

which allows us to write the defining relation of the quantum operators as

$$[X, P] = i\hbar \{x, p\} = i\hbar \quad (7.4.38)$$

The virtue of this viewpoint is that its generalization to the “quantization” of a system of N degrees of freedom is apparent:

Postulate II (For N Degrees of Freedom). The Cartesian coordinates x_1, \dots, x_N and momenta p_1, \dots, p_N of the classical description of a system with N degrees of freedom now become Hermitian operators $X_1, \dots, X_N; P_1, \dots, P_N$ obeying the commutation rules

$$\begin{aligned} [X_i, P_j] &= i\hbar \{x_i, p_j\} = i\hbar \delta_{ij} \\ [X_i, X_j] &= i\hbar \{x_i, x_j\} = 0 \\ [P_i, P_j] &= i\hbar \{p_i, p_j\} = 0 \end{aligned} \quad (7.4.39)$$

Similarly $\omega(x, p) \rightarrow \omega(x \rightarrow X, p \rightarrow P) = \Omega$.

[We restrict ourselves to Cartesian coordinates to avoid certain subtleties associated with the quantization of non-Cartesian but canonical coordinates; see Exercise (7.4.10). Once the differential equations are obtained, we may abandon Cartesian coordinates in looking for the solutions.]

It is evident that the generalization provided towards the end of Section 4.2, namely,

$$\begin{array}{ccc} X_i & \xrightarrow{\hspace{1cm}} & x_i \\ & \scriptstyle X \text{ basis} & \\ P_i & \xrightarrow{\hspace{1cm}} & -i\hbar \frac{\partial}{\partial x_i} \\ & \scriptstyle X \text{ basis} & \end{array}$$

is a choice but not *the* choice satisfying the canonical commutation rules, Eq. (7.4.39), for the same reason as in the $N=1$ case.

Given the commutation relations between X and P , the ones among dependent operators follow from the repeated use of the relations

$$[\Omega, \Lambda \Gamma] = \Lambda [\Omega, \Gamma] + [\Omega, \Lambda] \Gamma$$

and

$$[\Omega \Lambda, \Gamma] = \Omega [\Lambda, \Gamma] + [\Omega, \Gamma] \Lambda$$

Since PB obey similar rules (Exercise 2.7.1) except for the lack of emphasis on ordering of the classical variables, it turns out that if

$$\{\omega(x, p), \lambda(x, p)\} = \gamma(x, p)$$

then

$$[\Omega(X, P), \Lambda(X, P)] = i\hbar\Gamma(X, P) \quad (7.4.40)$$

except for differences arising from ordering ambiguities; hence the formal similarity between classical and quantum mechanics, first encountered in Chapter 6.

Although the new form of postulate II provides a general, basis-independent specification of the quantum operators corresponding to classical variables, that is to say for “quantizing,” in practice one typically works in the X basis and also ignores the latitude in the choice of P_i and sticks to the traditional one, $P_i = -i\hbar \partial/\partial x_i$, which leads to the simplest differential equations. The solution to the oscillator problem, given just the commutation relations (and a little help from Dirac) is atypical.

*Exercise 7.4.1.** Compute the matrix elements of X and P in the $|n\rangle$ basis and compare with the result from Exercise 7.3.4.

*Exercise 7.4.2.** Find $\langle X \rangle$, $\langle P \rangle$, $\langle X^2 \rangle$, $\langle P^2 \rangle$, $\Delta X \cdot \Delta P$ in the state $|n\rangle$.

Exercise 7.4.3. (Virial Theorem).* The virial theorem in classical mechanics states that for a particle bound by a potential $V(r) = ar^k$, the average (over the orbit) kinetic and potential energies are related by

$$\bar{T} = c(k) \bar{V}$$

when $c(k)$ depends only on k . Show that $c(k) = k/2$ by considering a circular orbit. Using the results from the previous exercise show that for the oscillator ($k = 2$)

$$\langle T \rangle = \langle V \rangle$$

in the quantum state $|n\rangle$.

Exercise 7.4.4. Show that $\langle n|X^4|n\rangle = (\hbar/2m\omega)^2[3 + 6n(n+1)]$.

*Exercise 7.4.5.** At $t=0$ a particle starts out in $|\psi(0)\rangle = 1/2^{1/2}(|0\rangle + |1\rangle)$. (1) Find $|\psi(t)\rangle$; (2) find $\langle X(0) \rangle = \langle \psi(0)|X|\psi(0)\rangle$, $\langle P(0) \rangle$, $\langle X(t) \rangle$, $\langle P(t) \rangle$; (3) find $\langle \dot{X}(t) \rangle$ and $\langle \dot{P}(t) \rangle$ using Ehrenfest’s theorem and solve for $\langle X(t) \rangle$ and $\langle P(t) \rangle$ and compare with part (2).

*Exercise 7.4.6.** Show that $\langle a(t) \rangle = e^{-i\omega t} \langle a(0) \rangle$ and that $\langle a^\dagger(t) \rangle = e^{i\omega t} \langle a^\dagger(0) \rangle$.

Exercise 7.4.7. Verify Eq. (7.4.40) for the case

- (1) $\Omega = X$, $\Lambda = X^2 + P^2$
- (2) $\Omega = X^2$, $\Lambda = P^2$

The second case illustrates the ordering ambiguity.

*Exercise 7.4.8.** Consider the three angular momentum variables in classical mechanics:

$$l_x = yp_z - zp_y$$

$$l_y = zp_x - xp_z$$

$$l_z = xp_y - yp_x$$

- (1) Construct L_x , L_y , and L_z , the quantum counterparts, and note that there are no ordering ambiguities.
- (2) Verify that $\{l_x, l_y\} = l_z$ [see Eq. (2.7.3) for the definition of the PB].
- (3) Verify that $[L_x, L_y] = i\hbar L_z$.

Exercise 7.4.9 (Important). Consider the unconventional (but fully acceptable) operator choice

$X \rightarrow x$

$$P \rightarrow -i\hbar \frac{d}{dx} + f(x)$$

in the X basis.

- (1) Verify that the canonical commutation relation is satisfied.
- (2) It is possible to interpret the change in the operator assignment as a result of a unitary change of the X basis:

$$|x\rangle \rightarrow |\tilde{x}\rangle = e^{ig(x)/\hbar} |x\rangle = e^{ig(x)/\hbar} |x\rangle$$

where

$$g(x) = \int^x f(x') dx'$$

First verify that

$$\langle \tilde{x}|X|\tilde{x}'\rangle = x\delta(x-x')$$

i.e.,

$$X \xrightarrow{\text{new } X \text{ basis}} x$$

Next verify that

$$\langle \tilde{x}|P|\tilde{x}'\rangle = \left[-i\hbar \frac{d}{dx} + f(x) \right] \delta(x-x')$$

$$P \xrightarrow[\text{new } X \text{ basis}]{} -i\hbar \frac{d}{dx} + f(x)$$

This exercise teaches us that the “ X basis” is not unique; given a basis $|x\rangle$, we can get another $|\tilde{x}\rangle$, by multiplying by a phase factor which changes neither the norm nor the orthogonality. The matrix elements of P change with f , the standard choice corresponding to $f=0$. Since the presence of f is related to a change of basis, the invariance of the physics under a change in f (from zero to nonzero) follows. What is novel here is that we are changing from one X basis to another X basis rather than to some other Ω basis. Another lesson to remember is that two different differential operators $\omega(x, -i\hbar d/dx)$ and $\omega(x, -i\hbar d/dx + f)$ can have the same eigenvalues and a one-to-one correspondence between their eigenfunctions, since they both represent the same abstract operator $\Omega(X, P)$. \square

*Exercise 7.4.10.** Recall that we always quantize a system by promoting the Cartesian coordinates x_1, \dots, x_N ; and momenta p_1, \dots, p_N to operators obeying the canonical commutation rules. If non-Cartesian coordinates seem more natural in some cases, such as the eigenvalue problem of a Hamiltonian with spherical symmetry, we first set up the differential equation in Cartesian coordinates and *then* change to spherical coordinates (Section 4.2). In Section 4.2 it was pointed out that if \mathcal{H} is written in terms of non-Cartesian but canonical coordinates $q_1 \dots q_N; p_1 \dots p_N; \mathcal{H}(q_i \rightarrow q_i, p_i \rightarrow -i\hbar \partial/\partial q_i)$ does not generate the correct Hamiltonian H , even though the operator assignment satisfies the canonical commutation rules. In this section we revisit this problem in order to explain some of the subtleties arising in the direct quantization of non-Cartesian coordinates without the use of Cartesian coordinates in intermediate stages.

- (1) Consider a particle in two dimensions with

$$\mathcal{H} = \frac{p_x^2 + p_y^2}{2m} + a(x^2 + y^2)^{1/2}$$

which leads to

$$H \rightarrow \frac{-\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + a(x^2 + y^2)^{1/2}$$

in the coordinate basis. Since the problem has rotational symmetry we use polar coordinates

$$\rho = (x^2 + y^2)^{1/2}, \quad \phi = \tan^{-1}(y/x)$$

in terms of which

$$H \xrightarrow[\text{coordinate basis}]{} \frac{-\hbar^2}{2m} \left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \right) + a\rho \quad (7.4.41)$$

Since ρ and ϕ are not mixed up as x and y are [in the $(x^2 + y^2)^{1/2}$ term] the polar version can be more readily solved.

