **transitions** 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 portion of the Hamiltonian is small compared to the time-independent part, it can be treated as a perturbation. My purpose in this chapter is to develop time-dependent

perturbation theory, and study its most important application: the emission or absorption of radiation by an atom—a process known in the old Bohr theory as a **quantum jump**.

9.1 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 H_0:

$$H_0\psi_a = E_a\psi_a \quad \text{and} \quad H_0\psi_b = E_b\psi_b, \tag{9.1}$$

and they are orthonormal:

$$\langle\psi_a|\psi_b\rangle = \delta_{ab}. \tag{9.2}$$

Any state can be expressed as a linear combination of them; in particular,

$$\Psi(0) = c_a\psi_a + c_b\psi_b. \tag{9.3}$$

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 exponential factor:

$$\Psi(t) = c_a\psi_a e^{-iE_at/\hbar} + c_b\psi_b e^{-iE_bt/\hbar}. \tag{9.4}$$

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

$$|c_a|^2 + |c_b|^2 = 1. \tag{9.5}$$

9.1.1 The Perturbed System

Now suppose we turn on a time-dependent perturbation $H'(t)$. 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:

$$\Psi(t) = c_a(t)\psi_a e^{-iE_at/\hbar} + c_b(t)\psi_b e^{-iE_bt/\hbar}. \tag{9.6}$$

[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 that 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,

so that $c_a(0) = 1$ and $c_b(0) = 0$, 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,

$$H\Psi = i\hbar\frac{\partial\Psi}{\partial t}, \quad \text{where} \quad H = H_0 + H'(t). \tag{9.7}$$

From Equations 9.6 and 9.7, we find

$$c_a[H_0\psi_a]e^{-iE_at/\hbar} + c_b[H_0\psi_b]e^{-iE_bt/\hbar} + c_a[H'\psi_a]e^{-iE_at/\hbar} + c_b[H'\psi_b]e^{-iE_bt/\hbar} = i\hbar\left[\dot{c}_a\psi_a e^{-iE_at/\hbar} + \dot{c}_b\psi_b e^{-iE_bt/\hbar} + c_a\psi_a\left(-\frac{iE_a}{\hbar}\right)e^{-iE_at/\hbar} + c_b\psi_b\left(-\frac{iE_b}{\hbar}\right)e^{-iE_bt/\hbar}\right].$$

In view of Equation 9.1, the first two terms on the left cancel the last two terms on the right, and hence

$$c_a[H'\psi_a]e^{-iE_at/\hbar} + c_b[H'\psi_b]e^{-iE_bt/\hbar} = i\hbar\left[\dot{c}_a\psi_a e^{-iE_at/\hbar} + \dot{c}_b\psi_b e^{-iE_bt/\hbar}\right]. \tag{9.8}$$

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 9.2):

$$c_a\langle\psi_a|H'|\psi_a\rangle e^{-iE_at/\hbar} + c_b\langle\psi_a|H'|\psi_b\rangle e^{-iE_bt/\hbar} = i\hbar\dot{c}_a e^{-iE_at/\hbar}.$$

For short, we define

$$H'_{ij} \equiv \langle\psi_i|H'|\psi_j\rangle; \tag{9.9}$$

note that the Hermiticity of H' entails $H'_{ji} = (H'_{ij})^*$. Multiplying through by $-(i/\hbar)e^{iE_at/\hbar}$, we conclude that

$$\dot{c}_a = -\frac{i}{\hbar}\left[c_aH'_{aa} + c_bH'_{ab}e^{-i(E_b-E_a)t/\hbar}\right]. \tag{9.10}$$

Similarly, the inner product with ψ_b picks out $\dot{c}_b$:

$$c_a\langle\psi_b|H'|\psi_a\rangle e^{-iE_at/\hbar} + c_b\langle\psi_b|H'|\psi_b\rangle e^{-iE_bt/\hbar} = i\hbar\dot{c}_b e^{-iE_bt/\hbar},$$

and hence

$$\dot{c}_b = -\frac{i}{\hbar}\left[c_bH'_{bb} + c_aH'_{ba}e^{i(E_b-E_a)t/\hbar}\right]. \tag{9.11}$$

Equations 9.10 and 9.11 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 H' vanish (see Problem 9.4 for the more general case in which the diagonal terms are *not* zero):

$$H'_{aa} = H'_{bb} = 0. \tag{9.12}$$

In that case the equations simplify:

$$\boxed{\begin{aligned} \dot{c}_a &= -\tfrac{i}{\hbar} H'_{ab} e^{-i\omega_0 t} c_b, \\ \dot{c}_b &= -\tfrac{i}{\hbar} H'_{ba} e^{i\omega_0 t} c_a, \end{aligned}} \qquad [9.13]$$

where

$$\omega_0 \equiv \frac{E_b - E_a}{\hbar}. \qquad [9.14]$$

(We'll assume that $E_b \geq E_a$, so $\omega_0 \geq 0$.)

***Problem 9.1** A hydrogen atom is placed in a (time-dependent) electric field $\mathbf{E} = E(t)\hat{k}$. Calculate all four matrix elements H'_{ij} of the perturbation $H' = -eEz$ 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. As a result, 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-level configuration— assuming transitions to higher excited states can be ignored.

***Problem 9.2** Solve Equation 9.13 for the case of a *time-independent* perturbation, assuming that $c_a(0) = 1$ and $c_b(0) = 0$. Check that $|c_a(t)|^2 + |c_b(t)|^2 = 1$. *Note:* 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 9.3 Suppose the perturbation takes the form of a delta function (in time):

$$H' = U\delta(t - t_0);$$

assume that $U_{aa} = U_{bb} = 0$, and let $U_{ab} \equiv \alpha$. 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 probability ($P_{a\to b}$) that a transition occurs? *Hint*: Refer to Problem 2.24. *Answer*: $P_{a\to b} = (|\alpha|^2/\hbar^2)/(1 + |\alpha|^2/4\hbar^2)^2$.

9.1.2 Time-Dependent Perturbation Theory

So far, everything is *exact*: We have made no assumption about the *size* of the perturbation. But if H' is "small", we can solve Equations 9.13 by a process of successive approximations, as follows. Suppose the particle starts out in the lower state:

$$c_a(0) = 1, \quad c_b(0) = 0. \tag{9.15}$$

If there were *no perturbation at all*, they would stay this way forever:

Zeroth Order:

$$c_a^{(0)}(t) = 1, \quad c_b^{(0)}(t) = 0. \tag{9.16}$$

To calculate the first-order approximation, we insert these values on the right side of Equation 9.13:

First Order:

$$\frac{dc_a}{dt} = 0 \quad \Rightarrow \quad c_a^{(1)}(t) = 1;$$

$$\frac{dc_b}{dt} = -\frac{i}{\hbar} H'_{ba} e^{i\omega_0 t} \quad \Rightarrow \quad c_b^{(1)} = -\frac{i}{\hbar} \int_0^t H'_{ba}(t') e^{i\omega_0 t'} \, dt'. \tag{9.17}$$

Now we insert *these* expressions on the right to obtain the *second*-order approximation:

Second Order:

$$\frac{dc_a}{dt} = -\frac{i}{\hbar} H'_{ab} e^{-i\omega_0 t} \left(-\frac{i}{\hbar}\right) \int_0^t H'_{ba}(t') e^{i\omega_0 t'} \, dt' \quad \Rightarrow$$

$$c_a^{(2)}(t) = 1 - \frac{1}{\hbar^2} \int_0^t H'_{ab}(t') e^{-i\omega_0 t'} \left[\int_0^{t'} H'_{ba}(t'') e^{i\omega_0 t''} \, dt'' \right] dt', \tag{9.18}$$

while c_b is unchanged, $c_b^{(2)}(t) = c_b^{(1)}(t)$. [Notice that in my notation $c_a^{(2)}(t)$ *includes* the zeroth order term; the second-order *correction* would be the integral term alone.]

In principle, we could continue this ritual indefinitely, always inserting the n^{th}-order approximation into the right side of Equation 9.13 and solving for the $(n+1)^{\text{th}}$ order. Notice that c_a is modified in every *even* order, and c_b in every *odd* order. Incidentally, 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 9.5). However, $|c_a^{(1)}(t)|^2 + |c_b^{(1)}(t)|^2$ *is* equal to 1 *to first order in* H', which is all we can expect from a first-order approximation. And the same goes for the higher orders.

**Problem 9.4 Suppose you *don't* assume that $H'_{aa} = H'_{bb} = 0$.

(a) 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 H'.

(b) There is a nicer way to handle this problem. Let

$$d_a \equiv e^{\frac{i}{\hbar}\int_0^t H'_{aa}(t')dt'} c_a, \quad d_b \equiv e^{\frac{i}{\hbar}\int_0^t H'_{bb}(t')dt'} c_b. \tag{9.19}$$

Show that

$$\dot{d}_a = -\frac{i}{\hbar} e^{i\phi} H'_{ab} e^{-i\omega_0 t} d_b; \quad \dot{d}_b = -\frac{i}{\hbar} e^{-i\phi} H'_{ba} e^{i\omega_0 t} d_a, \tag{9.20}$$

where

$$\phi(t) \equiv \frac{1}{\hbar}\int_0^t [H'_{aa}(t') - H'_{bb}(t')]\, dt'. \tag{9.21}$$

So the equations for d_a and d_b are identical in structure to Equation 9.13 (with an extra factor $e^{i\phi}$ tacked onto H').

(c) 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 9.5 Solve Equation 9.13 to second order in perturbation theory, for the general case $c_a(0) = a$, $c_b(0) = b$.

**Problem 9.6 Calculate $c_a(t)$ and $c_b(t)$, to second order, for the perturbation in Problem 9.2. Compare your answer with the exact result.

9.1.3 Sinusoidal Perturbations

Suppose the perturbation has sinusoidal time dependence:

$$H'(\mathbf{r}, t) = V(\mathbf{r})\cos(\omega t), \tag{9.22}$$

so that

$$H'_{ab} = V_{ab}\cos(\omega t), \tag{9.23}$$

where

$$V_{ab} \equiv \langle \psi_a | V | \psi_b \rangle. \tag{9.24}$$

(As before, I'll assume that 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) we have (Equation 9.17)

$$c_b(t) \cong -\frac{i}{\hbar} V_{ba} \int_0^t \cos(\omega t') e^{i\omega_0 t'}\, dt' = -\frac{iV_{ba}}{2\hbar}\int_0^t \left[e^{i(\omega_0+\omega)t'} + e^{i(\omega_0-\omega)t'}\right] dt'$$

$$= -\frac{V_{ba}}{2\hbar}\left[\frac{e^{i(\omega_0+\omega)t}-1}{\omega_0+\omega}+\frac{e^{i(\omega_0-\omega)t}-1}{\omega_0-\omega}\right]. \qquad [9.25]$$

This is 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

$$\omega_0 + \omega \gg |\omega_0 - \omega|. \qquad [9.26]$$

This is not much of a limitation, since perturbations at *other* frequencies have a negligible probability of causing a transition *anyway*.[1] Dropping the first term, we have

$$c_b(t) \cong -\frac{V_{ba}}{2\hbar}\frac{e^{i(\omega_0-\omega)t/2}}{\omega_0-\omega}\left[e^{i(\omega_0-\omega)t/2}-e^{-i(\omega_0-\omega)t/2}\right]$$

$$= -i\frac{V_{ba}}{\hbar}\frac{\sin[(\omega_0-\omega)t/2]}{\omega_0-\omega}e^{i(\omega_0-\omega)t/2}. \qquad [9.27]$$

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

$$P_{a\to b}(t) = |c_b(t)|^2 \cong \frac{|V_{ab}|^2}{\hbar^2}\frac{\sin^2[(\omega_0-\omega)t/2]}{(\omega_0-\omega)^2}. \qquad [9.28]$$

The most remarkable feature of this result is that, as a function of time, the transition probability *oscillates* sinusoidally (Figure 9.1). 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 $n = 1, 2, 3, \ldots$, 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 9.7 it is shown that this "flopping" is not an artifact of perturbation theory—it also occurs in the *exact* solution, though the flopping *frequency* is modified somewhat.

As I noted earlier, the probability of a transition is greatest when the driving frequency is close to the "natural" frequency ω_0. This is illustrated in Figure 9.2, where $P_{a\to b}$ 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 9.7 you will see that the *exact* result never exceeds 1.)

[1] In the following sections we will be applying this theory to the case of *light*, for which $\omega \sim 10^{14}$ Hz, so the denominator in *both* terms is huge, except (for the second one) in the neighborhood of ω_0.

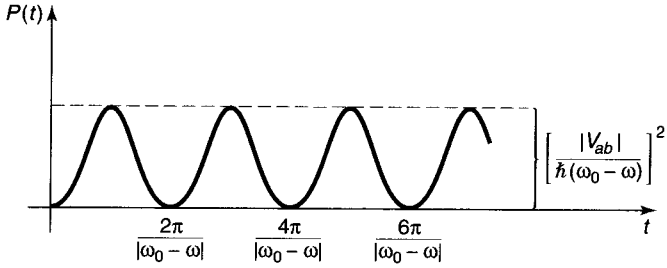

Figure 9.1: Transition probability as a function of time, for a sinusoidal perturbation (Equation 9.28).

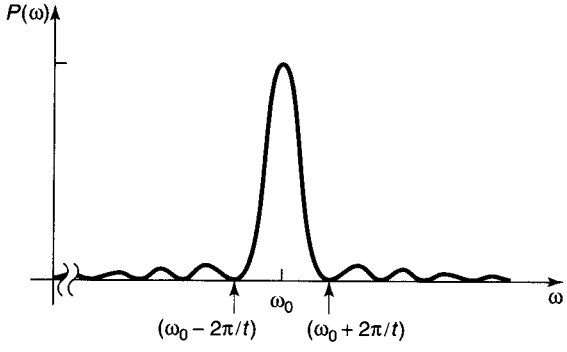

Figure 9.2: Transition probability as a function of driving frequency (Equation 9.28).

****Problem 9.7** The first term in Equation 9.25 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 $H' = (V/2)e^{-i\omega t}$, which is to say,

$$H'_{ba} = \frac{V_{ba}}{2}e^{-i\omega t}, \quad H'_{ab} = \frac{V_{ab}}{2}e^{i\omega t}. \qquad [9.29]$$

[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 9.25 for $c_a(t)$.] If you make this so-called **rotating wave approximation** at the *beginning* of the calculation, Equation 9.13 can be solved *exactly*, with no need for perturbation theory and no assumption about the strength of the field.

(a) Solve Equation 9.13 in the rotating wave approximation (Equation 9.29) 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**,

$$\omega_r \equiv \frac{1}{2}\sqrt{(\omega - \omega_0)^2 + (|V_{ab}|/\hbar)^2}. \qquad [9.30]$$

(b) Determine the transition probability, $P_{a\to b}(t)$, and show that it never exceeds 1. Confirm that $|c_a(t)|^2 + |c_b(t)|^2 = 1$.

(c) Check that $P_{a\to b}(t)$ reduces to the perturbation theory result (Equation 9.28) 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?

9.2 EMISSION AND ABSORPTION OF RADIATION

9.2.1 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 9.3). 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[2]; the atom, then, is exposed to a sinusoidally oscillating electric field

$$\mathbf{E} = E_0 \cos(\omega t)\,\hat{k} \qquad [9.31]$$

(for the moment I'll assume that the light is monochromatic and polarized along the z-direction). The perturbing Hamiltonian is[3]

$$H' = -q\,E_0 z \cos(\omega t), \qquad [9.32]$$

where q is the charge of the electron.[4] Evidently[5]

$$H'_{ba} = -\wp\,E_0 \cos(\omega t), \quad \text{where} \quad \wp \equiv q\langle\psi_b|z|\psi_a\rangle. \qquad [9.33]$$

[2]For visible light $\lambda \sim 5000$ Å, while the diameter of an atom is around 1 Å, so this approximation is reasonable; but it would *not* be for X-rays. Problem 9.20 explores the effect of spatial variation in the field.

[3]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 electro*static* 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 an atom).

[4]As usual, we assume that the nucleus is heavy and stationary; it is the wave function of the *electron* we are concerned with.

[5]The letter $\wp$ 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 $\wp$ to avoid confusion with momentum). In fact, $\wp$ is the off-diagonal matrix element of the z-component of the dipole moment operator $q\mathbf{r}$.

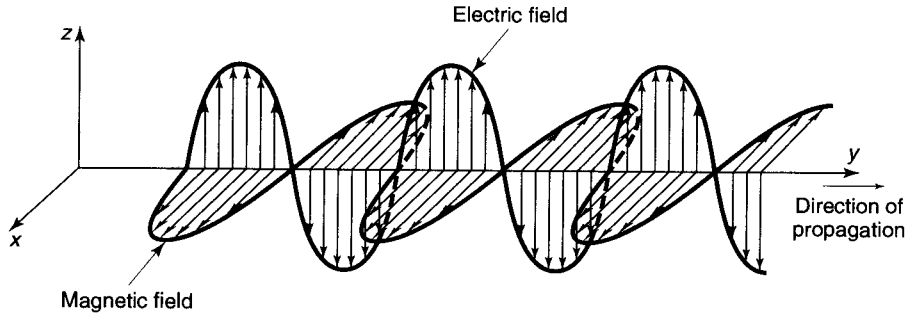

Figure 9.3: An electromagnetic wave.

Typically, ψ is an even or odd function of z; in either case $z|\psi|^2$ is odd, and integrates to zero (see Problem 9.1 for some examples). This licenses our usual assumption that the *diagonal* matrix elements of H' vanish. Thus the interaction of light with matter is governed by precisely the kind of oscillatory perturbation we studied in Section 9.1.3, with

$$V_{ba} = -\wp E_0. \qquad [9.34]$$

9.2.2 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 9.28, which (in view of Equation 9.34) takes the form

$$P_{a\to b}(t) = \left(\frac{|\wp|E_0}{\hbar}\right)^2 \frac{\sin^2[(\omega_0-\omega)t/2]}{(\omega_0-\omega)^2}. \qquad [9.35]$$

In this process, the atom absorbs energy $E_b - E_a = \hbar\omega_0$ from the electromagnetic field. We say that it has "absorbed a photon" (Figure 9.4a). [As I mentioned earlier, the word *"photon"* really belongs to **quantum electrodynamics** (the quantum theory of the electromagnetic field), whereas we are treating the field itself *classically*. But this terminology is convenient, as long as you don't read more into it than is really there.]

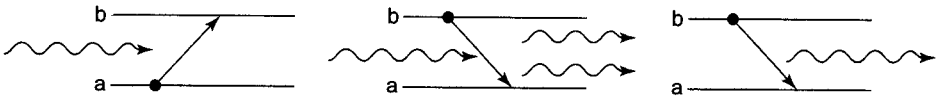

Figure 9.4: Three ways in which light interacts with atoms: (a) absorption, (b) stimulated emission, (c) spontaneous emission.

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 $P_{b\to a} = |c_a(t)|^2$, the probability of a transition *down* to the *lower* level:

$$P_{b\to a}(t) = \left(\frac{|\wp| E_0}{\hbar}\right)^2 \frac{\sin^2[(\omega_0 - \omega)t/2]}{(\omega_0 - \omega)^2}. \qquad [9.36]$$

(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 9.25 we 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 discovered 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 9.4b). This raises the possibility of *amplification*, for if I could obtain a bottle of atoms, all in the upper state, and trigger it with a single incident photon, a chain reaction would occur, with the first photon producing 2, and these 2 producing 4, 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, of course, the principle behind the **laser** (*l*ight *a*mplification by *s*timulated *e*mission of *r*adiation). 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 *produces* 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 9.4c). 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,[6] but there is a lovely argument due to Einstein[7] 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. 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.

9.2.3 Incoherent Perturbations

The energy density in an electromagnetic wave is[8]

$$u = \frac{\epsilon_0}{2} E_0^2, \qquad [9.37]$$

where E_0 is (as before) the amplitude of the electric field. So the transition probability (Equation 9.36) is (not surprisingly) proportional to the energy density of the fields:

$$P_{b\to a}(t) = \frac{2u}{\epsilon_0 \hbar^2} |\wp|^2 \frac{\sin^2[(\omega_0 - \omega)t/2]}{(\omega_0 - \omega)^2}. \qquad [9.38]$$

But this is for a **monochromatic** perturbation, consisting of a single frequency ω. In many applications the system is exposed to electromagnetic waves at a whole *range* of frequencies; in that case $u \to \rho(\omega)d\omega$, where $\rho(\omega)d\omega$ is the energy density in the

---

[6]For an especially nice treatment, see Rodney Loudon, *The Quantum Theory of Light*, 2nd ed. (Oxford: Clarendon Press, 1983).

[7]Einstein's paper was published in 1917, well before the Schrödinger equation. Quantum electrodynamics comes into the argument via the Planck blackbody formula (Equation 5.112), which dates from 1900.

