

Useful Reference Book of Basic Quantum Mechanics

Richard W. Robinett

ONE

Alternative Methods of Solution and Approximation Methods

It is a common practice to approach quantum mechanics through the study of a few, exactly soluble examples using the Schrödinger equation in position-space. The number of potential energy functions for which such closed-form solutions are available is, however, quite small. Luckily, many of them actually correspond reasonably well to actual physical systems; examples include the infinite well as a model of a free particle in a “box” (Chapters 5 and 7), the harmonic oscillator (Chapter 9), the rigid rotator (Chapter 16), and the Coulomb potential for the hydrogen atom (Chapter 17).

Nonetheless, it is important to recognize that other methods can be used to study the properties of a quantum system. Some of them are quite different from the Schrödinger equation approach, and many are amenable for use as numerical and approximation methods in problems for which analytic solutions are not available.

In this chapter, we focus on several methods which can be used to study the spectrum of energy eigenvalues and wavefunctions for time-independent systems and (more briefly) on the effects of time-dependent perturbations, not only as calculational tools for possible numerical analysis, but also as examples of very different ways of approaching quantum mechanical problems. We can make several general comments:

- Many (but by no means all) of the alternative approaches discussed here are most useful for the study of the ground state of the system. Because the structure of matter is ultimately determined by the lowest energy configuration, the determination of the properties of the ground state is arguably the most important; it is the very clearly “first among equals”.
- Any method which is to be used as a numerical approximation technique should be capable of increased precision (usually at the cost of increased

calculational difficulty) as well as providing an estimate of the errors made in the approximation. We will not focus extensively on these questions, but the reader should always keep in mind how each method can be extended in precision, as well as the possible effort involved in doing so.

- As our ultimate goal is to understand the physics behind the equations, we may well have to rethink what it means to “solve” a problem when we approach it numerically. For example, do we need an analytic functional form for $\psi(x)$ or is an array of numbers or an interpolating function enough? How precisely do we need to know the energy eigenvalues? When are we “done”?
- Finally, the use of numerical methods is often nicely complementary to the study of analytic examples. One often looks at a problem in a much different way when one approaches it expecting to write a computer program to “solve” it, and such new insights can be valuable. For example, the study of chaotic dynamics in classical mechanics owes much of its success to the application of numerical (as opposed to analytic) techniques to otherwise familiar problems.

In each section, we first discuss the formalism of each method and then give an example of its possible use as a computational tool.

10.1 Numerical Integration

Classical and quantum mechanics share the fact that their fundamental mathematical descriptions are given by second-order differential equations, Newton’s law for a point particle

$$m \frac{d^2 x(t)}{dt^2} = F(x) \quad (10.1)$$

and the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \quad (10.2)$$

We are used to thinking of Eqn. (10.1) as being completely deterministic,¹ in that, if we are given the appropriate initial conditions, namely, $x_0 = x(0)$ and $v_0 = \dot{x}(0)$, the future time development of $x(t)$ is then predicted. To see how a particle “uses” Eqn. (10.1) to “know where it should be” at later times, we can

¹ We ignore any complications such as the extreme sensitivity to initial conditions present in chaotic systems.

use a conceptually simple method² to integrate Newton's law directly. We first approximate the acceleration (the second derivative) via

$$\begin{aligned}
 \frac{d^2x(t)}{dt^2} &= \ddot{x}(t) = \lim_{\delta \rightarrow 0} \left(\frac{\dot{x}(t + \delta) - \dot{x}(t)}{\delta} \right) \\
 &= \lim_{\delta \rightarrow 0} \left[\frac{\lim_{\delta \rightarrow 0} ((x(t + 2\delta) - x(t + \delta))/\delta) - \lim_{\delta \rightarrow 0} ((x(t + \delta) - x(t))/\delta)}{\delta} \right] \\
 &= \lim_{\delta \rightarrow 0} \left(\frac{x(t + 2\delta) - 2x(t + \delta) + x(t)}{\delta^2} \right) \\
 &\approx \frac{x(t + 2\delta) - 2x(t + \delta) + x(t)}{\delta^2}.
 \end{aligned} \tag{10.3}$$

With this approximation, Newton's law can be written as

$$x(t + 2\delta) \approx 2x(t + \delta) - x(t) + \delta^2 \frac{F(x(t))}{m} \tag{10.4}$$

which is now a *difference equation* for $x(t)$, evaluated at the discretized times $t = n\delta$.

Since

$$v_0 = \dot{x}(0) \approx \frac{x(\delta) - x(0)}{\delta} \quad \text{we have} \quad x(\delta) \approx x(0) + \delta \dot{x}(0) \tag{10.5}$$

and the values of $x(t)$ at the first two of the discretized times, $n = 0, 1$, are fixed by the initial conditions; for later times, the $x(t = n\delta)$ with $n \geq 2$ are then determined by Eqn. (10.4).

Example 10.1. Numerical integration of the classical harmonic oscillator

The classical equation for a mass and a spring is of the form

$$\ddot{x}(t) = -\omega^2 x(t) \tag{10.6}$$

where $\omega = \sqrt{K/m}$. For any numerical problem, we must specialize to definite values, for both the physical parameters of the problem, and for the initial conditions; as an example, we choose

$$\omega = 2\pi, \quad x(0) = 1, \quad \text{and} \quad v(0) = \dot{x}(0) = 0 \tag{10.7}$$

² Much more powerful techniques, such as the Runge–Kutta method, are discussed in all textbooks dealing with numerical methods.

(Continued)

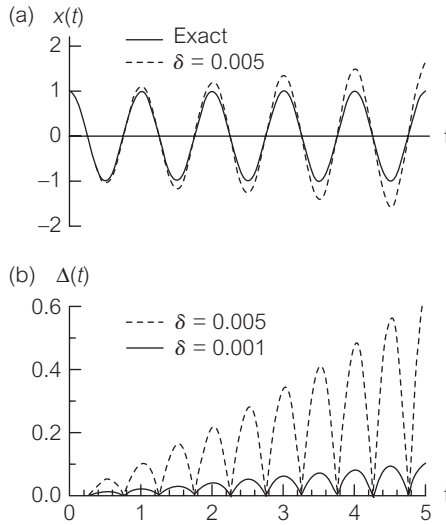


Figure 10.1. The exact (solid) and numerical (dashed) solutions of the harmonic oscillator differential equation are shown in (a). The differences between the numerical and exact solutions versus time for two different step sizes, δ , are plotted in (b).

which has the exact solution $x(t) = \cos(2\pi t)$. In Fig. 10.1(a), we show the result of a numerical solution of Eqn. (10.6) (dashed curve), using Eqn. (10.4), to be compared to the exact solution (solid curve). In Fig. 10.1(b) the difference between the numerical and exact solutions, $\Delta(t)$, is seen to increase with t , but it is also smaller for smaller step sizes, δ , as expected.

The same strategy can be used to solve the Schrödinger by approximating Eqn. (10.2) as

$$\psi(x + 2\delta) \approx 2\psi(x + \delta) - \psi(x) + \delta^2 \left[\frac{2m}{\hbar^2} (V(x) - E) \right] \psi(x) \quad (10.8)$$

and using

$$\psi(0) \quad \text{and} \quad \psi(\delta) \approx \psi(0) + \psi'(0)\delta \quad (10.9)$$

In this sense, Eqn. (10.2) is just as deterministic as Newton's laws; the chief differences are:

- The choice of $x = 0$ as the “initial” value is arbitrary.
- The differential equation can (and should) be integrated “to the left” as well to obtain $\psi(x)$ for $x < 0$.

- Most importantly, the Schrödinger equation can be integrated (solved) for any value of the energy eigenvalue, E ; the solutions so obtained, however, will not necessarily be physically acceptable, that is, square-integrable.

To illustrate the usefulness of this approach to the isolation of energy eigenvalues and their corresponding eigenfunctions, we restrict ourselves to the special case of a symmetric potential for reasons that will become clear. In that case, we know that the solutions will also be eigenfunctions of parity and hence satisfy

$$\text{even solutions: } \psi(0) = \text{arbitrary} \quad \text{and} \quad \psi'(0) = 0 \quad (10.10)$$

and

$$\text{odd solutions: } \psi(0) = 0 \quad \text{and} \quad \psi'(0) = \text{arbitrary} \quad (10.11)$$

The arbitrariness in $\psi(0)$ or $\psi'(0)$ present at this point is eventually removed when the wavefunction is properly normalized, but that is separate from the solution of the Schrödinger equation itself. The overall normalization does not affect the *shape* of the solution.

We now focus on the behavior of the wavefunction at large $|x|$ for various values of E . For the even case, for example, we can start at $x = 0$ with an arbitrary value of $\psi(0)$, use the oddness of $\psi'(x)$ to determine $\psi(\delta) = \psi(0)$, and then use Eqn. (10.8) to numerically integrate to arbitrarily large values of $x = n\delta$; we find the generic behavior shown in Fig. 10.2(a). If we call the lowest even energy eigenvalue $E_1^{(+)}$, then for values of $E < E_1^{(+)}$, the solutions diverge as $\psi(x) \rightarrow +\infty$ as $x \rightarrow +\infty$. When $E \gtrsim E_1^{(+)}$, the solutions are still poorly behaved at infinity, but now diverge with the opposite sign. Clearly, the energy of the physically acceptable square-integrable ground state solution lies between E_a and E_b ; this behavior is familiar from our study of the harmonic oscillator and Fig. 9.3.

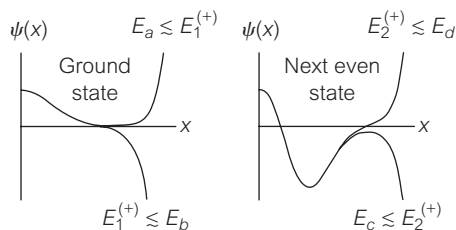


Figure 10.2. Numerical solutions of the Schrödinger equation for a symmetric potential. The energy parameters $E_a < E_1^{(+)} < E_b$ bracket the true ground state energy; $E_c < E_2^{(+)} < E_d$ bracket the first-excited even state.

Once such a pair of energy values which brackets the “acceptable” ground state solution is found, one can determine $E_1^{(+)}$ with increasing *precision* by a systematic exploration in the interval (E_a, E_b) , finding values of E , which bracket the “true value” with decreasing error; the resulting *accuracy* of the estimated value of E_0 , however, will still depend on the integration method used (Q10.2). As the energy parameter E is increased further, additional changes in sign of the wavefunction at infinity are encountered (Fig. 10.1(b)), and the energy spectrum can be systematically mapped out by finding pairs of energy values which bracket a “sign change.”

Example 10.2. Energy eigenvalues for the harmonic oscillator

The numerical solution of the Schrödinger equation for the harmonic oscillator potential is easy to implement using Eqn. (10.8) provided the problem is put into dimensionless form as in Section 9.2.1, namely

$$\frac{d^2\psi(y)}{dy^2} = (y^2 - \epsilon)\psi(y) \quad (10.12)$$

where the dimensionless eigenvalues are $\epsilon_n = 2E_n/\hbar\omega = (2n + 1)$. The even states have $\epsilon_n = 1, 5, 9, \dots$ and so forth. Values of E_a, E_b , which bracket the ground state ($E_1^{(+)}$) and first even excited state ($E_2^{(+)}$) energies for several values of δ are given by

| δ | $E_a < E_1^{(+)} < E_b$ | $E_a < E_2^{(+)} < E_b$ |
|----------|-------------------------|-------------------------|
| 0.1 | (1.191, 1.192) | (5.510, 5.511) |
| 0.01 | (1.0171, 1.0172) | (5.0431, 5.0432) |
| 0.001 | (1.00169, 1.00170) | (5.00423, 5.00424) |

so that the effect of decreasing the step size on the reliability of the results is clear.

It is useful to keep in mind that before applying any numerical technique to a new problem, it is best to “test” it on a well-understood example if at all possible.

Once an approximate energy eigenvalue is found, the wavefunction for each energy eigenvalue is obtained from the numerical integration as the collection of points $\psi(x = n\delta)$, and can be fit to a smooth function using interpolation techniques if desired; in any case, it can be normalized and used to extract further information about the quantum system. The odd states are found in a similar way (Q10.4) by making use of Eqn. (10.11).

10.2 The Variational or Rayleigh–Ritz Method

Many branches of physics can be formulated in terms of a simple minimum principle using the methods of the calculus of variations. Examples include minimum surface problems (bubble problems and the like), Fermat’s formulation of geometrical optics using a principle of least time and, perhaps most importantly, the principle of least action approach to classical mechanics.

In each case, the object of study is a *functional*, so-called because it takes as its argument a function and returns a number as its output. The *classical action* in mechanics, $S[x(t)]$, is just such an example; it takes any possible classical path, $x(t)$, and returns the numerical value

$$S[x(t)] = \int_{t_a}^{t_b} dt \left(\frac{1}{2} m \dot{x}^2(t) - V(x(t)) \right) \quad (10.13)$$

and the trajectory realized in nature is the unique path which minimizes Eqn. (10.13).

It is perhaps then not surprising that quantum mechanics can also be formulated in such a manner. We will first discuss just such an approach and then discuss how it can be applied as a calculational tool to approximate energy eigenvalues and wavefunctions.

Consider a Hamiltonian, \hat{H} , defining the bound state spectra of some system. We assume that it will have a discrete spectrum of bound state energies, E_n , with corresponding, already normalized wavefunctions $\psi_n(x)$. We can define an *energy functional* for *any* trial wavefunction, $\psi(x)$, via

$$\begin{aligned} E[\psi] &\equiv \langle \psi | \hat{H} | \psi \rangle = \langle \psi | \hat{T} | \psi \rangle + \langle \psi | V(x) | \psi \rangle \\ &= \int_{-\infty}^{+\infty} dx \psi^*(x) \hat{H} \psi(x) \\ &= \frac{1}{2m} \int_{-\infty}^{+\infty} dx \psi^*(x) \hat{p}^2 \psi(x) + \int_{-\infty}^{+\infty} dx V(x) |\psi(x)|^2 \end{aligned} \quad (10.14)$$

This is defined whether $\psi(x)$ is an eigenfunction or not. It is often convenient to use the “alternative” form of the average value of kinetic energy (as in

Eqn. (4.65)), that is,

$$\langle \psi | \hat{T} | \psi \rangle = \langle \hat{T} \rangle = \frac{\hbar^2}{2m} \int_{-\infty}^{+\infty} dx \left| \frac{d\psi(x)}{dx} \right|^2 \quad (10.15)$$

If for some reason the trial wavefunction ψ is not already normalized, we can simply write

$$E[\psi] \equiv \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \quad (10.16)$$

It is easy to see that this functional simply returns an energy eigenvalue when its argument is a normalized eigenstate, since

$$E[\psi_n] = \int_{-\infty}^{+\infty} dx \psi_n^*(x) \hat{H} \psi_n(x) = \int_{-\infty}^{+\infty} dx \psi_n^*(x) E_n \psi_n(x) = E_n \quad (10.17)$$

For a general wavefunction, $\psi(x)$, we assume we can use the expansion theorem and write $\psi(x) = \sum_{n=0}^{\infty} a_n \psi_n(x)$ and we find that

$$\begin{aligned} E[\psi] &= \langle \psi | \hat{H} | \psi \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_m^* a_n \langle \psi_m | \hat{H} | \psi_n \rangle \\ &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_m^* a_n E_n \delta_{n,m} \\ &= \sum_{n=0}^{\infty} |a_n|^2 E_n \end{aligned} \quad (10.18)$$

This derivation is similar to that of Section 6.4 for the average value of the energy operator, \hat{E} , in a general state, but the quantity that appears in the energy functional here is the expectation value of the appropriate Hamiltonian for the problem, which in general acts only on spatial degrees of freedom.

We assume that the energy eigenvalues are ordered, that is, $\cdots \geq E_2 \geq E_1 \geq E_0$ so that

$$E[\psi] = \sum_{n=0}^{\infty} |a_n|^2 E_n \geq \sum_{n=0}^{\infty} |a_n|^2 E_0 = E_0 \quad \text{or} \quad E[\psi] \geq E_0 \quad (10.19)$$

because the expansion coefficients, when squared, sum to unity, since $\psi(x)$ is assumed normalized. The lower bound is only “saturated” when $\psi(x) = \psi_0(x)$ in which case $a_n = \delta_{n,0}$ and only the ground state energy term contributes.

This is then the desired minimum principle, namely that:

- The energy functional, $E[\psi]$, defined via Eqn. (10.14), always gives an energy at least as large as the true ground state energy, that is, $E[\psi] \geq E_0$ for all ψ .

To use this property as a calculational tool, we first note that if the wavefunction used in the functional has an arbitrary parameter, for example, $\psi(x) = \psi(x; a)$, then the energy functional yields a function of one variable, namely

$$E[\psi(x; a)] = E(a) \quad (10.20)$$

An example of this would be the family of Gaussian variational wavefunctions, $\psi(x; a) = \exp(-x^2/2a^2)/\sqrt{a\sqrt{\pi}}$, with a variable width.

Because the functional satisfies the minimum principle for each value of the parameter, one can minimize the variational function $E(a)$ and be assured that the resulting minimum is still greater than the true ground state energy. Thus, one can find the trial wavefunction, in the one parameter family considered, which has the lowest energy. The minimizing wavefunction accomplishes this by somehow “adjusting” to as similar as possible to the exact ground state solution. This approach is similar in spirit to the “zero-point energy” argument of P1.14, but is more powerful because:

- The guaranteed lower bound of Eqn. (10.19) provides a method of assessing the reliability of the approximations.

— Of two variational estimates of the ground state energy, the lower one is always closer to the true value.

In this context, “lower is always better” as we know that we can never “undershoot” E_0 on the negative side.

- It also provides an approximation to the wavefunction as well as to the energy; one can then use it to estimate expectation values and to find the approximate momentum-space wavefunction. As an aside, because the argument leading to Eqn. (10.18) is not specific to a position-space representation, one can also use the variational method with momentum-space wavefunctions (P10.5).

For illustrative purposes, we will sometimes calculate $|a_0|^2$ as a measure of the “overlap” of the trial solution with the exact ground state wavefunction (if known); it can be used as a quantitative measure of the similarity of any two functions. We reiterate, however, that the trial wavefunction of a given class which minimizes the energy is not necessarily the one which has the largest overlap with the true ground state wavefunction, that is, it does not necessarily maximize $|a_0|^2$ (see, for example, P10.9).

Example 10.3. Variational estimate for the harmonic oscillator I

As an example of the method, consider approximating the ground state energy and eigenfunction of the simple harmonic oscillator by using the family of trial wavefunctions $\psi(x; a) = \exp(-x^2/2a^2)/\sqrt{a\sqrt{\pi}}$ mentioned above. Because the true ground state solution is also Gaussian, we expect to find the exact answer. We have to evaluate Eqn. (10.14) with $V(x) = m\omega^2 x^2/2$ and we find that

$$E[\psi(x; a)] = E(a) = \langle \hat{T} \rangle + \langle V(x) \rangle = \frac{\hbar^2}{4ma^2} + \frac{1}{4}m\omega^2 a^2 \quad (10.21)$$

Minimizing this expression we find

$$\frac{dE(a)}{da} = -\frac{\hbar^2}{2ma^3} + \frac{1}{2}m\omega^2 a = 0 \quad (10.22)$$

which yields $a_{\min} = \sqrt{\hbar/m\omega}$ and $E(a_{\min}) = \hbar\omega/2$ as expected.

Example 10.4. Variational estimate for the harmonic oscillator II

To illustrate the principle in the case where the form of the ground state wavefunction is not known, consider as a trial wavefunction for the simple harmonic oscillator (SHO) the wavefunction

$$\psi(x; a) = \begin{cases} 0 & \text{for } |x| > a \\ N(a^2 - x^2)^2 & \text{for } |x| < a \end{cases} \quad (10.23)$$

where the variational parameter is again a and the normalization constant is given by $N = \sqrt{315/256a^9}$. A similar calculation to the one above shows (P10.3) that the energy function is

$$E[\psi] = E(a) = \frac{3\hbar^2}{2ma^2} + \frac{m\omega^2 a^2}{22}. \quad (10.24)$$

This has a minimum value at $a_{\min}^2 = \sqrt{33}\hbar/m\omega$ yielding

$$E(a_{\min}) = \frac{\hbar\omega}{2} \sqrt{\frac{12}{11}} = (0.522)\hbar\omega \quad (10.25)$$

which is only 4.4% greater than the exact value.

The trial wavefunctions, along with their corresponding energies for several choices of a are shown in Fig. 10.3 along with the value of $|a_0|^2$. We plot in Fig. 10.4 the fractional difference between the variational energy and the exact ground state value $(E(\text{var}) - E(\text{exact}))/E(\text{exact})$ as well as the probability that the variational wavefunction is *not* in the ground state, that is, $1 - |a_0|^2$, versus the variational parameter a . We note

(Continued)

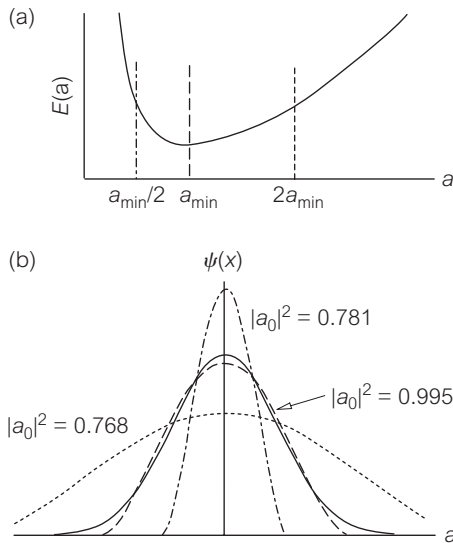
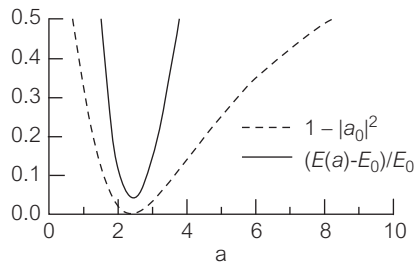


Figure 10.3. (a) The variational energy, $E(a)$ versus a , showing the value of a_{min} which minimizes the energy functional (vertical dashed line) and two other values (dotted and dot-dash). (b) The corresponding variational wavefunctions (same dashing) along with the exact ground state (solid curve). Values of the overlap, given by $|a_0|^2$, for each variational waveform are also shown.

Figure 10.4. The fractional energy error (solid curve) and the degree of “non-overlap” (dashed curve) versus variational parameter a for Example 10.4. This illustrates that first-order changes in the wavefunction give second-order changes in the energy functional.



that variations in the parameter a seem to have a much larger effect on the energy functional than on the wavefunction itself.

To formalize this last observation further, let us imagine making small variations around the exact ground state wavefunction, $\psi_0(x)$, parameterized by $\psi_0(x) \rightarrow \psi_0(x) + \lambda\phi(x)$ so that $\phi(x)$ represents a first-order change in the wavefunction; we use λ to keep track of the expansion.

Consider then the energy functional (where we use Eqn. (10.16) since the new wavefunction is not properly normalized) and we find

$$\begin{aligned}
 E[\psi_0 + \lambda\phi] &= \frac{\langle \psi_0 + \lambda\phi | \hat{H} | \psi_0 + \lambda\phi \rangle}{\langle \psi_0 + \lambda\phi | \psi_0 + \lambda\phi \rangle} \\
 &= \frac{E_0 \langle \psi_0 | \psi_0 \rangle + \lambda(\langle \psi_0 | \hat{H} | \phi \rangle + \langle \phi | \hat{H} | \psi_0 \rangle) + \lambda^2 \langle \phi | \hat{H} | \phi \rangle}{\langle \psi_0 | \psi_0 \rangle + \lambda \langle \psi_0 | \phi \rangle + \lambda \langle \phi | \psi_0 \rangle + \lambda^2 \langle \phi | \phi \rangle} \\
 &= E_0 \left(\frac{1 + \lambda(\langle \psi_0 | \phi \rangle + \langle \phi | \psi_0 \rangle) + \lambda^2 \langle \phi | \hat{H} | \phi \rangle / E_0}{1 + \lambda(\langle \psi_0 | \phi \rangle + \langle \phi | \psi_0 \rangle) + \lambda^2 \langle \phi | \phi \rangle} \right) \\
 &= E_0 (1 + \mathcal{O}(\lambda^2))
 \end{aligned} \tag{10.26}$$

since the zeroth-order and $\mathcal{O}(\lambda)$ terms are identical. This shows that, in general,

- *First-order changes* ($\mathcal{O}(\lambda)$) in the trial wavefunction, away from the true ground state solution, give rise to *second-order changes* ($\mathcal{O}(\lambda^2)$) in the corresponding energy functional.

This fact is reflected in Fig. 10.4 as the fractional change in energy does seem to vary quadratically with deviations away from the minimum value of the variational parameter, while the deviation in the wavefunction itself (as measured by $1 - |a_0|^2$) seems to vary much more weakly on a . This is a typical feature of problems involving the calculus of variations.

If the variational method is to be useful as an approximation method there should be some possibility of further refinement of the estimation of the ground state energy. This can be accomplished by simply taking as a trial wavefunction one with a larger number of variational parameters. For example, one might consider

$$\psi(x; a, b) = e^{-x^2/2a^2} (1 + bx^2) \tag{10.27}$$

which has an additional parameter, b , but which reduces to the original choice in some limit (namely, $b = 0$). In this case, we are guaranteed to have

$$E(a_{\min}) \equiv E[\psi(x; a_{\min})] > E(a_{\min}, b_{\min}) \equiv E[\psi(x; a_{\min}, b_{\min})] \geq E_0 \tag{10.28}$$

because any variational energy must be larger than the true ground state and because the minimum with nonzero values of b will be at least as small as for

$b = 0$. The new minimum value will be determined by

$$\frac{\partial E(a, b)}{\partial a} = \frac{\partial E(a, b)}{\partial b} = 0 \tag{10.29}$$

By adding more and more variational parameters, we can allow the trial wavefunction to conform as closely as possible to the exact ground state.

Example 10.5. Variational estimate for the harmonic oscillator III

We illustrate the improvement possible with multi-parameter trial wavefunctions by using the (unnormalized) function

$$\psi(x; a, b) = \begin{cases} 0 & \text{for } |x| > a \\ (a^2 - x^2)^2(1 + bx^2) & \text{for } |x| < a \end{cases} \tag{10.30}$$

as a trial solution for the ground state of the SHO. We plot in Fig. 10.5 a contour plot of $E(a, b)$ versus a, b ; the small star on the dotted line indicates the minimum for the $b = 0$ case, while the small $+$ indicates the new global minimum which does indeed have somewhat lower energy. The values of the exact, one-parameter, and two-parameter fits for various quantities are shown below:

| quantity | exact | $\psi(x; a)$ | $\psi(x; a, b)$ |
|------------------------------|-------|--------------|-----------------|
| $E_0/(\hbar\omega/2)$ | 1 | 1.0445 | 1.0198 |
| $ a_0 ^2$ | 1 | 0.9951 | 0.9977 |
| $\langle x^2 \rangle/\rho^2$ | 1/2 | 0.5222 | 0.5099 |
| $\langle x^4 \rangle/\rho^4$ | 3/4 | 0.6923 | 0.6884 |

(10.31)

The energy is lower and the overall fit is better ($|a_0|^2$ is closer to 1) than in the one-parameter case; it is clear, however, that various higher moments (i.e. average values of x^{2n}) are never fit very well with this particular form which is not surprising given its lack of a realistic ‘tail’ for large $|x|$.

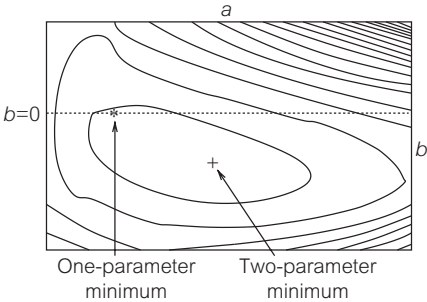


Figure 10.5. Contour plot of two-parameter variational energy $E(a, b)$ versus (a, b) ; the dotted line corresponds to the one-parameter family, $E(a)$ versus a . The minimum energy for the two-parameter family of trial wavefunctions is lower than for the one-parameter set.

The process can be continued with as many variational parameters as one can handle, presumably improving the agreement with experiment, if not providing much more useful insight into the basic physics.³ One of the most famous *tour de force* calculations of this type are variational calculations of the ground state of the helium atom which use trial wavefunctions with hundreds of parameters.⁴

In some cases, it is also possible to extend the variational method to give rigorous lower bounds for excited states as well as the ground state. Suppose, for example, that one could choose a trial wavefunction which was somehow known to be orthogonal to the true ground state, that is, $\langle \psi | \psi_0 \rangle = 0$; this would guarantee that $a_0 = 0$ in the expansion theorem. The standard argument would then give

$$\psi = \sum_{n=1}^{\infty} a_n \psi_n \implies E[\psi] = \sum_{n=1}^{\infty} |a_n|^2 E_n \geq \sum_{n=1}^{\infty} |a_n|^2 E_1 = E_1 \quad (10.32)$$

Thus, all the trial wavefunctions in this restricted class would have energies at least as great as the first excited state.

Various symmetries of the problem can often be used to restrict the form of the trial wavefunction so as to satisfy this constraint. For example, in a one-dimensional problem with a symmetric potential, $V(x) = V(-x)$, we know that the ground state will be an even function; therefore any odd trial wavefunction will have $a_0 = 0$ and hence satisfy Eqn. (10.32) and give a good estimate of the first excited state energy. Less prosaically, in three-dimensional problems with spherical symmetry, the ground state will have no angular momentum (i.e. $l = 0$) and excited states with higher values of l are automatically orthogonal to the ground state.

10.3 The WKB method

The variational method is best suited to evaluating the properties of the ground state solution, that is, for $n = 0$. It is useful to have a complementary approach, which is more appropriate for the quasi-classical regime where $n \gg 1$; we have argued that this limit is also attained, in some sense, when $\hbar \rightarrow 0$. Such an

³ It is said that, when confronted with the result of an impressive numerical calculation, Eugene Wigner said “*It is nice to know that the computer understands the problem. But I would like to understand it too.*”; See Nussenzveig (1992).

⁴ See Bethe and Jackiw (1968) or Park (1992) for discussions.

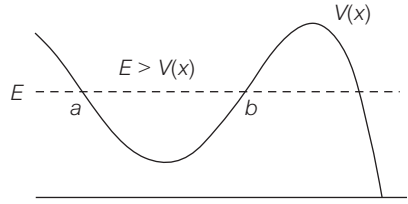


Figure 10.6. Generic potential with classical turning points for the WKB approximation.

approach was first discussed in the context of quantum mechanics by Wentzel, Kramers, and Brillouin⁵ and is therefore often called the WKB method.

10.3.1 WKB Wavefunctions

Motivated by the simple form for a free-particle de Broglie wave, that is,

$$\psi(x) = A e^{i2\pi x/\lambda} = A e^{ikx} = A e^{ipx/\hbar} \quad (10.33)$$

we attempt a solution of the time-independent Schrödinger equation of the form

$$\psi(x) = A(x) e^{iF(x)/\hbar} \quad (10.34)$$

where $A(x)$ and $F(x)$ are an amplitude and phase term, respectively. We retain the explicit factor of \hbar and will use it to parameterize the smallness of various terms. We also assume, for the moment, that we are in the classically allowed region so that $E > V(x)$, as in Fig. 10.6, so that $a < x < b$.

With the *ansatz*⁶ in Eqn. (10.34), the Schrödinger equation becomes

$$\begin{aligned} 0 = A(x) & \left[\frac{1}{2m} \left(\frac{dF(x)}{dx} \right)^2 - (E - V(x)) \right] \\ & - \hbar \left(\frac{i}{2m} \right) \left[2 \frac{dA(x)}{dx} \frac{dF(x)}{dx} + A(x) \frac{d^2 F(x)}{dx^2} \right] \\ & - \hbar^2 \left[\frac{1}{2m} \frac{d^2 A(x)}{dx^2} \right] \end{aligned} \quad (10.35)$$

At this point, we can either consider \hbar as an arbitrary small parameter and set the first two terms (of order $\mathcal{O}(\hbar^0)$ and $\mathcal{O}(\hbar^1)$, respectively) separately to zero or else we can require that both the real and imaginary parts of Eqn. (10.35) are satisfied; in either case, we neglect the last term (being of order $\mathcal{O}(\hbar^2)$) and discuss the validity of this approximation below.

⁵ It was also studied independently by Jeffries; the name WKBJ approximation is therefore sometimes used.

⁶ The German term *ansatz*, often defined as *formulation* or *setup*, is often taken to mean something like “assumed form of the solution” in the context of a physics or math problem.

The $\mathcal{O}(\hbar^0)$ equation (or real part) is easily written as

$$\frac{dF(x)}{dx} = \pm \sqrt{2m(E - V(x))} \equiv \pm p(x) \quad (10.36)$$

or

$$F(x) = \pm \int^x p(x) dx \quad (10.37)$$

where $p(x)$ is simply the classical momentum. The $\mathcal{O}(\hbar^1)$ (or imaginary part) then gives

$$2 \frac{dA(x)}{dx} p(x) + A(x) \frac{dp(x)}{dx} = 0 \quad (10.38)$$

which can be multiplied on both sides by $A(x)$ to obtain

$$\left(2A(x) \frac{dA(x)}{dx} \right) p(x) + [A(x)]^2 \frac{dp(x)}{dx} \equiv \frac{d}{dx} ([A(x)]^2 p(x)) = 0 \quad (10.39)$$

or

$$[A(x)]^2 p(x) = C \quad (10.40)$$

where C is a constant. The two linearly independent solutions (corresponding to right (+) and left-moving (−) waves) are then given by

$$\psi_{\pm}(x) = \frac{C_{\pm}}{\sqrt{p(x)}} e^{\pm i \int^x p(x) dx / \hbar} \propto \frac{1}{\sqrt{v(x)}} e^{\pm i \int^x k(x) dx} \quad (10.41)$$

where $p(x) = \hbar k(x)$ defines the “local wavenumber” $k(x)$, and $v(x)$ is the local speed. This remarkably simple solution has several obvious features:

- The corresponding probability density, $|\psi(x)|^2$, satisfies

$$|\psi(x)|^2 \propto \frac{1}{p(x)} \propto \frac{1}{v(x)} \quad (10.42)$$

which is exactly of the form of the *classical probability distribution* first discussed in Section 5.1. This implies (recall Fig. 9.7) that the wavefunctions for the low-lying energy levels will be *poorly* described by the WKB solutions; the quantum wavefunctions for large quantum numbers will, however, approach these semiclassical solutions when suitably locally averaged.

- The phase of the wavefunction can be written as

$$\int^x k(x) dx = \int^x d\phi(x) \quad (10.43)$$

where

$$d\phi(x) = k(x) dx = \frac{2\pi}{\lambda(x)} dx \quad \text{or} \quad \frac{d\phi}{2\pi} = \frac{dx}{\lambda(x)} \quad (10.44)$$

Thus, as the particle moves a distance dx , or a fraction of a “local” wavelength, $df = dx/\lambda(x)$, through the potential, it acquires a phase $d\phi = 2\pi df$.

- The solutions can be easily extended to the case where $E < V(x)$, that is, in the classically disallowed regions by appropriate changes in sign giving

$$\psi_{\pm} = \frac{\tilde{C}_{\pm}}{\sqrt{p(x)}} \exp\left(\pm \sqrt{2m/\hbar^2} \int^x \sqrt{V(x) - E} dx\right) \quad (10.45)$$

which are the exponentially suppressed solutions discussed in Section 8.2.2; these give rise to quantum tunneling effects. The WKB wavefunction thus has features of both the classical probability distribution, arising from averaging over the trajectory, and the quantum wavefunction.

With this form of the solution, we can examine the effect of neglecting the $\mathcal{O}(\hbar^2)$ term in Eqn. (10.35). Taking the ratio of the last term to the first we find something of the order

$$\frac{\hbar^2}{F'(x)^2} \left(\frac{A''(x)}{A(x)} \right) \propto \frac{\hbar^2}{p(x)^2} \frac{1}{l^2} \propto \frac{1}{[lk(x)]^2} \propto \left(\frac{\lambda(x)}{2\pi l} \right)^2 \quad (10.46)$$

where l is a typical distance scale over which $E - V(x)$ changes. Thus, if the “local de Broglie wavelength,” $\lambda(x)$, is much shorter than the distance scale over which the potential changes, the semiclassical approximation is a good one. This is obviously not the case near the classical turning points where the explicit $1/\sqrt{p(x)}$ factors in Eqn. (10.41) actually diverge, indicating that the solution is poorly behaved there.

To obtain a complete description of the wavefunction, the solutions inside and outside the well must be smoothly matched onto each other. The formalism for doing this is not beyond the level of this text, but we choose to only quote the results.⁷ For example, we can take linear combinations of the complex exponential solutions near the left turning point to write

$$\psi_L(x) = \frac{A_L}{\sqrt{p(x)}} \cos\left(\int_a^x k(x) dx - C_L\pi\right) \quad (10.47)$$

⁷ See, for example, Park (1992); I also like the discussion in Migdal and Krainov (1969).

For an infinite wall type boundary condition, it is easy to see that $C_L = 1/2$ since we require that

$$\psi_L(a) \propto \cos(-C_L\pi) = 0 \quad (10.48)$$

For a smoother potential, one for which one can approximate $V(x)$ near $x = a$ by a linear function,⁸ the appropriate value of C_L turns out to be $1/4$. This can be interpreted as saying that the quantum wavefunction penetrates $\pi/4 = 2\pi/8$ or $\sim 1/8$ of a local wavelength into the classically disallowed region.

10.3.2 WKB Quantized Energy Levels

One of the most useful results arising from the WKB method is a semiclassical estimate for the quantized energy levels in a potential. Matching the WKB wavefunctions at each of the two classical turning points yields two, presumably equivalent descriptions of $\psi(x)$ inside the well, namely

$$\psi_L(x) = \frac{A_L}{\sqrt{p(x)}} \cos\left(\int_a^x k(x) dx - C_L\pi\right) \quad (10.49)$$

and

$$\psi_R(x) = \frac{A_R}{\sqrt{p(x)}} \cos\left(\int_x^b k(x) dx - C_R\pi\right) \quad (10.50)$$

If these two solutions are to agree, we must clearly have $|A_L| = |A_R|$; then comparing the arguments of the cosines we find that

$$\int_a^b k(x) dx - (C_L + C_R)\pi = n\pi \quad \text{for } n = 0, 1, 2, \dots \quad (10.51)$$

This implies that

$$\int_a^b k(x) dx = (n + C_L + C_R)\pi \quad \text{for } n = 0, 1, 2, \dots \quad (10.52)$$

or

$$\int_a^b \sqrt{2m(E - V(x))} dx = (n + C_L + C_R)\pi\hbar \quad (10.53)$$

Recalling that $k(x) = 2\pi/\lambda(x)$, we see that Eqn. (10.52) is simply a more sophisticated version of “fitting an integral number of de Broglie half-wavelengths in a box” and generalizes the Bohr–Sommerfeld quantization condition. The value of n can be seen to count the number of nodes in the quantum wavefunction.

⁸ In this case, the solution which interpolates between the inside and outside can be described by an Airy function (See Appendix E.2).

Example 10.6. Infinite well and harmonic oscillator

For the standard infinite well, we have $C_L = C_R = 1/2$ (infinite wall boundary conditions) and $k(x) = \sqrt{2mE/\hbar^2}$, so that the WKB quantization condition gives

$$\int_0^a \sqrt{2mE} dx = (n+1)\pi \quad (10.54)$$

or

$$E_n = \frac{\hbar^2(n+1)^2\pi^2}{2ma^2} \quad \text{for } n = 0, 1, 2, \dots \quad (10.55)$$

which is the exact answer.

More interestingly, the WKB quantization also gives the correct answer for the harmonic oscillator. In that case we have

$$p(x) = \sqrt{2m(E - m\omega^2 x^2/2)} \quad \text{or} \quad k(x) = \frac{m\omega}{\hbar} \sqrt{A^2 - x^2} \quad (10.56)$$

where $E = m\omega^2 A^2/2$. Since $C_L = C_R = 1/4$ in this case (noninfinite walls) we find

$$\int_{-A}^{+A} k(x) dx = \frac{m\omega}{\hbar} \int_{-A}^{+A} \sqrt{A^2 - x^2} dx = (n+1/2)\pi \quad (10.57)$$

or

$$E_n = (n+1/2)\hbar\omega. \quad (10.58)$$

Once again, the harmonic oscillator problem can be solved exactly with seemingly every method brought to bear.

We have noted that we have dropped terms of order $\mathcal{O}(\hbar^2)$ or $1/(lk)^2$; since typically we find $k_n \propto n/l$, we expect the WKB estimates of the energies to have errors of order $\mathcal{O}(1/n^2)$. This is consistent with our keeping the C_L, C_R terms in Eqn. (10.52), which, in this language, are of order $\mathcal{O}(1/n)$.

10.4 Matrix Methods

The variational method relies on the expansion of a general quantum state in terms of energy eigenstates. In this section, we describe a matrix approach, which also uses the algebraic structure inherent in the Schrödinger equation, but in a rather different way.

Suppose that we have solved for the energy eigenstates of some Hamiltonian operator, \hat{H} . We call them $\psi_n(x)$ where we let the label n start with $n = 1$ for

notational convenience; with this labeling, the ground state is $\psi_1(x)$, the first excited state $\psi_2(x)$, and so on.

We know that a general wavefunction can be expanded in such eigenstates via $\psi(x) = \sum_{n=1}^{\infty} a_n \psi_n(x)$ and from Section 6.4 we know that the information content in $\psi(x)$ and the expansion coefficients, $\{a_n\}$, is the same. We can write the collected $\{a_n\}$ as an (infinite-dimensional) vector \mathbf{a}

$$\psi(x) \Longleftrightarrow \{a_n\} \Longleftrightarrow \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} \Longleftrightarrow \mathbf{a} \quad (10.59)$$

where we demand that $\sum_n |a_n|^2 = 1$ for proper normalization. In this language, individual energy eigenstates are written as

$$\psi_1(x) \Longleftrightarrow \{a_1 = 1, a_{n>1} = 0\} \Longleftrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \Longleftrightarrow \mathbf{e}_1 \quad (10.60)$$

and so forth. The set of vectors, \mathbf{e}_i , corresponding to eigenfunctions are said to form a *basis* for the infinite-dimensional vector space; they are like the unit vectors of a more physical vector space. We then have $\mathbf{a} = \sum_i a_i \mathbf{e}_i$.

The Schrödinger equation $\hat{H}\psi = E\psi$ can be written in the form

$$\hat{H} \left(\sum_m a_m \psi_m(x) \right) = E \left(\sum_m a_m \psi_m(x) \right) \quad (10.61)$$

so that if we multiply both sides by $\psi_n^*(x)$ (on the left, as usual) and integrate we find that

$$\sum_m \langle \psi_n | \hat{H} | \psi_m \rangle a_m = E \sum_m \langle \psi_n | \psi_m \rangle a_m = E \sum_m \delta_{n,m} a_m = E a_n \quad (10.62)$$

We then choose to identify

$$\langle \psi_n | \hat{H} | \psi_m \rangle \equiv H_{nm} \quad (10.63)$$

with the n, m th element of a *matrix* \mathbf{H} in which case the Schrödinger equation takes the form of a matrix eigenvalue problem (see Appendix F.1), namely

$$\begin{pmatrix} H_{11} & H_{12} & H_{13} & \cdots \\ H_{21} & H_{22} & H_{23} & \cdots \\ H_{31} & H_{32} & H_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} = E \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} \quad (10.64)$$

or

$$\mathbf{H}\mathbf{a} = E\mathbf{a} \quad (10.65)$$

for short. The H_{nm} are called the *matrix elements* of the Hamiltonian and are said to form a *matrix representation* of the operator \hat{H} .

We note that \mathbf{H} can always be evaluated using a particular set of basis vectors, namely the eigenfunctions of \hat{H} itself. In this particular case, the matrix takes an especially simple form namely, because

$$H_{nm} = \langle \psi_n | \hat{H} | \psi_m \rangle = \langle \psi_n | E_m | \psi_m \rangle = E \delta_{nm} \quad (10.66)$$

so that the matrix \mathbf{H} is *diagonal*. Thus, Eqn. (10.64) takes the form

$$\begin{pmatrix} E_1 & 0 & 0 & \cdots \\ 0 & E_2 & 0 & \cdots \\ 0 & 0 & E_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} = E \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \end{pmatrix} \quad (10.67)$$

The only way Eqn. (10.67) can be satisfied is if

$$\det(\mathbf{H} - E \mathbf{1}) = \det \begin{vmatrix} E_1 - E & 0 & 0 & \cdots \\ 0 & E_2 - E & 0 & \cdots \\ 0 & 0 & E_3 - E & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} = 0 \quad (10.68)$$

where $\mathbf{1}$ is the unit matrix. This is equivalent to

$$(E_1 - E)(E_2 - E)(E_3 - E) \cdots = \prod_{n=1}^{\infty} (E_n - E) = 0 \quad (10.69)$$

so that the energy eigenvalues are simply the E_n , as we knew; the corresponding eigenvectors of the matrix equation are then simply the \mathbf{e}_i (Why?).

We then say that:

- The matrix representation of a Hamiltonian, when evaluated using its eigenfunctions as a basis, is diagonal and the diagonal entries are just its energy eigenvalues.

Matrix representations of other operators can also be generated. For example, the position and momentum operators, x and \hat{p} , have matrix counterparts denoted by \mathbf{x} and \mathbf{p} and are defined via

$$\mathbf{x}_{nm} = \langle \psi_n | x | \psi_m \rangle \quad \text{and} \quad \mathbf{p}_{nm} = \langle \psi_n | \hat{p} | \psi_m \rangle \quad (10.70)$$

Such matrix representations satisfy the usual rules of matrix algebra, namely

$$(\mathbf{x}^2)_{nm} = \sum_k \mathbf{x}_{nk} \mathbf{x}_{km} \quad (10.71)$$

or more explicitly

$$\begin{pmatrix} \mathbf{x}_{11}^2 & \mathbf{x}_{12}^2 & \cdots \\ \mathbf{x}_{21}^2 & \mathbf{x}_{22}^2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} & \cdots \\ \mathbf{x}_{21} & \mathbf{x}_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} & \cdots \\ \mathbf{x}_{21} & \mathbf{x}_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad (10.72)$$

The matrix representation for the kinetic energy operator is, for example, $\mathbf{T} = \mathbf{p}^2/2m$ or

$$\mathbf{T}_{nm} = \frac{1}{2m} \sum_k \mathbf{p}_{nk} \mathbf{p}_{km} \quad (10.73)$$

Example 10.7. Matrix representation of the harmonic oscillator

We can make use of the results of Chapter 9 to evaluate many of these matrix representations for the specific case of the harmonic oscillator. Using the standard energy eigenvalues we find that

$$\mathbf{H} = \frac{\hbar\omega}{2} \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 3 & 0 & \cdots \\ 0 & 0 & 5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (10.74)$$

Using the results in Section 9.2.1, we then find that

$$\mathbf{x}_{nm} = \sqrt{\frac{\hbar}{2m\omega}} \left(\delta_{n,m-1} \sqrt{m} + \delta_{n,m+1} \sqrt{m+1} \right) \quad (10.75)$$

and we also quote the result

$$\begin{aligned} \mathbf{x}_{nm}^2 = \frac{\hbar}{2m\omega} & \left(\delta_{n,m+2} \sqrt{(m+1)(m+2)} + (2n+1)\delta_{n,m} \right. \\ & \left. + \delta_{n,m-2} \sqrt{m(m-1)} \right) \end{aligned} \quad (10.76)$$

We can check that the matrix equation

$$\mathbf{x}_{nm}^2 = \sum_k \mathbf{x}_{nk} \mathbf{x}_{km} \quad (10.77)$$

(Continued)

holds explicitly by comparing

$$\frac{\hbar}{2m\omega} \begin{pmatrix} 1 & 0 & \sqrt{1 \cdot 2} & 0 & \dots \\ 0 & 3 & 0 & \sqrt{2 \cdot 3} & \dots \\ \sqrt{1 \cdot 2} & 0 & 5 & 0 & \dots \\ 0 & \sqrt{2 \cdot 3} & 0 & 7 \dots & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (10.78)$$

$$\sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \cdot \sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Similar results hold for p and p^2 and one can show (P10.16) that

$$H_{nm} = \frac{1}{2m} p_{nm}^2 + \frac{m\omega^2}{2} x_{nm}^2 \quad (10.79)$$

holds as a matrix equation.

The *average* or *expectation value* of an operator in any state can also be written in this language. For example, we have

$$\begin{aligned} \langle x \rangle &= \langle \psi | x | \psi \rangle = \left\langle \sum_n a_n \psi_n \left| x \right| \sum_m a_m \psi_m \right\rangle \\ &= \sum_{n,m} a_n^* \langle \psi_n | x | \psi_m \rangle a_m \\ &= \sum_{n,m} a_n^* x_{nm} a_m \end{aligned} \quad (10.80)$$

with similar expressions for other operators. We can also easily include the time-dependence for any state via

$$\mathbf{a}(t) \Longleftrightarrow \sum_i a_i \mathbf{e}_i e^{-iE_i t/\hbar} \Longleftrightarrow \begin{pmatrix} a_1 e^{-iE_1 t/\hbar} \\ a_2 e^{-iE_2 t/\hbar} \\ \vdots \end{pmatrix} \quad (10.81)$$

The expectation value of the energy operator can be checked to satisfy

$$\langle \hat{E} \rangle_t = \sum_n |a_n|^2 E_n \quad (10.82)$$

independent of time because the energy matrix is diagonal; other average values have less trivial time-dependence (P10.17) in agreement with earlier examples.

Thus far we have only considered the case in which we already know the energy eigenfunctions and eigenvalues of the Hamiltonian operator \hat{H} . In this instance, the discussion above is interesting, but provides little new information; we have just provided yet another representation of the solution space. If, on the other hand, we did not know the stationary states we could still proceed as follows:

1. Pick a convenient set of energy eigenfunctions to *some* problem, called $\zeta_n(x)$; we immediately know that they form a complete set so that the expansion theorem will work.
2. Evaluate the Hamiltonian matrix using this set of basis functions, that is calculate

$$\mathbf{H}_{nm} \equiv \langle \zeta_n | \hat{H} | \zeta_m \rangle = \frac{1}{2m} \langle \zeta_n | \hat{p}^2 | \zeta_m \rangle + \langle \zeta_n | V(x) | \zeta_m \rangle \quad (10.83)$$

In this case, \mathbf{H} will no longer be diagonal.

3. The Schrödinger equation in matrix form is still an eigenvalue problem of the form in Eqn. (10.64); its eigenvalues are determined by the condition that $\det(\mathbf{H} - E \mathbf{1}) = 0$.
4. If the eigenvalues are labeled via E_i and the corresponding eigenvectors by $\mathbf{a}^{(i)}$, the position-space wavefunctions are given by $\psi_i(x) = \sum_n^\infty a_n^{(i)} \zeta_n(x)$.

Since finding the *exact* eigenvalues and eigenvectors of an infinite-dimensional matrix is only possible in very special cases, to use this method as a real calculational tool we most often restrict ourselves to a truncated version of the problem. More specifically, we try to diagonalize the $N \times N$ submatrix in the upper left-hand corner for some finite value of N . As N is made larger, we expect to obtain an increasingly good representation of the exact result. Because there exist powerful techniques for diagonalizing large matrices, especially if they happen to have large numbers of vanishing components (so-called sparse matrices), this technique is well suited for numerical computations.

Example 10.8. Infinite well plus δ -function I: Matrix methods

As an example of this method, consider the potential discussed in P8.8, namely, a symmetric infinite well defined via

$$V(x) = \begin{cases} 0 & \text{for } |x| < a \\ +\infty & \text{for } |x| > a \end{cases} \quad (10.84)$$

plus a δ -function potential spike at the origin,

$$V_g(x) = g\delta(x) \quad (10.85)$$

This problem can be solved exactly and hence is useful as a testing ground for various approximation techniques. (For a thorough discussion, see Lapidus (1987).)

We know that the odd states are unaffected by $V_g(x)$ since they all possess nodes at $x = 0$. We thus only consider the even states only for which the energy eigenvalue condition can be written as

$$\lambda = -2y \cot(y) \quad (10.86)$$

where

$$\lambda \equiv \frac{2mag}{\hbar^2} \quad \text{and} \quad E = \frac{\hbar^2 y^2}{2ma^2} \quad (10.87)$$

We naturally choose as a set of basis functions the even solutions of the symmetric well *without* the δ -function potential, that is,

$$\psi_n(x) = \frac{1}{\sqrt{a}} \cos\left(\frac{(n-1/2)\pi x}{a}\right) \quad (10.88)$$

Evaluating the Hamiltonian matrix with this basis set, we find

$$H = \begin{pmatrix} \hbar^2\pi^2/8ma^2 + g/a & g/a & g/a & \cdots \\ g/a & 9\hbar^2\pi^2/8ma^2 + g/a & g/a & \cdots \\ g/a & g/a & 25\hbar^2\pi^2/8ma^2 + g/a & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (10.89)$$

or

$$H = \left(\frac{\hbar^2\pi^2}{8ma^2}\right) \begin{pmatrix} 1 + \epsilon & \epsilon & \epsilon & \cdots \\ \epsilon & 9 + \epsilon & \epsilon & \cdots \\ \epsilon & \epsilon & 25 + \epsilon & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (10.90)$$

where $\epsilon \equiv 4\lambda/\pi^2$. Since we rely on matrix diagonalization methods ("canned" packages exist in many programming languages which will find the eigenvalues and eigenvectors of

(Continued)

matrices), we must choose some specific numerical values. For $\lambda = 5$, the exact even energy eigenvalues (in terms of $\hbar^2\pi^2/8ma^2$) obtained from Eqn. (10.86) are

$$2.2969 \text{ (1)} \quad 10.8048 \text{ (9)} \quad 26.9303 \text{ (25)} \quad 50.9743 \text{ (49)} \quad (10.91)$$

where the terms in parentheses are the values without the δ -function term.

Using an available package (in this case *Mathematica*®), we find the eigenvalues for increasingly large $N \times N$ truncated basis sets:

$$\begin{array}{ccccc} 1 \times 1 & 3.02642 & \text{---} & \text{---} & \text{---} \\ 2 \times 2 & 2.54241 & 11.5104 & \text{---} & \text{---} \\ 3 \times 3 & 2.44832 & 11.1412 & 27.4898 & \text{---} \\ 4 \times 4 & 2.40672 & 11.0335 & 27.2396 & 51.4298 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 10 \times 10 & 2.33857 & 10.8861 & 27.0277 & 51.07991 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 50 \times 50 & 2.30506 & 10.8204 & 26.9487 & 50.9937 \end{array} \quad (10.92)$$

It does seem that the eigenvalues of the truncated set approach the exact values at N is increased. Such programs also give the eigenvectors as well; we display the components corresponding to the ground state solution below:

$$\begin{array}{cc} 1 \times 1 & (1) \\ 2 \times 2 & (0.97264, -0.23232) \\ 3 \times 3 & (0.97451, -0.21543, -0.06258) \\ 4 \times 4 & (0.97575, -0.20818, -0.06075, -0.02946) \end{array} \quad (10.93)$$

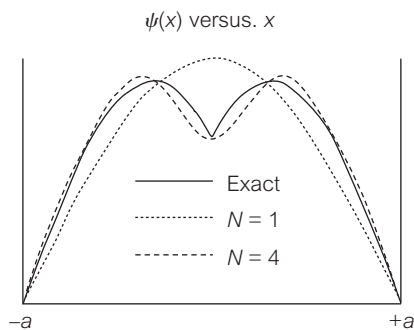


Figure 10.7. The exact (solid) and two approximate solutions for Example 10.8.

Using these values, we illustrate in Fig. 10.7 the approximations to the ground state wavefunction for the first four approximations, comparing them to the exact solution, with the cusp expected from the singular δ -function; the convergence to the exact solution is not particularly rapid.

10.5 Perturbation Theory

We now turn to what is undoubtedly the most widely used approximation method we will discuss, perturbation theory. We are certainly used to the notion of the systematic expansion of some quantity in terms of a small parameter; a familiar example is the series expansion of a function,

$$f(x) = f(0) + f'(0)x + \frac{1}{2!}f''(0)x^2 + \dots \quad (10.94)$$

Such an expansion may well formally converge for all values of x (such as for the series for $\exp(x)$), but is often most useful as a calculational tool when $|x| \ll 1$.

Perturbation theory extends this notion to quantum mechanics in cases where the system under study can be described by an “unperturbed” Hamiltonian, \hat{H}_0 , for which the energy eigenstates can be obtained exactly, that is,

$$\hat{H}_0 \psi_n^{(0)} = E_n^{(0)} \psi_n^{(0)} \quad (10.95)$$

We began this chapter with the observation that many important systems such as the hydrogen atom or the harmonic oscillator can actually be solved exactly. One can then imagine “turning on” an additional perturbing interaction, \hat{H}' , which will change the spectrum and wavefunctions; examples include the addition of an electric field acting on a charged particle (via a term $\hat{H}' = V(x) = -qEx$ in one dimension) or a magnetic field acting on a magnetic moment ($\hat{H}' = -\mu \cdot \mathbf{B}$). (While we will most often consider the case where the perturbation is a (small) additional potential energy function, other cases are possible (P10.20).) We can then write

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}' \quad (10.96)$$

where we introduce a dimensionless parameter λ (which can be set equal to unity at the end of the calculation) to act as an expansion parameter. Our goal is then to solve the Schrödinger equation for the complete system,

$$\hat{H} \psi_n = E_n \psi_n \quad (10.97)$$

as a series in λ .

We focus in the next two sections on time-independent problems, but briefly discuss problems involving perturbations which evolve in time in Section 10.5.3.

10.5.1 Nondegenerate States

We will begin by making the assumption that the energy levels of the unperturbed system are all distinct, that is, that there are no degeneracies where $E_n^{(0)} \approx E_l^{(0)}$

for some pair l, n . Then, as we imagine $\lambda \rightarrow 0$, we can unambiguously write

$$\psi_n \xrightarrow{\lambda \rightarrow 0} \psi_n^{(0)} \quad \text{and} \quad E_n \xrightarrow{\lambda \rightarrow 0} E_n^{(0)} \quad (10.98)$$

and make a unique identification of each perturbed state with its unperturbed counterpart. Motivated by these assumptions, we first write

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (10.99)$$

as a series in λ . Then, since the unperturbed eigenstates form a complete set, we always have

$$\psi_n = \sum_{j=0}^{\infty} a_{nj} \psi_j^{(0)} = a_{nn} \psi_n^{(0)} + \sum_j' a_{nj} \psi_j^{(0)} \quad (10.100)$$

where \sum_j' denotes the infinite sum with the $j = n$ term removed. The new (perturbed) eigenstates can always be written as a linear combination of the old (unperturbed) eigenstates.

The coefficients have slightly different expansions in λ ,

$$a_{nn} = a_{nn}^{(0)} + \lambda a_{nn}^{(1)} + \lambda^2 a_{nn}^{(2)} + \dots \quad (10.101)$$

$$a_{nj} = \lambda a_{nj}^{(1)} + \lambda^2 a_{nj}^{(2)} + \dots \quad \text{for } j \neq n \quad (10.102)$$

because Eqn. (10.98) implies that

$$\lim_{\lambda \rightarrow 0} a_{nj} = \delta_{nj} \quad (10.103)$$

We can constrain the expansion coefficients of Eqn. (10.100) further by noting that the normalization condition

$$\sum_{j=1}^{\infty} |a_{nj}|^2 = 1 \quad \text{for all } n \quad (10.104)$$

implies that

$$1 = |a_{nn}|^2 + \sum_j' |a_{nj}|^2 = |a_{nn}|^2 + \sum_j' \left(\lambda a_{nj}^{(1)} + \dots \right)^2 = |a_{nn}|^2 + \mathcal{O}(\lambda^2) \quad (10.105)$$

so that

$$a_{nn} \approx 1 \quad \text{to } \mathcal{O}(\lambda^2) \quad \text{which implies that } a_{nn}^{(1)} = 0 \quad (10.106)$$

The Schrödinger equation can now be written (to $\mathcal{O}(\lambda^2)$) in the form

$$\begin{aligned} (\hat{H}_0 + \lambda \hat{H}') \left(\psi_n + \lambda \sum_j' a_{nj} \psi_j^{(0)} + \dots \right) = \\ (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots) \left(\psi_n + \lambda \sum_j' a_{nj} \psi_j^{(0)} + \dots \right) \end{aligned} \quad (10.107)$$

We first multiply Eqn. (10.107) by $(\psi_n^{(0)})^*$ on the left and integrate to obtain

$$\begin{aligned} \left\langle \psi_n^{(0)} \left| \hat{H}_0 + \lambda \hat{H}' \right| \left(\psi_n + \lambda \sum_j' a_{nj} \psi_j^{(0)} + \dots \right) \right\rangle \\ = (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots) \left\langle \psi_n^{(0)} \left| \left(\psi_n + \lambda \sum_j' a_{nj} \psi_j^{(0)} + \dots \right) \right\rangle \right\end{aligned} \quad (10.108)$$

Equating powers of λ and making extensive use of the orthogonality of the unperturbed wavefunctions, namely that $\langle \psi_n^{(0)} | \psi_j^{(0)} \rangle = \delta_{nj}$, we find

$$\mathcal{O}(\lambda^0) : E_n^{(0)} = \langle \psi_n^{(0)} | \hat{H}_0 | \psi_n^{(0)} \rangle \quad (10.109)$$

$$\mathcal{O}(\lambda^1) : E_n^{(1)} = \langle \psi_n^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle \equiv H'_{nn} \quad (10.110)$$

$$\mathcal{O}(\lambda^2) : E_n^{(2)} = \sum_j' a_{nj}^{(1)} \langle \psi_n^{(0)} | \hat{H}' | \psi_j^{(0)} \rangle \equiv \sum_j' a_{nj}^{(1)} H'_{nj} \quad (10.111)$$

These expressions all require the *matrix elements* of the *perturbing Hamiltonian*, evaluated using the *unperturbed eigenfunctions*, H'_{nk} .