The question we address is the following: why not *start* with \mathcal{H} expressed in terms of polar coordinates and the conjugate momenta

$$p_\rho = \mathbf{e}_\rho \cdot \mathbf{p} = \frac{x p_x + y p_y}{(x^2 + y^2)^{1/2}}$$

(where \mathbf{e}_ρ is the unit vector in the radial direction), and

$$p_\phi = x p_y - y p_x \quad (\text{the angular momentum, also called } l_z)$$

i.e.,

$$\mathcal{H} = \frac{p_\rho^2}{2m} + \frac{p_\phi^2}{2m\rho^2} + a\rho \quad (\text{verify this})$$

and directly promote all classical variables ρ, p_ρ, ϕ , and p_ϕ to quantum operators obeying the canonical commutations rules? Let's do it and see what happens. If we choose operators

$$P_\rho \rightarrow -i\hbar \frac{\partial}{\partial \rho}$$

$$P_\phi \rightarrow -i\hbar \frac{\partial}{\partial \phi}$$

that obey the commutation rules, we end up with

$$H \xrightarrow[\text{coordinate basis}]{} \frac{-\hbar^2}{2m} \left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \right) + a\rho \quad (7.4.42)$$

which disagrees with Eq. (7.4.41). Now this in itself is not serious, for as seen in the last exercise the same physics may be hidden in two different equations. In the present case this isn't true: as we will see, the Hamiltonians in Eqs. (7.4.41) and (7.4.42) do not have the same eigenvalues.[‡] We know Eq. (7.4.41) is the correct one, since the quantization procedure in terms of Cartesian coordinates has empirical support. What do we do now?

(2) A way out is suggested by the fact that although the choice $P_\rho \rightarrow -i\hbar \partial/\partial \rho$ leads to the correct commutation rule, it is not Hermitian! Verify that

$$\begin{aligned} \langle \psi_1 | P_\rho | \psi_2 \rangle &= \int_0^\infty \int_0^{2\pi} \psi_1^* \left(-i\hbar \frac{\partial \psi_2}{\partial \rho} \right) \rho d\rho d\phi \\ &\neq \int_0^\infty \int_0^{2\pi} \left(-i\hbar \frac{\partial \psi_1}{\partial \rho} \right)^* \psi_2 \rho d\rho d\phi \\ &= \langle \mathbf{P}_\rho \psi_1 | \psi_2 \rangle \end{aligned}$$

(You may assume $\rho \psi_1^* \psi_2 \rightarrow 0$ as $\rho \rightarrow 0$ or ∞ . The problem comes from the fact that $\rho d\rho d\phi$ and not $d\rho d\phi$ is the measure for integration.)

[‡] What we will see is that $P_\rho = -i\hbar d/d\rho$, and hence the H constructed with it, are non-Hermitian.

Show, however, that

$$P_\rho \rightarrow -i\hbar \left(\frac{\partial}{\partial \rho} + \frac{1}{2\rho} \right) \quad (7.4.43)$$

is indeed Hermitian and also satisfies the canonical commutation rule. The angular momentum $P_\phi \rightarrow -i\hbar \partial/\partial\phi$ is Hermitian, as it stands, on single-valued functions: $\psi(\rho, \phi) = \psi(\rho, \phi + 2\pi)$.

(3) In the Cartesian case we saw that adding an arbitrary $f(x)$ to $-i\hbar \partial/\partial x$ didn't have any physical effect, whereas here the addition of a function of ρ to $-i\hbar \partial/\partial\rho$ seems important. Why? [Is $f(x)$ completely arbitrary? Mustn't it be real? Why? Is the same true for the $-i\hbar/2\rho$ piece?]

(4) Feed in the new momentum operator P_ρ and show that

$$H \xrightarrow[\text{coordinate basis}]{} \frac{-\hbar^2}{2m} \left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} - \frac{1}{4\rho^2} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \right) + a\rho$$

which still disagrees with Eq. (7.4.41). We have satisfied the commutation rules, chosen Hermitian operators, and yet do not get the right quantum Hamiltonian. The key to the mystery lies in the fact that \mathcal{H} doesn't determine H uniquely since terms of order \hbar (or higher) may be present in H but absent in \mathcal{H} . While this ambiguity is present even in the Cartesian case, it is resolved by symmetrization in all interesting cases. With non-Cartesian coordinates the ambiguity is more severe. There are ways of constructing H given \mathcal{H} (the path integral formulation suggests one) such that the substitution $P_\rho \rightarrow -i\hbar(\partial/\partial\rho + 1/2\rho)$ leads to Eq. (7.4.41). In the present case the quantum Hamiltonian corresponding to

$$\mathcal{H} = \frac{p_\rho^2}{2m} + \frac{p_\phi^2}{2m\rho^2} + a\rho$$

is given by

$$H \xrightarrow[\text{coordinate basis}]{} \mathcal{H} \left(\rho \rightarrow \rho, p_\rho \rightarrow -i\hbar \left[\frac{\partial}{\partial \rho} + \frac{1}{2\rho} \right]; \phi \rightarrow \phi, p_\phi \rightarrow -i\hbar \frac{\partial}{\partial \phi} \right) - \frac{\hbar^2}{8m\rho^2} \quad (7.4.44)$$

Notice that the additional term is indeed of nonzero order in \hbar .

We will not get into a discussion of these prescriptions for generating H since they finally reproduce results more readily available in the approach we are adopting. \square

7.5. Passage from the Energy Basis to the X Basis

It was remarked in the last section that although the $|n\rangle$ basis was ideally suited for evaluating the matrix elements of operators between oscillator eigenstates, the amplitude for finding the particle in a state $|n\rangle$ at the point x could not be readily computed: it seemed as if one had to find the eigenkets $|x\rangle$ of the operators X [Eq. (7.4.32)] and then take the inner product $\langle x|n\rangle$. But there is a more direct way to get $\psi_n(x) = \langle x|n\rangle$.

We start by projecting the equation defining the ground state of the oscillator

$$a|0\rangle = 0 \quad (7.5.1)$$

on the X basis:

$$\begin{aligned} |0\rangle \rightarrow \langle x|0\rangle &= \psi_0(x) \\ a &= \left(\frac{m\omega}{2\hbar} \right)^{1/2} X + i \left(\frac{1}{2m\omega\hbar} \right)^{1/2} P \\ &\rightarrow \left(\frac{m\omega}{2\hbar} \right)^{1/2} x + \left(\frac{\hbar}{2m\omega} \right)^{1/2} \frac{d}{dx} \end{aligned} \quad (7.5.2)$$

In terms of $y = (m\omega/\hbar)^{1/2}x$,

$$a = \frac{1}{2^{1/2}} \left(y + \frac{d}{dy} \right) \quad (7.5.3)$$

For later use we also note that (since d/dy is *anti-Hermitian*),

$$a^\dagger = \frac{1}{2^{1/2}} \left(y - \frac{d}{dy} \right) \quad (7.5.4)$$

In the X basis Eq. (7.5.1) then becomes

$$\left(y + \frac{d}{dy} \right) \psi_0(y) = 0 \quad (7.5.5)$$

or

$$\frac{d\psi_0(y)}{\psi_0(y)} = -y dy$$

or

$$\psi_0(y) = A_0 e^{-y^2/2}$$

or

$$\psi_0(x) = A_0 \exp\left(-\frac{m\omega x^2}{2\hbar}\right)$$

or (upon normalizing)

$$= \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \exp\left(-\frac{m\omega x^2}{2\hbar}\right) \quad (7.5.6)$$

By projecting the equation

$$|n\rangle = \frac{(a^\dagger)^n}{(n!)^{1/2}} |0\rangle$$

onto the X basis, we get the *normalized* eigenfunctions

$$\langle x|n\rangle = \psi_n \left[x = \left(\frac{\hbar}{m\omega} \right)^{1/2} y \right] = \frac{1}{(n!)^{1/2}} \left[\frac{1}{2^{1/2}} \left(y - \frac{d}{dy} \right) \right]^n \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-y^2/2} \quad (7.5.7)$$

A comparison of the above result with Eq. (7.3.22) shows that

$$H_n(y) = e^{y^2/2} \left(y - \frac{d}{dy} \right)^n e^{-y^2/2} \quad (7.5.8)$$

We now conclude our rather lengthy discussion of the oscillator. If you understand this chapter thoroughly, you should have a good grasp of how quantum mechanics works.

Exercise 7.5.1. Project Eq. (7.5.1) on the P basis and obtain $\psi_0(p)$.

Exercise 7.5.2. Project the relation

$$a|n\rangle = n^{1/2}|n-1\rangle$$

on the X basis and derive the recursion relation

$$H_n(y) = 2nH_{n-1}(y)$$

using Eq. (7.3.22).