[8]See, for example, D. Halliday and R. Resnick, *Fundamentals of Physics*, 3rd ed., extended (New York: John Wiley & Sons, 1988), Section 38-5. 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.

frequency range $d\omega$, and the net transition probability takes the form of an integral:[9]

$$P_{b\to a}(t) = \frac{2}{\epsilon_0\hbar^2}|\wp|^2 \int_0^\infty \rho(\omega) \left\{ \frac{\sin^2[(\omega_0-\omega)t/2]}{(\omega_0-\omega)^2} \right\} d\omega. \qquad [9.39]$$

Ordinarily, the term in curly brackets is sharply peaked about ω_0 (Figure 9.2), whereas $\rho(\omega)$ is relatively broad; in that case we may as well replace $\rho(\omega)$ by $\rho(\omega_0)$ and take it outside the integral:

$$P_{b\to a}(t) \cong \frac{2|\wp|^2}{\epsilon_0\hbar^2}\rho(\omega_0) \int_0^\infty \frac{\sin^2[(\omega_0-\omega)t/2]}{(\omega_0-\omega)^2} d\omega. \qquad [9.40]$$

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

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2}\, dx = \pi, \qquad [9.41]$$

we find

$$P_{b\to a}(t) \cong \frac{\pi|\wp|^2}{\epsilon_0\hbar^2}\rho(\omega_0)t. \qquad [9.42]$$

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*:

$$R_{b\to a} = \frac{\pi}{\epsilon_0\hbar^2}|\wp|^2\rho(\omega_0). \qquad [9.43]$$

So far, we have assumed that the perturbing wave is coming in along the x-direction (Figure 9.3) and polarized in the z-direction. But we shall be 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 $|\wp|^2$, is the *average* of $|\hat{n}\cdot\boldsymbol{\wp}|^2$, where

$$\boldsymbol{\wp} \equiv q\langle\psi_b|\mathbf{r}|\psi_a\rangle \qquad [9.44]$$

(generalizing Equation 9.33), and the average is over both polarizations ($\hat{n}$) and over all incident directions. This averaging can be carried out as follows.

---

[9]Equation 9.39 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.

Polarization: For propagation in the z-direction, the two possible polarizations are $\hat{\imath}$ and $\hat{\jmath}$, so the polarization average (subscript p) is

$$(\hat{n}\cdot\boldsymbol{\wp})_p^2 = \frac{1}{2}[(\hat{\imath}\cdot\boldsymbol{\wp})^2 + (\hat{\jmath}\cdot\boldsymbol{\wp})^2] = \frac{1}{2}(\wp_x^2 + \wp_y^2) = \frac{1}{2}\wp^2\sin^2\theta, \qquad [9.45]$$

where θ is the angle between $\boldsymbol{\wp}$ and the direction of propagation.

Propagation direction: Now let's set the polar axis along $\boldsymbol{\wp}$ and integrate over all propagation directions to get the polarization–propagation average (subscript pp):

$$(\hat{n}\cdot\boldsymbol{\wp})_{pp}^2 = \frac{1}{4\pi}\int\left[\frac{1}{2}\wp^2\sin^2\theta\right]\sin\theta\,d\theta\,d\phi = \frac{\wp^2}{4}\int_0^\pi \sin^3\theta\,d\theta = \frac{\wp^2}{3}. \qquad [9.46]$$

So the transition rate for stimulated emission from state b to state a, under the influence of incoherent, unpolarized light incident from all directions, is

$$\boxed{R_{b\to a} = \frac{\pi}{3\epsilon_0\hbar^2}|\boldsymbol{\wp}|^2\rho(\omega_0),} \qquad [9.47]$$

where $\boldsymbol{\wp}$ is the matrix element of the electric dipole moment between the two states (Equation 9.44) and $\rho(\omega_0)$ is the energy density in the fields, per unit frequency, evaluated at $\omega_0 = (E_b - E_a)/\hbar$.[10]

9.3 SPONTANEOUS EMISSION

9.3.1 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,[11] so that the number of particles leaving the upper state by this process, per unit time, is $N_b A$.[12] The transition rate for stimulated emission, as we have seen (Equation 9.47), is proportional to the energy density of the electromagnetic field—call it $B_{ba}\rho(\omega_0)$. 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,

$$\frac{dN_b}{dt} = -N_b A - N_b B_{ba}\rho(\omega_0) + N_a B_{ab}\rho(\omega_0). \qquad [9.48]$$

---

[10] This is a special case of Fermi's **Golden Rule** for time-dependent perturbation theory.

[11] Normally I'd use R for a transition rate, but out of deference to *der Alte* everyone follows Einstein's notation in this context.

[12] Assume that N_a and N_b are very large, so we can treat them as continuous functions of time, and ignore statistical fluctuations.

Suppose that 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

$$\rho(\omega_0) = \frac{A}{(N_a/N_b)B_{ab} - B_{ba}}. \quad [9.49]$$

On the other hand, we know from elementary statistical mechanics[13] that the number of particles with energy E, in thermal equilibrium at temperature T, is proportional to the **Boltzmann factor**, $\exp(-E/k_BT)$, so

$$\frac{N_a}{N_b} = \frac{e^{-E_a/k_BT}}{e^{-E_b/k_BT}} = e^{\hbar\omega_0/k_BT}, \quad [9.50]$$

and hence

$$\rho(\omega_0) = \frac{A}{e^{\hbar\omega_0/k_BT}B_{ab} - B_{ba}}. \quad [9.51]$$

But Planck's blackbody formula Equation 5.112 tells us the energy density of thermal radiation:

$$\rho(\omega) = \frac{\hbar}{\pi^2c^3}\frac{\omega^3}{e^{\hbar\omega/k_BT} - 1}; \quad [9.52]$$

comparing the two expressions, we conclude that

$$B_{ab} = B_{ba} \quad [9.53]$$

and

$$A = \frac{\omega^3\hbar}{\pi^2c^3}B_{ba}. \quad [9.54]$$

Equation 9.53 confirms what we already knew: that 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 9.54, 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 9.47 we read off

$$B_{ba} = \frac{\pi}{3\epsilon_0\hbar^2}|\mathbf{p}|^2, \quad [9.55]$$

and it follows that the spontaneous emission rate is

$$\boxed{A = \frac{\omega^3|\mathbf{p}|^2}{3\pi\epsilon_0\hbar c^3}.} \quad [9.56]$$

[13]See, for example, Charles Kittel and Herbert Kroemer, *Thermal Physics*, 2nd ed. (New York: Freeman, 1980), Chapter 3.

Problem 9.8 As a mechanism for downward transitions, spontaneous emission competes with thermally stimulated emission (stimulated emission for which Planck 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?

9.3.2 The Lifetime of an Excited State

Equation 9.56 is our fundamental result; it gives the transition rate for spontaneous emission. Suppose, now, that you have a bottle full of atoms, with $N_b(t)$ of them in 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:

$$dN_b = -AN_b\,dt \qquad [9.57]$$

(assuming there is no mechanism to replenish the supply).[14] Solving for $N_b(t)$, we find

$$N_b(t) = N_b(0)e^{-At}; \qquad [9.58]$$

evidently the number remaining in the excited state decreases exponentially, with a time constant

$$\tau = \frac{1}{A}. \qquad [9.59]$$

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 9.56) gives the transition rate for $\psi_b \to \psi_a$ regardless of any other allowed states (see Problem 9.14). Typically, an excited atom has many different **decay modes** (that is, ψ_b can decay to a large number of different lower-energy states, $\psi_{a_1}, \psi_{a_2}, \psi_{a_3}, \ldots$). In that case the transition rates *add*, and the net lifetime is

$$\tau = \frac{1}{A_1 + A_2 + A_3 + \cdots}. \qquad [9.60]$$

Example. 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 2.50) and decays by spontaneous emission to state $|n'\rangle$. From Equation 9.44, we have

$$\mathbf{\mathcal{P}} = q\langle n|x|n'\rangle\hat{\imath}.$$

[14] 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.

You calculated the matrix elements of x back in Problem 3.50:

$$\langle n|x|n'\rangle = \sqrt{\frac{\hbar}{2m\overline{\omega}}}(\sqrt{n'}\delta_{n,n'-1} + \sqrt{n}\delta_{n',n-1}),$$

where $\overline{\omega}$ is the natural frequency of the oscillator (I use the overbar to distinguish it from the frequency of the emitted radiation, although as we'll see in a moment the two turn out to be equal, and at that point I'll drop the bar). But we're talking about *emission*, so n' must be *lower* than n; for our purposes, then,

$$\mathbf{p} = q\sqrt{\frac{n\hbar}{2m\overline{\omega}}}\delta_{n',n-1}\hat{\imath}. \qquad [9.61]$$

Evidently transitions occur only to states one step lower on the "ladder," and the frequency of the photon emitted is

$$\omega = \frac{E_n - E_{n'}}{\hbar} = \frac{(n+1/2)\hbar\overline{\omega} - (n'+1/2)\hbar\overline{\omega}}{\hbar} = (n-n')\overline{\omega} = \overline{\omega}. \qquad [9.62]$$

Not surprisingly, the system radiates at the classical oscillator frequency. The transition rate (Equation 9.56) is

$$A = \frac{nq^2\omega^2}{6\pi\epsilon_0 mc^3}, \qquad [9.63]$$

and the lifetime of the n^{th} stationary state is

$$\tau_n = \frac{6\pi\epsilon_0 mc^3}{nq^2\omega^2}. \qquad [9.64]$$

Meanwhile, each radiated photon carries an energy $\hbar\omega$, so the *power* radiated is $A\hbar\omega$:

$$P = \frac{q^2\omega^2}{6\pi\epsilon_0 mc^3}(n\hbar\omega),$$

or, since the energy of an oscillator in the n^{th} state is $E = (n+1/2)\hbar\omega$,

$$P = \frac{q^2\omega^2}{6\pi\epsilon_0 mc^3}\left(E - \frac{1}{2}\hbar\omega\right). \qquad [9.65]$$

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**:[15]

$$P = \frac{q^2a^2}{6\pi\epsilon_0 c^3}. \qquad [9.66]$$

[15] See, for example, David J. Griffiths *Introduction to Electrodynamics*, 2nd ed. (Englewood Cliffs. NJ: Prentice-Hall, 1989), Section 9.1.4.

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,

$$P = \frac{q^2 x_0^2 \omega^4}{12\pi\epsilon_0 c^3}.$$

But the *energy* of the oscillator is $E = (1/2)m\omega^2 x_0^2$, so $x_0^2 = 2E/m\omega^2$, and hence

$$P = \frac{q^2\omega^2}{6\pi\epsilon_0 m c^3} E. \qquad [9.67]$$

This is the average power radiated by a *classical* oscillator with energy E. In the classical limit ($\hbar \to 0$) the classical and quantum formulas agree[16]; however, the quantum formula (Equation 9.65) protects the ground state: If $E = (1/2)\hbar\omega$ the oscillator does not radiate.

Problem 9.9 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 9.10** 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 scan them before you start calculating. *Answer*: 1.60×10^{-9} seconds for all except ψ_{200}, which is infinite.

9.3.3 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 9.10 (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 may specify the states with the usual quantum numbers n, l, and m, and the matrix elements are

$$\langle n'l'm'|\mathbf{r}|nlm\rangle.$$

---

[16] In fact, if we express P in terms of the energy *above the ground state*, the two formulas are *identical*.

Clever exploitation of the angular momentum commutation relations and the hermiticity of the angular momentum operators yields a set of powerful constraints on this quantity.

Selection rules involving m and m': Consider first the commutators of L_z with x, y, and z, which we worked out in Chapter 4 (see Equation 4.122):

$$[L_z, x] = i\hbar y, \quad [L_z, y] = -i\hbar x, \quad [L_z, z] = 0. \tag{9.68}$$

From the third of these it follows that

$$0 = \langle n'l'm'|[L_z, z]|nlm\rangle = \langle n'l'm'|(L_z z - zL_z)|nlm\rangle$$

$$= \langle n'l'm'|[(m'\hbar)z - z(m\hbar)]|nlm\rangle = (m' - m)\hbar\langle n'l'm'|z|nlm\rangle.$$

Conclusion:

$$\text{Either } m' = m, \text{ or else } \langle n'l'm'|z|nlm\rangle = 0. \tag{9.69}$$

So unless $m' = m$, the matrix elements of z are always zero.

Meanwhile, from the commutator of L_z with x we get

$$\langle n'l'm'|[L_z, x]|nlm\rangle = \langle n'l'm'|(L_z x - xL_z)|nlm\rangle$$

$$= (m' - m)\hbar\langle n'l'm'|x|nlm\rangle = i\hbar\langle n'l'm'|y|nlm\rangle.$$

Conclusion:

$$(m' - m)\langle n'l'm'|x|nlm\rangle = i\langle n'l'm'|y|nlm\rangle \tag{9.70}$$

So you never have to compute matrix elements of y—you can always get them from the corresponding matrix elements of x.

Finally, the commutator of L_z with y yields

$$\langle n'l'm'|[L_z, y]|nlm\rangle = \langle n'l'm'|(L_z y - yL_z)|nlm\rangle$$

$$= (m' - m)\hbar\langle n'l'm'|y|nlm\rangle = -i\hbar\langle n'l'm'|x|nlm\rangle.$$

Conclusion:

$$(m' - m)\langle n'l'm'|y|nlm\rangle = -i\langle n'l'm'|x|nlm\rangle \tag{9.71}$$

In particular, combining Equations 9.70 and 9.71,

$$(m' - m)^2\langle n'l'm'|x|nlm\rangle = i(m' - m)\langle n'l'm'|y|nlm\rangle = \langle n'l'm'|x|nlm\rangle,$$

and hence

$$\text{either } (m' - m)^2 = 1, \text{ or else } \langle n'l'm'|x|nlm\rangle = \langle n'l'm'|y|nlm\rangle = 0. \tag{9.72}$$

From Equations 9.69 and 9.72 we obtain the **selection rule** for m:

$$\boxed{\text{No transitions occur unless } \Delta m = \pm 1 \text{ or } 0.} \qquad [9.73]$$

This is an easy result to understand if you remember that the photon carries spin 1, and hence *its* value of m is 1, 0, or -1[17]; conservation of (the z-component of) angular momentum requires that the atom give up whatever the photon takes away.

Selection rules involving l and l': In Problem 9.11 you are asked to derive the following commutation relation:

$$\left[L^2, [L^2, \mathbf{r}]\right] = 2\hbar^2(\mathbf{r}L^2 + L^2\mathbf{r}). \qquad [9.74]$$

As before, we sandwich this commutator between $\langle n'l'm'|$ and $|nlm\rangle$ to derive the selection rule:

$$\langle n'l'm'|\left[L^2, [L^2, \mathbf{r}]\right]|nlm\rangle = 2\hbar^2\langle n'l'm'|(\mathbf{r}L^2 + L^2\mathbf{r})|nlm\rangle$$

$$= 2\hbar^4[l(l+1) + l'(l'+1)]\langle n'l'm'|\mathbf{r}|nlm\rangle = \langle n'l'm'|(L^2[L^2, \mathbf{r}] - [L^2, \mathbf{r}]L^2)|nlm\rangle$$

$$= \hbar^2[l'(l'+1) - l(l+1)]\langle n'l'm'|[L^2, \mathbf{r}]|nlm\rangle$$

$$= \hbar^2[l'(l'+1) - l(l+1)]\langle n'l'm'|(L^2\mathbf{r} - \mathbf{r}L^2)|nlm\rangle$$

$$= \hbar^4[l'(l'+1) - l(l+1)]^2\langle n'l'm'|\mathbf{r}|nlm\rangle. \qquad [9.75]$$

Conclusion:

$$\text{Either } 2[l(l+1) + l'(l'+1)] = [l'(l'+1) - l(l+1)]^2$$

$$\text{or else } \langle n'l'm'|\mathbf{r}|nlm\rangle = 0. \qquad [9.76]$$

But

$$[l'(l'+1) - l(l+1)] = (l' + l + 1)(l' - l)$$

and

$$2[l(l+1) + l'(l'+1)] = (l' + l + 1)^2 + (l' - l)^2 - 1,$$

so the first condition in Equation 9.76 can be written

$$[(l' + l + 1)^2 - 1][(l' - l)^2 - 1] = 0. \qquad [9.77]$$

---

[17] When the polar axis is along the direction of propagation, the middle value does not occur, and if you are only interested in the *number* of linearly independent photon states, the answer is 2, not 3. However, in this case the photon need not be going in the z-direction, and all three values are possible.

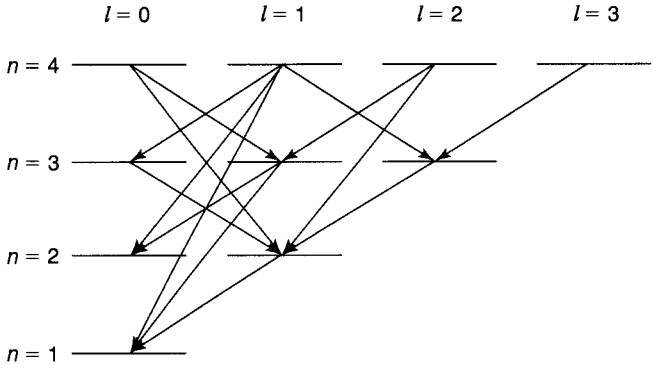

Figure 9.5: Allowed decays for the first four Bohr levels in hydrogen.

The first factor *cannot* be zero (unless $l' = l = 0$—this loophole is closed in Problem 9.12), so the condition simplifies to $l' = l \pm 1$. Thus we obtain the selection rule for l:

$$\boxed{\text{No transitions occur unless } \Delta l = \pm 1.} \qquad [9.78]$$

Again, this result (though scarcely trivial to *derive*) is easy to *interpret*: The photon carries spin 1, so the rules for addition of angular momentum would allow $l' = l + 1$, $l' = l$, or $l' = l - 1$ (for electric dipole radiation the middle possibility—though permitted by conservation of angular momentum—does not occur).

Evidently not all transitions to lower-energy states can proceed by spontaneous emission; some are forbidden by the selection rules. The scheme of allowed transitions for the first four Bohr levels in hydrogen is shown in Figure 9.5. Note that the $2S$ state (ψ_{200}) is "stuck": It cannot decay, because there is no lower-energy state with $l = 1$. 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, either by collisions or by what are (misleadingly) called **forbidden** transitions (Problem 9.20), or by multiphoton emission.

Problem 9.11 Prove the commutation relation Equation 9.74. *Hint*: First show that

$$[L^2, z] = 2i\hbar(xL_y - yL_x - i\hbar z).$$

Use this and the fact that $\mathbf{r} \cdot \mathbf{L} = \mathbf{r} \cdot (\mathbf{r} \times \mathbf{p}) = 0$ to demonstrate that

$$\left[L^2, [L^2, z]\right] = 2\hbar^2(zL^2 + L^2 z).$$

The generalization from z to $\mathbf{r}$ is trivial.

Problem 9.12 Plug the "loophole" in Equation 9.78 by showing that if $l' = l = 0$ then $\langle n'l'm'|\mathbf{r}|nlm\rangle = 0$.

Problem 9.13 An electron in the $n = 3$, $l = 0$, $m = 0$ 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:

$$|300\rangle \to |nlm\rangle \to |n'l'm'\rangle \to \ldots \to |100\rangle.$$

(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. When there is more than one decay route open, the transition *rates* add.

FURTHER PROBLEMS FOR CHAPTER 9

Problem 9.14 Develop time-dependent perturbation theory for a multilevel system, starting with the generalization of Equations 9.1 and 9.2:

$$H_0\psi_n = E_n\psi_n, \quad \langle\psi_n|\psi_m\rangle = \delta_{nm}. \tag{9.79}$$

At time $t = 0$ we turn on a perturbation $H'(t)$, so that the total Hamiltonian is

$$H = H_0 + H'(t). \tag{9.80}$$

(a) Generalize Equation 9.6 to read

$$\Psi(t) = \sum c_n(t)\psi_n e^{-iE_nt/\hbar}, \tag{9.81}$$

and show that

$$\dot{c}_m = -\frac{i}{\hbar}\sum_n c_n H'_{mn} e^{i(E_m - E_n)t/\hbar}, \tag{9.82}$$

where

$$H'_{mn} \equiv \langle\psi_m|H'|\psi_n\rangle. \tag{9.83}$$

(b) If the system starts out in the state ψ_N, show that (in first-order perturbation theory)

$$c_N(t) \cong 1 - \frac{i}{\hbar}\int_0^t H'_{NN}(t')\,dt' \tag{9.84}$$

and

$$c_m(t) \cong -\frac{i}{\hbar}\int_0^t H'_{mN}(t')e^{i(E_m-E_N)t'/\hbar}\,dt', \quad (m \neq N). \qquad [9.85]$$

(c) For example, suppose 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*:

$$4|H'_{mN}|^2\,\frac{\sin^2[(E_N - E_m)t/2\hbar]}{(E_N - E_m)^2}. \qquad [9.86]$$

(d) Now suppose H' is a sinusoidal function of time: $H' = V\cos(\omega t)$. 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

$$P_{N\to m} = |V_{mN}|^2\frac{\sin^2[(E_N - E_m \pm \hbar\omega)t/2\hbar]}{(E_N - E_m \pm \hbar\omega)^2}. \qquad [9.87]$$

(e) Suppose a multilevel system is immersed in incoherent electromagnetic radiation. Using Section 9.2.3 as a guide, show that the transition rate for stimulated emission is given by the same formula (Equation 9.47) as for a two-level system.

Problem 9.15 For the examples in Problem 9.14 (c) and (d), calculate $c_m(t)$, to first order. Check the normalization condition:

$$\sum_m |c_m(t)|^2 = 1, \qquad [9.88]$$

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 9.16 A particle starts out (at time $t = 0$) in the N^{th} state of the infinite square well. Now water leaks into the well, and then drains out again, so that the bottom is at uniform potential $V_0(t)$, with $V_0(0) = V_0(T) = 0$.

(a) Solve the *exact* equation (Equation 9.82) for $c_m(t)$, and show that the wave function changes *phase*, but no transitions to other states 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.

Note: 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. See Problem 1.13.

***Problem 9.17** 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

$$V(x) = \begin{cases} V_0, & \text{if } 0 \le x \le a/2, \\ 0, & \text{if } a/2 < x \le a, \\ \infty, & \text{otherwise,} \end{cases}$$

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 energy is now E_2.

Problem 9.18 Justify the following version of the energy-time uncertainty principle (due to Landau): $\Delta E \Delta t \ge \hbar/2$, where Δt is the time it takes to execute a transition involving an energy change ΔE, under the influence of a constant perturbation (see Problem 9.14c.) Explain more precisely what ΔE and Δt mean in this context.