The $\mathcal{O}(\lambda^0)$ term ($E_n^{(0)}$) simply reproduces the unperturbed energy spectrum. The equation for $E_n^{(1)}$ in Eqn. (10.110) is a very important result as it states that:

- The first-order shift in the energy of level n due to a (small) perturbation is given by the diagonal matrix element of the perturbing Hamiltonian, H'_{nn} , evaluated with the unperturbed wavefunctions, that is,

$$E_n^{(1)} = \langle \psi_n^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle = H'_{nn} \quad (10.112)$$

which we repeat because of its extreme importance.

Using Eqn. (10.112) we see that it can sometimes happen that the first-order energy shift vanishes identically because of symmetry. For example, a charged

particle in the infinite symmetric well subject to a weak electric field, given by a potential of the form $V(x) = -qEx$, would have a first-order energy shift given by

$$E_n^{(1)} = \begin{cases} -qE\langle u_n^{(+)} | x | u_n^{(+)} \rangle & \text{for even states} \\ -qE\langle u_n^{(-)} | x | u_n^{(-)} \rangle & \text{for odd states} \end{cases} \quad (10.113)$$

which vanishes for all states. In such cases, the second-order term $E_n^{(2)}$ is the leading correction.

The form of Eqn. (10.111) also suggests the more general result:

- The k th order correction to the energy levels requires knowledge of the $(k - 1)$ th order wavefunctions

so that to determine $E^{(2)}$ we require the leading-order expansion coefficients, $a_{nj}^{(1)}$.

To obtain information on the expansion coefficients, we multiply Eqn. (10.107) by $(\psi_j^{(0)})^*$ with $j \neq n$ and integrate. The $\mathcal{O}(\lambda^0)$ terms are absent, while the $\mathcal{O}(\lambda^1)$ terms require that

$$a_{nk}^{(1)} = \frac{\langle \psi_k^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle}{(E_n^{(0)} - E_k^{(0)})} = \frac{H'_{nk}}{(E_n^{(0)} - E_k^{(0)})} \quad (10.114)$$

The first-order wavefunction thus receives contributions from every state for which the *off-diagonal* matrix elements are nonvanishing, that is, $H'_{nk} \neq 0$. Combining Eqns. (10.114) and (10.111), we find that the second-order corrections to the energies are given by

$$\begin{aligned} E_n^{(2)} &= \sum_j' a_{nk}^{(1)} H'_{nk} \\ &= \sum_k' \left[\frac{\langle \psi_k^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle}{(E_n^{(0)} - E_k^{(0)})} \right] \langle \psi_n^{(0)} | \hat{H}' | \psi_k^{(0)} \rangle \\ &= \sum_k' \frac{|\langle \psi_n^{(0)} | \hat{H}' | \psi_k^{(0)} \rangle|^2}{(E_n^{(0)} - E_k^{(0)})} \\ E_n^{(2)} &= \sum_k' \frac{|H'_{nk}|^2}{(E_n^{(0)} - E_k^{(0)})} \end{aligned} \quad (10.115)$$

In the last step, we have made use of the fact that \hat{H}' is Hermitian to write $H'_{nk} = (H'_{kn})^*$; this form makes it clear that the second-order shift in energy is manifestly real. This result has many interesting consequences:

- The second-order shift depends on the off-diagonal matrix elements, but inversely weighted by the “distance in energy” to the state in question; thus, in general, states nearby in energy have a larger effect.
- This form also implies that the spacing in energy levels must be larger than the matrix elements of the perturbation for the expansion to be valid, that is, we demand that

$$H'_{nk} \ll |E_n^{(0)} - E_k^{(0)}| \quad (10.116)$$

This shows that degenerate energy levels must clearly be handled in a different way.

- States with energy below (above) a given level induce a second-order energy shift which is positive (negative); this effect is often referred to as “level repulsion”.
- In particular, the second-order shift in the ground state energy is clearly always negative as all the other states lie above it. For many problems for which there are large numbers of levels, one can argue heuristically that the second-order shift for any fixed energy level will be negative due to the large number of states above it; this can be motivated on more physical grounds⁹ and is often observed.

The second-order expansion coefficients (the $a_{nj}^{(2)}$) are too complicated to reproduce here but, for reference, we state without proof that the result for the third-order shift in energies is

$$E_n^{(3)} = \sum_k' \sum_j' \frac{H'_{nk} H_{kj} H'_{jn}}{(E_n^{(0)} - E_k^{(0)})(E_n^{(0)} - E_j^{(0)})} - H'_{nn} \sum_k' \frac{|H'_{nk}|^2}{(E_n^{(0)} - E_k^{(0)})^2} \quad (10.117)$$

We see that the work required to continue the perturbation theory expansion increases rapidly, so that often only the first- and second-order corrections are calculated. We now turn to some examples.

⁹ See the nice discussion by Saxon (1968).

Example 10.9. Harmonic oscillator with applied electric field

The problem of a charged oscillator in a constant electric field, described by the Hamiltonian

$$\hat{H} = \frac{1}{2}m\omega^2 x^2 - Fx \quad (10.118)$$

(where $F = qE$) was investigated in P9.9 where it was shown that it could be solved exactly; the resulting energy spectrum is

$$E_n = (n + 1/2)\hbar\omega - \frac{F^2}{2m\omega^2} \quad (10.119)$$

Let us approach this problem by considering the electric field interaction to be a small perturbation about the unperturbed oscillator, that is, $\hat{H}' = -Fx$; we can then use F as an expansion coefficient to count powers in perturbation theory. We have $E_n^{(0)} = (n + 1/2)\hbar\omega$, of course, while the first-order correction vanishes (because of symmetry) since

$$E_n^{(1)} = \langle \psi_n | \hat{H}' | \psi_n \rangle = -qE \langle \psi_n | x | \psi_n \rangle = 0 \quad (10.120)$$

The second-order correction is given by

$$E_n^{(2)} = F^2 \sum_k' \frac{|\langle \psi_n | x | \psi_k \rangle|^2}{(E_n^{(0)} - E_k^{(0)})} = F^2 \sum_k' \frac{|\langle n | x | k \rangle|^2}{(n - k)\hbar\omega} \quad (10.121)$$

Using the results in Section 9.2.2, we know that

$$\langle n | x | k \rangle = \sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n} \delta_{k,n-1} + \sqrt{n+1} \delta_{k,n+1} \right) \quad (10.122)$$

Inserting this result into Eqn. (10.121), we find that

$$\begin{aligned} E_n^{(2)} &= \frac{F^2}{\hbar\omega} \left(\frac{\hbar}{2m\omega} \right) \sum_k' \frac{(n\delta_{k,n-1} + (n+1)\delta_{k,n+1})}{(n-k)} \\ &= \frac{F^2}{2m\omega^2} \left(\frac{n}{1} + \frac{(n+1)}{-1} \right) \\ &= -\frac{F^2}{2m\omega^2} \end{aligned} \quad (10.123)$$

which reproduces the exact answer. One would then expect that all of the higher-order corrections to the energy would then vanish identically and one can confirm explicitly (P10.21) that the third-order correction in Eqn. (10.117) is indeed zero in this case. It is also an example where the second-order corrections are, in fact, negative for all energy levels.

This does not imply, however, that the expansion coefficients have a similarly simple series behavior. To see this, we can make use of the exact ground state solution to the complete problem (see P9.9 again)

$$\psi(x; F) = \frac{1}{\sqrt{\rho\sqrt{\pi}}} e^{-(x-x_0)^2/2\rho^2} \quad (10.124)$$

(Continued)

where $x_0 = F/m\omega^2$ and $\rho = \sqrt{\hbar/m\omega}$. The expansion coefficient a_{00} is then given by

$$a_{00} \equiv \int [\psi(x; F=0)]^* \psi(x; F) dx = e^{-x_0^2/4\rho^2} = e^{-F^2/F_0^2} \quad (10.125)$$

where $F_0 \equiv 2\sqrt{\hbar m\omega}$. Expanding a_{00} in powers of F we find that

$$a_{00} = 1 - \frac{F^2}{F_0^2} + \frac{1}{2} \frac{F^4}{F_0^4} + \dots \quad (10.126)$$

Thus, while the perturbation series for the energies terminates at second-order, the expansion coefficients require the full series to converge to the exact answer.

Example 10.10. Infinite well plus δ -function II: Perturbation theory

Consider the problem, discussed in P8.8, of the symmetric infinite square well potential plus a δ -function potential at the origin; in this case, let the δ -function constitute the perturbation so that $\hat{H}' = g\delta(x)$.

For the odd case, the explicit application of the boundary conditions for the full problem require that the $u_n^{(-)}(x)$ vanish at the origin and gives the same energy eigenvalue condition as for the infinite well alone. This can be confirmed to any order in perturbation theory since all of the relevant matrix elements in Eqns (10.112), (10.115), and (10.117) vanish explicitly. (See P10.22.)

For the even case, the exact eigenvalue condition was given by $\lambda = -2y \cot(y)$ where $\lambda \equiv 2mag/\hbar^2$ and $E = \hbar^2 y^2/2ma^2$. Focusing only on the ground state, this eigenvalue condition can be expanded to second order (see P10.23) to yield

$$y = \frac{\pi}{2} + \frac{\lambda}{\pi} - 2\frac{\lambda^2}{\pi^3} + \mathcal{O}(\lambda^3) = \frac{\pi}{2} \left(1 + \frac{2\lambda}{\pi^2} - \frac{4\lambda^2}{\pi^4} \right) + \dots \quad (10.127)$$

or

$$\begin{aligned} E_1^{(+)} &\approx \frac{\hbar^2}{2ma^2} \frac{\pi^2}{4} \left(1 + \frac{2\lambda}{\pi^2} - \frac{4\lambda^2}{\pi^4} + \dots \right)^2 \\ &\approx \frac{\hbar^2 \pi^2}{8ma^2} + \lambda \left(\frac{\hbar^2}{2ma^2} \right) - \lambda^2 \left(\frac{\hbar^2}{2\pi^2 ma^2} \right) + \dots \end{aligned} \quad (10.128)$$

The first-order perturbation result for even states is simply

$$(E_n^{(+)})^{(1)} = \langle u_n^{(+)} | g\delta(x) | u_n^{(+)} \rangle = \frac{g}{a} = \lambda \left(\frac{\hbar^2}{2ma^2} \right) \quad (10.129)$$

(Continued)

independent of n for all even states; this obviously agrees with the explicit expansion of the eigenvalue condition for the ground state in Eqn. (10.128). The second-order correction is then

$$\begin{aligned}
 E_n^{(+)(2)} &= \sum_{k=2}^{\infty} \frac{|\langle u_n^{(+)} | g\delta(x) | u_k^{(+)} \rangle|^2}{(\hbar^2\pi^2/8ma^2)((1 - (2k - 1)^2))} \\
 &= -\lambda^2 \frac{2\hbar^2}{m\pi^2 a^2} S \quad \text{where } S \equiv \sum_{k=2}^{\infty} \frac{1}{(2k - 1)^2 - 1} = \frac{1}{4} \\
 &= -\lambda^2 \left(\frac{\hbar^2}{2\pi^2 ma^2} \right) \tag{10.130}
 \end{aligned}$$

which also agrees with the expansion of the exact result. (See Epstein (1960) for a nice discussion of the subtleties of this problem.)

While the technical details of the calculation are beyond our level, it is appropriate to note here that one of the most spectacularly successful predictions in all of physics makes use of perturbation theory. The magnetic moment of both the electron and the muon can be calculated in the theory of quantum electrodynamics,¹⁰ using more advanced perturbation theory methods. A recent theoretical result for the electron magnetic moment (expressed as a dimensionless number) is

$$g_e(\text{theory}) = 2.0023193048 \text{ (8)} \tag{10.131}$$

where the uncertainty is indicated in the last significant digit. Amazingly, it can also be measured to a similar precision with the result

$$g_e(\text{experiment}) = 2.0023193048 \text{ (4)} \tag{10.132}$$

10.5.2 Degenerate Perturbation Theory

When two (or more) energy levels of the unperturbed system are degenerate, any linear combinations of the corresponding wavefunctions, $\psi_n^{(0)}(x)$ and $\psi_l^{(0)}(x)$, still give same energy eigenvalue. (Such combinations can still be made orthogonal to each other, of course.) This implies, however, that the unique identification of each perturbed state with an unperturbed counterpart as in Eqn. (10.98) is not possible. The breakdown of the perturbation method in this case is clearly signaled by the appearance of small energy denominators

¹⁰ See Perkins (2000) and references therein for a discussion at an undergraduate level.

in Eqn. (10.114); states nearby in energy can play an important role and have to be considered on a more equal footing.

In this limit, it is convenient to return to the matrix formulation of the eigenvalue problem and consider the 2×2 submatrix involving the states in question, namely

$$\begin{pmatrix} E_n^{(0)} + \lambda H'_{nn} & \lambda H'_{nl} \\ \lambda H'_{ln} & E_l^{(0)} + \lambda H'_{ll} \end{pmatrix} \begin{pmatrix} a_n \\ a_l \end{pmatrix} = E \begin{pmatrix} a_n \\ a_l \end{pmatrix} \quad (10.133)$$

where the a_n, a_l are the expansion coefficients. In general, for a case with N degenerate levels, the corresponding $N \times N$ submatrix must be considered.

This system of linear equations (for the $a_{n,l}$) will only have a nontrivial solution if the appropriate determinant vanishes, that is,

$$\det \begin{pmatrix} E_n^{(0)} + \lambda H'_{nn} - E & \lambda H'_{nl} \\ \lambda H'_{ln} & E_l^{(0)} + \lambda H'_{ll} - E \end{pmatrix} = 0 \quad (10.134)$$

The special case of an exact degeneracy where $E_n^{(0)} = E_l^{(0)} \equiv \mathcal{E}$ is easiest to treat; in this case, the energy eigenvalues are determined by the condition

$$(E - [\mathcal{E} + \lambda H'_{nn}]) (E - [\mathcal{E} + \lambda H'_{ll}]) - \lambda^2 H'_{ln} H'_{nl} = 0 \quad (10.135)$$

or

$$E_{\pm} = \mathcal{E} + \frac{\lambda}{2} (H'_{nn} + H'_{ll}) \pm \frac{\lambda}{2} \sqrt{(H'_{nn} - H'_{ll})^2 + 4H'_{ln} H'_{nl}} \quad (10.136)$$

The first term is obviously the (common) value of the unperturbed energy while the second is the average of the first-order energy shifts in each level, consistent with the non-degenerate case; the third term, however, can split the two levels and generally removes the degeneracy.

Substituting the result of Eqn. (10.136) into the matrix equation Eqn. (10.133), we find that the expansion coefficients are given by

$$\frac{a_n^{(\pm)}}{a_l^{(\pm)}} = \frac{2H'_{nl}}{(H'_{ll} - H'_{nn}) + \sqrt{(H'_{nn} - H'_{ll})^2 + 4H'_{ln} H'_{nl}}} \quad (10.137)$$

so that the appropriate (unnormalized) eigenfunctions are given by

$$\psi^{(\pm)}(x) \propto a_n^{(\pm)} \psi_n(x) + a_l^{(\pm)} \psi_l(x) \quad (10.138)$$

The actual energy splitting in Eqn. (10.136) clearly depends on λ , that is, on the *magnitude* of the perturbation; the appropriate linear combinations, however, do not, but are determined by the *form* of the perturbation, that is the *relative* sizes of the matrix elements H'_{nn} , H'_{ll} , and H'_{ln} (P10.26).

Because degeneracy of energy levels is far more common in multi-particle or multidimensional systems, we postpone presenting examples until later chapters.

10.5.3 Time-Dependent Perturbation Theory

The problem of determining the future time-development of a given quantum mechanical state, if one knows the exact energy eigenvalues and eigenfunctions of the system, is usually straightforward, since if we have $\hat{H}^{(0)}\psi_n^{(0)} = E_n\psi_n^{(0)}$, then

$$\psi^{(0)}(x, 0) = \sum_{n=0}^{\infty} a_n \psi_n^{(0)} \longrightarrow \psi^{(0)}(x, t) = \sum_{n=0}^{\infty} a_n e^{-iE_n t/\hbar} \psi_n^{(0)} \quad (10.139)$$

A similar problem arises in *time-dependent perturbation theory* when the system is subject to a (small) time-dependent change, so that the resulting Hamiltonian is given by

$$\hat{H} = \hat{H}^{(0)} + \lambda \hat{H}'(t), \quad (10.140)$$

just as in Eqn. (10.96), but now with \hat{H}' depending explicitly on time. We wish to see how the introduction of the perturbing potential changes the time-development of a quantum state, since we now must satisfy

$$\left[\hat{H}^{(0)} + \lambda \hat{H}'(t) \right] \psi(x, t) = \hat{H} \psi(x, t) = i\hbar \frac{\partial \psi(x, t)}{\partial t} \quad (10.141)$$

Because the eigenfunctions of $\hat{H}^{(0)}$ still form a complete set, we can always write

$$\psi(x, t) = \sum_{n=0}^{\infty} a_n(t) e^{-iE_n t/\hbar} \psi_n^{(0)} \quad (10.142)$$

where we now assume that the expansion coefficients, the $a_n(t)$, themselves depend on time. (We can choose this form without loss of generality, thereby defining the $a_n(t)$, but this form is convenient since for the case of no perturbation, it reduces to the standard result in Eqn. (10.139).) We can substitute a solution of the form in Eqn. (10.142) into the time-dependent Schrödinger equation (Eqn. (10.141)), giving the respective left- and right-hand sides

$$\hat{H} \psi(x, t) = \sum_n E_n a_n(t) e^{-iE_n t/\hbar} \psi_n^{(0)} + \sum_n a_n(t) e^{-iE_n t/\hbar} \left[\lambda \hat{H}' \psi_n \right] \quad (10.143)$$

$$\begin{array}{ccc} \Downarrow & & \Downarrow \\ i\hbar \frac{\partial \psi(x, t)}{\partial t} = i\hbar \sum_n \left[\frac{da_n(t)}{dt} - \frac{iE_n}{\hbar} a_n(t) \right] e^{-iE_n t/\hbar} \psi_n & & \end{array} \quad (10.144)$$

Equating these two, cancelling the terms proportional to E_n , and multiplying (on the left, as usual) by $(\psi_m^{(0)})^*$ and integrating, we can make use of the orthonormality properties of the eigenstates to write

$$i\hbar \frac{da_m(t)}{dt} e^{-iE_m t/\hbar} = \sum_n a_n(t) e^{-iE_n t/\hbar} \langle \psi_m^{(0)} | \hat{H}'(t) | \psi_n^{(0)} \rangle \quad (10.145)$$

or

$$\begin{aligned} \frac{da_m(t)}{dt} &= -\frac{i}{\hbar} \sum_n a_n(t) e^{i(E_m - E_n)t/\hbar} \langle \psi_m^{(0)} | \hat{H}'(t) | \psi_n^{(0)} \rangle \\ &\equiv -\frac{i}{\hbar} \sum_n a_n(t) e^{i\omega_{m,n}t} H'_{mn} \end{aligned} \quad (10.146)$$

where

$$\omega_{m,n} \equiv \frac{(E_m - E_n)}{\hbar} \quad \text{and} \quad H'_{mn} \equiv \langle \psi_m^{(0)} | \hat{H}'(t) | \psi_n^{(0)} \rangle \quad (10.147)$$

and H'_{nm} can be called in this context a *transition matrix element*. We recall that the frequency $\omega_{m,n}$ sets the characteristic timescale for any two-state system (Section 4.6). At this stage, the infinite set of coupled equations implied by Eqn. (10.146) is still completely equivalent to the time-dependent Schrödinger equation for the perturbed system, as no approximations have been made.

We now specialize to the case where the initial state of the system is that of an energy eigenstate of the unperturbed Hamiltonian, namely, we assume that $\psi(x, 0) = \psi_k^{(0)}$. We then expect that for a small perturbation the expansion coefficients, $a_m(t)$, for states with $m \neq k$ will be small (since $a_{m \neq k}(t = 0) = \delta_{m,k} = 0$ to begin with), while the single $a_k(t)$ corresponding to the original eigenstate will be of order unity, namely

$$a_m(t) = \begin{cases} \mathcal{O}(\lambda) \ll 1 & \text{for } m \neq k \\ \mathcal{O}(1) & \text{for } m = k \end{cases} \quad (10.148)$$

Using this approximation in Eqn. (10.146), we find that

$$\frac{da_m(t)}{dt} \approx -\frac{i}{\hbar} e^{i\omega_{m,k}t} H'_{mk} \quad (10.149)$$

or

$$a_m(t) = -\frac{i}{\hbar} \int_{t_0}^t e^{i\omega_{m,k}t} H'_{mk}(t) dt \quad (10.150)$$

for $m \neq k$, if we assume that the perturbation is “turned on” at time $t = t_0$. This form is important since it implies that the probability that the particle will be

found in the m th state of the unperturbed system at a later time is given by

$$P_{k \rightarrow m}(t) = |a_m(t)|^2 = \frac{1}{\hbar^2} \left| \int_{t_0}^t e^{i\omega_{m,k}t} \mathbf{H}'_{mk}(t) dt \right|^2 \quad (10.151)$$

We expect that the chance of being excited (or decaying) into a different final state (the *transition probability*, $P_{k \rightarrow m}(t)$) will depend both on the transition matrix element (which includes information on the perturbation *and* its ability to “connect” the initial and final states) as well as on the detailed history of how the perturbation is applied in time, via the integral, which is weighted by the oscillatory exponential factor present in all two-state systems.

Example 10.11. Harmonic oscillator in a time-dependent electric field

Consider a particle of mass m and charge q in the ground state, $|0\rangle$, of the harmonic oscillator potential $V(x) = m\omega^2 x^2/2$. It is subject to an external time-dependent electric field of the form

$$\hat{H}'(t) = (-q\mathcal{E}_0 x) e^{-t^2/2\tau^2} \quad (10.152)$$

The perturbation is allowed to act over the time interval $(-\infty, +\infty)$ and we wish to evaluate the probability that the particle is eventually found in any excited state, $|n\rangle$.

The expansion coefficient from Eqn. (10.150) in this case is given by

$$a_m(t) = \frac{i}{\hbar} \langle n | q\mathcal{E}_0 x | 0 \rangle \int_{-\infty}^{+\infty} e^{i\omega_{n,0}t} e^{-t^2/2\tau^2} dt \quad (10.153)$$

The off-diagonal oscillator matrix element is given by Eqn. (9.50) as

$$\langle n | x | 0 \rangle = \sqrt{\frac{\hbar}{2m\omega}} \delta_{n,1} \quad (10.154)$$

while the integral over time is a standard Gaussian form. Combining these results, we find that the transition probability to any excited state is given by

$$P_{0 \rightarrow n}(t) = |a_n(t)|^2 = \frac{1}{\hbar^2} \left[\frac{q^2 \mathcal{E}_0^2 \hbar}{2m\omega} \delta_{n,1} \right] \left[2\pi\tau^2 e^{-(\omega_{n,0}\tau)^2} \right] \quad (10.155)$$

This result does show that transition probabilities in time-dependent perturbation theory can depend on *what* the perturbation is, to *which* states it is trying to couple, and *how* it is applied.

- The explicit form of the electric field potential in this case, for example, has implied that only adjacent levels can be excited (to this order of perturbation theory, at least) which is reminiscent of the selection rules discussed in Section 16.3.3. For an initial state $|k\rangle$, only the final states $|k \pm 1\rangle$ would be populated by this perturbation. The transition probability does, of course, also depend on the strength of the external field.

(Continued)

- The transition probability in this case is peaked at a value of $\omega_{n,0}\tau = 1$ so that application of the external perturbing potential at a rate (given by the timescale τ) which matches the natural frequency or period of the system gives the biggest effect. For the special case of the oscillator, $\omega_{n,0} \delta_{n,1} = \omega = 2\pi/T_{cl}$ is precisely the classical periodicity of the problem, so perturbations which occur on timescales much longer than this allow the system to accommodate itself to the change. This connection to the classical periodicity is much more general as we recall (P1.16 and Section 12.7)) that the classical period of a quantum system can be written as $T_{cl} = 2\pi\hbar/|dE_n/dn| \sim 2\pi/|\Delta E_{n,m}|$.

We can examine the result of applying the perturbation slowly over a long timescale in some detail, by writing Eqn. (10.150) in a way which includes information on the rate at which \hat{H}' is applied more directly. For example, using an identity and an integration-by-parts (IBP) trick, we have

$$\begin{aligned}
 a_m(t) &= -\frac{i}{\hbar} \int_{t_0}^t e^{i\omega_{m,k}t} H'_{mk}(t) dt \\
 &= -\frac{1}{\hbar\omega_{m,k}} \int_{t_0}^t \frac{d}{dt} \left[e^{i\omega_{m,k}t} \right] H'_{mk}(t) dt \\
 &\stackrel{\text{IBP}}{=} -\frac{1}{\hbar\omega_{m,k}} \left[-\int_{t_0}^t e^{i\omega_{m,k}t} \left(\frac{dH'_{mk}}{dt} \right) dt + e^{i\omega_{m,k}t} H'_{mk}(t) \right] \quad (10.156)
 \end{aligned}$$

If the rate at which the perturbation is applied, as encoded in the dH'_{mk}/dt term, is slow enough that that term can be neglected, we have the simplified result

$$a_m(t) = \frac{H'_{mk}(t)}{(E_k - E_m)} e^{i(E_m - E_k)t/\hbar} \quad (10.157)$$

Since $a_k(t) \sim \mathcal{O}(1) \cdot e^{-iE_k t/\hbar}$, the time-dependence of the perturbed state is simply

$$\begin{aligned}
 \psi(x, t) &= e^{-iE_k t/\hbar} \psi_k^{(0)} + \sum_{m \neq k} \left[\frac{H'_{mk}(t)}{(E_m - E_k)} e^{i(E_m - E_k)t/\hbar} \right] e^{-iE_m t/\hbar} \psi_m^{(0)} \\
 &= \left(\psi_k^{(0)} + \sum_{m \neq k} \frac{H'_{mk}(t)}{(E_m - E_k)} \psi_m^{(0)} \right) e^{-iE_k t/\hbar} \quad (10.158)
 \end{aligned}$$

We stress that the time-dependence of this state is simply that of a single eigenstate, while the form of the (spatial) wavefunction is that due to a first-order perturbation theory treatment of $\hat{H}'(t)$, as in Eqn. (10.114), evaluated at time t .

This implies that in the limit of slow application of the perturbation, the individual eigenstates of the old system “morph” into the corresponding eigenstates of the new system, in a state-by-state or quantum-number by quantum-number manner; thus the ground state of the new system corresponds to the ground state of the old system, the first excited state to the first excited state, and so forth. This limit of a slowly acting perturbation is called the *adiabatic approximation* and was used (implicitly) in P5.21.

The opposite extreme, where the system undergoes a very rapid change, is called the *sudden approximation* and in this case we have a discontinuous change in the Hamiltonian of the system. For example, if we have $\hat{H}_1 \rightarrow \hat{H}_2$ at $t = 0$, the eigenfunctions of the system can be written as

$$\hat{H}_1 \psi_n = E_n \psi_n \quad t < 0 \text{ (original system)} \quad (10.159)$$

$$\hat{H}_2 \phi_n = \mathcal{E}_n \phi_n \quad t > 0 \text{ (new system)} \quad (10.160)$$

In this case, if the initial state was that of an eigenstate of the original system (ψ_n) for $t < 0$, then in the new universe of solutions we have the standard expansion theorem result that

$$\psi(x, 0) = \psi_n = \sum_k c_k \phi_k \quad \text{with expansion coefficient } c_k = \langle \phi_k | \psi_n \rangle \quad (10.161)$$

and the future time-dependence of the original eigenstate in the new system will be given by

$$\psi(x, t) = \sum_k c_k e^{-i\mathcal{E}_k t/\hbar} \phi_k \quad (10.162)$$

We have also made implicit use of this method in earlier problems (Example 6.3 and P9.6) with a more physical use discussed in P17.8.

10.6 Questions and Problems

- Q10.1.** If you are given a numerical solution of the Schrödinger equation in the form of a list of values at discrete points, that is, $\psi(x = n\epsilon)$, how would you normalize the solution? How would you find $\langle x \rangle$? How about $\langle \hat{p} \rangle$? How would calculate the momentum-space wavefunction, $\phi(p)$?
- Q10.2.** Distinguish carefully between the *precision* and the *accuracy* of a measurement. For example, for a given approximation method, you can imagine determining

the range over which the solutions change their sign at infinity more and more precisely; is this increased precision or increased accuracy? Do you think that decreasing the step size ϵ or using a better integration method results in increased accuracy or increased precision?

- Q10.3.** Assume that you have a program which numerically integrates the Schrödinger equation and that you have found two energy values, E_a and E_b , which bracket an acceptable (i.e. square-integrable) solution of the SE. Describe an efficient strategy to get arbitrarily close to the “real” energy. Hint: If someone tells you they have a number between 1 and 1000, what is the optimal strategy to find their number using the minimum number of “yes–no” questions.
- Q10.4.** If you have a program which solves the Schrödinger equation for the even solutions in a symmetric potential, what lines of code would you have to change to let it solve for the odd solutions?
- Q10.5.** Assume that you have a quantum mechanical system with quantized energies E_i and probability densities $|\psi_i(x)|^2$. Suppose that you can add a δ -function perturbation at an arbitrary location. How could you then “map out” the wavefunction using the observed shifts in energy. Assume that the δ -function strengths are small enough that first-order perturbation theory can be used. This approach was followed by Salis *et al.* (1997).
- P10.1. Numerical integration—Newton’s law.** Pick some simple technique designed to numerically integrate second order differential equations, perhaps even the simple one used in Section 10.1. Write a short program (using a computer language, programmable calculator, or even a spread sheet program) to solve Newton’s laws for a general potential or force law.

(a) Apply it to the differential equation

$$\frac{d^2x(t)}{dt^2} = -x(t) \quad \text{where } x(0) = 1 \quad \text{and} \quad \dot{x}(0) = 0 \quad (10.163)$$

Compare your results for decreasing step size with the exact solution (which is, of course, $x(t) = \cos(t)$.) Try to reproduce Fig. 10.1.

(b) Try the same thing for the equation

$$\frac{d^2x(t)}{dt^2} = +x(t) \quad \text{where } x(0) = 1 \quad \text{and} \quad \dot{x}(0) = -1 \quad (10.164)$$

and also for the initial conditions $x(0) = +1$ and $\dot{x}(0) = +1$. What are the exact solutions, how well does your program work in these two cases, and why?

- P10.2. Numerical integration—The Schrödinger equation.** Using your experience from P10.1, modify your program to solve the Schrödinger equation for a symmetric potential.

- (a) Apply it to the case of the harmonic oscillator written in dimensionless coordinates as

$$\frac{d^2\psi(y)}{dy^2} - y^2\psi(y) = -\epsilon\psi(y) \quad (10.165)$$

where ϵ are dimensionless eigenvalues. Try to reproduce the values in Example 10.2. Repeat for the odd case where the eigenvalues are $\epsilon = 3, 7, 11, \dots$

- (b) Apply your program to the case of a quartic potential, that is, $V(x) = Cx^4$. Write the Schrödinger equation in dimensionless variables and find the first two even and odd energy eigenvalues. Use your previous experience with the oscillator case to estimate the errors in your calculation.
- P10.3.** Show that the trial wavefunction in Example 10.4 yields the energy function in Eqn. (10.24). Try the problem with the nonzero piece of the wavefunction given by $N(a^2 - x^2)^n$ with $n = 1, 3, 4$ as well and compare your results.
- P10.4.** Estimate the ground state energy of the SHO by using the family of trial wavefunctions

$$\psi(x; a) = \sqrt{\frac{1}{a}} e^{-|x|/a} \quad (10.166)$$

Why is your answer so much worse than that using the cut-off polynomial expression of Example 10.4?

- P10.5.** The momentum-space wavefunction corresponding to P10.4 is

$$\phi(p) = \sqrt{\frac{2p_0}{\pi}} \left(\frac{p_0}{p^2 + p_0^2} \right) \quad (10.167)$$

where $p_0 \equiv \hbar/a$. Evaluate the energy functional in momentum space using this trial wavefunction for the SHO and show that you get the same result (for the energy and trial parameter) as in position space.

- P10.6.** Estimate the energy of the first excited state of the SHO potential by using a trial wavefunction of the form

$$\psi(x; a) = \begin{cases} 0 & \text{for } |x| > a \\ Nx(a^2 - x^2)^2 & \text{for } |x| < a \end{cases} \quad (10.168)$$

Is your answer guaranteed to be larger than the real answer?

- P10.7.** Use a Gaussian trial wavefunction to estimate the ground state energy for the quartic potential, $V(x) = gx^4$. Show that your answer is

$$E_{\min} = \left(\frac{3}{4}\right)^{4/3} \left(\frac{\hbar^4 g}{m}\right)^{1/3} \quad (10.169)$$

Compare this to the “exact” answer (determined by numerical integration) which has the prefactor 0.668.

- P10.8.** (a) Estimate the ground state energy of the symmetric infinite well using the family of trial wavefunctions

$$\psi(x) = \begin{cases} 0 & \text{for } |x| > a \\ N(a^\lambda - |x|^\lambda) & \text{for } |x| < a \end{cases} \quad (10.170)$$

where λ is the variational parameter and you must determine the normalization constant N

- (b) Estimate the energy of the first excited state by using the wavefunction in (a) multiplied by x to make an appropriate odd trial wavefunction. You will, of course, have to renormalize the wavefunction,

- P10.9.** (a) Estimate the ground state energy of the symmetric infinite well by using the wavefunction

$$\psi(x) = \begin{cases} 0 & \text{for } |x| > a \\ N(a^2 - x^2) & \text{for } |x| < a \end{cases} \quad (10.171)$$

Evaluate $E(\text{var})/E_0 - 1$ and $1 - |a_0|^2$ for this state. Note that this has no variational parameter.

- (b) Now consider the family of trial functions

$$\psi(x; b) = \begin{cases} 0 & \text{for } |x| > a \\ N'(a^2 - x^2)(1 + bx^2/L^2) & \text{for } |x| < a \end{cases} \quad (10.172)$$

which does have an additional parameter. Calculate both the variational energy $E(b) = E[\psi(x; b)]$ and $1 - |a_0|^2$. Find the values of b , which minimize each of these two quantities and show that they are slightly different. Specifically, show that

$$b_{\min} = \left(\frac{504 - 51\pi^2}{7\pi^2 - 72} \right) \approx -0.223216 \quad (10.173)$$

for the overlap maximum, while

$$b_{\min} = \frac{(-98 + 8\sqrt{133})}{26} \approx -0.22075 \quad (10.174)$$

for the energy minimum. This demonstrates that while there is a strong correlation between the wavefunction which minimizes the variational energy and the one which maximizes the overlap with the ground state wavefunction, the two criteria are ultimately independent.

- P10.10.** Using the variational method, show that any purely attractive potential in one dimension has at least one bound state. By purely attractive, we mean that $V(x) \leq 0$ for all x . We also assume that $V(x) \rightarrow 0$ as $x \rightarrow \pm\infty$. Hint: Show that we can find a (perhaps very shallow and narrow) finite square well

potential, $V_0(x)$, which satisfies $0 > V_0(x) > V(x)$ for all x and use the fact that a finite square well always has at least one bound state.

- P10.11.** Show that the matching of solutions leading to the WKB quantization condition implies that $C_L = C_R(-1)^n$.
- P10.12.** Apply the WKB quantization condition to the symmetric linear potential, $V(x) = F|x|$. The “exact” answers for the lowest lying even (+) and odd (−) states are given by

$$E_i^{(\pm)} = y_i^{(\pm)} \left(\frac{\hbar^2 F^2}{2m} \right)^{1/3} \quad (10.175)$$

where

$$\begin{aligned} y_1^{(+)} &= 1.0188 & y_1^{(-)} &= 2.3381 \\ y_2^{(+)} &= 3.2482 & y_2^{(-)} &= 4.0879 \\ y_3^{(+)} &= 4.8201 & y_3^{(-)} &= 5.5206 \\ y_4^{(+)} &= 6.1633 & y_4^{(-)} &= 6.7867 \end{aligned} \quad (10.176)$$

Does the agreement get better with increasing n as expected? Can you plot the WKB estimates and “exact” answers in such a way as to demonstrate that?

- P10.13.** Apply the WKB quantization condition to the “half-harmonic oscillator” potential, namely,

$$V(x) = \begin{cases} +\infty & \text{for } x < 0 \\ m\omega^2 x^2/2 & \text{for } x > 0 \end{cases} \quad (10.177)$$

What are the appropriate values of C_L , C_R and what are the WKB energies? What are the exact results for this problem? Hint: Recall P9.10.

- P10.14.** Apply the WKB quantization condition to estimate the bound state energies of the potential

$$V(x) = -\frac{V_0}{\cosh(x/a)^2} \quad (10.178)$$

- (a) Show that your results can be written in the form

$$E_n = -\left(\sqrt{V_0} - (n + 1/2)\sqrt{\frac{\hbar^2}{2ma^2}} \right)^2 \quad (10.179)$$

Hint: You might use the integral

$$\int_0^A \frac{\sqrt{A^2 - u^2}}{1 + u^2} du = \frac{\pi}{2} \left(\sqrt{1 + A^2} - 1 \right) \quad (10.180)$$

- (b) If $V_0 \gg \hbar^2/2ma^2$, show that your result approximates the harmonic oscillator approximation for this potential.
- (c) One might think that one could take the limit $V_0 \rightarrow \infty$ and $a \rightarrow 0$ in such a way as to reproduce an attractive δ -function potential. Discuss the WKB approximation to the energy levels in this limit; if it works, does it reproduce the result of Section 8.1.2? If it does not, why?

P10.15. Harmonic oscillator matrix elements.

- (a) Evaluate \mathbf{p}_{nm} for the harmonic oscillator using the methods in Example 10.7. Using your result, show that the commutator $[x, p] = i\hbar$ holds as a matrix equation.
- (b) Evaluate \mathbf{p}_{nm}^2 and use your result to show that Eqn. (10.79) holds by evaluating both sides as matrices.

P10.16. Infinite well matrix elements.

- (a) Evaluate the matrix elements \mathbf{p}_{nm} using the ‘standard’ infinite well energy eigenstates as a basis. How would you show that

$$\mathbf{H}_{nm} = \frac{1}{2m} \sum_k \mathbf{p}_{nk} \mathbf{p}_{km} \quad (10.181)$$

is the (diagonal) Hamiltonian matrix.

- (b) Evaluate the matrix elements \mathbf{x}_{nm} . Can you show that $[\mathbf{x}, \mathbf{p}] = i\hbar$ holds as a matrix equation?

P10.17. What is the expectation value, $\langle \hat{x} \rangle_t$, for a state vector in a matrix representation for general t , that is, how does Eqn. (10.80) generalize to $t \neq 0$? What does your expression look like for a state with only two components?

P10.18. Show that the wavefunction to second order in perturbation theory (assuming no degeneracies) is given by

$$\begin{aligned} \psi_n^{(2)} = & \sum_m' \sum_k' \frac{\mathbf{H}_{mk}' \mathbf{H}_{kn}'}{(E_n^{(0)} - E_k^{(0)})(E_n^{(0)} - E_m^{(0)})} \psi_m^{(0)} - \sum_m' \frac{\mathbf{H}_{mn}' \mathbf{H}_{mn}}{(E_n^{(0)} - E_m^{(0)})^2} \psi_m^{(0)} \\ & - \frac{1}{2} \psi_n^{(0)} \sum_m' \frac{|\mathbf{H}_{mn}'|^2}{(E_n^{(0)} - E_m^{(0)})^2} \end{aligned} \quad (10.182)$$

P10.19. We have seen in P6.4 that a constant shift in the potential energy function, that is, $V(x) \rightarrow V(x) + V_0$ can have no effect on the observable physics. Consider such a shift as a perturbation and evaluate (i) the first-, second-, and third-order changes in the energy of any state using Eqns. (10.112), (10.115), and (10.117) and (ii) the first-order shift in the wavefunction using Eqn. (10.114) and discuss your results. You can also use the results of P10.18 to check the second-order shift in the wavefunction.

P10.20. Relativistic effects in perturbation theory. The nonrelativistic series for the kinetic energy in Eqn. (1.8) is given by

$$T = \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \cdots \quad (10.183)$$

Using first-order perturbation theory, value the effect of the second term in this expansion (with p replaced by the operator \hat{p}) on the n th level of a harmonic oscillator.

P10.21. Referring to Example 10.9, use Eqn. (10.117) to show that the third-order energy shift to the energy levels vanishes, as expected since the exact result is of second order.

P10.22. Referring to Example 10.10, evaluate the first-, second-, and third-order shifts in energy for the odd states due to the $\delta(x)$ perturbation and show that they vanish.

P10.23. Derive the expansion in Eqn. (10.127) for the ground state solution in Example 10.10 by writing

$$y = \frac{\pi}{2} + a\lambda + b\lambda^2 + \cdots \quad (10.184)$$

substituting this into the exact eigenvalue, $\lambda = -2y \cot(y)$ and equating powers of λ .

P10.24. A particle of mass m in a harmonic oscillator potential $V(x) = m\omega^2 x^2/2$ is subject to a small perturbing potential of the same type, namely, $V'(x) = \lambda x^2$.

(a) Show that the energy spectrum can be derived exactly with the result

$$E'_n = \left(n + \frac{1}{2}\right) \hbar \bar{\omega} \quad (10.185)$$

where $\bar{\omega} = \omega \sqrt{1 + 2\lambda/m\omega^2}$. Expand this for small λ to $\mathcal{O}(\lambda^2)$ for comparison with part (b).

(b) Evaluate the first- and second-order shifts in energy using Eqns (10.112) and (10.115) and compare your results to the exact answer in part (a). You will find the matrix elements of $\langle n|x^2|k\rangle$ in Example 10.7 useful.

P10.25. Anharmonic oscillator in perturbation theory. Evaluate the effect of a small anharmonic term of the form

$$V'(x) = -\lambda kx^3 \quad (10.186)$$

on the spectrum of the harmonic oscillator in first- and second-order perturbation theory. You may find the following matrix element useful:

$$\begin{aligned} \langle \psi_n | x^3 | \psi_k \rangle = & \left(\frac{\hbar}{2m\omega} \right)^{3/2} \left(\sqrt{(n+1)(n+2)(n+3)} \delta_{k,n+3} + 3(n+1)^{3/2} \delta_{k,n+1} \right. \\ & \left. + 3n^{3/2} \delta_{k,n-1} + \sqrt{n(n-1)(n-2)} \delta_{k,n-3} \right) \end{aligned} \quad (10.187)$$

P10.26. Degenerate states in perturbation theory

- (a) Show that the first-order shifts in energy in Eqn. (10.136) are real as they should be. Do the $a_{n,l}$ have to be real?
- (b) Using Eqn. (10.137), show that the linear combinations $\psi^{(+)}(x)$ and $\psi^{(-)}(x)$ are always orthogonal.
- (c) Discuss the energy levels and mixing of eigenstates in the case where $H'_{nm} = H'_{nl} = 0$ but $H'_{ln} = (H'_{nl})^* \neq 0$; show that the eigenfunctions are “completely mixed.”
- (d) Discuss the case where the degenerate states are not connected in the Hamiltonian to lowest order, that is for which $H'_{ln} = (H'_{nl})^* = 0$.

P10.27. Consider a particle of mass m and charge q in the ground state of the symmetric infinite well of Section 5.2.3. It is subject to a time-dependent electric field of the form

$$\hat{H}'(t) = (-q\mathcal{E}_0 x)e^{-|t|/\tau} \quad (10.188)$$

Find the probability that the particle will be excited to the first excited state if the perturbation is allowed to act over the time range $(-\infty, +\infty)$. Repeat for the probability that it is excited to the second-excited state.

P10.28. Consider a system in a general eigenstate ψ_k , which is subject to a time-dependent harmonic perturbation of the form

$$\hat{H}'(t) = 2V(x) \cos(\omega t) \quad (10.189)$$

If this perturbation is turned on at $t = 0$ and then removed at $t > 0$, find the probability that the system is in a new state, ψ_m . In the limit of long times, what states are most likely to be connected by this perturbation? Discuss what this might have to do with the emission or absorption of radiation.

TWO

Multiparticle Systems

Just as in classical mechanics, the study of single particle systems in one dimension provides invaluable experience in the use of quantum mechanical concepts and techniques. To extend these ideas to more realistic applications (to be able to enter the Quantum World of Part II) we also need to develop the formalisms necessary to handle multiparticle systems. In this chapter we will discuss separable systems (in Section 14.2), the important special case of two-body systems (Section 14.3), notation for spin-1/2 wavefunctions (Section 14.4), and lastly, in Section 14.5, the important constraints placed on quantum wavefunctions of multiparticle systems due to indistinguishability.

14.1 Generalities

In classical mechanics, Newton's laws for a multiparticle system have the form

$$m_i \frac{d^2 x_i(t)}{dt^2} = \mathcal{F}_i(x_i) + \sum_{j \neq i} F_{ij}(x_i - x_j) \quad \text{for } i = 1, 2, \dots, N \quad (14.1)$$

where we have specialized to the case of external forces, $\mathcal{F}_i(x_i)$, which act on each particle separately and mutual two-body interactions, $F_{ij} = F_{ij}(x_i - x_j)$; the functional form of the F_{ij} is consistent with Newton's third law which requires that $F_{ij} = -F_{ji}$. When solved self-consistently, these equations predict the time-dependence of the coordinates of all of the particles, $x_i(t)$, once the initial conditions are specified.

The time-development of the corresponding quantum system is dictated by the *multiparticle Hamiltonian* operator

$$\hat{H} = \sum_i \frac{\hat{p}_i^2}{2m_i} + \sum_i \mathcal{V}_i(x_i) + \sum_{i>j} V_{ij}(x_i - x_j) \quad (14.2)$$

Here the external and two-body potentials give the corresponding forces via

$$\mathcal{F}_i(x_i) = -\frac{\partial \mathcal{V}_i(x_i)}{\partial x_i} \quad \text{and} \quad F_{ij}(x_i - x_j) = -\frac{\partial V_{ij}(x_i - x_j)}{\partial x_i} \quad (14.3)$$

and the restriction $i > j$ on the double sum is to avoid double counting. The momentum operator corresponding to each coordinate is $\hat{p}_i = (\hbar/i)\partial/\partial x_i$ and one has $[x_j, \hat{p}_k] = i\hbar\delta_{jk}$.

This acts on a multiparticle wavefunction $\psi(x_1, x_2, \dots, x_n; t)$ as the time-development operator and generalizes the Schrödinger equation to

$$\hat{H}\psi(x_1, x_2, \dots, x_N; t) = i\hbar \frac{\partial}{\partial t} \psi(x_1, x_2, \dots, x_N; t) \quad (14.4)$$

If none of the potentials actually depend on time, the usual exponential time-dependence is found so that

$$\psi(x_1, x_2, \dots, x_N; t) = \psi_E(x_1, x_2, \dots, x_N) e^{-iEt/\hbar} \quad (14.5)$$

where ψ_E satisfies the *time-independent multiparticle Schrödinger equation*

$$\hat{H}\psi_E(x_1, x_2, \dots, x_N) = E\psi_E(x_1, x_2, \dots, x_N) \quad (14.6)$$

The multiparticle wavefunction is then associated with a probability amplitude, so that

$$P(x_1, x_2, \dots, x_N; t) = |\psi(x_1, x_2, \dots, x_N; t)|^2 \quad (14.7)$$

is a *multivariable probability density*. A more concrete definition is:

- The quantity $|\psi(x_1, x_2, \dots, x_n; t)|^2 dx_1 dx_2, \dots, dx_n$ is the probability that a measurement of the positions of the N particles, at time t , would find

particle 1 in the interval $(x_1, x_1 + dx_1)$

and

particle 2 in the interval $(x_2, x_2 + dx_2)$

and

\vdots

and

particle N in the interval $(x_N, x_N + dx_N)$

Just as with any multivariable probability distribution, this implies that the wavefunction must be normalized so that

$$\int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 \cdots \int_{-\infty}^{+\infty} dx_n |\psi(x_1, x_2, \dots, x_n; t)|^2 = 1 \quad (14.9)$$

and the total probability of measuring “something” is unity and not, for example, N since we have more than one particle; we emphasize that $|\psi|^2$ does not “count the number of particles,” but rather specifies the probability of the entire system of particles being in a particular configuration.

All of the usual conditions on the smoothness and convergence of one-dimensional wavefunctions can be easily generalized, as can the expressions for expectation values; for example, one has

$$\begin{aligned} \langle \hat{O} \rangle_t &= \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 \cdots \int_{-\infty}^{+\infty} dx_n \\ &\times \psi^*(x_1, x_2, \dots, x_N; t) \hat{O} \psi(x_1, x_2, \dots, x_N; t) \end{aligned} \quad (14.10)$$

for the average value of any operator \hat{O} .

Example 14.1. Correlations in two-particle wavefunctions

Consider a two-particle wavefunction

$$\psi(x_1, x_2) = N e^{-(ax_1^2 + 2bx_1x_2 + cx_2^2)} \quad (14.11)$$

where we must have $a, c > 0$ in order for the wavefunction to be normalizable. The normalization constant is determined by the condition that

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 |\psi(x_1, x_2)|^2 \\ &= N^2 \int_{-\infty}^{+\infty} dx_2 e^{-(c-b^2/a)x_2^2} \int_{-\infty}^{+\infty} dx_1 e^{-a(x_1 + bx_2/a)^2} \\ &= N^2 \frac{\pi}{\sqrt{ac - b^2}} \end{aligned} \quad (14.12)$$

This form (obtained by completing the square in the exponent) is useful in that it shows that one must also have $ac - b^2 > 0$ in order for the state to be acceptable. It is then easy to show that $\langle x_1 \rangle = \langle x_2 \rangle = 0$, while the *covariance* (measuring a correlation between the two particles) is given by

$$\langle (x_1 - \langle x_1 \rangle)(x_2 - \langle x_2 \rangle) \rangle = \langle x_1 x_2 \rangle = -\frac{b}{(ac - b^2)} \quad (14.13)$$

This shows that the positions of the coordinates are correlated with each other, so that a measurement of one provides nontrivial information on the other. When $b \rightarrow 0$, we note that the wavefunction becomes the (uncorrelated) product of two Gaussians, and is simply a product wavefunction.

14.2 Separable Systems

A great simplification occurs if the mutual interactions of the particles can be ignored ($V_{ij} = 0$), because the Hamiltonian then takes the form

$$\hat{H} = \sum_i \frac{\hat{p}_i^2}{2m_i} + \sum_i \mathcal{V}(x_i) = \sum_i \left(\frac{\hat{p}_i^2}{2m_i} + \mathcal{V}(x_i) \right) \equiv \sum_i \hat{H}_i \quad (14.14)$$

With no mutual interactions present to give rise to dynamical correlations, it is natural to assume a product solution of the form

$$\psi_E(x_1, x_2, \dots, x_N) = \psi_1(x_1) \psi_2(x_2) \cdots \psi_N(x_N). \quad (14.15)$$

The time-independent Schrödinger equation, Eqn. (14.6), can then be written as

$$\begin{aligned} E [\psi_1(x_1) \psi_2(x_2) \cdots \psi_N(x_N)] &= [\hat{H}_1 \psi_1(x_1)] \psi_2(x_2) \cdots \psi_N(x_N) \\ &+ \psi_1(x_1) [\hat{H}_2 \psi_2(x_2)] \cdots \psi_N(x_N) \quad (14.16) \\ &+ \cdots + \psi_1(x_1) \psi_2(x_2) \cdots [\hat{H}_N \psi_N(x_N)] \end{aligned}$$

Using the usual separation of variables trick, we divide both sides by Eqn. (14.15) and find that

$$\frac{[\hat{H}_1 \psi_1(x_1)]}{\psi_1(x_1)} + \frac{[\hat{H}_2 \psi_2(x_2)]}{\psi_2(x_2)} + \cdots + \frac{[\hat{H}_N \psi_N(x_N)]}{\psi_N(x_N)} = E. \quad (14.17)$$

This is only consistent if

$$\hat{H}_i \psi_i(x_i) = E_i \psi_i(x_i) \quad \text{for } i = 1, 2, \dots, N \quad (14.18)$$

where $E_1 + E_2 + \cdots + E_N = E$; we then have to solve N “versions” of the one-dimensional problem. Several comments can be made:

- If each component wavefunction is properly normalized, then the product solution is also, since

$$\int_{-\infty}^{+\infty} dx_1 \cdots \int_{-\infty}^{+\infty} dx_N |\psi(x_1, x_2, \dots, x_N)|^2 = \prod_i \left[\int_{-\infty}^{+\infty} dx_i |\psi_i(x_i)|^2 \right] = 1 \quad (14.19)$$

- The overall time-dependence can also be factorized since

$$e^{-iEt/\hbar} = e^{-iE_1 t/\hbar} \cdots e^{-iE_n t/\hbar} \quad (14.20)$$

which implies that

$$\psi(x_1, x_2, \dots, x_N; t) = \psi_1(x_1, t) \psi_2(x_2, t) \cdots \psi_N(x_N, t) \quad (14.21)$$

This is potentially useful as wave packets for each particle can be constructed using superposition techniques, so that products of such wave packets will also be valid solutions for the noninteracting case.

Example 14.2. Degeneracy in two-particle systems

Consider two particles of the same mass m confined to the standard infinite well; for the moment, we neglect any mutual interactions.¹ The general solution to this two-particle system is

$$\psi_{(n_1, n_2)}(x_1, x_2) = u_{(n_1)}(x_1) u_{(n_2)}(x_2) \quad \text{where } u_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) \quad (14.22)$$

with the corresponding energy spectrum

$$E_{(n_1, n_2)} = \frac{\hbar^2 \pi^2}{2ma^2} (n_1^2 + n_2^2) \quad (14.23)$$

The ground state energy is $\hbar^2 \pi^2 / ma^2$, corresponding to $(n_1, n_2) = (1, 1)$ and is unique. The first excited state, given by the two choices $(1, 2)$ and $(2, 1)$, is doubly degenerate, and the corresponding wavefunctions can be written as $\psi_\alpha = \psi_{(1,2)}(x_1, x_2)$ and $\psi_\beta = \psi_{(2,1)}(x_1, x_2)$; these two choices are not unique because we can invoke the linearity of the Schrödinger equation to show that any (appropriately orthogonal) linear combination of these two is also a solution with energy $E_{(1,2)} = 5\hbar^2 \pi^2 / 2ma^2 = E_{(2,1)}$.

We can now use this example to illustrate the methods of degenerate perturbation theory, as outlined in Section 10.5.2. We add a small mutual interaction term given by $V'(x_1, x_2) = g\delta(x_1 - x_2)$ where positive (negative) g corresponds to a repulsive (attractive) interaction between the two particles. Referring to Eqn. (10.134), we require the various matrix elements of the perturbing interaction; for example,

$$\begin{aligned} H'_{\alpha\alpha} &= \langle \psi_\alpha | V(x_1 - x_2) | \psi_\alpha \rangle \\ &= \int_0^a dx_1 \int_0^a dx_2 [u_1(x_1) u_2(x_2)] (g\delta(x_1 - x_2)) [u_1(x_1) u_2(x_2)] \\ &= g \left(\frac{2}{a}\right)^2 \int_0^a dx_1 \sin^2\left(\frac{\pi x_1}{a}\right) \sin^2\left(\frac{2\pi x_1}{a}\right) \\ H'_{\alpha\alpha} &= \frac{g}{a} \end{aligned} \quad (14.24)$$

¹ If they are both in the same one-dimensional well, this implies that they are somewhat “ghostlike” as they must be able to “pass through” each other.

(Continued)

with identical answers for $H'_{\beta\beta}$ and $H'_{\alpha\beta}$. The condition determining the (split) energy eigenvalues (Eqn. (10.134)) then reads

$$\det \begin{pmatrix} \mathcal{E} + g/a - E & g/a \\ g/a & \mathcal{E} + g/a - E \end{pmatrix} = 0 \quad (14.25)$$

where $\mathcal{E} = E_{(2,1)} = E_{(1,2)}$ is the initially degenerate energy level. The resulting polynomial equation is easily solved and yields

$$E^{(\pm)} = \mathcal{E} + \frac{g}{a} \pm \frac{g}{a} \quad (14.26)$$

The two energy eigenvalues and corresponding (normalized) eigenstates are given by

$$E^{(-)} = \mathcal{E} \quad \psi^{(-)}(x_1, x_2) = \frac{1}{\sqrt{2}} (\psi_{(1,2)}(x_1, x_2) - \psi_{(2,1)}(x_1, x_2)) \quad (14.27)$$

$$E^{(+)} = \mathcal{E} + \frac{2g}{a} \quad \psi^{(+)}(x_1, x_2) = \frac{1}{\sqrt{2}} (\psi_{(1,2)}(x_1, x_2) + \psi_{(2,1)}(x_1, x_2)) \quad (14.28)$$

The antisymmetric combination state, corresponding to $E^{(-)}$, is unshifted in energy because the $\psi^{(-)}$ wavefunction vanishes where the perturbation has any effect, namely, for $x_1 = x_2$. The symmetric solution has a larger probability of having $x_1 = x_2$ than do either ψ_α or ψ_β individually and it can “feel” the effect of the perturbation; the energy of the symmetric state is therefore increased or decreased depending on the sign of g .

14.3 Two-Body Systems

While much of classical mechanics is concerned with the motion of single particles under the influence of external forces, many standard problems, especially in gravitation, are concerned with the motion of two bodies subject only to their mutual interaction. While general methods of solution for the N -body problem² (with $N \geq 3$) do not exist, a simple change of variables is often enough to transform Newton's equations for two particles into an effective one-particle problem which can then be approached using a variety of familiar techniques.

Such techniques are perhaps even more important in quantum mechanics where many of the “textbook” examples are two-body systems; examples include diatomic molecules, the hydrogen atom, the deuteron (proton–neutron bound state), and quarkonia (quark–antiquark bound states). In these cases, we are often more interested in probing the (sometimes unknown) force between the particles, so that not having to deal with the complications of many particles

² See, for example, Symon (1971).

is extremely important. Even though it is implemented in a very different way, the same coordinate transformation “trick” works in both classical and quantum mechanics and we begin our study by reviewing the classical case.

14.3.1 Classical Systems

The classical equations of motion for a two-particle system with no external forces and only mutual two-body interactions are

$$m_1 \ddot{x}_1(t) = F_{21}(x_1 - x_2) \quad \text{and} \quad m_2 \ddot{x}_2(t) = F_{12}(x_1 - x_2) \quad (14.29)$$

and we recall that $F \equiv F_{21} = -F_{12}$ from Newton’s third law. Two combinations of these equations then immediately suggest themselves and naturally select out a new set of variables. If, for example, we add the two equations in Eqn. (14.29), we obtain

$$\begin{aligned} 0 &= F_{21} + F_{12} = m_1 \ddot{x}_1(t) + m_2 \ddot{x}_2(t) \\ &= (m_1 + m_2) \left(\frac{m_1 \ddot{x}_1(t) + m_2 \ddot{x}_2(t)}{m_1 + m_2} \right) \\ 0 &= M \ddot{X}(t) \end{aligned} \quad (14.30)$$

where we define the *total mass*, $M = m_1 + m_2$, and the *center-of-mass coordinate*

$$X(t) = \frac{m_1 x_1(t) + m_2 x_2(t)}{m_1 + m_2} \quad (14.31)$$

We note that Eqn. (14.30) gives the standard result that if there are no net external forces, the center-of-mass of a system moves at constant speed. A related variable is the *total momentum*, given by

$$P(t) = M \dot{X}(t) = m_1 v_1(t) + m_2 v_2(t) = p_1(t) + p_2(t) \quad (14.32)$$

so that Eqn. (14.30) also shows that the total momentum is conserved.