Exercise 7.5.3. Starting with

$$a + a^\dagger = 2^{1/2}y$$

and

$$(a + a^\dagger)|n\rangle = n^{1/2}|n-1\rangle + (n+1)^{1/2}|n+1\rangle$$

and Eq. (7.3.22), derive the relation

$$H_{n+1}(y) = 2yH_n(y) - 2nH_{n-1}(y)$$

Exercise 7.5.4. Thermodynamics of Oscillators.* The Boltzmann formula

$$P(i) = e^{-\beta E(i)} / Z$$

where

$$Z = \sum_i e^{-\beta E(i)}$$

gives the probability of finding a system in a state i with energy $E(i)$, when it is in thermal equilibrium with a reservoir of absolute temperature $T = 1/\beta k$, $k = 1.4 \times 10^{-16}$ ergs/ $^\circ$ K; being Boltzman's constant. (The "probability" referred to above is in relation to a classical ensemble of similar systems and has nothing to do with quantum mechanics.)

(1) Show that the thermal average of the system's energy is

$$\bar{E} = \sum_i E(i) P(i) = \frac{-\partial}{\partial \beta} \ln Z$$

(2) Let the system be a classical oscillator. The index i is now continuous and corresponds to the variables x and p describing the state of the oscillator, i.e.,

$$i \rightarrow x, p$$

and

$$\sum_i \rightarrow \iint dx dp$$

and

$$E(i) \rightarrow E(x, p) = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2$$

Show that

$$Z_{\text{cl}} = \left(\frac{2\pi}{\beta m\omega^2} \right)^{1/2} \left(\frac{2\pi m}{\beta} \right)^{1/2} = \frac{2\pi}{\omega\beta}$$

and that

$$\bar{E}_{\text{cl}} = \frac{1}{\beta} = kT$$

Note that E_{cl} is independent of m and ω .

(3) For the quantum oscillator the quantum number n plays the role of the index i . Show that

$$Z_{\text{qu}} = e^{-\beta \hbar \omega / 2} (1 - e^{-\beta \hbar \omega})^{-1}$$

and

$$\bar{E}_{\text{qu}} = \hbar \omega \left(\frac{1}{2} + \frac{1}{e^{\beta \hbar \omega} - 1} \right)$$

(4) It is intuitively clear that as the temperature T increases (and $\beta = 1/kT$ decreases) the oscillator will get more and more excited and eventually (from the correspondence principle)

$$\bar{E}_{\text{qu}} \xrightarrow{T \rightarrow \infty} \bar{E}_{\text{cl}}$$

Verify that this is indeed true and show that “large T ” means $T \gg \hbar \omega / k$.

(5) Consider a crystal with N_0 atoms, which, for small oscillations, is equivalent to $3N_0$ decoupled oscillators. The mean thermal energy of the crystal \bar{E}_{crystal} is \bar{E}_{cl} or \bar{E}_{qu} summed over all the normal modes. Show that if the oscillators are treated classically, the specific heat per atom is

$$C_{\text{cl}}(T) = \frac{1}{N_0} \frac{\partial \bar{E}_{\text{crystal}}}{\partial T} = 3k$$

which is independent of T and the parameters of the oscillators and hence the same for all crystals.[‡] This agrees with experiment at high temperatures but not as $T \rightarrow 0$. Empirically,

$$\begin{aligned} C(T) &\rightarrow 3k & (T \text{ large}) \\ &\rightarrow 0 & (T \rightarrow 0) \end{aligned}$$

Following Einstein, treat the oscillators quantum mechanically, assuming for simplicity that they all have the same frequency ω . Show that

$$C_{\text{qu}}(T) = 3k \left(\frac{\theta_E}{T} \right)^2 \frac{e^{\theta_E/T}}{(e^{\theta_E/T} - 1)^2}$$

where $\theta_E = \hbar \omega / k$ is called the *Einstein temperature* and varies from crystal to crystal. Show that

$$\begin{aligned} C_{\text{qu}}(T) &\xrightarrow{T \gg \theta_E} 3k \\ C_{\text{qu}}(T) &\xrightarrow{T \ll \theta_E} 3k \left(\frac{\theta_E}{T} \right)^2 e^{-\theta_E/T} \end{aligned}$$

Although $C_{\text{qu}}(T) \rightarrow 0$ as $T \rightarrow 0$, the exponential falloff disagrees with the observed $C(T) \rightarrow_{T \rightarrow 0} T^3$ behavior. This discrepancy arises from assuming that the frequencies of all

[‡] More precisely, for crystals whose atoms behave as point particles with no internal degrees of freedom.

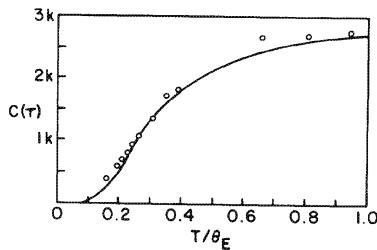


Figure 7.3. Comparison of experiment with Einstein's theory for the specific heat in the case of diamond. (θ_E is chosen to be 1320 K.)

normal modes are equal, which is of course not generally true. [Recall that in the case of two coupled masses we get $\omega_1 = (k/m)^{1/2}$ and $\omega_{11} = (3k/m)^{1/2}$.] This discrepancy was eliminated by Debye.

But Einstein's simple picture by itself is remarkably successful (see Fig. 7.3).

The Path Integral Formulation of Quantum Theory

We consider here an alternate formulation of quantum mechanics invented by Feynman in the forties.[‡] In contrast to the Schrödinger formulation, which stems from Hamiltonian mechanics, the Feynman formulation is tied to the Lagrangian formulation of mechanics. Although we are committed to the former approach, we discuss in this chapter Feynman's alternative, not only because of its aesthetic value, but also because it can, in a class of problems, give the full propagator with tremendous ease and also give valuable insight into the relation between classical and quantum mechanics.

8.1. The Path Integral Recipe

We have already seen that the quantum problem is fully solved once the propagator is known. Thus far our practice has been to first find the eigenvalues and eigenfunctions of H , and then express the propagator $U(t)$ in terms of these. In the path integral approach one computes $U(t)$ directly. For a single particle in one dimension, the procedure is the following.

To find $U(x, t; x', t')$:

- (1) Draw all paths in the x - t plane connecting (x', t') and (x, t) (see Fig. 8.1).
- (2) Find the action $S[x(t)]$ for each path $x(t)$.
- (3)
$$U(x, t; x', t') = A \sum_{\text{all paths}} e^{iS[x(t)]/\hbar} \quad (8.1.1)$$

where A is an overall normalization factor.

[‡] The nineteen forties that is, and in his twenties. An interesting account of how he was influenced by Dirac's work in the same direction may be found in his Nobel lectures. See, *Nobel Lectures—Physics*, Vol. III, Elsevier Publication, New York (1972).

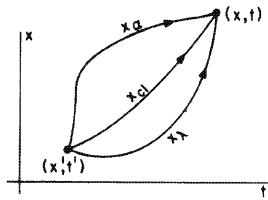


Figure 8.1. Some of the paths that contribute to the propagator. The contribution from the path $x(t)$ is $Z = \exp\{iS[x(t)]/\hbar\}$.

8.2. Analysis of the Recipe

Let us analyze the above recipe, postponing for a while the proof that it reproduces conventional quantum mechanics. The most surprising thing about it is the fact that every path, including the classical path, $x_{cl}(t)$, gets the same weight, that is to say, a number of unit modulus. How are we going to regain classical mechanics in the appropriate limit if the classical path does not seem favored in any way?

To understand this we must perform the sum in Eq. (8.1.1). Now, the correct way to sum over all the paths, that is to say, path integration, is quite complicated and we will discuss it later. For the present let us take the heuristic approach. Let us first pretend that the continuum of paths linking the end points is actually a discrete set. A few paths in the set are shown in Fig. 8.1.

We have to add the contributions $Z_a = e^{iS[x_a(t)]/\hbar}$ from each path $x_a(t)$. This summation is done schematically in Fig. 8.2. Since each path has a different action, it contributes with a different phase, and the contributions from the paths essentially cancel each other, until we come near the classical path. Since S is stationary here, the Z 's add constructively and produce a large sum. As we move away from $x_{cl}(t)$, destructive interference sets in once again. It is clear from the figure that $U(t)$ is dominated by the paths near $x_{cl}(t)$. Thus the classical path is important, not because it contributes a lot by itself, but because in its vicinity the paths contribute coherently.

How far must we deviate from x_{cl} before destructive interference sets in? One may say crudely that coherence is lost once the phase differs from the stationary value $S[x_{cl}(t)]/\hbar \equiv S_{cl}/\hbar$ by about π . This in turn means that the action for the coherence paths must be within $\hbar\pi$ of S_{cl} . For a macroscopic particle this means a very tight constraint on its path, since S_{cl} is typically $\approx 1 \text{ erg sec} \approx 10^{27}\hbar$, while for an electron there is quite a bit of latitude. Consider the following example. A free particle leaves the origin at $t=0$ and arrives at $x=1 \text{ cm}$ at $t=1 \text{ second}$. The classical path is

$$x = t \quad (8.2.1)$$

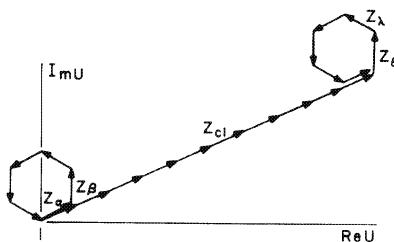
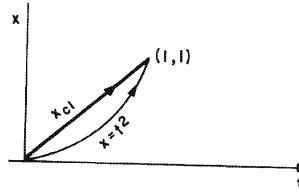


Figure 8.2. Schematic representation of the sum $\sum Z_a$. Paths near $x_{cl}(t)$ contribute coherently since S is stationary there, while others cancel each other out and may be ignored in the first approximation when we calculate $U(t)$.