***Problem 9.19** An electron is at rest at the origin, in the presence of a magnetic field whose *magnitude* (B_0) is constant but whose *direction* rides around at constant angular velocity ω on the lip of a cone of opening angle α:

$$\mathbf{B}(t) = B_0[\sin\alpha\cos(\omega t)\hat{\imath} + \sin\alpha\sin(\omega t)\hat{\jmath} + \cos\alpha\hat{k}]. \qquad [9.89]$$

(a) Construct the 2×2 Hamiltonian matrix (Equation 4.158) for this system.

(b) Find the *exact* solution to the (time-dependent) Schrödinger equation, assuming the particle starts out with spin up. *Hint*: You can do it from scratch, or by noting that in this case the rotating wave approximation is *exact*, and refering to Problem 9.7. *Answer*:

$$\chi(t) = \begin{pmatrix} [\cos(\lambda t/2) + i[(\omega + \omega_1\cos\alpha)/\lambda]\sin(\lambda t/2)]\, e^{-i\omega t/2} \\ i[(\omega_1\sin\alpha)/\lambda]\sin(\lambda t/2) e^{i\omega t/2} \end{pmatrix}, \qquad [9.90]$$

where

$$\omega_1 \equiv -eB_0/m \quad \text{and} \quad \lambda \equiv \sqrt{\omega^2 + \omega_1^2 + 2\omega\omega_1\cos\alpha}. \qquad [9.91]$$

(c) Now treat the same problem by (first-order) time-dependent perturbation theory: use Equation 9.17 to calculate the (approximate) probability of a transition from spin up (the initial state) to spin down, as a function of time, and compare the exact answer (from part b). State the criterion on the strength of the field that determines whether perturbation theory is applicable in this case.

***Problem 9.20** In Equation 9.31 we assumed that the atom is so small (in comparison to the wavelength of light) that spatial variations in the field can be ignored. The *true* electric field would be

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0\cos(\mathbf{k}\cdot\mathbf{r} - \omega t). \qquad [9.92]$$

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:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0[\cos(\omega t) + (\mathbf{k} \cdot \mathbf{r}) \sin(\omega t)]. \quad [9.93]$$

The first term gives rise to the **allowed** (**electric dipole**) transitions we considered in the text; the second gives rise 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[18]).

(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*:

$$R_{b \to a} = \frac{q^2 \omega^5}{\pi \epsilon_0 \hbar c^5} |\langle a|(\hat{n} \cdot \mathbf{r})(\hat{k} \cdot \mathbf{r})|b\rangle|^2. \quad [9.94]$$

(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

$$R = \frac{\hbar q^2 \omega^3 n(n-1)}{15\pi \epsilon_0 m^2 c^5}. \quad [9.95]$$

Find the *ratio* of the "forbidden" rate to the "allowed" rate, and comment on the terminology. (*Note*: w is the frequency of the *photon*, not the *oscillator*.)

(c) Show that the $2S \to 1S$ 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 a two-photon emission, and the lifetime is about a tenth of a second.[19])

Problem 9.21 We have encountered stimulated emission, (stimulated) absorption, and spontaneous emission ... how come there's no such thing as spontaneous *absorption*?

---

[18] For a systematic treatment (including the role of the magnetic field), see David Park, *Introduction to the Quantum Theory*, 3rd ed. (New York: McGraw-Hill, 1992), Chapter 11.

[19] See Masataka Mizushima, *Quantum Mechanics of Atomic Spectra and Atomic Structure*, New York: Benjamin, 1970), Section 5.6.

CHAPTER 10

THE ADIABATIC APPROXIMATION

10.1 THE ADIABATIC THEOREM

10.1.1 Adiabatic Processes

Imagine a perfect pendulum, with no friction or air resistance, oscillating back and forth in a vertical plane. If I grab the support and shake it in a jerky manner, the bob will swing around in a wild chaotic fashion. But if I *very gently and steadily* move the support (Figure 10.1), 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 in the external conditions* characterizes 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 an oscillating platform, for example, T_e would be the period of the *platform's* motion). An adiabatic process is one for which $T_e \gg T_i$.[1]

The basic strategy for analyzing an adiabatic process is first to solve the problem with the external parameters held *fixed*, and only at the *end* of the calculation allow them to change with time. For example, the classical period of a pendulum of (constant) length L is $2\pi\sqrt{L/g}$; if the length is now gradually *changing*, the

[1] For an interesting discussion of classical adiabatic processes, see Frank S. Crawford, *Am. J. Phys.* **58**, 337 (1990).

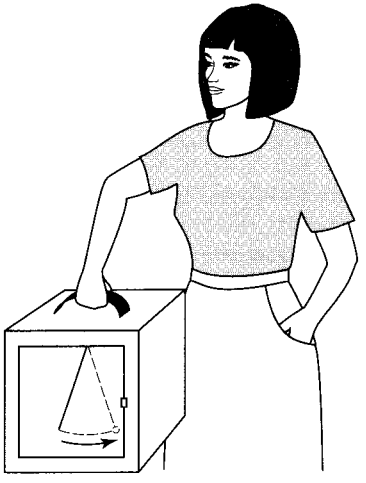

Figure 10.1: Adiabatic motion: If the case is transported very gradually, the pendulum inside keeps swinging with the same amplitude, in a plane parallel to the original one.

period will presumably be $2\pi\sqrt{L(t)/g}$. When you stop to think about it, we actually *use* the adiabatic approximation (implicitly) all the time without noticing it. A case in point is our discussion of the hydrogen molecule ion (Section 7.3). 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 7.10). 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**.

In quantum mechanics, the essential content of the **adiabatic approximation** can be cast in the form of a theorem. Suppose that the Hamiltonian changes *gradually* from some initial form H^i to some final form H^f (Figure 10.2). The **adiabatic theorem** states that if the particle was initially in the nth eigenstate of H^i, it will be

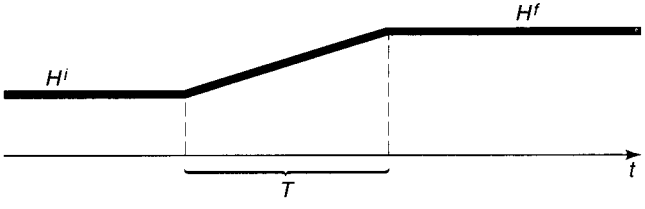

Figure 10.2: A model for adiabatic change in the Hamiltonian, from H^i to H^f.

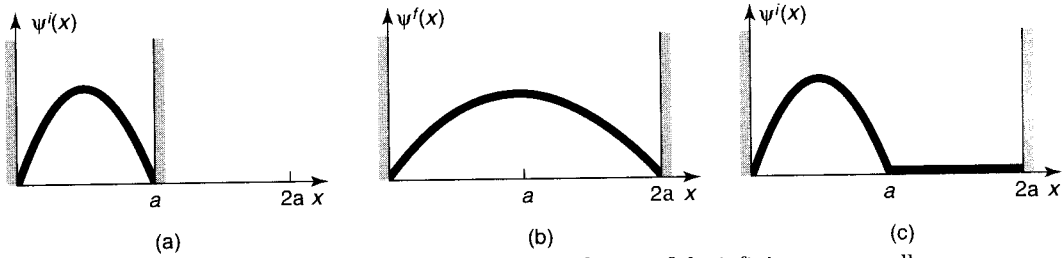

Figure 10.3: (a) Particle starts out in the ground state of the infinite square well. (b) If the wall moves *slowly*, the particle remains in the ground state. (c) If the wall moves *rapidly*, the particle is left (momentarily) in its initial state.

carried (under the Schrödinger equation) into the nth eigenstate of H^f. (I assume that the spectrum is discrete and nondegenerate throughout the transition from H^i to H^f, 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.)

For example, suppose we prepare a particle in the ground state of the infinite square well (Figure 10.3a):

$$\psi^i(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi}{a}x\right). \qquad [10.1]$$

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 10.3b):

$$\psi^f(x) = \sqrt{\frac{1}{a}} \sin\left(\frac{\pi}{2a}x\right) \qquad [10.2]$$

(apart, perhaps, from a phase factor). Notice that we're not talking about a *small* change in the Hamiltonian (as in perturbation theory)—this one is a *huge* change. All we require is that it happen *slowly*. By contrast, if the well expands *suddenly*, the resulting state is still $\psi^i(x)$ (Figure 10.3c), which is a complicated linear combination of eigenstates of the new Hamiltonian (Problem 3.48).

***Problem 10.1** The case of an infinite square well whose right wall expands at a *constant* velocity (v) can be solved *exactly*.[2] A complete set of solutions is

$$\Phi_n(x,t) \equiv \sqrt{\frac{2}{w}} \sin\left(\frac{n\pi}{w}x\right) e^{i(mvx^2 - 2E_n^i at)/2\hbar w}, \qquad [10.3]$$

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 nth allowed

---

[2] S. W. Doescher and M. H. Rice, *Am. J. Phys.* **37**, 1246 (1969).

energy of the *original* well (width a). The *general* solution is a linear combination of the Φ's:

$$\Psi(x,t) = \sum_{n=1}^{\infty} c_n \Phi_n(x,t); \qquad [10.4]$$

the coefficients c_n are *independent of* t.

(a) Check that Equation 10.3 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

$$c_n = \frac{2}{\pi} \int_0^{\pi} e^{-i\alpha z^2} \sin(nz) \sin(z)\, dz, \qquad [10.5]$$

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 $e^{-i\alpha z^2} \cong 1$ 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

$$\theta(t) = -\frac{1}{\hbar} \int_0^t E_1(t')\, dt', \qquad [10.6]$$

where $E_n(t) \equiv n^2\pi^2\hbar^2/2mw^2$ is the n^{th} *instantaneous* eigenvalue, at time t. Comment on this result.

10.1.2 Proof of the Adiabatic Theorem

The adiabatic theorem is simple to state, and it *sounds* plausible, but it is not easy to prove.[3] Suppose the time-dependent part of the Hamiltonian can be written in the form[4]

$$H'(t) = Vf(t), \qquad [10.7]$$

where $f(t)$ is a function that starts out zero (at $t = 0$) and increases to 1 (at $t = T$), Figure 10.4. Assume that the particle starts out in the n^{th} eigenstate of the original Hamiltonian:

$$\Psi(0) = \psi_n^i \qquad [10.8]$$

and evolves into some state $\Psi(t)$. Our problem is to show that if the function $f(t)$ rises *very gradually*, then the probability that the particle, at time T, is in the nth eigenstate of the *final* Hamiltonian (ψ_n^f) is 1. More precisely, we must demonstrate that

$$|\langle \Psi(T) | \psi_m^f \rangle|^2 = \begin{cases} 1, & \text{if } m = n, \\ 0, & \text{if } m \neq n. \end{cases} \qquad [10.9]$$

(Of course, if the first of these is true, the second *has* to be, and vice versa. But it is not clear at this stage which condition will be easier to prove.)

Assume for the moment that V is *small*, so we can use first-order time-independent perturbation theory to determine ψ_m^f. From Equation 6.12,

$$\psi_m^f \cong \psi_m + \sum_{k \neq m} \frac{V_{km}}{E_m - E_k} \psi_k, \qquad [10.10]$$

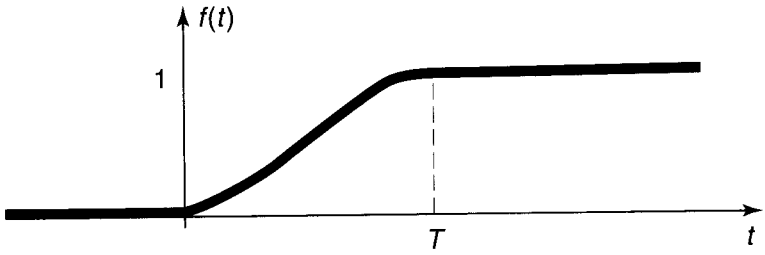

Figure 10.4: The function $f(t)$, in Equation 10.7.

[3]The adiabatic theorem is usually attributed to Ehrenfest, who studied adiabatic processes in early versions of the quantum theory. The first proof in modern quantum mechanics was given by Born and Fock, *Zeit. f. Physik* **51**, 165 (1928). Other proofs will be found in Messiah, *Quantum Mechanics* (New York: John Wiley & Sons, 1962), Vol. II, Chapter XVII, Section 12, and J-T. Hwang and Philip Pechukas, *J. Chem. Phys.* **67**, 4640 (1977). The argument given here is suggested by Gasiorowicz, *Quantum Physics* (New York: John Wiley & Sons, 1974), Chapter 22, Problem 6.

[4]The assumption that H' is the *product* of an operator (V) and a (real) function of t is not necessary for the theorem itself, but it does make the proof less cumbersome. In Section 10.1.3 we will encounter a case in which the different matrix elements of H' have *different* (complex) time dependences. As long as the adiabatic approximation (in the form of Equation 10.15) holds for *each* of them, the adiabatic theorem itself is valid.

where

$$V_{km} \equiv \langle \psi_k | V | \psi_m \rangle. \tag{10.11}$$

(To simplify the notation, I'll drop the superscript i on eigenfunctions and eigenvalues of the initial Hamiltonian; these are the "unperturbed" states for the problem.)

Meanwhile, we use first-order time-*dependent* perturbation theory to determine $\Psi(T)$. From Equation 9.81,

$$\Psi(t) = \sum_l c_l(t) \psi_l e^{-iE_l t/\hbar}, \tag{10.12}$$

where (Equation 9.84)

$$c_n(t) \cong 1 - \frac{i}{\hbar} V_{nn} \int_0^t f(t')\, dt' \tag{10.13}$$

and (Equation 9.85)

$$c_l(t) \cong -\frac{i}{\hbar} V_{ln} \int_0^t f(t') e^{i(E_l - E_n)t'/\hbar}\, dt', \quad l \neq n. \tag{10.14}$$

This last integral can be evaluated using integration by parts. Note that

$$e^{i(E_l - E_n)t'/\hbar} = \frac{-i\hbar}{E_l - E_n} \frac{d}{dt'} \left[e^{i(E_l - E_n)t'/\hbar} \right],$$

so

$$c_l(t) \cong -\frac{V_{ln}}{E_l - E_n} \int_0^t f(t') \frac{d}{dt'} \left[e^{i(E_l - E_n)t'/\hbar} \right] dt'$$

$$= -\frac{V_{ln}}{E_l - E_n} \left\{ f(t) e^{i(E_l - E_n)t/\hbar} - \int_0^t \frac{df}{dt'} e^{i(E_l - E_n)t'/\hbar}\, dt' \right\}.$$

[I dropped the lower limit in the first term, because $f(0) = 0$.] Now comes the adiabatic approximation: We want $f(t)$ to be a *very gradual* function, so that df/dt is extremely small. Specifically, we assume that

$$\frac{df}{dt} \ll \frac{|E_l - E_n|}{\hbar} f; \tag{10.15}$$

then the last term makes a negligible contribution to $c_l(t)$, and we conclude that

$$\Psi(T) \cong \left[\left(1 - i \frac{V_{nn} A}{\hbar} \right) \psi_n - \sum_{l \neq n} \frac{V_{ln}}{E_l - E_n} \psi_l \right] e^{-iE_n T/\hbar}, \tag{10.16}$$

where A is the area under the graph of $f(t)$, from 0 to T.

Putting together Equations 10.10 and 10.16, and exploiting the orthonormality of the initial eigenfunctions, we find that

$$\langle \Psi(T)|\psi_n^f\rangle = \left[1 + i\frac{V_{nn}A}{\hbar} + \sum_{k\neq n}\frac{|V_{kn}|^2}{(E_n - E_k)^2}\right] e^{iE_nT/\hbar}, \qquad [10.17]$$

while, for $m \neq n$,

$$\langle \Psi(T)|\psi_m^f\rangle = \left\{\left[1 + i\frac{V_{nn}A}{\hbar}\right]\frac{V_{nm}}{E_m - E_n} - \frac{V_{nm}}{E_m - E_n} + \sum_{n\neq k\neq m}\frac{V_{nk}V_{km}}{(E_n - E_k)(E_m - E_k)}\right\} e^{iE_nT/\hbar}$$

$$= \left[\frac{iAV_{nn}V_{nm}}{\hbar(E_m - E_n)} + \sum_{n\neq k\neq m}\frac{V_{nk}V_{km}}{(E_n - E_k)(E_m - E_k)}\right] e^{iE_nT/\hbar}. \qquad [10.18]$$

But wait: These wave functions were only accurate to *first* order in V, so the *second*-order terms in Equations 10.17 and 10.18 are spurious (we have already thrown away quantities of comparable size). To *first* order, we have

$$\langle \Psi(T)|\psi_m^f\rangle = \begin{cases} \left[1 + i\frac{V_{nn}A}{\hbar}\right] e^{iE_nT/\hbar}, & m = n, \\ 0 & m \neq n. \end{cases} \qquad [10.19]$$

It follows that

$$|\langle \Psi(T)|\psi_n^f\rangle|^2 = 1, \qquad [10.20]$$

while (for $m \neq n$)

$$|\langle \Psi(T)|\psi_m^f\rangle|^2 = 0. \qquad [10.21]$$

Ostensibly, *either* of these would suffice to establish the desired result (Equation 10.9). However, Equation 10.20 is only accurate to *first* order (in V), whereas Equation 10.21 is accurate to *second* order (and for that matter to *third* order as well).[5] In truth, Equation 10.20 tells us *nothing* (it would be valid also for a *nonadiabatic* transformation); the *crucial* point is the cancellation of the first-order terms in Equation 10.18, for this tells us that there will be no transitions to other states.[6]

---

[5] See Problem 9.15 for a discussion of the analogous situation in ordinary perturbation theory.

[6] In this context the word "*transition*" means from an eigenstate ψ_n^i of the *initial* Hamiltonian (H^i) to a different eigenstate ψ_m^f of the *final* Hamiltonian (H^f). The adiabatic theorem says that if the Hamiltonian changes *gradually* from H^i to H^f, there will be *no* such transitions. By contrast, in the previous chapter we were always dealing with eigenstates of the *same* (unperturbed) Hamiltonian. At the end of the process the perturbation was (explicitly or implicitly) turned off, and a "transition" meant from one eigenstate of the unperturbed Hamiltonian to another eigenstate of the unperturbed Hamiltonian. The transition amplitudes were of *first* order in H' (Equations 9.17 and 9.85) and the transition probabilities of *second* order (for example, Equations 9.28, 9.86, and 9.87). The essence of the adiabatic theorem (as we shall see in the next paragraph) is that the transition amplitudes are only *second* order, and the transition probabilities *fourth* order in the (small) perturbation.

This shows that if the change in the Hamiltonian is both adiabatic *and* very small (so that first-order perturbation theory can be applied), then there will be no transitions. But what if the change, though gradual, is *not* small? In that case we chop the interval T into N subintervals, so that the change in the Hamiltonian during a single subinterval (ΔV) is of order V/N; if N it large, then ΔV is small, and we can apply the previous argument to each subinterval. If the transition amplitude (Equation 10.18) were *first* order in the perturbation, then the *total* transition amplitude would go like

$$N\left(\frac{V}{N}\right) \to V \qquad [10.22]$$

(N steps, each making a contribution proportional to ΔV). The net result would be of order V, and if V is large, so too would be the transition amplitude. But in *fact* the transition amplitude is *second* order, so the total goes like

$$N\left(\frac{V}{N}\right)^2 \to \frac{V^2}{N} \qquad [10.23]$$

In the limit as $N \to \infty$, the transition amplitude goes to zero, *regardless* of the size of V. QED

Problem 10.2 In the beginning of this chapter, I characterized an adiabatic process informally as one for which $T_e \gg T_i$. How is this related to the precise condition (Equation 10.15) required in the proof (in other words, what are T_e and T_i here)?

10.1.3 An Example

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 10.5):

$$\mathbf{B}(t) = B_0[\sin\alpha\cos(\omega t)\hat{\imath} + \sin\alpha\sin(\omega t)\hat{\jmath} + \cos\alpha\hat{k}. \qquad [10.24]$$

The Hamiltonian (Equation 4.158) is

$$H(t) = \frac{e}{m}\mathbf{B}\cdot\mathbf{S} = \frac{e\hbar B_0}{2m}[\sin\alpha\cos(\omega t)\sigma_x + \sin\alpha\sin(\omega t)\sigma_y + \cos\alpha\sigma_z]$$

$$= -\frac{\hbar\omega_1}{2}\begin{pmatrix} \cos\alpha & e^{-i\omega t}\sin\alpha \\ e^{i\omega t}\sin\alpha & -\cos\alpha \end{pmatrix}, \qquad [10.25]$$

where

$$\omega_1 \equiv -\frac{eB_0}{m}. \qquad [10.26]$$

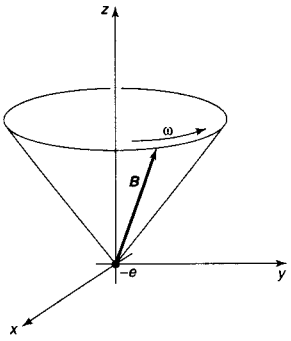

Figure 10.5: Magnetic field sweeps around on a cone, at angular velocity ω, Equation 10.24.