If we now divide both sides of Eqns (14.29) by the respective masses, and then subtract, we find

$$F \left(\frac{1}{m_1} + \frac{1}{m_2} \right) = \frac{F_{21}}{m_1} - \frac{F_{12}}{m_2} = \ddot{x}_1(t) - \ddot{x}_2(t) \quad (14.33)$$

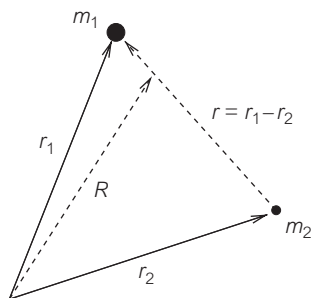
or

$$F(x) = \mu \ddot{x}(t) \quad (14.34)$$

where we have defined the *reduced mass* via

$$\frac{1}{\mu} \equiv \frac{1}{m_1} + \frac{1}{m_2} \quad \text{or} \quad \mu = \frac{m_1 m_2}{m_1 + m_2} \quad (14.35)$$

Figure 14.1. Center-of-mass (\mathbf{R}) and relative ($\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$) coordinates for a two-body system in two-dimensions, which generalizes Eqn. (14.31). You should be able to estimate the ratio m_1/m_2 from the figure.



and the *relative coordinate* via

$$\mathbf{x}(t) = \mathbf{x}_1(t) - \mathbf{x}_2(t) \quad (14.36)$$

while $F(x) \equiv F_{21}(x_1 - x_2)$. The nontrivial dynamics of the system is then described by Eqn. (14.34); the “interesting” physics is all contained in the relative coordinate which describes the motion of a fictitious particle of effective mass μ . The change of variables can also be inverted to give

$$\mathbf{x}_1(t) = \mathbf{X}(t) + \frac{m_2}{M}\mathbf{x}(t) \quad \text{and} \quad \mathbf{x}_2(t) = \mathbf{X}(t) - \frac{m_1}{M}\mathbf{x}(t) \quad (14.37)$$

so that the motion of each particle can be extracted if so desired. The same variable change works in more realistic two- and three-dimensional systems and we visualize the new coordinates in two-dimensions in Fig. 14.1.

14.3.2 Quantum Case

The quantum version of the two-body problem requires us to solve the two-particle Schrödinger equation given by

$$\hat{H}\psi(x_1, x_2) = E\psi(x_1, x_2) \quad (14.38)$$

where the Hamiltonian is given by

$$\begin{aligned} \hat{H} &= \frac{\hat{p}_1^2}{2m_1} + \frac{\hat{p}_2^2}{2m_2} + V(x_1 - x_2) \\ &= -\frac{\hbar^2}{2m_1} \frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m_2} \frac{\partial^2}{\partial x_2^2} + V(x_1 - x_2) \end{aligned} \quad (14.39)$$

if there are no external forces. The change to center-of-mass and relative coordinates,

$$\mathbf{X} = \frac{m_1\mathbf{x}_1 + m_2\mathbf{x}_2}{m_1 + m_2} \quad \text{and} \quad \mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2 \quad (14.40)$$

(where we drop the t dependence, as these are no longer classical coordinates, but quantum mechanical labels) is trivially implemented for the potential energy

term where $V(x_1 - x_2) = V(x)$. For the kinetic energy operators, we need to rewrite

$$\hat{p}_i = \frac{\hbar}{i} \frac{\partial}{\partial x_i} \quad \text{for } i = 1, 2 \quad (14.41)$$

in terms of the spatial derivatives of the X, x coordinates which give the momentum operators corresponding to those variables, namely

$$\hat{P} \equiv \frac{\hbar}{i} \frac{\partial}{\partial X} \quad \text{and} \quad \hat{p} \equiv \frac{\hbar}{i} \frac{\partial}{\partial x}. \quad (14.42)$$

This requires the chain rule relation

$$\frac{\partial}{\partial x_{1,2}} = \frac{\partial X}{\partial x_{1,2}} \frac{\partial}{\partial X} + \frac{\partial x}{\partial x_{1,2}} \frac{\partial}{\partial x} \quad (14.43)$$

which, using Eqn. (14.40), gives

$$\hat{p}_1 = \frac{m_1}{M} \hat{P} + \hat{p} \quad \text{and} \quad \hat{p}_2 = \frac{m_2}{M} \hat{P} - \hat{p} \quad (14.44)$$

We note that adding these two equations gives $\hat{P} = \hat{p}_1 + \hat{p}_2$, now as an operator relation. The kinetic energy operators in the Hamiltonian can now be written as

$$\begin{aligned} \frac{\hat{p}_1^2}{2m_1} + \frac{\hat{p}_2^2}{2m_2} &= \frac{1}{2m_1} \left(\frac{m_1}{M} \hat{P} + \hat{p} \right)^2 + \frac{1}{2m_2} \left(\frac{m_2}{M} \hat{P} - \hat{p} \right)^2 \\ &= \frac{1}{2m_1} \left(\frac{m_1^2}{M^2} \hat{P}^2 + \frac{m_1}{M} (\hat{p}\hat{P} + \hat{P}\hat{p}) + \hat{p}^2 \right) \\ &\quad + \frac{1}{2m_2} \left(\frac{m_2^2}{M^2} \hat{P}^2 - \frac{m_2}{M} (\hat{p}\hat{P} + \hat{P}\hat{p}) + \hat{p}^2 \right) \\ \frac{\hat{p}_1^2}{2m_1} + \frac{\hat{p}_2^2}{2m_2} &= \frac{\hat{P}^2}{2M} + \frac{\hat{p}^2}{2\mu} \end{aligned} \quad (14.45)$$

While we have been careful with the ordering of \hat{P} and \hat{p} , it is easy to see that $[\hat{P}, \hat{p}] = 0$. This can also be used to confirm that the total momentum of the system is a constant since \hat{P} commutes with the Hamiltonian operator, that is,

$$[\hat{H}, \hat{P}] = \frac{1}{2M} [\hat{P}^2, \hat{P}] + \frac{1}{2\mu} [\hat{p}^2, \hat{P}] + [V(x), \hat{P}] = 0 \quad (14.46)$$

We emphasize that this is true only under our assumption that the potential is only a function of the relative coordinate; if there are external forces, an equivalent classical result is obtained (P14.3).

Most importantly, in the new coordinates, the Hamiltonian is now separable since

$$\hat{H} = \frac{\hat{P}^2}{2M} + \left(\frac{\hat{p}^2}{2\mu} + V(x) \right) = \hat{H}_X + \hat{H}_x \quad (14.47)$$

so a product wavefunction of the form

$$\left[\Psi(X) e^{-iE_X t/\hbar} \right] \left[\psi(x) e^{-iE_x t/\hbar} \right] \quad (14.48)$$

satisfying

$$\hat{H}_X \Psi(X) = E_X \Psi(X) \quad \text{and} \quad \hat{H}_x \psi(x) = E_x \psi(x) \quad (14.49)$$

is a solution. The center-of-mass equation has trivial plane wave solutions of the form

$$\Psi_P(X) = e^{i(PX - P^2 t/2M)/\hbar} \quad (14.50)$$

(with P a number) from which wave packets, representing the constant velocity motion of the center-of-mass, can be constructed.

Example 14.3. Reduced mass effects in two-particle systems

A simple model of a one-dimensional diatomic molecule consists of two masses m_1, m_2 interacting via the potential $V(x_1 - x_2) = K(x_1 - x_2 - l)^2/2$ where l is the equilibrium separation of the two masses. The equation for the relative coordinate,

$$\left(\frac{\hat{p}^2}{2\mu} + V(x - l) \right) \psi(x) = E_x \psi(x) \quad (14.51)$$

has the standard harmonic oscillator solutions $\psi(x) = \phi_n(x - l)$ with quantized energy levels given by $E_x^{(n)} = (n + 1/2)\hbar\omega$ where $\omega = \sqrt{K/\mu}$. The dependence on μ of the zero-point energy of vibrational states in diatomic molecules was mentioned in Section 9.3.

In systems where $m_1 \approx m_2$, the reduced mass is roughly $\mu \approx m_1/2 \approx m_2/2$ and its effect is obviously important to include. In cases such as the hydrogen atom, where one has $m_1 = m_e \ll m_p = m_2$ and $\mu = m_e/(1 + m_e/m_p) \approx m_e$, the effect is much smaller (since $m_e/m_p \approx 1/2000$) and is sometimes not stressed sufficiently. The discrete energy spectrum of a hydrogen-like atom (a single electron interacting via a Coulomb force with a nucleus of charge Z) is given by Eqn. (1.42) as

$$E_n = -\frac{1}{2}\mu c^2 Z^2 \alpha^2 \frac{1}{n^2} \quad (14.52)$$

where the reduced mass μ now properly appears. The frequencies of the photons emitted in a transition are

$$\hbar\omega_{nl} = \hbar 2\pi f_{nl} = E_n - E_l = \frac{1}{2}\mu c^2 Z^2 \alpha^2 \left(\frac{1}{n^2} - \frac{1}{l^2} \right) \quad (14.53)$$

The dependence on μ implies that the corresponding lines in atoms with the same value of Z , but with different nuclear masses (i.e. isotopes) will have slightly differing wavelengths. This effect was utilized in the discovery of the “*hydrogen isotope of mass 2*,” now known as deuterium. (See P14.6 for details.)

14.4 Spin Wavefunctions

In describing the quantum state of a particle, we have concentrated on the wavefunctions corresponding to observable quantities such as position ($\psi(x)$), momentum ($\phi(p)$), or energy eigenvalues ($\{a_n; n = 0, 1 \dots\}$). The “spin-up” and “spin-down” label necessary to describe spin-1/2 particles must also be included in the multiparticle wavefunctions for such particles. A more comprehensive discussion of spin in quantum mechanics is given in the next chapter, but we introduce here, for convenience, some of the basic formalism for spin-1/2 particles.

A convenient (matrix) representation of the spin operator (quantized along some convenient direction, often the z -axis) is given by

$$\mathbf{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (14.54)$$

which acts on a (complex) *spinor wavefunction*

$$\chi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (14.55)$$

In order to be normalized, such spinors must satisfy

$$(\chi)^\dagger \chi = (\alpha^*, \beta^*) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = |\alpha|^2 + |\beta|^2 = 1 \quad (14.56)$$

The eigenvectors and eigenvalues of \mathbf{S}_z are seen to be

$$\chi^+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{with} \quad \mathbf{S}_z \chi^+ = +\frac{\hbar}{2} \chi^+ \quad (14.57)$$

$$\chi^- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{with} \quad \mathbf{S}_z \chi^- = -\frac{\hbar}{2} \chi^- \quad (14.58)$$

Since \mathbf{S}_z is a Hermitian (matrix) operator (note that $\mathbf{S}_z^\dagger = \mathbf{S}_z$), it is not surprising that its eigenvalues, $\pm\hbar/2$, are real. For the same reason, the eigenfunctions (in

this case eigenvectors) also satisfy the usual orthonormality conditions since

$$(\chi^{(+)})^\dagger \chi^{(+)} = 1 = (\chi^{(-)})^\dagger \chi^{(-)} \quad \text{and} \quad (\chi^{(+)})^\dagger \chi^{(-)} = 0 = (\chi^{(-)})^\dagger \chi^{(+)} \quad (14.59)$$

and there is a corresponding expansion theorem, written as

$$\chi = \begin{pmatrix} a^{(+)} \\ a^{(-)} \end{pmatrix} = a^{(+)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a^{(-)} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a^{(+)} \chi^{(+)} + a^{(-)} \chi^{(-)} \quad (14.60)$$

This form makes it clear that $|a^{(+)}|^2 (|a^{(+)}|^2)$ is the probability that a measurement of the spin (projected onto the z -axis) will yield a value of $+\hbar/2$ ($-\hbar/2$); it is also consistent with the expectation value

$$\langle \chi | S_z | \chi \rangle = (a^{(+)*}, a^{(-)*}) \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a^{(+)} \\ a^{(-)} \end{pmatrix} = \frac{\hbar}{2} (|a^{(+)}|^2 - |a^{(-)}|^2) \quad (14.61)$$

A spin-1/2 particle can then carry information on its spin state in its quantum wavefunction, $\psi(x, \chi)$, and inner products between different quantum states must be generalized. For example, the overlap “integral” of $\psi_a(x) \chi_a$ and $\psi_b(x) \chi_b$ will be

$$\langle \psi_a | \psi_b \rangle = \left[\int_{-\infty}^{+\infty} dx \psi_a^*(x) \psi_b(x) \right] [(\chi_a)^\dagger \chi_b] \quad (14.62)$$

and wavefunctions can be orthogonal because of different spin dependences. An expansion in energy and spin eigenstates might then have the form

$$\psi(x, \chi) = \sum_n \left(a_n^{(+)} u_n(x) \chi^{(+)} + a_n^{(-)} u_n(x) \chi^{(-)} \right) \quad (14.63)$$

Example 14.4. Expansion in energy and spin eigenstates

Consider a spin-1/2 particle in a harmonic oscillator potential described by the wavefunction

$$\psi(x, \chi) = N \left(3\psi_0(x) \chi^{(+)} - (2+i)\psi_1(x) \chi^{(-)} + \sqrt{6}\psi_1(x) \chi^{(+)} \right) \quad (14.64)$$

The normalization constant can be determined by the requirement that

$$\sum_n (|a_n^{(+)}|^2 + |a_n^{(-)}|^2) = 1 \quad \text{so that } N = 1/\sqrt{20} \quad (14.65)$$

The average value of the energy is

$$\langle \hat{E} \rangle = \left(\frac{9}{20} \right) \left(\frac{1}{2} \hbar \omega \right) + \left(\frac{5+6}{20} \right) \left(\frac{3}{2} \hbar \omega \right) = \frac{21}{20} \hbar \omega \quad (14.66)$$

(Continued)

while the expectation value of S_z is

$$\langle S_z \rangle = \frac{\hbar}{2} \left(\frac{9+6}{20} \right) - \frac{\hbar}{2} \left(\frac{5}{20} \right) = \frac{\hbar}{4} \quad (14.67)$$

The combined probability that a measurement will find $S_z = +\hbar/2$ and $E = 3\hbar\omega/2$ is $P = 3/10$.

For multiparticle wavefunctions, we have to specify the position and spin label for each particle,³ for example, $\psi(x_1, \chi_1; x_2, \chi_2; \dots; x_n, \chi_n)$. We will often denote all of the relevant labels for a given particle by simply specifying the common numerical index, that is $\psi(1; 2; \dots; N)$.

For example, for two (noninteracting) electrons in an infinite well, the following wavefunctions will all turn out to be physically acceptable:

$$\psi_A(1; 2) = \frac{1}{\sqrt{2}} \left(\chi_1^{(-)} \chi_2^{(+)} - \chi_1^{(+)} \chi_2^{(-)} \right) u_1(x_1) u_1(x_2) \quad (14.68)$$

$$\psi_B(1; 2) = \chi_1^{(+)} \chi_2^{(+)} \frac{1}{\sqrt{2}} (u_1(x_1) u_2(x_2) - u_2(x_1) u_1(x_2)) \quad (14.69)$$

$$\psi_C(1; 2) = \frac{1}{\sqrt{2}} \left(u_1(x_1) u_2(x_2) \chi_1^{(+)} \chi_2^{(-)} - u_2(x_1) u_1(x_2) \chi_1^{(-)} \chi_2^{(+)} \right) \quad (14.70)$$

The inner product for spin-states for different particles is generalized to be

$$\langle \chi_1, \chi_2 | \chi_1, \chi_2 \rangle = \langle \chi_1 | \chi_1 \rangle \langle \chi_2 | \chi_2 \rangle = (|\alpha_1|^2 + |\beta_1|^2)(|\alpha_2|^2 + |\beta_2|^2) = 1 \quad (14.71)$$

You should now be able to show that the wavefunctions in Eqns (14.68)–(14.70) are properly normalized and mutually orthogonal and be able to calculate the energy of each state.

14.5 Indistinguishable Particles

Thus far, we have focused for the most part on exploring the consequences of a wave description of particles, its implications for observable phenomena, and the connection between the classical and quantum limit; we have thus concentrated on what we have termed “ \hbar physics.” But there is another dichotomy between the extreme quantum and classical limits of particles which was introduced

³ We do not consider other possible degrees of freedom which might be labeled, such as “isospin” for nucleons and “color” for quarks.

in Chapter 7 which we can describe as “indistinguishability physics”; hereafter, indistinguishable will be abbreviated IND for convenience.

Unlike a set of billiard balls which have different colors and even distinct numeric labels, each of the electrons in an atom is seemingly equivalent to every other electron. In the same way, all protons are effectively identical, all neutrons are equivalent, and so forth; no experiment has yet been able to discern any measureable differences between individual electrons, individual protons, and so on. The same statement holds for other particles other than just the “building blocks of nature,” such as photons and all other “elementary” particles.

We can roughly define:

- A set of indistinguishable (IND) particles is one in which the interchange of any two particles has no observable effect on any property of the system.

The notion of indistinguishability raises an interesting question when one considers the total number of wavefunctions which can have the same total energy, that is, the degeneracy. For simplicity, say we have a product wavefunction of N IND particles of the form

$$\psi(1; 2; \dots; N) = \phi_a(1)\phi_a(2) \cdots \phi_a(N) \quad (14.72)$$

where each particle is in the same quantum state; the total energy is simply $E_{\text{tot}} = NE_a$. The assumption of indistinguishability means any permutation of the indices will give a state with the same energy. Because of the special form of Eqn. (14.72), however, the resulting exchanges do not change the wavefunction and there is only one distinct wavefunction with this energy.

Contrast to this is the situation where all N particles are in totally different one-particle configurations,

$$\psi(1; 2; \dots; N) = \phi_{a_1}(1)\phi_{a_2}(2) \cdots \phi_{a_N}(N) \quad \text{with } a_1 \neq a_2 \neq \cdots \neq a_N \quad (14.73)$$

with energy $E_{\text{tot}} = E_{a_1} + E_{a_2} + \cdots + E_{a_N}$. One can use N different labels for the first state ϕ_{a_1} , leaving $N - 1$ for the second state ϕ_{a_2} , and so forth; there are thus $N!$ different permutations of the labels, each of which gives a state of the same total energy and, in this case, all $N!$ wavefunctions are distinct. For example, with three particles in the $n = 0, 1, 2$ levels of the harmonic oscillator, we might have (ignoring spin labels) the $3! = 6$ different states

$$\begin{aligned} &\phi_0(x_1)\phi_1(x_2)\phi_2(x_3), \quad \phi_0(x_1)\phi_1(x_3)\phi_2(x_2), \quad \phi_0(x_2)\phi_1(x_1)\phi_2(x_3) \\ &\phi_0(x_2)\phi_1(x_3)\phi_2(x_1), \quad \phi_0(x_3)\phi_1(x_1)\phi_2(x_2), \quad \phi_0(x_3)\phi_1(x_2)\phi_2(x_1) \end{aligned} \quad (14.74)$$

For *distinguishable* particles, each of these choices corresponds to a different physical system since one can, by definition, tell the particles apart; the probability density for a hydrogen atom with an electron “here” and a proton “there” is obviously different from the exchanged system. For IND particles, however, we have the possibility of $N!$ wavefunctions which supposedly all describe the same (presumably unique) physical system and the obvious question is:

- Which one (if any) of these $N!$ choices is the appropriate wavefunction?

To help answer this question, we first formalize the notion of interchange by defining the *exchange operator*, $\hat{\mathcal{E}}_{ij}$, which has the effect of exchanging particles i and j with *all* their appropriate labels. Recalling the notation

$$\psi(x_1, \chi_1; x_2, \chi_2; \dots; x_N, \chi_N) \equiv \psi(1; 2; \dots; N) \quad (14.75)$$

we define the exchange operator such that

$$\hat{\mathcal{E}}_{ij}\psi(1; 2; \dots; i; \dots; j; \dots; N) = \psi(1; 2; \dots; j; \dots; i; \dots; N) \quad (14.76)$$

so that $x_i \leftrightarrow x_j$ and $\chi_i \leftrightarrow \chi_j$.

If, for example, one has two IND spin-1/2 particles in the infinite well with wavefunction

$$\psi(1; 2) = u_4(x_1)u_7(x_2)\chi_1^{(+)}\chi_2^{(-)} \quad (14.77)$$

we will have

$$\psi(2; 1) = \hat{\mathcal{E}}_{12}\psi(1; 2) = u_4(x_2)u_7(x_1)\chi_1^{(-)}\chi_2^{(+)} \quad (14.78)$$

The two-particle wavefunctions in Eqns (14.68)–(14.70) are easily seen to be antisymmetric under the action of $\hat{\mathcal{E}}_{12}$.

There are $N(N-1)$ distinct $\hat{\mathcal{E}}_{ij}$ which, by themselves, exchange labels pairwise; the complete set of all the $N!$ permutations of the particle indices is generated by taking products of the individual $\hat{\mathcal{E}}_{ij}$. For example, when $N = 3$, we have

$$\psi(3; 2; 1) = \hat{\mathcal{E}}_{13}\psi(1; 2; 3) \quad (14.79)$$

while

$$\psi(2; 3; 1) = \hat{\mathcal{E}}_{12}\psi(1; 3; 2) = \hat{\mathcal{E}}_{12}\hat{\mathcal{E}}_{23}\psi(1; 2; 3) \quad (14.80)$$

The set of the $\hat{\mathcal{E}}_{ij}$ and their products forms a *permutation group* (P14.7).

We can establish several important properties of the exchange operators using two-particle systems as an example, for ease of notation:

- The exchange operator is Hermitian. We show this explicitly for the position degree of freedom by noting that

$$\begin{aligned}
 \langle \hat{\mathcal{E}}_{12} \rangle^* &= \left[\int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 \psi^*(x_1, x_2) \hat{\mathcal{E}}_{ij} \psi(x_1, x_2) \right]^* \\
 &= \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 \psi^*(x_2, x_1) \psi(x_1, x_2) \\
 &= \int_{-\infty}^{+\infty} dy_2 \int_{-\infty}^{+\infty} dy_1 \psi^*(y_1, y_2) \psi(y_2, y_1) \\
 &= \int_{-\infty}^{+\infty} dy_2 \int_{-\infty}^{+\infty} dy_1 \psi^*(y_1, y_2) \hat{\mathcal{E}}_{12} \psi(y_1, y_2) \\
 \langle \hat{\mathcal{E}}_{12} \rangle^* &= \langle \hat{\mathcal{E}}_{12} \rangle
 \end{aligned} \tag{14.81}$$

where a simple relabeling of the dummy integration variables is used; the similar proof for spin wavefunctions is discussed in P14.8.

- The operator $\hat{\mathcal{E}}_{ij}$ certainly commutes with the many-body Hamiltonian because

$$\begin{aligned}
 [\hat{H}, \hat{\mathcal{E}}_{12}] \psi(1; 2) &= \hat{H} \hat{\mathcal{E}}_{12} \psi(1; 2) - \hat{\mathcal{E}}_{12} \hat{H} \psi(1; 2) \\
 &= \hat{H} \psi(2; 1) - \hat{\mathcal{E}}_{ij} E_{(12)} \psi(1; 2) \\
 &= (E_{(21)} - E_{(12)}) \psi(1; 2) = 0
 \end{aligned} \tag{14.82}$$

since the energy is an observable which should be unchanged by interchange. We then know that the states of the system are simultaneous eigenfunctions of both the energy and *all* the exchange operators.

- The square of the exchange operator is just the identity since

$$\left(\hat{\mathcal{E}}_{12} \right)^2 \psi(1; 2) = \hat{\mathcal{E}}_{12} \hat{\mathcal{E}}_{12} \psi(1; 2) = \hat{\mathcal{E}}_{12} \psi(2; 1) = \psi(1; 2) \tag{14.83}$$

Just as with the parity operator (Section 6.6), this fact implies that the eigenvalues of the exchange operator are ± 1 , corresponding to states which are symmetric (+1) and antisymmetric (−1) under interchange of any two particles, that is

$$\text{symmetric: } \psi(2; 1) = \hat{\mathcal{E}}_{12} \psi(1; 2) = +\psi(1; 2) \tag{14.84}$$

or

$$\text{antisymmetric: } \psi(2; 1) = \hat{\mathcal{E}}_{12} \psi(1; 2) = -\psi(1; 2). \tag{14.85}$$

This is certainly consistent with the fact that the probability distribution of two exchanged wavefunctions should be the same under exchange, namely that

$$P(2; 1) = |\psi(2; 1)|^2 = |\pm \psi(1; 2)|^2 = P(1; 2) \quad (14.86)$$

- These last two points, taken together, imply a very powerful constraint on the total wavefunction (by which we mean both position and spin degrees of freedom), namely:

The wavefunction of N IND particles must be either totally symmetric (S) or totally antisymmetric (A) under the exchange of any two of the IND particles, that is either

$$\text{totally symmetric: } \hat{\mathcal{E}}_{ij}\psi_S(1; 2; \dots; N) = +\psi_S(1; 2; \dots; N) \quad (14.87)$$

or

$$\text{totally antisymmetric: } \hat{\mathcal{E}}_{ij}\psi_A(1; 2; \dots; N) = -\psi_A(1; 2; \dots; N) \quad (14.88)$$

for all possible pairs (i, j) .

This important result is the key ingredient in determining the correct form of the quantum wavefunction for a system of IND particles. Using the states in Eqn. (14.74) as an example, we see that none of them satisfy either (14.87) or (14.88) by themselves. One can see, however, that the linear combinations

$$\begin{aligned} \psi_S(1; 2; 3) = C_S [& \phi_0(x_1)\phi_1(x_2)\phi_2(x_3) + \phi_0(x_1)\phi_1(x_3)\phi_2(x_2) \\ & + \phi_0(x_2)\phi_1(x_1)\phi_2(x_3) + \phi_0(x_2)\phi_1(x_3)\phi_2(x_1) \\ & + \phi_0(x_3)\phi_1(x_1)\phi_2(x_2) + \phi_0(x_3)\phi_1(x_2)\phi_2(x_1)] \end{aligned} \quad (14.89)$$

and

$$\begin{aligned} \psi_A(1; 2; 3) = C_A [& \phi_0(x_1)\phi_1(x_2)\phi_2(x_3) - \phi_0(x_1)\phi_1(x_3)\phi_2(x_2) \\ & - \phi_0(x_2)\phi_1(x_1)\phi_2(x_3) + \phi_0(x_2)\phi_1(x_3)\phi_2(x_1) \\ & + \phi_0(x_3)\phi_1(x_1)\phi_2(x_2) - \phi_0(x_3)\phi_1(x_2)\phi_2(x_1)] \end{aligned} \quad (14.90)$$

are respectively symmetric and antisymmetric under the interchange of any two labels; the constants C_S , C_A are determined, of course, by the overall normalization.

A general prescription for the construction of such properly symmetrized or antisymmetrized linear combinations is easy to generate. For the symmetric combination, one can take

$$\text{completely symmetric: } \psi_S(1; 2; \dots; N) = C_S \sum_P \psi(1; 2; \dots; N) \quad (14.91)$$

where \sum_P denotes the sum over all possible permutations of the N indices; this form certainly reproduces Eqn. (14.89). While there can be as many as $N!$ terms in this sum, if the IND particles are not all in different quantum states, the number of terms can be far less. For example, if the N particles are all in the same state given by Eqn. (14.72), the permutations in (14.91) are all identical and we find

$$\psi_S(1; 2 \dots; N) = C_S N! \psi(1; 2; \dots; N). \quad (14.92)$$

In this language, the antisymmetric combination can be written schematically as

$$\text{completely antisymmetric: } \psi_A(1; 2; \dots; N) = C_A \sum_P (-1)^P \psi(1; 2; \dots; N). \quad (14.93)$$

- Here n_P is the *number of two-particle permutations or exchanges* which are required to achieve the overall permutation denoted by P starting from the canonical ordering $(1; 2 \dots; N)$; this factor gives the alternating signs required by the antisymmetry.

As an example of this last case, we note that

$$\begin{aligned} \psi(2; 1; 3) &= \hat{\mathcal{E}}_{12} \psi(1; 2; 3) \implies n_P = 1, (-1)^{n_P} = -1 \\ \psi(3; 1; 2) &= \hat{\mathcal{E}}_{13} \hat{\mathcal{E}}_{23} \psi(1; 2; 3) \implies n_P = 2, (-1)^{n_P} = +1 \end{aligned} \quad (14.94)$$

as in Eqn. (14.90). This form also implies that:

- No two particles described by a totally antisymmetric wavefunction can be in the same quantum state,

which we can see as follows. Suppose particles i and j were in the same quantum state, namely, $i = j$; then since the overall wavefunction must be antisymmetric under $\hat{\mathcal{E}}_{ij}$, we have

$$\psi(1; 2; \dots; j; \dots; i; \dots; N) = -\psi(1; 2; \dots; i; \dots; j; \dots; N) \text{ by antisymmetry}$$

\Downarrow

$$\begin{aligned} \psi(1; 2; \dots; i; \dots; i; \dots; N) &= -\psi(1; 2; \dots; i; \dots; i; \dots; N) \text{ since } i = j, \text{ implying} \\ \psi(1; 2; \dots; i; \dots; i; \dots; N) &= 0 \end{aligned} \quad (14.95)$$

The more careful way of stating this result is that:

- The wavefunction (and hence the probability density) for two particles in a completely antisymmetric state to occupy the same quantum “niche” vanishes.

For the case of noninteracting particles where the wavefunction can be written in product form, the antisymmetric combination can be written in an especially simple form using a determinant, namely

$$\psi_A(1; 2; \dots; N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \phi_{a_1}(1) & \phi_{a_2}(1) & \cdots & \phi_{a_N}(1) \\ \phi_{a_1}(2) & \phi_{a_2}(2) & \cdots & \phi_{a_N}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{a_1}(N) & \phi_{a_2}(N) & \cdots & \phi_{a_N}(N) \end{pmatrix} \quad (14.96)$$

The “recipe” for constructing this matrix is to put the N (necessarily different) single-particle wavefunctions in succeeding *columns* while the particle state labels $1, 2, \dots, N$ are then inserted in different *rows*. The overall antisymmetry of the wavefunction is guaranteed by the linear algebra result that the exchange of any two rows (or columns) of a matrix introduces a factor of (-1) in the determinant. The overall normalization constant is correct provided each ϕ_{a_i} is properly normalized (P14.9). This form is called a *Slater determinant* and is useful even when the particles interact with each other as it can be used as a trial wavefunction for a variational calculation.

A similar shorthand notation for the *symmetric* state is

$$\psi_S(1; 2; \dots; N) = C_S \det \begin{pmatrix} \phi_{a_1}(1) & \phi_{a_2}(1) & \cdots & \phi_{a_N}(1) \\ \phi_{a_1}(2) & \phi_{a_2}(2) & \cdots & \phi_{a_N}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{a_1}(N) & \phi_{a_2}(N) & \cdots & \phi_{a_N}(N) \end{pmatrix}_+ \quad (14.97)$$

where the $+$ subscript indicates that the determinant should be taken with all positive signs. For symmetric states, more than one particle can be in a given state, not all of the resulting terms will necessarily be different, and the normalization must be determined case by case; an example is Eqn. (14.92) and P14.9.

The requirement that IND particles have totally symmetric or antisymmetric wavefunctions has therefore reduced the possible ambiguity in the number of quantum states describing the same physics from being as large as $N!$ possible choices to only two. The physical property of the particles in question that determines which choice is actually realized in nature is their intrinsic angular momentum or spin, specifically whether the particles have integral spin ($J = 0, 1, 2, \dots$) or half-integral spin ($1/2, 3/2, 5/2, \dots$); the former are called *bosons* while the latter are known as *fermions*. This distinction is the content of the *spin-statistics theorem* which states that:

- The total wavefunction (including both spin and position information) of a system of indistinguishable bosons (fermions) must be symmetric (antisymmetric) under the interchange of any two particles.

The exclusion principle, as stated in Section 7.1 that “*no two electrons may be in the same quantum state*” is seen to be an immediate consequence of this result from Eqn. (14.95). The same result must then hold for neutrons and protons in nuclear systems, quarks inside nucleons, and all other particles with half-integral spin. Besides yielding a “no-go” theorem for what is not allowed, the spin-statistics theorem provides a prescription for the construction of the appropriate wavefunction for a system of IND particles via Eqns (14.91) and (14.93) and, as such, is a much more powerful statement about how nature organizes itself.

Example 14.5. Two electrons with spin in a box

Consider two noninteracting electrons in an infinite well potential. The ground state of the system is achieved when both particles are in the lowest allowed energy state with $E_{\text{tot}} = 2E_1$. This is allowed provided their spins are different (as in Fig. 14.2(a)), in which case the Slater determinant wavefunction is

$$\begin{aligned}\psi(1; 2) &= \frac{1}{\sqrt{2!}} \det \begin{pmatrix} u_1(x_1)\chi_1^{(+)} & u_1(x_1)\chi_1^{(-)} \\ u_1(x_2)\chi_2^{(+)} & u_1(x_2)\chi_2^{(-)} \end{pmatrix} \\ &= u_1(x_1)u_2(x_2) \frac{1}{\sqrt{2}} \left(\chi_1^{(+)}\chi_2^{(-)} - \chi_2^{(+)}\chi_1^{(-)} \right) \quad (14.98)\end{aligned}$$

The antisymmetric wavefunction with both particles in the ground state *and* with spins aligned vanishes, as in Eqn. (14.95).

For the first excited state of the system, one electron can be “elevated” to the next energy level so that $E_{\text{tot}} = E_1 + E_2$ and both spin configurations in Fig. 14.2(b) and (c) are possible; the corresponding wavefunctions are then

$$\begin{aligned}\psi(1; 2) &= \frac{1}{\sqrt{2!}} \det \begin{pmatrix} u_1(x_1)\chi_1^{(+)} & u_2(x_1)\chi_1^{(-)} \\ u_1(x_2)\chi_2^{(+)} & u_2(x_2)\chi_2^{(-)} \end{pmatrix} \quad (14.99) \\ &= \frac{1}{\sqrt{2}} \left(u_1(x_1)u_2(x_2)\chi_1^{(+)}\chi_2^{(-)} - u_2(x_1)u_1(x_2)\chi_1^{(-)}\chi_2^{(+)} \right)\end{aligned}$$

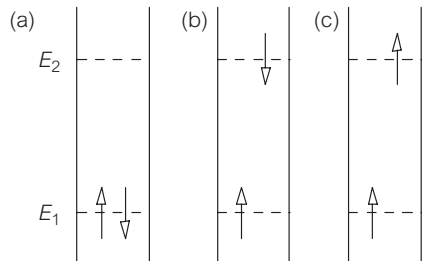


Figure 14.2. Allowed states of two indistinguishable spin-1/2 particles in an infinite well; Case (a) is the ground state, while (b) and (c) show two possible first excited states.

(Continued)

and

$$\begin{aligned}\psi(1; 2) &= \frac{1}{\sqrt{2!}} \det \begin{pmatrix} u_1(x_1)\chi_1^{(+)} & u_2(x_1)\chi_1^{(+)} \\ u_1(x_2)\chi_2^{(+)} & u_2(x_2)\chi_2^{(+)} \end{pmatrix} \\ &= \chi_1^{(+)} \chi_1^{(+)} \frac{1}{\sqrt{2}} (u_1(x_1)u_2(x_2) - u_2(x_1)u_1(x_2))\end{aligned}\quad (14.100)$$

with a similar state with both spins down also possible. These are just the wavefunctions of Eqns (14.68)–(14.70).

We know that the presence of mutual interactions between particles will induce dynamical correlations between them which are reflected in their quantum wavefunctions. What is more surprising is that IND particles exhibit such correlations even when they do not interact, simply due to the requirement of indistinguishability; these can be called an effective “Fermi repulsion” and “Bose attraction” which we illustrate in the next example.

Example 14.6. Correlations due to indistinguishability

Consider two particles of mass m in a harmonic oscillator potential; assume that we somehow know that there is one particle in the ground state (ψ_0) and one in the first excited state (ψ_1). We define the three wavefunctions

$$\psi_D(x_1, x_2) = \psi_0(x_1)\psi_1(x_2) \quad (14.101)$$

$$\psi_B(x_1, x_2) = \frac{1}{\sqrt{2}} (\psi_0(x_1)\psi_1(x_2) + \psi_1(x_1)\psi_0(x_2)) \quad (14.102)$$

$$\psi_F(x_1, x_2) = \frac{1}{\sqrt{2}} (\psi_0(x_1)\psi_1(x_2) - \psi_1(x_1)\psi_0(x_2)) \quad (14.103)$$

where

$$\psi_0(x) = \sqrt{\frac{1}{\rho\sqrt{\pi}}} e^{-x^2/2\rho^2} \quad \text{and} \quad \psi_1(x) = \sqrt{\frac{2}{\rho\sqrt{\pi}}} \left(\frac{x}{\rho}\right) e^{-x^2/2\rho^2} \quad (14.104)$$

are the appropriate SHO position-space wavefunctions, and the labels stand for distinguishable (D), boson (B), and fermion (F), respectively. We imagine for example that the bosons have no spin so that their wavefunction must be symmetric in the position coordinate, while the fermions have a symmetric spin wavefunction (which we do not exhibit) implying that their position-space wavefunction must be odd under exchange. There is, of course, another distinguishable wavefunction with $1 \leftrightarrow 2$.

(Continued)

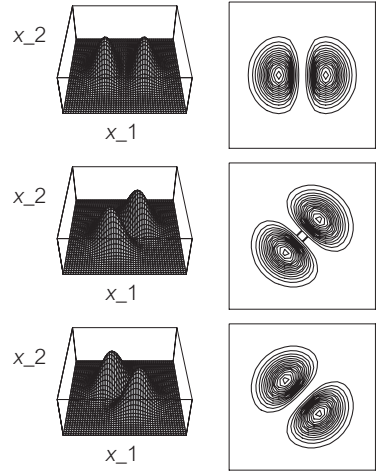


Figure 14.3. Three-dimensional plots (left) and contour plots (right) of $|\psi(x_1, x_2)|^2$ versus x_1, x_2 for the case of (top) distinguishable particles, (middle) indistinguishable bosons, and (bottom) indistinguishable fermions.

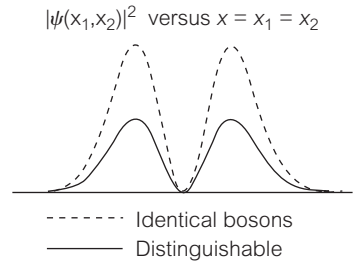


Figure 14.4. A “slice” through the contour plots of Fig. 14.3 along the $x_1 = x_2$ direction. The fermion probability distribution vanishes (Pauli principle), and the identical boson configuration is twice as likely as the indistinguishable particle state.

To see the correlations contained in these wavefunctions, we calculate expectation values and note that $\langle x_1 \rangle = \langle x_2 \rangle = 0$ for all cases but:

$$\begin{aligned}
 D : \langle x_1^2 \rangle &= \rho^2/2 & \langle x_2^2 \rangle &= 3\rho^2/2 & \langle x_1 x_2 \rangle &= 0 & \langle (x_1 - x_2)^2 \rangle &= 2\rho^2 \\
 B : \langle x_1^2 \rangle &= \rho^2 & \langle x_2^2 \rangle &= \rho^2 & \langle x_1 x_2 \rangle &= +\rho^2/2 & \langle (x_1 - x_2)^2 \rangle &= \rho^2 \\
 F : \langle x_1^2 \rangle &= \rho^2 & \langle x_2^2 \rangle &= \rho^2 & \langle x_1 x_2 \rangle &= -\rho^2/2 & \langle (x_1 - x_2)^2 \rangle &= 3\rho^2
 \end{aligned} \quad (14.105)$$

The two particles described by the fermion (boson) wavefunction are, on average, farther apart (closer together) than if they were distinguishable particles. This is illustrated in Fig. 14.3(a–c) for the three cases where we plot $|\psi(x_1, x_2)|^2$ versus x_1, x_2 . A “slice” through these plots along the line $x_1 = x_2$ is shown in Fig. 14.4 where we see that the totally symmetric wavefunction is twice as probable to be found with the particles in the same state; the totally antisymmetric wavefunction, of course, vanishes identically in this case.

(Continued)

We can further explore the implications of these correlations by “turning on” a mutual interaction between the two particles of the form

$$V'(x_2 - x_2; d) \equiv \frac{\lambda}{d\sqrt{\pi}} \exp\left(-\frac{(x_2 - x_2)^2}{2d^2}\right) \quad (14.106)$$

This function has the nice property that

$$\lim_{d \rightarrow 0} [V(x_2 - x_2; d)] = \lambda \delta(x_2 - x_2) \quad (14.107)$$

so that the particles only interact when they are on “top of each other” in this limit. This form is also convenient as an estimate of the effect of this interaction can be made using first order perturbation theory and the necessary overlap integrals can all be done analytically (P14.10). The shift in energy due to this perturbation at this order can be written as

$$\begin{aligned} E_D^{(1)} = \Delta E_D &= \frac{\lambda}{\sqrt{\pi}} \frac{\rho^2 + d^2}{(2\rho^2 + d^2)^{3/2}} = V_b \left[\frac{1 + z^2}{(2 + z^2)^{3/2}} \right] \\ E_B^{(1)} = \Delta E_B &= \frac{\lambda}{\sqrt{\pi}} \frac{2\rho^2 + d^2}{(2\rho^2 + d^2)^{3/2}} = V_b \left[\frac{2 + z^2}{(2 + z^2)^{3/2}} \right] \\ E_F^{(2)} = \Delta E_F &= \frac{\lambda}{\sqrt{\pi}} \frac{d^2}{(2\rho^2 + d^2)^{3/2}} = V_b \left[\frac{z^2}{(2 + z^2)^{3/2}} \right] \end{aligned} \quad (14.108)$$

where $V_b \equiv \lambda/b\sqrt{\pi}$ and $z \equiv d/\rho$; we plot these results in Fig. 14.5. One sees that when $d \ll \rho$ ($z \ll 1$ or “range of mutual interaction” \ll “particle separation”), the perturbation “samples” the various wavefunctions in a region where the correlations are dramatic, and the resulting energy shifts are very different; when $d \gg \rho$ ($z \gg 1$) the effects of indistinguishability are less important.

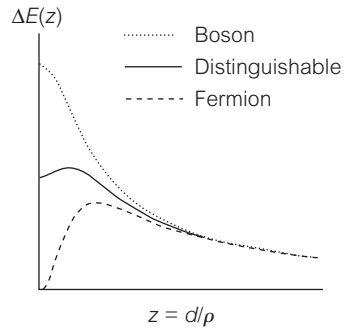


Figure 14.5. Energy shift, $\Delta E(z)$, versus the range of the mutual interaction, $z = d/\rho$. Results for the distinguishable (solid), indistinguishable boson (dotted), and indistinguishable fermion (dashed) states are shown.

14.6 Questions and Problems

Q14.1. What are the *dimensions* of an N -particle wavefunction in one dimension? in three dimensions?

Q14.2. Does the notion of probability flux generalize to multiparticle wavefunctions?

Q14.3. Referring to Fig. 14.1, show how to *estimate* the ratio m_1/m_2 .

P14.1. **Multiparticle momentum-space wavefunctions.**

(a) Show that the definition

$$\phi(p_1, p_2) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 e^{-i(p_1 x_1 + p_2 x_2)/\hbar} \psi(x_1, x_2) \quad (14.109)$$

is appropriate for the momentum-space wavefunction corresponding to a two-particle position-space $\psi(x_1, x_2)$; specifically, show that $\phi(p_1, p_2)$ is normalized if $\psi(x_1, x_2)$ is, that the appropriate inverse relation holds, and anything else you think is important.

(b) Evaluate the momentum-space wavefunction corresponding to $\psi(x_1, x_2)$ in Example 14.1 and show that it is proportional to

$$\phi(p_1, p_2) = \exp\left(-\frac{(cp_1^2 - 2bp_1p_2 + ap_2^2)}{2\hbar^2(ac - b^2)}\right) \quad (14.110)$$

Normalize this wavefunction, show that this form has the right limit as $b \rightarrow 0$ and $ac - b^2 \rightarrow 0$, and interpret your results.

(c) Find the covariance for the variables p_1, p_2 and show that it is opposite in sign to that for x_1, x_2 and interpret your result.

P14.2. Consider two *distinguishable* and noninteracting particles of mass m moving in the same harmonic oscillator potential.

(a) What is the ground state energy E_0 and wavefunction $\psi_0(x_1, x_2)$ of the two-particle system? Is the ground state energy degenerate?

(b) Show that the first excited state E_1 is doubly degenerate and write down the two wavefunctions, $\psi_1^{a,b}(x_1, x_2)$. Consider a δ -function interaction between the particles as in Example 14.2, $V'(x_1 - x_2) = g\delta(x_1 - x_2)$ -as a small perturbation. Find the perturbed energies and eigenfunctions.

(c) Show that the second excited state is triply degenerate and write down the possible wavefunctions. Show if a $g\delta(x_1 - x_2)$ mutual interaction is added

that the resulting eigenstates are given by

$$\begin{aligned}\psi_2^a(x_1, x_2) &= \frac{1}{\sqrt{2}} [\psi_0(x_1)\psi_2(x_2) - \psi_2(x_1)\psi_0(x_2)] \\ \psi^b(x_1, x_2) &= \frac{1}{2} [\psi_0(x_1)\psi_2(x_2) - \sqrt{2}\psi_1(x_1)\psi_1(x_2) + \psi_2(x_1)\psi_0(x_2)] \\ \psi^c(x_1, x_2) &= \frac{1}{2} [\psi_0(x_1)\psi_2(x_2) + \sqrt{2}\psi_1(x_1)\psi_1(x_2) + \psi_2(x_1)\psi_0(x_2)]\end{aligned}$$

where the $\psi_n(x)$ are the one-particle SHO eigenfunctions. Show that these states are mutually orthogonal. If the mutual interaction is repulsive ($g > 0$), which state has the highest energy? the lowest energy? What are the energy eigenvalues? Hint: For this three-state system you have to diagonalize a 3×3 matrix. You presumably know one linear combination which would be unaffected by the perturbation and hence one eigenvector and eigenfunction. Extracting this one helps you solve for the other two.

P14.3. Consider the operator representing the total momentum of a multiparticle system, namely

$$\hat{P} = \hat{p}_1 + \hat{p}_2 + \cdots + \hat{p}_N \quad (14.111)$$

The time-dependence of the expectation value of this operator will be given (recall Section 12.5) by

$$\frac{d}{dt} \langle \hat{P} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{P}] \rangle \quad (14.112)$$

where \hat{H} is now the multiparticle Hamiltonian in Eqn. (14.2). Show that this can be written as

$$\frac{d}{dt} \langle \hat{P} \rangle = \langle \mathcal{F}_1 + \mathcal{F}_2 + \cdots + \mathcal{F}_N \rangle = \langle \mathcal{F}_{\text{tot}} \rangle \quad (14.113)$$

where \mathcal{F}_{tot} corresponds to the total external force; if this vanishes, the total momentum is conserved as in the classical case. Note: This result implies that the effects of the two-body mutual forces cancel and you should show that

$$\frac{\partial V_{ij}}{\partial x_i} + \frac{\partial V_{ij}}{\partial x_j} = 0 \quad (14.114)$$

P14.4. In changing to center-of-mass and relative coordinates for a two-body system, the wavefunction in terms of the original x_1, x_2 coordinates can be recovered by using

$$\Psi(X) \phi(x) = \Psi(X(x_1, x_2)) \phi(x(x_1, x_2)) \longrightarrow \psi(x_1, x_2) \quad (14.115)$$

We have stressed that a probability interpretation requires not only the value of $|\psi|^2$, but also the measure, so it is more relevant to compare

$$|\Psi(X(x_1, x_2))\phi(x(x_1, x_2))|^2 dX dx \stackrel{?}{\longleftrightarrow} |\psi(x_1, x_2)|^2 dx_1 dx_2 \quad (14.116)$$

Use the Jacobian of the transformation from x_1, x_2 to x, X to show that $dX dx = dx_1 dx_2$. Hint: The Jacobian of a coordinate transformation from (w, v) to (x, y) is given by

$$dw dv = \left| \det \begin{pmatrix} \partial w / \partial x & \partial w / \partial y \\ \partial v / \partial x & \partial v / \partial y \end{pmatrix} \right| dx dy \quad (14.117)$$

(As a test, recall the change from Cartesian to polar coordinates and show that $dx dy = r dr d\theta$.)

P14.5. Consider two *distinguishable* particles of mass m moving in the potential

$$V(x_1, x_2) = \frac{1}{2}m\omega^2 x_1^2 + \frac{1}{2}m\omega^2 x_2^2 \quad (14.118)$$

- (a) Using the fact that the potential is separable in the x_1, x_2 coordinates, find the energy spectrum and ground state wavefunction. Since this is a product wavefunction, there can be no correlations between the two particles; show that $\text{cov}(x_1, x_2)$ vanishes in any state. Recall that

$$\text{cov}(x_1, x_2) \equiv \langle (x_1 - \langle x_1 \rangle)(x_2 - \langle x_2 \rangle) \rangle \quad (14.119)$$

- (b) Show that the potential also separates when expressed in center-of-mass (X) and relative (x) coordinates and find the energy spectrum and ground state wavefunction in these coordinates. Show that the degeneracy is the same in each representation and that the ground state wavefunctions agree.
- (c) Add an additional mutual interaction of the form

$$V'(x_1, x_1) = V'(x_1 - x_2) = \lambda(x_1 - x_2)^2 \quad (14.120)$$

and find the energy spectrum exactly using center-of-mass and relative coordinates. Are the energy levels changed in the way you expect from the form of V' ?

- (d) When such a mutual interaction is present, we expect the positions of the particles to be correlated. Evaluate $\text{cov}(x_1, x_2)$ for the ground state wavefunction and show that it is proportional to λ when λ is small. Convince yourself that the correlation should be positive (negative) when $\lambda > 0$ (< 0). Hint: Use the coordinate transformations to show things like

$$\langle x_1 \rangle = \left\langle X + \frac{1}{2}x \right\rangle = 0 + 0 \quad (14.121)$$

and

$$\langle x_1^2 \rangle = \langle X^2 \rangle + \langle Xx \rangle + \frac{1}{4}\langle x^2 \rangle \quad (14.122)$$

P14.6. Reduced mass effects in “heavy hydrogen.”

- (a) The energy levels of hydrogen-like atoms with a proton nucleus (ordinary hydrogen or H^1) will be slightly different from those with a deuteron nucleus (so-called “heavy hydrogen” or deuterium, H^2) which has roughly twice the mass of a proton, due to reduced mass effects. Show that the shift in wavelength of a given line for H^2 relative to H^1 is roughly

$$\frac{\Delta\lambda}{\lambda} \approx m_e \left(\frac{1}{M_D} - \frac{1}{m_p} \right) \quad (14.123)$$

where $M_D \approx 2m_p$. Evaluate this fractional change numerically.

- (b) The original discovery of deuterium was made by looking for such shifts in the *visible*, atomic Balmer spectra of hydrogen. The original paper⁴ says that

When with ordinary hydrogen, the times of exposure required to just record the strong H^1 lines were increased 4000 times, very faint lines appeared at the calculated positions for the H^2 lines” . . . “on the short wave-length side and separated from them by between 1 and 2 Å.”

Using the result of part (a), quantitatively explain the wavelength shifts observed. Estimate the relative abundance of H^2 and H^1 in normal hydrogen.

- P14.7. Permutation groups.** We have seen that permutations play an important role in the physics of IND particles and in this problem you are asked to study some of the properties of the permutation group, using the case of three particles as an example. Consider three ($N = 3$) objects, labeled a , b , and c which can be in the three positions 1, 2, and 3; there are then $N! = 3! = 6$ different ways in which the labels can be placed. These permutations can all be obtained from one standard labeling, say (a, b, c) by the action of 6 permutation operators;

$$\begin{aligned} 1(a, b, c) &\longrightarrow (a, b, c) \\ (12)(a, b, c) &\longrightarrow (b, a, c) \\ (13)(a, b, c) &\longrightarrow (c, b, a) \\ (23)(a, b, c) &\longrightarrow (a, c, b) \\ (231)(a, b, c) &\longrightarrow (b, c, a) \\ (312)(a, b, c) &\longrightarrow (c, a, b) \end{aligned}$$

The element **1** is the identity operator. The natural multiplication on these group elements is obtained by letting the permutations act in order, for example,

$$(g_1 \cdot g_2)(a, b, c) \implies g_1(g_2(a, b, c)) \quad (14.124)$$

⁴ Urey, Brickwedde, and Murphy (1932).

so that, for example,

$$(12) \cdot (231) = (13) \quad \text{and} \quad (23) \cdot (13) = (312) \quad (14.125)$$

Complete the multiplication table below and show that these elements form a group satisfying all of the requirements in the definition of Appendix F.2.

| | 1 | (12) | (13) | (23) | (231) | (312) |
|-------|-------|------|-------|------|-------|-------|
| 1 | 1 | (12) | (13) | (23) | (231) | (312) |
| (12) | (12) | | | | (13) | |
| (13) | (13) | | | | | |
| (23) | (23) | | (312) | | | |
| (231) | (231) | | | | | |
| (312) | (312) | | | | | |

P14.8. (a) Generalize the proof in Eqn. (14.81) to show that the exchange operator $\hat{\mathcal{E}}_{ij}$ is Hermitian by showing that $\langle \hat{\mathcal{E}}_{ij} \rangle$ is real when evaluated with any multiparticle position space wavefunction.

(b) Show that the expectation value of $\hat{\mathcal{E}}_{12}$ is real when evaluated between spin-states, that is, show that

$$\langle \chi_1; \chi_2 | \hat{\mathcal{E}}_{12} | \chi_1; \chi_2 \rangle = (\alpha_1^*, \beta_1^*)(\alpha_2^*, \beta_2^*) \hat{\mathcal{E}}_{12} \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix} = |\alpha_1 \alpha_2^* + \beta_1 \beta_2^*|^2 \quad (14.126)$$

is real.

P14.9. (a) Show that the Slater determinant in Eqn. (14.96) is properly normalized.

(b) Four spinless particles move in the same harmonic oscillator potential; two are in the ground state and two in the first excited state. Write down the normalized wavefunction for this system.

P14.10. Confirm the results in Eqn. (14.108).

P14.11. Consider two IND spin-1/2 particles which interact via the potential $V(x_1 - x_2) = k(x_1 - x_2)^2/2$.

(a) Ignoring spin for the moment, show that the position-space wavefunctions can be written as

$$\psi(x_1, x_2) \propto e^{iP(x_1+x_2)/2\hbar} \psi_n(x_1 - x_2) \quad (14.127)$$

where the ψ_n are the harmonic oscillator eigenstates.

(b) Show that under the exchange $1 \leftrightarrow 2$ that these solutions satisfy

$$\hat{\mathcal{E}}_{12} \psi(x_1, x_2) = (-1)^n \psi(x_1, x_2) \quad (14.128)$$

- (c) Add the appropriate symmetric or antisymmetric spin wavefunctions and find the allowed states of the system.

P14.12. Consider the following very simplified model of the lithium 7 nucleus (${}^7\text{Li}$), consisting of 3 protons and 4 neutrons in a one-dimensional infinite well of width a . Assume that the protons and neutrons do not interact with each other.

- (a) What is the ground state energy for this system?
- (b) Write down the normalized wavefunction for the ground state.
- (c) What is the energy of the first excited state?

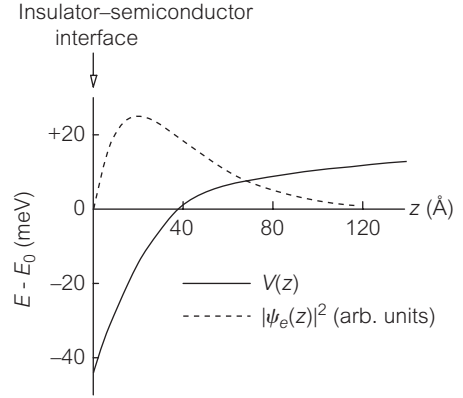
Three

Two-Dimensional Quantum Mechanics

One-dimensional (1D) systems provide examples of many of the most important features of quantum mechanics, but it is also instructive to consider two-dimensional (hereafter 2D or planar) systems for several reasons:

- Systems with two spatial degrees of freedom provide more opportunities to study multivariable probability concepts, separation of coordinates techniques, and new mathematical methods and special functions. They also allow for the visualization of many quantum phenomena which arise in more realistic three-dimensional (3D) systems, but which are obviously difficult to plot in 3D.
- Two-dimensional systems naturally exhibit symmetries not present in 1D systems, and provide a glimpse of the intimate connection between symmetries and the degeneracy of energy levels.
- Two-dimensional systems allow one to study rotational motion and its symmetries as well as the properties of angular momentum, both quantum mechanically and in its approach to the classical limit. For example, charged particles in a uniform magnetic field classically can undergo circular planar orbits (see Section 18.5); a similar quantum system of electrons in two-dimensions in a uniform \mathbf{B} field has important implications for the understanding of the so-called quantum Hall effect in condensed matter physics.
- Finally, and perhaps most importantly, while often considered of pedagogical use only, 2D systems of particles are rapidly becoming of more practical importance as their realization in surface physics becomes increasingly easy. It is now possible, using modern crystal growth techniques such as molecular-beam epitaxy and other methods, to fabricate semiconductor nanostructures, artificially created patterns of atoms whose atomic composition and sizes are

Figure 15.1. Semischematic representation of the potential energy, $V(z)$ (solid curve), and electron wavefunction, $|\psi_e(z)|^2$ (dashed curve), versus the distance from the surface, z , for a 2D electron gas near an insulator–semiconductor interface, indicating the approximate localization distances and energy scales. Adapted from von Klitzing (1987).



controllable at the nanometer scale, which is comparable to interatomic distances. At such length scales, quantum effects obviously become increasingly important. Even more dramatically, scanning tunnel microscopy (STM) techniques can now be used to manipulate individual atoms and molecules¹ with atomic scale precision. The quantum corral shown in Fig. 1.3 was constructed in this manner.

As an example of such a system, consider a 2D electron gas, bound to a surface or interface between surfaces by the potential shown in Fig. 15.1. The typical localization scale (determined, say, by the thickness of the interface layer) might be $L \sim 40 \text{ Å}$; this implies quantized energies in the direction perpendicular to the surface (here the z direction) of the order of

$$E_n \sim \frac{\hbar^2 \pi^2 n^2}{2m_e L^2} \sim n^2 25 \text{ meV} \quad (15.1)$$

so that the energy required to excite the electrons in this direction would be roughly $\Delta E \sim 75 \text{ meV}$. This can be compared to other typical energy scales for the two-dimensional electron gas which might be either:

- The thermal energy, $k_B T \lesssim 25 \text{ meV}$, for temperatures at, or below, room temperature (300 K) or
- The Fermi energy of the system. Recall from P7.2 that

$$E_F^{(2D)} = \frac{\hbar^2 \pi}{m_e} n_e^{(2D)} \quad (15.2)$$

¹ See, for example, Stroscio and Eigler (1991).

where $n_e^{(2D)} = N_{\text{tot}}/L^2$ is the electron surface density in two dimensions. Using a typical value for $n_e^{(2D)}$, this can be written in the form

$$E_F^{(2D)} = 2.5 \text{ meV} \left(\frac{n_e^{(2D)}}{10^{12} \text{ e}^-/\text{cm}^2} \right) \quad (15.3)$$

and we see that typical 2D collisions will not be able to “excite” electrons in the z direction, provided the density is not too high; thus the electron wavefunctions will stay effectively localized within the interface or on the surface (as in Fig. 15.1), and one can study an effectively 2D problem.