Figure 8.3. Two possible paths connecting $(0, 0)$ and $(1, 1)$. The action on the classical path $x = t$ is $m/2$, while on the other, it is $2m/3$.



Consider another path

$$x = t^2 \quad (8.2.2)$$

which also links the two space-time points (Fig. 8.3.)

For a classical particle, of mass, say 1 g, the action changes by roughly $1.6 \times 10^{26}\hbar$, and the phase by roughly 1.6×10^{26} rad as we move from the classical path $x = t$ to the nonclassical path $x = t^2$. We may therefore completely ignore the nonclassical path. On the other hand, for an electron whose mass is $\approx 10^{-27}$ g, $\delta S \approx \hbar/6$ and the phase change is just around a sixth of a radian, which is well within the coherence range $\delta S/\hbar \lesssim \pi$. It is in such cases that assuming that the particle moves along a well-defined trajectory, $x_{\text{cl}}(t)$, leads to conflict with experiment.

8.3. An Approximation to $U(t)$ for a Free Particle

Our previous discussions have indicated that, to an excellent approximation, we may ignore all but the classical path and its neighbors in calculating $U(t)$. Assuming that each of these paths contributes the same amount $\exp(iS_{\text{cl}}/\hbar)$, since S is stationary, we get

$$U(t) = A' e^{iS_{\text{cl}}/\hbar} \quad (8.3.1)$$

where A' is some normalizing factor which “measures” the number of paths in the coherent range. Let us find $U(t)$ for a free particle in this approximation and compare the result with the exact result, Eq. (5.1.10).

The classical path for a free particle is just a straight line in the x - t plane:

$$x_{\text{cl}}(t'') = x' + \frac{x - x'}{t - t'} (t'' - t') \quad (8.3.2)$$

corresponding to motion with uniform velocity $v = (x - x')/(t - t')$. Since $\mathcal{L} = mv^2/2$ is a constant,

$$S_{\text{cl}} = \int_{t'}^t \mathcal{L} dt'' = \frac{1}{2} m \frac{(x - x')^2}{t - t'}$$

so that

$$U(x, t; x', t') = A' \exp \left[\frac{im(x - x')^2}{2\hbar(t - t')} \right] \quad (8.3.3)$$

To find A' , we use the fact that as $t - t'$ tends to 0, U must tend to $\delta(x - x')$. Comparing Eq. (8.3.3) to the representation of the delta function encountered in Section 1.10 (see footnote on page 61),

$$\delta(x - x') \equiv \lim_{\Delta \rightarrow 0} \frac{1}{(\pi \Delta^2)^{1/2}} \exp \left[-\frac{(x - x')^2}{\Delta^2} \right]$$

(valid even if Δ is imaginary) we get

$$A' = \left[\frac{m}{2\pi\hbar i(t - t')} \right]^{1/2}$$

so that

$$U(x, t; x', 0) \equiv U(x, t; x') = \left(\frac{m}{2\pi\hbar it} \right)^{1/2} \exp \left[\frac{im(x - x')^2}{2\hbar t} \right] \quad (8.3.4)$$

which is the exact answer! We have managed to get the exact answer by just computing the classical action! However, we will see in Section 8.6 that only for potentials of the form $V = a + bx + cx^2 + d\dot{x} + ex\dot{x}$ is it true that $U(t) = A(t) e^{iS_0/\hbar}$. Furthermore, we can't generally find $A(t)$ using $U(x, 0; x') = \delta(x - x')$ since A can contain an arbitrary dimensionless function f such that $f \rightarrow 1$ as $t \rightarrow 0$. Here $f \equiv 1$ because we can't construct a nontrivial dimensionless f using just m , \hbar , and t (check this).

8.4. Path Integral Evaluation of the Free-Particle Propagator

Although our heuristic analysis yielded the exact free-particle propagator, we will now repeat the calculation without any approximation to illustrate path integration.

Consider $U(x_N, t_N; x_0, t_0)$. The peculiar labeling of the end points will be justified later. Our problem is to perform the path integral

$$\int_{x_0}^{x_N} e^{iS[x(t)]/\hbar} \mathcal{D}[x(t)] \quad (8.4.1)$$

where

$$\int_{x_0}^{x_N} \mathcal{D}[x(t)]$$

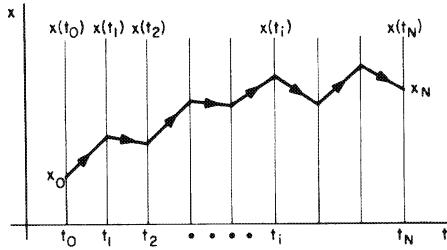


Figure 8.4. The discrete approximation to a path $x(t)$. Each path is specified by $N-1$ numbers $x(t_1), \dots, x(t_{N-1})$. To sum over paths we must integrate each x_i from $-\infty$ to $+\infty$. Once all integrations are done, we can take the limit $N \rightarrow \infty$.

is a symbolic way of saying "integrate over all paths connecting x_0 and x_N (in the interval t_0 and t_N)."
 Now, a path $x(t)$ is fully specified by an infinity of numbers $x(t_0), \dots, x(t), \dots, x(t_N)$, namely, the values of the function $x(t)$ at every point t in the interval t_0 to t_N . To sum over all paths we must integrate over all possible values of these infinite variables, except of course $x(t_0)$ and $x(t_N)$, which will be kept fixed at x_0 and x_N , respectively. To tackle this problem, we follow the idea that was used in Section 1.10: we trade the function $x(t)$ for a discrete approximation which agrees with $x(t)$ at the $N+1$ points $t_n = t_0 + n\varepsilon$, $n=0, \dots, N$, where $\varepsilon = (t_N - t_0)/N$. In this approximation each path is specified by $N+1$ numbers $x(t_0), x(t_1), \dots, x(t_N)$. The gaps in the discrete function are interpolated by straight lines. One such path is shown in Fig. 8.4. We hope that if we take the limit $N \rightarrow \infty$ at the end we will get a result that is insensitive to these approximations.[‡] Now that the paths have been discretized, we must also do the same to the action integral. We replace the continuous path definition

$$S = \int_{t_0}^{t_N} \mathcal{L}(t) dt = \int_{t_0}^{t_N} \frac{1}{2} m \dot{x}^2 dt$$

by

$$S = \sum_{i=0}^{N-1} \frac{m}{2} \left(\frac{x_{i+1} - x_i}{\varepsilon} \right)^2 \quad (8.4.2)$$

where $x_i = x(t_i)$. We wish to calculate

$$\begin{aligned} U(x_N, t_N; x_0, t_0) &= \int_{x_0}^{x_N} \exp\{iS[x(t)]/\hbar\} \mathcal{D}[x(t)] \\ &= \lim_{\substack{N \rightarrow \infty \\ \varepsilon \rightarrow 0}} A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[\frac{i}{\hbar} \frac{m}{2} \sum_{i=0}^{N-1} \frac{(x_{i+1} - x_i)^2}{\varepsilon} \right] \\ &\quad \times dx_1 \cdots dx_{N-1} \end{aligned} \quad (8.4.3)$$

[‡] We expect that the abrupt changes in velocity at the points $t_0 + n\varepsilon$ that arise due to our approximation will not matter because \mathcal{L} does not depend on the acceleration or higher derivatives.

It is implicit in the above that x_0 and x_N have the values we have chosen at the outset. The factor A in the front is to be chosen at the end such that we get the correct scale for U when the limit $N \rightarrow \infty$ is taken.