The normalized eigenspinors of $H(t)$ are

$$\chi_+(t) = \begin{pmatrix} \cos(\alpha/2) \\ e^{i\omega t}\sin(\alpha/2) \end{pmatrix} \quad [10.27]$$

and

$$\chi_-(t) = \begin{pmatrix} \sin(\alpha/2) \\ -e^{i\omega t}\cos(\alpha/2) \end{pmatrix}; \quad [10.28]$$

they represent spin up and spin down, respectively, *along the instantaneous direction of* $\mathbf{B}(t)$ (see Problem 4.31). The corresponding eigenvalues are

$$E_\pm = \mp\frac{\hbar\omega_1}{2}. \quad [10.29]$$

Suppose the electron starts out with spin up, along $\mathbf{B}(0)$:

$$\chi(0) = \begin{pmatrix} \cos(\alpha/2) \\ \sin(\alpha/2) \end{pmatrix}. \quad [10.30]$$

The exact solution to the time-dependent Schrödinger equation is (Problem 10.3)

$$\chi(t) = \begin{pmatrix} \left[\cos(\lambda t/2) + i\frac{(\omega_1+\omega)}{\lambda}\sin(\lambda t/2)\right]\cos(\alpha/2)e^{-i\omega t/2} \\ \left[\cos(\lambda t/2) + i\frac{(\omega_1-\omega)}{\lambda}\sin(\lambda t/2)\right]\sin(\alpha/2)e^{i\omega t/2} \end{pmatrix}, \quad [10.31]$$

where

$$\lambda \equiv \sqrt{\omega^2 + \omega_1^2 + 2\omega\omega_1\cos\alpha}, \quad [10.32]$$

or, writing it as a linear combination of χ_+ and χ_-,

$$\chi(t) = \left[\cos\left(\frac{\lambda t}{2}\right) + i\frac{(\omega_1 + \omega\cos\alpha)}{\lambda}\sin\left(\frac{\lambda t}{2}\right)\right]e^{-i\omega t/2}\chi_+(t) + i\left[\frac{\omega}{\lambda}\sin\alpha\sin\left(\frac{\lambda t}{2}\right)\right]e^{-i\omega t/2}\chi_-(t). \quad [10.33]$$

Evidently the (exact) probability of a transition to spin down (along the current direction of **B**) is

$$|\langle\chi(t)|\chi_-(t)\rangle|^2 = \left[\frac{\omega}{\lambda}\sin\alpha\sin\left(\frac{\lambda t}{2}\right)\right]^2. \qquad [10.34]$$

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 $\lambda \cong \omega_1$, and therefore

$$|\langle\chi(t)|\chi_-(t)\rangle|^2 \cong \left[\frac{\omega}{\omega_1}\sin\alpha\sin\left(\frac{\lambda t}{2}\right)\right]^2 \to 0, \qquad [10.35]$$

as advertised. The magnetic field leads the electron around by its nose, with the spin always pointing in the direction of **B**. By contrast, if $\omega \gg \omega_1$ then $\lambda \cong \omega$, and the system bounces back and forth between spin up and spin down (Figure 10.6).

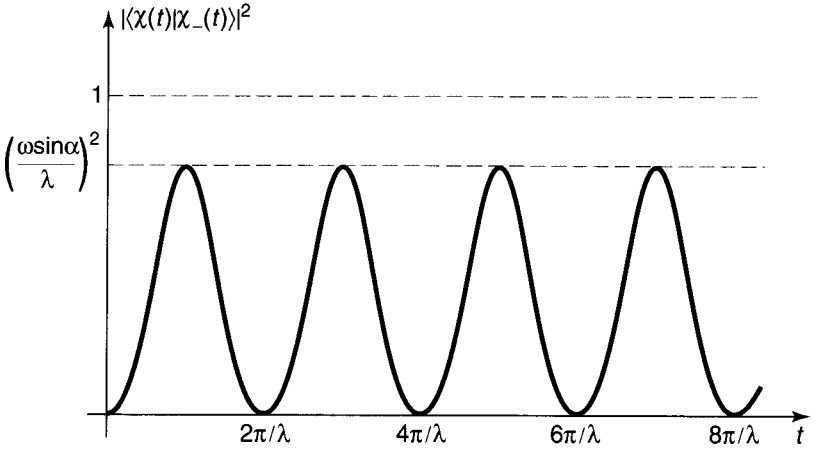

Figure 10.6: Plot of the transition probability, Equation 10.34, in the *nonadiabatic* regime ($\omega \ll \omega_1$).

Problem 10.3 Check that Equation 10.31 satisfies the time-dependent Schrödinger equation for the Hamiltonian (Equation 10.25). *Note*: This is the same as Problem 9.19(b), except that now the electron starts out with spin up along **B**, whereas in Equation 9.90 it started out with spin up along z. Also confirm Equation 10.33, and show that the sum of the squares of the coefficients is 1, as required for proper normalization.

10.2 Berry's Phase

10.2.1 Nonholonomic Processes

Let us return to the classical model I used (in Section 10.1.1) to develop the notion of an *adiabatic* process: a perfectly frictionless pendulum whose support is carried around from place to place. I claimed that as long as the motion of the support is *very slow*, compared to the period of the pendulum (so that the pendulum executes many oscillations before the support has moved appreciably), it will continue to swing in the same plane (or one parallel to it), with the same amplitude (and, of course, the same frequency).

But what if I took this ideal pendulum up to the North Pole, and set it swinging—say, in the direction of Portland (Figure 10.7). (For the moment, I'll pretend the earth is not rotating.) Very gently (that is, *adiabatically*), I carry it down the longitude line passing through Portland, and on beyond, down to the equator. At this stage it is swinging north-south. Now I carry it (still swinging north-south) partway around the equator. And finally, I carry it back up to the North Pole, along the new longitude line. It is clear that 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. Now Θ is equal to the *solid angle* (Ω) subtended (at the center of the earth) by the path around which I carried the pendulum. For this path surrounds a fraction $\Theta/2\pi$ of the northern hemisphere, so its area is $A = (1/2)(\Theta/2\pi)4\pi R^2 = \Theta R^2$ (where R is the radius of the earth), and hence

$$\Theta = A/R^2 \equiv \Omega. \qquad [10.36]$$

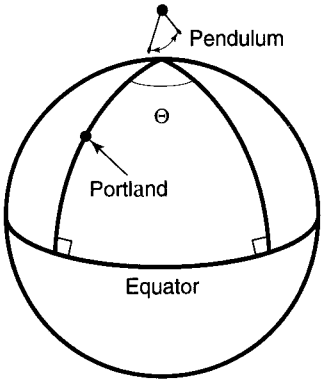

Figure 10.7: Itinerary for adiabatic transport of a pendulum on the surface of the earth.

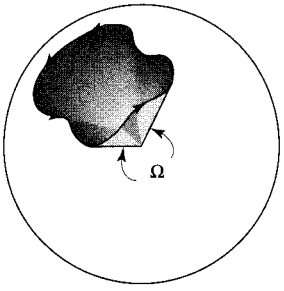

Figure 10.8: Arbitrary path on the surface of a sphere, subtending a solid angle Ω.

This is a particularly nice way to express the answer, because it turns out to be independent of the *shape* of the path (Figure 10.8).[7]

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 10.9) is

$$\Omega = \int \sin\theta \, d\theta d\phi = 2\pi(-\cos\theta)\Big|_0^{\theta_0} = 2\pi(1 - \cos\theta_0). \qquad [10.37]$$

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,[8] but is seen in this context to admit a purely *geometrical* interpretation.

A system such as this, which does not return to its original state when transported around a closed loop, is said to be **nonholonomic**. (The "transport" in question need not involve physical *motion*: What we have in mind is that the external parameters of

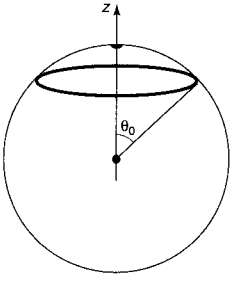

Figure 10.9: Path of a Foucault pendulum in the course of one day.

[7] You can prove this for yourself, if you are interested. Think of the circuit as being made up of tiny segments of great circles (geodesics on the sphere); the pendulum makes a fixed angle with each geodesic segment, so the net angular deviation is related to the sum of the vertex angles of the spherical polygon.

[8] See, for example, Jerry B. Marion, *Classical Dynamics*, 2nd ed. (New York: Academic Press, 1970), Section 11.4. Geographers measure latitude (λ) up from the equator, rather than down from the pole, so $\cos\theta_0 = \sin\lambda$.

the system are changed in some fashion that eventually returns them to their initial values.) Nonholonomic systems are ubiquitous—in a sense, every cyclical engine is a nonholonomic device: At the end of each cycle the car has moved forward a bit, or a weight has been lifted slightly, or something. The idea has even been applied to the locomotion of microbes in fluids at low Reynolds number.[9] My project for the next section is to study the *quantum mechanics of nonholonomic, adiabatic processes*. The essential question is this: How does the final state differ from the initial state, if the parameters in the Hamiltonian are carried adiabatically around some closed cycle?

10.2.2 Geometric Phase

If the Hamiltonian is *independent* of time, then a particle which starts out in the nth eigenstate $\psi_n(x)$,

$$H\psi_n(x) = E_n\psi_n(x),$$

remains in the n^{th} eigenstate, simply picking up a phase factor:

$$\Psi_n(x,t) = \psi_n(x)e^{-iE_nt/\hbar}. \qquad [10.38]$$

If the Hamiltonian *changes* with time, then the eigenfunctions and eigenvalues themselves are time dependent:

$$H(t)\psi_n(x,t) = E_n(t)\psi_n(x,t). \qquad [10.39]$$

But the adiabatic theorem tells us that when H changes very *gradually*, a particle which starts out in the nth eigenstate will *remain* in the n^{th} eigenstate—picking up at most a time-dependent phase factor—even as the eigenfunction itself evolves. That is to say,

$$\Psi_n(x,t) = \psi_n(x,t)e^{-\frac{i}{\hbar}\int_0^t E_n(t')dt'}e^{i\gamma_n(t)}. \qquad [10.40]$$

The term

$$\theta_n(t) \equiv -\frac{1}{\hbar}\int_0^t E_n(t')\,dt' \qquad [10.41]$$

is known as the **dynamic phase**; it generalizes the "standard" factor $(-E_nt/\hbar)$ to the case where E_n is a function of time. (You will have encountered dynamical phase factors already, if you worked Problems 9.16 and and 10.1.) Any "extra" phase, $\gamma_n(t)$, is called the **geometric phase**. At the moment we don't know what it *is*, or what physical significance (if any) it carries; all we can say is that the adiabatic theorem does not *rule out* such a factor, since the particle is still "in the n^{th} eigenstate", *whatever* the value of γ_n. [More precisely, a measurement of the energy at time t would be certain to return the value $E_n(t)$.] Indeed, since the eigenvalue equation (Equa-

[9]The pendulum example is an application of **Hannay's angle**, which is the classical analog to Berry's phase. For a collection of papers on both subjects, see Alfred Shapere and Frank Wilczek, eds., *Geometric Phases in Physics*, (Singapore: World Scientific, 1989).

tion 10.39) and the normalization condition only determine $\psi_n(x, t)$ up to an arbitrary phase, and since this arbitrary phase could in principle be chosen independently at each instant of time [though in practice we shall always take $\psi_n(x, t)$ to be a *smooth* function of t], we *have* to allow for an arbitrary phase factor in Equation [10.40].[10] Notice, incidentally, that energy is not conserved here. Of *course* not: Whoever is changing the Hamiltonian is pumping energy into or out of the system.

If we plug Equation 10.40 into the (time-dependent) Schrödinger equation,

$$i\hbar\frac{\partial\Psi}{\partial t} = H(t)\Psi, \qquad [10.42]$$

there emerges a simple formula for the time development of the geometric phase:

$$i\hbar\left[\frac{\partial\psi_n}{\partial t}e^{i\theta_n}e^{i\gamma_n} - \frac{i}{\hbar}E_n\psi_n e^{i\theta_n}e^{i\gamma_n} + i\frac{d\gamma_n}{dt}\psi_n e^{i\theta_n}e^{i\gamma_n}\right]$$
$$= [H\psi_n]e^{i\theta_n}e^{i\gamma_n} = E_n\psi_n e^{i\theta_n}e^{i\gamma_n},$$

whence

$$\frac{\partial\psi_n}{\partial t} + i\psi_n\frac{d\gamma_n}{dt} = 0. \qquad [10.43]$$

Taking the inner product with ψ_n (which I assume has been normalized), we obtain

$$\frac{d\gamma_n}{dt} = i\langle\psi_n|\frac{\partial\psi_n}{\partial t}\rangle. \qquad [10.44]$$

Now $\psi_n(x, t)$ depends on t because there is some parameter $R(t)$ in the Hamiltonian that is changing with time. [In Problem 10.1, $R(t)$ would be the width of the infinite square well, whose right wall is expanding.] Thus

$$\frac{\partial\psi_n}{\partial t} = \frac{\partial\psi_n}{\partial R}\frac{dR}{dt}, \qquad [10.45]$$

so that

$$\frac{d\gamma_n}{dt} = i\langle\psi_n|\frac{\partial\psi_n}{\partial R}\rangle\frac{dR}{dt},$$

and hence

$$\gamma_n(t) = i\int_0^t\langle\psi_n|\frac{\partial\psi_n}{\partial R}\rangle\frac{dR}{dt'}\,dt' = i\int_{R_i}^{R_f}\langle\psi_n|\frac{\partial\psi_n}{\partial R}\rangle\,dR, \qquad [10.46]$$

where R_i and R_f are the initial and final values of $R(t)$. In particular, if the Hamiltonian returns to its original form after time T, so that $R_f = R_i$, then $\gamma_n(T) = 0$ —nothing very interesting *there*!

---

[10] For this reason, most people assumed until quite recently that the geometric phase was of no conceivable physical significance. It was Michael Berry's inspiration to realize that if you carry the Hamiltonian around a closed *cycle*, bringing it back to its original form at the end, the relative phase at the beginning and at the end of the process is a *nonarbitrary* quantity, with profound physical implications.

However, I assumed (in Equation 10.45) that there is only *one* parameter in the Hamiltonian that is changing. Suppose there are N of them: $R_1(t)$, $R_2(t)$, ..., $R_N(t)$; in that case

$$\frac{\partial \psi_n}{\partial t} = \frac{\partial \psi_n}{\partial R_1}\frac{dR_1}{dt} + \frac{\partial \psi_n}{\partial R_2}\frac{dR_2}{dt} + \cdots + \frac{\partial \psi_n}{\partial R_N}\frac{dR_N}{dt} = (\nabla_R \psi_n) \cdot \frac{d\mathbf{R}}{dt}, \quad [10.47]$$

where $\mathbf{R} \equiv (R_1, R_n, \ldots, R_N)$, and ∇_R is the gradient with respect to these parameters. This time we have

$$\gamma_n(t) = i \int_{\mathbf{R}_i}^{\mathbf{R}_f} \langle \psi_n | \nabla_R \psi_n \rangle \cdot d\mathbf{R}, \quad [10.48]$$

and if the Hamiltonian returns to its original form after a time T, the net geometric phase change is

$$\boxed{\gamma_n(T) = i \oint \langle \psi_n | \nabla_R \psi_n \rangle \cdot d\mathbf{R}.} \quad [10.49]$$

This is a *line* integral around a closed loop in parameter space, and it is *not*, in general, zero. Equation 10.49 was first obtained by Berry in 1984,[11] and $\gamma_n(T)$ is called **Berry's phase**. Notice that $\gamma_n(T)$ depends *only on the path taken*, not on how *fast* that path is traversed (provided, of course, that it is slow enough to validate the adiabatic hypothesis). By contrast, the accumulated *dynamic* phase,

$$\theta_n(T) = -\frac{1}{\hbar}\int_0^T E_n(t')\,dt',$$

depends critically on the elapsed time.

The derivation of Berry's phase raises several questions, which I would like to address before turning to some examples and applications.

1. Is $\gamma_n(t)$ real? If it's *not*, then $e^{i\gamma_n}$ is not a phase factor at all, but an *exponential* factor, and the normalization of Ψ_n (in Equation 10.40) is lost. Since the time-dependent Schrödinger equation conserves probability, it must preserve normalization. It would be comforting to *check* this, by showing explicitly that Equation 10.48 yields a real γ_n. In fact, this is very easy to do. First note that

$$\nabla_R \langle \psi_n | \psi_n \rangle = 0 \quad [10.50]$$

(because by assumption ψ_n is normalized). So

$$\langle \nabla_R \psi_n | \psi_n \rangle + \langle \psi_n | \nabla_R \psi_n \rangle = \langle \psi_n | \nabla_R \psi_n \rangle^* + \langle \psi_n | \nabla_R \psi_n \rangle = 0.$$

Since $\langle \psi_n | \nabla_R \psi_n \rangle$ plus its complex conjugate is zero, it follows that

$$\langle \psi_n | \nabla_R \psi_n \rangle \quad \text{is imaginary,} \quad [10.51]$$

and hence, by Equation 10.48, $\gamma_n(t)$ is real. [Incidentally, if ψ_n itself is *real*, then so

[11] M. V. Berry, *Proc. R. Soc. Lond.* A **392**, 45 (1984), reprinted in Shapere and Wilczek, (footnote 9). It is astonishing, in retrospect, that such a fundamental result escaped notice for 60 years.

(obviously) is $\langle\psi_n|\nabla_R\psi_n\rangle$—this quantity is both real and imaginary, and it must therefore be *zero*. Evidently the geometric phase vanishes whenever the eigenfunctions (of $H(t)$) are real.]

2. Is Berry's phase measurable? We are accustomed to thinking that the phase of the wave function is arbitrary—physical quantities involve $|\Psi|^2$, and the phase factor cancels out. But $\gamma_n(T)$ *can* be measured, if (for example) we take a beam of particles (all in the state Ψ) and split it in two, so that one beam passes through an adiabatically changing potential, while the other does not. When the two beams are recombined, the total wave function has the form

$$\Psi = \frac{1}{2}\Psi_0 + \frac{1}{2}\Psi_0 e^{i\Gamma}, \qquad [10.52]$$

where Ψ_0 is the "direct" beam wave function, and Γ is the *extra* phase (in part dynamic, and in part geometric) acquired by the beam subjected to the varying H). In this case

$$|\Psi|^2 = \frac{1}{4}|\Psi_0|^2\left(1 + e^{i\Gamma}\right)\left(1 + e^{-i\Gamma}\right)$$

$$= \frac{1}{2}|\Psi_0|^2(1 + \cos\Gamma) = |\Psi_0|^2\cos^2(\Gamma/2). \qquad [10.53]$$

So by looking for points of constructive and destructive interference (where Γ is an even or odd multiple of π, respectively), one can easily measure Γ. (Berry, and other early writers, worried that the geometric phase might be swamped by a larger dynamic phase, but it has proved possible to arrange things so as to separate out the two contributions.)

3. Where does the derivation invoke the adiabatic hypothesis? At first glance, the argument going from Equation 10.40 to Equation 10.48 appears to have proved altogether *too* much. Why doesn't the derivation work *in reverse*, showing that as long as $\gamma_n(t)$ is given by Equation 10.48, the expression in Equation 10.40 satisfies the Schrödinger equation *exactly—whether or not the process is adiabatic*?(!) (This would, of course, be nonsense; it would imply that the adiabatic theorem is empty: No transitions *ever* occur, even if the change in the Hamiltonian is far from gradual.) The answer is that the step following Equation 10.43, in which we take the inner product, *cannot* in general be reversed: Although Equation 10.43 implies Equation 10.44, Equation 10.44 does *not* imply Equation 10.43. In fact, there is a serious fraud in the derivation, which I did not confess at the time because it somewhat spoils the beauty of the argument. The truth is that although Equation 10.44 is correct, Equation 10.43 is *not*.[12] For Equation 10.40 is only true in the extreme adiabatic limit—the

---

[12]Indeed, if you take Equation 10.43 at face value, it can be solved directly for γ_n:

$$\frac{d\gamma_n}{dt} = i\frac{\partial}{\partial t}(\ln\psi_n) \Rightarrow \psi_n(x,t) = \phi_n(x)e^{-i\gamma_n(t)},$$

and hence (going back to Equation 10.40),

$$\Psi_n(x,t) = \phi_n(x)e^{-\frac{i}{\hbar}\int_0^t E_n(t')dt'}.$$

The geometric phase, in effect, *soaks up* the time dependence acquired by the eigenfunction $\psi_n(x,t)$ as a consequence of the change in H. But this is completely false, as we shall see in the examples.

exact solution would contain admixtures of other states:

$$\Psi_n(x,t) = \psi_n(x,t)e^{i\theta_n(t)}e^{i\gamma_n(t)} + \epsilon \sum_{m\neq n} c_m(t)\psi_m(x,t), \qquad [10.54]$$

where $\epsilon \equiv T_i/T_e$ characterizes the departure from adiabaticity (it goes to zero in the adiabatic limit). Inclusion of this term modifies Equation 10.43, to read

$$\frac{\partial \psi_n}{\partial t} + i\psi_n \frac{d\gamma_n}{dt} = -e^{-i\theta_n}e^{-i\gamma_n}\epsilon \sum_{m\neq n}\left[\left(\frac{i}{\hbar}c_m E_m + \frac{dc_m}{dt}\right)\psi_m + c_m \frac{\partial \psi_m}{\partial t}\right]. \qquad [10.55]$$

Both terms on the left are first order in ϵ (if the Hamiltonian didn't change at *all*, both $\partial\psi_n/\partial t$ and γ_n would be zero), but *so are the first two terms on the right.* The *final* term is *second* order, so it can legitimately be ignored, but dropping the first two (as I did, implicitly, in my derivation of Equation 10.43), is illegal. For consistency (noting, while I'm at it, that γ_n is already first order, so $e^{i\gamma_n} \cong 1$ on the right), I *should* have written

$$\frac{\partial \psi_n}{\partial t} + i\psi_n \frac{d\gamma_n}{dt} = -e^{-i\theta_n}\epsilon \sum_{m\neq n}\left(\frac{i}{\hbar}c_m E_m + \frac{dc_m}{dt}\right)\psi_m, \qquad [10.56]$$

instead of Equation 10.43. Fortunately, the inner product (with ψ_n) kills the extra term, and that's how it comes about that Equation 10.44 is correct, even though Equation 10.43, from which it was obtained, was *not.* (See Problem 10.7.)