15.1 2D Cartesian Systems

The simplest example of a quantum system in two-dimensions is one described by Cartesian coordinates with a Hamiltonian of the form

$$\hat{H} = \frac{1}{2m} \left(\hat{p}_x^2 + \hat{p}_y^2 \right) + V(x, y) \quad (15.4)$$

where

$$\hat{p}_x = \frac{\hbar}{i} \frac{\partial}{\partial x} \quad \text{and} \quad \hat{p}_y = \frac{\hbar}{i} \frac{\partial}{\partial y} \quad (15.5)$$

with a wavefunction

$$\psi(x, y; t) = \psi(x, y) e^{-iEt/\hbar} \quad (15.6)$$

satisfying the time-independent Schrödinger equation

$$\hat{H}\psi(x, y) = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \psi(x, y) + V(x, y)\psi(x, y) = E\psi(x, y) \quad (15.7)$$

The probability density (now in two-dimensions) is given by $|\psi(x, y, t)|^2$, which must satisfy

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy |\psi(x, y; t)|^2 = 1 \quad (15.8)$$

and average or expectation values are calculated in the usual way via

$$\langle \hat{O} \rangle_t \equiv \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \psi^*(x, y; t) \hat{O} \psi(x, y; t) \quad (15.9)$$

If we consider *separable potentials* of the form

$$V(x, y) = V_x(x) + V_y(y) \quad (15.10)$$

we can assume a factorized form for the wavefunction

$$\psi(x, y; t) = X(x) Y(y) e^{-i(E_x + E_y)t/\hbar} \quad (15.11)$$

where each coordinate satisfies its own, 1D Schrödinger equation

$$\frac{\hat{p}_x^2}{2m} X(x) + V_x(x) X(x) = E_x X(x) \quad (15.12)$$

$$\frac{\hat{p}_y^2}{2m} Y(y) + V_y(y) Y(y) = E_y Y(y) \quad (15.13)$$

and the total energy of the system is given by $E = E_x + E_y$.

15.1.1 2D Infinite Well

A simple and instructive case is that of the 2D infinite well (or square box) with walls at $x = 0, L$ and $y = 0, L$; this is of the form above as we can define

$$V_{1D}(z; L) = \begin{cases} 0 & \text{for } 0 < z < L \\ \infty & \text{otherwise} \end{cases} \quad (15.14)$$

in which case the 2D potential is of the form

$$V_{2D}(x, y) = V_{1D}(x; L) + V_{1D}(y; L) \quad (15.15)$$

The fully normalized solutions can be written in the form

$$u_{(n,m)}(x, y) = u_n(x) u_m(y) = \frac{2}{L} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi y}{L}\right) \quad (15.16)$$

with

$$E_{(n,m)} = E_n + E_m = \frac{\hbar^2 \pi^2}{2mL^2} (n^2 + m^2) \quad (15.17)$$

and the spectrum is illustrated in Fig. 15.2. The wavefunctions for several sets of n, m are shown in Fig. 15.3 illustrating the wave properties of the system. We note that if the 1D infinite well corresponds to waves on a string with fixed ends, then this case corresponds to the vibrations of a square drumhead. Such wavefunctions are not only of pedagogical interest, as very similar patterns of “standing electron waves” can be directly observed on surfaces using STM techniques.²

² For one of the first experimental realizations, see Crommie, Lutz, and Eigler (1993).

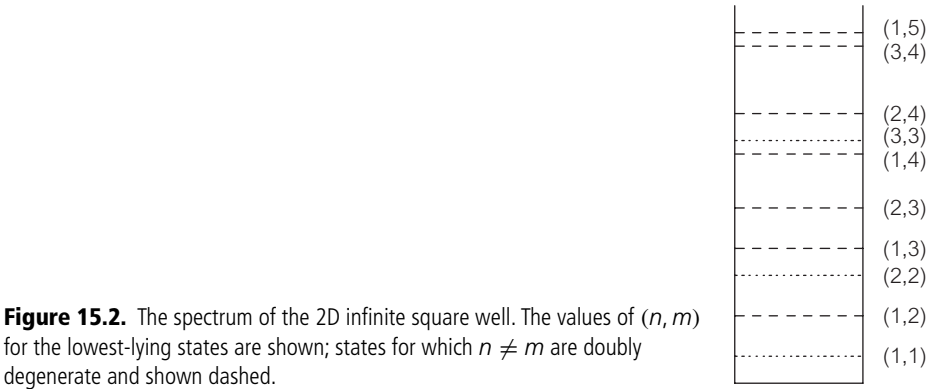


Figure 15.2. The spectrum of the 2D infinite square well. The values of (n, m) for the lowest-lying states are shown; states for which $n \neq m$ are doubly degenerate and shown dashed.

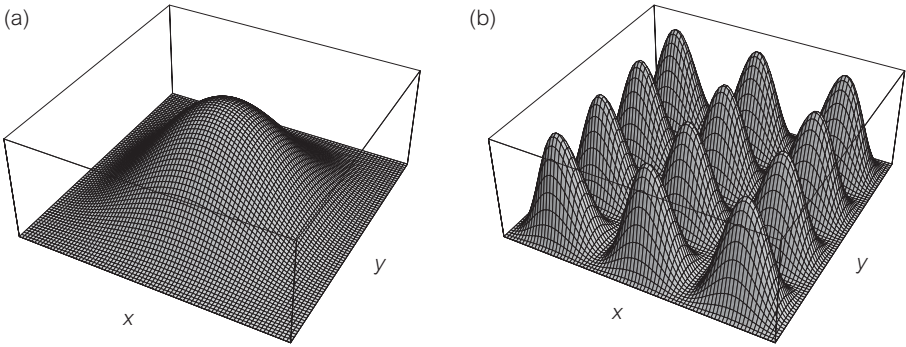


Figure 15.3. Plots of $|u_{(n,m)}(x, y)|^2$ versus (x, y) for the 2D infinite square well for (a) $(n, m) = (1, 1)$ and (b) $(n, m) = (3, 4)$.

Particle-like, wave packet solutions for this separable potential can be formed by using the linearity of the Schrödinger equation to write

$$\psi_{\text{WP}}(x, y; t) = \psi_{\text{WP}}(x; t)\psi_{\text{WP}}(y, t) \quad (15.18)$$

where one has

$$\psi_{\text{WP}}(x, t) = \sum_{n=1}^{\infty} a_n^{(x)} u_n(x) e^{-iE_n^{(x)} t/\hbar} \quad (15.19)$$

and similarly for y . For simplicity, we can use Gaussian weighting factors

$$a_n^{(x)} = e^{-(p_n - p_0^{(x)})^2 \alpha^2 / 2} e^{-ip_n x_0} \quad (15.20)$$

with $p_n = \hbar n/L$, as in Section 5.4.2, or the more rigorous version in Example 6.4. We can localize the packet initially such that it is centered at (x_0, y_0) with a central value of momentum $\mathbf{p}_0 = (p_0^{(x)}, p_0^{(y)})$. The ballistic propagation of such a wave packet with elastic collisions from the walls is illustrated in Fig. 15.4 where we

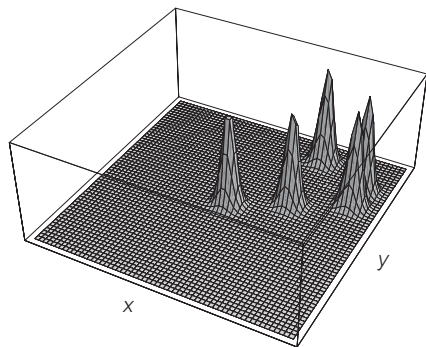


Figure 15.4. Series of snapshots (taken at equal time intervals) showing the propagation of a quasi-Gaussian wavepacket in a two-dimensional infinite square well; the packet is initially localized in the center of the well, with initial momenta $p_0^{(x)} = 2p_0^{(y)}$.

have chosen $(x_0, y_0) = (L/2, L/2)$ and $p_{(0)}^x = 2p_{(0)}^y$. Thus, the system can exhibit both wave- and particle-like behavior in a fashion similar to the 1D case. (2D infinite well potentials of this type, of arbitrary shapes or “footprints,” are often called *quantum billiard* systems.)

The energy eigenfunction solutions also form a complete set in that the time-dependence of any allowable wavefunction in the 2D infinite well can be written in the form

$$\psi(x, y; t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{(n,m)} u_{(n,m)}(x, y) e^{-iE_{(n,m)}t/\hbar} \quad (15.21)$$

where $|a_{(n,m)}|^2$ is the probability that a measurement of the energy associated with $\psi(x, y)$ will yield the value $E_{n,m}$. The form in Eqn. (15.18) is then a special case of the time-development of such a solution.

The most interesting new feature of this system is the fact that more than one independent energy eigenstate corresponds to the same energy level, at least for $n \neq m$, where the exchange $n \leftrightarrow m$ gives the same energy. Such a system is said to exhibit *degeneracy* and we say that:

- The quantum value of an observable quantity is *degenerate* when two (or more) independent eigenfunctions of an operator yield the *same* eigenvalue.

In this case, the degeneracy is easily traced to the symmetry of the potential (and kinetic energies) since the exchange of labels $x \leftrightarrow y$ has no observable effect on the system, leading naturally to a doubly degenerate set of levels (when $n \neq m$.) We can formalize this notion by introducing an *exchange operator*, $\hat{E}_{(x,y)}$, defined via

$$\hat{E}_{(x,y)} f(x, y) \equiv f(y, x) \quad (15.22)$$

which can be easily shown to be Hermitian even though its classical connection is far from obvious; this operator is very similar to the multiparticle exchange operator, \mathcal{E}_{ij} , used in Chapter 14.

Following the discussion of Section 12.5 on conserved quantities, we note that the statement above that the exchange $x \leftrightarrow y$ “has no effect on the system” can be associated with the fact that this operator commutes with the Hamiltonian, that is,

$$\left(\hat{H} \hat{E}_{(x,y)} - \hat{E}_{(x,y)} \hat{H} \right) \psi(x, y) = 0 \quad \text{or} \quad [\hat{H}, \hat{E}_{(x,y)}] = 0 \quad (15.23)$$

This implies that there will be simultaneous eigenfunctions of *both* \hat{H} and $\hat{E}_{x,y}$. By invoking the same arguments used previously for both the parity and exchange operators (mostly the fact that $(\hat{E}_{(x,y)})^2 = 1$), the eigenvalues (eigenfunctions) of the exchange operator can be seen to be ± 1 (even-odd functions under exchange), that is,

$$\hat{E}_{(x,y)} f^{(+)}(x, y) = +f^{(+)}(x, y) \quad \text{and} \quad \hat{E}_{(x,y)} f^{(-)}(x, y) = -f^{(-)}(x, y) \quad (15.24)$$

The $u_{(n,m)}(x, y)$ solutions individually do not, however, immediately satisfy this requirement. We note that any linear combination of degenerate energy eigenstates will also be an energy eigenstate with the same energy eigenvalue since

$$\begin{aligned} \hat{H} \left(\sum_E a_E \psi_E(x) \right) &= \sum_E a_E \hat{H} \psi_E(x) \\ &= \sum_E E a_E \psi_E(x) \\ &= E \left(\sum_E a_E \psi_E(x) \right) \end{aligned} \quad (15.25)$$

We are thus free to take appropriate linear combinations of degenerate solutions provided they remain orthogonal, and it is easy to see that the required combinations are

$$u_{(n,m)}^{(\pm)}(x, y) \equiv \frac{1}{\sqrt{2}} (u_{(n,m)}(x, y) \pm u_{(m,n)}(x, y)) \quad (15.26)$$

which do satisfy Eqn. (15.24). (See P15.8 for an application of these linear combinations.)

It is something of a “folk-theorem” (meaning roughly a statement which is universally accepted as being true, but difficult to state precisely and to prove in

each case) that:

- Most degeneracies are necessarily a result of *some* symmetry (sometimes not obvious) of the system under consideration.

The study of symmetries in quantum mechanics (under the guise of group theory) has had profound applications in atomic, nuclear, and elementary particle physics.

15.1.2 2D Harmonic Oscillator

Another separable Cartesian system is described by an isotropic harmonic oscillator, that is, a 2D mass and spring, defined by the potential energy

$$V(x, y) = \frac{1}{2}K(x^2 + y^2) = \frac{1}{2}m\omega^2(x^2 + y^2). \quad (15.27)$$

The product wavefunctions are given by

$$\psi_{(n,m)}(x, y) = \psi_n(x)\psi_m(y) \quad (15.28)$$

where the $\psi_n(x)$ are the solutions of Section 9.2.2. The energy spectrum is given by

$$E_{n,m} = E_n + E_m = (n + m + 1)\hbar\omega = (N + 1)\hbar\omega \quad (15.29)$$

which is illustrated in Fig. 15.5. As before, the energy levels with $n \leftrightarrow m$ are degenerate. The total degeneracy, that is, the number of distinct states, N_s , with energy value labeled by N is $N_s = N$; this is much larger than expected solely on the basis of the $x \leftrightarrow y$ symmetry. The enlarged degeneracy is partly due to the fact that the system also exhibits a symmetry under *rotations*, since the potential can also be written in the circularly symmetric way

$$V(x, y) = \frac{1}{2}K(x^2 + y^2) = \frac{1}{2}Kr^2 = V(r) \quad (15.30)$$

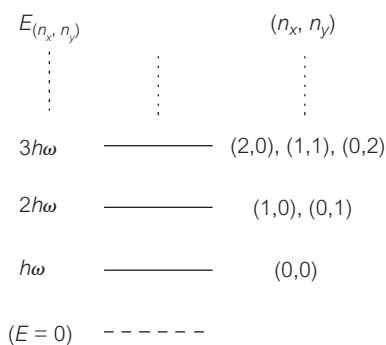


Figure 15.5. Energy spectrum and degeneracies for the 2D simple harmonic oscillator potential using Cartesian coordinates. Values of (n_x, n_y) for each level are shown.

and the system can then be separated in cylindrical coordinates as well, as we will see in Section 15.3.3.

Two-dimensional wave packets, *a la* Eqn. (15.18), can also be constructed in this case using, for example, the special Gaussian packet of Section 12.6.2, and can be shown to undergo semiclassical motion (P15.12). This is especially interesting in the case of an *non-isotropic spring*, that is, a potential of the form

$$V(x, y) = \frac{1}{2}K_x x^2 + \frac{1}{2}K_y y^2 \quad (15.31)$$

which is still separable. In this case, the natural vibration frequencies are different, $\omega_{x,y} = \sqrt{K_{x,y}/m}$, and the “trajectories” of the wave packets will, in general, not be periodic. In the special case where the frequencies are *commensurate*, namely, rational multiples of each other, that is

$$\frac{\omega_x}{\omega_y} = \frac{p}{q} \quad (15.32)$$

(with p, q integers) the classical motion *is* periodic and the quantum wave packets can reproduce the classical “Lissajous figures” discussed in many classical mechanics texts.³

15.2 Central Forces and Angular Momentum

15.2.1 Classical Case

Cartesian coordinates may not be the most natural set of variables for the study of many systems, and this is especially true for 2D systems described by a cylindrically symmetric potential of the form

$$V(\mathbf{r}) = V(r, \theta) = V(r) \quad (15.33)$$

In this case, the classical force is given by

$$\begin{aligned} \mathbf{F}(\mathbf{r}) &= -\nabla V(r, \theta) \\ &= -\frac{\partial V(r, \theta)}{\partial r} \hat{\mathbf{r}} - \frac{1}{r} \frac{\partial V(r, \theta)}{\partial \theta} \hat{\boldsymbol{\theta}} \\ &= -\frac{dV(r)}{dr} \hat{\mathbf{r}} \\ \mathbf{F}(\mathbf{r}) &= F(r) \hat{\mathbf{r}} \end{aligned} \quad (15.34)$$

³ See, for example, Marion and Thornton (2004).

where $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ are unit vectors in the radial and tangential directions, respectively. Thus, for a central potential, the force is directed radially toward (or away from) the origin.

The corresponding classical torque, $\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$, then vanishes and the relation

$$\frac{d\mathbf{L}}{dt} = \boldsymbol{\tau} = 0 \quad (15.35)$$

guarantees that the classical angular momentum, \mathbf{L} , is conserved, that is, is constant in time. Because of its importance as an additional conserved quantity (along with the total energy), we will discuss the quantum version of angular momentum extensively, in 2D in Section 15.2.2 and in 3D in Chapter 16.

The classical equations of motion for the particle in polar coordinates can be derived from those in Cartesian coordinates, namely,

$$\mathbf{F}(\mathbf{r}) = F(r) \hat{\mathbf{r}} = m\mathbf{a}(t) \quad (15.36)$$

giving

$$x: F(r) \left(\frac{x(t)}{r(t)} \right) = m\ddot{x}(t) \quad (15.37)$$

$$y: F(r) \left(\frac{y(t)}{r(t)} \right) = m\ddot{y}(t) \quad (15.38)$$

and by using the relations

$$x(t) = r(t) \cos(\theta(t)) \quad \text{and} \quad y(t) = r(t) \sin(\theta(t)) \quad (15.39)$$

to obtain

$$\frac{F(r)}{m} \cos(\theta) = \ddot{r} \cos(\theta) - 2\dot{r} \sin(\theta) \dot{\theta} - r \cos(\theta) \dot{\theta}^2 - r \sin(\theta) \ddot{\theta} \quad (15.40)$$

$$\frac{F(r)}{m} \sin(\theta) = \ddot{r} \sin(\theta) + 2\dot{r} \cos(\theta) \dot{\theta} - r \sin(\theta) \dot{\theta}^2 + r \cos(\theta) \ddot{\theta}. \quad (15.41)$$

The linear combination of equations given by $(15.41) \times (r \cos(\theta)) - (15.40) \times (r \sin(\theta))$ implies that

$$0 = 2\dot{r}r\dot{\theta} + r^2\ddot{\theta} = \frac{d}{dt} (r^2\dot{\theta}) = 0 \quad (15.42)$$

which is another statement of conservation of angular momentum as $L_z = r p = mrv = mr^2\dot{\theta}$ in polar coordinates. Using this identification, the other obvious combination, $(15.41) \sin(\theta) + (15.40) \cos(\theta)$, then gives the dynamical equation of motion for the radial coordinate,

$$F(r) = m\ddot{r} - mr\dot{\theta}^2 = m\ddot{r} - \frac{L_z^2}{mr^3} \quad (15.43)$$

which is the standard Newton's law result, including the familiar “centrifugal force” term. The derivation makes clear that this term is solely due to the proper accounting of the rotational motion and not to any “fictitious force.”

Most importantly for connections to quantum mechanics, the total energy will be constant for a conservative potential, so we can write

$$E = \frac{1}{2} m \mathbf{v}^2(t) + V(x, y) \quad (15.44)$$

and the chain rule, and Eqn. (15.39), gives

$$\mathbf{v}^2 = v_x^2 + v_y^2 = \dot{x}^2 + \dot{y}^2 = \dot{r}^2 + r^2 \dot{\theta}^2 \quad (15.45)$$

so that the total energy can be written as

$$E = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2) + V(r, \theta) = \frac{m \dot{r}^2}{2} + \frac{L_z^2}{2mr^2} + V(r, \theta) \quad (15.46)$$

The second term can be put in the form

$$T_{rot} = \frac{L_z^2}{2mr^2} = \frac{1}{2} mr^2 \dot{\theta}^2 = \frac{1}{2} I \omega^2 \quad (15.47)$$

where $I = mr^2$ is the rotational moment of inertia for a point mass; this makes it clear that it represents the rotational kinetic energy. It is this form for the energy which can be most easily generalized to a quantum mechanical Hamiltonian.

15.2.2 Quantum Angular Momentum in 2D

To extend the notion of angular momentum to quantum mechanical operators, it is most natural to start from the classical definition

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (15.48)$$

so that the relevant component for 2D motion is the angular momentum about the z -axis, namely

$$L_z = x p_y - y p_x \quad (15.49)$$

Motivated by the position representation of operators, we replace the classical momentum components by their operator analogs and define

| | | |
|-----------|--------------------|---------|
| classical | quantum mechanical | |
| | | (15.50) |

$$L_z \longrightarrow \hat{L}_z \equiv x \hat{p}_y - y \hat{p}_x = \frac{\hbar}{i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)$$

It is easy to show that \hat{L}_z is Hermitian (P4.21), and that it is also the infinitesimal generator of rotations around the z -axis (P15.13), just as \hat{p}_x is responsible for translations along the x -axis.

To express this more naturally in polar coordinates, we again use the defining relations

$$x = r \cos(\theta) \quad \text{and} \quad y = r \sin(\theta) \quad (15.51)$$

or their inverses

$$r = \sqrt{x^2 + y^2} \quad \text{and} \quad \tan(\theta) = \frac{y}{x} \quad (15.52)$$

and the chain rule to find

$$\frac{\partial}{\partial x} = \sin(\theta) \frac{\partial}{\partial r} + \frac{\cos(\theta)}{r} \frac{\partial}{\partial \theta} \quad (15.53)$$

$$\frac{\partial}{\partial y} = \cos(\theta) \frac{\partial}{\partial r} - \frac{\sin(\theta)}{r} \frac{\partial}{\partial \theta} \quad (15.54)$$

This then gives

$$\hat{L}_z = x\hat{p}_y - y\hat{p}_x = \frac{\hbar}{i} \frac{\partial}{\partial \theta} \quad (15.55)$$

and θ, \hat{L}_z can be seen (P15.13) to have many similarities with the conjugate pair x, \hat{p}_x .

The *eigenfunctions of angular momentum* (in 2D), labeled $\Theta_m(\theta)$, are then determined by the equation

$$\frac{\hbar}{i} \frac{d\Theta_m(\theta)}{d\theta} = \hat{L}_z \Theta_m(\theta) = L_z \Theta_m(\theta) = m\hbar \Theta_m(\theta) \quad (15.56)$$

(operator \hat{L}_z) (eigenvalue L_z)

which yields

$$\Theta_m(\theta) = \frac{e^{im\theta}}{\sqrt{2\pi}} \quad (15.57)$$

where we have chosen to write the dimensionful angular momentum eigenvalue, L_z , in terms of the natural unit of \hbar . The normalization constant is chosen so as to satisfy

$$1 = \int_0^{2\pi} d\theta |\Theta_m(\theta)|^2 \quad (15.58)$$

which is natural for an angular wavefunction.

One difference between these solutions and momentum eigenstates arises because of the periodic nature of the variable θ ; this presumably requires us to identify coordinates separated by $\theta = 2\pi$ as representing the same physical point,⁴ that is,

$$\Theta_m(\theta + 2\pi) = \Theta_m(\theta) \implies e^{i2\pi m} = 1 \quad (15.59)$$

or $m = 0, \pm 1, \pm 2, \dots$ and the angular momentum is quantized

$$L_z = 0, \pm\hbar, \pm 2\hbar, \dots \quad (15.60)$$

This quantization once again arises because of the need to impose (appropriate) boundary conditions. It also guarantees that eigenfunctions corresponding to different eigenvalues are orthogonal, namely

$$\langle n|m \rangle = \int_0^{2\pi} \Theta_n^*(\theta) \Theta_m(\theta) d\theta = \frac{1}{2\pi} \int_0^{2\pi} (e^{in\theta})^* e^{im\theta} d\theta = \delta_{n,m} \quad (15.61)$$

These complex angular wavefunctions correspond most closely to the plane wave solutions (traveling waves) for momentum; linear combinations can yield $\Theta(\theta) = \sin(m\theta)$ or $\cos(m\theta)$, which are more like standing waves, and which may be more appropriate for some bound state problems, or for visualization purposes.

The appropriate Hamiltonian operator in Cartesian coordinates,

$$\frac{\hat{p}_x^2 + \hat{p}_y^2}{2m} + V(x, y) = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + V(x, y) = -\frac{\hbar^2}{2m} \nabla^2 + V(x, y) \quad (15.62)$$

can be written in polar coordinates by expressing the 2D gradient squared in terms of (r, θ) by extending the chain rule arguments used above. One finds (P15.14) that

$$\frac{\partial^2 \psi(x, y)}{\partial x^2} + \frac{\partial^2 \psi(x, y)}{\partial y^2} = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \psi(r, \theta)}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \psi(r, \theta)}{\partial \theta^2} \quad (15.63)$$

The Hamiltonian operator in polar coordinates can thus be written as

$$\hat{H}_{(r, \theta)} = -\frac{\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} \right) + V(r, \theta) \quad (15.64)$$

where we have expanded the radial derivative $(\partial/\partial r)$ operators. We will also henceforward write the mass as μ to avoid confusion with the angular momentum quantum number m ; this will also be appropriate for two-body

⁴ This is true for the angular momentum associated with the orbital motion of particles; for the case of intrinsic angular momentum (spin), see Section 16.4.

problems where the use of the reduced mass μ is natural. The angular derivative term can be written in the form

$$\hat{T}_\theta = -\frac{\hbar^2}{2\mu r^2} \frac{\partial^2}{\partial \theta^2} = \frac{1}{2\mu r^2} \hat{L}_z^2 \quad (15.65)$$

which is indeed the obvious quantum operator analog of the classical energy of rotation in Eqn. (15.47). The appropriate Schrödinger equation is then

$$\hat{H}_{(r,\theta)} \psi(r, \theta) = E \psi(r, \theta) \quad (15.66)$$

along with the normalization condition for the corresponding probability density,

$$1 = \int_0^\infty r \, dr \int_0^{2\pi} d\theta |\psi(r, \theta)|^2 \quad (15.67)$$

This condition is associated with the fact that the probability of finding the particle simultaneously in the small coordinate intervals $(r, r + dr)$ and $(\theta, \theta + d\theta)$ is

$$d\text{Prob}(r, \theta) = |\psi(r, \theta)|^2 r \, dr \, d\theta \quad (15.68)$$

and we will see that the additional factor of r in the “measure” is important.

The case of central force motion for which the potential has no angular dependence, that is, $V(\mathbf{r}) = V(r)$, is the most important and we note that:

- In this case, $[\hat{H}, \hat{L}_z] = 0$, so that the energy eigenfunctions will also be eigenfunctions of the (planar) angular momentum; this fact also implies that the angular momentum will be a conserved quantity.
- The Schrödinger equation is separable in this case, so we can assume solutions of the form $\psi(r, \theta) = R(r) \Theta_m(\theta)$.

Performing the separation of variables in the Schrödinger equation, we find that

$$\begin{aligned} \frac{r^2}{R(r)} \left\{ -\left(\frac{d^2 R(r)}{dr^2} + \frac{1}{r} \frac{dR(r)}{dr} \right) + \frac{2\mu}{\hbar^2} (V(r) - E) R(r) \right\} &= \frac{1}{\Theta_m(\theta)} \frac{d^2 \Theta_m(\theta)}{d\theta^2} \\ &= -m^2 \end{aligned} \quad (15.69)$$

so that the Schrödinger equation for the radial wavefunction (the quantum analog of Eqn. (15.43)) is

$$-\frac{\hbar^2}{2\mu} \left(\frac{d^2 R(r)}{dr^2} + \frac{1}{r} \frac{dR(r)}{dr} \right) + \left(V(r) + \frac{\hbar^2 m^2}{2\mu r^2} \right) R(r) = E R(r) \quad (15.70)$$

Given the already defined normalization properties of the $\Theta_m(\theta)$, the radial wavefunction must satisfy

$$1 = \int_0^\infty r \, dr |R(r)|^2 \quad (15.71)$$

A simple example of 2D motion for which this formulation is useful is the case of a mass connected to a light, but rigid, rod of length r_0 , free to rotate around the origin; such a system is sometimes called a *rigid rotator* or *rotor*. In this case, the Hamiltonian is simply $\hat{H} = \hat{L}_z^2/2\mu r_0^2$ with eigenfunctions given by the $\Theta_m(\theta)$ and quantized energies given by $E_m = \hbar^2 m^2/2\mu r_0^2$. Solutions corresponding to $\pm m$ have the same quantized energies corresponding, in turn, to the equivalence of clockwise versus counterclockwise motion. The same result can be inferred from the complete radial Schrödinger equation Eqn. (15.70) if we assume that there is no potential ($V(r, \theta) = 0$) and that the radius is fixed (so that spatial derivatives of $R(r)$ vanish). We turn to less trivial examples of rotational motion in the next section.

15.3 Quantum Systems with Circular Symmetry

15.3.1 Free Particle

The Schrödinger equation for a free particle in polar coordinates reads

$$-\frac{\hbar^2}{2\mu} \left(\frac{d^2 R(r)}{dr^2} + \frac{1}{r} \frac{dR(r)}{dr} \right) + \frac{\hbar^2 m^2}{2\mu r^2} R(r) = ER(r) \quad (15.72)$$

which can be written in terms of the dimensionless variable $z = kr$ (where $k = \sqrt{2\mu E/\hbar^2}$) as

$$\frac{d^2 R(z)}{dz^2} + \frac{1}{z} \frac{dR(z)}{dz} + \left(1 - \frac{m^2}{z^2} \right) R(z) = 0 \quad (15.73)$$

which can be recognized from the mathematical literature as *Bessel's equation* (see Appendix E.4.) Similarly to the case of a free particle in one dimension, it has two linearly independent solutions for each value of m^2 , the so-called regular solution, $J_{|m|}(z)$, standardly labeled *cylindrical Bessel functions* of order $|m|$ (or Bessel functions of the first kind), and the irregular solutions, $Y_{|m|}(z)$, (Neumann functions or Bessel functions of the second kind); we will explore the mathematical properties and physical meaning of these solutions in this section.

We can exhibit the behavior of the solutions for large z by noting that the equation in this limit becomes approximately

$$\frac{d^2 R(z)}{dz^2} = -R(z) \quad (15.74)$$

so that the behavior is oscillatory, that is, $R(z) \rightarrow \sin(z)$, $\cos(z)$ or $\exp(\pm iz)$. We can do better by assuming a solution of the form

$$R(z) \longrightarrow z^\alpha \cos(z) \quad (15.75)$$

where we assume that $\alpha < 0$, and substitution into Eqn. (15.73) implies that (P15.15) the next order term ($z^{\alpha-1}$) also vanishes when $\alpha = -1/2$. These results help justify the well-known asymptotic expansions

$$J_{|m|}(z) \longrightarrow \sqrt{\frac{2}{\pi z}} \cos\left(z - \frac{|m|\pi}{2} - \frac{\pi}{4}\right) [1 + O(1/z^2)] \quad (15.76)$$

$$Y_{|m|}(z) \longrightarrow \sqrt{\frac{2}{\pi z}} \sin\left(z - \frac{|m|\pi}{2} - \frac{\pi}{4}\right) [1 + O(1/z^2)] \quad (15.77)$$

which are valid for $z \gg 0$. This behavior has an immediate physical interpretation as the *probability density times measure* gives

$$\begin{aligned} d\text{Prob}(r, \theta) &= |R(r)\Theta_{|m|}(\theta)|^2 r \, dr \, d\theta \\ &\propto \cos\left(kr - \frac{|m|\pi}{2} - \frac{\pi}{4}\right)^2 dr \, d\theta \end{aligned} \quad (15.78)$$

implying that, in a spatially averaged sense, there is a uniform distribution of probability corresponding to constant speed motion everywhere in the plane; compare this to the case of the 2D free particle in Cartesian coordinates (P15.1) and the corresponding probability distribution. This can also be contrasted with the 1D case of the unstable harmonic oscillator (Section 9.5) where $\psi(x) \propto 1/\sqrt{x}$ as well, but which there corresponded to an (exponentially) accelerating particle. This is a reminder of the importance of the coordinate measure in the implementation of a probability interpretation. In Fig. 15.6, we plot $|J_0(kr)|^2$, both with and without the extra factor of r , to show the effect.

A similar analysis can be used to examine the $r \rightarrow 0$ behavior and we assume a series solution of the form

$$R(z) \rightarrow \sum_{s=\beta}^{\infty} a_s z^s = a_\beta z^\beta + a_{\beta+1} z^{\beta+1} + \dots \quad (15.79)$$

with β to be determined. Once again, substitution into the differential equation Eqn. (15.73) yields the condition

$$\beta^2 = m^2 \quad \text{or} \quad \beta = \pm|m| \quad (15.80)$$

Figure 15.6. Plots of $|J_0(z)|^2$ (solid) and $z|J_0(z)|^2$ (dotted) versus z showing the effect of the “measure” for the free-particle wavefunction in two dimensions in polar coordinates. The horizontal dashed line corresponds to a probability distribution for constant speed motion in the plane.

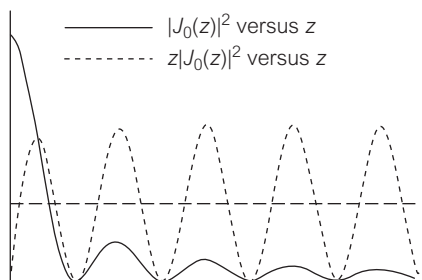
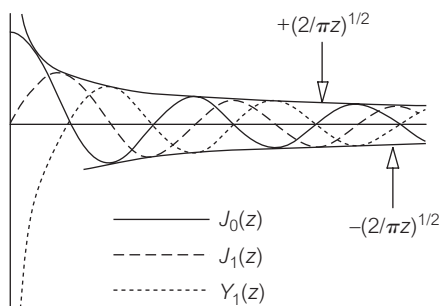


Figure 15.7. Plots of the regular $J_{0,1}(z)$ and irregular $Y_1(z)$ solutions of Bessel’s equation, showing the small and large z behavior. The behavior for large z is consistent with Eqns (15.76) and (15.77).



The regular, that is, well-behaved at the origin, Bessel functions are conventionally chosen to have $\beta = +|m|$, while the ill-behaved Neumann functions are described by $\beta = -|m|$ near the origin.⁵ (The behavior of the two functions is somewhat similar to the exponentially growing and decaying solutions found in tunneling problems, the rotational kinetic term $\hbar^2 m^2 / 2\mu r^2$ playing the role of an “angular momentum barrier” in this case; see P15.16.)

Because of its divergence at the origin, it is often necessary to exclude this solution by hand, that is, use the freedom to pick its coefficient in the most general solution to vanish. We plot $J_{0,1}(z)$ and $Y_1(z)$ in Fig. 15.7 for illustration.

The small r behavior of the Bessel functions solution is also intuitively physical. The probability of being ‘near’ the origin when the particle is in a state of angular momentum $\pm m\hbar$ is given by

$$d\text{Prob} \propto (kr)^{2|m|+1} \quad \text{when } r \rightarrow 0 \quad (15.81)$$

where the additional factor of r comes from the measure. This suppression can be understood as arising from the centrifugal barrier term in Eqn. (15.65), which demands a large cost in energy to be near the origin for rotating particles. The behavior of the first few Bessel functions, $J_m(x)$, for $m = 0, 1, 2, 3$ for small

⁵ The behavior of $Y_0(z)$ near the origin requires special treatment as $m = 0$ in that case corresponds to a logarithmic behavior; specifically $Y_0(z) \rightarrow 2J_0(z)\log(\gamma z/2)/\pi$.

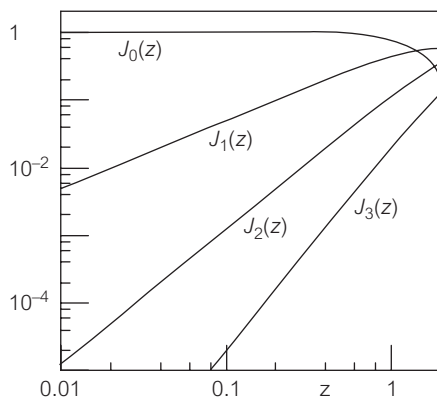


Figure 15.8. Plots of $J_m(z)$ versus z , on logarithmic scales, showing the power-law behavior of $J_m(z) \propto z^m$ for small z , arising from the centrifugal barrier.

argument is shown in Fig. 15.8 for illustration and the plot on semilog paper demonstrates the increasingly large power law behavior near the origin.

15.3.2 Circular Infinite Well

A simple use of the free-particle wavefunctions arises in the study of the infinite circular well,⁶ defined by the potential

$$V(r) = \begin{cases} 0 & \text{for } r < R \\ \infty & \text{for } r > R \end{cases} \quad (15.82)$$

Inside the well, where the particle is free, the solutions are

$$\psi(r, \theta) = J_{|m|}(kr) \Theta_m(\theta) \quad (15.83)$$

or, perhaps more appropriately for bound states,

$$J_{|m|}(kr) \sin(m\theta) \quad \text{and} \quad J_{|m|}(kr) \cos(m\theta) \quad (15.84)$$

where we have excluded the irregular $Y_{|m|}(kr)$ solutions for the reasons discussed above. The boundary conditions at the edge of the well are satisfied for all values of θ provided that $J_m(kR) = 0$. If we label the n th zero of the m th Bessel function by $a_{(n,m)}$, we can see that the corresponding radial wavefunction will have $n_r = n - 1$ radial nodes; we note yet again that imposition of the the boundary conditions has determined the quantized energies, in this case giving

$$E_{n,m} = \frac{\hbar^2 k_{(n,m)}^2}{2\mu} = \frac{\hbar^2 a_{(n,m)}^2}{2\mu R^2} \quad (15.85)$$

⁶ For a discussion of the visualization of the solutions for this problem in both quantum and classical mechanics, see Robinett (1996a).

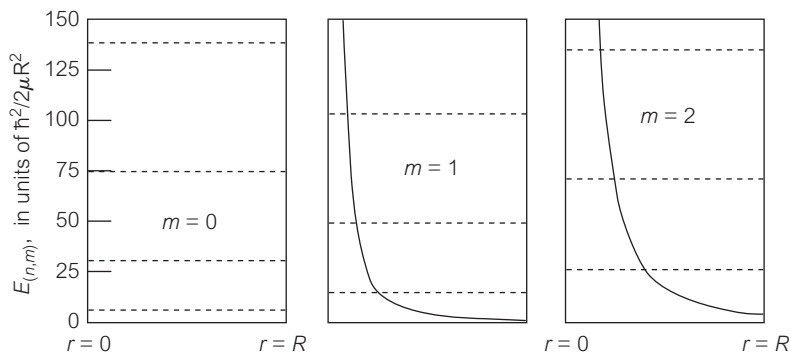


Figure 15.9. Part of the energy spectrum for the infinite circular well obtained using the Bessel function zeros in Eqn. (15.86). The solid curves correspond to the rotational kinetic energy (or centrifugal barrier) term, $\hbar^2 m^2 / 2\mu r^2$, in the Schrödinger equation. Note the resulting increase in energy of corresponding levels as m is increased.

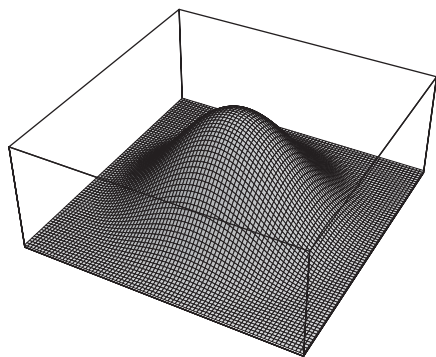


Figure 15.10. Plot of $|\psi(r, \theta)|^2$ for the ground state of the circular infinite well with $(n_r, m) = (0, 0)$.

The notation n_r is motivated by the fact that it counts the number of nodes in the radial wavefunction. Some of the lowest-lying zeroes are given by

$$\begin{array}{llll}
 m = 0: & 2.40483 & 5.52008 & 8.65373 \quad \dots \\
 m = 1: & 3.83171 & 7.01559 & 10.1735 \quad \dots \\
 m = 2: & 5.13562 & 8.41724 & \dots \\
 m = 3: & 6.38016 & 9.76102 & \dots
 \end{array} \tag{15.86}$$

and part of the resulting energy spectrum is shown in Fig. 15.9. Each state with $m \neq 0$ state is doubly degenerate because the two values of $\pm|m|$, corresponding to rotations in opposite senses, give the same energy.

To see connections to both wave physics and classical particle motion, we plot, in Figs 15.10–15.12, $|\psi(r, \theta)|^2$ for several cases:

- In Fig. 15.10, we show the ground state corresponding to $m = 0$ and the first radial zero, that is $a_{(0,0)} = 2.404$, with no rotational kinetic energy and the

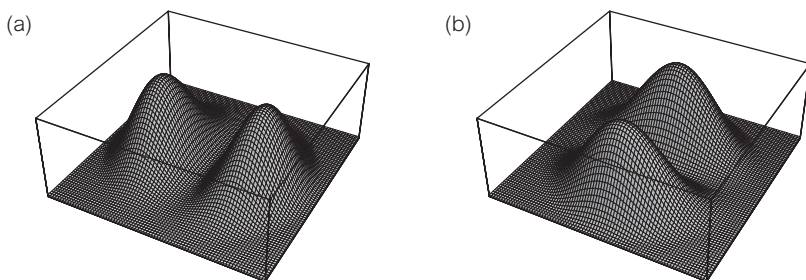


Figure 15.11. Plot of $|\psi(r, \theta)|^2$ for the lowest-lying $m = 1$ states; both the $\sin(\theta)$ (a) and $\cos(\theta)$ (b) cases are plotted to help visualize the double degeneracy.

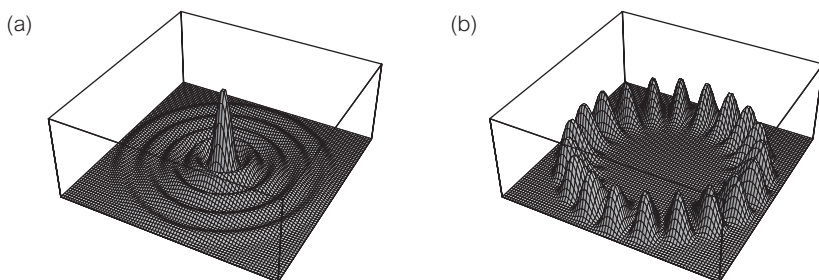


Figure 15.12. Plot of $|\psi(r, \theta)|^2$ for “radial” and “angular” states: Case (a) corresponds to a radially excited state with $(n_r, m) = (4, 0)$, with no angular momentum, while (b) is for $(n_r, m) = (0, 10)$ with large angular momentum (and hence rotational kinetic energy) and little radial kinetic energy (no radial nodes).

least amount of radial kinetic energy. The similarity to the shape of a circular drumhead is obvious.

- Figures 15.11(a) and (b) show $|\psi(r, \theta)|^2$ for the lowest lying states with $|m| = 1$ ($a_{(0,1)} = 3.8318\dots$), with both the $\cos(\theta)$ and the $\sin(\theta)$ solutions plotted to illustrate their similarity and to help visualize their degeneracy.
- Figure 15.12(a) shows a radially excited state (large n_r) but still with no angular momentum, $(n_r, m) = (4, 0)$, with $a_{(4,0)} = 14.43$. This corresponds to a classical particle bouncing back and forth through the origin in a particle interpretation or a spherically symmetric wave reflecting from the walls. We note that “corrals” of heavy atoms can be constructed which approximate infinite circular potential wells on surfaces and the measured electron densities in such a configuration closely matches these predictions; an example was shown in Fig. 1.3.
- Finally, Fig. 15.12(b) shows a state with large angular momentum ($m = 10$) but with the smallest radial quantum number possible, specifically $a_{(0,10)}$ with

$a_{(0,10)} = 14.48$. This corresponds most closely to a particle in such a well undergoing uniform circular motion and the expected peaking of the quantum wavefunction near the walls (classically the particle would, after all, roll around the inner edge) is apparent. This might be the quantum equivalent of a “roulette wheel.” The total energies of the (4, 0) and (0, 10) states differ by less than 1%, but the distribution between radial and rotational kinetic energy is clearly very different.

It is, in principle, possible to construct localized wave packets and track the quasi-classical ballistic motion of particles bouncing in the circular well,⁷ but we will not consider that here due to the technical complexity.

15.3.3 Isotropic Harmonic Oscillator

We return to the isotropic harmonic oscillator in two-dimensions, defined by the potential

$$V(r) = \frac{1}{2}Kr^2 = \frac{1}{2}\mu\omega^2r^2 \quad (15.87)$$

with the corresponding Schrödinger equation

$$-\frac{\hbar^2}{2\mu} \left(\frac{d^2 R(r)}{dr^2} + \frac{1}{r} \frac{dR(r)}{dr} \right) + \left(\frac{\hbar^2 m^2}{2\mu r^2} + \frac{\mu\omega^2}{2} r^2 \right) R(r) = ER(r) \quad (15.88)$$

A standard change of variables, $r = \rho y$ with $\rho^2 = \hbar/\mu\omega$, reduces this to

$$\frac{d^2 R(y)}{dy^2} + \frac{1}{y} \frac{dR(y)}{dy} + \left(\epsilon - y^2 - \frac{m^2}{y^2} \right) R(y) = 0 \quad (15.89)$$

with the dimensionless energy eigenvalue $\epsilon = 2E/\hbar\omega$. The large y -dependence can be extracted as in Section 9.2.1, while the behavior near the origin is guaranteed to be of the form $y^{|m|}$ from Eqn. (15.80). We are thus led to write

$$R(y) = y^{|m|} e^{-y^2/2} G(y) \quad (15.90)$$

leading to

$$\frac{d^2 G(y)}{dy^2} + \left(\frac{2|m|+1}{y} - 2y \right) \frac{dG(y)}{dy} + (\epsilon - 2 - 2|m|) G(y) = 0 \quad (15.91)$$

A somewhat less obvious change of variables to $z = y^2$ then yields the differential equation

$$\frac{d^2 G(z)}{dz^2} + \frac{dG(z)}{dz} \left(\frac{|m|+1}{z} - 1 \right) + G(z) \left(\frac{\epsilon - 2(|m|+1)}{4z} \right) = 0 \quad (15.92)$$

⁷ See Doncheski *et al.*, (2003) for details.

which, to the mathematically sophisticated can be recognized as *Laguerre's equation* as discussed in Appendix E.7. A series solution of the form $G(z) = \sum_{s=0}^{\infty} b_s z^s$ yields the recursion relation

$$\frac{b_{s+1}}{b_s} = \frac{(s + (|m| + 1)/2 - \epsilon/4)}{(s + 1)(s + |m| + 1)} \longrightarrow \frac{1}{s} \quad (15.93)$$

in the limit of large s implying (again, as in Section 9.2.1) that $G(z) \sim e^z \sim e^{y^2}$, which would be inconsistent with the desired behavior in Eqn. (15.90). Once again we find the series must terminate, yielding a polynomial of finite degree in z . The quantized energies are given in terms of the maximum power of this polynomial, $s_{\max} = n_r$, by the condition

$$\epsilon = 2|m| + 2 + 4s_{\max} \equiv 2|m| + 2 + 4n_r \quad (15.94)$$

and we note that n_r also counts the number of radial nodes of the resulting polynomial. This leads to the quantized energies

$$E_{n_r, m} = \hbar\omega (|m| + 2n_r + 1) \quad (15.95)$$

and the corresponding $G(z)$ are *generalized Laguerre polynomials*, denoted as $L_{n_r}^{(|m|)}(z)$. The first few of these are given here for later use:

$$\begin{aligned} L_0^{(k)}(z) &= 1 \\ L_1^{(k)}(z) &= 1 + k - z \\ L_2^{(k)}(z) &= \frac{1}{2} (2 + 3k + k^2 - 2z(k + 2) + z^2) \end{aligned} \quad (15.96)$$

The resulting constant polynomials in the case of $n_r = 0$ are especially important for the classical limit.

The complete (but unnormalized) solutions in polar coordinates are then given by

$$\psi_{n_r, m}(r, \theta) \propto r^{|m|} e^{-r^2/2\rho^2} L_{n_r}^{(|m|)}(r^2/\rho^2) e^{im\theta}. \quad (15.97)$$

The energy spectrum and degeneracies thusly derived from polar coordinates are shown in Fig. 15.13 and we see that the degeneracies agree with those found using Cartesian coordinates. The wavefunctions for a given energy level in the two different schemes are necessarily linear combinations of each other, which can be shown explicitly in simple cases (P15.24).

A particularly easy classical limit to exhibit in this case is that corresponding to uniform circular motion in which case one would use $|m| \gg 1$ and $n_r = 0$ corresponding to the minimum possible radial kinetic energy. In this case, the

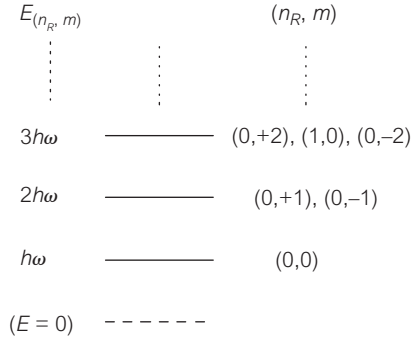


Figure 15.13. Energy spectrum for the 2D harmonic oscillator problem as obtained in polar coordinates showing the same energy level degeneracy as in Fig. 15.5; values of (n_r, m) are shown for each level.

Laguerre polynomials, $L_0^{|m|}(z)$, are constants so the radial probability density is proportional to

$$P(r) = |\psi_{0,m}(r, \theta)|^2 \sim r^{2|m|} e^{-r^2/\rho^2} \quad (15.98)$$

which has a maximum value when

$$0 = \frac{dP(r)}{dr} = \left(2|m|r^{2|m|-1} - \frac{2r^{2|m|+1}}{\rho^2} \right) e^{-r^2/\rho^2} \quad \text{or} \quad r_{\max}^2 = |m|\rho^2. \quad (15.99)$$

Recalling the definition ρ , we find that

$$r_{\max}^2 = \frac{|m|\hbar}{\mu\omega} = \frac{L_z}{\mu\omega} \quad (15.100)$$

The classical circular orbit, of constant radius r_0 , is determined by Eqn. (15.43) where we take $\ddot{r}(t) = 0$ implying that

$$-\mu\omega^2 r_0 = -Kr_0 = F(r_0) = -\frac{L_z^2}{\mu r_0^3} \quad \text{or} \quad r_0^2 = \frac{L_z}{\sqrt{\mu K}} = \frac{L_z}{\mu\omega} \quad (15.101)$$

in agreement with Eqn. (15.100), and with the correspondence principle. Somewhat surprisingly, the form of this Schrödinger equation for a 2D simple harmonic oscillator (SHO) and its solutions are very similar to that for a charged particle in a uniform magnetic field (Section 18.5), which partly motivates our detailed study of it here.

15.4 Questions and Problems

Q15.1. Estimate the zero-point energy and spread in position of an electron bound to a (horizontal) surface because of gravity. Assume for simplicity that the potential

is given by

$$V(z) = \begin{cases} \infty & \text{for } z < 0 \\ m_e g z & \text{for } z > 0 \end{cases} \quad (15.102)$$

What does this potential look like superimposed on Fig. 15.1? Repeat for the case of a neutron in the earth's gravitational potential. Can you imagine how you might see experimental evidence of *Quantum states of neutrons in Earth's gravitational field*⁸?

- Q15.2. What would a plot of $|\phi(p_x, p_y)|^2$ versus (p_x, p_y) look like for the $(n, m) = (1, 1)$ state in the 2D infinite square well? How about for $(n, m) = (10, 10)$ or $(15, 30)$?
- Q15.3. What would $|\phi(p_x, p_y; t)|^2$ look like for the “bouncing 2D wave packet” in Fig. 15.4 as a function of time?
- Q15.4. Recall the wavefunctions in the circular well shown in Fig. 15.12 (a) and (b) corresponding to “radial” and “angular” motion, respectively. For the “radial” case in Fig. 15.12(a), it seems that there is a much larger probability of finding the particle near the origin than elsewhere. Is this consistent with what you know about the corresponding classical “motion”? The momentum space distributions for these two cases can be evaluated (numerically) and the resulting distributions are plotted in Fig. 15.14(a) and (b). Explain why they have the form they do; no numerical calculations are required, simply use your physical intuition, and think about what the momentum vectors in 2D would look like for the two cases of “radial” and “angular” motion. The small “bump” at $(p_x, p_y) = (0, 0)$ for the “radial” case seems to imply that there will be a reasonable chance of finding the particle with vanishing total momentum; is that true?

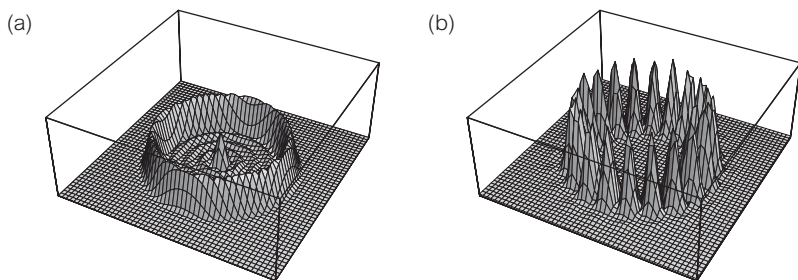


Figure 15.14. Momentum-space probability distributions, $|\phi(p_x, p_y)|^2$ versus p_x, p_y , corresponding to the “radial” (a) and “angular” (b) wavefunctions in the circular infinite well of Fig. 15.12(a) and (b), respectively. Why do they look rather similar?

⁸ See the paper of the same name by Nesvizhevsky *et al.* (2002).

- Q15.5. How would you solve the 2D harmonic oscillator problem using raising and lowering operators? Go as far as you can in generalizing the arguments of Section 13.2.
- Q15.6. Are there any uncertainty relations in two-dimensions for pairs of variables like x, \hat{p}_y ? Can you find an example of a state which has $\Delta x \Delta p_y = 0$ or arbitrarily small?
- Q15.7. What are the appropriate commutation relations for θ and \hat{L}_z ? Is there an associated uncertainty principle? Is there any problem⁹ due to the fact that θ is only defined up to multiples of 2π , so that $\Delta\theta$ cannot be arbitrarily large?
- Q15.8. “*Can you hear the shape of a drum*”¹⁰? The solution of the 2D Schrödinger equation for infinite wall boundaries of various shapes (“footprints”) has many similarities with finding the allowed frequencies of vibration of 2D drum-heads of the same shape, namely, solving the wave equation, with vanishing amplitudes on the variously shaped edges. To what extent do you think that knowing the “spectrum” of allowed energy eigenvalues (or allowed vibratory modes) allows you to determine the shape of the boundary?
- P15.1. (a) Find the plane wave solutions of the time-independent free-particle Schrödinger equation in two-dimensions using Cartesian components. Show that they can be written in the form

$$\psi(\mathbf{r}; t) = e^{i(k_x x - k_y y - \omega t)} = e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (15.103)$$

and find the dispersion relation relating \mathbf{k} and ω .

- (b) Explicitly construct a localized Gaussian wave packet with central momentum value $\mathbf{p}_0 = (p_0^x, p_0^y)$ with initial central position $\mathbf{r}_0 = (x_0, y_0)$. Calculate $\langle x \rangle_t$ and $\langle y \rangle_t$ for this state.

- P15.2. Consider a 2D potential given by

$$V(x, y) = \begin{cases} 0 & \text{for } y < 0 \\ V_0 > 0 & \text{for } y > 0 \end{cases} \quad (15.104)$$

which is a 2D step-up potential.

- (a) To examine plane wave scattering from such a step, consider a solution of the form

$$\psi(\mathbf{r}; t) = \begin{cases} Ie^{i(\mathbf{k}_1 \cdot \mathbf{r} - \omega t)} + Re^{i(\mathbf{k}'_1 \cdot \mathbf{r} - \omega t)} & \text{for } y < 0 \\ Te^{i(\mathbf{k}_2 \cdot \mathbf{r} - \omega t)} & \text{for } y > 0 \end{cases} \quad (15.105)$$

where the wavevectors $\mathbf{k}_1, \mathbf{k}'_1, \mathbf{k}_2$ are defined in Fig. 15.15. Match the wavefunction along the $z = 0$ boundary to find a relation between θ_1, θ'_1 , and θ_2 and compare to Snell’s law of refraction.

⁹ See, for example, Roy and Sannigrahi (1979).

¹⁰ Which is the appropriately speculative title of a famous article by M. Kac (1966).

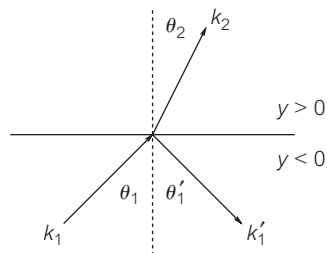


Figure 15.15. Wave numbers and reflection–refraction angles for 2D step-up potential.

- (b) Match the derivatives of the wavefunction at the boundary (which direction?) to determine the reflection and transmission probabilities, namely, $|R/I|^2$ and $|T/I|^2$. Show that your results reduce to those in Section 11.2 for normal incidence, namely, $\theta_1 = 0$.
- (c) For a given angle of incidence, θ , what is the minimum incident energy below which all of the incident particles will be reflected.
- (d) Discuss how the notion of probability flux in Eqn. (4.32) can be generalized to two dimensions and how conservation of flux is realized in this problem.

P15.3. (a) Find the plane wave solutions to the Schrödinger equation for the 2D potential

$$V(x, y) = \begin{cases} 0 & \text{for } x > 0 \text{ and } y > 0 \\ +\infty & \text{otherwise} \end{cases} \quad (15.106)$$

which is like a “corner (90°) reflector.” What would you expect for the behavior of a wave packet incident on such a potential from various angles and for various incident energies?

- (b) Can you construct “mirror” or “image” type solutions, by analogy with those discussed in Section 3.3 for one dimension? How many “images” do you need? Hint: Use your intuition from optics.
- (c) Are there angles besides 90° between the two infinite wall barriers for which “mirror” or “image” solutions are easily obtained? For example, how about 45° or 60° ?

P15.4. Wave packet for projectile motion:

- (a) Write down the Schrödinger equation describing a particle moving in a vertical plane subject to a constant downward gravitational force, and show that it is separable.
- (b) Use previously obtained results for the free particle (Section 3.2.2) and uniformly accelerating (Section 4.7.2) wave packets to write down a wave packet solution, $\psi_{WP}(x, y; t)$, for this problem.
- (c) Evaluate $\langle x \rangle_t$ and $\langle y \rangle_t$ for the wave packet.

P15.5. Consider a wavefunction in the 2D infinite square box of Section 15.1.1 given at $t = 0$ by

$$\psi(x, y; 0) = Nx(L - x)y(L - y) \quad (15.107)$$

- (a) Find N such that $\psi(x, y, 0)$ is normalized.
- (b) What is the probability that a measurement of the energy of a particle described by this state would yield the ground state energy of the system at time $t = 0$? What is this probability at later times?

P15.6. Show that the exchange operator $\hat{E}_{x,y}$ in Eqn. (15.22) is Hermitian and that its eigenvalues are ± 1 .

P15.7. Accidental degeneracies?

- (a) Find the energy eigenvalues for a particle in an infinite *rectangular* well with sides of lengths $L_1 \neq L_2$. Show that, in general, that no degenerate energy levels.
- (b) Show that if L_1 and L_2 are commensurate, that is, if $L_1/L_2 = p/q$ is a ratio of integers, that two different levels characterized by pairs of integers (n_1, n_2) and $(pn_2/q, qn_1/p)$ can be degenerate. Show that an example is when $L_1 = 2L_2$ and the pairs $(4, 1)$ and $(2, 2)$ give rise to degenerate energy states. This phenomenon is often called “accidental degeneracy” as it is not due to any obvious symmetry. (Exchange symmetry is not an obviously useful idea for this asymmetric box.)
- (c) For the special cases discussed in (b), consider a “bigger” square infinite well of size $L = qL_1 = pL_2$ on a side and show that the original box “fits into” the lower left-hand corner. Show that the degenerate wavefunctions in the original box, when extended to the larger box, are simply the standard degenerate pairs discussed in the text. This is illustrated in Fig. 15.16 for the explicit example in part (b).

P15.8. Isosceles triangle infinite well. Imagine the 2D infinite well defined by Eqn. (15.15) cut in half diagonally by the inclusion of another infinite wall along the $x = y$ direction, as in Fig. 15.17. The lower half of the potential is now an infinite well, but with an isosceles triangle ($45^\circ - 45^\circ - 90^\circ$) footprint.

- (a) Show that one of the linear combinations solutions in Eqn. (15.26) not only satisfies the Schrödinger equation for the new well, but also satisfies all of the relevant boundary conditions.
- (b) What is the energy eigenvalue spectrum for this shape? Evaluate the 30 lowest-lying energy levels. Do you find any degeneracies?

P15.9. Show that there are N different states of the 2D SHO which have energy $(N + 1)\hbar\omega$, that is, calculate the degeneracy of each level. Use the solution in Cartesian components.

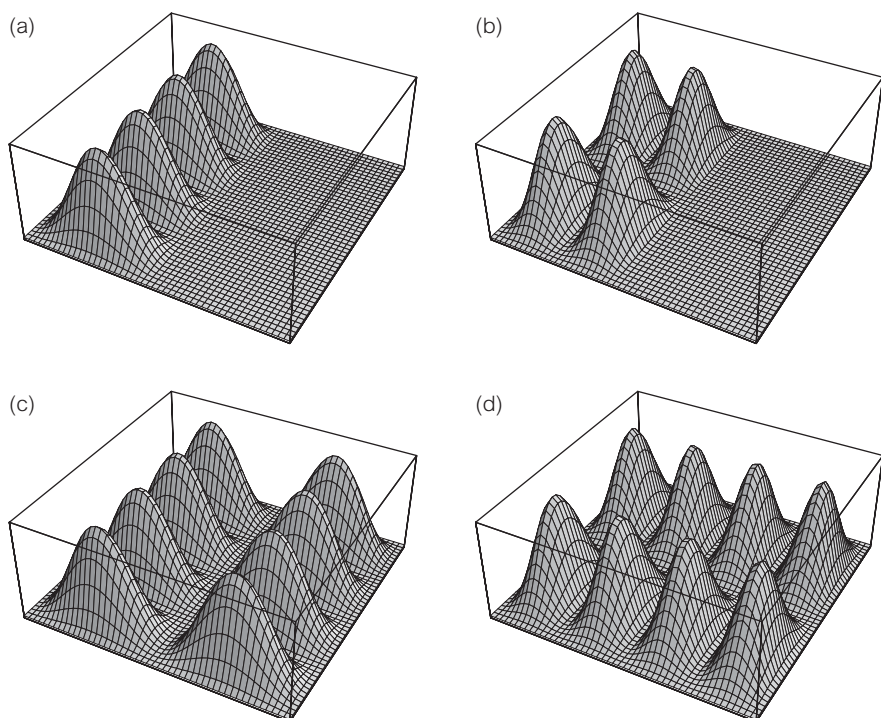


Figure 15.16. Accidental degeneracies in 2D rectangular boxes. Cases (a) and (b) correspond to the degenerate $(n_x, n_y) = (4, 1)$ and $(2, 2)$ levels in a box with $L_x = 2L_y$; (c) and (d) then correspond to the same levels in the related “extended” square box with $L = L_x = 2L_y$ where the levels are “naturally” degenerate.