Let us first switch to the variables

$$y_i = \left(\frac{m}{2\hbar\varepsilon} \right)^{1/2} x_i$$

We then want

$$\lim_{N \rightarrow \infty} A' \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left[- \sum_{i=0}^{N-1} \frac{(y_{i+1} - y_i)^2}{i} \right] dy_1 \cdots dy_{N-1} \quad (8.4.4)$$

where

$$A' = A \left(\frac{2\hbar\varepsilon}{m} \right)^{(N-1)/2}$$

Although the multiple integral looks formidable, it is not. Let us begin by doing the y_1 integration. Considering just the part of the integrand that involves y_1 , we get

$$\int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{i} [(y_2 - y_1)^2 + (y_1 - y_0)^2] \right\} dy_1 = \left(\frac{i\pi}{2} \right)^{1/2} e^{-(y_2 - y_0)^2 / 2i} \quad (8.4.5)$$

Consider next the integration over y_2 . Bringing in the part of the integrand involving y_2 and combining it with the result above we compute next

$$\begin{aligned} & \left(\frac{i\pi}{2} \right)^{1/2} \int_{-\infty}^{\infty} e^{-(y_3 - y_2)^2 / i} \cdot e^{-(y_2 - y_0)^2 / 2i} dy_2 \\ &= \left(\frac{i\pi}{2} \right)^{1/2} e^{-(2y_3^2 + y_0^2) / 2i} \left(\frac{2\pi i}{3} \right)^{1/2} e^{(y_0 + 2y_3)^2 / 6i} \\ &= \left[\frac{(i\pi)^2}{3} \right]^{1/2} e^{-(y_3 - y_0)^2 / 3i} \end{aligned} \quad (8.4.6)$$

By comparing this result to the one from the y_1 integration, we deduce the pattern: if we carry out this process $N-1$ times so as to evaluate the integral in Eq. (8.4.4), it will become

$$\frac{(i\pi)^{(N-1)/2}}{N^{1/2}} e^{-(y_N - y_0)^2 / Ni}$$

or

$$\frac{(i\pi)^{(N-1)/2}}{N^{1/2}} e^{-m(x_N - x_0)^2/2\hbar\varepsilon Ni}$$

Bringing in the factor $A(2\hbar\varepsilon/m)^{(N-1)/2}$ from up front, we get

$$U = A \left(\frac{2\pi\hbar\varepsilon i}{m} \right)^{N/2} \left(\frac{m}{2\pi\hbar\varepsilon Ni} \right)^{1/2} \exp \left[\frac{im(x_N - x_0)^2}{2\hbar N\varepsilon} \right]$$

If we now let $N \rightarrow \infty$, $\varepsilon \rightarrow 0$, $N\varepsilon \rightarrow t_N - t_0$, we get the right answer provided

$$A = \left[\frac{2\pi\hbar\varepsilon i}{m} \right]^{-N/2} \equiv B^{-N} \quad (8.4.7)$$

It is conventional to associate a factor $1/B$ with each of the $N-1$ integrations and the remaining factor $1/B$ with the overall process. In other words, we have just learnt that the precise meaning of the statement “integrate over all paths” is

$$\int \mathcal{D}[x(t)] = \lim_{\substack{\varepsilon \rightarrow 0 \\ N \rightarrow \infty}} \frac{1}{B} \int_{-\infty}^{\infty} \int \int \dots \int_{-\infty}^{\infty} \frac{dx_1}{B} \cdot \frac{dx_2}{B} \dots \frac{dx_{N-1}}{B}$$

where

$$B = \left(\frac{2\pi\hbar\varepsilon i}{m} \right)^{1/2} \quad (8.4.8)$$

8.5. Equivalence to the Schrödinger Equation

The relation between the Schrödinger and Feynman formalisms is quite similar to that between the Newtonian and the least action formalisms of mechanics, in that the former approach is local in time and deals with time evolution over infinitesimal periods while the latter is global and deals directly with propagation over finite times.

In the Schrödinger formalism, the change in the state vector $|\psi\rangle$ over an infinitesimal time ε is

$$|\psi(\varepsilon)\rangle - |\psi(0)\rangle = \frac{-i\varepsilon}{\hbar} H |\psi(0)\rangle \quad (8.5.1)$$

which becomes in the X basis

$$\psi(x, \varepsilon) - \psi(x, 0) = \frac{-i\varepsilon}{\hbar} \left[\frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x, 0) \right] \psi(x, 0) \quad (8.5.2)$$

to first order in ε . To compare this result with the path integral prediction to the same order in ε , we begin with

$$\psi(x, \varepsilon) = \int_{-\infty}^{\infty} U(x, \varepsilon; x') \psi(x', 0) dx' \quad (8.5.3)$$

The calculation of $U(\varepsilon)$ is simplified by the fact that there is no need to do any integrations over intermediate x 's since there is just one slice of time ε between the start and finish. So

$$U(x, \varepsilon; x') = \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} \exp \left\{ i \left[\frac{m(x-x')^2}{2\varepsilon} - \varepsilon V\left(\frac{x+x'}{2}, 0\right) \right] / \hbar \right\} \quad (8.5.4)$$

where the $(m/2\pi\hbar i\varepsilon)^{1/2}$ factor up front is just the $1/B$ factor from Eq. (8.4.8). We take the time argument of V to be zero since there is already a factor of ε before it and any variation of V with time in the interval 0 to ε will produce an effect of second order in ε . So

$$\begin{aligned} \psi(x, \varepsilon) &= \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} \int_{-\infty}^{\infty} \exp \left[\frac{i m (x-x')^2}{2\varepsilon\hbar} \right] \exp \left[-\frac{i\varepsilon}{\hbar} V\left(\frac{x+x'}{2}, 0\right) \right] \\ &\quad \times \psi(x', 0) dx' \end{aligned} \quad (8.5.5)$$

Consider the factor $\exp[i m (x-x')^2 / 2\varepsilon\hbar]$. It oscillates very rapidly as $(x-x')$ varies since ε is infinitesimal and \hbar is so small. When such a rapidly oscillating function multiplies a smooth function like $\psi(x', 0)$, the integral vanishes for the most part due to the random phase of the exponential. Just as in the case of the path integration, the only substantial contribution comes from the region where the phase is stationary. In this case the only stationary point is $x=x'$, where the phase has the minimum value of zero. In terms of $\eta=x'-x$, the region of coherence is, as before,

$$\frac{m\eta^2}{2\varepsilon\hbar} \lesssim \pi$$

or

$$|\eta| \lesssim \left(\frac{2\varepsilon\hbar\pi}{m} \right)^{1/2} \quad (8.5.6)$$

Consider now

$$\begin{aligned} \psi(x, \varepsilon) &= \left(\frac{m}{2\pi\hbar i\varepsilon} \right)^{1/2} \int_{-\infty}^{\infty} \exp(i m \eta^2 / 2\varepsilon\hbar) \cdot \exp \left[-\left(\frac{i}{\hbar} \right) \varepsilon V\left(x + \frac{\eta}{2}, 0\right) \right] \\ &\quad \times \psi(x + \eta, 0) d\eta \end{aligned} \quad (8.5.7)$$

We will work to first order in ε and therefore to second order in η [see Eq. (8.5.6) above]. We expand

$$\begin{aligned}\psi(x + \eta, 0) &= \psi(x, 0) + \eta \frac{\partial \psi}{\partial x} + \frac{\eta^2}{2} \frac{\partial^2 \psi}{\partial x^2} + \dots \\ \exp\left[-\left(\frac{i}{\hbar}\right)\varepsilon V\left(x + \frac{\eta}{2}, 0\right)\right] &= 1 - \frac{i\varepsilon}{\hbar} V\left(x + \frac{\eta}{2}, 0\right) + \dots \\ &= 1 - \frac{i\varepsilon}{\hbar} V(x, 0) + \dots\end{aligned}$$

since terms of order $\eta\varepsilon$ are to be neglected. Equation (8.5.7) now becomes

$$\begin{aligned}\psi(x, \varepsilon) &= \left(\frac{m}{2\pi\hbar i\varepsilon}\right)^{1/2} \int_{-\infty}^{\infty} \exp\left(\frac{im\eta^2}{2\hbar\varepsilon}\right) \left[\psi(x, 0) - \frac{i\varepsilon}{\hbar} V(x, 0) \psi(x, 0) \right. \\ &\quad \left. + \eta \frac{\partial \psi}{\partial x} + \frac{\eta^2}{2} \frac{\partial^2 \psi}{\partial x^2} \right] d\eta\end{aligned}$$

Consulting the list of Gaussian integrals in Appendix A.2, we get

$$\begin{aligned}\psi(x, \varepsilon) &= \left(\frac{m}{2\pi\hbar i\varepsilon}\right)^{1/2} \left[\psi(x, 0) \left(\frac{2\pi\hbar i\varepsilon}{m}\right)^{1/2} - \frac{\hbar\varepsilon}{2im} \left(\frac{2\pi\hbar i\varepsilon}{m}\right)^{1/2} \frac{\partial^2 \psi}{\partial x^2} \right. \\ &\quad \left. - \frac{i\varepsilon}{\hbar} \left(\frac{2\pi\hbar i\varepsilon}{m}\right)^{1/2} V(x, 0) \psi(x, 0) \right]\end{aligned}$$

or

$$\psi(x, \varepsilon) - \psi(x, 0) = \frac{-i\varepsilon}{\hbar} \left[\frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x, 0) \right] \psi(x, 0) \quad (8.5.8)$$

which agrees with the Schrödinger prediction, Eq. (8.5.1).

8.6. Potentials of the Form $V = a + bx + cx^2 + dx\dot{x} + ex\dot{x}^2$ [‡]

We wish to compute

$$U(x, t; x') = \int_{x'}^x e^{iS[x(t')]/\hbar} \mathcal{D}[x(t')] \quad (8.6.1)$$

[‡] This section may be omitted without loss of continuity.