When the parameter space is three dimensional, $\mathbf{R} = (R_1, R_2, R_3)$, Berry's formula (Equation 10.49) is reminiscent of the expression for **magnetic flux** in terms of the vector potential $\mathbf{A}$. The flux, Φ, through a surface S bounded by a curve C (Figure 10.10), is

$$\Phi \equiv \int_S \mathbf{B}\cdot d\mathbf{a}. \qquad [10.57]$$

If we write the magnetic field in terms of the vector potential ($\mathbf{B} = \nabla \times \mathbf{A}$), and apply Stokes' theorem:

$$\Phi = \int_S (\nabla \times \mathbf{A})\cdot d\mathbf{a} = \oint_C \mathbf{A}\cdot d\mathbf{r}. \qquad [10.58]$$

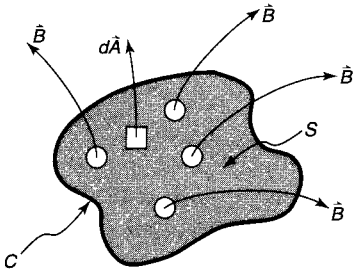

Figure 10.10: Magnetic flux through a surface S bounded by the closed curve C.

Thus Berry's phase can be thought of as the "flux" of a "magnetic field"

$$\text{"}\mathbf{B}\text{"} = i\nabla_R \times \langle \psi_n | \nabla_R \psi_n \rangle, \qquad [10.59]$$

through the (closed-loop) trajectory in parameter space. In the three-dimensional case, then, Berry's phase can be written as a surface integral,

$$\gamma_n(T) = i \int [\nabla_R \times \langle \psi_n | \nabla_R \psi_n \rangle] \cdot d\mathbf{a}. \qquad [10.60]$$

The magnetic analogy can be carried much further, but for our purposes Equation 10.60 is merely a convenient alternative expression for $\gamma_n(T)$.

***Problem 10.4**

(a) Use Equation 10.46 to calculate the geometric phase change when the infinite square well expands adiabatically from width w_1 to width w_2. Comment on this result.

(b) If the expansion occurs at a constant rate ($dw/dt = v$), what is the dynamic phase change for this process?

(c) If the well now contracts back to its original size, what is Berry's phase for the cycle?

Problem 10.5 The delta-function well (Equation 2.96) supports a single bound state (Equation 2.111). Calculate the geometric phase change when α gradually increases from α_1 to α_2. If the increase occurs at a constant rate ($d\alpha/dt = c$), what is the dynamic phase change for this process?

Problem 10.6 As I noted in the text (and Problems 10.4 and 10.5 confirm), if $\psi_n(x, t)$ is *real*, the geometric phase vanishes. You might try to beat this rap by tacking an unnecessary (but perfectly legal) phase factor onto the eigenfunctions: $\psi_n'(x, t) \equiv e^{i\phi_n}\psi_n(x, t)$, where $\phi_n(\mathbf{R})$ 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 10.40. And for a closed loop it gives *zero*. *Moral*: For nonzero Berry's phase, you need (1) more than one time-dependent parameter in the Hamiltonian, and (2) a Hamiltonian that yields nontrivially complex eigenfunctions.

10.2.3 An Example

The classic example of Berry's phase is an electron at the origin, subjected to a magnetic field of constant magnitude but changing direction. Consider first the special case (analyzed in Section 10.1.3) in which $\mathbf{B}(t)$ precesses around at a constant angular

velocity ω, making a fixed angle α with the z-axis. The *exact* solution (for an electron that starts out with "spin up" along **B**) is given by Equation 10.33. In the adiabatic regime, $\omega \ll \omega_1$,

$$\lambda = \omega_1\sqrt{1 + 2\frac{\omega}{\omega_1}\cos\alpha + \left(\frac{\omega}{\omega_1}\right)^2} \cong \omega_1\left(1 + \frac{\omega}{\omega_1}\cos\alpha\right) = \omega_1 + \omega\cos\alpha, \quad [10.61]$$

and Equation 10.33 becomes

$$\chi(t) \cong e^{i\omega_1 t/2}e^{i(\omega\cos\alpha)t/2}e^{-i\omega t/2}\chi_+(t) + i\left[\frac{\omega}{\omega_1}\sin\alpha\sin\left(\frac{\omega_1 t}{2}\right)\right]e^{-i\omega t/2}\chi_-(t). \quad [10.62]$$

As $\omega/\omega_1 \to 0$ the second term drops out completely, and the result matches the expected adiabatic form (Equation 10.40). The dynamic phase is

$$\theta_+(t) = -\frac{1}{\hbar}\int_0^t E_+(t')\,dt' = \frac{\omega_1 t}{2} \quad [10.63]$$

(where $E_+ = -\hbar\omega_1/2$ is taken from Equation 10.29), so the geometric phase is

$$\gamma_+(t) = (\cos\alpha - 1)\frac{\omega t}{2}. \quad [10.64]$$

For a complete cycle $T = 2\pi/\omega$, and therefore Berry's phase is

$$\gamma_+(T) = \pi(\cos\alpha - 1). \quad [10.65]$$

Now consider the more general case, in which the tip of the magnetic field vector sweeps out an *arbitrary* closed curve on the surface of a sphere of radius $r = B_0$ (Figure 10.11). The eigenstate representing spin up along $\mathbf{B}(t)$ has the form (see Problem 4.31)

$$\chi_+ = \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\sin(\theta/2) \end{pmatrix}, \quad [10.66]$$

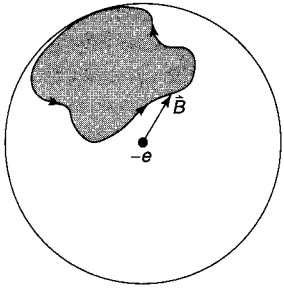

Figure 10.11: Magnetic field of constant magnitude but changing direction sweeps out a closed loop.

where θ and ϕ (the spherical coordinates of $\mathbf{B}$) are now *both* functions of time. Looking up the gradient in spherical coordinates, we find

$$\nabla\chi_+ = \frac{\partial\chi_+}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial\chi_+}{\partial\theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial\chi_+}{\partial\phi}\hat{\phi}$$

$$= \frac{1}{r}\begin{pmatrix} -(1/2)\sin(\theta/2) \\ (1/2)e^{i\phi}\cos(\theta/2) \end{pmatrix}\hat{\theta} + \frac{1}{r\sin\theta}\begin{pmatrix} 0 \\ ie^{i\phi}\sin(\theta/2) \end{pmatrix}\hat{\phi}. \quad [10.67]$$

Hence

$$\langle\chi_+|\nabla\chi_+\rangle = \frac{1}{2r}\Big[-\sin(\theta/2)\cos(\theta/2)\,\hat{\theta} + \sin(\theta/2)\cos(\theta/2)\,\hat{\theta}$$

$$+ 2i\frac{\sin^2(\theta/2)}{\sin\theta}\,\hat{\phi}\Big] = i\frac{\sin^2(\theta/2)}{r\sin\theta}\,\hat{\phi}. \quad [10.68]$$

For Equation 10.60 we need the *curl* of this quantity:

$$\nabla\times\langle\chi_+|\nabla\chi_+\rangle = \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\left(\frac{i\sin^2(\theta/2)}{r\sin\theta}\right)\right]\hat{r} = \frac{i}{2r^2}\hat{r}. \quad [10.69]$$

According to Equation 10.60, then,

$$\gamma_+(T) = -\frac{1}{2}\int\frac{1}{r^2}\,\hat{r}\cdot d\mathbf{a}. \quad [10.70]$$

The integral is over the area on the sphere swept out by $\mathbf{B}$ in the course of the cycle, so $d\mathbf{a} = r^2\,d\Omega\,\hat{r}$, and hence

$$\gamma_+(T) = -\frac{1}{2}\int d\Omega = -\frac{1}{2}\Omega, \quad [10.71]$$

where Ω is the solid angle subtended at the origin. This is a delightfully simple result, and tantalizingly reminiscent of the classical problem with which we began the discussion (transport of a frictionless pendulum around a closed path on the surface of the earth). It says that if you take a magnet, and lead the electron's spin around adiabatically in an arbitrary closed path, the net (geometric) phase change will be minus one half the solid angle swept out by the magnetic field vector. In view of Equation 10.37, the general result is consistent with the special case Equation 10.65, as of course it *had* to be.

Problem 10.7 Consider, once again, the special case of the precessing field (Section 10.1.3).

(a) Use the eigenspinor (Equation 10.27) to determine $\langle\chi_+|(\partial\chi_+/\partial t)\rangle$, and put the result into Equation 10.44, for an alternative derivation of Equation 10.64.

(b) Show that Equation 10.43 does *not* work, in this case. Use Equation 10.62 to determine c_- (in Equation 10.54). Confirm that the last term in Equation 10.55 is second order in ω (don't forget the $\epsilon = \omega/\omega_1$ out front). Show that $\gamma_+(t)$ (Equation 10.64) does satisfy the *corrected* version of Equation 10.43, Equation 10.56.

***Problem 10.8** Work out the analog to Equation 10.71 for a particle of spin 1. *Answer*: $-\Omega$ (for spin s the result is $-s\Omega/2$).

10.2.4 The Aharonov-Bohm Effect

In classical electrodynamics the potentials (φ and $\mathbf{A}$)[13] are not directly measurable—the physical quantities are the electric and magnetic fields:

$$\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}. \tag{10.72}$$

The fundamental laws of the theory (Maxwell's equations and the Lorentz force law) make no reference to potentials, which are (from a logical point of view) no more than convenient but dispensible scaffolding for getting a better purchase on the real structure (the fields). Indeed, you're perfectly free to *change* the potentials:

$$\varphi \to \varphi' = \varphi - \frac{\partial \Lambda}{\partial t}, \quad \mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla\Lambda, \tag{10.73}$$

where Λ is an arbitrary function of position and time; this is called a gauge transformation, and it has no effect at all on the fields.

In quantum mechanics the potentials play a more significant role, for the Hamiltonian (Equation 4.201) is expressed in terms of φ and $\mathbf{A}$, not $\mathbf{E}$ and $\mathbf{B}$:

$$H = \frac{1}{2m}\left(\frac{\hbar}{i}\nabla - q\mathbf{A}\right)^2 + q\varphi. \tag{10.74}$$

Nevertheless, the theory is still invariant under gauge transformations (see Problem 4.53), and it was taken for granted until quite recently that there could be no electromagnetic influences in regions where $\mathbf{E}$ and $\mathbf{B}$ are zero—any more than there can be in the classical theory. But in 1959 Aharonov and Bohm[14] showed that the vector potential *can* affect the quantum behavior of a charged particle that never encounters an electromagnetic field. I'll work out a simple example first, then discuss the

---

[13] I'm sorry, but we have reached a notational impasse: It is customary in quantum mechanics to use the letter V for *potential energy*, but in electrodynamics the same letter is reserved for the scalar potential. To avoid confusion I'll use φ for the scalar potential. See Problems 4.51, 4.52, and 4.53 for background on this material.

[14] Y. Aharonov and D. Bohm, *Phys. Rev.* **115**, 485 (1959). For a significant precursor, see W. Ehrenberg and R. E. Siday, *Proc. Phys. Soc. London* **B62**, 8 (1949).

Aharanov-Bohm effect itself, and finally indicate how it can be thought of as an example of Berry's phase.

Imagine a particle constrained to move in a circle of radius b (a bead on a wire ring, if you like). Along the axis runs a solenoid of radius $a < b$, carrying a magnetic field **B** (see Figure 10.12). If the solenoid is extremely long, the field inside is uniform, and the field outside is zero. But the vector potential outside the solenoid is *not* zero; in fact (adopting the convenient gauge condition $\nabla \cdot \mathbf{A} = 0$),[15]

$$\mathbf{A} = \frac{\Phi}{2\pi r}\hat{\phi}, \quad (r > a), \tag{10.75}$$

where $\Phi = \pi a^2 B$ is the **magnetic flux** through the solenoid. Meanwhile, the solenoid is uncharged, so the scalar potential φ is zero. In this case the Hamiltonian (Equation 10.74) becomes

$$H = \frac{1}{2m}\left[-\hbar^2\nabla^2 + q^2A^2 + 2i\hbar q\mathbf{A}\cdot\nabla\right]. \tag{10.76}$$

But the wave function depends only on the azimuthal angle ϕ, ($\theta = \pi/2$ and $r = b$) so $\nabla \to (\hat{\phi}/b)(d/d\phi)$, and the Schrödinger equation reads

$$\frac{1}{2m}\left[-\frac{\hbar^2}{b^2}\frac{d^2}{d\phi^2} + \left(\frac{q\Phi}{2\pi b}\right)^2 + i\frac{\hbar q\Phi}{\pi b^2}\frac{d}{d\phi}\right]\psi(\phi) = E\psi(\phi). \tag{10.77}$$

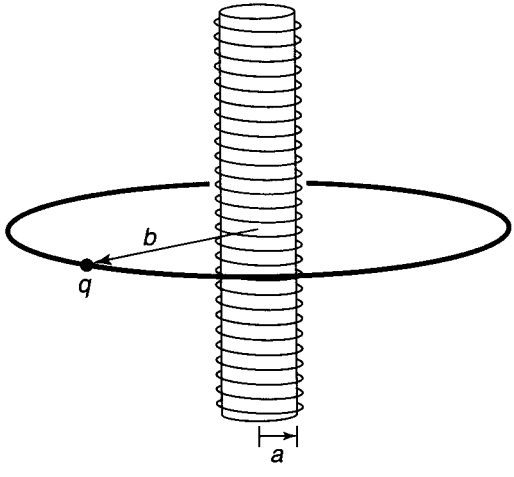

Figure 10.12: Charged bead on a circular ring through which a long solenoid passes.

This is a linear differential equation with constant coefficients:

$$\frac{d^2\psi}{d\phi^2} - 2i\beta\frac{d\psi}{d\phi} + \epsilon\psi = 0, \tag{10.78}$$

[15] See, for instance, D. J. Griffiths, *Introduction to Electrodynamics*, 2nd ed. (Englewood Cliffs, NJ: Prentice Hall, 1989), Equation 5.65.

where

$$\beta \equiv \frac{q\Phi}{2\pi\hbar} \quad \text{and} \quad \epsilon \equiv \frac{2mb^2E}{\hbar^2} - \beta^2. \tag{10.79}$$

Solutions are of the form

$$\psi = Ae^{i\lambda\phi}, \tag{10.80}$$

with

$$\lambda = \beta \pm \sqrt{\beta^2 + \epsilon} = \beta \pm \frac{b}{\hbar}\sqrt{2mE}. \tag{10.81}$$

Continuity of $\psi(\phi)$, at $\phi = 2\pi$, requires that λ be an *integer*:

$$\beta \pm \frac{b}{\hbar}\sqrt{2mE} = n, \tag{10.82}$$

and it follows that

$$E_n = \frac{\hbar^2}{2mb^2}\left(n - \frac{q\Phi}{2\pi\hbar}\right)^2, \quad (n = 0, \pm 1, \pm 2, \ldots). \tag{10.83}$$

The solenoid lifts the twofold degeneracy of the bead on a ring (Problem 2.43): Positive n, representing a particle traveling in the *same* direction as the current in the solenoid, has a somewhat *lower* energy (assuming q is positive) than negative n, describing a particle traveling in the *opposite* direction. And, more important, the allowed energies clearly depend on the field inside the solenoid, *even though the field at the location of the particle is zero.*[16]

More generally, suppose a particle is moving through a region where **B** is zero (so $\nabla \times \mathbf{A} = 0$), but **A** itself is *not*. (I'll assume that **A** is static, although the method can be generalized to time-dependent potentials.) The (time-dependent) Schrödinger equation,

$$\left[\frac{1}{2m}\left(\frac{\hbar}{i}\nabla - q\mathbf{A}\right)^2 + V\right]\Psi = i\hbar\frac{\partial\Psi}{\partial t}, \tag{10.84}$$

with potential energy V—which may or may not include an electrical contribution $q\varphi$—can be simplified by writing

$$\Psi = e^{ig}\Psi', \tag{10.85}$$

where

$$g(\mathbf{r}) \equiv \frac{q}{\hbar}\int_{\mathcal{O}}^{\mathbf{r}} \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}', \tag{10.86}$$

[16]It is a peculiar property of **superconducting** rings that the enclosed flux is *quantized*: $\Phi = (2\pi\hbar/q)n'$, where n' is an integer. In this case the effect is undetectable, since $E_n = (\hbar^2/2mb^2)(n + n')^2$, and $(n + n')$ is just another integer. (Incidentally, the charge q here turns out to be *twice* the charge of an electron; the superconducting electrons are locked together in pairs.) However, **flux quantization** is enforced by the *superconductor* (which induces circulating currents to make up the difference), not by the solenoid or the electromagnetic field, and it does not occur in the example considered here.

and $\mathcal{O}$ is some (arbitrarily chosen) reference point. Note that this definition makes sense *only* when $\nabla \times \mathbf{A} = 0$ throughout the region in question—otherwise the line integral depends entirely on the *path* taken from $\mathcal{O}$ to $\mathbf{r}$, and hence does not define a function of $\mathbf{r}$. In terms of Ψ', the gradient of Ψ is

$$\nabla\Psi = e^{ig}(i\nabla g)\Psi' + e^{ig}(\nabla\Psi');$$

but $\nabla g = (q/\hbar)\mathbf{A}$, so

$$\left(\frac{\hbar}{i}\nabla - q\mathbf{A}\right)\Psi = \frac{\hbar}{i}e^{ig}\nabla\Psi', \qquad [10.87]$$

and it follows that

$$\left(\frac{\hbar}{i}\nabla - q\mathbf{A}\right)^2\Psi = -\hbar^2 e^{ig}\nabla^2\Psi'. \qquad [10.88]$$

Putting this into Equation 10.84, and canceling the common factor of e^{ig}, we are left with

$$-\frac{\hbar^2}{2m}\nabla^2\Psi' + V\Psi' = i\hbar\frac{\partial\Psi'}{\partial t}. \qquad [10.89]$$

Evidently Ψ' satisfies the Schrödinger equation *without* $\mathbf{A}$. If we can solve Equation 10.89, correcting for the presence of a (curl-free) vector potential is trivial: You just tack on the phase factor e^{ig}.

Aharonov and Bohm proposed an experiment in which a beam of electrons is split in two, and passed either side of a long solenoid, before being recombined (Figure 10.13). The beams are kept well away from the solenoid itself, so they encounter only regions where $\mathbf{B} = 0$. But $\mathbf{A}$, which is given by Equation 10.75, is *not* zero, and (assuming V is the same on both sides), the two beams arrive with *different phases*:

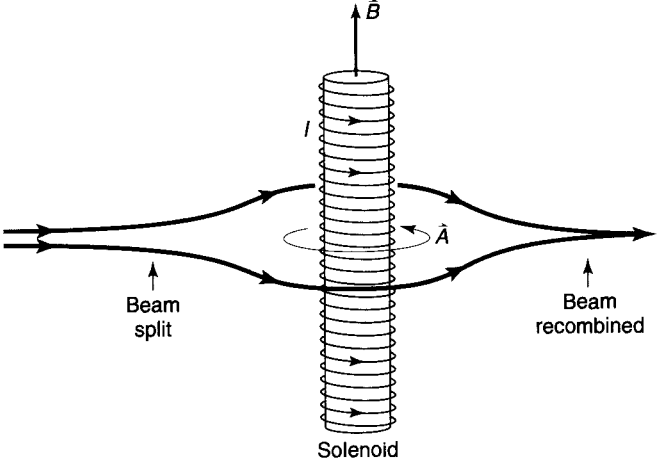

Figure 10.13: The Aharonov-Bohm effect: electron beam splits, with half passing either side of a long solenoid.

$$g = \frac{q}{\hbar}\int \mathbf{A}\cdot d\mathbf{r} = \frac{q\Phi}{2\pi\hbar}\int\left(\frac{1}{r}\hat{\phi}\right)\cdot(r\hat{\phi}\,d\phi) = \pm\frac{q\Phi}{2\hbar}. \qquad [10.90]$$

The plus sign applies to the electrons traveling in the same direction as **A**—which is to say, in the same direction as the current in the solenoid. The beams arrive *out of phase* by an amount proportional to the magnetic flux their paths encircle:

$$\text{phase difference} = \frac{q\Phi}{\hbar}. \qquad [10.91]$$

This phase shift leads to measurable interference (as in Equation 10.53), and has been confirmed experimentally by Chambers and others.[17]

The Aharonov-Bohm effect can be regarded as an example of geometric phase, as Berry himself noted in his first paper. Suppose the charged particle is confined to a box (which is centered at point **R** outside the solenoid) by a potential $V(\mathbf{r}-\mathbf{R})$—see Figure 10.14. (In a moment we're going to transport the box around the solenoid, so **R** will become a function of time, but for now it is just some fixed vector.) The eigenfunctions of the Hamiltonian are determined by

$$\left\{\frac{1}{2m}\left[\frac{\hbar}{i}\nabla - q\mathbf{A}(\mathbf{r})\right]^2 + V(\mathbf{r}-\mathbf{R})\right\}\psi_n = E_n\psi_n. \qquad [10.92]$$

We have already learned how to solve equations of this form:

$$\psi_n = e^{ig}\psi_n', \qquad [10.93]$$

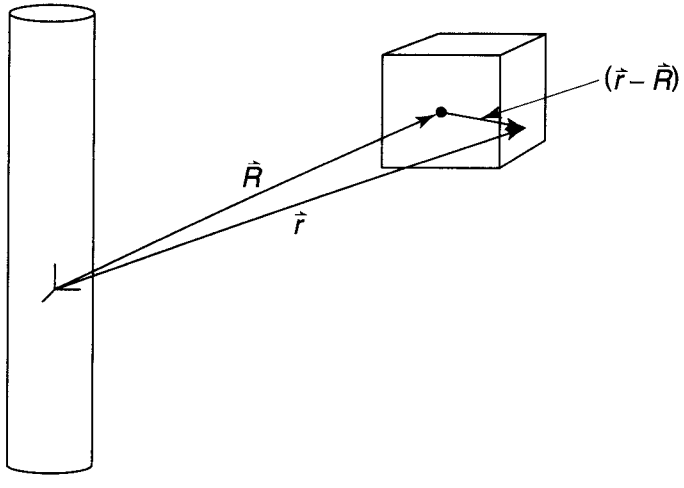

Figure 10.14: Particle confined to a box, by a potential $V(\mathbf{r}-\mathbf{R})$.