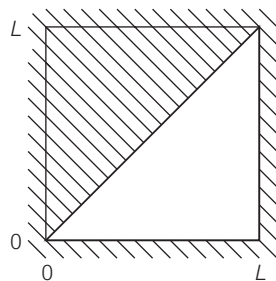


Figure 15.17. Isosceles (45° – 45° – 90°) infinite well footprint made by cutting a 2D square well in half along a diagonal.

- P15.10.** Evaluate the expectation values of x , y , \hat{p}_x , and \hat{p}_y in any energy eigenstate of the 2D SHO of the form in Eqn. (15.28). Show that the expectation value of \hat{L}_z also vanishes in any such state. How then can we have states of the 2D SHO with definite nonzero values of quantized angular momentum?
- P15.11.** Investigate possible accidental degeneracies in the energy spectrum of the non-isotropic 2D harmonic oscillator. What happens, for example, when $k_x = 4k_y$?

If you find any degeneracies, can you find any similarities with the discussion of P15.7?

P15.12. 2D harmonic oscillator wave packets.

- Using the explicit 1D SHO wave packets in Section 12.6.2, write down the probability distribution for a wave packet undergoing uniform circular motion in the isotropic 2D harmonic oscillator potential. Hint: Choose a wave packet for the x coordinate with appropriate initial position, but vanishing initial momentum and oppositely for the y coordinate packet.
- Calculate $\langle x \rangle_t$, $\langle y \rangle_t$, $\langle \hat{p}_x \rangle_t$, and $\langle \hat{p}_y \rangle_t$ for this state and show that they behave as expected.
- Calculate the expectation value of the angular momentum $\langle \hat{L}_z \rangle$ and show explicitly that it is conserved.
- Show that you can write the probability density for a wave packet representing counterclockwise motion in the form

$$|\psi(r, \theta; t)|^2 = \frac{1}{\pi L^2(t)} \exp\left(-(r^2 - 2rx_0 \cos(\theta - \omega t) + x_0^2)\right) \quad (15.108)$$

- Repeat parts (a)–(c) for a wave packet representing a more general elliptical classical path.
- Repeat parts (a)–(c) for a wave packet under the influence of a “non-isotropic” spring of the form Eqn. (15.31).

P15.13. Angular momentum operator.

- Show that

$$e^{i\alpha \hat{L}_z} f(\theta) = f(\theta + \alpha) \quad (15.109)$$

- Calculate $[\hat{L}_z, \theta]$.

P15.14. Using the defining relations Eqn. (15.51) or (15.52), derive Eqns (15.53), (15.54), and (15.55) and show that

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} \quad (15.110)$$

P15.15. Substitute the trial solution Eqn. (15.75) into Bessel’s equation and show that $\alpha = -1/2$ gives the next to leading behavior for large z .

P15.16. The *long-distance* (typically exponentially suppressed) behavior of quantum wavefunctions was discussed in Section 8.2.2. Use the same ideas, but with the rotational kinetic energy term, to derive the *short-distance* behavior of the solutions of the free-particle Schrödinger equation with circular symmetry. Specifically, assume that

$$\psi(r) \sim \exp\left(\pm \sqrt{\frac{2\mu}{\hbar^2}} \int^r \sqrt{V(r)} dr\right) \quad \text{where} \quad V(r) = \frac{m^2 \hbar^2}{2\mu r^2} \quad (15.111)$$

and show that the suppression due to “tunneling” into the angular momentum barrier becomes a power law behavior.

P15.17. Classical probability distributions for the circular infinite well.

- (a) Use the methods outlined in Section 5.1 to derive the classical radial probability distribution for a particle in the circular infinite well. Hint: use the energy relation in the form

$$E = \frac{1}{2}\mu\dot{r}^2 + \frac{L^2}{2\mu r^2} \quad (15.112)$$

to find an expression relating dr and dt , giving $P_{cl}(r) dr \propto dt$.

- (b) Show that your result can be written in the form

$$P_{CL}(r) = \frac{r}{\sqrt{R^2 - R_{\min}^2} \sqrt{r^2 - R_{\min}^2}} \quad (15.113)$$

where $R_{\min} = \sqrt{L^2/2\mu E}$ is the distance of closest approach. Discuss the limiting cases of purely radial and purely angular motion.

- (c) Calculate $\langle r \rangle$ and Δr for discuss their behavior in the same limiting cases.

P15.18. Consider the wavefunction in the infinite circular well of radius a given by

$$\psi(r, \theta) = N(a - r) \sin^2(\theta) \quad (15.114)$$

- (a) Find N such that ψ is properly normalized.
- (b) What is the probability that a measurement of L_z in this state will yield $0\hbar$?; $\pm 1\hbar$?; $\pm 2\hbar$?; any other value? You should be able to obtain definite numerical answers for this part. Hint: What is the appropriate expansion theorem?
- (c) What is the probability that a measurement of the position of the particle will find it in the inner half of the circle, that is, with $r < a/2$?
- (d) What is the probability that a measurement of the energy finds this particle to be in the ground state of the well? Your answer will be in terms of integrals with Bessel functions which you do not need to evaluate numerically.

P15.19. For the infinite circular well of radius R , the angular wavefunctions, $\Theta_m(\theta)$, ensure that eigenstates with different values of m will be orthogonal. For a given value of m , show that one must have

$$\int_0^R dr r J_m(k_{(n_1, m)} r) J_m(k_{(n_2, m)} r) = 0 \quad \text{if} \quad n_1 \neq n_2 \quad (15.115)$$

If you have access to and expertise with an all-purpose computer mathematics package such as *Mathematica*®, confirm this by numerical integration for several cases if you can.

P15.20. Centrifugal force in the circular infinite well. Classically, a particle undergoing uniform circular motion would require a force given by

$$F_c(R) = \frac{\mu v^2}{R} = \mu \omega^2 R = \frac{\mu R^2 \omega^2}{R} = \frac{2T_{\text{rot}}}{R} \quad (15.116)$$

to keep it in motion. Calculate the force exerted by the rapidly spinning ball in the infinite well problem by considering the change in energy when the wall is slowly moved outward a small amount dR , that is,

$$dW = \frac{\hbar^2 a_{(n_r, m)}^2}{2\mu R^2} - \frac{\hbar^2 a_{(n_r, m)}^2}{2\mu (R + dR)^2} = F \cdot dR \quad (15.117)$$

Show that the force exerted by the wall on the particle is consistent with the classical result.

P15.21. Variational calculation for the circular infinite well.

- (a) Make a variational estimate of the ground state energy of the circular infinite well with radius a . Since this state will necessarily have $m = 0$, this amounts to evaluating the energy functional

$$E[\psi] = \frac{\langle \psi | \hat{H}_r | \psi \rangle}{\langle \psi | \psi \rangle} \quad (15.118)$$

where

$$\hat{H}_r = -\frac{\hbar^2}{2\mu} \left(\frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} \right) \quad (15.119)$$

Assume a trial wavefunction of the form $\psi(r, \theta) = R(r) = a^\lambda - r^\lambda$ where λ is used as the variational parameter. Compare your answer to the exact ground state energy obtained from Eqn. (15.85) and the table of Bessel function zeroes in Eqn. (15.87).

- (b) Because the angular wavefunctions $\Theta_m(\theta)$ form an orthogonal set, one can actually make a rigorous variational estimate for the lowest energy state for every value of m (Why?). Recalling the required behavior of the wavefunction near the origin, use a trial wavefunction of the form

$$\psi_{0,m}(r, \theta) = R_{(0,m)}(r) e^{im\theta} \quad \text{where} \quad R_{(0,m)}(r) = (a^\lambda - r^\lambda) r^m \quad (15.120)$$

Show that the variational energy is given by

$$E_{\text{var}}(\lambda; m) = \frac{\hbar^2}{2\mu a^2} \frac{(1+m)(1+\lambda+m)(2+\lambda+2m)}{(\lambda+m)} \quad (15.121)$$

with a minimum value

$$E_{\text{var}}^{\text{min}}(m) = \frac{\hbar^2}{2\mu a^2} \left[\frac{(1+m)(4+2m+(3+m)\sqrt{2+m})}{\sqrt{2+m}} \right] \quad (15.122)$$

Compare this to your answer in part (a) and to the exact answers for $m = 1, 2, 3, 4$ in Eqn. (15.87)

- (c) Sketch the variational radial wavefunctions versus r for increasing values of m and note that the probability is increasingly peaked near the boundary. Using these approximate wavefunctions, find the value of r at which the probability, that is, $r|R_{0,m}(r)|^2$ is peaked, and show that it approaches a as $m \rightarrow \infty$.

P15.22. Semicircular infinite circular well. Consider a particle of mass μ in a “half” infinite circular potential well defined by

$$V(r, \theta) = \begin{cases} 0 & \text{for } 0 < \theta < \pi \text{ and } r < R \\ \infty & \text{otherwise} \end{cases} \quad (15.123)$$

- (a) Find the allowed energies and wavefunctions in terms of those of the “full” infinite circular well. Discuss the degeneracy of each level (if any).
 (b) Show that the angular momentum operator, \hat{L}_z , is still Hermitian, and discuss why.
 (c) Show that \hat{L}_z no longer commutes with the Hamiltonian, so that $\langle \hat{L}_z \rangle_t$ need no longer be constant in time.
 (d) Consider the wavefunction

$$\psi(r, \theta; 0) = AJ_0(k_{(0,1)}r) \sin(\theta) + BJ_0(k_{(0,2)}r) \sin(2\theta) \quad (15.124)$$

Find the wavefunction for later times and evaluate $\langle \hat{L}_z \rangle_t$ and show that it is not constant. Show that any initial wavefunction with only even m or only odd m components will have a constant (and vanishing) angular momentum.

P15.23. 2D annular infinite well. Consider a 2D potential with circular symmetry corresponding to two infinite walls at $r = b, a$, defined via

$$V(r) = \begin{cases} \infty & \text{for } r < b \\ 0 & \text{for } b < r < a \\ \infty & \text{for } r > a \end{cases} \quad (15.125)$$

- (a) Find the allowed solutions and derive the condition which determines the energy eigenvalues for each value of m . Hint: The origin is excluded from the region where solutions are to be considered.
 (b) Estimate the ground state energy when $\delta = a - b \ll a$, that is, the two walls are very close together. Is there any similarity in this limit to a long rectangular potential of dimensions $2\pi a \times \delta$?

P15.24. 2D harmonic oscillator solutions in polar coordinates.

- (a) Confirm that the $L_{n_r}^{(k)}(z)$ given in Eqn. (15.97) actually solve Eqn. (15.92); for the case of $n_r = 0$, show that the solutions are trivial.
- (b) Show that the degeneracy of the level with energy $E = N_s \hbar \omega$ is N_s .
- (c) Compare the wavefunctions for the states with $E = (N+1)\hbar\omega$ in Cartesian and polar coordinates and show explicitly that they can be written as linear combinations of each other for the cases $N = 0, 1, 2$.

P15.25. Consider a particle of mass m moving in the 2D “one-quarter” isotropic harmonic oscillator potential

$$V(r, \theta) = \begin{cases} Kr^2/2 & \text{for } 0 < \theta < \pi/2 \\ \infty & \text{otherwise} \end{cases} \quad (15.126)$$

This corresponds (very roughly) to a particle in two-dimensions, tied to a spring, which is tethered at a square corner.

- (a) What are the allowed energy levels and wavefunctions in such a potential? How do they compare to the “full” harmonic oscillator?

P15.26. Consider a particle of mass m moving in a modified 2D harmonic potential given by

$$V(x, y) = \frac{K}{2} (x^2 + y^2) + \lambda xy \quad (15.127)$$

- (a) Show that the energy eigenvalue spectrum for this potential can be obtained explicitly. Hint: Make a change of variables from x, y to \bar{x}, \bar{y} , using a simple rotation, in order to eliminate the cross-term. Are there conditions on the allowed values of K, λ for this problem to be well-posed?
- (b) Assuming that $\lambda \ll K$, write down the three lowest-energy energy eigenvalues.
- (c) Now, treating the λxy term as a perturbation, and using the standard oscillator problem as the unperturbed system, calculate the first-, second-, and third-order corrections to the *ground state* energy using Eqns (10.112), (10.115), and (10.117), respectively. Hint: Use the matrix element results in Section 9.2.2. Compare your answers to the exact result as a series in λ .
- (d) Finally, treating the next two lowest-lying states using degenerate perturbation theory (why?), find how they split in energy, and compare again to the exact result.

Four

Gravity and Electromagnetism in Quantum Mechanics

In introductory physics courses, we are introduced to (usually at some length) two of the four fundamental forces of nature,¹ namely, gravity and electromagnetism. In this chapter, we discuss some examples of the description of these forces and their effects in quantum mechanical terms.

18.1 Classical Gravity and Quantum Mechanics

The observation that gravity is the most obvious force in the macroscopic world² while electromagnetism is the dominant interaction at the microscopic level of atoms and molecules is easily explained by two facts:

1. The gravitational force between an electron and a proton is 40 orders of magnitude smaller than the electrostatic attraction between them (P18.1), independent of their separation.
2. The electric charges of the electron and proton, while of opposite sign, are experimentally remarkably close in magnitude; the best experimental limits³ imply that $|Q_p + Q_e|/e \leq 10^{-21}$. This implies that normal matter, in bulk, is electrically neutral allowing gravitational interactions to dominate since its interaction strength (i.e. its mass) increases with the size of the system, while the net charge is roughly zero.

¹ The other important interactions, the so-called weak and strong interactions, are of most relevance in the subatomic domain, and are discussed in any good text on subatomic (i.e. nuclear or elementary particle) physics; see, for example, Perkins (1996).

² Everyone, after all, has fallen down.

³ Taken from the Review of Particle Properties (2002).

These facts suggest that it might not be possible to find a terrestrial system in which gravity and quantum mechanics simultaneously play an important and fundamental role.⁴

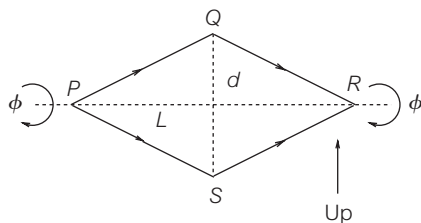
Gravitational effects are occasionally used to novel effect in laboratory atomic systems; one notable example⁵ is the “atomic fountain” which “launches” cesium atoms upward in an apparatus used to make precise frequency measurements, and uses the weak pull of gravity to increase the interaction time. This type of geometry is currently being used as one of the primary atomic frequency standards in the United States; this is, in turn, an integral part of the global position system (SPS). Other electrically neutral particles, most notably neutrons, have been used in a variety of experiments which test both the “classical mechanics” of subatomic particles and the quantum effects of local gravity on the quantum mechanical wavefunction, and that is the main topic of this section.

Beams of reactor neutrons have been measured⁶ to “fall” in the earth’s gravitational field, following parabolic trajectories just like textbook projectiles.⁷ This is purely a classical effect and would be a “wave packet” limit of a particle in a linear potential (as in P15.4).

A much more interesting quantum interference effect was observed using *two* beams of neutrons which were split and then allowed to recombine, as in Fig. 18.1. All such interference experiments rely on the wave phenomenon of the addition of different wave amplitudes, differing by a phase; the electron interference patterns in Fig. 1.2 are an example where the phase difference arises from a difference in path length, and we review this effect in Section 18.8.

In one version of this experiment,⁸ the paths followed by the two neutron beams differ in that they find themselves in regions of different gravitational potential energy. The resulting phase difference is most easily estimated by using

Figure 18.1. Geometry for experiment showing interference effects as a result of the quantum phase of the neutron wavefunction due to the earth’s gravitational potential. The neutrons follow the two paths PQR and PSR , which can be rotated about the horizontal axis PR .



⁴ Recall, however, the compact astrophysical objects discussed in Section 7.4.3.

⁵ See Gibble and Chu (1993).

⁶ See Dabbs *et al.* (1965) who even measure the local acceleration of gravity using neutron fall to be $g_{\text{exp}} = 975.4 \pm 3.1 \text{ cm/s}^2$, compared to the known local value of $g = 979.74 \text{ cm/s}^2$ at the site of the experiment.

⁷ The usual approximation of neglecting air resistance is, in this case, presumably an excellent one!

⁸ See Colella and Overhauser (1980) for details and the original references.

WKB-type wavefunctions (as in Section 10.3.1), that is, we use

$$\psi(x_1) \sim \psi(x_2) \exp\left(i \int_{x_1}^{x_2} p(x) dx / \hbar\right) \quad (18.1)$$

for the phase of the neutron wavefunction. If we consider the gravitational potential, $V(y) = mgy$ (where $y = y(x) = (d/L)x$ along the path) as a small perturbation, we have

$$p(x) = \sqrt{2m(E - V(x))} \approx \sqrt{2mE} \left(1 - \frac{1}{2} \frac{V(x)}{E} + \dots\right) \quad (18.2)$$

since we assume that $V(x) \ll E$. The first term will be the same for both upper and lower arms (provided the path lengths are identical), so that the interference comes from the second term and depends on the path followed through the potential. We then have the resulting total phase increase in going along the upper arm PQR given by

$$\phi_{\text{PQR}} = -\frac{1}{\hbar} \sqrt{\frac{m}{2E}} 2 \int_0^L mgy(x) dx = -\frac{m^2 g}{\hbar p} \frac{dL}{2} \quad (18.3)$$

where we use $E = p^2/2m$. The total difference in phase between the upper (PQR) and lower (PSR) paths can then be written as

$$\Delta\phi = -\frac{m^2 g A \lambda}{2\pi \hbar^2} \quad (18.4)$$

where $A = 2dL$ is the area of the PQRS surface, and we have used the deBroglie relation. Using values which are relevant to the real apparatus, namely, $A \approx (2-3 \text{ cm})^2$, and thermal neutrons ($E \approx k_B T \approx 0.03 \text{ eV}$), we find that $\Delta\phi \approx 28$; this would correspond to $\Delta\phi/2\pi \approx 5$ fringe shifts in a standard interference experiment. To eliminate systematic effects and make the interference pattern more obvious, the neutron counting rate was measured at point R (as this measured the total recombined amplitude squared) as the device was rotated around the axis PR. As the angle is changed, the effective path difference and relative phase is altered; this resulted in the pattern shown in Fig. 18.2, which illustrates the effect. More sophisticated versions of this experimental technique⁹ have even proved sensitive to the earth's rotation, and succeeded in measuring the resulting Coriolis effect on the neutron's path.

⁹ See Staudenmann *et al.* (1980).

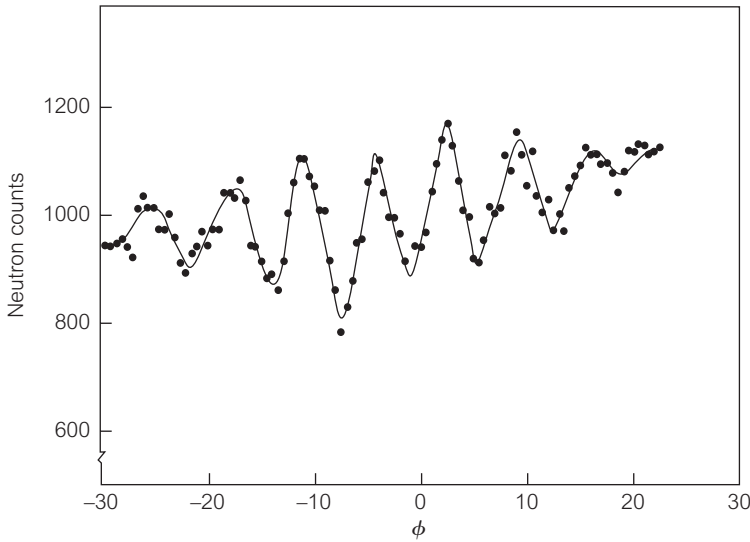


Figure 18.2. Interference pattern from neutron interferometer showing the effect of gravity on the phase of the neutron wavefunction; data are from Colella, Overhauser, and Werner *et al.* (1975).

18.2 Electromagnetic Fields

The classical equations governing the motion of a charged particle moving under the influence of external electric and magnetic fields are given by Newton's law of motion with the Lorentz force law, namely,

$$\begin{array}{ll} \text{Newton's law} & \text{Lorentz force} \\ m\ddot{\mathbf{a}}(t) = \mathbf{F} & = q [\mathbf{E}(\mathbf{r}, t) + \mathbf{v}(t) \times \mathbf{B}(\mathbf{r}, t)] \end{array} \quad (18.5)$$

We will learn below how to incorporate the interactions of electromagnetic (hereafter EM) fields with charged particles into a quantum description of matter, but we first review the classical description of the EM fields themselves.

18.2.1 Classical Electric and Magnetic Fields

One of the major intellectual triumphs of classical physics¹⁰ was the coherent presentation and extension of the then-known laws of electricity and magnetism

¹⁰ Feynman (1963) has said “there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics. The American Civil War will pale into provincial insignificance in comparison with this important scientific event of the same decade.”

by Maxwell. In modern language, these can be written in the following form:

$$\nabla \cdot \mathbf{E}(\mathbf{r}, t) = \frac{1}{\epsilon} \rho(\mathbf{r}, t) \quad \text{Gauss' law} \quad (18.6)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \quad \text{"no-name" law} \quad (18.7)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial}{\partial t} \mathbf{B}(\mathbf{r}, t) \quad \text{Faraday's law} \quad (18.8)$$

$$\nabla \times \mathbf{B}(\mathbf{r}, t) = \mu \mathbf{J}(\mathbf{r}, t) + \mu \epsilon \frac{\partial}{\partial t} \mathbf{E}(\mathbf{r}, t) \quad \text{Ampere's law} \quad (18.9)$$

The electric (\mathbf{E}) and magnetic (\mathbf{B}) fields are determined by the local charge (ρ) and current (\mathbf{J}) densities; as a simple example, recall that Gauss' law implies that the electric field arising from a point electric charge in vacuum is given by

$$\mathbf{E}(\mathbf{r}) = \frac{q}{4\pi\epsilon_0} \frac{\hat{\mathbf{r}}}{r^2} = \frac{Kq}{r^2} \hat{\mathbf{r}} \quad (18.10)$$

giving the familiar inverse-square law for the Coulomb force. The corresponding equation for the magnetic field, Eqn. (18.7), implies that there are no point magnetic charges. We also note that Maxwell's equations are often written in an even more compact form using the auxiliary fields $\mathbf{D} = \epsilon \mathbf{E}$ and $\mathbf{B} = \mu \mathbf{H}$ using the fundamental constants of electricity (the electric permittivity, ϵ) and magnetism (the magnetic permeability, μ).

The important concept of the conservation of electric charge is implicitly contained in Maxwell's equations, and can be recovered by taking $\nabla \cdot$ (Ampere's law) and using Gauss' law to find

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = 0 \quad (18.11)$$

This so-called *equation of continuity* has the identical form as that arising from the conservation of probability in Section 4.2.

One of the most important consequences of Maxwell's equations is the connection with the EM wave equation. One can, for example, take $\nabla \times$ (Ampere's law), and use Eqns (18.7) and (18.8), and the vector identity

$$\nabla \times (\nabla \times \mathbf{B}) = \nabla (\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} \quad (18.12)$$

to show that

$$\frac{\partial^2 \mathbf{B}(\mathbf{r}, t)}{\partial t^2} = \frac{1}{\epsilon_0 \mu_0} \nabla^2 \mathbf{B}(\mathbf{r}, t) = c^2 \nabla^2 \mathbf{B}(\mathbf{r}, t) \quad (18.13)$$

in vacuum and in the absence of charges; the electric field $\mathbf{E}(\mathbf{r}, t)$ satisfies an identical wave equation. The derivation of the connection between the three

fundamental constants $c = 1/\sqrt{\epsilon_0\mu_0}$ was one of the main results of Maxwell's contribution. Plane wave solutions of Eqn. (18.13) of the form

$$\mathbf{B}(\mathbf{r}, t) = \mathbf{B}_0 e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \quad \text{and} \quad \mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \quad (18.14)$$

must satisfy $\omega^2 = |\mathbf{k}|^2 c^2$. They must also, however, be consistent with each of individual Maxwell equations; Faraday's law, for example, implies that

$$\mathbf{k} \times \mathbf{E}_0 = -\omega \mathbf{B}_0 \quad (18.15)$$

while Ampere's law requires that

$$\mathbf{k} \times \mathbf{B}_0 = c^2 \omega \mathbf{E}_0 \quad (18.16)$$

These imply that the electric and magnetic fields in an EM wave must be *transverse*, that is, perpendicular to the direction of propagation determined by \mathbf{k} , and must have magnitudes related by $|\mathbf{B}_0| = |\mathbf{E}_0|/c$.

A final important result describing the flow of energy can be obtained by taking $\mathbf{B} \cdot$ (Faraday's law) $-\mathbf{E} \cdot$ (Ampere's law) (for simplicity, in the case where there is no current density) where one finds

$$\frac{\partial u(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{S}(\mathbf{r}, t) = 0 \quad (18.17)$$

The quantity $u(\mathbf{r}, t)$ is given by

$$u(\mathbf{r}, t) = u_E(\mathbf{r}, t) + u_B(\mathbf{r}, t) = \frac{1}{2} \epsilon |\mathbf{E}(\mathbf{r}, t)|^2 + \frac{1}{2\mu} |\mathbf{B}(\mathbf{r}, t)|^2 \quad (18.18)$$

and can be shown to correspond to the local energy density stored in the electric and magnetic fields. The so-called *Poynting vector*, given by

$$\mathbf{S}(\mathbf{r}, t) = \frac{1}{\mu} \mathbf{E}(\mathbf{r}, t) \times \mathbf{B}(\mathbf{r}, t) \quad (18.19)$$

then describes the rate of energy flow per unit time per unit area; \mathbf{S} also has the units of *intensity*, namely, power per unit area. Equation (18.17) then expresses the manner in which energy “flows” through a system. Using the relativistic connection between energy and momentum, we also note that \mathbf{S}/c has the units of (momentum per unit area per unit time) or (force per area) or pressure and can be used to describe *radiation pressure*. The related quantity, \mathbf{S}/c^2 , has the units of *momentum density* or momentum per unit volume; we can then write

$$\mathbf{P}_{\text{rad}} = \frac{\mathbf{S}}{c} \quad \text{and} \quad \frac{d\mathbf{p}}{dV} = \frac{\mathbf{S}}{c^2} \quad (18.20)$$

Example 18.1. Laser fields

Modern high-power lasers used in atomic physics research can achieve intensities of the order $I = 10^{22} \text{ W/m}^2$. The strength of the electric field component corresponding to this intensity is given by

$$|\mathbf{S}| = \frac{1}{\mu_0} |\mathbf{E}| |\mathbf{B}| = \frac{1}{\mu_0 c} |\mathbf{E}|^2 = I \quad (18.21)$$

so that $|\mathbf{E}| \sim 2 \times 10^{12} \text{ N/C}$; this can also be expressed in somewhat different units as $|\mathbf{E}| \sim 200 \text{ V/\AA}$. A standard value to which to compare this field strength is the magnitude of the electric field felt by the electron in the hydrogen atom, namely, the field arising from a charge e at a distance of roughly $a_0 \sim 0.53 \text{ \AA}$. This “typical” atomic field strength is approximately

$$E_c \equiv \frac{Ke}{a_0^2} = \left(\frac{Ke^2}{\hbar c} \right) \left(\frac{\hbar c}{a_0^2} \right) \frac{1}{e} \approx \left(\frac{1}{137} \right) \left(\frac{1973 \text{ eV} \cdot \text{\AA}}{(0.53 \text{ \AA})^2} \right) \frac{1}{e} \approx 51 \frac{\text{V}}{\text{\AA}} \quad (18.22)$$

Another useful relation arises when one thinks of the laser pulse as an ensemble of photons, each carrying quantized energy $\hbar\omega$ at the speed of light. The Poynting vector can then be written in the form

$$|\mathbf{S}| = n_\gamma \hbar\omega c \quad (18.23)$$

where n_γ is the number density (number per unit volume) of photons. For a laser of the intensity above, operating at $\lambda = 1000 \text{ nm} = 10^4 \text{ \AA}$, corresponding to a photon energy $\hbar\omega \approx 1 \text{ eV}$, the relation Eqn. (18.23) implies that there are roughly $2 \times 10^{22} \text{ } \gamma/\text{m}^3$. An atom in such a laser field can find itself immersed in a relatively dense “photon gas.”

Example 18.2. Uncertainty principle constraints on E and B fields

It is clear from the discussions above that the fact that EM radiation can carry momentum is not a consequence of any particle-like (i.e. photon) interpretation, but arises naturally from classical considerations. We can, however, combine the expression for the momentum density in EM fields,

$$\frac{d\mathbf{p}}{dV} = \frac{1}{\mu_0 c^2} \mathbf{E} \times \mathbf{B} \quad (18.24)$$

with the standard Heisenberg uncertainty principle to “motivate” (not prove) an interesting limit to the measurability of EM field strengths. If we take one component of Eqn. (18.24) and consider the momentum in some small volume, δV , we can write

$$\frac{p_x}{\delta V} \sim (\mathbf{E} \times \mathbf{B})_x = \frac{1}{\mu_0 c^2} (E_y B_z - E_z B_y) \quad (18.25)$$

(Continued)

We can argue that the *uncertainty* in this momentum component will be of the same order as the uncertainties in the values of the field components, so that

$$\left(\delta V \frac{1}{\mu_0 c^2} \Delta E_y \Delta B_z \right) \Delta x \sim \Delta p_x \Delta x \gtrsim \frac{\hbar}{2} \quad (18.26)$$

where the last bound comes from the quantum mechanical uncertainty principle. Taken together, these imply that

$$\Delta E_y \Delta B_z \gtrsim \frac{\hbar}{\mu_0 c^2} \frac{1}{\Delta x \delta V} \quad (18.27)$$

This interesting result can be derived in a somewhat more careful way by examining the ways in which one might attempt to simultaneously measure any two such components of the EM fields¹¹

It suggests that the values of the EM fields become increasingly uncertain as one attempts to measure them on increasingly smaller distance scales; the values of the **E** and **B** fields, and hence the energy contained in the EM field, can be thought of as fluctuating wildly at short distances.

In classical mechanics, one often finds it useful to use the concept of a potential energy function $V(\mathbf{r})$ instead of the Newtonian force $\mathbf{F}(\mathbf{r})$; the connection between the two is $\mathbf{F}(\mathbf{r}) = -\nabla V(\mathbf{r})$. This relation makes it clear that different choices of the “zero of potential” can have no physical meaning as $V(\mathbf{r}) \rightarrow V(\mathbf{r}) + V_0$ yields the same measurable force.

A similar, but more subtle and deep situation arises in electrodynamics where one can express the (physical) electric and magnetic fields in terms of scalar ($\phi(\mathbf{r}, t)$) and vector ($\mathbf{A}(\mathbf{r}, t)$) potentials via

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t) \quad (18.28)$$

$$\mathbf{E}(\mathbf{r}, t) = -\nabla \phi(\mathbf{r}, t) - \frac{\partial}{\partial t} \mathbf{A}(\mathbf{r}, t) \quad (18.29)$$

It is easy to show that the fields produced by the potentials ϕ, \mathbf{A} will be identical to those produced by the new potentials ϕ', \mathbf{A}' provided they are related by the transformations

$$\mathbf{A}'(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}, t) + \nabla f(\mathbf{r}, t) \quad (18.30)$$

$$\phi'(\mathbf{r}, t) = \phi(\mathbf{r}, t) - \frac{\partial}{\partial t} f(\mathbf{r}, t) \quad (18.31)$$

¹¹ See Landé (1951).

where $f(\mathbf{r}, t)$ is an arbitrary scalar function. There are thus an infinite number of different EM potentials which correspond to a given configuration of measurable fields. Such a change in potentials is called a *gauge transformation*, and will be seen to play an important role in the quantum mechanical treatment of charged particle interactions.

Example 18.3. Scalar and vector potentials

We will consider below the familiar cases of charged particles acted on by uniform electric and magnetic fields, and we consider the EM potentials, ϕ and \mathbf{A} , which can describe these cases.

A uniform electric field, given by $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0$, can be obtained from potentials

$$\phi(\mathbf{r}, t) = -\mathbf{E}_0 \cdot \mathbf{r} \quad \text{and} \quad \mathbf{A}(\mathbf{r}, t) = 0 \quad (18.32)$$

which is the standard choice. However, the same field can be described by the potentials

$$\phi'(\mathbf{r}, t) = 0 \quad \text{and} \quad \mathbf{A}'(\mathbf{r}, t) = -\mathbf{E}_0 t \quad (18.33)$$

It is easy to see that the transformation relating these two sets of potentials via Eqns (18.30) and (18.31) is generated by $f(\mathbf{r}, t) = -\mathbf{E}_0 \cdot \mathbf{r}t$.

A uniform magnetic field in the z direction, $\mathbf{B} = B_0 \hat{\mathbf{z}} = (0, 0, B_0)$ can be obtained from the potentials

$$\phi(\mathbf{r}, t) = 0 \quad \text{and} \quad \mathbf{A}(\mathbf{r}, t) = \frac{B_0}{2}(-y, x, 0) \quad (18.34)$$

One of the many other possible choices is

$$\phi'(\mathbf{r}, t) = 0 \quad \text{and} \quad \mathbf{A}'(\mathbf{r}, t) = B_0(-y, 0, 0) \quad (18.35)$$

and these two sets can be seen to be gauge equivalent with $f(\mathbf{r}, t) = -B_0 xy/2$.

18.2.2 E and B Fields in Quantum Mechanics

The standard procedure we have adopted to extend the equations of motion of classical mechanics to the quantum Schrödinger equation (in position space) has been through a generalization of the Hamiltonian formulation of classical mechanics. For the free particle in one dimension this consisted of simply making the identification

$$H_{\text{classical}} = \frac{p^2}{2m} \quad \Longrightarrow \quad \hat{H} = \frac{\hat{p}^2}{2m} \quad (18.36)$$

We show in Appendix G that the classical Hamiltonian function appropriate for a charged particle (of charge q) acted on by external electric and magnetic fields

(now in three dimensions) is given by

$$H_{\text{classical}} = \frac{1}{2m} (\mathbf{p} - q\mathbf{A}(\mathbf{r}, t))^2 + q\phi(\mathbf{r}, t). \quad (18.37)$$

This choice for $H_{\text{classical}}$ reproduces the correct equations of motion in Eqn. (18.5).

The corresponding quantum mechanical Hamiltonian is obtained by replacing the momentum variable by its operator counterpart thereby giving the Schrödinger equation

$$\hat{H}\psi(\mathbf{r}, t) = \hat{E}\psi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) \quad (18.38)$$

where

$$\hat{H} = \frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A}(\mathbf{r}, t))^2 + q\phi(\mathbf{r}, t) \quad (18.39)$$

One must, of course, be careful of the ordering of any differential operators, so we find

$$\begin{aligned} [\hat{\mathbf{p}} - q\mathbf{A}(\mathbf{r}, t)]^2 \psi(\mathbf{r}) &= -\hbar^2 \nabla^2 \psi(\mathbf{r}, t) + iq\hbar \nabla \cdot [\mathbf{A}(\mathbf{r}, t) \psi(\mathbf{r}, t)] \\ &\quad + iq\hbar \mathbf{A}(\mathbf{r}, t) \cdot [\nabla \psi(\mathbf{r}, t)] + q^2 [\mathbf{A}(\mathbf{r}, t) \cdot \mathbf{A}(\mathbf{r}, t)] \psi(\mathbf{r}, t) \end{aligned} \quad (18.40)$$

The classical equations of motion depend on the \mathbf{E} and \mathbf{B} fields themselves, and are obviously invariant under any gauge transformation. The Hamiltonian which now appears in the Schrödinger equation, however, depends on the EM potentials explicitly, and does change its form under such a transformation. Specifically, under the change in potentials given by Eqns (18.30) and (18.31), the original Hamiltonian, Eqn. (18.39), is replaced by

$$\begin{aligned} \hat{H}' &= \frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A}'(\mathbf{r}, t))^2 + q\phi'(\mathbf{r}, t) \\ &= \frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A}(\mathbf{r}, t) - q\nabla f(\mathbf{r}, t))^2 \\ &\quad + q\phi'(\mathbf{r}, t) - q \frac{\partial}{\partial t} f(\mathbf{r}, t) \end{aligned} \quad (18.41)$$

which does not manifestly imply the same physical solutions. It is not hard to show, however, that if one also simultaneously changes the original wavefunction $\psi(\mathbf{r}, t)$ by a (possibly time- and space-dependent) phase factor, namely

$$\psi'(\mathbf{r}, t) = \psi(\mathbf{r}, t) e^{iqf(\mathbf{r}, t)/\hbar} \quad (18.42)$$

then

$$\begin{aligned}\hat{H}\psi(\mathbf{r}, t) &= \hat{E}\psi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) \\ \Downarrow \\ \hat{H}'\psi'(\mathbf{r}, t) &= \hat{E}\psi'(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \psi'(\mathbf{r}, t)\end{aligned}\tag{18.43}$$

The probability densities corresponding to ψ' and ψ are identical because

$$|\psi'(\mathbf{r}, t)|^2 = |\psi(\mathbf{r}, t)|^2\tag{18.44}$$

and the gauge transformation makes no change in the observable physics of the system. (Recall P6.4 where a change in the “zero of potential” is discussed, and a similar result is found, namely, a simple change of phase.) The solutions obtained in different gauges may well look quite different, but must correspond to the same physical energy eigenvalues and be related via Eqn. (18.42); it is often useful to use this freedom of gauge to choose the ϕ and \mathbf{A} which make the problem most tractable.

In the next two sections we deal with several special cases corresponding to uniform electric and magnetic fields where a problem can be solved explicitly in separate gauges, and where the connections between the solutions are easily confirmed; these examples also show that different properties of the solution of a given physical problem may be more apparent in one gauge or another.

18.3 Constant Electric Fields

We first consider the action of a uniform electric field, for simplicity in the $+x$ direction, that is, $\mathbf{E}_0 = E_0\hat{\mathbf{x}}$, on an otherwise free particle of charge q . The classical solutions correspond to free-particle motion in the y, z directions and uniform acceleration in the x direction, namely

$$x(t) = \frac{qE_0}{2m}t^2 + v_{0x}t + x_0\tag{18.45}$$

$$y(t) = v_{0y}t + y_0\tag{18.46}$$

$$z(t) = v_{0z}t + z_0\tag{18.47}$$

The corresponding quantum problem can be defined using the standard gauge choice in Eqn. (18.32), by the Hamiltonian operator

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} - qE_0x = \left(\frac{\hat{p}_x^2}{2m} - eE_0x \right) + \frac{\hat{p}_y^2}{2m} + \frac{\hat{p}_z^2}{2m}\tag{18.48}$$

The system is clearly separable and the wavefunction can be written in the form

$$\psi(\mathbf{r}, t) = \psi(x, t) e^{i(p_y y - p_y^2 t/2m)/\hbar} e^{i(p_z z - p_z^2 t/2m)/\hbar} \quad (18.49)$$

The problem of uniform acceleration in one dimension was discussed in Section 4.7.2 where a Gaussian wave packet solution was found with the form

$$\begin{aligned} \psi(x, t) &= \frac{1}{\sqrt{\beta\sqrt{\pi}(1 + it/t_0)}} e^{iFt(x - Ft^2/6m)/\hbar} \\ &\times \exp\left(-(x - Ft^2/2m)^2/2\beta^2(1 + it/t_0)\right) \end{aligned} \quad (18.50)$$

In Eqn. (18.50), $F = qE_0$, and the initial Gaussian wave packet is obviously

$$\psi(x, 0) = \frac{1}{\sqrt{\beta\sqrt{\pi}}} e^{-x^2/2\beta^2} \quad (18.51)$$

In solving this problem, we have explicitly used the most familiar gauge in which

$$\phi(\mathbf{r}, t) = -Ex \quad \text{and} \quad \mathbf{A}(\mathbf{r}, t) = 0 \quad (18.52)$$

but we can just as well use the gauge equivalent set

$$\phi(\mathbf{r}, t) = 0 \quad \text{and} \quad \hat{A}(\mathbf{r}, t) = -E_0 \hat{x}t \quad (18.53)$$

These two are related by the gauge function $f(x, t) = -E_0 xt$.

In the new gauge, the y and z dependence is unchanged, but the 1D Schrödinger equation for the x behavior now reads

$$\begin{aligned} \hat{H}\psi'(x) &= \left(\hat{H}_0 + t\hat{H}_1 + t^2\hat{H}_2\right)\psi'(x, t) \\ &= \left(\frac{\hat{p}_x^2}{2m} + \frac{qE_0 t}{m}\hat{p}_x + \frac{1}{2m}(qE_0 t)^2\right)\psi'(x, t) \\ &= i\hbar \frac{\partial}{\partial t} \psi'(x, t) \end{aligned} \quad (18.54)$$

where the terms linear in \hat{p}_x simplify because the vector potential does not depend on position. As in Section 12.5, we can formally solve this problem by integration of the initial value problem and find

$$\psi'(x, t) = e^{-i(\hat{H}_0 t + \hat{H}_1 t^2/2 + \hat{H}_2 t^3/3)/\hbar} \psi'(x, 0) \quad (18.55)$$

In this special case, the three component pieces, $\hat{H}_{0,1,2}$ commute with each other so that their order in the exponential is unimportant and the time-development operator can be written in the simpler form

$$\begin{aligned} \psi'(x, t) &= e^{-i(\hat{H}_2 t^3/3)/\hbar} e^{-i(\hat{H}_1 t^2/2)/\hbar} e^{-i(\hat{H}_0 t)/\hbar} \psi'(x, 0) \\ &= e^{-iF^2 t^3/6m\hbar} e^{-iFt^2 \hat{p}_x/2m\hbar} e^{-i\hat{p}_x^2 t/2m\hbar} \psi'(x, 0) \end{aligned} \quad (18.56)$$

and we can consider the effect of each operator in turn. The time-development operator for the free particle, namely, $\hat{H}_0 = \hat{p}_x^2/2m$ has a simple effect when acting on a Gaussian initial state (Example 12.4), that is,

$$\psi'(x, t) = e^{-i\hat{H}_0 t/\hbar} \left(\frac{1}{\sqrt{\beta}\sqrt{\pi}} e^{-x^2/2\beta^2} \right) = \frac{1}{\sqrt{\beta(1+it/t_0)}\sqrt{\pi}} e^{-x^2/2\beta^2(1+it/t_0)} \quad (18.57)$$

corresponding to spreading, but no translation of the central value. The effect of the second operator can be made clear by recalling (P12.6) that the momentum operator generates translations in space via

$$e^{ia\hat{p}/\hbar} \psi(x, t) = \psi(x+a, t) \quad (18.58)$$

so that the spreading wave packet of Eqn. (18.57) is translated via

$$x \longrightarrow x - \frac{qE_0 t^2}{2m} = x - \frac{Ft^2}{2m} \quad (18.59)$$

The final factor, $e^{-i\hat{H}_2 t^3/3\hbar}$, is a simple phase. Taken together, these three operators acting on the initial wavefunction give

$$\psi'(x, t) = \frac{1}{\sqrt{\beta}\sqrt{\pi}(1+it/t_0)} e^{-i(F^2 t^3/6m)/\hbar} e^{-(x-Ft^2/2m)^2/2\beta^2(1+it/t_0)} \quad (18.60)$$

where $F = qE_0$. This explicitly satisfies

$$\psi'(x, t) = e^{-iqf(x,t)/\hbar} \psi(x, t) = e^{-iFxt/\hbar} \psi(x, t) \quad (18.61)$$

as expected from the gauge transformation.

18.4 Atoms in Electric Fields: The Stark Effect

One of the simplest ways to probe and manipulate the energy level structure and wavefunction of an atom is to place it in a constant external electric field. We will consider the quantum version of this problem for hydrogen where the resulting pattern of energy level shifts is called the *Stark effect*; we begin, however, by considering a simplified classical model to gain some intuition.

18.4.1 Classical Case

As a simple model of a (classical) hydrogen atom, consider a positive point charge $+e$, representing the proton, embedded in a uniformly charged spherical

cloud of radius a_0 (the Bohr radius) with total charge $-e$. The resulting charge density is

$$\rho(\mathbf{r}) = \begin{cases} -3e/4\pi a_0^3 & \text{for } r < a_0 \\ 0 & \text{for } r > a_0 \end{cases} \quad (18.62)$$

Gauss's law can then be used to derive (as in P17.23) the corresponding electric field giving

$$\mathbf{E}(\mathbf{r}) = \begin{cases} -Ker/a_0^3 & \text{for } r < a_0 \\ -Ke\hat{\mathbf{r}}/r^2 & \text{for } r > a_0 \end{cases} \quad (18.63)$$

with corresponding potential energy

$$V(\mathbf{r}) = \begin{cases} -Ke^2/2a_0(3 - r^2/a_0^2) & \text{for } r < a_0 \\ -Ke^2/r & \text{for } r > a_0 \end{cases} \quad (18.64)$$

If an external electric field E_0 (say in the z direction) is applied, the electron cloud and “nucleus” will shift in opposite directions by a net amount r_0 until a new equilibrium situation is achieved as in Fig. 18.3. The new system has the separation of charge characteristic of an electric dipole and the “atom” has been *polarized* by the external field. If we assume that the spherical shape of the electron cloud is unchanged for sufficiently small external fields, the net displacement can be determined by balancing the forces on the proton giving

$$0 = F_{\text{external}} + F_{\text{cloud}} = eE_0 - \frac{Ke^2 r_0}{a_0^3} \quad \text{or} \quad r_0 = E_0 a_0^3 / Ke \quad (18.65)$$

The resulting induced *dipole moment* of the system is $p \equiv er_0 = E_0 a_0^3 / K$. The new situation can also be visualized in terms of the potential energy function in

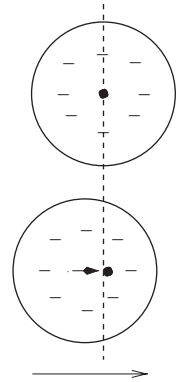


Figure 18.3. Classical picture of polarization of a charge distribution by an external electric field.

Applied electric field induces dipole moment

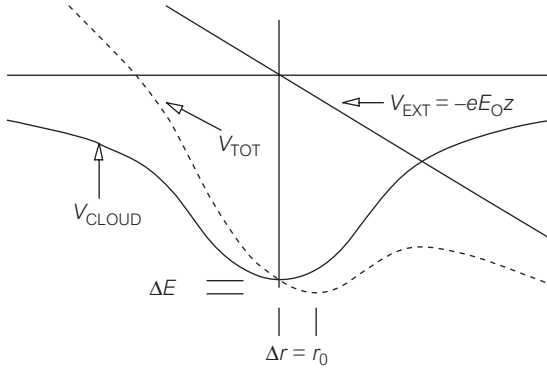


Figure 18.4. Energy shift due to applied electric field.

Fig. 18.4 where the new minimum is apparent. It is easy to show that the new system has a potential energy, which is lower by an amount

$$\Delta V = \Delta E = -\frac{E_0^2 a_0^3}{2K} \quad (18.66)$$

The interaction energy of an electric dipole in an external field is given by $-\mathbf{p} \cdot \mathbf{E}_0$ so an atom with a permanent dipole moment would have an energy shift linear in the applied field. For the situation above, there is only an induced dipole moment so that energy shift is necessarily quadratic in E_0 ; the *electric dipole polarizability* (labeled α) is often defined from the observed energy shift ΔE via

$$\Delta E = -\frac{1}{2}\alpha E_0^2 \quad (18.67)$$

The simple model above then implies that $\alpha \propto a_0^3/K$ with a numerical coefficient which varies with the assumed charge distribution. We note that this energy shift has the dimensions given by product of the energy density in the external field ($u_E = \epsilon|\mathbf{E}|^2/2$) times a typical atomic or molecular volume $V \sim a_0^3$. This result also gives the form for the interaction potential of a charged particle with a neutral atom or molecule, the so-called *polarization potential*, via

$$V_{\text{POL}}(r) \sim -\frac{1}{2}\alpha \left(\frac{Ke}{r^2}\right)^2 \propto (\alpha K)\frac{K^2 e^2}{r^4} \propto \frac{K^2 e^2 a_0^3}{r^4} \quad (18.68)$$

Note that the introduction of the external (linear) potential $V_{\text{ext}} = -eE_0z$ makes the potential “turn over” for sufficiently large values of $|z|$. The existence of a potential minimum even for arbitrarily large values of E_0 guarantees that there will always be a classical bound state; quantum effects, such as tunneling and zero-point energy, will make the problem in real atoms more interesting.

18.4.2 Quantum Stark Effect

We now turn to the quantum description of hydrogen atom in an external field. Being somewhat careful initially, we note that the potential felt by the two charged particles (proton and electron) in an external field in the $+z$ direction is

$$V(\mathbf{r}_e, \mathbf{r}_p) = -eE_0 z_p + eE_0 z_e = eE_0(z_e - z_p) = eE_0 z \quad (18.69)$$

where $z = z_{\text{rel}} = z_e - z_p$ is the relative coordinate. The complete Hamiltonian is now

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2\mu} - \frac{Ke^2}{r} + eE_0 z \quad (18.70)$$

For sufficiently weak fields, we can use perturbation theory¹² to treat the additional term representing the external field; we will consider its effects on both the ground state and first excited states as examples of second-order and degenerate state perturbation theory, respectively.

For the ground state, $\psi_{100}(\mathbf{r})$, the first-order shift in energy is given by

$$E_1^{(1)} = \langle \psi_{100} | eE_0 z | \psi_{100} \rangle = eE_0 \int d\mathbf{r} |\psi_{100}(\mathbf{r})|^2 z = 0 \quad (18.71)$$

from familiar parity arguments. This is also consistent with the classical arguments in which we expect no energy shift linear in the applied field. We must then consider the second-order shift given by

$$E_1^{(2)} = \sum_{n=2}^{\infty} \sum_{l=0}^{n-1} \sum_{m=-l}^{+l} \frac{|\langle \psi_{nlm} | eE_0 z | \psi_{100} \rangle|^2}{E_1^{(0)} - E_n^{(0)}} \quad (18.72)$$

Writing $z = r \cos(\theta)$, we require the matrix elements

$$\langle \psi_{nlm} | z | \psi_{100} \rangle = \int d\mathbf{r} [R_{n,l}(r) Y_{l,m}^*(\theta, \phi)] [r \cos(\theta)] [R_{1,0}(r) Y_{0,0}(\theta, \phi)] \quad (18.73)$$

The angular integration can be easily performed since

$$Y_{0,0} = \frac{1}{\sqrt{4\pi}} \quad \text{while} \quad \cos(\theta) = \sqrt{\frac{4\pi}{3}} Y_{1,0}(\theta, \phi) \quad (18.74)$$

which contributes a factor of

$$\int d\Omega Y_{l,m}^*(\theta, \phi) \frac{1}{\sqrt{3}} Y_{1,0}(\theta, \phi) = \frac{1}{\sqrt{3}} \delta_{l,1} \delta_{m,0} \quad (18.75)$$

¹² It has been said (Condon and Shortley, 1951) that the treatment of the Stark effect in hydrogen was the first application of perturbation theory in quantum mechanics.

when we use the orthonormality properties of the spherical harmonics. This result is in the form of a selection rule. It also simplifies the overlap integrals required for the radial piece of Eqn. (18.73) as we now only require

$$\int_0^\infty dr r^2 R_{n,l}(r) r R_{1,0}(r) \delta_{1,0} \delta_{m,0} = \int_0^\infty dr r^3 R_{n,1}(r) R_{1,0}(r) \quad (18.76)$$

However, the complete evaluation of the second-order shift requires not only the bound states in Eqn. (18.72), but also the continuum ($E_k > 0$) states as well, giving a contribution in schematic form of

$$E_1^{(2)} = \int_{E_k > 0} \frac{|\langle \psi_k | e E_0 z | \psi_{100} \rangle|^2}{E_1^{(0)} - E_k} dk \quad (18.77)$$

where $\psi_k(\mathbf{r})$ denotes the unbound (scattering) states.¹³ This complicates the explicit evaluation of the second-order shift, but luckily other methods, including the use of parabolic coordinates¹⁴ (in which the the problem is still separable) or subtle tricks¹⁵ exist which give the complete result in closed form as

$$E_1^{(2)} = -\frac{9}{4} \frac{E_0^2 a_0^3}{K} \quad (18.78)$$

consistent with our classical arguments.

For the first excited state of hydrogen, corresponding to $n = 2$, we have four states which are degenerate in energy in the absence of an external field,

$$n = 2, l = 0 \quad \psi_{200} = \psi_A \quad (18.79)$$

and

$$n = 2, l = 1, m = \begin{cases} +1 & \psi_{2,1,+1} \equiv \psi_B \\ 0 & \psi_{2,1,0} \equiv \psi_C \\ -1 & \psi_{2,1,-1} \equiv \psi_D \end{cases} \quad (18.80)$$

Using the formalism of degenerate state perturbation theory (Section 10.5.2), we then are required, in principle, to diagonalize the 4×4 matrix corresponding

¹³ See Ruffa (1973) for an explicit calculation of just this contribution.

¹⁴ See, for example, Bethe and Salpeter (1957).

¹⁵ A number of textbooks discuss the method due to Dalgarno and Lewis (1955).

to the equation

$$\begin{pmatrix} E_2^{(0)} + \langle V \rangle_{AA} & \langle V \rangle_{AB} & \langle V \rangle_{AC} & \langle V \rangle_{AD} \\ \langle V \rangle_{BA} & E_2^{(0)} + \langle V \rangle_{BB} & \langle V \rangle_{BC} & \langle V \rangle_{BD} \\ \langle V \rangle_{CA} & \langle V \rangle_{CB} & E_2^{(0)} + \langle V \rangle_{CC} & \langle V \rangle_{CD} \\ \langle V \rangle_{DA} & \langle V \rangle_{DB} & \langle V \rangle_{DC} & E_2^{(0)} + \langle V \rangle_{DD} \end{pmatrix} \begin{pmatrix} \psi_A \\ \psi_B \\ \psi_C \\ \psi_D \end{pmatrix} = E_2^{(1)} \begin{pmatrix} \psi_A \\ \psi_B \\ \psi_C \\ \psi_D \end{pmatrix} \quad (18.81)$$

where $V = eE_0 z$. The problem simplifies considerably since the matrix elements connecting $\psi_B = \psi_{2,1,+1}$ to the other states through this interaction vanish since

$$\langle \psi_B | z | \psi_{A,C,D} \rangle = 0 \text{ because they have different } m \text{ values} \quad (18.82)$$

while

$$\langle \psi_B | z | \psi_B \rangle = 0 \text{ because of parity} \quad (18.83)$$

The same is true of the matrix elements for $\psi_D = \psi_{2,1,-1}$; these two states effectively decouple from the problem, and their energy levels are unchanged. That leaves the 2×2 subspace corresponding to $\psi_{A,C}$. The matrix elements connecting these two also simplify as

$$\langle \psi_A | z | \psi_A \rangle = 0 = \langle \psi_C | z | \psi_C \rangle \quad (18.84)$$

by parity, while

$$\langle \psi_A | z | \psi_C \rangle = -3a_0 \quad (18.85)$$

by direct calculation (P18.8). So, in the $\psi_{A,C}$ subspace, we have the determinant

$$\det \begin{pmatrix} E_2^{(0)} - E_2^{(1)} & -3eE_0 a_0 \\ -3eE_0 a_0 & E_2^{(0)} - E_2^{(1)} \end{pmatrix} = 0 \quad (18.86)$$

or

$$E_2^{(1)} = E_2^{(0)} \pm 3eE_0 a_0 \quad (18.87)$$

and the energies of the $\psi_{2,0,0}$ and $\psi_{2,1,0}$ states are split as shown in Fig. 18.5. The two solutions labeled by \pm correspond to the (normalized) wavefunctions

$$\psi^{(+)} = \frac{1}{\sqrt{2}} (\psi_{2,0,0} - \psi_{2,1,0}) \quad \text{and} \quad \psi^{(-)} = \frac{1}{\sqrt{2}} (\psi_{2,0,0} + \psi_{2,1,0}) \quad (18.88)$$

Figure 18.5. $V(z) = eE_0z$ times $|\psi^\pm(0, 0, z)|^2$ for the linear combination states shifted up, (+), and down, (−), in energy in the linear Stark effect.

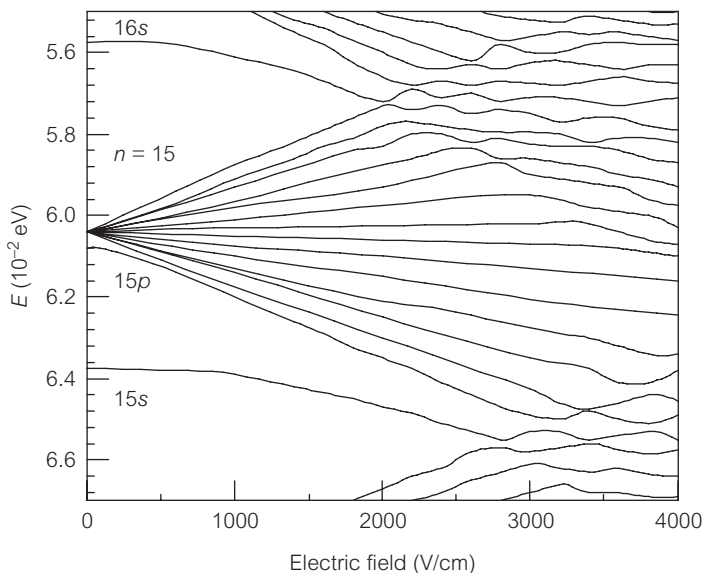
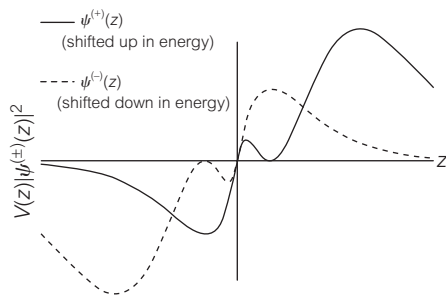


Figure 18.6. Energy versus applied electric field for highly excited states of lithium ($m = 0$ states) illustrating the first-order (linear in E) and second-order (quadratic in E) Stark effects. Data taken from Zimmerman *et al.* (1979).

To visualize this result, in Fig. 18.5 we plot $V(z)|\psi^{(\pm)}(0, 0, z)|^2$ versus z to show the distribution of potential energy in the two cases; the + (−) combinations are obviously shifted up (down) in energy as expected.

It is important to stress that the presence of an energy shift which is linear or first order in E_0 , depends critically on the presence of degenerate energy levels; linear combinations of the $\psi_{2,1,0}$ and $\psi_{2,0,0}$ states are required to produce the dipole moments which can interact via $E = -\mathbf{p} \cdot \mathcal{E}$, and these states must be degenerate (or nearly so) for this mixing to occur.

To exemplify these remarks further, in Fig. 18.6 we show part of the energy spectrum of lithium Rydberg atoms as a function of the strength of the applied field. As discussed in Section 17.3.1, the states with the lowest values of l are

shifted down in energy due to their interactions with the inner electron core; in this case the $15s$ ($l=0$) state is lowered in energy much more than the $15p$ ($l=1$) state while the remaining ones with $l = 2, 3, \dots, 14$ are still almost degenerate. We note the following distinctive features:

- For small field strengths (less than roughly 2500 V/cm), the energy shift due to the external field for the $15s$ state (as well as the $16s$ state) is quadratic instead of linear due to its “isolation” from any nearby degenerate states. The polarizability of these states can be estimated from the data (P18.14)
- For quite small fields, the $15p$ state also has a quadratic energy shift but as the Stark level shifts of the $n = 15$ states become comparable to the initial splitting of the $15p$ state from the rest, it “joins in” and contributes to the linear Stark effect pattern. States do not have to be exactly degenerate to require use of degenerate perturbation theory¹⁶ as discussed in Section 10.5.2.
- The remaining $n = 15$ states show the standard Stark effect of a linear energy shift with a large number of splittings due to the large degeneracy.
- For large enough fields, the perturbation theory predictions become unreliable and the pattern of energy levels becomes highly complex. Note however, the many avoided level crossings which are characteristic of “level repulsion,” as mentioned in Section 10.5.