Let us write every path as

$$x(t'') = x_{\text{cl}}(t'') + y(t'') \quad (8.6.2)$$

It follows that

$$\dot{x}(t'') = \dot{x}_{\text{cl}}(t'') + \dot{y}(t'') \quad (8.6.3)$$

Since all the paths agree at the end points, $y(0) = y(t) = 0$. When we slice up the time into N parts, we have for intermediate integration variables

$$x_i \equiv x(t''_i) = x_{\text{cl}}(t''_i) + y(t''_i) \equiv x_{\text{cl}}(t''_i) + y_i$$

Since $x_{\text{cl}}(t''_i)$ is just some constant at t''_i ,

$$dx_i = dy_i$$

and

$$\int_{x'}^x \mathcal{D}[x(t'')] = \int_0^0 \mathcal{D}[y(t'')] \quad (8.6.4)$$

so that Eq. (8.6.1) becomes

$$U(x, t; x') = \int_0^0 \exp \left\{ \frac{i}{\hbar} S[x_{\text{cl}}(t'') + y(t'')] \right\} \mathcal{D}[y(t'')] \quad (8.6.5)$$

The next step is to expand the functional S in a Taylor series about x_{cl} :

$$\begin{aligned} S[x_{\text{cl}} + y] &= \int_0^t \mathcal{L}(x_{\text{cl}} + y, \dot{x}_{\text{cl}} + \dot{y}) dt'' \\ &\equiv \int_0^t \left[\mathcal{L}(x_{\text{cl}}, \dot{x}_{\text{cl}}) + \left(\frac{\partial \mathcal{L}}{\partial x} \Big|_{x_{\text{cl}}} y + \frac{\partial \mathcal{L}}{\partial \dot{x}} \Big|_{x_{\text{cl}}} \dot{y} \right) \right. \\ &\quad \left. + \frac{1}{2} \left(\frac{\partial^2 \mathcal{L}}{\partial x^2} \Big|_{x_{\text{cl}}} y^2 + 2 \frac{\partial^2 \mathcal{L}}{\partial x \partial \dot{x}} \Big|_{x_{\text{cl}}} y \dot{y} + \frac{\partial^2 \mathcal{L}}{\partial \dot{x}^2} \Big|_{x_{\text{cl}}} \dot{y}^2 \right) \right] dt'' \end{aligned} \quad (8.6.6)$$

The series terminates here since \mathcal{L} is a quadratic polynomial.

The first piece $\mathcal{L}(x_{\text{cl}}, \dot{x}_{\text{cl}})$ integrates to give $S[x_{\text{cl}}] \equiv S_{\text{cl}}$. The second piece, linear in y and \dot{y} , vanishes due to the classical equation of motion. In the last piece, if we recall

$$\mathcal{L} = \frac{1}{2} m \dot{x}^2 - a - bx - cx^2 - d \dot{x} - ex \dot{x} \quad (8.6.7)$$

we get

$$\frac{1}{2} \frac{\partial^2 \mathcal{L}}{\partial x^2} = -c \quad (8.6.8)$$

$$\frac{\partial^2 \mathcal{L}}{\partial x \partial \dot{x}} = -e \quad (8.6.9)$$

$$\frac{1}{2} \frac{\partial^2 \mathcal{L}}{\partial \dot{x}^2} = m \quad (8.6.10)$$

Consequently Eq. (8.6.5) becomes

$$U(x, t; x') = \exp\left(\frac{iS_{cl}}{\hbar}\right) \int_0^0 \exp\left[\frac{i}{\hbar} \int_0^t \left(\frac{1}{2} m \dot{y}^2 - cy^2 - ey \dot{y}\right) dt''\right] \times \mathcal{D}[y(t'')] \quad (8.6.11)$$

Since the path integral has no memory of x_{cl} , it can only depend on t . So

$$U(x, t; x') = e^{iS_{cl}/\hbar} A(t) \quad (8.6.12)$$

where $A(t)$ is some unknown function of t . Now if we were doing the free-particle problem, we would get Eq. (8.6.11) with $c=e=0$. In this case we know that [see Eq. (8.3.4)]

$$A(t) = \left(\frac{m}{2\pi\hbar i t}\right)^{1/2} \quad (8.6.13)$$

Since the coefficient b does not figure in Eq. (8.6.11), it follows that the same value of $A(t)$ corresponds to the linear potential $V=a+bx$ as well. For the harmonic oscillator, $c=\frac{1}{2}m\omega^2$, and we have to do the integral

$$A(t) = \int_0^0 \exp\left[i/\hbar \int_0^t \frac{1}{2} m(\dot{y}^2 - \omega^2 y^2)\right] dt'' \mathcal{D}[y(t'')] \quad (8.6.14)$$

The evaluation of this integral is discussed in the book by Feynman and Hibbs referred to at the end of this section. Note that even if the factor $A(t)$ in $\psi(x, t)$ is not known, we can extract all the probabilistic information at time t .

Notice the ease with which the Feynman formalism yields the full propagator in these cases. Consider in particular the horrendous alternative of finding the eigenfunctions of the Hamiltonian and constructing from them the harmonic oscillator propagator.

The path integral method may be extended to three dimensions without any major qualitative differences. In particular, the form of U in Eq. (8.6.12) is valid for potentials that are at most quadratic in the coordinates and the velocities. An

interesting problem in this class is that of a particle in a uniform magnetic field. For further details on the subject of path integral quantum mechanics, see R. P. Feynman and A. R. Hibbs, *Path Integrals and Quantum Mechanics*, McGraw-Hill (1965), and Chapter 21.

*Exercise 8.6.1.** Verify that

$$U(x, t; x', 0) = A(t) \exp(iS_{\text{cl}}/\hbar), \quad A(t) = \left(\frac{m}{2\pi\hbar i t} \right)^{1/2}$$

agrees with the exact result, Eq. (5.4.31), for $V(x) = -fx$. Hint: Start with $x_{\text{cl}}(t'') = x_0 + v_0 t'' + \frac{1}{2}(f/m)t''^2$ and find the constants x_0 and v_0 from the requirement that $x_{\text{cl}}(0) = x'$ and $x_{\text{cl}}(t) = x$.

Exercise 8.6.2. Show that for the harmonic oscillator with

$$\mathcal{L} = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} m \omega^2 x^2$$

$$U(x, t; x') = A(t) \exp \left\{ \frac{im\omega}{2\hbar \sin \omega t} [(x^2 + x'^2) \cos \omega t - 2xx'] \right\}$$

where $A(t)$ is an unknown function. (Recall Exercise 2.8.7.)

Exercise 8.6.3. We know that given the eigenfunctions and the eigenvalues we can construct the propagator:

$$U(x, t; x', t') = \sum_n \psi_n(x) \psi_n^*(x') e^{-iE_n(t-t')/\hbar} \quad (8.6.15)$$

Consider the reverse process (since the path integral approach gives U directly), for the case of the oscillator.

(1) Set $x = x' = t' = 0$. Assume that $A(t) = (m\omega/2\pi i \hbar \sin \omega t)^{1/2}$ for the oscillator. By expanding both sides of Eq. (8.6.15), you should find that $E = \hbar\omega/2, 5\hbar\omega/2, 9\hbar\omega/2, \dots$, etc. What happened to the levels in between?

(2) (Optional). Now consider the extraction of the eigenfunctions. Let $x = x'$ and $t' = 0$. Find $E_0, E_1, |\psi_0(x)|^2$, and $|\psi_1(x)|^2$ by expanding in powers of $\alpha = \exp(i\omega t)$.

*Exercise 8.6.4.** Recall the derivation of the Schrödinger equation (8.5.8) starting from Eq. (8.5.4). Note that although we chose the argument of V to be the midpoint $x + x'/2$, it did not matter very much: any choice $x + \alpha\eta$, (where $\eta = x' - x$) for $0 \leq \alpha \leq 1$ would have given the same result since the difference between the choices is of order $\eta\varepsilon \simeq \varepsilon^{3/2}$. All this was thanks to the factor ε multiplying V in Eq. (8.5.4) and the fact that $|\eta| \simeq \varepsilon^{1/2}$, as per Eq. (8.6.5).

Consider now the case of a vector potential which will bring in a factor

$$\exp\left[\frac{iq\varepsilon}{\hbar c} \frac{x-x'}{\varepsilon} A(x+\alpha\eta)\right] \equiv \exp\left[-\frac{iq\varepsilon}{\hbar c} \frac{\eta}{\varepsilon} A(x+\alpha\eta)\right]$$

to the propagator for one time slice. (We should really be using vectors for position and the vector potential, but the one-dimensional version will suffice for making the point here.) Note that ε now gets canceled, in contrast to the scalar potential case. Thus, going to order ε to derive the Schrödinger equation means going to order η^2 in expanding the exponential. This will not only bring in an A^2 term, but will also make the answer sensitive to the argument of A in the linear term. Choose $\alpha = 1/2$ and verify that you get the one-dimensional version of Eq. (4.3.7). Along the way you will see that changing α makes an order ε difference to $\psi(x, \varepsilon)$ so that we have no choice but to use $\alpha = 1/2$, i.e., use the *midpoint prescription*. This point will come up in Chapter 21.

The Heisenberg Uncertainty Relations

9.1. Introduction

In classical mechanics a particle in a state (x_0, p_0) has associated with it well-defined values for any dynamical variable $\omega(x, p)$, namely, $\omega(x_0, p_0)$. In quantum theory, given a state $|\psi\rangle$, one can only give the probabilities $P(\omega)$ for the possible outcomes of a measurement of Ω . The probability distribution will be characterized by a mean or expectation value

$$\langle \Omega \rangle = \langle \psi | \Omega | \psi \rangle \quad (9.1.1)$$

and an uncertainty about this mean:

$$(\Delta \Omega) = [\langle \psi | (\Omega - \langle \Omega \rangle)^2 | \psi \rangle]^{1/2} \quad (9.1.2)$$

There are, however, states for which $\Delta \Omega = 0$, and these are the eigenstates $|\omega\rangle$ of Ω .