[17] R. G. Chambers, *Phys. Rev. Lett.* **5**, 3 (1960).

where[18]

$$g \equiv \frac{q}{\hbar}\int_{\mathbf{R}}^{\mathbf{r}} \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}' \qquad [10.94]$$

and ψ' satisfies the same eigenvalue equation, only with $\mathbf{A} \to 0$:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}-\mathbf{R})\right]\psi_n' = E_n\psi_n'. \qquad [10.95]$$

Notice that ψ_n' is a function only of the combination $(\mathbf{r}-\mathbf{R})$, not (like ψ_n) of $\mathbf{r}$ and $\mathbf{R}$ separately.

Now let's carry the box around the solenoid (in this case the process doesn't even have to be adiabatic). To determine Berry's phase, we must first evaluate the quantity $\langle\psi_n|\nabla_R\psi_n\rangle$. Noting that

$$\nabla_R\psi_n = \nabla_R\left[e^{ig}\psi_n'(\mathbf{r}-\mathbf{R})\right] = -i\frac{q}{\hbar}\mathbf{A}(\mathbf{R})e^{ig}\psi_n'(\mathbf{r}-\mathbf{R}) + e^{ig}\nabla_R\psi_n'(\mathbf{r}-\mathbf{R}),$$

we find

$$\begin{aligned}\langle\psi_n|\nabla_R\psi_n\rangle &= \int e^{-ig}[\psi_n'(\mathbf{r}-\mathbf{R})]^* e^{ig}\left[-i\frac{q}{\hbar}\mathbf{A}(\mathbf{R})\psi_n'(\mathbf{r}-\mathbf{R}) + \nabla_R\psi_n'(\mathbf{r}-\mathbf{R})\right]d^3\mathbf{r} \\ &= -i\frac{q}{\hbar}\mathbf{A}(\mathbf{R}) - \int[\psi_n'(\mathbf{r}-\mathbf{R})]^*\nabla\psi_n'(\mathbf{r}-\mathbf{R})\,d^3\mathbf{r}. \qquad [10.96]\end{aligned}$$

The ∇ with no subscript denotes the gradient with respect to $\mathbf{r}$, and I used the fact that $\nabla_R = -\nabla$, when acting on a function of $(\mathbf{r}-\mathbf{R})$. But the last integral is $i/\hbar$ times the expectation value of momentum, in an eigenstate of the Hamiltonian $-(\hbar^2/2m)\nabla^2 + V$, which we know from Section 2.1 is *zero*. So

$$\langle\psi_n|\nabla_R\psi_n\rangle = -i\frac{q}{\hbar}\mathbf{A}(\mathbf{R}). \qquad [10.97]$$

Putting this into Berry's formula (Equation 10.49), we conclude that

$$\gamma_n(T) = \frac{q}{\hbar}\oint\mathbf{A}(\mathbf{R})\cdot d\mathbf{R} = \frac{q}{\hbar}\int(\nabla\times\mathbf{A})\cdot d\mathbf{a} = \frac{q\Phi}{\hbar}, \qquad [10.98]$$

which neatly confirms the Aharonov-Bohm result (Equation 10.91), and reveals that the Aharonov-Bohm effect is a particular instance of geometric phase.[19]

---

[18]It is convenient to set the reference point $\mathcal{O}$ at the center of the box, for this guarantees that we recover the original phase convention for ψ_n when we complete the journey around the solenoid. If you use a fixed point in space, for example, you'll have to readjust the phase "by hand", at the far end; this leads to exactly the same answer, but it's a crude way to do it. In general, when choosing the phase convention for the eigenfunctions in Equation 10.39, you want to make sure that $\psi_n(x, T) = \psi_n(x, 0)$ so that no spurious phase changes are introduced.

[19]Incidentally, in this case the analogy between Berry's phase and magnetic flux (Equation 10.59) becomes *almost* an identity: "$\mathbf{B}$" = $(q/\hbar)\mathbf{B}$.

What are we to make of the Aharonov-Bohm effect? Evidently our classical preconceptions are simply *mistaken*: There *can* be electromagnetic effects in regions where the fields are zero. Note, however, that this does not make **A** itself measurable—only the enclosed *flux* comes into the final answer, and the theory remains gauge invariant.

Problem 10.9

(a) Derive Equation 10.76, from Equation 10.74.

(b) Derive Equation 10.88, starting with Equation 10.87.

FURTHER PROBLEMS FOR CHAPTER 10

* * ***Problem 10.10** 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; notice that $f(t)$ has the dimensions of *length*]. The Hamiltonian is

$$H(t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}m\omega^2 x^2 - m\omega^2 x f(t). \qquad [10.99]$$

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.[20]

(a) Determine the *classical* position of the oscillator, assuming it started out at rest at the origin [$x_c(0) = \dot{x}_c(0) = 0$]. *Answer*:

$$x_c(t) = \omega \int_0^t f(t') \sin[\omega(t - t')]\, dt'. \qquad [10.100]$$

(b) Show that the solution to the (time-dependent) Schrödinger equation for this oscillator, assuming it started out in the nth state of the *undriven* oscillator [$\Psi(x, 0) = \psi_n(x)$, where $\psi_n(x)$ is given by Equation 2.50], can be written as

$$\Psi(x, t) = \psi_n(x - x_c) e^{\frac{i}{\hbar}\left[-(n+\frac{1}{2})\hbar\omega t + m\dot{x}_c(x - \frac{x_c}{2}) + \frac{m\omega^2}{2}\int_0^t f(t')x_c(t')dt'\right]}. \qquad [10.101]$$

(c) Show that the eigenfunctions and eigenvalues of $H(t)$ are

$$\psi_n(x, t) = \psi_n(x - f); \quad E_n(t) = \left(n + \frac{1}{2}\right)\hbar\omega - \frac{1}{2}m\omega^2 f^2. \qquad [10.102]$$

[20] See Y. Nogami, *Am. J. Phys.* **59**, 64 (1991) and references therein.

(d) Show that in the adiabatic approximation the classical position (Equation 10.100) reduces to $x_c(t) \cong f(t)$. *Hint*: Use the integration-by-parts trick of Section 10.1.2. State the precise criterion—analogous to Equation 10.15—for adiabaticity.

(e) Confirm the adiabatic theorem for this example, by using the results in (c) and (d) to show that

$$\Psi(x, t) \cong \psi_n(x, t)e^{i\theta_n(t)}e^{i\gamma_n(t)}. \qquad [10.103]$$

Check that the dynamic phase has the correct form (Equation 10.41). Is the geometric phase what you would expect?

***Problem 10.11** In time-dependent perturbation theory, we used the completeness of the unperturbed eigenfunctions (of H_0) to expand $\Psi(x, t)$ (see Equation 9.81). But we could as well use the instantaneous eigenfunctions of $H(t)$ (Equation 10.39), since they, too, are complete:

$$\Psi(x, t) = \sum_n c_n(t)\psi_n(x, t)e^{i\theta_n}, \qquad [10.104]$$

where $\theta_n(t)$ is given by Equation 10.41. We can use this expansion to develop an **adiabatic series**, whose leading term is the adiabatic approximation itself and whose successive terms represent the *departure* from perfect adiabaticity.

(a) Insert Equation 10.104 into the (time-dependent) Schrödinger equation, and obtain the following formula for the coefficients:

$$\dot{c}_m = -\sum_n \langle\psi_m|\frac{\partial\psi_n}{\partial t}\rangle c_n e^{i(\theta_n-\theta_m)}. \qquad [10.105]$$

(b) Suppose the system starts out in the N^{th} state; in the adiabatic approximation, it *remains* in the N^{th} state, picking up (at most) a time-dependent geometric phase (compare Equations 10.40 and 10.104):

$$c_n(t) = \delta_{nN}e^{i\gamma_N(t)}. \qquad [10.106]$$

Substitute this into the right side of Equation 10.105, and obtain the "first correction" to adiabaticity:

$$c_m(t) = c_m(0) - \int_0^t \langle\psi_m|\frac{\partial\psi_N}{\partial t'}\rangle e^{i\gamma_N}e^{i(\theta_N-\theta_m)}\,dt'. \qquad [10.107]$$

This enables us to calculate transition probabilities in the *nearly* adiabatic regime. To develop the "second correction," we would insert Equation 10.107 on the right side of Equation 10.105, and so on.

(c) As an example, apply Equation 10.107 to the driven-oscillator (Problem 10.10). Show that (in the near-adiabatic approximation) transitions are possible only to the two immediately adjacent levels, for which

$$c_{N+1}(t) = i\sqrt{\frac{m\omega}{2\hbar}}\sqrt{N+1}\int_0^t \dot{f}(t')e^{i\omega t'}\,dt',$$

$$c_{N-1}(t) = i\sqrt{\frac{m\omega}{2\hbar}}\sqrt{N}\int_0^t \dot{f}(t')e^{-i\omega t'}\,dt'.$$

(The transition *probabilities* are the absolute squares of these, of course.)

CHAPTER 11

SCATTERING

11.1 INTRODUCTION

11.1.1 Classical Scattering Theory

Imagine a particle incident on some scattering center (say, a proton fired at a heavy nucleus). It comes in with an energy E and an **impact parameter** b, and it emerges at some **scattering angle** θ—see Figure 11.1. (I'll assume for simplicity that the target is azimuthally symmetrical, so the trajectory remains in one plane, and that the target is very heavy, so the recoil is negligible.) The essential problem of classical scattering

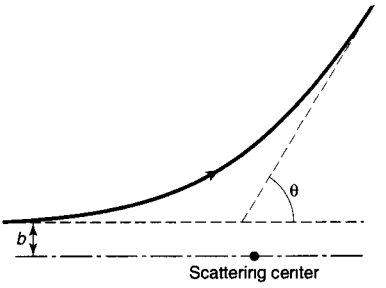

Figure 11.1 The classical scattering problem, showing the impact parameter b and the scattering angle θ.

theory is this: *Given the impact parameter, calculate the scattering angle.* Ordinarily, of course, the smaller the impact parameter, the greater the scattering angle.

Example: 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 11.2). In terms of the angle α, the impact parameter is $b = R \sin\alpha$, and the scattering angle is $\theta = \pi - 2\alpha$, so

$$b = R \sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right) = R \cos\left(\frac{\theta}{2}\right). \qquad [11.1]$$

Evidently

$$\theta = \begin{cases} 2\cos^{-1}(b/R), & \text{if } b \leq R, \\ 0, & \text{if } b \geq R. \end{cases} \qquad [11.2]$$

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 11.3). 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]:

$$\boxed{d\sigma = D(\theta)\, d\Omega.} \qquad [11.3]$$

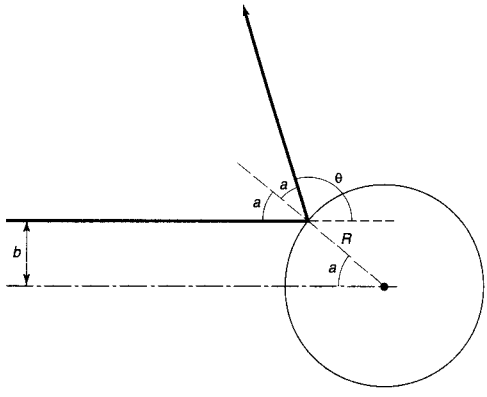

Figure 11.2: Elastic hard-sphere scattering.

[1]This is terrible language: D isn't a *differential*—if anything, it's a *derivative*—and it isn't a cross-section. To my ear, the words "differential cross-section" would apply more properly 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 $d\sigma/d\Omega$, but I think it will be less confusing if we give the differential cross-section its own symbol.

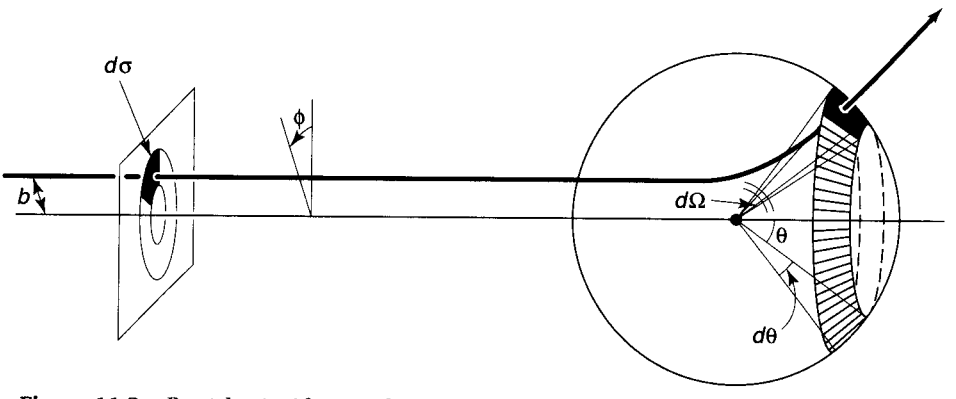

Figure 11.3: Particles incident in the area $d\sigma$ scatter into the solid angle $d\Omega$.

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

$$D(\theta) = \frac{b}{\sin\theta}\left|\frac{db}{d\theta}\right|. \qquad [11.4]$$

(Since θ is typically a *decreasing* function of b, the derivative is actually negative—hence the absolute value sign.)

Example: Hard-sphere scattering (continued). In the case of hard-sphere scattering (Equation 11.1),

$$\frac{db}{d\theta} = -\frac{1}{2}R\sin\left(\frac{\theta}{2}\right), \qquad [11.5]$$

so

$$D(\theta) = \frac{R\cos(\theta/2)}{\sin\theta}\left(\frac{R\sin(\theta/2)}{2}\right) = \frac{R^2}{4}. \qquad [11.6]$$

This example is unusual in that the differential cross-section is actually independent of θ.

The **total cross-section** is the *integral* of $D(\theta)$ over all solid angles:

$$\boxed{\sigma \equiv \int D(\theta)\, d\Omega;} \qquad [11.7]$$

roughly speaking, it is the total area of incident beam that is scattered by the target. For example, in the case of the hard sphere,

$$\sigma = (R^2/4)\int d\Omega = \pi R^2, \qquad [11.8]$$

which is just what we would expect: It's the cross-sectional area of the sphere; BBs 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):

$$\mathcal{L} \equiv \text{number of incident particles per unit area, per unit time.} \qquad [11.9]$$

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

$$D(\theta) = \frac{1}{\mathcal{L}}\frac{dN}{d\Omega}. \qquad [11.10]$$

This is often taken as the *definition* of the differential cross-section, because it makes reference only to quantities easily measured in the laboratory: If the detector accepts particles scattering into a solid angle $d\Omega$, we simply count the *number* recorded, per unit time, divide by $d\Omega$, and normalize to the luminosity of the incident beam.

***Problem 11.1** Consider the problem of **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. *Note*: This is not easy, and you might want to refer to a book on classical mechanics, such as Jerry B. Marion, *Classical Dynamics of Particles and Systems*, 2nd ed. (New York: Academic Press, (1970)), Section 9.5. *Answer*: $b = (q_1q_2/8\pi\epsilon_0 E)\cot(\theta/2)$.

(b) Determine the differential scattering cross-section. *Answer*:

$$D(\theta) = \left[\frac{q_1q_2}{16\pi\epsilon_0 E\sin^2(\theta/2)}\right]^2. \qquad [11.11]$$

(c) Show that the total cross-section for Rutherford scattering is *infinite*.

11.1.2 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 11.4).[2] That is, we will look for solutions to the Schrödinger equation of the general form

$$\psi(r,\theta) \approx A\left\{e^{ikz} + f(\theta)\frac{e^{ikr}}{r}\right\}, \quad \text{for large } r. \qquad [11.12]$$

(The spherical wave must carry 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:

$$k \equiv \frac{\sqrt{2mE}}{\hbar}. \qquad [11.13]$$

(As before, I shall assume the target is azimuthally symmetrical; in the more general case the amplitude f of the outgoing spherical wave could depend on ϕ as well as θ.)

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 11.5)

$$dP = |\psi_{\text{incident}}|^2\, dV = |A|^2 (v\, dt)\, d\sigma.$$

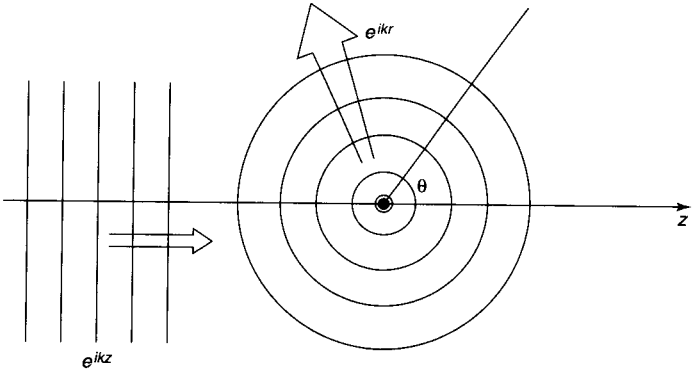

Figure 11.4: Scattering of waves; incoming plane wave generates outgoing spherical wave.

[2]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 classical *particles*, and you could even think of Figure 11.4 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. In that case we'd write the wave function in the *real* form

$$A\{\cos(kz) + f(\theta)\cos(kr + \delta)/r\},$$

and $f(\theta)$ would represent the amplitude of the scattered sound wave in the direction θ.

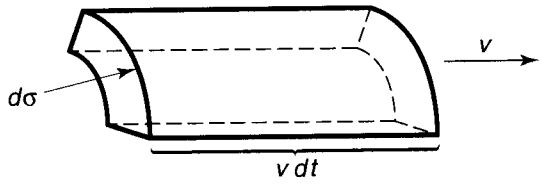

Figure 11.5: The volume dV of incident beam that passes through area $d\sigma$ in time dt.

But this is equal to the probability that the particle later emerges 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 $d\sigma = |f|^2\,d\Omega$, so

$$\boxed{D(\theta) = \frac{d\sigma}{d\Omega} = |f(\theta)|^2.} \qquad [11.14]$$

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 next sections we will study two techniques for calculating the scattering amplitude: **partial wave analysis** and the **Born approximation**.

Problem 11.2 Construct the analogs to Equation 11.12 for one-dimensional and two-dimensional scattering.

11.2 PARTIAL WAVE ANALYSIS

11.2.1 Formalism

As we found in Chapter 4, the Schrödinger equation for a spherically symmetrical potential $V(r)$ admits the separable solutions

$$\psi(r, \theta, \phi) = R(r)Y_l^m(\theta, \phi), \qquad [11.15]$$

where Y_l^m is a spherical harmonic (Equation 4.32) and $u(r) = rR(r)$ satisfies the "radial equation" (Equation 4.37):

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2} + \left[V(r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right]u = Eu. \qquad [11.16]$$

At *very* large r the potential goes to zero, and the centrifugal term is negligible, so

$$\frac{d^2u}{dr^2} \approx -k^2u.$$

The general solution is

$$u(r) = Ce^{ikr} + De^{-ikr};$$

the first term represents an *outgoing* spherical wave, and the second an *incoming* one—for the scattered wave, we evidently want $D = 0$. At very large r, then,

$$R(r) \approx \frac{e^{ikr}}{r},$$

as we already deduced (on qualitative grounds) in the previous section (Equation 11.12).

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 11.6). In the intermediate region (where V can be ignored but the centrifugal term cannot),[3] the radial equation becomes

$$\frac{d^2u}{dr^2} - \frac{l(l+1)}{r^2}u = -k^2u, \qquad [11.17]$$

and the general solution (as we found in Section 4.1.3) is a linear combination of spherical Bessel functions (Equation 4.45):

$$u(r) = Arj_l(kr) + Brn_l(kr). \qquad [11.18]$$

However, neither j_l (which is something 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**:

$$h_l^{(1)}(x) \equiv j_l(x) + in_l(x); \quad h_l^{(2)}(x) \equiv j_l(x) - in_l(x). \qquad [11.19]$$

The first few spherical Hankel functions are listed in Table 11.1. At large r, $h_l^{(1)}(kr)$ (the "Hankel function of the first kind") goes like e^{ikr}/r, whereas $h_l^{(2)}(kr)$ (the "Hankel function of the second kind") goes like e^{-ikr}/r; for outgoing waves we evidently need spherical Hankel functions of the *first* kind:

$$R(r) = Ch_l^{(1)}(kr). \qquad [11.20]$$

---

[3] What follows does not apply to the Coulomb potential, since $1/r$ goes to zero more slowly than $1/r^2$, as $r \to \infty$, and the centrifugal term does *not* dominate in this region. In this sense the Coulomb potential is not "localized."

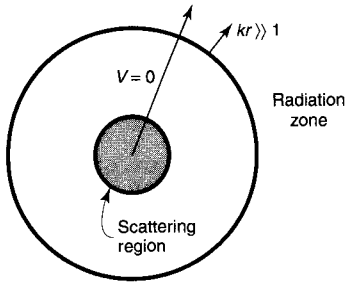

Figure 11.6: Scattering from a localized potential: the scattering region (shaded), the intermediate region (where $V = 0$), and the radiation zone (where $kr \gg 1$).