We briefly mention two other effects which arise in the quantum description of atomic energy levels in an external electric field:

- The combined Coulomb plus external field potential felt by an electron (illustrated in Fig. 18.7) implies that there are no absolutely stable bound states in such a potential as there is always a possibility of quantum tunneling. One

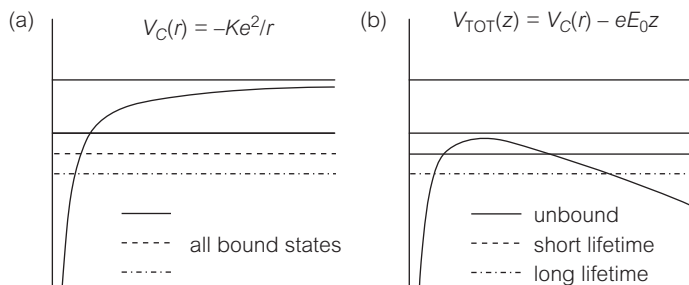


Figure 18.7. Potential energy due to Coulomb plus external electric fields showing the possibility of field ionization and tunneling.

¹⁶ For a nice discussion of degeneracy effects and dipole moments, see Gasiorowicz (1996).

can derive estimates (P18.15) of the tunneling probability which confirm that the lifetimes in modest fields are much longer than the age of the universe, consistent with intuition. If, however, the external field is increased enough then the potential “turns over” allowing states to become unbound, leading to so-called *field ionization*. Simple estimates of the field necessary to unbind the n th state seem to be confirmed by experiment.

- The emission of photons via radiative transitions from one state to another forms the basis of atomic and molecular spectroscopy, which, in turn, provides the evidence for the quantized energy levels calculated in quantum mechanics. The related (inverse) process of single-photon absorption is conceptually the same; electrons absorb photons of energy $\hbar\omega = |E_n - E_{n'}|$ in transitions from one level to another or are ionized from the n -th level if $E_\gamma = \hbar\omega > |E_n|$. Sufficiently intense beams of photons of energy *less than* $|E_n|$ can still ionize atoms¹⁷ via the process of *multiphoton ionization* (MPI) which is illustrated schematically in Fig. 18.8. The absorption of the “first” photon would not normally allow the electron to be freed, but if the number density of photons is high enough so that there is a chance of a second (or third and so forth) interaction, the ionization can occur as a multistep process. A similar process is called *above threshold ionization* (ATI) where the final state energy of the electron is measured to be

$$E_e = -|E_{\text{binding}}| + n\hbar\omega \quad (18.89)$$

with $n > 1$ implying that more than one photon absorption has been used to “kick” the electron into the continuum of free-particle states. Clearly the probability for a k -step ionization process scales with the intensity, I , as I^k . The effect is illustrated schematically in Fig. 18.9 where the same laser beam

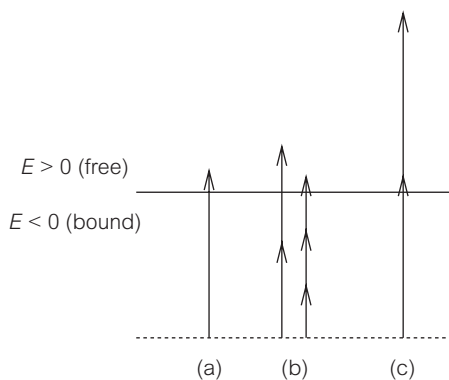
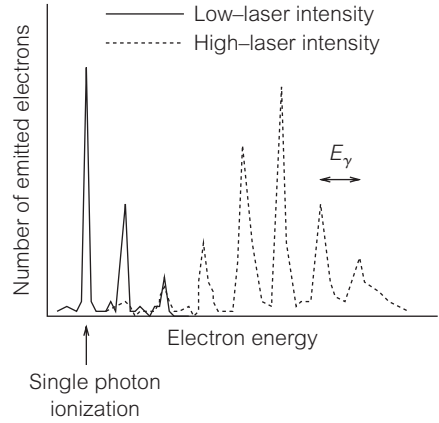


Figure 18.8. Schematic diagram showing photon absorption leading to (a) single photon ionization, (b) multiphoton ionization (MPI), and (c) above threshold ionization (ATI).

¹⁷ For a nice review of multiphoton processes in atoms, see Delone and Krainov (1994).

Figure 18.9. Number of emitted photoelectrons versus their kinetic energy for two laser intensities; for low laser intensities (solid), single photon ionization dominates, while for larger laser intensities (dashed), the absorption of more than one photon becomes appreciable. The distance between peaks is $E_\gamma = \hbar\omega$ indicating that the kinetic energies are given by $mv^2/2 = nE_\gamma - E_0$.



is used in both (a) and (b) (meaning that the frequency ω or energy of each photon E_γ is unchanged), but the intensity is increased near the critical value for the onset of the multiphoton process.

18.5 Constant Magnetic Fields

We next consider the problem of a charged particle (charge q) in a uniform magnetic field in the $+z$ direction. The classical solutions consist of helical trajectories, that is, uniform translational motion in the direction parallel to the field \mathbf{B}_0 and uniform rotational motion in the plane perpendicular to \mathbf{B}_0 ; for a positive charge q , the circular motion is in the clockwise direction when viewed from above ($z > 0$), as in Fig. 18.10.

To discuss the quantum version, we choose the gauge where

$$\mathbf{A}(\mathbf{r}, t) = \frac{B_0}{2}(-y, x, 0) \quad (18.90)$$

and consider other choices in the problems. The Hamiltonian from Eqn. (18.39) can be written as

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2\mu} - \frac{q}{2\mu}(\hat{\mathbf{p}} \cdot \mathbf{A} + \mathbf{A} \cdot \hat{\mathbf{p}}) + \frac{1}{2\mu} \left(\frac{qB_0}{2} \right)^2 (x^2 + y^2) \quad (18.91)$$

The middle terms can be combined and written in the form

$$-\frac{q}{\mu} \frac{\hbar B_0}{i} \frac{1}{2} \left(-\frac{\partial}{\partial y} x + \frac{\partial}{\partial x} y \right) = -\frac{qB_0}{2\mu} \hat{L}_z \quad (18.92)$$

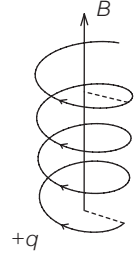


Figure 18.10. Classical helical motion of a positively charged particle in a uniform magnetic field.

This form is clearly a generalization of the classical result

$$\hat{H} = -\mathbf{M} \cdot \mathbf{B} \quad (18.93)$$

where the magnetic moment is given by

$$\mathbf{M} = g \left(\frac{q}{2\mu} \right) \mathbf{L} \quad (18.94)$$

with the classical value of $g = 1$, as found in P16.12 for a rotating particle.

We can define the so-called *Larmor frequency*

$$\omega_L = \frac{qB_0}{2\mu} = \frac{\omega_c}{2} \quad (18.95)$$

which is half the *cyclotron frequency* which corresponds to the classical frequency of circular motion determined via the Lorentz force law, namely

$$F = \mu a_c \quad \longrightarrow \quad qvB_0 = \mu \frac{v^2}{r} = \mu \omega_c^2 r \quad \text{or} \quad \omega_c = \frac{qB_0}{\mu} \quad (18.96)$$

We then write the Hamiltonian in the form

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2\mu} - \omega_L \hat{L}_z + \frac{1}{2} \mu \omega_L^2 (x^2 + y^2) \quad (18.97)$$

The corresponding Schrödinger equation is most conveniently solved in cylindrical coordinates where we assume a solution of the form

$$\psi(\mathbf{r}) = \psi(r, \theta, z) = R(r) e^{im\theta} e^{ik_z z} \quad (18.98)$$

This solution is an eigenfunction of both the \hat{L}_z and z kinetic energy terms which together have contributions to the total energy given by $-m\hbar\omega_L$ and $\hbar^2 k_z^2 / 2\mu$, respectively. The remaining two-dimensional problem is essentially the planar harmonic oscillator (in polar coordinates) discussed extensively in Section 15.3.3. The resulting solutions are given by

$$\psi(r, \theta, z) \propto r^{|m|} e^{-r^2/2\rho^2} L_{n_r}^{|m|}(r^2/\rho^2) e^{im\theta} e^{ik_z z} \quad (18.99)$$

where $\rho^2 = \hbar/\mu\omega_L$ with energy eigenvalues

$$E_{n_r, m} = \hbar\omega_L (2n_r + |m| - m + 1) + \frac{\hbar^2 k_z^2}{2\mu} \quad (18.100)$$

The equally spaced harmonic-oscillator-like energy states are often called *Landau levels*.

As discussed for the two-dimensional oscillator, the classical limit of uniform circular motion in the plane corresponds to the minimum number of radial nodes ($n_r = 0$). In that case, large values of $|m|$ of either sign could give macroscopic energies corresponding to motion in either clockwise or counterclockwise directions. Here, Eqn. (18.100) implies that only $m = -|m|$, that is, $L_z = -|m|\hbar$ or clockwise motion gives that limit, as expected for the motion of a positive charge.

Using the wavefunction solutions in Eqn. (18.99) for $n_r = 0$ and large $|m|$, one can show that the probability density is strongly peaked at a (planar) radius coordinate given by

$$r^2 = \frac{|m|\hbar}{\mu\omega_L} = \frac{|L_z|}{\mu\omega_L} \quad (18.101)$$

To compare this to the classical case, we can use arguments similar to those in Section 1.4 and Section 15.3.3, and the Lorentz force law, to find

$$\frac{\mu v^2}{r} = \mu a_c = F = qvB_0 \quad (18.102)$$

The relationship between velocity and momentum in the presence of a magnetic vector potential is generalized (P18.7) to be

$$\mu \mathbf{v} = \mathbf{p} - q\mathbf{A} \quad (18.103)$$

so that

$$\mu (\mathbf{r} \times \mathbf{v})_z = (\mathbf{r} \times \mathbf{p})_z - q (\mathbf{r} \times \mathbf{A})_z \quad (18.104)$$

or

$$-\mu r v = -|L_z| - \frac{1}{2} q B_0 r^2 \quad (18.105)$$

Equations (18.102) and (18.105) can then be combined to give

$$r^2 = \frac{2|L_z|}{qB_0} = \frac{|L_z|}{\mu\omega_L} \quad (18.106)$$

which is consistent with the large quantum number limit.

The energies of the system now form a continuous spectrum because of the motion in the z direction. Even for a fixed value of k_z , however, there is still

an infinite degeneracy in the quantized energy levels corresponding to the fact that $|m| - m = 0$ for all values of positive m . This degeneracy has its origin in the arbitrariness in the initial conditions for the planar orbits; the x and y coordinates of the center of the circular orbit are not specified even if the total energy and k_z are. This is strictly true only for a region of uniform field which is of infinite extent; if the field is confined in the x and y directions by a box with sides L , not all circular orbits will “fit into” the box. In this more realistic case, the degeneracy, N_d , of each level¹⁸ can be shown to be

$$N_d = \frac{qB_0 L^2}{\pi \hbar} \quad (18.107)$$

which scales as the area as expected.

18.6 Atoms in Magnetic Fields

It has been said that “*Magnetism is inseparable from quantum mechanics. . .*”¹⁹ since it can be shown that systems interacting via purely classical mechanics in statistical equilibrium can exhibit no magnetic moments, even in response to externally applied magnetic fields. Of the many possible manifestations of magnetic effects in quantum mechanical systems, in this section, we consider only three:

1. The problem of a one-electron atom subject to an external \mathbf{B} field (the Zeeman effect).
2. The interactions of the electron spin in an atom due to the “internal” magnetic field caused by the orbital motion of the atomic constituents themselves (the so-called spin-orbit coupling).
3. The magnetic couplings of two spin magnetic moments in an atom, giving rise to so-called hyperfine splittings in atoms and contributing to spin–spin level shifts in other systems.

18.6.1 The Zeeman Effect: External \mathbf{B} Fields

To study the effect of a constant external magnetic field (oriented in the z -direction for definiteness) on a one-electron atom, we consider the

¹⁸ For a careful discussion of the boundary conditions which give rise to this estimate, see Peierls (1955).

¹⁹ See Kittel (1971); He goes on to say “and were the value of \hbar to go to zero, the loss to the science of magnetism is one of the catastrophes that would overwhelm the universe.”

Hamiltonian for such a system (in a standard gauge) given by

$$\begin{aligned}\hat{H} &= \left(\frac{\hat{\mathbf{p}}^2}{2m_e} - \frac{KZe^2}{r} \right) + \frac{eB}{2m_e} \hat{L}_z + \frac{e^2 B^2}{8m_e} (x^2 + y^2) \\ &= \hat{H}_{\text{Coul}} + \hat{H}_1 + \hat{H}_2\end{aligned}\quad (18.108)$$

where $q = -e$ and we assume $\mu \approx m_e$. This is strictly valid only for a spinless electron as we have ignored the coupling of its intrinsic angular momentum to the external field; this complication is discussed in P18.22.

For sufficiently small applied fields, we can initially neglect the term quadratic in \mathbf{B} , the so-called *diamagnetic term*; in this case we see that the eigenfunctions of the standard Coulomb problem remain solutions since the spherical harmonics are also eigenfunctions of \hat{L}_z . The shift in energy due to the external field is then given by

$$E_B^{(1)} Y_{l,m} = \hat{H}_1 Y_{l,m} = \left[\left(\frac{e\hbar}{2m_e} B \right) m \right] Y_{l,m} \quad (18.109)$$

or

$$E_B^{(1)} = \left[\frac{e\hbar}{2m_e} B \right] m = (\mu_e B) m = m\hbar\omega_L \quad (18.110)$$

We have written this in two complementary forms:

- One form implicitly use the *Bohr magneton*, $\mu_e = e\hbar/2m_e$, whose numerical value²⁰ is $\mu_e = 5.788 \times 10^{-5} \text{ eV/T}$; this emphasizes its identification as being the interaction energy of the magnetic moment due to orbital motion with the external field.
- The second form is written in terms which are reminiscent of the quantized Landau energy level spacing for a charged particle discussed in Section 18.5.

The introduction of \hat{H}_1 reduces the (unexpectedly large) symmetry of the pure Coulomb problem by picking out a specific axis, and the energy spectrum is correspondingly ‘disrupted’; L_z , however, remains a constant of the motion and the problem is still exactly soluble. The inclusion of \hat{H}_2 completely destroys the symmetry, but for small fields the effect of \hat{H}_2 can then be estimated via perturbation theory using the standard solutions for hydrogen-like atoms; using the angular and radial overlap integrals in Eqn. (16.83) and Eqn. (17.30), and the fact that

$$x^2 + y^2 = r^2 \sin^2(\theta) \quad (18.111)$$

²⁰ In this section, as elsewhere, the appropriate MKSA unit of magnetic field is the Tesla (T) which is related to the Gauss (G) via $1 \text{ T} = 10^4 \text{ G}$.

we find that

$$E_B^{(2)} \approx \frac{e^2 B^2 a_0^2}{8m_e} F(n, l, m) \quad (18.112)$$

where

$$F(n, l, m) = \left[n^2(5n^2 + 1 - 3l(l+1)) \frac{(l^2 + l - 1 - m^2)}{(2l-1)(2l+3)} \right] \quad (18.113)$$

when evaluated in the state $R_{n,l,m} = R_{n,l}(r)Y_{l,m}(\theta, \phi)$ and where a_0 is the Bohr radius. The prefactor in Eqn. (18.112) can be written in the form

$$\frac{e^2 B^2 a_0^2}{8m_e} = \frac{1}{4} \frac{(\hbar\omega_L)^2}{|E_l|} \quad (18.114)$$

where the E_n are the Bohr energy levels for the Coulomb problem. To see how this scales with n , we can choose for definiteness $l = n-1$ and $m=0$ and find that

$$E_B^{(2)} \approx \frac{e^2 B^2 a_0^2}{16m_e} n^4 \quad (18.115)$$

In the limit when the effects of \hat{H}_2 can no longer be considered as a small perturbation, other calculational tools (such as variational methods or diagonalization of matrices) must be employed to estimate the energy eigenvalues using the full Hamiltonian of Eqn. (18.108).

Considering only $E_B^{(1)}$, we see that for a given value of l , the $2l+1$ degenerate levels corresponding to different values of m are split by equal amounts and the pattern of energy level shifts is shown in Fig. 18.11. Despite the seemingly large number of possible new transitions, the selection rule for the magnetic quantum

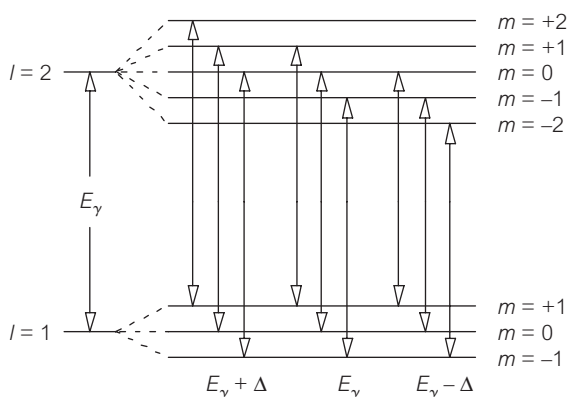


Figure 18.11. Energy level splittings for P ($l=1$) and D ($l=2$) states illustrating the linear Zeemaeffect.

number for dipole radiation (see Section 16.3.3), namely, $\Delta m = +1, 0, -1$, and the uniform splitting imply that the absorption or emission line corresponding to the transition in the $B = 0$ case (E_γ) is only split into three distinct lines ($E_\gamma + \mu_e B \Delta m$). This pattern of shifts in spectral lines is called the (*linear*) *ordinary Zeeman effect*, and it is observed for small field strengths in atoms in which the total electronic spin is zero; the more usual case where one must also consider the electron spin exhibits more structure, and is called the *anomalous Zeeman effect*.

To better understand the relative magnitudes of the energy splittings induced by $\hat{H}_{1,2}$, we first note that the difference between adjacent energy levels in hydrogen goes as

$$\Delta E_n \sim 13.6 \text{ eV} \left(\frac{1}{(n-1)^2} - \frac{1}{n^2} \right) \sim 27 \text{ eV} \frac{1}{n^3} \quad (18.116)$$

The maximum splitting due to \hat{H}_1 between the uppermost ($m = +l$) and lowermost ($m = -l$) scales roughly as

$$\Delta E_B^{(1)} \sim (5.8 \times 10^{-5} \frac{\text{eV}}{\text{T}}) B (2n) \quad (18.117)$$

since $n \geq l$. This implies that

$$\frac{\Delta E_B^{(1)}}{\Delta E_n} \sim \left(\frac{B}{2.3 \times 10^5 \text{ T}} \right) n^4 \quad (18.118)$$

Typical values of the magnetic field range from $B \sim 0.5 \text{ G} = 5 \times 10^{-5} \text{ T}$ (corresponding to the earth's intrinsic field) or less (if shielding is provided) to $B \sim 5 - 10 \text{ T}$, which is now standardly available in laboratories. For low-lying states ($n = 1, 2, \dots$), we see that the (linear) Zeeman splittings are always much smaller than the differences between energy levels; however, Eqn. (18.118) implies that the two become comparable in a 1 Tesla field for $n \sim 22$. We can also write the ratio in Eqn. (18.118) in a more symbolic fashion in the form

$$\frac{\Delta E_B^{(1)}}{\Delta E_n} = \left(\frac{B}{B_c} \right) n^4 \quad \text{with} \quad B_c \equiv \frac{\hbar}{ea_0^2} = \frac{\Phi_B}{\pi a_0^2} = 2.3 \times 10^5 \text{ T} \quad (18.119)$$

where $\Phi_B = \pi \hbar / e$ is written in terms of fundamental constants, and has the units of a magnetic flux.

For such large fields (B) or principal quantum numbers (n), we must consider the effect of \hat{H}_2 , and we can use the perturbation theory estimate of Eqn. (18.114) to write

$$\frac{\Delta E_B^{(2)}}{\Delta E_n} \approx \left(\frac{B}{4B_c} \right)^2 n^7 \quad (18.120)$$

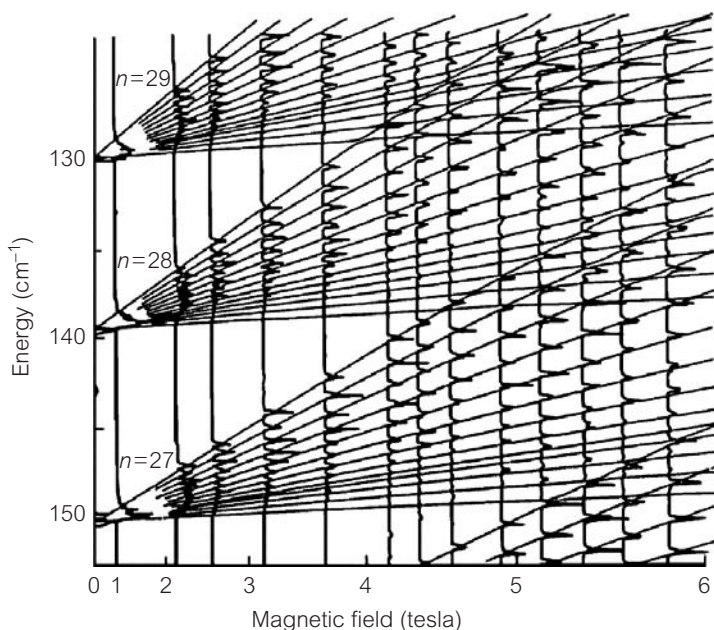


Figure 18.12. Energy versus applied magnetic field squared (note the horizontal scale!) illustrating the effect of diamagnetic \mathbf{H}_2 term in the Hamiltonian; data taken from Zimmerman *et al.* (1978) from measurements of lithium atoms.

This condition implies that the *quadratic Zeeman effect* (with energy shifts varying as B^2) will dominate over the more familiar linear Zeeman effect for highly excited Rydberg states (n large) or for sufficiently large magnetic field strengths²¹

As an example of the effect of the diamagnetic term in the Hamiltonian, \hat{H}_2 , in the large field limit, we show in Fig. 18.12 the spectra of highly excited ($n = 27 - 29$) sodium Rydberg states in magnetic fields up to $B = 6\text{ T}$. The experiment was performed by selectively choosing specific m states so that the linear Zeeman effect is absent. The theoretical predictions (obtained by matrix diagonalization of the Hamiltonian) are shown in Fig. 18.12 (solid lines) along with the experimental excitation curves (turn your head 90°) at various values of the external field. The energy (note the units) is plotted versus B^2 (note the horizontal scale also) to show the quadratic (parabolic) dependence on B ; note also the avoided level crossings.

²¹ It is estimated that when $B > 100\text{ T}$ one has $E_B^{(2)} > E_B^{(1)}$. Such large values of the magnetic field can be found at the surfaces of collapsed astrophysical objects such as white dwarf or neutron stars, and quadratic Zeeman effects due to them have likely been seen in spectral lines from such objects.

18.6.2 Spin-Orbit Splittings: Internal B Fields

Even in the absence of externally applied magnetic fields, the charged particles (and their associated magnetic moments) in an atom are still subject to magnetic forces due to the “motion” of the other charges. To understand the physical origin of this effect, we consider first the classical problem a charged particle in uniform circular motion, as in Fig. 18.13; the circulating charge is equivalent to a current loop and the magnetic field at the center of the loop can be calculated classically using the Biot–Savart law. The small contribution from one “part of the circuit” can be integrated to obtain the familiar result

$$d\mathbf{B} = \frac{\mu_0}{4\pi} \frac{Id\mathbf{l} \times \mathbf{r}}{r^3} = \frac{\mu_0}{4\pi} \frac{dq\mathbf{v} \times \mathbf{r}}{r^3} \Rightarrow \mathbf{B}_{\text{cent}} = \frac{\mu_0}{4\pi} \frac{qv}{r^2} \hat{\mathbf{k}} \quad (18.121)$$

If we

- Associate the circulating charge $q = Ze$ with the nuclear charge (remember that the electron and nucleus are in orbit around each other so the electron “sees” a positively charged current loop),
- Recall that the fundamental constant of magnetism can be rewritten using $c^2 = 1/\mu_0\epsilon_0$, and
- Use the classical relation for the angular momentum $|\mathbf{L}| = rp = rmv$,

we can estimate that an atomic electron “sees” an effective magnetic field of the order

$$\mathbf{B}_{\text{cent}} = \frac{\mu_0}{4\pi} \frac{Ze}{mr^3} \mathbf{L} = \frac{Ze^2}{4\pi\epsilon_0} \frac{1}{mc^2 r^3} \mathbf{L} \quad (18.122)$$

We can make an initial estimate of the magnitude of this field by letting $|\mathbf{L}| \sim \hbar$ and $r \sim a_0$ and find that $B_{\text{cent}} \sim 12 \text{ T}$; this is much larger than typical external fields, and can cause sizable level splittings.

The magnetic moment associated with the electron, namely

$$\mathbf{M} = \frac{gq}{2m} \mathbf{S} = -\frac{e}{m_e} \mathbf{S} \quad (18.123)$$

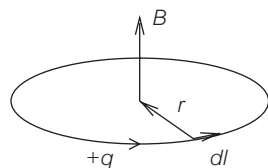


Figure 18.13. Classical picture of the magnetic field at center of current loop.

will then have an interaction energy in the atomic field given by

$$E = -\mathbf{M} \cdot \mathbf{B} = \frac{ZKe^2}{m^2 c^2} \frac{1}{r^3} \mathbf{L} \cdot \mathbf{S} \quad (18.124)$$

where $K = 1/4\pi\epsilon_0$ as usual.

This classical estimate of the *spin-orbit (SO) interaction* (so-called because it exhibits the coupling between the electron spin and the orbital motion) neglects other important relativistic effects (note the explicit factors of c which appear), and the corresponding quantum mechanical expression is most convincingly derived by a nonrelativistic reduction of the Dirac equation for the electron. When this is done, the precise expression for the spin-orbit coupling term differs by only a factor of $1/2$ from our simple estimate, giving the spin-orbit coupling term

$$\hat{H}_{SO} = \frac{ZKe^2}{2m^2 c^2} \frac{1}{r^3} \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \quad (18.125)$$

To evaluate the effect of this interaction on the spectrum of a one-electron atom, we first need to calculate the effect of the $\mathbf{L} \cdot \mathbf{S}$ term. The total angular momentum due to orbital and spin contributions will be given by

$$l + 1/2 \implies j = l + 1/2, \quad l - 1/2 \quad (18.126)$$

with wavefunctions discussed in P16.16. The more formal relation among the corresponding operators

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (18.127)$$

can be manipulated to yield

$$\mathbf{J}^2 = (\mathbf{L} + \mathbf{S})^2 \implies \mathbf{L} \cdot \mathbf{S} = \frac{1}{2} (\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) \quad (18.128)$$

Recalling that S is fixed, we then have

$$\mathbf{L} \cdot \mathbf{S} = \frac{1}{2} \left(j(j+1) - l(l+1) - \frac{3}{4} \right) \hbar^2 \quad (18.129)$$

or

$$\mathbf{L} \cdot \mathbf{S} = \begin{cases} l/2 & \text{for } j = l + 1/2 \\ -(l+1)/2 & \text{for } j = l - 1/2 \end{cases} \quad (18.130)$$

which implies that states with $j = l + 1/2$ ($l - 1/2$) are shifted up (down) in energy. For states with $l > 0$, the original $2(2l+1)$ degenerate levels are split into two distinct sets of states with $2(l+1/2) + 1 = 2l+2$ and $2(l-1/2) + 1 = 2l$ levels,

respectively. For $l=0$, only $j=1/2$ is allowed and there is no splitting. This pattern of energy level shifts is sometimes called *Lande's interval rule*.

The fact that \hat{H}_{SO} is relativistic in origin suggests that its effects may be small for low Z (recall Eqn. (17.16)), and allows us to use perturbation theory to estimate its size. We require the expression

$$\left\langle \frac{1}{r^3} \right\rangle_{n,l} = \left(\frac{Z^3}{a_0^3} \right) \frac{2}{n^3 l(2l+1)(l+1)} \quad (18.131)$$

The splittings due to this spin-orbit coupling can then be written as

$$\Delta E_{n,l}^{SO} = \frac{Ke^2}{2m_e^2 c^2} \frac{Z^4 \hbar^2}{a_0^3} \frac{1}{n^3 l(2l+1)(l+1)} \begin{cases} +l & \text{for } j = l + 1/2 \\ -(l+1) & \text{for } j = l - 1/2 \end{cases} \quad (18.132)$$

We note that this expression is not obviously well-defined for $l=0$, being a ratio of the form $0/0$. A careful derivation using the Dirac equation shows that the s -state energy levels are, in fact, shifted by just the amount predicted by Eqn. (18.132) by simply canceling the factors of l for the $j=l+1/2$ case.²²

Using the expression for the unperturbed energy levels, we can rewrite Eqn. (18.132) in the more compact form

$$\frac{\Delta E_{n,l}^{SO}}{|E_n|} = (Z\alpha)^2 \frac{1}{nl(2l+1)(l+1)} \begin{cases} +l & \text{for } j = l + 1/2 \\ -(l+1) & \text{for } j = l - 1/2 \end{cases} \quad (18.133)$$

Several comments can be made:

- The factor of $(Z\alpha)^2$ in Eqn. (18.133) implies that spin-orbit splittings are typically $Z^2 \times 10^{-5}$ smaller than the unperturbed energy levels; this justifies our use of perturbation theory.
- At this same level of precision, we must also include relativistic corrections to the non-relativistic approximation for the kinetic energy (as in Section 1.4), which we standardly use. We can use the results of P17.12 in the form

$$\frac{\Delta E_{n,l}^{rel}}{|E_n|} = -\frac{(Z\alpha)^2}{n} \left[\frac{2}{2l+1} - \frac{3}{4n} \right] \quad (18.134)$$

²² This result may seem counter-intuitive as we would expect no spin-orbit contribution for s -states for which there is no angular momentum; in fact, a more careful evaluation of the effect of Eqn. (18.125) for the case of $l=0$, not based on perturbation theory, shows that $\langle 1/r^3 \rangle$ is indeed finite, so that it gives no effect on s -states. There is, however, another term which arises in the nonrelativistic reduction of the Dirac equation (which has no simple classical correspondence) which is only nonvanishing for s -states, and which gives exactly the effect as Eqn. (18.132) which we then take as *operationally* correct for all values of l ; for a thorough discussion, see Condon and Shortley (1951).

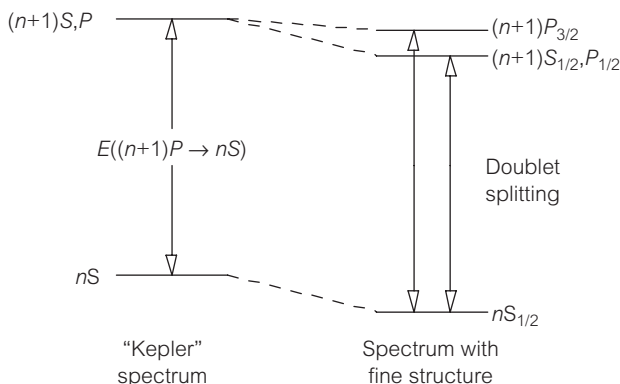


Figure 18.14. Splittings of energy levels due to fine structure (spin-orbit plus relativistic kinetic energy effects) on S and P states; the “doublet” splitting relevant for $(n+1)P \rightarrow nS$ transitions is due to the spin-orbit coupling only.

and taken together with Eqn. (18.133), these two relativistic corrections successfully reproduce the results of the Dirac equation analysis up to order $(Z\alpha)^4$, namely

$$\frac{\Delta E_{n,l}^{rel} + \Delta E_{n,l}^{SO}}{|E_n|} = -\frac{(Z\alpha)^2}{n} \left[\frac{2}{2j+1} - \frac{3}{4n} \right] \quad (18.135)$$

for both $j = l \pm 1/2$. The spin splittings due to these combined relativistic effects are shown schematically in Fig. 18.14 for adjacent S and P states. The spectroscopic notation nL_j is used to distinguish states of different principle quantum number (n), orbital angular momentum ($L = S, P, D, F, \dots$ for $l = 0, 1, 2, 3, \dots$) and total angular momentum $j = l \pm 1/2$.

- The splitting of the $(n+1)P \rightarrow nS$ transition line into two distinct lines is due to the energy difference between the $P_{3/2,1/2}$ energy levels, which, in turn arises solely from the spin-orbit coupling since the relativistic correction is the same for fixed l, n . This splitting is responsible for the distinctive pattern of *doublets* in the line spectra of alkali atoms; the pair of closely spaced lines called the *sodium doublet*²³ is perhaps the most famous example, and corresponds to the transition of the outermost electron from the $4P \rightarrow 3S$ level.

Spin-orbit coupling terms are also present in nuclear systems where they play an important role in understanding the observed pattern of the so-called magic numbers implied by nuclear shell structure. Just as with atomic systems, large

²³ This fact is the basis for the only suggestion for a quantum mechanics experiment in this book, namely, looking at the emission spectrum of sodium (using a diffraction grating) by appropriately igniting some table salt, that is, sodium chloride; see Crawford (1968) for more concrete suggestions.

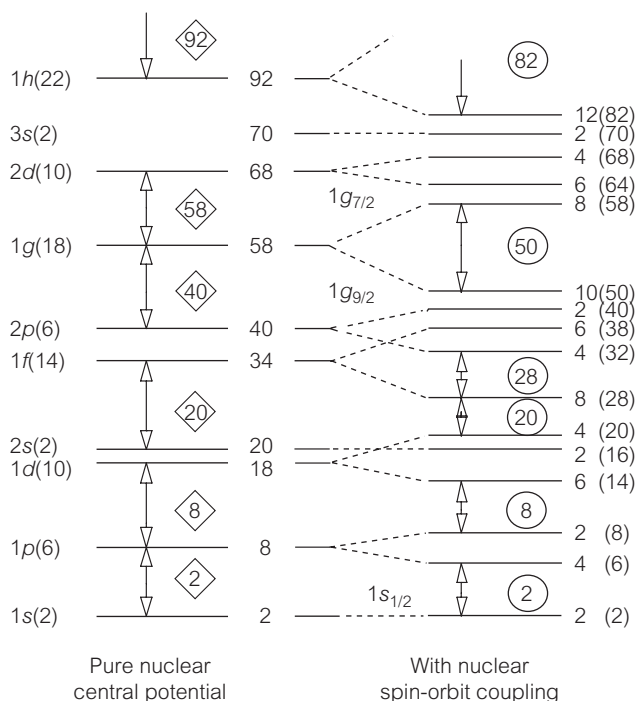


Figure 18.15. Energy level scheme for typical central nuclear potential (left) and including spin-orbit interactions (right). The inclusion of the spin-orbit couplings reproduces the observed “magic numbers”.

energy gaps between relatively closely spaced sets of quantized energy levels in nuclei can give rise to especially stable configurations as the available states are “filled” with neutrons and protons; many observable properties are correlated with nuclei whose Z or N is equal to 2, 8, 20, 28, 50, 82, 126.

Calculations for various models described by purely central potentials, $V(\mathbf{r}) = V(r)$, as in Fig. 18.15, give levels consistent with the first few such “magic numbers,” but fail to reproduce the observed pattern for heavier nuclei. If one includes a strong spin-orbit coupling²⁴ of the form $V_{SO}(r) \mathbf{L} \cdot \mathbf{S}$ (with the appropriate sign for $V_{SO}(r)$), the level structure is changed dramatically for larger angular momentum states and nicely reproduces observed nuclear shell structure. For example, the lowest-lying $l = 4$ state, here labeled $1g(18)$, can accommodate $2(2 \cdot 4 + 1) = 18$ spin-1/2 neutrons or protons; the total angular momentum of a single nucleon in such a state is $J = L + S = 4 + 1/2 = 7/2$ (8 states) and $J = 9/2$ (10 states), and the fine-structure interaction splits these two combinations into different shells.

²⁴ This observation was made independently by Maria Goeppert Mayer and H. D. Jensen for which they shared the Nobel prize in 1963.

18.6.3 Hyperfine Splittings: Magnetic Dipole–Dipole Interactions

The magnetic moment of an atomic electron can also interact with the magnetic dipole field associated with the nuclear magnetic moment. In the context of hydrogen (and other atomic systems) the resulting spin–spin interactions give rise to small level shifts called *hyperfine splittings* (hereafter *h.f.s.*), and are also present in other “hydrogenic” bound state systems.

The magnetic field from a *point* dipole, \mathbf{M} , is given by

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \left(\frac{3(\mathbf{M} \cdot \mathbf{r}) \mathbf{r}}{r^5} - \frac{\mathbf{M}}{r^3} + \frac{8\pi}{3} \mathbf{M} \delta(\mathbf{r}) \right) \quad (18.136)$$

The first two terms are the familiar result derived in most standard texts on electricity and magnetism as the dipole field of a distant current loop; the third term is more subtle,²⁵ and arises when one considers point dipoles. The corresponding classical energy of one dipole in the field of another is then simply

$$\begin{aligned} E &= -\mathbf{M} \cdot \mathbf{B} = \hat{H}_{\text{dip}} \\ &= -\frac{\mu_0}{4\pi} \left(\frac{3(\mathbf{M}_N \cdot \mathbf{r})(\mathbf{M}_e \cdot \mathbf{r})}{r^5} - \frac{\mathbf{M}_N \cdot \mathbf{M}_e}{r^3} + \frac{8\pi}{3} \mathbf{M}_N \cdot \mathbf{M}_e \delta(\mathbf{r}) \right) \end{aligned} \quad (18.137)$$

where we have specialized to nuclear (N) and electronic (e) moments given by

$$\mathbf{M}_N = g_N \frac{Ze}{2M} \mathbf{S}_N \quad \text{and} \quad \mathbf{M}_e = -g_e \frac{e}{2m_e} \mathbf{S}_e \approx -\frac{e}{m_e} \mathbf{S}_e \quad (18.138)$$

where M , Z , and g_N are the nuclear mass, charge, and gyromagnetic ratio, respectively. The Hamiltonian describing this dipole–dipole interaction is given by this form where $\mathbf{S}_{e,N}$ are associated with spin operators.

The effect of this spin–spin interaction on the ground state of a hydrogen-like atom can be estimated using perturbation theory via $\Delta E_{1,0}^{(h.f.s)} = \langle \psi_{1,0,0} | \hat{H}_{\text{dip}} | \psi_{1,0,0} \rangle$, which we can evaluate using several observations:

1. The expectation value of the “standard” dipole term in Eqn. (18.137) vanishes when evaluated in the spherically symmetric ground state.
2. The “point dipole” term then contributes a factor of

$$\langle \psi_{1,0,0} | \delta(\mathbf{r}) | \psi_{1,0,0} \rangle = |\psi_{1,0,0}(0)|^2 = \frac{1}{\pi} \left(\frac{Z}{a_0} \right)^3 \quad (18.139)$$

3. The magnetic permeability can be replaced in favor of $\mu_0 = 1/(c^2 \epsilon_0)$.

²⁵ For a nice discussion, see Griffiths (1999).

4. The product of spin operators can be performed using the standard trick, namely

$$\mathbf{S}_e \cdot \mathbf{S}_N = \frac{1}{2} (\mathbf{S}^2 - \mathbf{S}_e^2 - \mathbf{S}_N^2) \quad (18.140)$$

where $\mathbf{S} = \mathbf{S}_e + \mathbf{S}_N$ in general. For the case of hydrogen when the total spin can be $S = 1/2 + 1/2 = 0, 1$, we have

$$\mathbf{S}_e \cdot \mathbf{S}_p = \hbar^2 \begin{cases} 1(1+1) - 3/4 - 3/4 = 1/2 & \text{for } S = 1 \\ 0 - 3/4 - 3/4 = -3/2 & \text{for } S = 0 \end{cases} \quad (18.141)$$

These can be combined to find the spin-dependent shifts to the ground state levels of hydrogen, namely

$$\Delta E_{1,0}^{(h.f.s)} = \frac{Ke^2}{2m_e^2 c^2} \frac{Z^4 \hbar^2}{a_0^3} \left[g_p \frac{4}{3} \frac{m_e}{m_p} \right] \begin{cases} +1 & \text{for } S = 1 \\ -3 & \text{for } S = 0 \end{cases} \quad (18.142)$$

which we have put in a form which can be compared more readily to the expression for fine-structure splitting in Eqn. (18.132). We note that:

- The $S = 0$ or 1S_0 state is split down relative to the $S = 1$ or 3S_1 state (where the notation $^S L_J$ is used).
- This hyperfine splitting is suppressed relative to the fine-structure effects by the factor in square brackets, namely, $m_e/M \lesssim 1/1800$ and motivates its name.
- The photon energy corresponding to transitions from the $S = 1$ to $S = 0$ states is roughly $E_\gamma = 5.9 \times 10^{-6}$ eV which corresponds to a frequency and wavelength of roughly $f \approx 1420$ Hz and $\lambda \approx 21$ cm, respectively. Interstellar hydrogen atoms can undergo collisional excitations which populate the (slightly) excited $S = 1$ state from the ground state ($S = 0$), and the resulting 21 cm radio emission line is extensively used by astronomers to map the concentrations of hydrogen gas; for example, Doppler-shifted line profiles of the 21 cm line spectrum have been used to map out the spiral-arm structure of our galaxy. This physical feature of the most basic atom in nature has been described²⁶ as “a unique, objective standard frequency, which must be known to every observer in the universe”; and has given rise to suggestions²⁷ that it be used (or at least monitored) for interstellar communications.

²⁶ See Cocconi and Morrison (1959).

²⁷ For a guide to the scientific literature on search for extraterrestrial civilizations, see Kuiper (1989).

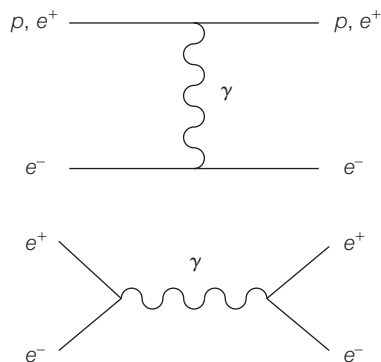


Figure 18.16. Feynman diagrams leading to hyperfine splittings for e^+e^- bound states.

The spin-spin level splittings due to magnetic dipole–dipole interactions are present in other systems as well. In the e^+e^- positronium system, for example, the effect is present with two important changes:

- Because both particles have the same mass, the hyperfine splitting is not suppressed relative to spin-orbit or fine-structure effects as the m_e/M_p term in Eqn. (18.142) is now of order unity.
- An additional physical mechanism contributes to the effective interaction in Eqn. (18.137) due to the possibility of “annihilation”-type interactions as, shown in Fig. 18.16.

18.7 Spins in Magnetic Fields

18.7.1 Measuring the Spinor Nature of the Neutron Wavefunction

We have focused on the interaction of charged particles with \mathbf{E} and \mathbf{B} fields. Electrically neutral particles can have nontrivial EM interactions via the coupling of a magnetic dipole moment with external fields. This was discussed in Section 16.4 in connection with the spinor wavefunction of spin-1/2 particles where it was shown that such an interaction can provide a “handle” on the precession of the spin vector. This fact was the basis for the observation²⁸ that the phase behavior of the neutron spinor wavefunction, with its predicted phase change of -1 on rotation by 2π , could be tested in such systems.

²⁸ First made by Bernstein (1967) and Aharonov and Susskind (1967).

The angle through which a magnetic moment precesses in a field in a small time dt is given by Eqns (16.179) and (16.180) as

$$d\theta = \omega_L dt = \frac{\mu B}{\hbar} dt \quad (18.143)$$

This can be integrated over any path over which the neutron moves with constant speed v to yield

$$\theta = \int d\theta = \frac{\mu_n}{\hbar} \int B dt = \frac{g\mu}{v\hbar} \int_{\text{path}} \mathbf{B} \cdot d\mathbf{l} \quad (18.144)$$

and v is the classical speed. Recall that the neutron's magnetic moment is roughly $\mu_n = -1.93 \mu_N$ where $\mu_N = e\hbar/2m_n$.

After having its magnetic moment precess through the angle θ , the spinor wavefunction becomes

$$\begin{pmatrix} \alpha^+(\theta) \\ \alpha^-(\theta) \end{pmatrix} = \begin{pmatrix} e^{+i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix} \begin{pmatrix} \alpha^+(0) \\ \alpha^-(0) \end{pmatrix} = \begin{pmatrix} \alpha^+(0)e^{+i\theta/2} \\ \alpha^-(0)e^{-i\theta/2} \end{pmatrix} \quad (18.145)$$

Now consider the experiment described schematically by Fig. 18.17 in which a beam of neutrons is split between two paths, one which traverses a magnetic field, and one in a field-free region. If the initial neutron beam is unpolarized (equal amounts of “up” and “down”), then the neutrons along the path ABD experience no change in phase and

$$\psi_{ABD} = \psi_A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (18.146)$$

The pieces of the neutron spin wavefunction which do traverse the magnetic field pick up phases given by

$$\psi_{ACD}(\theta) = \begin{pmatrix} e^{+i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix} \psi_A = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{+i\theta/2} \\ e^{-i\theta/2} \end{pmatrix} \quad (18.147)$$

When the beams are recombined, the total wavefunction is given by

$$\psi_{TOT}(\theta) = \psi_{ABD}(\theta) + \psi_{ACD}(\theta) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 + e^{+i\theta/2} \\ 1 + e^{-i\theta/2} \end{pmatrix} \quad (18.148)$$

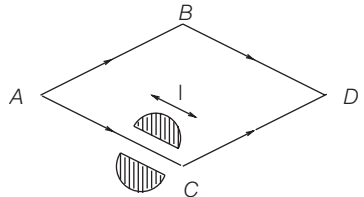


Figure 18.17. Schematic representation of experiment designed to test the phase change on rotation of the neutron's spinor wavefunction. The neutrons following the path ACD traverse a region of magnetic field, while those which follow ABD do not.

so that the probability density is proportional to

$$|\psi_{TOT}(\theta)|^2 = \frac{1}{2} \left(\left| 1 + e^{+i\theta/2} \right|^2 + \left| 1 + e^{+i\theta/2} \right|^2 \right) \propto 1 + \cos\left(\frac{\theta}{2}\right) \quad (18.149)$$

The counting rate at a detector at point D measures the beam intensity and hence this probability; if one scales the intensity obtained with various amounts of “magnetic path length” $\int_{\text{path}} \mathbf{B} \cdot d\mathbf{l}$ to that with the field turned off, one should then have

$$\frac{I(\theta)}{I(0)} = \frac{|\psi_{TOT}(\theta)|^2}{|\psi_{TOT}(0)|^2} = \frac{1 + \cos(\theta/2)}{2} \quad (18.150)$$

This formula exhibits the typical spin-1/2 phase change upon rotation by 2π ; constructive interference corresponding to a return to the same phase is only seen after a rotation of the magnetic moment by an angle of $\theta = 4\pi$.

Two sets of experiments using neutron interferometers (as in Section 18.1) were performed²⁹ soon after the predictions were made and results from one of them are shown in Fig. 18.18. Numerical values for this data set are analyzed in P18.27.

18.7.2 Spin Resonance

The steady precession of a magnetic moment around the direction of a static field can exhibit striking resonance effects if a time-varying magnetic field is applied

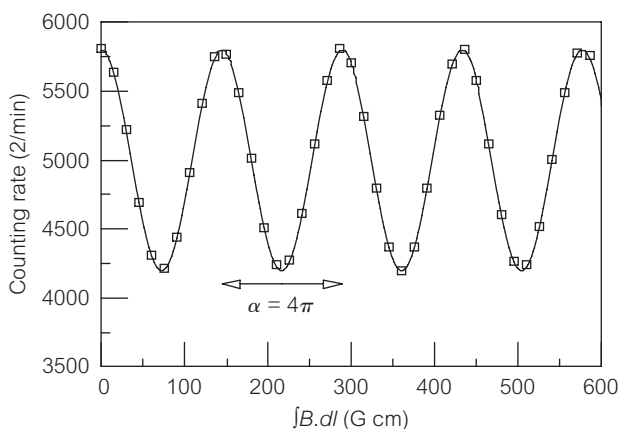


Figure 18.18. Measured neutron intensity versus magnetic path length illustrating the “sign flip” for a spin-1/2 wavefunction rotated through 2π . Data from Rauch *et al.* (1975).

²⁹ See Werner *et al.* (1975) and Rauch *et al.* (1975).

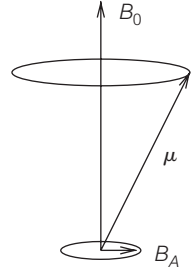


Figure 18.19. Magnetic moment, μ , precessing around static field \mathbf{B}_0 plus additional, rotating magnetic field \mathbf{B}_A .

at right angles to the original field. This can be understood in two relatively simple ways:

- We illustrate in Fig. 18.19 a large, uniform field in the z direction and the precession of the magnetic moment \mathbf{M} (and hence spin direction) around it. The presence of a small, rotating \mathbf{B}_A field will, in general, have little effect as the torque it induces on \mathbf{M} will average out to zero if it rotates at a different rate (or even different direction) than does \mathbf{M} ; if, however, the magnetic field “stays in phase” with the moment by rotating at a rate $\omega = \omega_{\text{precess}} = g\mu B/\hbar$, its torque can act continuously and induce dramatic changes in the rotational motion.
- We have seen (Section 16.4) that in a uniform field, the energy levels of the spin-1/2 particle are given by

$$E_{\pm} = \pm \hbar \omega_{\text{phase}} = \pm \frac{g\mu B}{2} \quad (18.151)$$

which implies a splitting between the parallel and antiparallel configurations of spins given by $\Delta E = g\mu B$; this implies that EM radiation of energy $E_{\gamma} = \hbar\omega = \Delta E$ or angular frequencies $\omega = g\mu B/\hbar$ can preferentially be absorbed.

In either viewpoint, the spin system will exhibit resonance behavior when subjected to external magnetic fields when $\omega = \omega_{\text{precess}}$, and we will analyze the dynamical equations of motion for the spinor wavefunction in a somewhat formal way in order to derive the exact form of the resonance response.

We assume a large, static field (\mathbf{B}_0) in the z direction and a rotating field of arbitrary magnitude (\mathbf{B}_A) at right angles to it, given by

$$\mathbf{B}_0 = (0, 0, B_0) \quad (18.152)$$

$$\mathbf{B}_A = (B_A \cos(\omega t), -B_A \sin(\omega t), 0) \quad (18.153)$$

The spin Hamiltonian is then given by

$$\begin{aligned}\hat{H} &= -\frac{g\mu}{2} \mathbf{B} \cdot \boldsymbol{\sigma} = -\frac{g\mu}{2} \left(B_A^{(x)} \sigma_x + B_A^{(y)} \sigma_y + B_A^{(z)} \sigma_z \right) \\ &= -\frac{g\mu}{2} \begin{pmatrix} B_0 & B_A e^{+i\omega t} \\ B_A e^{-i\omega t} & -B_0 \end{pmatrix}\end{aligned}\quad (18.154)$$

where the spin matrices are defined in Eqn. (16.166). The corresponding Schrödinger equation for the spinor coordinates is

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \alpha_+(t) \\ \alpha_-(t) \end{pmatrix} = \hat{H} \begin{pmatrix} \alpha_+(t) \\ \alpha_-(t) \end{pmatrix} \quad (18.155)$$

which can be written in the form

$$\begin{pmatrix} \dot{\alpha}_+(t) \\ \dot{\alpha}_-(t) \end{pmatrix} = \frac{i}{2} \begin{pmatrix} \omega_0 & \omega_A e^{+i\omega t} \\ \omega_A e^{-i\omega t} & -\omega_0 \end{pmatrix} \begin{pmatrix} \alpha_+(t) \\ \alpha_-(t) \end{pmatrix} \quad (18.156)$$

where we have defined two new frequencies,

$$\omega_0 = \frac{g\mu B_0}{\hbar} \quad \text{and} \quad \omega_A = \frac{g\mu B_A}{\hbar} \quad (18.157)$$

in addition to the frequency ω describing the time rate of change of the external field. These then give the coupled differential equations

$$2\dot{\alpha}_+(t) = i \left(\omega_0 \alpha_+(t) + \omega_A e^{+i\omega t} \alpha_-(t) \right) \quad (18.158)$$

$$2\dot{\alpha}_-(t) = i \left(\omega_A e^{-i\omega t} \alpha_+(t) - \omega_0 \alpha_-(t) \right) \quad (18.159)$$

We assume that the spinor is initially in the “up” direction so that the

$$\begin{pmatrix} \alpha_+(0) \\ \alpha_-(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (18.160)$$

and attempt to find the solutions of Eqn. (18.156) subject to these initial conditions.

Motivated by the time-dependence of the spinor wavefunction when $B_A = 0$, we attempt a solution of the forms

$$\alpha_+(t) = A_+ e^{i\omega_+ t} \quad \text{and} \quad \alpha_-(t) = A_- e^{i\omega_- t} \quad (18.161)$$

so that Eqns (18.158) and (18.159) become

$$2\omega_+ A_+ = \omega_0 A_+ + \omega_A A_- e^{i(\omega_+ + \omega_- - \omega_+)t} \quad (18.162)$$

$$2\omega_- A_- = \omega_A A_+ e^{-i(\omega_+ + \omega_- - \omega_-)t} - \omega_0 A_- \quad (18.163)$$

The time-dependence can be eliminated if we choose

$$\omega_+ = \omega_- + \omega \quad (18.164)$$

which gives now the coupled *algebraic* equations in matrix form

$$\begin{pmatrix} (2\omega_+ - \omega_0) & -\omega_A \\ -\omega_A & (2\omega_- + \omega_0) \end{pmatrix} \begin{pmatrix} A_+ \\ A_- \end{pmatrix} = 0 \quad (18.165)$$

For a consistent solution, the determinant of coefficients must vanish and if we combine this condition with Eqn. (18.164) we find that

$$\omega_- = -\frac{\omega}{2} \pm \frac{1}{2} \sqrt{(\omega - \omega_0)^2 + \omega_A^2} = -\frac{\omega}{2} \pm \Delta\omega \quad (18.166)$$

so that $\omega_+ = +\omega/2 \pm \Delta\omega$.

As with any set of coupled second-order differential equations, the result for $\alpha_-(t)$ will consist of a linear combination of the two independent solutions, namely

$$\alpha_-(t) = A e^{-i\omega t/2} e^{+i\Delta\omega t} + B e^{-i\omega t/2} e^{-i\Delta\omega t} \quad (18.167)$$

The initial condition $\alpha_-(0) = 0$ implies that $B = -A$ while its derivative is given by $\dot{\alpha}_-(0) = 2iA\Delta\omega$. Substituting these values into Eqn. (18.159) and using the other initial condition on the spinor upper component, namely, $\alpha_+(0) = 1$, implies that

$$2iA = \frac{i\omega_A}{2\Delta\omega} \quad (18.168)$$

Our “big result” is then that the total time-dependence is given by

$$\alpha_-(t) = \frac{i\omega_A}{2\Delta\omega} \sin(\Delta\omega t) e^{-i\omega t/2} \quad (18.169)$$

This spinor amplitude gives information on the probability that the spin will have “flipped” to a state anti-parallel to B_0 at later times since

$$\text{Prob}(\text{spin down})(t) = |\alpha_-(t)|^2 = \left[\frac{\omega_A^2}{\omega_A^2 + (\omega - \omega_0)^2} \right] \sin^2(\Delta\omega t) \quad (18.170)$$

Observations on this result include:

- The prefactor in Eqn. (18.170), namely,

$$\frac{\omega_A^2}{\omega_A^2 + (\omega - \omega_0)^2} \quad (18.171)$$

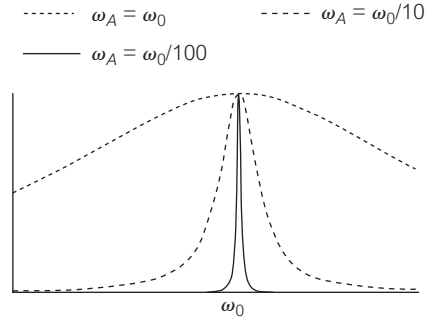


Figure 18.20. Probability of a spin-flip versus frequency of applied field ω showing Lorentzian line shape for spin-resonance; different values of ω_A are shown indicating the effect of the magnitude of B_A on the “sharpness” of the resonance peak.

exhibits the *Lorentzian line shape* discussed in P4.9 which is typical of resonance phenomena, and we will illustrate its form for various values of ω_A in Fig. 18.20. The spin-flip amplitude has its maximum value when

$$\omega = \omega_0 = \frac{g\mu B_0}{\hbar} = \omega_{\text{precess}} \quad (18.172)$$

as expected, but we also note that the “sharpness” of the resonance is proportional to ω_A . In order to measure the resonant frequency as precisely as possible (which is the hallmark of the method), we want to make B_A small which often means that the sample must be shielded from stray fields not under the control of the experimenter (the earth’s field, for example, or stray radio-frequency (RF) signals).

- When applied to an unpaired electron spin, this method is called *electron spin resonance* or *ESR* and for typical external field strengths of $B_0 \approx 0.3 \text{ T}$, the resonant frequencies, wavelengths, and photon energies are

$$f \approx 10 \text{ GHz}, \quad \lambda \approx 3 \text{ cm}, \quad \text{and} \quad E_\gamma = \Delta E \approx 4 \times 10^{-5} \text{ eV} \quad (18.173)$$

so that ESR usually requires microwave techniques. When applied to the nuclear magnetic moments of free or unpaired protons (or other nuclei with nonvanishing spins) the technique is called *nuclear magnetic resonance* or *NMR*; the corresponding resonance parameters for protons in a 1 T field are

$$f \approx 40 \text{ MHz}, \quad \lambda \approx 7 \text{ m}, \quad \text{and} \quad E_\gamma = \Delta E \approx 2 \times 10^{-7} \text{ eV} \quad (18.174)$$

which shows that RFs are required for the time-dependent fields.

- Even with large field strengths B_0 , the energy “gain” from being parallel to the field is much smaller than ordinary thermal fluctuations so that there is only a small excess of “spin-up” (N_+) versus “spin-down” (N_-) states in a typical sample. For example, the population ratio in the two states is given by their

respective Boltzmann factors, namely

$$\frac{N_+}{N_-} = \frac{e^{-E_+/kT}}{e^{-E_-/kT}} = e^{-\Delta E/kT} \quad (18.175)$$

where k is Boltzmann's constant and T is the temperature. At room temperatures, where $kT \approx 1/40$ eV, this implies that

$$\left(\frac{N_+ - N_-}{N_+ + N_-} \right) \sim \frac{1}{2} \left(1 - e^{-\Delta E/kT} \right) \sim \frac{\Delta E}{2kT} \approx 4 \times 10^{-6} \quad (18.176)$$

for the proton values above.

- Electron spin resonance and NMR have been extensively applied in the fields of nuclear and solid state physics as well as in chemistry. The precision with which frequency measurements can be made can then be translated into very accurate determinations of nuclear magnetic moments. Nuclei with known moments can then be used to probe the electronic environment in solids or determine the structure of complex molecules. The technique is perhaps best known, however, when it is applied to probing the protons in the human body where it is known as *magnetic resonance imaging* or *MRI*³⁰

18.8 The Aharonov–Bohm Effect

We have so far considered the EM interactions of charged particles and magnetic dipoles in regions where there is a nonvanishing \mathbf{B} field. We now describe a strikingly “quantum” phenomenon in which the phase of the wavefunction of a charged particle is changed even in a field-free region; the phase can be calculated in terms of the vector potential \mathbf{A} and shows that the potentials themselves can play an important role.

We have stressed that a given configuration of electric and magnetic fields can be derived from an arbitrarily large number of different EM potentials, ϕ and \mathbf{A} . This fact even carries over to the case where there are no magnetic fields present. The restriction

$$\nabla \times \mathbf{A}(\mathbf{r}, t) = \mathbf{B}(\mathbf{r}, t) = 0 \quad (18.177)$$

in a field-free region can still be satisfied by the gradient of any scalar function, namely, $\mathbf{A}(\mathbf{r}, t) = \nabla f(\mathbf{r}, t)$; this fact is also consistent with the gauge transformation Eqns (18.30) and (18.31) starting with a vanishing \mathbf{A} . This relation can be

³⁰ For a review of the physical principles and clinical applications of MRI, see Partain *et al.* (1988).

inverted to give

$$f(\mathbf{r}, t) = \int^{\mathbf{r}} d\mathbf{r}' \cdot \mathbf{A}(\mathbf{r}', t) \quad (18.178)$$

so that the phase factor connecting a free particle wavefunction to its gauge equivalent partner can be written as

$$\psi'(\mathbf{r}, t) = e^{iqf(\mathbf{r}, t)/\hbar} [\psi_{\text{free}}(\mathbf{r}, t)] = \exp\left(iq \int^{\mathbf{r}} d\mathbf{r}' \cdot \mathbf{A}(\mathbf{r}', t)/\hbar\right) \psi_{\text{free}}(\mathbf{r}, t) \quad (18.179)$$

The phase factor is seemingly arbitrary as it depends on the choice of gauge, and is unmeasurable by itself as $|\psi'|^2 = |\psi|^2$.