If we consider two Hermitian operators Ω and Λ , they will generally have some uncertainties $\Delta \Omega$ and $\Delta \Lambda$ in an arbitrary state. In the next section we will derive the Heisenberg uncertainty relations, which will provide a lower bound on the product of uncertainties, $\Delta \Omega \cdot \Delta \Lambda$. Generally the lower bound will depend not only on the operators but also on the state. Of interest to us are those cases in which the lower bound is independent of the state. The derivation will make clear the conditions under which such a relation will exist.

9.2. Derivation of the Uncertainty Relations

Let Ω and Λ be two Hermitian operators, with a commutator

$$[\Omega, \Lambda] = i\Gamma \quad (9.2.1)$$

You may readily verify that Γ is also Hermitian. Let us start with the uncertainty product in a normalized state $|\psi\rangle$:

$$(\Delta\Omega)^2(\Delta\Lambda)^2 = \langle\psi|(\Omega - \langle\Omega\rangle)^2|\psi\rangle\langle\psi|(\Lambda - \langle\Lambda\rangle)^2|\psi\rangle \quad (9.2.2)$$

where $\langle\Omega\rangle = \langle\psi|\Omega|\psi\rangle$ and $\langle\Lambda\rangle = \langle\psi|\Lambda|\psi\rangle$. Let us next define the pair

$$\begin{aligned}\hat{\Omega} &= \Omega - \langle\Omega\rangle \\ \hat{\Lambda} &= \Lambda - \langle\Lambda\rangle\end{aligned} \quad (9.2.3)$$

which has the same commutator as Ω and Λ (verify this). In terms of $\hat{\Omega}$ and $\hat{\Lambda}$

$$\begin{aligned}(\Delta\Omega)^2(\Delta\Lambda)^2 &= \langle\psi|\hat{\Omega}^2|\psi\rangle\langle\psi|\hat{\Lambda}^2|\psi\rangle \\ &= \langle\hat{\Omega}\psi|\hat{\Omega}\psi\rangle\langle\hat{\Lambda}\psi|\hat{\Lambda}\psi\rangle\end{aligned} \quad (9.2.4)$$

since

$$\hat{\Omega}^2 = \hat{\Omega}\hat{\Omega} = \hat{\Omega}^\dagger\hat{\Omega}$$

and

$$\hat{\Lambda}^2 = \hat{\Lambda}\hat{\Lambda} = \hat{\Lambda}^\dagger\hat{\Lambda} \quad (9.2.5)$$

If we apply the Schwartz inequality

$$|V_1|^2|V_2|^2 \geq |\langle V_1|V_2\rangle|^2 \quad (9.2.6)$$

(where the equality sign holds only if $|V_1\rangle = c|V_2\rangle$, where c is a constant) to the states $|\hat{\Omega}\psi\rangle$ and $|\hat{\Lambda}\psi\rangle$, we get from Eq. (9.2.4),

$$(\Delta\Omega)^2(\Delta\Lambda)^2 \geq |\langle\hat{\Omega}\psi|\hat{\Lambda}\psi\rangle|^2 \quad (9.2.7)$$

Let us now use the fact that

$$\langle\hat{\Omega}\psi|\hat{\Lambda}\psi\rangle = \langle\psi|\hat{\Omega}^\dagger\hat{\Lambda}|\psi\rangle = \langle\psi|\hat{\Omega}\hat{\Lambda}|\psi\rangle \quad (9.2.8)$$

to rewrite the above inequality as

$$(\Delta\Omega)^2(\Delta\Lambda)^2 \geq |\langle\psi|\hat{\Omega}\hat{\Lambda}|\psi\rangle|^2 \quad (9.2.9)$$

Now, we know that the commutator has to enter the picture somewhere. This we arrange through the following identity:

$$\begin{aligned}\hat{\Omega}\hat{\Lambda} &= \frac{\hat{\Omega}\hat{\Lambda} + \hat{\Lambda}\hat{\Omega}}{2} + \frac{\hat{\Omega}\hat{\Lambda} - \hat{\Lambda}\hat{\Omega}}{2} \\ &= \frac{1}{2}[\hat{\Omega}, \hat{\Lambda}]_+ + \frac{1}{2}[\hat{\Omega}, \hat{\Lambda}]\end{aligned} \quad (9.2.10)$$

where $[\hat{\Omega}, \hat{\Lambda}]_+$ is called the *anticommutator*. Feeding Eq. (9.2.10) into the inequality (9.2.9), we get

$$(\Delta\Omega)^2(\Delta\Lambda)^2 \geq |\langle \psi | \frac{1}{2}[\hat{\Omega}, \hat{\Lambda}]_+ + \frac{1}{2}[\hat{\Omega}, \hat{\Lambda}] | \psi \rangle|^2 \quad (9.2.11)$$

We next use the fact that

(1) since $[\hat{\Omega}, \hat{\Lambda}] = i\Gamma$, where Γ is Hermitian, the expectation value of the commutator is pure imaginary;

(2) since $[\hat{\Omega}, \hat{\Lambda}]_+$ is Hermitian, the expectation value of the anticommutator is real.

Recalling that $|a+ib|^2 = a^2 + b^2$, we get

$$\begin{aligned} (\Delta\Omega)^2(\Delta\Lambda)^2 &\geq \frac{1}{4}|\langle \psi | [\hat{\Omega}, \hat{\Lambda}]_+ | \psi \rangle + i\langle \psi | \Gamma | \psi \rangle|^2 \\ &\geq \frac{1}{4}\langle \psi | [\hat{\Omega}, \hat{\Lambda}]_+ | \psi \rangle^2 + \frac{1}{4}\langle \psi | \Gamma | \psi \rangle^2 \end{aligned} \quad (9.2.12)$$

This is the general uncertainty relation between any two Hermitian operators and is evidently state dependent. Consider now canonically conjugate operators, for which $\Gamma = \hbar$. In this case

$$(\Delta\Omega)^2(\Delta\Lambda)^2 \geq \frac{1}{4}\langle \psi | [\hat{\Omega}, \hat{\Lambda}]_+ | \psi \rangle^2 + \frac{\hbar^2}{4} \quad (9.2.13)$$

Since the first term is positive definite, we may assert that for *any* $|\psi\rangle$

$$(\Delta\Omega)^2(\Delta\Lambda)^2 \geq \hbar^2/4$$

or

$$\Delta\Omega \cdot \Delta\Lambda \geq \hbar/2 \quad (9.2.14)$$

which is the celebrated uncertainty relation. Let us note that the above inequality becomes an equality only if

$$(1) \quad \hat{\Omega}|\psi\rangle = c\hat{\Lambda}|\psi\rangle$$

and

$$(9.2.15)$$

$$(2) \quad \langle \psi | [\hat{\Omega}, \hat{\Lambda}]_+ | \psi \rangle = 0$$

9.3. The Minimum Uncertainty Packet

In this section we will find the wave function $\psi(x)$ which saturates the lower bound of the uncertainty relation for X and P . According to Eq. (9.2.15) such a state is characterized by

$$(P - \langle P \rangle)|\psi\rangle = c(X - \langle X \rangle)|\psi\rangle \quad (9.3.1)$$

and

$$\langle \psi | (P - \langle P \rangle)(X - \langle X \rangle) + (X - \langle X \rangle)(P - \langle P \rangle) | \psi \rangle = 0 \quad (9.3.2)$$

where $\langle P \rangle$ and $\langle X \rangle$ refer to the state $|\psi\rangle$, implicitly defined by these equations. In the X basis, Eq. (9.3.1) becomes

$$\left(-i\hbar \frac{d}{dx} - \langle P \rangle \right) \psi(x) = c(x - \langle X \rangle) \psi(x)$$

or

$$\frac{d\psi(x)}{\psi(x)} = \frac{i}{\hbar} [\langle P \rangle + c(x - \langle X \rangle)] dx \quad (9.3.3)$$

Now, whatever $\langle X \rangle$ may be, it is always possible to shift our origin (to $x = \langle X \rangle$) so that in the new frame of reference $\langle X \rangle = 0$. In this frame, Eq. (9.3.3) has the solution

$$\psi(x) = \psi(0) e^{i\langle P \rangle x/\hbar} e^{icx^2/2\hbar} \quad (9.3.4)$$

Let us next consider the constraint, Eq. (9.3.2), which in this frame reads

$$\langle \psi | (P - \langle P \rangle)X + X(P - \langle P \rangle) | \psi \rangle = 0$$

If we now exploit Eq. (9.3.1) and its adjoint, we find

$$\begin{aligned} \langle \psi | c^* X^2 + c X^2 | \psi \rangle &= 0 \\ (c + c^*) \langle \psi | X^2 | \psi \rangle &= 0 \end{aligned}$$

from which it follows that c is pure imaginary:

$$c = i|c| \quad (9.3.5)$$

Our solution, Eq. (9.3.4) now becomes

$$\psi(x) = \psi(0) e^{i\langle P \rangle x/\hbar} e^{-|c|x^2/2\hbar}$$

In terms of

$$\begin{aligned} \Delta^2 &= \hbar/|c| \\ \psi(x) &= \psi(0) e^{i\langle P \rangle x/\hbar} e^{-x^2/2\Delta^2} \end{aligned} \quad (9.3.6)$$

where Δ^2 , like $|c|$, is arbitrary. If the origin were not chosen to make $\langle X \rangle$ zero, we would have instead

$$\psi(x) = \psi(\langle X \rangle) e^{i\langle P \rangle (x - \langle X \rangle)/\hbar} e^{-(x - \langle X \rangle)^2/2\Delta^2} \quad (9.3.7)$$

Thus the *minimum uncertainty wave function* is a Gaussian of arbitrary width and center. This result, for the special case $\langle X \rangle = \langle P \rangle = 0$, was used in the quest for the state that minimized the expectation value of the oscillator Hamiltonian.