Thus the exact wave function, in the exterior region [where $V(r) = 0$], is

$$\psi(r, \theta, \phi) = A \left\{ e^{ikz} + \sum_{l,m} C_{l,m} h_l^{(1)}(kr) Y_l^m(\theta, \phi) \right\}. \qquad [11.21]$$

Now, for *very large* r, the Hankel function goes like $(-i)^{l+1} e^{ikr}/kr$ (Table 11.1), so

$$\psi(r, \theta, \phi) \approx A \left\{ e^{ikz} + f(\theta, \phi) \frac{e^{ikr}}{r} \right\}, \qquad [11.22]$$

where

$$f(\theta, \phi) = \frac{1}{k} \sum_{l,m} (-i)^{l+1} C_{l,m} Y_l^m(\theta, \phi). \qquad [11.23]$$

This confirms more rigorously the general structure postulated in Equation 11.12, and tells us how to compute the scattering amplitude, $f(\theta, \phi)$, in terms of the **partial wave amplitudes** $C_{l,m}$. Evidently the differential cross-section is

$$D(\theta, \phi) = |f(\theta, \phi)|^2 = \frac{1}{k^2} \sum_{l,m} \sum_{l',m'} (i)^{l-l'} C_{l,m}^* C_{l',m'} (Y_l^m)^* Y_{l'}^{m'}, \qquad [11.24]$$

Table 11.1: Spherical Hankel functions, $h_l^{(1)}(x)$ and $h_l^{(2)}(x)$.

$h_0^{(1)} = -i \frac{e^{iz}}{z}$	$h_0^{(2)} = i \frac{e^{-iz}}{z}$
$h_1^{(1)} = \left(-\frac{i}{z^2} - \frac{1}{z}\right) e^{iz}$	$h_1^{(2)} = \left(\frac{i}{z^2} - \frac{1}{z}\right) e^{-iz}$
$h_2^{(1)} = \left(-\frac{3i}{z^3} - \frac{3}{z^2} + \frac{i}{z}\right) e^{iz}$	$h_2^{(2)} = \left(\frac{3i}{z^3} - \frac{3}{z^2} - \frac{i}{z}\right) e^{-iz}$
$h_l^{(1)} \to \frac{1}{x} \exp\left\{+i\left[x - \frac{\pi}{2}(l+1)\right]\right\}$	for $x >> 1$
$h_l^{(2)} \to \frac{1}{x} \exp\left\{-i\left[x - \frac{\pi}{2}(l+1)\right]\right\}$	

and the total cross-section is

$$\sigma = \frac{1}{k^2}\sum_{l,m}\sum_{l',m'}(i)^{l-l'}C^*_{l,m}C_{l',m'}\int (Y_l^m)^* Y_{l'}^{m'}\,d\Omega = \frac{1}{k^2}\sum_{l,m}|C_{l,m}|^2. \quad [11.25]$$

(I used the orthonormality of the spherical harmonics, Equation 4.33, in the last step.)

In the previous paragraph I kept the possible ϕ dependence because it cost me nothing. But if (as is ordinarily the case) V is independent of ϕ, then only terms with $m = 0$ survive (remember, $Y_l^m \sim e^{im\phi}$). Now (from Equations 4.27 and 4.32)

$$Y_l^0(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta), \quad [11.26]$$

where P_l is the lth Legendre polynomial. So for the case of azimuthal symmetry, the exact wave function (in the exterior region) is

$$\boxed{\psi(r,\theta) = A\left\{e^{ikz} + \sum_{l=0}^{\infty}\sqrt{\frac{2l+1}{4\pi}}C_l h_l^{(1)}(kr)P_l(\cos\theta)\right\};} \quad [11.27]$$

the scattering amplitude is

$$\boxed{f(\theta) = \frac{1}{k}\sum_{l=0}^{\infty}(-i)^{l+1}\sqrt{\frac{2l+1}{4\pi}}C_l P_l(\cos\theta);} \quad [11.28]$$

and the total cross-section is

$$\sigma = \frac{1}{k^2}\sum_{l=0}^{\infty}|C_l|^2. \quad [11.29]$$

11.2.2 Strategy

All that remains is to determine the partial wave amplitudes C_l for the potential in question. This is accomplished by solving the Schrödinger equation in the *interior* region [where $V(r)$ is distinctly *non*zero] and matching this to the exterior solution (Equation 11.27), using the appropriate boundary conditions. But first I need to do a little cosmetic work, because as it stands my notation is hybrid: I used *spherical* coordinates for the scattered wave, but *Cartesian* coordinates for the incident wave. Before proceeding, it is useful 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

$$\sum_{l,m}[A_{l,m}\, j_l(kr) + B_{l,m} n_l(kr)]Y_l^m(\theta, \phi).$$

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 [$n_l(kr)$ blows up at $r = 0$], and since $z = r\cos\theta$ has no ϕ dependence, only $m = 0$ terms occur. The expansion of a plane wave in terms of spherical waves is sometimes called **Rayleigh's formula**[4]:

$$e^{ikz} = \sum_{l=0}^{\infty} i^l (2l + 1)\, j_l(kr)\, P_l(\cos\theta). \qquad [11.30]$$

Thus the wave function, in the exterior region, can be written in the more consistent form

$$\psi(r, \theta) = A \sum_{l=0}^{\infty} \left[i^l (2l + 1)\, j_l(kr) + \sqrt{\frac{2l+1}{4\pi}} C_l h_l^{(1)}(kr) \right] P_l(\cos\theta). \qquad [11.31]$$

Example: Hard-sphere scattering. Suppose

$$V(r) = \begin{cases} \infty, & \text{for } r \le a, \\ 0, & \text{for } r > a. \end{cases} \qquad [11.32]$$

The boundary condition, then, is

$$\psi(a, \theta) = 0, \qquad [11.33]$$

so

$$\sum_{l=0}^{\infty} \left[i^l (2l + 1)\, j_l(ka) + \sqrt{\frac{2l+1}{4\pi}} C_l h_l^{(1)}(ka) \right] P_l(\cos\theta) = 0 \qquad [11.34]$$

for all θ, from which it follows (Problem 11.3) that

$$C_l = -i^l \sqrt{4\pi(2l + 1)}\, \frac{j_l(ka)}{h_l^{(1)}(ka)}. \qquad [11.35]$$

In particular, the total cross-section is

$$\sigma = \frac{4\pi}{k^2} \sum_{l=0}^{\infty} (2l + 1) \left| \frac{j_l(ka)}{h_l^{(1)}(ka)} \right|^2. \qquad [11.36]$$

---

[4]For a guide to the proof, see George Arfken, *Mathematical Methods for Physicists*, 3rd ed. (Orlando, FL: Academic Press, 1985), Exercise 12.4.7, page 665.

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 4.3, we note that $n_l(z)$ is much larger than $j_l(z)$, for small z, so

$$\frac{j_l(z)}{h_l^{(1)}(z)} = \frac{j_l(z)}{j_l(z) + in_l(z)} \approx -i\frac{j_l(z)}{n_l(z)}$$

$$\approx -i\frac{2^l l! z^l/(2l+1)!}{-(2l)!z^{-l-1}/2^l l!} = \frac{i}{2l+1}\left[\frac{2^l l!}{(2l)!}\right]^2 z^{2l+1}, \qquad [11.37]$$

and hence

$$\sigma \approx \frac{4\pi}{k^2}\sum_{l=0}^{\infty}\frac{1}{2l+1}\left[\frac{2^l l!}{(2l)!}\right]^4 (ka)^{4l+2}.$$

But we're assuming $ka \ll 1$, so the higher powers are negligible—in the low-energy approximation the scattering is dominated by the $l = 0$ term. (This means that the differential cross-section is independent of θ, just as it was in the classical case.) Evidently

$$\sigma \approx 4\pi a^2, \qquad [11.38]$$

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.

Problem 11.3 Derive Equation 11.35, starting with Equation 11.34.

****Problem 11.4** 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 $l = 0$ term contributes significantly. (To simplify matters, throw out all $l \neq 0$ terms right from the start. The main problem, of course, is to determine C_0.) Express your answer in terms of the dimensionless quantity $\phi \equiv 2ma\alpha/\hbar^2$. *Answer*: $\sigma = 4\pi a^2\phi^2/(1+\phi)^2$.

11.3 THE BORN APPROXIMATION

11.3.1 Integral Form of the Schrödinger Equation

The time-independent Schrödinger equation,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi, \qquad [11.39]$$

can be written more succinctly as

$$(\nabla^2 + k^2)\psi = Q, \qquad [11.40]$$

where

$$k \equiv \frac{\sqrt{2mE}}{\hbar} \quad \text{and} \quad Q \equiv \frac{2m}{\hbar^2}V\psi. \qquad [11.41]$$

This has the superficial form 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":

$$(\nabla^2 + k^2)G(\mathbf{r}) = \delta^3(\mathbf{r}). \qquad [11.42]$$

Then we could express ψ as an integral:

$$\psi(\mathbf{r}) = \int G(\mathbf{r} - \mathbf{r}_0)Q(\mathbf{r}_0)\,d^3\mathbf{r}_0. \qquad [11.43]$$

For it is easy to show that this satisfies Schrödinger's equation, in the form of Equation 11.40:

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = \int \left[(\nabla^2 + k^2)G(\mathbf{r} - \mathbf{r}_0)\right] Q(\mathbf{r}_0)\,d^3\mathbf{r}_0$$

$$= \int \delta^2(\mathbf{r} - \mathbf{r}_0)Q(\mathbf{r}_0)\,d^3\mathbf{r}_0 = Q(\mathbf{r}).$$

$G(\mathbf{r})$ is called the **Green's function** for the Helmholtz equation. (In general, the Green's function for a given differential equation represents the "response" to a delta-function source.)

Our first task[5] is to solve Equation 11.42 for $G(\mathbf{r})$. This is most easily accomplished by taking the Fourier transform, which turns the *differential* equation into an *algebraic* equation. Let

$$G(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int e^{i\mathbf{s}\cdot\mathbf{r}} g(\mathbf{s})\,d^3\mathbf{s}. \qquad [11.44]$$

[5] *Warning:* You are approaching two pages of heavy analysis, including contour integration; if you wish, skip straight to the answer, Equation 11.55.

Then

$$(\nabla^2 + k^2)G(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int \left[(\nabla^2 + k^2)e^{i\mathbf{s}\cdot\mathbf{r}}\right] g(\mathbf{s})\, d^3\mathbf{s}.$$

But

$$\nabla^2 e^{i\mathbf{s}\cdot\mathbf{r}} = -s^2 e^{i\mathbf{s}\cdot\mathbf{r}}, \qquad [11.45]$$

and (see Equation 2.126)

$$\delta^3(\mathbf{r}) = \frac{1}{(2\pi)^3} \int e^{i\mathbf{s}\cdot\mathbf{r}}\, d^3\mathbf{s}, \qquad [11.46]$$

so Equation 11.42 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 follows[6] that

$$g(\mathbf{s}) = \frac{1}{(2\pi)^{3/2}(k^2 - s^2)}. \qquad [11.47]$$

Putting this back into Equation 11.44, we find

$$G(\mathbf{r}) = \frac{1}{(2\pi)^3} \int e^{i\mathbf{s}\cdot\mathbf{r}} \frac{1}{(k^2 - s^2)}\, d^3\mathbf{s}. \qquad [11.48]$$

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 11.7). Then $\mathbf{s}\cdot\mathbf{r} = sr\cos\theta$, the ϕ integral is trivial (2π), and the θ integral is

$$\int_0^\pi e^{isr\cos\theta} \sin\theta\, d\theta = -\frac{e^{isr\cos\theta}}{isr}\Big|_0^\pi = \frac{2\sin(sr)}{sr}. \qquad [11.49]$$

Thus

$$G(\mathbf{r}) = \frac{1}{(2\pi)^2}\frac{2}{r}\int_0^\infty \frac{s\sin(sr)}{k^2 - s^2} ds = \frac{1}{4\pi^2 r}\int_{-\infty}^\infty \frac{s\sin(sr)}{k^2 - s^2} ds. \qquad [11.50]$$

The remaining integral is not so simple. It pays to revert to exponential notation and factor the denominator:

$$G(\mathbf{r}) = \frac{i}{8\pi^2 r}\left\{\int_{-\infty}^\infty \frac{se^{isr}}{(s-k)(s+k)} ds - \int_{-\infty}^\infty \frac{se^{-isr}}{(s-k)(s+k)} ds\right\}$$

$$= \frac{i}{8\pi^2 r}(I_1 - I_2). \qquad [11.51]$$

These two integrals can be evaluated using **Cauchy's integral formula**:

$$\oint \frac{f(z)}{(z - z_0)} dz = 2\pi i f(z_0), \qquad [11.52]$$

---

[6]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 2.85.

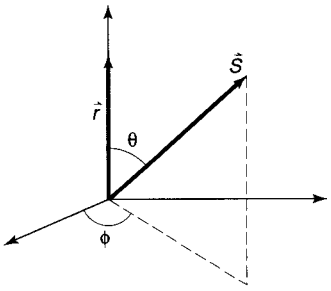

Figure 11.7: Convenient coordinates for the integral in equation [11.48].

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 11.8). (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.)

For each integral in Equation 11.51 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 11.9a). The contour encloses only the singularity at $s = +k$, so

$$I_1 = \oint \left[\frac{se^{isr}}{s+k}\right] \frac{1}{s-k}\, ds = 2\pi i \left[\frac{se^{isr}}{s+k}\right]\Bigg|_{s=k} = i\pi e^{ikr}. \qquad [11.53]$$

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 11.9b); 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):

$$I_2 = -\oint \left[\frac{se^{-isr}}{s-k}\right] \frac{1}{s+k}\, ds = -2\pi i \left[\frac{se^{-isr}}{s-k}\right]\Bigg|_{s=-k} = -i\pi e^{ikr}. \qquad [11.54]$$

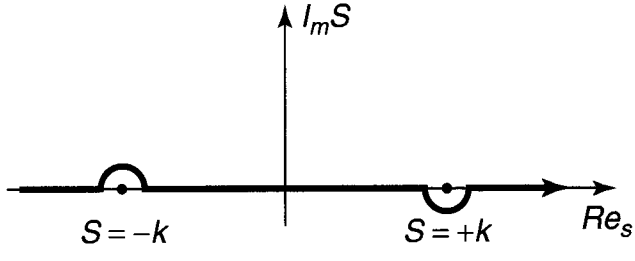

Figure 11.8: Skirting the poles in the contour integral (Equation 11.51).

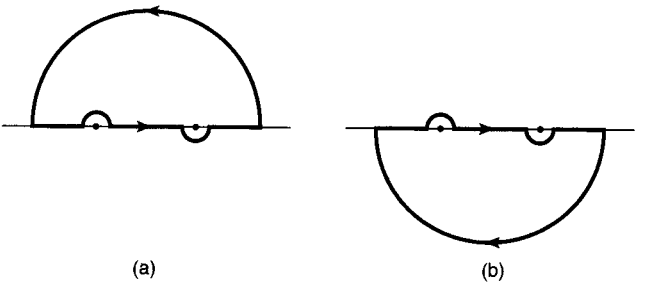

Figure 11.9: Closing the contour in equations [11.53] and [11.54].

Conclusion:

$$G(\mathbf{r}) = \frac{i}{8\pi^2 r}\left[\left(i\pi e^{ikr}\right) - \left(-i\pi e^{ikr}\right)\right] = -\frac{e^{ikr}}{4\pi r}. \qquad [11.55]$$

This, finally, is the Green's function for the Helmholtz equation—the solution to Equation 11.42. (If you got lost in all that analysis, you might want to *check* the result by direct differentiation—see Problem 11.5.) 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:

$$(\nabla^2 + k^2)G_0(\mathbf{r}) = 0; \qquad [11.56]$$

clearly, the result $(G + G_0)$ still satisfies Equation 11.42. 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 11.43, the general solution to the Schrödinger equation takes the form

$$\boxed{\psi(\mathbf{r}) = \psi_0(\mathbf{r}) - \frac{m}{2\pi\hbar^2}\int \frac{e^{ik|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|} V(\mathbf{r}_0)\psi(\mathbf{r}_0)\, d^3\mathbf{r}_0,} \qquad [11.57]$$

where ψ_0 satisfies the *free* particle Schrödinger equation,

$$(\nabla^2 + k^2)\psi_0 = 0. \qquad [11.58]$$

Equation 11.57 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 we can't do the integral unless we 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 11.5 Check that Equation 11.55 satisfies Equation 11.42, by direct substitution. *Hint:* $\nabla^2(1/r) = -4\pi\delta^3(\mathbf{r})$.[7]

Problem 11.6 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 $\kappa \equiv \sqrt{-2mE}/\hbar$).

11.3.2 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 11.57, so

$$|\mathbf{r} - \mathbf{r}_0|^2 = r^2 + r_0^2 - 2\mathbf{r}\cdot\mathbf{r}_0 \cong r^2\left(1 - 2\frac{\mathbf{r}\cdot\mathbf{r}_0}{r^2}\right), \qquad [11.59]$$

and hence

$$|\mathbf{r} - \mathbf{r}_0| \cong r - \hat{r}\cdot\mathbf{r}_0. \qquad [11.60]$$

Let

$$\mathbf{k} \equiv k\hat{r}; \qquad [11.61]$$

then

$$e^{ik|\mathbf{r}-\mathbf{r}_0|} \cong e^{ikr}e^{-i\mathbf{k}\cdot\mathbf{r}_0}, \qquad [11.62]$$

and therefore

$$\frac{e^{ik|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|} \cong \frac{e^{ikr}}{r}e^{-i\mathbf{k}\cdot\mathbf{r}_0}. \qquad [11.63]$$

[In the *denominator* we can afford to make the more radical approximation $|\mathbf{r} - \mathbf{r}_0| \cong r$; in the *exponent* we need to keep the next term. If this puzzles you, try writing out 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

$$\psi_0(\mathbf{r}) = Ae^{ikz}, \qquad [11.64]$$

representing an incident plane wave. For large r, then,

$$\psi(\mathbf{r}) \cong Ae^{ikz} - \frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}\int e^{-i\mathbf{k}\cdot\mathbf{r}_0}V(\mathbf{r}_0)\psi(\mathbf{r}_0)\,d^3\mathbf{r}_0. \qquad [11.65]$$

---

[7] See, for example, D. J. Griffiths, *Introduction to Electrodynamics*, 2nd ed. (Englewood Cliffs, NJ: Prentice Hall, 1989), p. 52.

This is in the standard form (Equation 11.12), and we can read off the scattering amplitude:

$$f(\theta, \phi) = -\frac{m}{2\pi\hbar^2 A} \int e^{-i\mathbf{k}\cdot\mathbf{r}_0} V(\mathbf{r}_0)\psi(\mathbf{r}_0)\, d^3\mathbf{r}_0. \qquad [11.66]$$

So far, this is *exact*. Now we invoke the **Born approximation**: Suppose the incoming plane wave is *not substantially altered by the potential*; then it makes sense to use

$$\psi(\mathbf{r}_0) \approx \psi_0(\mathbf{r}_0) = Ae^{ikz_0} = Ae^{i\mathbf{k}'\cdot\mathbf{r}_0}, \qquad [11.67]$$

where

$$\mathbf{k}' \equiv k\hat{z}, \qquad [11.68]$$

inside the integral. (This would be the *exact* wave function, if V were zero; it is essentially a *weak potential* approximation.) In the Born approximation, then,

$$\boxed{f(\theta, \phi) \cong -\frac{m}{2\pi\hbar^2} \int e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}_0} V(\mathbf{r}_0)\, d^3\mathbf{r}_0.} \qquad [11.69]$$

(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 11.10.)

In particular, for **low-energy** (long-wavelength) **scattering**, the exponential factor is essentially constant over the scattering region, and the Born approximation simplifies to

$$\boxed{f(\theta, \phi) \cong -\frac{m}{2\pi\hbar^2} \int V(\mathbf{r})\, d^3\mathbf{r}, \quad \text{(low energy)}.} \qquad [11.70]$$

(I dropped the subscript on $\mathbf{r}$, since there is no occasion for confusion at this point.)

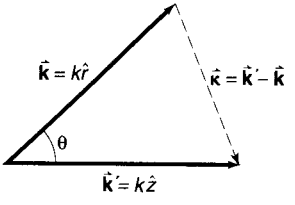

Figure 11.10: Two wave vectors in the Born approximation: $\mathbf{k}$ points in the *incident* direction, $\mathbf{k}'$ in the *scattered* direction.

Example: Low-energy soft-sphere[8] scattering. Suppose

$$V(\mathbf{r}) = \begin{cases} V_0, & \text{if } r \leq a, \\ 0, & \text{if } r > a. \end{cases} \quad [11.71]$$

In this case the low-energy scattering amplitude is

$$f(\theta, \phi) \cong -\frac{m}{2\pi\hbar^2} V_0 \left(\frac{4}{3}\pi a^3\right) \quad [11.72]$$

(independent of θ and ϕ), the differential cross-section is

$$\frac{d\sigma}{d\Omega} = |f|^2 \cong \left(\frac{2mV_0a^3}{3\hbar^2}\right)^2, \quad [11.73]$$

and the total cross-section is

$$\sigma \cong 4\pi \left(\frac{2mV_0a^3}{3\hbar^2}\right)^2. \quad [11.74]$$

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

$$\boldsymbol{\kappa} \equiv \mathbf{k}' - \mathbf{k}, \quad [11.75]$$

and let the polar axis for the $\mathbf{r}_0$ integral lie along $\boldsymbol{\kappa}$, so that

$$(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}_0 = \kappa r_0 \cos\theta_0. \quad [11.76]$$

Then

$$f(\theta) \cong -\frac{m}{2\pi\hbar^2} \int e^{i\kappa r_0 \cos\theta_0} V(r_0) r_0^2 \sin\theta_0 \, dr_0 \, d\theta_0 \, d\phi_0. \quad [11.77]$$

The ϕ_0 integral is trivial (2π), and the θ_0 integral is one we have encountered before (see Equation 11.49). Dropping the subscript on r, we are left with

$$f(\theta) \cong -\frac{2m}{\hbar^2\kappa} \int_0^\infty rV(r) \sin(\kappa r) \, dr, \text{ (spherical symmetry).} \quad [11.78]$$

[8] 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*.