We have already been reminded in Sections 18.1 and 18.7.1 that interference experiments are necessary to probe the relative phases of wavefunctions so, before proceeding further, we recall the wave physics behind electron (or any wave) interference experiments as illustrated in Fig. 18.21. The wavefunctions from the source, S , travel different path lengths and acquire different phase factors. When the electron beams are recombined the wave amplitude can be expressed as

$$\psi_{\text{tot}}^O = \psi_1^O + \psi_2^O = \psi_1^S e^{ikl_1} + \psi_2^S e^{ikl_2} = \left(\psi_1^S + \psi_2^S e^{i\theta}\right) e^{ikl_1} \quad (18.180)$$

where

$$\theta = k(l_2 - l_1) = \frac{2\pi}{\lambda} \Delta l \quad (18.181)$$

This implies the familiar result that a difference in path length equal to an integral number of wavelengths will give constructive interference.

Aharonov and Bohm³¹ suggested a conceptually simple version of such an interference experiment to probe the quantum phase induced by the vector potential which is illustrated in Fig. 18.22. A small region of magnetic field is present (in the circular area shown), but the particles move essentially classically

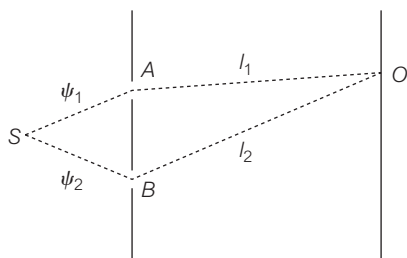


Figure 18.21. Path length geometry for classical wave interference.

³¹ See Aharonov and Bohm (1957).

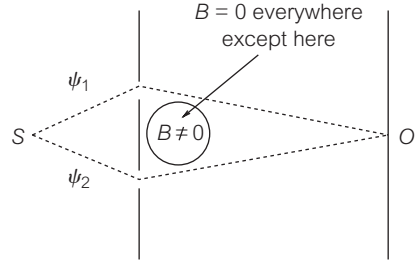


Figure 18.22. Schematic representation of Aharonov–Bohm experiments.

in field-free regions along the paths SAO and SBO . The geometrical path length along SAO and SBO are identical, but there can still be a phase difference in the two paths due to a variation in the vector potential traversed, despite the fact that neither particle is in a magnetic field. Similarly to Eqn. (18.180), we have

$$\psi_{tot}^O = \psi_1^O + \psi_2^O = \psi_1^S e^{i\theta_1} + \psi_2^S e^{i\theta_2} = \left(\psi_1^S + \psi_2^S e^{i(\theta_2 - \theta_1)} \right) e^{i\theta_1} \quad (18.182)$$

where

$$\theta_1 = \frac{q}{\hbar} \int_{SAO}^{\mathbf{r}} d\mathbf{r}' \cdot \mathbf{A}(\mathbf{r}', t) \quad \text{and} \quad \theta_2 = \frac{q}{\hbar} \int_{SBO}^{\mathbf{r}} d\mathbf{r}' \cdot \mathbf{A}(\mathbf{r}', t) \quad (18.183)$$

The phase difference which can give rise to interference effects can be written as

$$\begin{aligned} \Delta\theta &= \theta_2 - \theta_1 \\ &= \frac{q}{\hbar} \left(\int_{SAO}^{\mathbf{r}} d\mathbf{r}' \cdot \mathbf{A}(\mathbf{r}', t) - \int_{SBO}^{\mathbf{r}} d\mathbf{r}' \cdot \mathbf{A}(\mathbf{r}', t) \right) \\ &= \frac{q}{\hbar} \oint_{SAOBS} d\mathbf{r}' \cdot \mathbf{A}(\mathbf{r}', t) \end{aligned} \quad (18.184)$$

The line integral of \mathbf{A} around the closed path $SAOBS$ can be rewritten using Stoke's theorem in the form of an area integral

$$\Delta\theta = \frac{q}{\hbar} \int_{\text{area}} \nabla' \times \mathbf{A}(\mathbf{r}', t) \cdot d\mathbf{S} = \frac{q}{\hbar} \int_{\text{area}} \mathbf{B} \cdot d\mathbf{S} = \frac{q}{\hbar} \Phi_B \quad (18.185)$$

where Φ_B is the magnetic flux enclosed by the path. This result is striking as it says that

- The phase of the wavefunction does depend on the (gauge-dependent) vector potential \mathbf{A} in a nontrivial way, even in a region where the physical magnetic field vanishes.
- That phase dependence gives rise to observable interference effects.
- The quantity which actually determines the interference pattern, however, is the enclosed magnetic flux, which is a perfectly gauge invariant quantity.

The experimental verification of this prediction was first performed³² by the use of a micron size iron “whisker” containing the field.

18.9 Questions and Problems

- Q18.1. The word “gauge” is most often associated with a standard or scale of measurement, especially of length. Examples include the thickness of wire, the distance between rails of a railway, and the like. How do you think that such a word came to be applied to the transformations of EM fields in Eqns (18.30) and (18.31)? Any good encyclopedia of science (or the web) may provide the historical background.
- Q18.2. In the discussion of the Stark effect, we found that only states with $m = 0$ contributed to the second-order perturbation theory result. Discuss why this should be so, concentrating on the relationship between the perturbation eE_0z and \hat{L}_z .
- Q18.3. Suppose one wished to calculate the Stark shifts for a hydrogen-like ion. How would form of the perturbation change, and what additional approximations, if any, would one have to make? By what factor would the polarizability for the ground state of such ions change, that is, how would Eqn. (18.78) scale with Z ? How would you attempt to evaluate the polarizability of the helium atom?
- Q18.4. Evaluate the “generalized slope” of the best fit line in log–log plot in Fig. 17.15, and discuss the relationship between polarizability and ionization potential that it implies.
- Q18.5. Develop a sports metaphor for multiple photon ionization by considering how a single individual might not be able to throw (kick, hit) a ball through a given long distance, but that a team of individuals might, and how the likelihood would scale with team size.
- Q18.6. One is often told that a microwave oven works along these lines:
- A time-dependent external electric field interacts with the permanent dipole moment of the water molecules in the food causing them to rotate.
 - The resulting rotational kinetic energy is transferred to neighboring molecules, resulting in the desired increase in overall temperature.

For an oven operating at $f \approx 2.5$ GHz, evaluate the energy of a single microwave photon, and compare it to the minimum energy necessary to excite a typical rotational state as discussed in Section 16.2.2. Based on your answer, decide whether the microwave oven is a classical or a quantum device, and then explain briefly how you think it works.

³² See Chambers (1960).

- Q18.7.** The energy levels of the Rydberg states of lithium shown in Fig. 18.12 are labeled on the vertical axis as “Energy (cm^{-1})”. Show that the numerical values given there corresponding to the $n = 29, 28, 27$ states (for vanishing magnetic field) are consistent with what you know about the Balmer formula.
- Q18.8.** What kind of spectroscopic evidence would show that sunspots are correlated with regions of high magnetic field activity?
- Q18.9.** It was mentioned that the 21 cm hyperfine line of hydrogen might be useful as a “universal” standard of frequency. The Pioneer-10 and -11 spacecraft carried engraved plaques which were designed to be the first material artifacts of mankind designed to escape the solar system carrying a message (as opposed, for example, to EM transmissions). These “cosmic greeting cards” specified all distances and times in terms of this frequency.³³ What do you think the message consisted of? What would you have included in such a message? What fundamental concepts in physics would you think are important and possible to communicate to another intelligent civilization?
- Q18.10.** How would one use NMR phenomena as the basis for an imaging technique as in MRI? How does one know where in the body the RF photons from the external field were absorbed?
- P18.1. Gravitational versus electrostatic forces.** (a) Calculate the ratio of the gravitational force to the electrostatic force between a proton and an electron.
- (b) A rough limit on any possible difference between the magnitudes of the electron and proton charges can be inferred from cosmology; we will use the notation $|Q_p + Q_e| = \Delta e$ for any such difference. The electrostatic repulsion between two (supposedly neutral) hydrogen atoms could, in principle, overwhelm their gravitational attraction if Δe were too large; such a repulsion could make the universe expand at a rate which is much larger than observed. Show that demanding that any such repulsive force be less than 10% of the mutual gravitational attraction of two hydrogen atoms implies that $\Delta e/e \lesssim 10^{-18}$.
- P18.2. Neutron wave optics.**³⁴ (a) What is the wavelength λ and wave number k for thermal neutrons, that is, ones for which $E = p^2/2m = k_B T/2$ with $T = 300 \text{ K}$?
- (b) The behavior of neutrons in matter can be characterized by an effective *index of refraction* given by

$$n = \sqrt{1 - \frac{4\pi N b}{k^2}} \quad (18.186)$$

where N is the number of nuclei per unit volume and b is the so-called *scattering length*. Estimate the quantity $1 - n$ for thermal neutrons if $b \approx 5 F$ and

³³ For a brief review, see Robinett (2001).

³⁴ For a review of applications of neutron scattering to solid state physics, see Dobbrzynski and Blinowski (1994).

$$N \approx 10^{29} \text{ m}^{-3}.$$

(c) Using (b), estimate the “glancing angle” at which thermal neutrons incident on matter will experience total internal reflection. Can you imagine how this effect is used to provide a “clean” neutron beam far from a “dirty” reactor? Hint: How do fiber optic cables work?

(d) Neutrons are projected horizontally from a height H with velocity v and follow parabolic trajectories. Show that the critical height, H_c , at which all of the neutrons “skid” along the surface because of internal reflection is given by

$$H_c = \frac{2\pi m_n^2}{g\hbar^2} Nb \quad (18.187)$$

This effect is the basis for the *neutron-gravity refractometer*.

P18.3. Photon properties.

(a) How many photons per second enter your eye from a 100 W lightbulb one meter away; estimate the size of your iris in such a situation.

(b) What is the number density of photons (n_γ/m^3) one kilometer away from a 50,000 W AM radio station operating at $f = 1000$ kHz; assume that the antenna radiates its energy uniformly (which is not a very good approximation.)

(c) What is the radiation pressure (in N/m²) due to sunlight at the earth’s surface; the solar luminosity is roughly 4×10^{26} W. Compare it to atmospheric pressure. Why do you “feel” the sunlight’s energy but not its momentum?

(d) Solar radiation pressure is responsible for “sweeping” the solar system clean of dust particles below a certain size; estimate the radius of dust particles for which the radiation force overcomes the gravitational attraction of the sun. Is this pressure large enough to “push” anything else around, even in space? (See Clarke (1972) and Wright (1992) for some ideas on this subject.)

P18.4. Show that the vector potential for a uniform field in an arbitrary direction, \mathbf{B}_0 , can be written as $\mathbf{A} = -\mathbf{r} \times \mathbf{B}_0/2$.

P18.5. Show that the combined effects of a gauge transformation on the Hamiltonian via Eqn. (18.41) and on the wavefunction via Eqn. (18.42) reproduces the original Schrödinger equation.

P18.6. For a system described by the Hamiltonian

$$\hat{H} = \frac{1}{2\mu} (\mathbf{p} - q\mathbf{A})^2 \quad (18.188)$$

show that the probability current which satisfies the equation of continuity,

$$\frac{\partial}{\partial t} (\psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)) + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = 0 \quad (18.189)$$

is given by

$$\mathbf{J} = \frac{\hbar}{2\mu i} \left(\psi^* \nabla \psi - \psi \nabla \psi^* - \frac{2iq}{\hbar} \mathbf{A} |\psi|^2 \right) \quad (18.190)$$

P18.7. EM potentials, quantum mechanics, and the Lorentz force.

(a) Using the quantum mechanical Hamiltonian in Eqn. (18.39), show that the time-dependence of the expectation value of position is given by

$$m \frac{d}{dt} \langle x \rangle_t = \langle \hat{p}_x - qA_x \rangle_t \equiv \langle \hat{\Pi}_x \rangle_t \quad (18.191)$$

Hint: Use the general result in Eqn. (12.88), but note that since x does not depend explicitly on time, only the commutator term is relevant. The combination $\hat{\Pi} = \hat{\mathbf{p}} - \mathbf{A}$ is the appropriate mechanical momentum operator for problems with EM potentials.

(b) This last result seems to suggest that $d\langle x \rangle_t/dt$ might depend on the particular vector potential, \mathbf{A} , used to evaluate the expectation value. First show that $\langle x \rangle_t$ is invariant under gauge transformations given by Eqns (18.30), (18.31), and especially (18.42), because $|\psi'|^2 = |\psi|^2$. Then show that $d\langle x \rangle_t/dt$ is also invariant, but in a more interesting way.

(c) Show that the Hamiltonian can then be written in the form

$$\hat{H} = \frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A})^2 + q\phi \equiv \frac{\hat{\Pi}^2}{2m} + q\phi \quad (18.192)$$

but examine the commutation relations among the $\hat{\Pi}_{x,y,z}$. For example, show that $[\hat{\Pi}_x, \hat{\Pi}_y] = iq\hbar B_z$.

(d) Finally, show that

$$m \frac{d^2}{dt^2} \langle x \rangle_t = \frac{d}{dt} \langle \hat{p}_x - qA_x \rangle_t = q \left\langle E_x + \frac{1}{2m} \left\{ (\hat{\Pi} \times \mathbf{B})_x - (\mathbf{B} \times \hat{\Pi})_x \right\} \right\rangle_t \quad (18.193)$$

using the definitions of the \mathbf{E} and \mathbf{B} fields in terms of the potentials in Eqns (18.28) and (18.29). Hint: In this case, the vector potential \mathbf{A} *does* depend on time, so include both terms in Eqn. (12.88). Discuss why this is the appropriate generalization of the classical Lorentz force equation. Repeat part (b) to show that $d^2\langle x \rangle_t/dt^2$ is invariant under gauge transformations.

P18.8. Second-order Stark shift.

(a) Evaluate the contribution to the second-order Stark shift for hydrogen in Eqn. (18.72) from the $n = 2$ and $n = 3$ states. How much of the *total* second-order shift in Eqn. (18.78) is due to these two states?

(b) One can show that the contribution from each (bound) $\psi_{n,1,0}$ state to the ground state Stark shift in Eqn. 18.72) is given by

$$- \left[\frac{(2n)^9 (n-1)^{2n-6}}{3(n+1)^{2n+6}} \right] \frac{E_0^2 a_0^3}{K} \quad (18.194)$$

Compare your answers to part (a) for $n = 2, 3$ with this general expression. Write a short computer program to sum these contributions until you feel you have converged to a stable answer. What fraction of the *total* shift is due to the bound states alone?

(c) Evaluate the matrix element in Eqn. (18.85).

P18.9. Stark effect for electric field in arbitrary direction. In Section 18.4.2 we considered a uniform electric field in a specific direction, $V(\mathbf{r}) = e\mathcal{E}z$. Suppose the field is in an arbitrary direction, namely, $V(\mathbf{r}) = e\mathcal{E} \cdot \mathbf{r}$ where $\mathcal{E} = (\mathcal{E}_x, \mathcal{E}_y, \mathcal{E}_z)$.

(a) Why must the results for the first- and second-order shifts for the ground state be the same as in Eqns (18.71) and (18.78)?

(b) Show explicitly that this is so. Use the relations

$$x = -\sqrt{\frac{2\pi}{3}} r (Y_{1,1}(\theta, \phi) - Y_{1,-1}(\theta, \phi)) \quad (18.195)$$

$$y = i\sqrt{\frac{2\pi}{3}} r (Y_{1,1}(\theta, \phi) + Y_{1,-1}(\theta, \phi)) \quad (18.196)$$

(c) What happens to the shifts in the first excited states? Is the pattern of splitting the same? Are the linear combinations which split the same?

(d) Explicitly work out the splitting of the first excited state for a uniform field of the form $V(\mathbf{r}) = e\mathcal{E}z$

P18.10. Variational method for Stark effect. Use the variational method to estimate the ground state energy for the Stark Hamiltonian $\hat{H} = \hat{H}_{\text{Coul}} + e\mathcal{E}z$ using the trial wavefunction

$$\psi(r) = \cos(\theta) \psi_{1,0,0}(\mathbf{r}) + \sin(\theta) \psi_{2,1,0}(\mathbf{r}) \quad (18.197)$$

with θ as the variational parameter. Evaluate the energy shift due to the electric field by calculating $\Delta E = E(\mathcal{E} \neq 0) - E(\mathcal{E} = 0)$, show that it is proportional to \mathcal{E}^2 , and evaluate α using Eqn. (18.67).

P18.11. Work out the pattern of level splittings for the Stark effect for the $n = 3$ case of hydrogen. You will have to consider the nine possible states corresponding to s, p, d states.

P18.12. Stark effect for a three-dimensional harmonic oscillator.

(a) Consider a three-dimensional isotropic oscillator with potential

$$V(x, y, z) = \frac{1}{2}K(x^2 + y^2 + z^2) \quad (18.198)$$

with an external field given by $V(z) = e\mathcal{E}z$. Show that the energy eigenvalues and wavefunctions can be obtained exactly by generalizing P9.9.

(b) Use your results to obtain the polarization, α , for the ground state.

(c) What happens to the first and second excited state? Is there a linear Stark effect for any set of levels? Give a reason for your answer.

(d) If the oscillator is nonisotropic, so that

$$V(x, y, z) = \frac{1}{2}(K_x x^2 + K_y y^2 + K_z z^2) \quad (18.199)$$

show that the problem is still soluble exactly. Show that the energy shift due to the external field in an arbitrary direction, $V(\mathbf{r}) = e\mathbf{E} \cdot \mathbf{r}$, can be written in the form

$$\Delta E = -\frac{1}{2}\boldsymbol{\mathcal{E}} \cdot \boldsymbol{\alpha} \cdot \boldsymbol{\mathcal{E}} \quad (18.200)$$

where $\boldsymbol{\alpha}$ is now a *polarization tensor*.

- P18.13.** Consider a particle of mass m and charge e in a three-dimensional cubical box of side L . Estimate the dipole polarizability of the particle in the ground state. Is there a first-order Stark effect for the first excited state (which is triply degenerate)?
- P18.14.** Use the data in Fig. 18.6 to estimate the polarizability of the (nondegenerate) $15s$ state. Compare the value you obtain with the ground state value given by $\alpha = 9a_0^3/2K$ and explain any differences.
- P18.15.** Use the tunneling formula for field emission in Section 11.4.1 as a very rough guide to estimate the lifetime, due to tunneling, of the electron in a hydrogen atom placed in an external electric field \mathcal{E} . (This amounts to neglecting the effect of the Coulomb attraction.) Show that this gives

$$\tau_{\text{tunnel}} \sim \tau_H \exp\left(\frac{4\sqrt{2}}{3} \sqrt{\frac{m_e W}{\hbar^2}} \frac{W}{e\mathcal{E}}\right) \quad (18.201)$$

where τ_H is a characteristic atomic timescale. Obtain a numerical estimate of the lifetime by using $W \sim E_1 \sim 13.6 \text{ eV}$, $\mathcal{E} = 2000 \text{ V/m}$, and τ_H as the classical orbital period for hydrogen.

- P18.16.** An electron in an external electric field has the potential energy function

$$V(\mathbf{r}) = -\frac{Ke^2}{r} - e\mathcal{E}z \quad (18.202)$$

Use this to argue that a Rydberg atom in a state of effective quantum number n^* will be ionized by an external field of strength

$$\mathcal{E} = \frac{E_c}{16(n^*)^4} \quad (18.203)$$

where E_c is the “typical” atomic field in Eqn. (18.22).

P18.17. Multiphoton ionization.

- (a) Imagine that the values below have been obtained in an experiment looking for multiphoton ionization of hydrogen atoms from their ground state; they give the yield of emitted electrons versus the laser intensity, I .

| I (TW/cm ²) | Electrons (10^{10}) |
|---------------------------|-------------------------|
| 4 | 3 |
| 7 | 18 |
| 20 | 300 |
| 60 | 8000 |

From the “data” (which is not from any particular experiment), estimate the value of k in a power-law fit to the data of the form

$$\text{electron yield} \propto (\text{laser power})^k \quad (18.204)$$

- (b) Estimate how many photons are required for multiphoton ionization of hydrogen if the photon wavelength used was $\lambda = 2480\text{\AA}$ and compare to your answer in part (a).

P18.18. Uniform magnetic field in different gauges. The consider the problem of a charged particle in a uniform field in the $+z$ direction described by the vector potential $\mathbf{A} = (0, xB_0, 0)$.

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

$$\hat{H} = \frac{\hat{p}_x^2}{2\mu} + \frac{(qB_0)^2}{2\mu}x^2 + \frac{\hat{p}_y^2}{2\mu} - \frac{qB_0}{\mu}\hat{p}_y + \frac{\hat{p}_z^2}{2\mu} \quad (18.205)$$

- (b) Try a solution of the form

$$\psi(x, y, z) = X(x) e^{i(k_y y + k_z z)} \quad (18.206)$$

and show that the equation for $X(x)$ reduces to that for a shifted harmonic oscillator in as P9.9.

- (c) Find the quantized energy eigenvalues, and show that they give the same spectrum as Eqn. (18.100) with the same degeneracy.
- (d) The wavefunctions corresponding to this solution, and those for the case where $\mathbf{A} = B_0/2(-y, x, 0)$ should be related by a simple phase, as in Eqn. (18.42). Is it easy to see this relationship?

P18.19. Two-dimensional harmonic oscillator in a magnetic field. A particle of mass μ and charge $q > 0$ moves in a plane subject to a two-dimensional isotropic harmonic oscillator potential, $V(x, y) = K(x^2 + y^2)/2$. A uniform magnetic field in the $+z$ direction is also applied. Using the methods of Section 18.5, find the allowed energy eigenvalues.

P18.20. Velocity selectors. The motion of a charged particle in combined (but constant) electric \mathbf{E}_0 and magnetic \mathbf{B}_0 fields can be quite complicated in general. For a particle with initial velocity given by

$$\mathbf{v}_{\text{cross}} = \frac{\mathbf{E}_0 \times \mathbf{B}_0}{|\mathbf{B}_0|^2} \quad (18.207)$$

the electric and magnetic forces cancel, and the particle moves in a straight line trajectory. This is the basis for a classical “velocity selector.” Solve for the energy eigenvalue spectrum and wavefunctions for the quantum version of this problem for uniform electric and magnetic fields in the $-x$ and $+z$ directions, respectively; use the potentials

$$\phi(\mathbf{r}) = E_0 x \quad \text{and} \quad \mathbf{A}(\mathbf{r}) = (0, xB_0, 0) \quad (18.208)$$

Discuss to what extent any “velocity selection effect” is still present in the quantum system.

P18.21. Hydrogen atom in a different gauge. One almost always solves the hydrogen atom problem by (implicitly) using the gauge

$$\phi(\mathbf{r}, t) = \frac{Ke}{r} \quad \text{and} \quad \mathbf{A}(\mathbf{r}, t) = 0 \quad (18.209)$$

to describe the Coulomb field of the proton.

(a) Show that this configuration can also be derived by the potentials

$$\phi(\mathbf{r}, t) = 0 \quad \text{and} \quad \mathbf{A}(\mathbf{r}, t) = -\frac{Ke}{r^2} \mathbf{r} t \quad (18.210)$$

and find the gauge function $f(\mathbf{r}, t)$ which connects it to the choice in (a).

(b) Set up the corresponding Hamiltonian, \hat{H}' , and show explicitly that

$$\psi'_{nlm} = \exp(iqf(\mathbf{r}, t)/\hbar) \psi_{nlm} \quad (18.211)$$

solves the Schrödinger equation in the new gauge.

(c) Discuss how the semi-classical arguments in Section 1.4 are changed in this case; recall that when there is a vector potential, one has $m\mathbf{v} = \mathbf{p} - q\mathbf{A}$.

P18.22. The “anomalous” Zeeman effect. If we include the effect of electron spin and its coupling to the external magnetic field in the Zeeman effect, we need to consider the Hamiltonian

$$\hat{H}_1 = \frac{eB}{2m_e} (\mathbf{L}_z + 2\mathbf{S}_z) \quad (18.212)$$

if we assume that $g_e \approx 2.0$ for the electron. Evaluate the expectation value of this term in the appropriate coupled states in P16.16 for both $j = l + 1/2$ and $j = l - 1/2$ and show that your result can be expressed in the form

$$\Delta E_B^{(1)} = \frac{e\hbar B}{m_e} \frac{m_j}{(2l+1)} \begin{cases} l+1 & \text{for } j = l + 1/2 \\ l & \text{for } j = l - 1/2 \end{cases} \quad (18.213)$$

P18.23. Diamagnetism. The quadratic term in the Hamiltonian for a uniform magnetic field in the standard gauge is given by

$$\frac{e^2 B^2}{8m} (x^2 + y^2) \quad (18.214)$$

(a) Using perturbation theory, show that the corresponding change in energy is

$$\Delta E = \frac{e^2 B^2}{12m} \langle r^2 \rangle \quad (18.215)$$

for spherically symmetric distributions.

(c) Using the definition, $\Delta E = -\boldsymbol{\mu} \cdot \mathbf{B}$ to show that the induced magnetic dipole is

$$\boldsymbol{\mu} = -\frac{e^2 \langle r^2 \rangle}{6m} \mathbf{B} \quad (18.216)$$

An induced magnetic moment opposite to the applied field is termed *diamagnetic*.

(d) Can you derive this result classically by looking at the magnetic moment induced by the charges rotating in the uniform field?

P18.24. Islands of isomerism. Using the nuclear shell energy level diagram in Fig. 18.15, explain why there are many long-lived excited states (so-called *isomers*) for odd-A nuclei for values of either N or Z just below the observed magic numbers. Hint: The rate for radiative decays between two energy levels is proportional to $(kR)^{2l+1}$ where l is the change in angular momentum between the initial and final state; $E_i - E_f = \Delta E = E_\gamma = \hbar k c$ is the photon energy in the transition.

P18.25. Use the results of P16.25 to show that the “magic numbers” for the three-dimensional harmonic oscillator potential are 2, 8, 20, 40, 70, 112, 168, . . . , and so forth. Hint: Use the degeneracy of the spectrum to find the *total* number of spin-1/2 particles which can be accommodated upon the closing of each shell. Since the harmonic oscillator has such a simple spectrum, the shell structure is obvious.

P18.26. (a) Use the expression in Eqn. (18.142) to find the numerical value of ΔE , f , and λ for the energy, frequency, and wavelength of the photon emitted in the $^3S_1 \rightarrow ^3S_0$ hyperfine transition in hydrogen.

(b) Generalize your result to find the corresponding quantity for deuterium. You will need to know that (i) the deuteron magnetic moment is $\mu_D = 0.8798\mu_N$, (ii) $M_D \approx 2m_p$, and (iii) the deuteron spin is $S = 1$.

P18.27. Use the data in Fig. 18.18 to show that the angle θ through which the neutron magnetic moment must precess to change its phase by -1 is $352 \pm 19^\circ$. Use the

following numerical values:

1. The observed “period” in the magnetic path length is $\int_{\text{path}} \mathbf{B} \cdot d\mathbf{l} = 1.44 \pm 0.08 \text{ T} \cdot \text{m}$.
2. The gyromagnetic ratio of the neutron is $g_n = -1.93$.
3. The neutron magnetic moment is given by $g_n e \hbar / 2m_n$.
4. The wavelength of the neutrons used was $1.82 \pm 0.01 \text{ \AA}$. This combined with the mass of the neutron $1.67 \times 10^{-27} \text{ kg}$ gives the neutron velocity v .

P18.28. Flux quantization. Consider an electron moving in the magnetic field geometry of the Aharonov–Bohm effect in Section 18.8. Show that upon making one complete circuit around the localized magnetic field region that the electron wavefunction acquires a phase given by $\exp(i e \Phi / \hbar)$. Use the fact that the electron wavefunction must be single valued at any point to show that this implies that magnetic flux must be quantized, that is,

$$\Phi_n = \left(\frac{2\pi \hbar}{e} \right) n \quad \text{where } n = 0, 1, 2, \dots \quad (18.217)$$

APPENDIX A

Dimensions and MKS-type Units for Mechanics, Electricity and Magnetism, and Thermal Physics

Many introductory physics textbooks begin with a discussion of the dimensions and the small number of dimensionful units necessary to fully describe physical quantities. In this Appendix we briefly review and collect some of the dimensional equivalences which readers may find useful in solving problems or checking derivations. We then collect the values of important physical constants, in these units, in Appendix B.

For purely mechanical systems, the standard MKS set of units, using the meter (m), kilogram (kg), and second (s) as basic dimensionful quantities, is presumably familiar. We collect below some of the more important connections.

The extension to include a fundamental dimension related to electromagnetism (advocated by Giorgi, and included in the MKSA system) uses the Ampere (A) as the basic quantity, but for purposes of simplicity, we will here use the basic unit of charge, the Coulomb (C), noting that the substitution $C = A \cdot s$ allows one to translate. For thermal physics problems, the degree Kelvin (or kelvin, here written simply as K) is used. We will not require the two additional basic units (candela and mole) used in the Système Internationale (SI) d'Unités.

For completeness, we note that many of the standard electromagnetic relationships, in these units, are given through the textbook, namely

- the Coulomb force and potential are given by Eqns (1.28) and (1.29), respectively the Lorentz force law is given by Eqn. (18.5),
- Maxwell's equations are given by Eqns. (18.6)–(18.9),
- the Poynting vector is given by Eqn. (18.19),
- the scalar (ϕ) and vector (\mathbf{A}) potentials are given by Eqns (18.28) and (18.29)

Table A.1. Dimensions for quantities in mechanics

| Quantity | Symbol (Name) | Units |
|-------------------|-------------------|------------------------|
| Mass | | kg |
| Length | | m |
| Time | | s |
| Speed | v | m/s |
| Acceleration | a | m/s ² |
| Force | F (Newton) | kg · m/s ² |
| Energy | E (Joule) | N · m or J |
| Power | P (Watt) | J/s or N · m/s |
| Momentum | p | kg · m/s or N/s |
| Moment of inertia | I | kg · m ² |
| Angular momentum | L | kg · m ² /s |
| Wavelength | λ | m |
| Wave number | k, κ | 1/m |
| Frequency | f (Hertz) | 1/s |
| Angular frequency | $\omega = 2\pi f$ | rad/s |

Table A.2. Dimensions of quantities in electromagnetism and thermal physics

| Quantity | Symbol (Name) | Units |
|----------------------------|----------------|---|
| Charge | q (Coulomb) | C |
| Current | I (Ampere) | A = C/s |
| Electric potential | V (Volt) | J/C or V |
| Electric field | E | N/C or V/m |
| Electric dipole moment | p | C · m |
| Electric quadrupole moment | Q | C · m ² |
| Electric permittivity | ϵ | C ² /(N · m ²) |
| Capacitance | F (farad) | C/V or C ² /J |
| Resistance | Ω (Ohm) | V/A or J · s/C ² |
| Electrical resistivity | ρ | $\Omega \cdot \text{m}$ or J · s · m/C ² |
| Electrical conductivity | σ | A/(m · V) or C ² /(J · m · s) |
| Magnetic field (induction) | T (Tesla) | N · s/(C · m) |
| Magnetic permeability | μ | N · s ² /C ² |
| Magnetic flux | Φ_B | Tesla · m ² or N · m · /C ² |
| Magnetic dipole moment | m | A · m ² or C · m ² /s |
| Magnetic permeability | μ | N · s ² /C ² |
| Auxiliary magnetic field | H | A/m or C/(m · s) |
| Poynting vector | $E \times H$ | J/(m ² · s) |
| Degree (absolute scale) | Kelvin | K |

A.1 Problems

- PA.1 Dimensions for familiar quantities I.** Add entries to Table A.1 for (a) pressure, (b) density, and (c) torque.
- PA.2 Dimensions for familiar quantities II.** Add entries to Table A.2 for (a) thermal conductivity, (b) specific heat, and (c) entropy. (You may have to look up some definitions for these.)

APPENDIX B

Physical Constants, Gaussian Integrals, and the Greek Alphabet

B.1 Physical Constants

We collect values of various physical constants used in the text. For dimensionful quantities involved in electricity and magnetism, we consistently use the same MKSA or SI (“Système International”) units, as in Appendix A, unless specifically noted. In addition to the usual units of mass (*kg*), length (*m*), and time (*s*), for simplicity, we often express less familiar dimensionful quantities in terms of force (Newton, *N*), energy (Joule, *J*), and charge (Coulomb, *C*).

| | |
|----------------------------|---|
| Planck’s constant | $\hbar = 1.055 \times 10^{-34} \text{ J} \cdot \text{s}$ $= 6.582 \times 10^{-16} \text{ eV} \cdot \text{s}$ $h = 2\pi\hbar = 6.626 \times 10^{-34} \text{ J} \cdot \text{s}$ |
| speed of light (in vacuum) | $c = 2.9979 \times 10^8 \text{ m/s}$ $\hbar c = 1973 \text{ eV} \cdot \text{\AA}$ $= 197.3 \text{ MeV} \cdot \text{F}$ |
| electron mass | $m_e = 9.11 \times 10^{-31} \text{ kg}$ $m_e c^2 = 0.511 \text{ MeV}$ |
| proton mass | $m_p = 1.67 \times 10^{-27} \text{ kg}$ $m_p c^2 = 938.3 \text{ MeV}$ |
| neutron mass | $m_n c^2 = 939.6 \text{ MeV}$ |
| muon rest mass | $m_\mu c^2 = 105.7 \text{ MeV}$ |
| fundamental charge | $e = 1.60 \times 10^{-19} \text{ C}$ |
| electric permittivity | $\epsilon_0 = 8.85 \times 10^{-12} \text{ C}^2/(\text{N} \cdot \text{m}^2)$ $K = 1/4\pi\epsilon_0 = 8.98 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2$ |

(Continued)

| | |
|-------------------------|---|
| | $Ke^2 = 2.30 \times 10^{-28} \text{ J} \cdot \text{m}$ |
| | $= 14.4 \text{ eV} \cdot \text{\AA}$ |
| | $= 1.44 \text{ MeV} \cdot \text{F}$ |
| fine structure constant | $\alpha = Ke^2 \hbar c = 1/137.0$ |
| permeability constant | $\mu_0 = 4\pi \times 10^{-7} \text{ N} \cdot \text{s}^2/\text{C}^2$ |
| flux quantum | $\Phi_B = 4.14 \times 10^{-15} \text{ T} \cdot \text{m}^2$ |
| Boltzmann constant | $k_B = 1.38 \times 10^{-23} \text{ J}/(\text{molecule} \cdot \text{K})$ |
| | $= 8.617 \times 10^{-5} \text{ MeV/K}$ |
| thermal energy at | $k_B T = 1/39 \text{ eV}$ |
| $T = 300 \text{ K}$ | |
| Avogadro constant | $N_0 = 6.02 \times 10^{23} \text{ molecules/mole}$ |
| Gas constant | $R = N_0 k_B = 8.31 \text{ J}/(\text{mole} \cdot \text{K})$ |
| Bohr radius | $a_0 = 0.53 \text{ \AA}$ |
| Rydberg constant | $R_\infty = 1.10 \times 10^7 \text{ m}^{-1}$ |
| Rydberg energy | $E_0 = m_e c^2 \alpha^2 / 2 = 13.6 \text{ eV}$ |
| Gravitational constant | $G = 6.67 \times 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2$ |
| solar mass | $M_\odot = 1.99 \times 10^{30} \text{ kg}$ |
| earth mass | $M_e = 5.98 \times 10^{24} \text{ kg}$ |
| moon mass | $M_m = 7.36 \times 10^{22} \text{ kg}$ |
| mean earth-sun distance | $AU = 1.50 \times 10^{11} \text{ m}$ |

Some useful conversion factors are:

| |
|--|
| $1 \text{ \AA} = 10^{-10} \text{ m}$ |
| $1 \text{ F} = 10^{-15} \text{ m}$ |
| $1 \text{ ly (lightyear)} = 9.46 \times 10^{15} \text{ m}$ |
| $1 \text{ pc (parsec)} = 3.09 \times 10^{16} \text{ m}$ |
| $1 \text{ eV} = 1.69 \times 10^{-19} \text{ J}$ |
| $1 \text{ barn} = 10^{-28} \text{ m}^2$ |
| $1 \text{ G (Gauss)} = 10^{-4} \text{ T (Tesla)}$ |

Some often used prefixes for powers of ten are:

| | | | | | |
|-----|------|-----------|-------|-------|------------|
| P | peta | 10^{15} | m | milli | 10^{-3} |
| T | tera | 10^{12} | μ | micro | 10^{-6} |
| G | giga | 10^9 | n | nano | 10^{-9} |
| M | mega | 10^6 | p | pico | 10^{-12} |
| k | kilo | 10^3 | f | femto | 10^{-15} |

B.2 The Greek Alphabet

| | | | | | |
|---------|-----------|------------|---------|------------|------------|
| Alpha | <i>A</i> | α | Nu | <i>N</i> | ν |
| Beta | <i>B</i> | β | Xi | Ξ | ξ |
| Gamma | Γ | γ | Omicron | <i>O</i> | <i>o</i> |
| Delta | Δ | δ | Pi | Π | π |
| Epsilon | <i>E</i> | ϵ | Rho | <i>R</i> | ρ |
| Zeta | <i>Z</i> | ζ | Sigma | Σ | σ |
| Eta | <i>H</i> | η | Tau | <i>T</i> | τ |
| Theta | Θ | θ | Upsilon | Υ | υ |
| Iota | <i>I</i> | ι | Phi | Φ | ϕ |
| Kappa | <i>K</i> | κ | Chi | <i>X</i> | χ |
| Lambda | Λ | λ | Psi | Ψ | ψ |
| Mu | <i>M</i> | μ | Omega | Ω | ω |

B.3 Gaussian Probability Distribution

Finding the probability that a variable represented by a Gaussian probability density with mean value μ and standard deviation σ will have a value in some finite region (a, b) requires the evaluation of the “area under the curve” given by

$$\text{Prob}[x \in (a, b)] = \int_a^b dx \, P(x; \mu, \sigma) \tag{B.1}$$

where

$$P(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{B.2}$$

Any such problem can be “standardized” in terms of a dimensionless variable by writing

$$z = \frac{x - \mu}{\sigma} \tag{B.3}$$

where z measures the “distance” of x away from the mean μ , in units of σ . All of the information required to evaluate such probabilities can be tabulated once and for all in the form of a cumulative probability distribution using this standardized normal random variable by calculating

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt \tag{B.4}$$

This corresponds to the probability of finding the variable t anywhere in the interval $(-\infty, z)$. It is defined such that $F(0) = 0.5$, corresponding to half the probability being on either side of μ . The integral defining $F(z)$ can be evaluated numerically and values are shown in Table B.1. They can be extended to negative values of z by using $F(-z) = 1 - F(z)$. Finally, the probability of finding the standardized variable in the interval (z_{\min}, z_{\max}) is given by

$$\text{Prob}[z \in (z_{\min}, z_{\max})] = F(z_{\max}) - F(z_{\min}) \quad (\text{B.5})$$

Example B.1. Normal distributions

As an example of the use of Table B1, we can calculate the probability that a measurement of a variable corresponding to a Gaussian distribution with $\mu = 7$ and $\sigma = 2$ will find it in the interval $(5.8, 9.4)$. The corresponding range in the standardized variables are

$$z_{\min} = \frac{5.8 - 7}{2} = -0.6 \quad \text{and} \quad z_{\max} = \frac{9.4 - 7}{2} = 1.2 \quad (\text{B.6})$$

The probability in this interval is

$$\begin{aligned} \text{Prob}[z \in (-0.6, 1.2)] &= F(1.2) - F(-0.6) \\ &= 0.8849 - (1.0000 - 0.7257) = 0.6016 \end{aligned} \quad (\text{B.7})$$

or about 60% of the total.

Table B.1. Values of the Cumulative Gaussian Probability Distribution Defined by the Integral in Eqn. (B.4)

| z | $F(z)$ | z | $F(z)$ | z | $F(z)$ |
|-----|--------|-----|--------|-----|--------|
| 0.0 | 0.5000 | 1.0 | 0.8413 | 2.0 | 0.9722 |
| 0.1 | 0.5398 | 1.1 | 0.8643 | 2.1 | 0.9821 |
| 0.2 | 0.5793 | 1.2 | 0.8849 | 2.2 | 0.9861 |
| 0.3 | 0.6179 | 1.3 | 0.9032 | 2.3 | 0.9893 |
| 0.4 | 0.6554 | 1.4 | 0.9192 | 2.4 | 0.9918 |
| 0.5 | 0.6915 | 1.5 | 0.9332 | 2.5 | 0.9938 |
| 0.6 | 0.7257 | 1.6 | 0.9452 | 2.6 | 0.9953 |
| 0.7 | 0.7580 | 1.7 | 0.9554 | 2.7 | 0.9965 |
| 0.8 | 0.7881 | 1.8 | 0.9641 | 2.8 | 0.9974 |
| 0.9 | 0.8159 | 1.9 | 0.9713 | 2.9 | 0.9981 |
| 1.0 | 0.8413 | 2.0 | 0.9772 | 3.0 | 0.9987 |

B.4 Problems

- PB.1. Verify the probabilities of measuring a Gaussian distribution in the intervals $(\mu - \sigma, \mu + \sigma)$, $(\mu - 2\sigma, \mu + 2\sigma)$, and $(\mu - 3\sigma, \mu + 3\sigma)$ as discussed in Example 4.2. How far away from μ (in terms of σ) should one go (symmetrically) on either side to have half of the probability contained under the Gaussian integral?

APPENDIX C

Complex Numbers and Functions

Because the formalism of quantum mechanics requires the manipulation of complex variables, we review here some of the basic definitions and formulae governing their properties. The imaginary unit, i , is defined¹ via $i = \sqrt{-1}$, and a general *complex number* is given by

$$z = a + ib \quad (\text{C.1})$$

where a, b themselves have purely real values. The values a and b are called the *real* and *imaginary* parts of z , respectively; these are often written in the form

$$a = \text{Re}(z) \quad \text{and} \quad b = \text{Im}(z) \quad (\text{C.2})$$

Complex numbers obey standard algebraic relations. For example, if $z_{1,2} = a_{1,2} + ib_{1,2}$, we have addition and subtraction given by

$$z_1 \pm z_2 = (a_1 \pm a_2) + i(b_1 \pm b_2) \quad (\text{C.3})$$

while multiplication is carried out by using the distributive law, giving

$$z_1 z_2 = (a_1 + ib_1)(a_2 + ib_2) = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + a_2 b_1) \quad (\text{C.4})$$

More complicated functions can often be evaluated using series expansions. For example, for θ real we can write

$$\begin{aligned} e^{i\theta} &= 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \dots \\ &= \left(1 - \frac{\theta^2}{2} + \dots\right) + i\left(\theta - \frac{\theta^3}{3!} + \dots\right) \\ e^{i\theta} &= \cos(\theta) + i \sin(\theta) \end{aligned} \quad (\text{C.5})$$

using the series expansions in Appendix D.2.

¹ Engineers often use the notation $j = \sqrt{-1}$.

The *complex conjugate* of a complex number is defined via

$$z^* \equiv a - ib \quad (\text{C.6})$$

that is, by letting $i \rightarrow -i$. A useful relation is

$$|z|^2 \equiv zz^* = (a + ib)(a - ib) = a^2 + b^2 \quad (\text{C.7})$$

which defines the *modulus*, $|z|$, of a complex number via

$$|z| = \sqrt{a^2 + b^2} \quad (\text{C.8})$$

This quantity is the analog of the absolute value of a real number. We can make use of the identity (C.5) to also write

$$a + ib = z \equiv |z|e^{i\theta} = |z|\cos(\theta) + i|z|\sin(\theta) \quad (\text{C.9})$$

where θ is called the *phase* or *argument* of the complex number z ; it is given by

$$\tan(\theta) = \frac{b}{a} = \frac{\text{Im}(z)}{\text{Re}(z)} \quad (\text{C.10})$$

This form for complex numbers is useful as it shows that

$$|z_1 z_2| = |z_1| |z_2| \quad (\text{C.11})$$

We obviously have $z^* = |z|e^{-i\theta}$, so complex conjugation “flips the phase” of z , but keeps its modulus fixed. A complex number with $|z| = 1$, that is, of the form $z = e^{i\theta}$, is often said to be “just a phase”. A general complex number can be represented as a point (or vector) in the complex plane, as shown in Fig. C.1, and addition and subtraction can be given a vector interpretation.

Some useful formulae (for θ real) are

$$\cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{and} \quad \sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i} \quad (\text{C.12})$$

which are easily derived by combining Eqn. (C.5) and its complex conjugate. One also has

$$\cos(i\theta) = \cosh(\theta) \quad \text{and} \quad \sin(i\theta) = i \sinh(\theta) \quad (\text{C.13})$$

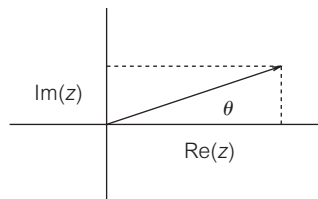


Figure C.1. Representation of a complex number, z , as a vector in the complex plane.

Other familiar trig identities are easily proved using complex notation. For example, one has

$$\begin{aligned}
 & 2 \sin\left(\frac{\alpha + \beta}{2}\right) \cos\left(\frac{\alpha - \beta}{2}\right) \\
 &= 2 \left(\frac{e^{i(\alpha+\beta)/2} - e^{-i(\alpha+\beta)/2}}{2i} \right) \left(\frac{e^{i(\alpha-\beta)/2} + e^{-i(\alpha-\beta)/2}}{2} \right) \\
 &= \sin(\alpha) + \sin(\beta)
 \end{aligned} \tag{C.14}$$

One also has

$$\sin(\alpha \pm \beta) = \sin(\alpha) \cos(\beta) \pm \cos(\alpha) \sin(\beta) \tag{C.15}$$

$$\cos(\alpha \pm \beta) = \cos(\alpha) \cos(\beta) \mp \sin(\alpha) \sin(\beta) \tag{C.16}$$

and the related special cases for double-angle formulae,

$$\sin(2\alpha) = 2 \sin(\alpha) \cos(\alpha) \tag{C.17}$$

$$\cos(2\alpha) = \cos^2(\alpha) - \sin^2(\alpha) \tag{C.18}$$

or

$$\sin^2(\alpha) = \frac{1}{2}[1 - \cos(2\alpha)] \tag{C.19}$$

$$\cos^2(\alpha) = \frac{1}{2}[1 + \cos(2\alpha)] \tag{C.20}$$

C.1 Problems

PC.1. Calculate the result of dividing two complex numbers; specifically, if

$$\operatorname{Re}(w) + i\operatorname{Im}(w) = w = \frac{z_1}{z_2} = \frac{a_1 + ib_1}{a_2 + ib_2} \tag{C.21}$$

find explicit expressions for $\operatorname{Re}(w)$, $\operatorname{Im}(w)$.

PC.2. Find a general expression for the modulus of $|w|$ if

$$w = \exp\left(\frac{1}{a + ib}\right) \tag{C.22}$$

where a, b are real; find a numerical value if $a = 2$ and $b = 1$.

PC.3. Verify the identity

$$|e^{i\alpha} + e^{i\beta}| = \left| 2 \cos\left(\frac{\alpha + \beta}{2}\right) \right| \tag{C.23}$$

if α, β are real. Derive a similar identity for $|e^{i\alpha} - e^{i\beta}|$.

PC.4. If $z = 2 - 3i$, find \sqrt{z} .

PC.5. Show that the following identities hold for any positive integer n ,

$$\begin{aligned} e^{-\pi i n^2 / 2} &= \begin{cases} +1 & \text{for } n \text{ even} \\ -i & \text{for } n \text{ odd} \end{cases} = \left(\frac{1-i}{2} \right) + \left(\frac{1+i}{2} \right) (-1)^n \\ &= \frac{1}{\sqrt{2}} \left[e^{-i\pi/4} + e^{+i\pi/4} (-1)^n \right] \end{aligned} \quad (\text{C.24})$$

APPENDIX D

Integrals, Summations, and Calculus Results

D.1 Integrals

In this section we collect many of the nontrivial indefinite and definite integrals which may be needed for most of the derivations or exercises in the text. Some of them are evaluated using sophisticated methods (such as contour integration, discussed briefly in Section D.4), but we are only interested in using this collection as a reference. The reader is urged to consult other mathematical handbooks or especially to make use of symbolic manipulation programs such as *Mathematica*[®] or *Maple*[®].

We begin by recalling that the simple rule for the differentiation of product functions

$$\frac{d}{dx}[f(x)g(x)] = \frac{df(x)}{dx}g(x) + f(x)\frac{dg(x)}{dx} \quad (\text{D.1})$$

is the basis for the integration by parts (or IBP) method which we use frequently, namely

$$\int_a^b dx \frac{df(x)}{dx} g(x) = - \int_a^b dx f(x) \frac{dg(x)}{dx} + [f(x)g(x)]_a^b \quad (\text{D.2})$$

Some standard indefinite integrals:

$$\int \frac{dx}{x^2 + a^2} = \frac{1}{a} \tan^{-1} \left(\frac{x}{a} \right) \quad (\text{D.3})$$

$$\int \frac{dx}{a^2 - x^2} = \frac{1}{2a} \log \left(\frac{a+x}{a-x} \right) \quad (a^2 > x^2) \quad (\text{D.4})$$

$$\int (\sin(ax)) (\sin(bx)) dx = \frac{\sin(a-b)x}{2(a-b)} - \frac{\sin(a+b)x}{2(a+b)} \quad (a^2 \neq b^2) \quad (\text{D.5})$$

$$\int (\cos(ax)) (\cos(bx)) dx = \frac{\sin(a-b)x}{2(a-b)} + \frac{\sin(a+b)x}{2(a+b)} \quad (a^2 \neq b^2) \quad (\text{D.6})$$

$$\int (\sin(ax)) (\cos(bx)) dx = -\frac{\cos(a-b)x}{2(a-b)} - \frac{\cos(a+b)x}{2(a+b)} \quad (a^2 \neq b^2) \quad (\text{D.7})$$

$$\int dx x \sin(ax) = \frac{1}{a^2} \sin(ax) - \frac{x}{a} \cos(ax) \quad (\text{D.8})$$

$$\int dx x \cos(ax) = \frac{1}{a^2} \cos(ax) + \frac{x}{a} \sin(ax) \quad (\text{D.9})$$

$$\int dx x^2 \sin(ax) = \frac{2x}{a^2} \sin(ax) - \frac{(a^2 x^2 - 2)}{a^3} \cos(ax) \quad (\text{D.10})$$

$$\int dx x^2 \cos(ax) = \frac{2x}{a^2} \cos(ax) + \frac{(a^2 x^2 - 2)}{a^3} \sin(ax) \quad (\text{D.11})$$

$$\begin{aligned} \int dx x^4 \sin(ax) &= \frac{4x(a^2 x^2 - 6)}{a^4} \sin(ax) \\ &\quad - \frac{(a^4 x^4 - 12a^2 x^2 + 24)}{a^5} \cos(ax) \end{aligned} \quad (\text{D.12})$$

$$\begin{aligned} \int dx x^4 \cos(ax) &= \frac{4x(a^2 x^2 - 6)}{a^4} \cos(ax) \\ &\quad + \frac{(a^4 x^4 - 12a^2 x^2 + 24)}{a^5} \sin(ax) \end{aligned} \quad (\text{D.13})$$

$$\int dx x \sin^2(ax) = \frac{x^2}{4} - \frac{x \sin(2ax)}{4a} - \frac{\cos(2ax)}{8a^2} \quad (\text{D.14})$$

$$\int dx x \cos^2(ax) = \frac{x^2}{4} + \frac{x \sin(2ax)}{4a} + \frac{\cos(2ax)}{8a^2} \quad (\text{D.15})$$

$$\int dx x^2 \sin^2(ax) = \frac{x^3}{6} - \left(\frac{x^2}{4a} - \frac{1}{8a^3} \right) \sin(2ax) - \frac{x \cos(2ax)}{4a^2} \quad (\text{D.16})$$

$$\int dx x^2 \cos^2(ax) = \frac{x^3}{6} + \left(\frac{x^2}{4a} - \frac{1}{8a^3} \right) \sin(2ax) + \frac{x \cos(2ax)}{4a^2} \quad (\text{D.17})$$

$$\int e^{ax} dx = \frac{1}{a} e^{ax} \quad (\text{D.18})$$

$$\int x e^{ax} dx = \frac{1}{a^2} (ax - 1) e^{ax} \quad (\text{D.19})$$

$$\int x^2 e^{ax} dx = \frac{1}{a^3} (a^2 x^2 - 2ax + 2) e^{ax} \quad (\text{D.20})$$

Some definite integrals:

$$\int_{-\infty}^{+\infty} \frac{\sin(x)}{x} dx = \pi \quad (\text{D.21})$$

$$\int_{-\infty}^{+\infty} \frac{\sin^2(x)}{x^2} dx = \pi \quad (\text{D.22})$$

$$\int_{-\infty}^{+\infty} \frac{\sin^4(x)}{x^2} dx = \frac{\pi}{2} \quad (\text{D.23})$$

$$\int_{-\infty}^{+\infty} \frac{(1 - \cos(x))}{x^2} dx = \pi \quad (\text{D.24})$$

$$\int_{-\infty}^{+\infty} \frac{(1 - \cos(x))^2}{x^2} dx = \pi \quad (\text{D.25})$$

$$\int_{-\infty}^{+\infty} \frac{\sin(x) \cos(x)}{x} dx = \frac{\pi}{2} \quad (\text{D.26})$$

$$\int_{-\infty}^{+\infty} \frac{\sin(x) \cos(mx)}{x} dx = \begin{cases} 0 & \text{for } |m| > 1 \\ \pi/2 & \text{for } |m| = 1 \\ \pi & \text{for } |m| < 1 \end{cases} \quad (\text{D.27})$$

$$\int_{-\infty}^{+\infty} \frac{\sin(x_1 - x) \sin(x_2 - x)}{(x - x_1)(x - x_2)} dx = \pi \frac{\sin(x_1 - x_2)}{(x_1 - x_2)} \quad (\text{D.28})$$

$$\int_0^{\infty} \frac{\cos(mx)}{x^2 + a^2} dx = \frac{\pi}{2|a|} e^{-|ma|} \quad (\text{D.29})$$

$$\int_0^{\infty} \frac{\cos(mx) \cos(nx)}{x^2 + a^2} dx = \frac{\pi}{a} \left(e^{-|(m-n)a|} + e^{-|(m+n)a|} \right) \quad (\text{D.30})$$

$$\int_0^{\infty} \frac{\sin(mx) \sin(nx)}{x^2 + a^2} dx = \frac{\pi}{a} \left(e^{-|(m-n)a|} - e^{-|(m+n)a|} \right) \quad (\text{D.31})$$

$$\int_0^{\infty} \cos(mx) e^{-ax} dx = \frac{a}{a^2 + m^2} \quad (a > 0) \quad (\text{D.32})$$

$$\int_0^{\infty} \sin(mx) e^{-ax} dx = \frac{m}{a^2 + m^2} \quad (a > 0) \quad (\text{D.33})$$

The following integrals make use of the Euler Gamma function ($\Gamma(z)$), the generalized factorial function, as discussed in Appendix C.9.

$$\int_0^{\infty} x^n e^{-x} dx = n! = \Gamma(n + 1) \quad (\text{D.34})$$

$$\int_0^{\infty} dx x^n e^{-(ax)^m} = \frac{1}{ma^{n+1}} \Gamma\left(\frac{n+1}{m}\right) \quad (\text{D.35})$$

$$\int_0^1 \frac{x^m dx}{\sqrt{1-x^n}} = \frac{\Gamma(1/2)\Gamma((m+1)/n)}{n\Gamma(1/2 + (m+1)/n)} \quad (\text{D.36})$$

$$\int_0^1 x^{m-1} (1-x)^{n-1} dx = \frac{\Gamma(n)\Gamma(n)}{\Gamma(m+n)} \quad (\text{D.37})$$

$$\int_0^{\infty} \frac{x^a dx}{(m+x^b)^c} = \frac{m^{(a+1-bc)/b}}{b} \left[\frac{\Gamma((a+1)/b) \Gamma(c - (a+1)/b)}{\Gamma(c)} \right] \quad (\text{D.38})$$

Integrals containing Gaussian terms of the form $\exp(-ax^2)$ are of special importance and we discuss their evaluation in slightly more detail. The standard trick for the evaluation of the basic integral

$$I \equiv \int_{-\infty}^{+\infty} dx \exp(-x^2) = \sqrt{\pi} \quad (\text{D.39})$$

is to consider

$$\begin{aligned} I^2 &= I \cdot I = \left(\int_{-\infty}^{+\infty} dx \exp(-x^2) \right) \cdot \left(\int_{-\infty}^{+\infty} dy \exp(-y^2) \right) \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dx dy \exp(-x^2 - y^2) \\ &= \int_0^{\infty} \int_0^{2\pi} r dr d\theta \exp(-r^2) \\ &= 2\pi \int_0^{\infty} dr r \exp(-r^2) \\ I^2 &= \pi \end{aligned} \quad (\text{D.40})$$

so that $I = \sqrt{\pi}$. The more general basic integral is

$$I(a) = \int_{-\infty}^{+\infty} dx \exp(-ax^2) = \sqrt{\frac{\pi}{a}} \quad (\text{D.41})$$

and a related one is

$$\begin{aligned} I(a, b) &\equiv \int_{-\infty}^{+\infty} dx \exp(-ax^2 - bx) \\ &= \int_{-\infty}^{+\infty} dx \exp(-a(x^2 + bx/a + b^2/4a^2 - b^2/4a)) \\ &= \exp(b^2/4a) \int_{-\infty}^{+\infty} dx \exp(-a(x + b/a)^2) \end{aligned}$$

$$I(a, b) = \exp(b^2/4a) \sqrt{\frac{\pi}{a}} \quad (\text{D.42})$$

where we have used a standard method of completing the square and shifting variables. One can generalize these expressions further by “differentiating under the integral sign” to obtain

$$\begin{aligned} J(a, b; n) &= \int_{-\infty}^{+\infty} dx x^n \exp(-ax^2 - bx) \\ &= \left(-\frac{\partial}{\partial b}\right)^n I(a, b) = \left(-\frac{\partial}{\partial b}\right)^n \left[\exp(b^2/4a) \sqrt{\frac{\pi}{a}} \right] \end{aligned} \quad (\text{D.43})$$

For example, one has

$$J(a, b; 1) \equiv \int_{-\infty}^{+\infty} x \exp(-ax^2 + bx) dx = \left(\frac{b}{2a}\right) \exp(b^2/4a) \sqrt{\frac{\pi}{a}} \quad (\text{D.44})$$

$$J(a, b; 2) \equiv \int_{-\infty}^{+\infty} x^2 \exp(-ax^2 + bx) dx = \left(\frac{b^2 + 2a}{4a^2}\right) \exp(b^2/4a) \sqrt{\frac{\pi}{a}} \quad (\text{D.45})$$

and so forth. For even values of $n = 2k$ we can also evaluate $J(a, b; n = 2k)$ by using

$$J(a, b; 2k) = \int_{-\infty}^{+\infty} x^{2k} e^{-ax^2 - bx} dx = \left(-\frac{\partial}{\partial a}\right)^k I(a, b) \quad (\text{D.46})$$

Integrals involving $\cos(kx)$ and $\sin(kx)$ terms and Gaussian integrands can also be done by using identities such as $\cos(kx) = [\exp(+ikx) + \exp(-ikx)]/2$ to obtain

$$\int_{-\infty}^{+\infty} e^{-ax^2 - bx} \cos(kx) dx = +\sqrt{\frac{\pi}{a}} e^{(b^2 - k^2)/4a} \cos(kb/2a) \quad (\text{D.47})$$

$$\int_{-\infty}^{+\infty} e^{-ax^2 - bx} \sin(kx) dx = -\sqrt{\frac{\pi}{a}} e^{(b^2 - k^2)/4a} \sin(kb/2a) \quad (\text{D.48})$$

Integrals containing $ax^2 + bx + c$ arise in the study of the classical limit of the hydrogen atom and elsewhere. If we define $X = ax^2 + bx + c$ and $q = 4ac - b^2$, one has

$$\int \frac{dx}{\sqrt{X}} = -\frac{1}{\sqrt{-a}} \sin^{-1} \left(\frac{2ax + b}{\sqrt{-q}} \right) \quad (a < 0) \quad (\text{D.49})$$

$$\int \frac{x dx}{\sqrt{X}} = \frac{\sqrt{X}}{a} - \frac{b}{2a} \int \frac{dx}{\sqrt{X}} \quad (\text{D.50})$$

$$\int \frac{x^2 dx}{\sqrt{X}} = \left(\frac{x}{2a} - \frac{3b}{4a^2} \right) \sqrt{X} + \frac{3b^2 - 4ac}{8a^2} \int \frac{dx}{\sqrt{X}} \quad (\text{D.51})$$

$$\int \frac{x^3 dx}{\sqrt{X}} = \left(\frac{x^2}{3a} - \frac{5bx}{12a^2} + \frac{5b^2}{8a^3} - \frac{2c}{3a^2} \right) \sqrt{X} + \left(\frac{3bc}{4a^2} - \frac{5b^3}{16a^3} \right) \int \frac{dx}{\sqrt{X}} \quad (\text{D.52})$$

$$\int \frac{dx}{x\sqrt{X}} = \frac{1}{\sqrt{-c}} \sin^{-1} \left(\frac{bx + 2c}{|x|\sqrt{-q}} \right) \quad (c < 0) \quad (\text{D.53})$$

D.2 Summations and Series Expansions

We collect here some useful results which evaluate the summations of certain finite and infinite series.

$$\sum_{k=1}^{k=N} x^k = \frac{1 - x^{N+1}}{1 - x} \quad (\text{D.54})$$

$$\sum_{k=1}^{k=N} k = \frac{N(N+1)}{2} \quad (\text{D.55})$$

$$\sum_{k=1}^{k=N} k^2 = \frac{N(N+1)(2N+1)}{6} \quad (\text{D.56})$$

The *Riemann zeta function* is defined via

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots = \sum_{n=1}^{\infty} \frac{1}{n^s} \quad (\text{D.57})$$

Some special cases are

$$\zeta(2) = \frac{\pi^2}{6} \quad \zeta(4) = \frac{\pi^4}{90} \quad \zeta(6) = \frac{\pi^6}{945} \quad \zeta(8) = \frac{\pi^8}{9450} \quad (\text{D.58})$$

One can also show that

$$\zeta_{\text{odd}}(s) \equiv \frac{1}{1} + \frac{1}{3^s} + \frac{1}{5^s} + \cdots = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^s} = \left(1 - \frac{1}{2^s}\right) \zeta(s) \quad (\text{D.59})$$

so that

$$\zeta_{\text{odd}}(2) = \frac{\pi^2}{8} \quad \text{and} \quad \zeta_{\text{odd}}(4) = \frac{\pi^4}{96} \quad (\text{D.60})$$

One also has:

$$S(x) = \sum_{k=1}^{\infty} \frac{1}{((2k-1)^2 - x^2)} = \frac{\pi}{2x} \tan\left(\frac{\pi x}{2}\right) \quad (\text{D.61})$$

$$T(x) = \sum_{k=1}^{\infty} \frac{1}{((2k-1)^2 - x^2)^2} = \frac{\pi}{16x^3} \left[\pi x \sec^2\left(\frac{\pi x}{2}\right) - 2 \tan\left(\frac{\pi x}{2}\right) \right] \quad (\text{D.62})$$

The *Taylor series expansion* of a (well-behaved) function $f(x)$ about the point $x = a$ is given by

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \cdots \\ &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n \end{aligned} \quad (\text{D.63})$$

where

$$f^{(n)}(a) = \left. \frac{d^n f(x)}{dx^n} \right|_{x=a} \quad (\text{D.64})$$

is the n th derivative of $f(x)$ evaluated at $z = a$. Familiar examples include:

$$\begin{aligned} (1 \pm x)^n &= 1 \pm nx + \frac{n(n-1)}{2!}x^2 \pm \frac{n(n-1)(n-2)}{3!}x^3 + \cdots \\ &= \sum_{k=1}^{\infty} (\pm 1)^k \frac{n!}{(n-k)!k!} x^k \quad \text{for } |x| < 1 \end{aligned} \quad (\text{D.65})$$

$$\sqrt{1 \pm x} = 1 \pm \frac{x}{2} - \frac{x^2}{8} \pm \frac{x^3}{16} + \cdots \quad \text{for } |x| < 1 \quad (\text{D.66})$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \text{for all real } x \quad (\text{D.67})$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k} \quad \text{for } -1 < x \leq +1 \quad (\text{D.68})$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^{2k-1}}{(2k-1)!} \quad \text{for all real } x \quad (\text{D.69})$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \quad \text{for all real } x \quad (\text{D.70})$$

$$\tan(x) = x + \frac{x^3}{3} + \frac{2x^5}{15} + \cdots \quad \text{for } |x| < \pi/2 \quad (\text{D.71})$$

A useful tool to investigate the convergence of a series expansion is the *ratio test*. If an infinite summation is defined via

$$S = \sum_{n=0}^{\infty} \rho_n \quad (\text{D.72})$$

the limit of successive ratios is defined via

$$\rho = \lim_{n \rightarrow \infty} \frac{\rho_{n+1}}{\rho_n} \quad (\text{D.73})$$

One then knows that

- The series *converges* (S is finite) if $\rho < 1$,
- The series *diverges* (S is infinite) if $\rho > 1$,
- The test is inconclusive (the series may either converge or diverge) if $\rho = 1$.