The angular dependence of f is carried by κ; from Figure 11.10 we see that

$$\kappa = 2k \sin(\theta/2). \qquad [11.79]$$

Example: Yukawa scattering. The **Yukawa potential** (which is a crude model for the binding force in an atomic nucleus) has the form

$$V(r) = \beta \frac{e^{-\mu r}}{r}, \qquad [11.80]$$

where β and μ are constants. The Born approximation gives

$$f(\theta) \cong -\frac{2m\beta}{\hbar^2 \kappa} \int_0^\infty e^{-\mu r} \sin(\kappa r)\, dr = -\frac{2m\beta}{\hbar^2(\mu^2 + \kappa^2)}. \qquad [11.81]$$

(You get to work out the integral for yourself in Problem 11.8.)

Example: 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

$$f(\theta) \cong -\frac{2mq_1q_2}{4\pi\epsilon_0 \hbar^2 \kappa^2}, \qquad [11.82]$$

or (using Equations 11.79 and 11.41),

$$f(\theta) \cong -\frac{q_1 q_2}{16\pi\epsilon_0 E \sin^2(\theta/2)}. \qquad [11.83]$$

The differential cross-section is the square of this:

$$\frac{d\sigma}{d\Omega} = \left[\frac{q_1 q_2}{16\pi\epsilon_0 E \sin^2(\theta/2)}\right]^2, \qquad [11.84]$$

which is precisely the Rutherford formula (Equation 11.11). It happens that for the Coulomb potential, classical mechanics, the Born approximation, and quantum field theory all yield the same result. In computer parlance, the Rutherford formula is amazingly "robust."

***Problem 11.7** Find the scattering amplitude, in the Born approximation, for soft-sphere scattering at arbitrary energy. Show that your formula reduces to Equation 11.72 in the low-energy limit.

Problem 11.8 Evaluate the integral in Equation 11.81 to confirm the expression on the right.

****Problem 11.9** Calculate the total cross-section for scattering from a Yukawa potential in the Born approximation. Express your answer as a function of E.

***Problem 11.10** For the potential in Problem 11.4,

(a) calculate $f(\theta)$, $D(\theta)$, and σ, in the low-energy Born approximation;

(b) calculate $f(\theta)$ for arbitrary energies, in the Born approximation;

(c) show that your results are consistent with the answer to Problem 11.4, in the appropriate regime.

11.3.3 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 11.11), and compute the transverse impulse that would be delivered to it in that case:

$$I = \int F_\perp \, dt. \qquad [11.85]$$

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

$$\theta \cong \tan^{-1}(I/p), \qquad [11.86]$$

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.

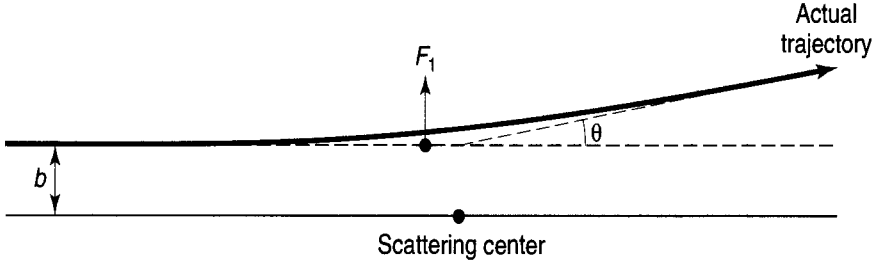

Figure 11.11: The impulse approximation assumes that the particle continues undeflected, and calculates the transverse momentum delivered.

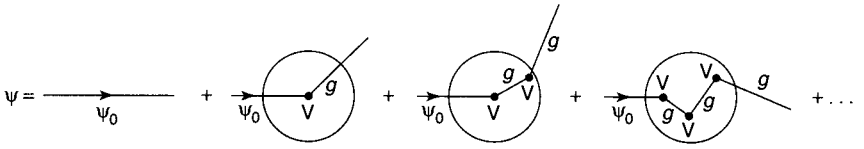

Figure 11.12: Diagrammatic interpretation of the Born series, Equation 11.91.

The integral form of the Schrödinger equation reads

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) + \int g(\mathbf{r} - \mathbf{r}_0) V(\mathbf{r}_0) \psi(\mathbf{r}_0)\, d^3\mathbf{r}_0, \qquad [11.87]$$

where ψ_0 is the incident wave,

$$g(\mathbf{r}) \equiv -\frac{m}{2\pi\hbar^2} \frac{e^{ikr}}{r} \qquad [11.88]$$

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,

$$\psi = \psi_0 + \int gV\psi. \qquad [11.89]$$

Suppose we take this expression for ψ, and plug it in under the integral sign:

$$\psi = \psi_0 + \int gV\psi_0 + \int gVgV\psi. \qquad [11.90]$$

Iterating this procedure, we obtain a formal series for ψ:

$$\psi = \psi_0 + \int gV\psi_0 + \int gVgV\psi_0 + \int gVgVgV\psi_0 + \cdots + \int (gV)^n \psi_0 + \cdots. \qquad [11.91]$$

In each term 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 clear now how one generates the higher-order corrections.

The Born series can be represented diagrammatically as shown in Figure 11.12. 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**.

Problem 11.11 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 11.1a) in the appropriate limit.

***Problem 11.12** Find the scattering amplitude for low-energy soft-sphere scattering in the *second* Born approximation. *Answer*: $-(2mV_0a^3/3\hbar^2)[1 - (4mV_0a^2/5\hbar^2)]$.

FURTHER PROBLEMS FOR CHAPTER 11

***Problem 11.13** Find the Green's function for the *one*-dimensional Schrödinger equation, and use it to construct the integral form (analogous to Equation 11.57). *Answer*:

$$\psi(x) = \psi_0(x) - \frac{im}{\hbar^2 k}\int_{-\infty}^{\infty} e^{ik|x-x_0|}V(x_0)\psi(x_0)\,dx_0. \qquad [11.92]$$

Problem 11.14 Use your result in Problem 11.13 to develop the Born approximation for one-dimensional scattering. That is, choose $\psi_0(x) = Ae^{ikx}$, and assume $\psi(x_0) \cong \psi_0(x_0)$ to evaluate the integral. Show that the reflection coefficient takes the form

$$R \cong \left(\frac{m}{\hbar^2 k}\right)^2 \left|\int_{-\infty}^{\infty} e^{2ikx}V(x)\,dx\right|^2. \qquad [11.93]$$

Problem 11.15 Use the one-dimensional Born approximation (Problem 11.14) to compute the transmission coefficient ($T = 1 - R$) for scattering from a delta function (Equation 2.96) and from a finite square well (Equation 2.127). Compare your results with the exact answers (Equations 2.123 and 2.151).

AFTERWORD

Now that you have (I hope) a sound understanding of what quantum mechanics *says*, I should 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 provides is the statistical distribution of all 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 question 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), you still cannot determine all of its features. Evidently there is some *other* information, external 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, 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**

[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 into 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 until a measurement intervenes, 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.

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.

A.1 The EPR Paradox

In 1935, Einstein, Podolsky, and Rosen[2] published the famous **EPR paradox**, which was designed to prove (on purely theoretical grounds) that the realist position is the only sustainable one. I'll describe a simplified version of the EPR paradox, due to David Bohm. 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 A.1). Now, the pion has spin zero, so conservation of angular momentum requires that the electron and positron are in the singlet configuration:

$$\frac{1}{\sqrt{2}}(\uparrow_-\downarrow_+ - \downarrow_-\uparrow_+). \qquad [A.1]$$

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 *way* 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.

To the realist, there's nothing surprising in 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 any such "spooky action-at-a-distance" (Einstein's words) preposterous. They concluded that the orthodox position is untenable; the

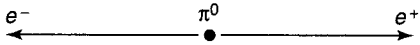

Figure A.1: Bohm's version of the EPR experiment: π^0 at rest decays into electron-positron pair.

[2] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).

electron and positron must have had well-defined spins all along, whether quantum mechanics can calculate them 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 somehow "travels" out 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 50-50 probability of finding *both* particles with spin up. Whatever one might think of such a theory in the abstract, the experiments are unambiguous: No such violation occurs—the correlation of the spins is perfect.

A.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* discription 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.[3] Over the years, a number of hidden variable theories have been proposed, to supplement quantum mechanics; 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.[4]

Bell suggested a generalization of the EPR/Bohm 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 **a**, and the second measures the spin of the positron along the direction **b** (Figure A.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	$+1$
$\vdots$	$\vdots$	$\vdots$

[3]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.

[4]Bell's original paper [*Physics* **1**, 195 (1964)] is a gem: brief, accessible, and beautifully written.

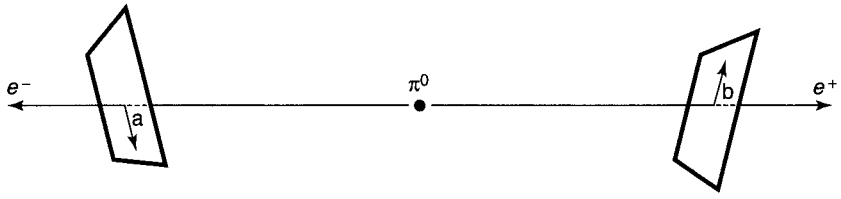

Figure A.2: Bell's version of the EPR-Bohm experiment: detectors independently oriented in directions **a** and **b**.

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:

$$P(\mathbf{a}, \mathbf{a}) = -1. \qquad [A.2]$$

By the same token, if they are *anti*-parallel ($\mathbf{b} = -\mathbf{a}$), then every product is $+1$, so

$$P(\mathbf{a}, -\mathbf{a}) = +1. \qquad [A.3]$$

For arbitrary orientations, quantum mechanics predicts

$$\boxed{P(\mathbf{a}, \mathbf{b}) = -\mathbf{a} \cdot \mathbf{b}} \qquad [A.4]$$

(see Problem 4.44). What Bell discovered is that this result is impossible in 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 gives the result of an electron measurement, and some other function $B(\mathbf{b}, \lambda)$ for the positron measurement. These functions can only[5] take on the values ± 1:

$$A(\mathbf{a}, \lambda) = \pm 1; \quad B(\mathbf{b}, \lambda) = \pm 1. \qquad [A.5]$$

[5]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. But never mind—the point of Bell's argument is to demonstrate that quantum mechanics is incompatible with *any* local deterministic theory—even one that bends over backward to be accommodating.

When the detectors are aligned, the results are perfectly (anti)correlated:

$$A(\mathbf{a}, \lambda) = -B(\mathbf{a}, \lambda), \qquad [A.6]$$

for all λ.

Now, the average of the product of the measurements is

$$P(\mathbf{a}, \mathbf{b}) = \int \rho(\lambda) A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda)\, d\lambda, \qquad [A.7]$$

where $\rho(\lambda)$ is the probability density for the hidden variable. [Like any probability density, it is 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 A.6, we can eliminate B:

$$P(\mathbf{a}, \mathbf{b}) = -\int \rho(\lambda) A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda)\, d\lambda. \qquad [A.8]$$

If $\mathbf{c}$ is any *other* unit vector,

$$P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c}) = -\int \rho(\lambda)\big[A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda) - A(\mathbf{a}, \lambda) A(\mathbf{c}, \lambda)\big]\, d\lambda. \qquad [A.9]$$

Or, since $[A(\mathbf{b}, \lambda)]^2 = 1$:

$$P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c}) = -\int \rho(\lambda)\big[1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)\big] A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda)\, d\lambda. \qquad [A.10]$$

But it follows from Equation A.5 that $-1 \leq [A(\mathbf{a}, \lambda) A(\mathbf{b}, \lambda)] \leq +1$, and $\rho(\lambda)[1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)] \geq 0$, so

$$|P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})| \leq \int \rho(\lambda)\big[1 - A(\mathbf{b}, \lambda) A(\mathbf{c}, \lambda)\big]\, d\lambda, \qquad [A.11]$$

or, more simply,

$$\boxed{\big|P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})\big| \leq 1 + P(\mathbf{b}, \mathbf{c}).} \qquad [A.12]$$

This is the famous **Bell inequality**. It holds for *any* local hidden variable theory (subject only to the minimal requirements of Equations A.5 and A.6), for we have made no assumptions whatever as to the nature or number of the hidden variables or their distribution (ρ).

But it is easy to show that the quantum mechanical prediction (Equation A.4) is incompatible with Bell's inequality. For example, suppose all three vectors lie in

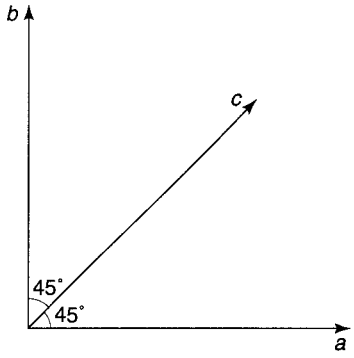

Figure A.3: An orientation of the detectors that demonstrates quantum violations of Bell's inequality.

a plane, and **c** makes a 45° angle with **a** and **b** (Figure A.3); in this case quantum mechanics says

$$P(\mathbf{a}, \mathbf{b}) = 0, \quad P(\mathbf{a}, \mathbf{c}) = P(\mathbf{b}, \mathbf{c}) = -0.707,$$

which is patently inconsistent with Bell's inequality:

$$0.707 \not\leq 1 - 0.707 = 0.293.$$

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.

Many experiments to test Bell's inequality were performed in the 1960s and 1970's, culminating in the work of Aspect, Grangier, and Roger.[6] 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 clearly incompatible with Bell's inequality.[7]

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

---

[6] A. Aspect, P. Grangier, and G. Roger, *Phys. Rev. Lett.* **49**, 91 (1982).

[7] 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. Recently, an improved version of Bell's theorem has been proposed in which a *single measurement* suffices to distinguish between the quantum prediction and that of any local hidden variable theory. See D. Greenberger, M. Horne, A. Shimony, and A. Zeilinger, *Am. J. Phys.* **58**, 1131, (1990) and N. David Mermin, *Am. J. Phys.* **58**, 731, (1990).

Figure A.4: The shadow of the bug moves across the screen at a velocity v' greater than c, provided that the screen is far enough away.

not, there remained the possibility of *nonlocal* hidden variable theories, to which Bell's theorem does not apply[8]). The real shock was the proof 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 alarmed at the idea of 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 A.4). However, the shadow does not carry any *energy*; nor can it transmit any *message* from one point to another on the screen. A person at point X cannot *cause anything to happen* at point Y by manipulating the passing shadow.

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.) The question is, are the superluminal influences predicted by quantum mechanics and detected by Aspect *causal*, in this sense,

[8]It is a curious twist of fate that the EPR paradox, which *assumed* locality to *prove* realism, led finally to the repudiation of locality and left the issue of realism undecided—the outcome (as Mermin 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 place. Still, some authors—notably Bell himself, in *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987)—argue that such theories offer the best hope of bridging the gap between the measured system and the measuring apparatus, and for supplying an intelligible mechanism for the collapse of the wave function.

or are they somehow ethereal enough (like the motion of the 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 correlation of 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 monitoring 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 was measured or not. For the lists of data compiled at the two ends, considered separately, are completely random. It is only when we *compare* the two lists later 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 infiuences *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.[9]

A.3 What is a Measurement?

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

[9] 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).

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?[10] And how can we tell when a measurement has occurred?

Schrödinger posed the essential question most starkly, in his famous **cat paradox**:[11]

> 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, the wave function of the cat has the schematic form

$$\psi = \frac{1}{\sqrt{2}}(\psi_{\text{alive}} + \psi_{\text{dead}}). \qquad \text{[A.13]}$$

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 it to be dead, then it's really *you* who killed it, by looking in the window.

Schrödinger regarded this as patent nonsense, and I think most physicists 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. How are we to reconcile this with the orthodox interpretation of quantum mechanics?

The most widely accepted answer is that the triggering of the Geiger counter constitutes the "measurement," in the sense of the statistical interpretation, not the intervention of a human observer. It is the essence of a measurement that some *macroscopic* system is affected (the Geiger counter, in this instance). The measurement occurs at the moment when the microscopic system (described by the laws of

---

[10]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, more recently, Leslie E. Ballentine, *Quantum Mechanics*, (Prentice Hall, Englewood Cliffs, NJ, 1990).

[11]E. Schrödinger, *Naturwiss.* **48**, 52 (1935); translation by Josef M. Jauch, *Foundations of Quantum Mechanics*, (Reading, MA: Addison-Wesley, 1968), p. 185.

quantum mechanics) interacts with the macroscopic system (described by the laws of classical mechanics) in such a way as to leave a permanent record. The macroscopic system itself is not permitted to occupy a linear combination of distinct states.[12]

I would not pretend that this is an entirely satisfactory resolution, but at least it avoids the stultifying solipsism of Wigner and others, who persuaded themselves that it is the intervention of human consciousness that constitutes a measurement in quantum mechanics. Part of the problem is the word "measurement" itself, which certainly carries an suggestion of human involvement. Heisenberg proposed the word "event", which might be preferable. But I'm afraid "measurement" is so ingrained by now that we're stuck with it. And, in the end, no manipulation of the terminology can completely exorcise this mysterious ghost.

A.4 The Quantum Zeno Paradox

The collapse of the wave function is undoubtedly the *most* peculiar feature of this whole story. It was introduced on purely theoretical grounds, to account for the fact that an immediately repeated measurement reproduces the same value. But surely such a radical postulate must carry directly observable consequences. In 1977 Misra and Sudarshan[13] proposed what they called the **quantum Zeno effect** as a dramatic experimental demonstration of the collapse of the wave function. Their idea was to take an unstable system (an atom in an excited state, say) and subject it to repeated measurements. Each observation collapses the wave function, resetting the clock, and it is possible by this means to delay indefinitely the expected transition to the lower state.[14]

Specifically, suppose a system starts out in the excited state ψ_2, which has a natural lifetime τ for transition to the ground state ψ_1. Ordinarily, for times substantially less than τ, the probability of a transition is proportional to t (see Equation 9.42); in fact, since the transition rate is $1/\tau$,

$$P_{2\to 1} = \frac{t}{\tau}. \qquad [A.14]$$

If we make a measurement after a time t, then, the probability that the system is still in the *upper* state is

$$P_1(t) = 1 - \frac{t}{\tau}. \qquad [A.15]$$

---

[12]Of course, in some ultimate sense the macroscopic system is *itself* described by the laws of quantum mechanics. But wave functions, in the first instance, describe individual elementary particles; the wave function of a macroscopic object would be a monstrously complicated composite, built out of all the wave functions of its 10^{23} constituent particles. Presumably somewhere in the statistics of large numbers macroscopic linear combinations become extremely improbable.

[13]B. Misra and E. C. G. Sudarshan, *J. Math. Phys.* **18**, 756 (1977).

[14]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**.

Suppose we *do* find it to be in the upper state. In that case the wave function collapses back to ψ_2, 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 evidently

$$\left(1 - \frac{t}{\tau}\right)^2 \approx 1 - \frac{2t}{\tau}, \qquad [A.16]$$

which is the same as it would have been had we never made the measurement at t. This is certainly what one would naively expect; if it were the whole story there would be nothing gained by observing the system, and there would be no quantum Zeno effect.

However, for *extremely* short times, the probability of a transition is *not* proportional to t, but rather to t^2 (see Equation 9.39)[15]:

$$P_{2\to 1} = \alpha t^2. \qquad [A.17]$$

In this case the probability that the system is still in the upper state after the two measurements is

$$\left(1 - \alpha t^2\right)^2 \approx 1 - 2\alpha t^2, \qquad [A.18]$$

whereas if we had never made the first measurement it would have been

$$1 - \alpha(2t)^2 \approx 1 - 4\alpha t^2. \qquad [A.19]$$

Evidently our observation of the system after time t *decreased* the net probability of a transition to the lower state!

Indeed, if we examine the system at n regular intervals, from $t = 0$ out to $t = T$ (that is, we make measurements at T/n, $2T/n$, $3T/n$, ..., T), the probability that the system is still in the upper state at the end is

$$\left(1 - \alpha(T/n)^2\right)^n \approx 1 - \frac{\alpha}{n}T^2, \qquad [A.20]$$

which goes to 1 in the limit $n \to \infty$: A *continuously* observed unstable system never decays at all! Some authors regard this as an absurd conclusion, and a proof that the collapse of the wave function is fallacious. However, their argument hinges on a rather loose interpretation of what constitutes "observation." If the track of a particle in a bubble chamber amounts to "continuous observation," then the case is closed, for such particles certainly do decay (in fact, their lifetime is not measureably extended by the presence of the detector). But such a particle is only intermittently interacting with the atoms in the chamber, and for the quantum Zeno effect to occur the successive measurements must be made *extremely* rapidly to catch the system in the t^2 regime.

---

[15] In the argument leading to linear time dependence, we assumed that the function $\sin^2(\Omega t/2)/\Omega^2$ in Equation 9.39 was a sharp spike. However, the *width* of the "spike" is of order $\Delta\omega = 4\pi/t$, and for *extremely* short t this approximation fails, and the integral becomes $(t^2/4)\int \rho(\omega)d\omega$.

As it turns out, the experiment is impractical for spontaneous transitions, but it can be done using *induced* transitions, and the results are in excellent agreement with the theoretical predictions.[16] Unfortunately, this experiment is not as compelling a confirmation of the collapse of the wave function as its designers hoped; the observed effect can be accounted for in other ways.[17]

****

In this book I have tried to present 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, ensemble models, and others—but I believe this one is conceptually the *simplest*, and certainly it is the one shared by most physicists today. 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.

---

[16] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland, *Phys. Rev. A* **41**, 2295 (1990).

[17] L. E. Ballentine, *Found. Phys.* **20**, 1329 (1990); T. Petrosky, S. Tasaki, and I. Prigogine, *Phys. Lett. A* **151**, 109 (1990).

INDEX

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

Y

Z