If the terms in the series (i) alternate in sign, (ii) decrease in magnitude (each one smaller than the one before it), and the terms approach zero, then the series is known to converge by *Liebniz's theorem*.

It is often useful to recall the definition of the (one-dimensional) integral as the “area under the curve.” The trapezoidal approximation to the area under $f(x)$ in the interval (a, b) is obtained by splitting the interval into N equal pieces of size $h = (b - a)/N$ which gives

$$\int_a^b dx f(x) \approx F_N(a, b) \equiv h \left(\frac{1}{2} f(a) + \sum_{n=1}^{N-1} f(a + nh) + \frac{1}{2} f(b) \right) \quad (\text{D.74})$$

This expression can form the basis for the simplest of numerical integration programs if necessary. The *Euler–Maclaurin formula* describes the difference between these two approximations via

$$\left[\int_a^b dx f(x) \right] - F_N(a, b) = -\frac{B_2}{2!} h^2 f'(x) \Big|_a^b - \frac{B_4}{4!} h^4 f'''(x) \Big|_a^b + \cdots \quad (\text{D.75})$$

where the B_n are the Bernoulli numbers the first few of which are

$$B_0 = 1 \quad B_2 = \frac{1}{6} \quad B_4 = -\frac{1}{30} \quad B_6 = \frac{1}{42} \quad (\text{D.76})$$

D.3 Assorted Calculus Results

The gradient-squared operator or *Laplacian operator* in rectangular (Cartesian), cylindrical (polar), or spherical coordinates is given by

$$\nabla^2 f(x, y, z) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \quad (\text{D.77})$$

$$\nabla^2 f(r, \theta, z) = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2} \quad (\text{D.78})$$

$$\begin{aligned} \nabla^2 f(r, \theta, \phi) = & \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin(\theta)} \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial f}{\partial \theta} \right) \\ & + \frac{1}{r^2 \sin^2(\theta)} \frac{\partial^2 f}{\partial \phi^2} \end{aligned} \quad (\text{D.79})$$

If one changes variables in a multidimensional integral, one must also apply the appropriate transformation in the “infinitesimal measure” as well. For example, if one changes variables via

$$x, y \implies u(x, y), v(x, y) \quad (\text{D.80})$$

then one has the relation

$$du \, dv = J(u, v; x, y) \, dx \, dy = \det \begin{pmatrix} \partial u / \partial x & \partial u / \partial y \\ \partial v / \partial x & \partial v / \partial y \end{pmatrix} dx \, dy \quad (\text{D.81})$$

with similar extensions to more dimensions; the function $J(u, v; x, y)$ is called the *Jacobian* of the transformation. You should be able to use this approach to derive the familiar result that $dx \, dy \, dz = d\mathbf{r} = r^2 dr \sin(\theta) d\theta d\phi$.

D.4 Real Integrals by Contour Integration

Large numbers of useful real integrals, especially ones involving integration over the entire real line, can be done using simple contour integration techniques, making use of complex variables. We *very* briefly review the rudimentary complex analysis and “tricks of the trade” needed to implement many such integrals, leaving detailed discussions to undergraduate texts on mathematical methods.

The basic result from complex analysis which is required is the *residue theorem* which simplifies the evaluation of integrals of complex functions about a closed curve C in the complex plane, using only knowledge of the structure of the poles

(and essential singularities) that are enclosed by C . The appropriate connection is given by

$$\oint_C f(z) dz = 2\pi i \sum_i \mathcal{R}_i \quad (\text{D.82})$$

where for a function $f(z)$ which has a pole of order n at $z = z_0$, the residue, \mathcal{R}_i , is given by

$$\mathcal{R}_i = \frac{1}{(n-1)!} \left\{ \left(\frac{d}{dz} \right)^{n-1} [(z - z_0)^n f(z)] \right\}_{z \rightarrow z_0} \quad (\text{D.83})$$

The closed contour C is assumed to have a counterclockwise orientation, while if it is completed in a clockwise direction, an overall minus sign is added to the right-hand side of Eqn. (D.82). A judicious choice of an appropriate contour is often all that is needed to use the residue theorem to aid in the evaluation of integrals along the real-axis, and we present two exemplary cases.

Simple poles: Consider the real integral

$$I_1 = \int_{-\infty}^{+\infty} \frac{dx}{(1+x^2)^3} \quad (\text{D.84})$$

which is clearly square-integrable and convergent. Generalize this to the complex line integral given by

$$\mathcal{I}_1 = \oint_C \frac{dz}{(1+z^2)^3} \quad (\text{D.85})$$

over the contour shown in Fig. C.2(a), considering the limit that $R \rightarrow \infty$ so that the semicircle C_1 eventually extends to infinity, while the part of C along the real line reproduces I_1 . For those values on the semicircle C_1 , we can write the complex variable z in the form

$$z = Re^{i\theta} \quad \text{so that} \quad dz = iRe^{i\theta} d\theta \quad \text{giving} \quad \frac{dz}{(1+z^2)^3} \rightarrow \frac{ie^{-5i\theta}}{R^5} \quad (\text{D.86})$$

which becomes arbitrarily small as $R \rightarrow \infty$.

On the one hand, the complex integral over C is the sum of the desired real integral and that over the semicircular curve C_1 in the form

$$\mathcal{I}_1 = \int_{-R}^{+R} \frac{dx}{(1+x^2)^3} + \int_{C_1 (R \rightarrow \infty)} \frac{dz}{(1+z^2)^2} \rightarrow I_1 \quad (\text{D.87})$$

in the limit that $R \rightarrow \infty$, since the contribution from C_1 vanishes. We can, however, also evaluate \mathcal{I}_1 using the residue theorem with $f(z) = 1/(1+z^2)^3$,

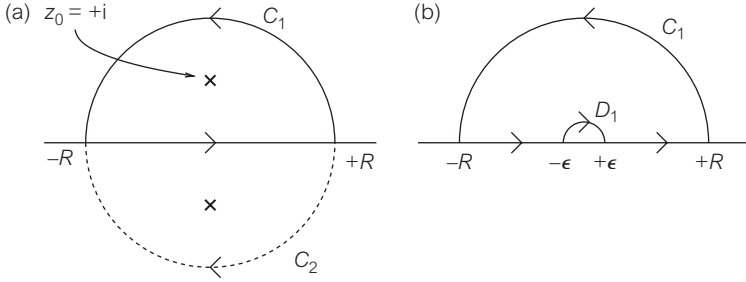


Figure D.1. Contours used in the evaluation of real integrals.

which has simple poles of order 3 at $z_0 = \pm i$. Using the contour C (semicircle C_1 plus real line) in Fig. D.1(a), which encloses the pole at $z_0 = +i$, we find

$$\begin{aligned} I_1 = \mathcal{I}_1 &= 2\pi i \frac{1}{2!} \left\{ \left(\frac{d}{dz} \right)^2 \left[(z-i)^3 \frac{1}{(z-i)^3(z+i)^3} \right] \right\}_{z \rightarrow +i} \\ &= \pi i \left(\frac{(-3)(-4)}{(2i)^5} \right) = \frac{3\pi}{8} \end{aligned} \quad (\text{D.88})$$

If we consider the related contour consisting of the semicircle C_2 (dashed contour) and the real line (enclosing the pole at $z_0 = -i$), we obtain the same result (recall the additional minus sign if the contour orientation is clockwise.)

Deformed contours: Consider the integral

$$I_2 = \int_{-\infty}^{+\infty} \frac{\sin(x)}{x} dx \quad (\text{D.89})$$

which despite appearances is everywhere finite and convergent. The integrand is well behaved at $x = 0$ (since $\lim_{x \rightarrow 0} \sin(x)/x = 1$) and while the large $|x|$ dependence of $1/x$ would yield a logarithmic divergence by itself, the alternation of signs due to the oscillatory $\sin(x)$ gives convergence; think of the integral as an infinite sum of terms (the positive and negative areas defining the area under the integrand) with alternating signs, and of decreasing magnitude, which guarantees so-called *conditional convergence*.

In this case we choose to write I_2 in terms of the imaginary part (Im) of an already complex integral in the form

$$I_2 = \text{Im} \left[\int_{-\infty}^{+\infty} \frac{e^{ix}}{x} dx \right] \equiv \text{Im}[\tilde{I}] \quad (\text{D.90})$$

and use contour integration over the deformed semicircle shown in Fig. C.2(b), where the region near $z = 0$ is treated more carefully, with a smaller semicircle of radius ϵ . The contour integral over C_1 , D_1 and the integrals over the ranges

$(-R, -\epsilon)$ and $(+\epsilon, +R)$ constitute the desired contour C , but since it encloses no poles we have

$$\mathcal{I}_2 \equiv \int_C \frac{e^{iz}}{z} dz = 2\pi i \sum_i \mathcal{R}_i = 0 \quad (\text{D.91})$$

via the residue theorem. The integral \mathcal{I}_2 can also be split up into the four contributions

$$\mathcal{I}_2 = \left[\int_{-R}^{-\epsilon} + \int_{+\epsilon}^{+R} \frac{e^{ix}}{x} dx \right] + \int_{C_1} \frac{e^{iz}}{z} dz + \int_{D_1} \frac{e^{iz}}{z} dz \quad (\text{D.92})$$

eventually taking the twin limits $\epsilon \rightarrow 0$ and $R \rightarrow \infty$. The first two terms give \tilde{I} in this limit, while for C_1 we use the same parameterization as in Eqn. (D.86) and it is easy to show that the contribution is exponentially suppressed as $\exp(-R \sin(\theta))$ as $R \rightarrow \infty$. Finally, for the contribution around D_1 , we use $z = \epsilon \exp(i\theta)$ which gives

$$\int_{D_1} \frac{e^{iz}}{z} dz = \int_{\pi}^0 \frac{e^{i\epsilon e^{i\theta}}}{\epsilon e^{i\theta}} (i\epsilon e^{i\theta} d\theta) \rightarrow i \int_{\pi}^0 d\theta = -i\pi \quad (\text{D.93})$$

as $\epsilon \rightarrow 0$. From Eqns (D.91) and (D.92) we then find that $0 = \mathcal{I}_2 = \tilde{I} + 0 - i\pi$ so that

$$\int_{-\infty}^{+\infty} \frac{e^{ix}}{x} dx = \tilde{I} = i\pi \quad \text{and} \quad \mathcal{I}_2 = \text{Im}[\tilde{I}] = \text{Im}[i\pi] = \pi \quad (\text{D.94})$$

A large number of other similar integrals can be done using the deformed contour shown in Fig. C.2(b). For example, you should be able to show that

$$\int_{-\infty}^{+\infty} \frac{[1 - \cos(y)]}{y^2} dy = \pi \quad (\text{D.95})$$

by taking the real part of a related integral, and extending it to the contour used above.

D.5 Plotting

The functional relationship between two variables is often best exemplified or analyzed (or even discovered in the first place) by plotting the “data” in an appropriate manner. In this section, we briefly recall some of the basics of plotting techniques; because linear relations are easiest to visualize, many standard tricks rely on graphing data in such a way as to yield a straight line.

For variables which are connected by an exponential relation one has

$$\text{exponential: } y = ae^{bx} \implies \ln(y) = \ln(a) + bx \quad (\text{D.96})$$

which suggests that one plot $\ln(y)$ versus x ; this gives a so-called *semilog plot*. A straight-line fit on such a plot implies an exponential relation, and the “generalized slope” is given by

$$b = \frac{(\ln(y_2) - \ln(y_1))}{(x_2 - x_1)} = \frac{\ln(y_2/y_1)}{(x_2 - x_1)} \quad (\text{D.97})$$

The value of a (or $\ln(a)$) plays the role of an “intercept” and can be extracted from any point on the line once b is known; if the point with $x = 0$ is included, then $y(0) = a$ is the most obvious choice.

For power-law relations of the form

$$\text{power-law: } y = cx^d \implies \ln(y) = \ln(c) + d \ln(x) \quad (\text{D.98})$$

it is best to graph $\ln(y)$ versus $\ln(x)$ giving a *log-log plot* where the “generalized slope” is now

$$d = \frac{(\ln(y_2) - \ln(y_1))}{(\ln(x_2) - \ln(x_1))} = \frac{\ln(y_2/y_1)}{\ln(x_2/x_1)} \quad (\text{D.99})$$

D.6 Problems

PD.1. Derive any of the integrals in Eqns (D.5)–(D.7) by using complex exponentials.

PD.2. Derive any of the integrals in Eqns (D.8)–(D.11) using IBP techniques.

PD.3. Derive the integral in Eqn. (D.8) by differentiating both sides of the relation

$$\int \cos(ax) dx = \frac{1}{a} \sin(ax) \quad (\text{D.100})$$

with respect to a .

PD.4. Evaluate $J(a, b; n)$ in Eqn. (D.43) for $n = 2, 4$ by differentiating with respect to b . Then evaluate those two cases using Eqn. (D.46) by differentiating with respect to a and confirm you get the same answers.

PD.5. Derive Eqn. (D.59).

PD.6. Evaluate the integral in Eqn. (D.24) using contour integration.

PD.7. At very low temperatures, the heat capacity (at constant volume) of metals is expected to given by an expression of the form

$$C_V = \gamma T + AT^3 \quad (\text{D.101})$$

Given experimental values for T and $C_V(T)$, what would be the best way to plot the data to confirm such a relation and to most easily extract γ and A ?

APPENDIX E

Special Functions

In this section, we collect some of the well-known properties of many of the special functions considered in this text. Most are quoted here without detailed proofs or derivations, while some have been discussed in a more physical context throughout the book.

E.1 Trigonometric and Exponential Functions

Although they are presumably familiar to all students, we briefly discuss the properties of the trigonometric and exponential functions. Since many of the special functions found in mathematical physics arise as solutions to similar differential equations, it is useful to recall here that:

- The differential equation

$$\frac{d^2 f(x)}{dx^2} = -k^2 f(x) \quad (\text{E.1})$$

has the (conventionally normalized) trig function solutions $f(x) = \sin(kx), \cos(kx)$, while

- The differential equation

$$\frac{d^2 f(x)}{dx^2} = +\kappa^2 f(x) \quad (\text{E.2})$$

has exponential solutions $f(x) = e^{\kappa x}, e^{-\kappa x}$.

The intuitive physical connections of these solutions with the oscillatory motion of a particle near a potential energy minimum (case (E.1)) versus the “runaway” (or damped) motion of a particle moved away from an unstable potential maximum (case (E.2)) can often be generalized to other differential equations to help understand the physical origin of the behavior of the solutions.

In each case above, as with any second-order ordinary differential equation, we obtain two, linearly independent solutions, $f_1(x)$, $f_2(x)$. The most general solution is then obtained by taking a linear combination $a_1f_1(x) + a_2f_2(x)$ and using the boundary conditions (in quantum mechanics) or initial conditions (in classical mechanics) to determine the arbitrary coefficients.

E.2 Airy Functions

The Airy differential equation is written in the form

$$\frac{d^2f(x)}{dx^2} = xf(x) \quad (\text{E.3})$$

Here we note the following:

- This problem is related to the quantum version of a particle moving under the influence of a uniform force.
- It also appears in the context of matching WKB-type (Chapter 10) solutions near classical turning points, where the potential energy function can be approximated (locally) as a linear potential.
- Using Eqn. (E.2) as a model, for $x > 0$, we expect exponentially damped or growing solutions; these should be consistent with the tunneling wavefunctions of Section 8.2.2.
- For $x < 0$ we expect oscillatory solutions with the period of oscillation *decreasing* for increasing $|x|$ as the “effective wave number” grows like $k^2 \sim |y|$.

The two linearly independent solutions are labeled $Ai(x)$ and $Bi(x)$ and are shown in Fig. E1. If we introduce the natural variable, $\zeta = 2x^{3/2}/3$, these solutions can be expanded for large values of $|x|$ as follows:

$$Ai(x) \longrightarrow \frac{1}{2\sqrt{\pi\sqrt{x}}} e^{-\zeta} \left[1 - \frac{c_1}{\zeta} + \dots \right] \quad (\text{E.4})$$

$$Ai(-x) \longrightarrow \frac{1}{\sqrt{\pi\sqrt{x}}} \left[\sin\left(\zeta + \frac{\pi}{4}\right) - \cos\left(\zeta + \frac{\pi}{4}\right) \frac{c_1}{\zeta} + \dots \right] \quad (\text{E.5})$$

$$Bi(x) \longrightarrow \frac{1}{\sqrt{\pi\sqrt{x}}} e^{\zeta} \left[1 + \frac{c_1}{\zeta} + \dots \right] \quad (\text{E.6})$$

$$Bi(-x) \longrightarrow \frac{1}{\sqrt{\pi\sqrt{x}}} \left[\cos\left(\zeta + \frac{\pi}{4}\right) + \sin\left(\zeta + \frac{\pi}{4}\right) \frac{c_1}{\zeta} + \dots \right] \quad (\text{E.7})$$

where $c_1 = 5/72$.

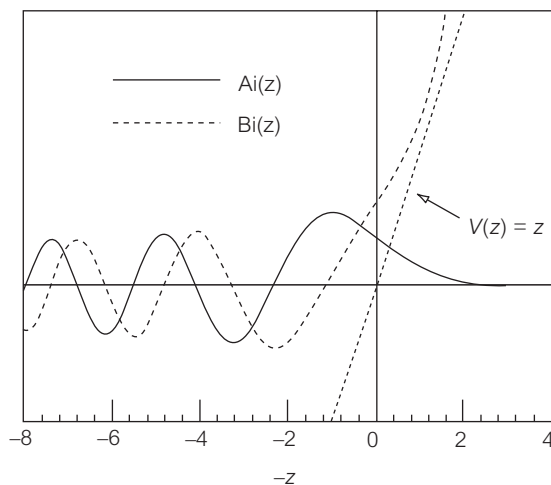


Figure E.1. Linearly independent solutions, $Ai(z)$ and $Bi(z)$, of the Airy differential equation.

E.3 Hermite Polynomials

The differential equation

$$\frac{d^2 h_n(z)}{dz^2} - 2z \frac{dh_n(z)}{dz} + 2n h_n(z) = 0 \quad (\text{E.8})$$

is called *Hermite's equation* and has the solutions given by *Rodrigues' formula*

$$h_n(z) = (-1)^n e^{z^2} \frac{d^n}{dz^n} \left(e^{-z^2} \right) \quad (\text{E.9})$$

which are polynomials of degree n . The solutions are defined over the interval $(-\infty, +\infty)$ and satisfy the normalization condition

$$\int_{-\infty}^{+\infty} [h_n(z)]^2 e^{-z^2} dz = 2^n n! \sqrt{\pi} \quad (\text{E.10})$$

The first few Hermite polynomials are given by

$$h_0(z) = 1 \quad h_1(z) = 2z \quad h_2(z) = 4z^2 - 2 \quad h_3(z) = 8z^3 - 12z \quad (\text{E.11})$$

E.4 Cylindrical Bessel Functions

The free-particle Schrödinger and wave equation in two dimensions (three dimensions), when written in polar (cylindrical) coordinates, leads to the equation

$$\frac{d^2 R_m(z)}{dz^2} + \frac{1}{z} \frac{dR_m(z)}{dz} + \left(1 - \frac{m^2}{z^2}\right) R_m(z) = 0 \quad (\text{E.12})$$

where we consider integral values of m . The solutions are generically called *cylindrical Bessel functions*, and for each $|m|$, the two linearly independent solutions are labeled $J_{|m|}(z)$ (Bessel functions of the first kind) or $Y_{|m|}(z)$ (Neumann functions or Bessel functions of the second kind). Their limiting behavior and properties are discussed and displayed graphically in Section 15.3.1.

E.5 Spherical Bessel Functions

The free-particle Schrödinger equation in three-dimensions written in spherical coordinates yields another version of Bessel's equation, namely

$$\frac{d^2 R_l(z)}{dz^2} + \frac{2}{z} \frac{dR_l(z)}{dz} + \left(1 - \frac{l(l+1)}{z^2}\right) R_l(z) = 0 \quad (\text{E.13})$$

with l an integer. Its solutions are the *spherical Bessel functions*, $j_l(z)$ and $n_l(z)$, which can be written in a standard form, in terms of the cylindrical Bessel functions

$$j_l(z) = \sqrt{\frac{\pi}{2z}} J_{l+1/2}(z) \quad \text{and} \quad n_l(z) = \sqrt{\frac{\pi}{2z}} Y_{l+1/2}(z) \quad (\text{E.14})$$

and are discussed in Section 16.6.

E.6 Legendre Polynomials

The (associated) Legendre's differential equation is written in the form

$$(1 - z^2) \frac{d^2 \Theta_{l,m}(z)}{dz^2} - 2z \frac{d\Theta_{l,m}(z)}{dz} + \left(l(l+1) - \frac{m^2}{(1 - z^2)}\right) \Theta_{l,m}(z) = 0 \quad (\text{E.15})$$

The solutions are the *associated Legendre polynomials* given by

$$P_l^m(z) = (-1)^m \frac{(1-z^2)^{m/2}}{2^l l!} \left(\frac{d}{dz} \right)^{l+m} (z^2-1)^l \quad (\text{E.16})$$

for $m > 0$ and extended to negative m via

$$P_l^{-m}(z) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(z) \quad (\text{E.17})$$

They are defined over the interval $(-1, 1)$ and the normalization is such that

$$\begin{aligned} \int_{-1}^{+1} dz P_l^m(z) P_{l'}^m(z) &= \int_0^\pi \sin(\theta) d\theta P_l^m(\cos(\theta)) P_{l'}^m(\cos(\theta)) \\ &= \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{l,l'} \end{aligned} \quad (\text{E.18})$$

The special case of $m = 0$ gives the *Legendre polynomials* which are defined via

$$P_l(z) \equiv P_l^{m=0}(z) \quad (\text{E.19})$$

which satisfy the differential equation

$$(1-z^2) \frac{d^2 P_l(z)}{dz^2} - 2z \frac{dP_l(z)}{dz} + l(l+1) P_l(z) = 0 \quad (\text{E.20})$$

E.7 Generalized Laguerre Polynomials

The differential equation

$$\frac{d^2 G(z)}{dz^2} + \left(\frac{\alpha-1}{z} + 1 \right) \frac{dG(z)}{dz} + nG(z) = 0 \quad (\text{E.21})$$

is called *Laguerre's equation* and has polynomial solutions labeled as

$$G(z) = L_n^{(\alpha)}(z) \quad (\text{E.22})$$

which can be generated using *Rodrigues' formula*

$$L_n^{(\alpha)}(z) = \frac{e^z}{n! z^\alpha} \left(\frac{d}{dz} \right)^n [z^{n+\alpha} e^{-z}] \quad (\text{E.23})$$

The solutions are defined over the interval $(0, +\infty)$ and satisfy the normalization condition

$$\int_0^{+\infty} dz z^\alpha e^{-z} [L_n^{(\alpha)}(z)]^2 = \frac{\Gamma(n+\alpha+1)}{n!} \quad (\text{E.24})$$

E.8 The Dirac δ -Function

The Dirac δ -function was introduced and discussed extensively in Section 2.4 and here we only list some additional properties and identities. We recall that

$$\int_a^b dx f(x) \delta(x - c) = \begin{cases} f(c) & \text{if } a < c < b \\ 0 & \text{otherwise} \end{cases} \quad (\text{E.25})$$

that is, the value of $f(x=c)$ is picked out from the integrand, or not, depending on whether the singularity is contained in the region of integration, or not. One can also derive (or justify) the following results.

$$\delta(ax) = \frac{1}{|a|} \delta(x) \quad (\text{E.26})$$

$$\begin{aligned} \delta(x^2 - a^2) &= \delta[(x - a)(x + a)] \\ &= \frac{1}{|x + a|} \delta(x - a) + \frac{1}{|x - a|} \delta(x + a) \\ \delta(x^2 - a^2) &= \frac{1}{2|a|} (\delta(x - a) + \delta(x + a)) \end{aligned} \quad (\text{E.27})$$

which is a special case of the more general relation

$$\delta[f(x)] = \sum_i \frac{\delta(x - x_i)}{|df/dx|_{x=x_i}} \quad (\text{E.28})$$

where the sum is over all possible roots of $f(x_i) = 0$. Finally, recall that the *step-* or *Heaviside-function* is defined via

$$\Theta(x - a) = \begin{cases} 0 & \text{for } x < a \\ 1 & \text{for } x > a \end{cases} \quad (\text{E.29})$$

and is given by

$$\Theta'(x - a) = \delta(x - a) \quad (\text{E.30})$$

One can show that $\delta(x)$ can be obtained by taking the limit of the family of functions

$$\delta_\epsilon(x) = \frac{\epsilon \sin^2(x/\epsilon)}{\pi x^2} \quad \text{as } \epsilon \rightarrow 0 \quad (\text{E.31})$$

E.9 The Euler Gamma Function

Using integration by parts techniques, it is easy to derive Eqn. (D.34), namely

$$\int_0^{\infty} dx x^n e^{-x} = n(n-1)(n-2) \cdots 3 \cdot 2 \cdot 1 \equiv n! \quad (\text{E.32})$$

where $n!$ is read as “ n -factorial.” The integral can be generalized to noninteger values of n , and the *Gamma function* is defined in this way via

$$\int_0^{\infty} dx x^{n-1} e^{-x} \equiv \Gamma(n) \quad \text{for } n \neq 0, -1, -2, -3, \dots \quad (\text{E.33})$$

For positive integers, it reduces to the factorial function

$$\Gamma(n) = (n-1)! \quad \text{for integral } n > 0 \quad (\text{E.34})$$

and also satisfies

$$\Gamma(n+1) = n\Gamma(n) \quad (\text{E.35})$$

$$\Gamma(n)\Gamma(1-n) = \frac{\pi}{\sin(n\pi)} \quad (\text{E.36})$$

Other special values are

$$\Gamma\left(\frac{1}{2}\right) = 2 \int_0^{\infty} e^{-t^2} dt = \sqrt{\pi} \quad (\text{E.37})$$

which can be evaluated since this is now a Gaussian integral. This can be combined with Eqn. (E.36) (for nonnegative integral n) to give

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^n} = \frac{(2n-1)!!}{2^n} \sqrt{\pi} \quad (\text{E.38})$$

which implicitly defines the double-factorial function. Finally, we note that *Stirling's formula* can be used to estimate the value of the factorial function for large argument, namely

$$\Gamma(n+1) = n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} + \cdots\right) \quad (\text{E.39})$$

E.10 Problems

PE.1. Show that the solutions to the Airy differential equation can be written in terms of cylindrical Bessel functions (satisfying Eqn. (E.12) of fractional ($n = \pm 1/3$))

order. For example, a standard result is that

$$Ai(-x) = \frac{1}{3}\sqrt{x} [J_{1/3}(y) + J_{-1/3}(y)] \quad (\text{E.40})$$

where $y = 2x^{3/2}/3$.

PE.2. Estimate the value of $20!$ using Stirling's formula and compare to the exact value.

APPENDIX F

Vectors, Matrices, and Group Theory

We collect here some of the most basic definitions and properties of real, finite-dimensional vectors and matrices. We intentionally ignore all of the subtleties regarding the precise definitions of vectors and tensors and supply only the “bare necessities”. Many comments on the generalization of these ideas to complex vectors and infinite dimensional spaces are given in the text. We then briefly describe some of the rudiments of group theory.

F.1 Vectors and Matrices

We will take vectors to be ordered N -tuples of numbers, for example,

$$\mathbf{x} = (x_1, x_2, \dots, x_N) \quad \text{and} \quad \mathbf{y} = (y_1, y_2, \dots, y_N) \quad (\text{E.1})$$

along with a *dot-* or *inner-product* of the form

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^N x_i y_i = x_1 y_1 + x_2 y_2 + \dots + x_N y_N \quad (\text{E.2})$$

The *norm* (or generalized length) of the vector is taken to be

$$|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}} \quad (\text{E.3})$$

A *matrix* will be defined to be a square $N \times N$ array of the form

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1N} \\ M_{21} & M_{22} & \cdots & M_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{N1} & M_{N2} & \cdots & M_{NN} \end{pmatrix} \quad (\text{E.4})$$

The *unit matrix* is given by

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \quad (\text{F.5})$$

Multiplication of a vector by a matrix *on the left* (as with operators) gives a vector, that is,

$$\mathbf{x}' = \mathbf{M} \cdot \mathbf{x} \quad (\text{F.6})$$

In component form one can write

$$(x')_i = (\mathbf{M} \cdot \mathbf{x})_i = \sum_{j=1}^N M_{ij} x_j \quad (\text{F.7})$$

or more explicitly

$$\begin{aligned} \mathbf{M} \cdot \mathbf{x} &= \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1N} \\ M_{21} & M_{22} & \cdots & M_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{N1} & M_{N2} & \cdots & M_{NN} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \\ &= \begin{pmatrix} M_{11}x_1 + M_{12}x_2 + \cdots + M_{1N}x_N \\ M_{21}x_1 + M_{22}x_2 + \cdots + M_{2N}x_N \\ \vdots \\ M_{N1}x_1 + M_{N2}x_2 + \cdots + M_{NN}x_N \end{pmatrix} \end{aligned} \quad (\text{F.8})$$

The product of two matrices is again a matrix with the component definition

$$(\mathbf{M} \cdot \mathbf{N})_{ik} = \sum_{j=1}^N M_{ij} N_{jk} \quad (\text{F.9})$$

or

the (ik) -th element of $\mathbf{M} \cdot \mathbf{N}$



(the i th row of \mathbf{M}) dotted into (the k th column of \mathbf{N})

The *transpose* of the matrix \mathbf{M} , labeled \mathbf{M}^T , is obtained by “reflecting” all of its elements along the diagonal (the $i = j$ components staying fixed), that is,

$$(\mathbf{M}^T)_{ij} = (\mathbf{M})_{ji} = M_{ji} \quad (\text{F.10})$$

so that

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}^T = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix} \quad (\text{F.11})$$

The generalization of this to complex matrices is the *adjoint* or *Hermitian conjugate* defined via

$$\mathbf{M}^\dagger = (\mathbf{M}^T)^* = (\mathbf{M}^*)^T \quad (\text{F.12})$$

which “flips” the matrix elements *and* takes their complex conjugate.

The equivalent of the expectation value of an operator in a quantum state is given by

$$\langle x|M|x \rangle \sim \mathbf{x} \cdot \mathbf{M} \cdot \mathbf{x} = \sum_{j=1}^N \sum_{k=1}^N x_j M_{jk} x_k \quad (\text{F.13})$$

A matrix transformation of the form Eqn. (F.6) generally changes the norm of the vector since

$$\begin{aligned} \mathbf{x}' \cdot \mathbf{x}' &= \sum_i x'_i x'_i = \sum_{i=1}^N \left(\sum_{j=1}^N M_{ij} x_j \right) \left(\sum_{k=1}^N M_{ik} x_k \right) \\ &= \sum_{j,k=1}^N x_j \left[\sum_i (M^T)_{jk} M_{ik} \right] x_k \\ &= \sum_{j,k=1}^N x_j P_{jk} x_k \\ &\neq \sum_j x_j x_j = \mathbf{x} \cdot \mathbf{x} \end{aligned} \quad (\text{F.14})$$

unless one has

$$\sum_i (M^T)_{ji} M_{ik} = P_{jk} = \delta_{j,k} \quad \text{or} \quad \mathbf{M}^T \cdot \mathbf{M} = \mathbf{P} = \mathbf{1} \quad (\text{F.15})$$

Matrices satisfying Eqn. (F.15) are said to be *orthogonal*.

Finally, the *determinant* of a matrix is a number formed from the elements of the matrix via

$$\det(\mathbf{M}) = \sum_{i_1, i_2, \dots, i_N=1}^N \epsilon_{(i_1, i_2, \dots, i_N)} M_{1, i_1} M_{2, i_2} \cdots M_{N, i_N} \quad (\text{F.16})$$

The *totally antisymmetric symbol*,¹ $\epsilon_{(i_1, i_2, \dots, i_N)}$, is defined via

$$\epsilon_{(i_1, i_2, \dots, i_N)} = \begin{cases} +1 & \text{if } (i_1, i_2, \dots, i_N) \text{ is an} \\ & \text{even permutation of } (1, 2, \dots, N) \\ -1 & \text{if it is an odd permutation} \\ 0 & \text{otherwise} \end{cases} \quad (\text{F.17})$$

It vanishes if any two of its indices are the same and is antisymmetric under the interchange of any pair of indices. Each term in the determinant then consists of a product of one element from each row, with appropriate signs. For example,

$$\det(\mathbf{A}) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad (\text{F.18})$$

and

$$\begin{aligned} \det(\mathbf{B}) &= \det \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \\ &= b_{11}b_{22}b_{33} + b_{12}b_{23}b_{31} + b_{13}b_{32}b_{21} - b_{13}b_{31}b_{22} \\ &\quad - b_{11}b_{23}b_{32} - b_{12}b_{21}b_{33} \end{aligned} \quad (\text{F.19})$$

One important property of determinants is that the interchange of any two rows (or columns) gives the same value, but with an additional factor of (-1) ; this follows from the definition in Eqn. (F.16) and the antisymmetry of the ϵ symbol.

Equations of the form

$$\mathbf{M} \cdot \mathbf{v}_\lambda = \lambda \mathbf{v}_\lambda \quad (\text{F.20})$$

are called *eigenvalue problems* and λ is the *eigenvalue* and \mathbf{v}_λ the corresponding *eigenvector*. We can also write this as

$$(\mathbf{M} - \lambda \mathbf{1}) \cdot \mathbf{v} = 0 \quad (\text{F.21})$$

or in matrix form as

$$\begin{pmatrix} M_{11} - \lambda & M_{12} & \cdots & M_{1N} \\ M_{21} & M_{22} - \lambda & \cdots & M_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{N1} & M_{N2} & \cdots & M_{NN} - \lambda \end{pmatrix} = 0 \quad (\text{F.22})$$

This is equivalent to a set of N linear equations in N unknowns and the condition for a solution to exist is that

$$\det(\mathbf{M} - \lambda \mathbf{1}) = 0 \quad (\text{F.23})$$

¹ It is also called the Levi-Civita symbol.

and this condition determines the allowed eigenvalues λ . Real matrices for which

$$\mathbf{M}^T = \mathbf{M} \quad (\text{F.24})$$

are called *symmetric*, while complex matrices for which

$$\mathbf{M}^\dagger = \mathbf{M} \quad (\text{F.25})$$

are called *Hermitian* and both have the properties:

- The eigenvalues of \mathbf{M} are real.
- The eigenvectors of \mathbf{M} corresponding to different eigenvalues are orthogonal.

Example F.1. Eigenvalues and eigenvectors of a simple matrix

The eigenvalues of the matrix

$$\mathbf{M} = \begin{pmatrix} 23 & -36 \\ -36 & 2 \end{pmatrix} \quad (\text{F.26})$$

are determined by the condition

$$\det \begin{pmatrix} 23 - \lambda & -36 \\ -36 & 2 - \lambda \end{pmatrix} = \lambda^2 - 25\lambda - 1250 = 0 \quad (\text{F.27})$$

which has solutions $\lambda_1 = 50$ and $\lambda_2 = -25$. The eigenvector corresponding to λ_1 can be found by insisting that

$$\begin{pmatrix} 23 - 50 & -36 \\ -36 & 2 - 50 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0 \quad (\text{F.28})$$

or

$$\mathbf{v}_1 = \begin{pmatrix} 4/5 \\ -3/5 \end{pmatrix} \quad (\text{F.29})$$

when normalized so that $\mathbf{v}_1 \cdot \mathbf{v}_1 = 1$. One similarly finds that

$$\mathbf{v}_2 = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix} \quad (\text{F.30})$$

(with $\mathbf{v}_2 \cdot \mathbf{v}_2 = 1$ by construction) and we confirm that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

Finally, some useful identities involving the scalar and cross-products of vectors are

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{C} \cdot \mathbf{A}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}) \quad (\text{F.31})$$

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) \quad (\text{F.32})$$

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}) \quad (\text{F.33})$$

F.2 Group Theory

We conclude with a brief definition of a mathematical *group*. A set of elements $G = \{g_1, g_2, \dots\}$ along with a binary operation (often called “group multiplication”) denoted by $g_1 \cdot g_2$ constitutes a group if it satisfies four conditions:

1. The product of any two group elements is also a group element, that is, $g_3 = g_1 \cdot g_2$ is from G if g_1, g_2 are also; the group is closed under multiplication.
2. The group multiplication is associative, namely

$$(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3) \quad (\text{F.34})$$

3. There is a unique group element, labeled I or the identity element, which satisfies

$$I \cdot g_i = g_i \cdot I = g_i \quad (\text{F.35})$$

for all $g_i \in G$.

4. Every group element, g_i , has a unique inverse element, labeled g_i^{-1} which satisfies

$$g_i \cdot g_i^{-1} = g_i^{-1} \cdot g_i = I \quad (\text{F.36})$$

The set of group elements can be finite or infinite. Groups for which the multiplication gives the same answer in either order, that is, for which $g_i \cdot g_j = g_j \cdot g_i$ for every pair of group elements is called a *commutative* or *Abelian group*.

F.3 Problems

PF.1. Show that $(\mathbf{A} \cdot \mathbf{B} \cdots \mathbf{Y} \cdot \mathbf{Z})^T = \mathbf{Z}^T \cdot \mathbf{Y}^T \cdots \mathbf{B}^T \cdot \mathbf{A}^T$.

PF.2. Show that the cross-product of two vectors can be written in the form

$$\mathbf{A} \times \mathbf{B} = \det \begin{pmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{pmatrix} \quad (\text{F.37})$$

PF.3. Find the eigenvalues and eigenvectors of the Hermitian matrix

$$\mathbf{M} = \begin{pmatrix} 4 & 3 + 2i \\ 3 - 2i & -5 \end{pmatrix} \quad (\text{F.38})$$

and show explicitly that the two eigenvectors are orthogonal. Note that the dot-product of two *complex* vectors is defined via $\mathbf{v}_1^* \cdot \mathbf{v}_2$.

APPENDIX G

Hamiltonian Formulation of Classical Mechanics

In this appendix, we briefly review some aspects of the Hamiltonian formulation of classical mechanics. The Hamiltonian function for a single particle described by a single coordinate, call it x , is a function of x and the so-called “conjugate momentum,” p_x . Examples of such coordinate pairs include

- x and $p_x = m\dot{x}$ (ordinary momentum) for translational motion and
- θ and $p_\theta = L_z = mr^2\dot{\theta}$ (angular momentum) for a rotational system.

The Hamiltonian is written as $H = H(x, p_x)$ and x and p_x are initially considered as independent variables.

The dynamical equations of motion for $x(t)$ and $p_x(t)$ are *Hamilton’s equations*, namely

$$\frac{dx}{dt} = \dot{x} = \frac{\partial H}{\partial p_x} \quad (\text{G.1})$$

$$-\frac{dp_x}{dt} = -\dot{p}_x = \frac{\partial H}{\partial x} \quad (\text{G.2})$$

To see the equivalence to Newtonian mechanics, note that the Hamiltonian function

$$H(x, p_x) = \frac{p_x^2}{2m} + V(x) \quad (\text{G.3})$$

gives the equations

$$\dot{x} = \frac{p_x}{m} \quad \text{and} \quad -\dot{p}_x = \frac{\partial V(x)}{\partial x} \equiv -F(x) \quad (\text{G.4})$$

or

$$m\ddot{x} = \dot{p}_x = F(x) \quad (\text{G.5})$$

The Hamiltonian function for the degrees of freedom of more than one particle (in one dimension) can be written

$$H = H(x_i, p_i) = \sum_i \frac{p_i^2}{2m_i} + \sum_i V_i(x_i) + \sum_{i>j} V_{ij}(x_i - x_j) \quad (\text{G.6})$$

with the corresponding equations

$$\frac{dx_i}{dt} = \dot{x}_i = \frac{\partial H}{\partial p_i} \quad (\text{G.7})$$

$$-\frac{dp_i}{dt} = -\dot{p}_i = \frac{\partial H}{\partial x_i} \quad (\text{G.8})$$

For two functions which depend on the coordinates of a multivariable problem (and possibly the time coordinate explicitly), $g = g(x_i, p_i; t)$ and $h = h(x_i, p_i; t)$, the *Poisson bracket* is defined via

$$[g, h] = \sum_k \left(\frac{\partial g}{\partial x_k} \frac{\partial h}{\partial p_k} - \frac{\partial g}{\partial p_k} \frac{\partial h}{\partial x_k} \right) \quad (\text{G.9})$$

Any such arbitrary function can depend on time either from an explicit t dependence or via the coordinates $x_i(t), p_i(t)$; a convenient way of exhibiting the time-development of a function is

$$\begin{aligned} \frac{dg}{dt} &= \frac{\partial g}{\partial t} + \sum_k \left(\frac{\partial g}{\partial x_k} \frac{dx_k}{dt} + \frac{\partial g}{\partial p_k} \frac{dp_k}{dt} \right) \\ &= \frac{\partial g}{\partial t} + \sum_k \left(\frac{\partial g}{\partial x_k} \frac{\partial H}{\partial p_i} - \frac{\partial g}{\partial p_k} \frac{\partial H}{\partial x_i} \right) \\ &= \frac{\partial g}{\partial t} + [g, H] \end{aligned} \quad (\text{G.10})$$

Note the similarity between this classical relation and Eqn. (12.88) for the time rate of change of expectation values of quantum operators.

Example G.1.

Using the Poisson bracket formalism, we can show that angular momentum is conserved for a central potential in three dimensions. The Hamiltonian function is given by

$$H = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + V(r) \quad (\text{G.11})$$

where $r = \sqrt{x^2 + y^2 + z^2}$. We note that the force is given $\mathbf{F} = -\nabla V(r)$ so that

$$\frac{\partial V(r)}{\partial x} = (-F) \frac{x}{r} \quad (\text{G.12})$$

and so forth.

(Continued)

Considering, for definiteness, the z component of angular momentum given by $L_z = x p_y - y p_x$, one can show that

$$\begin{aligned}
 \frac{dL_z}{dt} &= \frac{\partial L_z}{\partial t} + [L_z, H] \\
 &= \left[\frac{\partial L_z}{\partial x} \frac{\partial H}{\partial p_x} - \frac{\partial L_z}{\partial p_x} \frac{\partial H}{\partial x} \right] + \left[\frac{\partial L_z}{\partial y} \frac{\partial H}{\partial p_y} - \frac{\partial L_z}{\partial p_y} \frac{\partial H}{\partial y} \right] + \left[\frac{\partial L_z}{\partial z} \frac{\partial H}{\partial p_z} - \frac{\partial L_z}{\partial p_z} \frac{\partial H}{\partial z} \right] \\
 &= \left[(p_y) \left(\frac{p_x}{m} \right) - (-y) \left(-F \frac{x}{r} \right) \right] + \left[(-p_x) \left(\frac{p_y}{m} \right) - (x) \left(-F \frac{y}{r} \right) \right] \\
 \frac{dL_z}{dt} &= 0
 \end{aligned} \tag{G.13}$$

This relation also suggests that the Poisson bracket of two functions is the classical quantity which can be generalized in quantum mechanics to the commutator of two operators

$$[\hat{g}, \hat{h}] \equiv \hat{g}\hat{h} - \hat{h}\hat{g} \tag{G.14}$$

via

$$[g, h]_{\text{Poisson}} \longrightarrow [\hat{g}, \hat{h}] = i\hbar[g, h]_{\text{Poisson}} \tag{G.15}$$

The Hamiltonian for a charged particle acted on by electromagnetic fields is written in terms of the potentials, $\phi(\mathbf{r}, t)$ and $\mathbf{A}(\mathbf{r}, t)$ via

$$H(\mathbf{r}, \mathbf{p}) = \frac{1}{2m} (\mathbf{p} - q\mathbf{A}(\mathbf{r}, t))^2 + q\phi(\mathbf{r}, t) \tag{G.16}$$

To prove this requires one to show that the corresponding Hamilton's equations reproduce Newton's laws with the Lorentz force. The classical Hamiltonian in Eqn. (G.16) can be written more explicitly as

$$\begin{aligned}
 H &= \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) - \frac{q}{m} (p_x A_x + p_y A_y + p_z A_z) \\
 &\quad + \frac{q^2}{2m} (A_x^2 + A_y^2 + A_z^2) + q\phi
 \end{aligned} \tag{G.17}$$

Hamilton's equations for the x and p_x coordinates in this case become

$$\dot{x} = \frac{\partial H}{\partial p_x} = \frac{p_x}{m} - \frac{q}{m} A_x \quad \text{or} \quad m\dot{x} = p_x - qA_x \tag{G.18}$$

and

$$\begin{aligned}
 -\dot{p}_x &= \frac{\partial H}{\partial x} \\
 &= -\frac{q}{m} \left(p_x \frac{\partial A_x}{\partial x} + p_y \frac{\partial A_y}{\partial x} + p_z \frac{\partial A_z}{\partial x} \right) \\
 &\quad + \frac{q^2}{2m} \left(A_x \frac{\partial A_x}{\partial x} + A_y \frac{\partial A_y}{\partial x} + A_z \frac{\partial A_z}{\partial x} \right) + q \frac{\partial \phi}{\partial x}
 \end{aligned} \tag{G.19}$$

These can be combined by differentiating Eqn. (G.18) with respect to t provided one recalls that

$$m\ddot{x} = \dot{p}_x - q \frac{dA_x}{dt} = \dot{p}_x - q \left(\frac{\partial A_x}{\partial t} + \dot{x} \frac{\partial A_x}{\partial x} + \dot{y} \frac{\partial A_x}{\partial y} + \dot{z} \frac{\partial A_x}{\partial z} \right) \tag{G.20}$$

since $\mathbf{A} = \mathbf{A}(x(t), y(t), z(t); t)$ depends on time explicitly (through the t) and implicitly (through the time-dependent positions). The resulting equation for the x variable is then

$$m\ddot{x} = q \left(\left\{ -\frac{\partial \phi}{\partial x} - \frac{\partial A_x}{\partial t} \right\} + \left\{ \dot{y} \left[\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right] + \dot{z} \left[\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z} \right] \right\} \right) \tag{G.21}$$

or

$$m\ddot{\mathbf{x}} = q (\mathbf{E}_x + (\mathbf{v} \times \mathbf{B})_x) \tag{G.22}$$

since

$$\mathbf{E} = -\nabla \phi(\mathbf{r}, t) - \frac{\partial}{\partial t} \mathbf{A}(\mathbf{r}, t) \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A}(\mathbf{r}, t) \tag{G.23}$$

The Hamiltonian formalism also provides a way to describe classical probability distributions for position and momentum which can be compared to their quantum counterparts (as discussed in Sections 5.1 and 9.4). One starts with the notion of the *classical phase space*. For one particle in one dimension, this is the space of possible values of x and p , as in Fig. G1; for N particles in three dimensions, it is a $6N$ -dimensional space corresponding to the possible values of $\mathbf{r}_i, \mathbf{p}_i$. Given a set of initial conditions, the solutions obtained from Newton's (or Hamilton's) equations for $x(t)$ and $p(t)$ trace out a trajectory in the phase space. For example, for a harmonic oscillator with initial conditions $x(0) = A$ and $\dot{x}(0) = 0$, the solutions are obviously

$$x(t) = A \cos(\omega t) \quad \text{and} \quad p(t) = -\frac{A\omega}{m} \sin(\omega t) \tag{G.24}$$

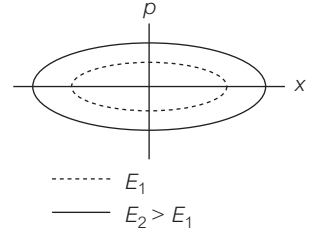


Figure G.1. Phase space diagram (plot of allowed values of p and x) for a single particle in a harmonic oscillator potential. The allowed “trajectories” in the parameter space in this case are determined by Eqn. (G.24).

which gives the elliptical path in phase space shown in Fig. G1. The form of this “trajectory” can be determined, even if we specify only the total energy, via the relation

$$E = H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \quad (\text{G.25})$$

The *phase space distribution*, $\rho(x, p)$, for a given value of E can then be written in the form

$$\rho(x, p) = K\delta(E - H(p, x)) \quad (\text{G.26})$$

where the normalization constant is determined by the condition that

$$\int dx \int dp \rho(x, p) = 1 \quad (\text{G.27})$$

The classical probability densities for position (x) or momentum (p) can be derived by integrating over the variable which is not specified; for example,

$$P_{\text{CL}}(x) = \int dp \rho(x, p) \quad \text{and} \quad P_{\text{CL}}(p) = \int dx \rho(x, p) \quad (\text{G.28})$$

As an example, consider the Hamiltonian with a general potential energy function $V(x)$

$$H = \frac{p^2}{2m} + V(x) \quad (\text{G.29})$$

If we write $p_0(x) = \sqrt{E - V(x)}$, the corresponding classical distribution in x will be given by

$$\begin{aligned} P_{\text{CL}}(x) &= K \int dp \delta \left(E - \left(\frac{p^2}{2m} - V(x) \right) \right) \\ &\propto \int dp \delta (p^2 - 2m(E - V(x))) \\ &\equiv \int dp \delta (p^2 - p_0^2(x)) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2|p_0(x)|} \int dp [\delta(p - p_0(x)) + \delta(p + p_0(x))] \\
&\propto \frac{1}{\sqrt{E - V(x)}}
\end{aligned} \tag{G.30}$$

since $p_0^2(x) = 2m(E - V(x))$. This is the same result obtained in Sections 5.1 and 9.4.1, using more intuitive methods.

G.1 Problems

PG.1. Write down Hamilton's equations for the angular variable pair θ and $p_\theta = mr^2\dot{\theta}$ where the Hamiltonian is

$$H = \frac{p_\theta^2}{2I} + V(\theta) \tag{G.31}$$

and show how the standard equations for rotational motion arise.

PG.2. Show that the Poisson bracket of the position and momentum coordinates of a multiparticle system satisfy

$$[x_i, x_j] = 0, \quad [p_i, p_j] = 0, \quad \text{and} \quad [x_i, p_j] = \delta_{ij} \tag{G.32}$$

PG.3. Consider a harmonic oscillator for which the classical Hamiltonian is

$$H = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{G.33}$$

Use Eqn. (G.10) to show that the function

$$\phi(x, p_x; t) = i(\log(A) - \log(x - ip/m\omega)) - \omega t \tag{G.34}$$

is actually independent of time for this system. What is the physical significance of this variable?

PG.4. For the classical Kepler problem, defined by the Hamiltonian

$$H = \frac{1}{2m}\mathbf{p}^2 - \frac{k}{r} \tag{G.35}$$

show that the *Lenz–Runge vector* defined by

$$\mathbf{R} = \frac{\hat{\mathbf{r}}}{r} - \left(\frac{1}{mk} \right) \mathbf{p} \times \mathbf{L} \tag{G.36}$$

is a constant of the motion, that is it is conserved. Do this by showing that $d\mathbf{R}/dt = 0$ using the Poisson bracket formalism.

PG.5. Using the Hamiltonian corresponding to a linear confining potential

$$H = \frac{p^2}{2m} + C|x| \quad (\text{G.37})$$

and Eqn. (G.28), derive $P_{\text{CL}}(p)$.

PG.6. What does the classical phase space diagram look like for one particle in the one-dimensional infinite well, that is, what is the analog of Fig. G1? For the particle in the potential of PG.5? For an unbound free particle? For a particle subject to a constant force given by $V(x) = -Fx$? For a particle bouncing up and down on a table, with no energy loss, under the influence of gravity?

Math identities and other results

Gaussian integrals:

$$\begin{aligned}
 I^{(0)}(a, b) &\equiv \int_{-\infty}^{+\infty} e^{-ax^2-bx} dx = \sqrt{\pi/a} \exp(b^2/4a) \\
 I^{(1)}(a, b) &\equiv \int_{-\infty}^{+\infty} x e^{-ax^2-bx} dx = (-2b/a) \sqrt{\pi/a} \exp(b^2/4a) \\
 I^{(2)}(a, b) &\equiv \int_{-\infty}^{+\infty} x^2 e^{-ax^2-bx} dx = [(b^2 + 2a)/4a^2] \sqrt{\pi/a} \exp(b^2/4a)
 \end{aligned}$$

Other integrals:

$$\begin{aligned}
 \int \sin^2(ax) dx &= x/2 - \sin(2ax)/4a \\
 \int \cos^2(ax) dx &= x/2 + \sin(2ax)/4a \\
 \int \sin(mx) \sin(nx) dx &= +\sin[(m-n)x]/2(m-n) - \sin[(m+n)x]/2(m+n) \\
 \int \cos(mx) \cos(nx) dx &= +\sin[(m-n)x]/2(m-n) + \sin[(m+n)x]/2(m+n) \\
 \int \sin(mx) \cos(nx) dx &= -\cos[(m-n)x]/2(m-n) - \cos[(m+n)x]/2(m+n)
 \end{aligned}$$

Trig identities:

$$\begin{aligned}
 \sin(a \pm b) &= \sin(a) \cos(b) \pm \cos(a) \sin(b) \\
 \cos(a \pm b) &= \cos(a) \cos(b) \mp \sin(a) \sin(b) \\
 \sin(a) \sin(b) &= \cos(a-b)/2 - \cos(a+b)/2 \\
 \cos(a) \cos(b) &= \cos(a-b)/2 + \cos(a+b)/2 \\
 \sin(a) \cos(b) &= \sin(a-b)/2 + \sin(a+b)/2 \\
 \cos(a) \sin(b) &= \sin(a+b)/2 - \sin(a-b)/2 \\
 \exp(\pm iz) &= \cos(z) \pm i \sin(z)
 \end{aligned}$$

Free-particle Gaussian wave packets (Chapters 3 and 4):

$$\begin{aligned}
 \phi_{(G)}(p, t) &= \phi_0(p) e^{-ip^2 t/2m\hbar} = \left[\sqrt{\frac{\alpha}{\sqrt{\pi}}} e^{-\alpha^2(p-p_0)^2/2} e^{-ipx_0/\hbar} \right] e^{-ip^2 t/2m\hbar} \\
 \psi_{(G)}(x, t) &= \frac{1}{\sqrt{\sqrt{\pi}\alpha\hbar(1+it/t_0)}} e^{ip_0(x-x_0)/\hbar} e^{-ip_0^2 t/2m\hbar} \\
 &\quad e^{-(x-x_0-p_0 t/m)^2/2(\alpha\hbar)^2(1+it/t_0)} \\
 |\psi_{(G)}(x, t)|^2 &= \frac{1}{\sqrt{\pi}\beta_t} e^{-(x-x_0-p_0 t/m)^2/\beta_t^2} \\
 \langle p \rangle_t &= p_0, \quad \langle p^2 \rangle_t = p_0^2 + \frac{1}{2\alpha^2}, \quad \text{and} \quad \Delta p_t = \Delta p_0 = \frac{1}{\alpha\sqrt{2}} \\
 \langle x \rangle_t &= x(t) \equiv x_0 + p_0 t/m, \quad \langle x^2 \rangle_t = [x(t)]^2 + \frac{\beta_t^2}{2}, \quad \text{and} \quad \Delta x_t = \frac{\beta_t}{\sqrt{2}}
 \end{aligned}$$

where

$$\beta_t \equiv \alpha\hbar\sqrt{1+t^2/t_0^2} \quad \text{and} \quad t_0 \equiv m\hbar\alpha^2 = \frac{m\hbar}{2(\Delta p_0)^2} = \frac{2m(\Delta x_0)^2}{\hbar}$$