9.4. Applications of the Uncertainty Principle

I now illustrate the use of the uncertainty principle by estimating the size of the ground-state energy and the spread in the ground-state wave function. It should be clear from this example that the success we had with the oscillator was rather atypical.

We choose as our system the hydrogen atom. The Hamiltonian for this system, assuming the proton is a spectator whose only role is to provide a Coulomb potential for the electron, may be written entirely in terms of the electron's variable as

$$H = \frac{P_x^2 + P_y^2 + P_z^2}{2m} - \frac{e^2}{(X^2 + Y^2 + Z^2)^{1/2}} \quad (9.4.1)\ddagger$$

Let us begin by mimicking the analysis we employed for the oscillator. We evaluate $\langle H \rangle$ in a normalized state $|\psi\rangle$:

$$\begin{aligned} \langle H \rangle &= \frac{\langle P_x^2 + P_y^2 + P_z^2 \rangle}{2m} - e^2 \left\langle \frac{1}{(X^2 + Y^2 + Z^2)^{1/2}} \right\rangle \\ &= \frac{\langle P_x^2 \rangle + \langle P_y^2 \rangle + \langle P_z^2 \rangle}{2m} - e^2 \left\langle \frac{1}{(X^2 + Y^2 + Z^2)^{1/2}} \right\rangle \end{aligned} \quad (9.4.2)$$

Since

$$\langle P_x^2 \rangle = \langle (\Delta P_x)^2 + \langle P_x \rangle^2 \rangle \quad \text{etc.}$$

the first step in minimizing $\langle H \rangle$ is to work only with states for which $\langle P_i \rangle = 0$. For such states

$$\langle H \rangle = \frac{(\Delta P_x)^2 + (\Delta P_y)^2 + (\Delta P_z)^2}{2m} - e^2 \left\langle \frac{1}{(X^2 + Y^2 + Z^2)^{1/2}} \right\rangle \quad (9.4.3)$$

\ddagger The operator $(X^2 + Y^2 + Z^2)^{-1/2}$ is just $1/r$ in the coordinate basis. We will occasionally denote it by $1/r$ even while referring to it in the abstract, to simplify the notation.

We cannot exploit the uncertainty relations

$$\Delta P_x \Delta X \geq \hbar/2, \text{ etc.}$$

yet since $\langle H \rangle$ is not a function of ΔX and ΔP . The problem is that $\langle (X^2 + Y^2 + Z^2)^{-1/2} \rangle$ is not simply related to ΔX , ΔY , and ΔZ . Now the handwaving begins. We argue that (see Exercise 9.4.2),

$$\left\langle \frac{1}{(X^2 + Y^2 + Z^2)^{1/2}} \right\rangle \simeq \frac{1}{\langle (X^2 + Y^2 + Z^2)^{1/2} \rangle} \quad (9.4.4)$$

where the \simeq symbol means that the two sides of Eq. (9.4.4) are not strictly equal, but of the same order of magnitude. So we write

$$\langle H \rangle \simeq \frac{(\Delta P_x)^2 + (\Delta P_y)^2 + (\Delta P_z)^2}{2m} - \frac{e^2}{\langle (X^2 + Y^2 + Z^2)^{1/2} \rangle}$$

Once again, we argue that

$$\langle (X^2 + Y^2 + Z^2)^{1/2} \rangle \simeq (\langle X^2 \rangle + \langle Y^2 \rangle + \langle Z^2 \rangle)^{1/2}$$

and get‡

$$\langle H \rangle \simeq \frac{(\Delta P_x)^2 + (\Delta P_y)^2 + (\Delta P_z)^2}{2m} - \frac{e^2}{(\langle X^2 \rangle + \langle Y^2 \rangle + \langle Z^2 \rangle)^{1/2}}$$

From the relations

$$\langle X^2 \rangle = (\Delta X)^2 + \langle X \rangle^2 \text{ etc.}$$

it follows that we may confine ourselves to states for which $\langle X \rangle = \langle Y \rangle = \langle Z \rangle = 0$ in looking for the state with the lowest mean energy. For such states

$$\langle H \rangle \simeq \frac{\Delta P_x^2 + \Delta P_y^2 + \Delta P_z^2}{2m} - \frac{e^2}{[(\Delta X)^2 + (\Delta Y)^2 + (\Delta Z)^2]^{1/2}}$$

For a problem such as this, with spherical symmetry, it is intuitively clear that the configuration of least energy will have

$$(\Delta X)^2 = (\Delta Y)^2 = (\Delta Z)^2$$

‡ We are basically arguing that the mean of the functions (of X , Y , and Z) and the functions of the mean ($\langle X \rangle$, $\langle Y \rangle$, and $\langle Z \rangle$) are of the same order of magnitude. They are in fact equal if there are no fluctuations around the mean and approximately equal if the fluctuations are small (recall the discussion toward the end of Chapter 6).

and

$$(\Delta P_x)^2 = (\Delta P_y)^2 = (\Delta P_z)^2$$

so that

$$\langle H \rangle \approx \frac{3(\Delta P_x)^2}{2m} - \frac{e^2}{3^{1/2} \Delta X} \quad (9.4.5)$$

Now we use

$$\Delta P_x \Delta X \geq \hbar/2$$

to get

$$\langle H \rangle \gtrsim \frac{3\hbar^2}{8m(\Delta X)^2} - \frac{e^2}{3^{1/2} \Delta X}$$

We now differentiate the right-hand side with respect to ΔX to find its minimum:

$$\frac{-6\hbar^2}{8m(\Delta X)^3} + \frac{e^2}{3^{1/2}(\Delta X)^2} = 0$$

or

$$\Delta X = \frac{3(3^{1/2})\hbar^2}{4me^2} \approx 1.3 \frac{\hbar^2}{me^2} \quad (9.4.6)$$

Finally,

$$\langle H \rangle \gtrsim \frac{-2me^4}{9\hbar^2} \quad (9.4.7)$$

What prevents us from concluding (as we did in the case of the oscillator), that the ground-state energy is $-2me^4/9\hbar^2$ or that the ground-state wave function is a Gaussian [of width $3(3^{1/2})\hbar^2/4me^2$] is the fact that Eq. (9.4.7) is an approximate inequality. However, the exact ground-state energy

$$E_g = -me^4/2\hbar^2 \quad (9.4.8)$$

differs from our estimate, Eq. (9.4.7), only by a factor ≈ 2 . Likewise, the true ground-state wave function is not a Gaussian but an exponential $\psi(x, y, z) = c \exp[-(x^2 + y^2 + z^2)^{1/2}/a_0]$, where

$$a_0 = \hbar^2/me^2$$

is called the *Bohr radius*. However, the ΔX associated with this wave function is

$$\Delta X = \hbar^2 / me^2 \quad (9.4.9)$$

which also is within a factor of 2 of the estimated ΔX in (9.4.6).

In conclusion, the uncertainty principle gives us a lot of information about the ground state, but not always as much as in the case of the oscillator.

*Exercise 9.4.1.** Consider the oscillator in the state $|n=1\rangle$ and verify that

$$\left\langle \frac{1}{X^2} \right\rangle \approx \frac{1}{\langle X^2 \rangle} \approx \frac{m\omega}{\hbar}$$

Exercise 9.4.2. (1) By referring to the table of integrals in Appendix A.2, verify that

$$\psi = \frac{1}{(\pi a_0^3)^{1/2}} e^{-r/a_0}, \quad r = (x^2 + y^2 + z^2)^{1/2}$$

is a normalized wave function (of the ground state of hydrogen). Note that in three dimensions the normalization condition is

$$\begin{aligned} \langle \psi | \psi \rangle &= \int \psi^*(r, \theta, \phi) \psi(r, \theta, \phi) r^2 dr d(\cos \theta) d\phi \\ &= 4\pi \int \psi^*(r) \psi(r) r^2 dr = 1 \end{aligned}$$

for a function of just r .

(2) Calculate $(\Delta X)^2$ in this state [argue that $(\Delta X)^2 = \frac{1}{3} \langle r^2 \rangle$] and regain the result quoted in Eq. (9.4.9).

(3) Show that $\langle 1/r \rangle \approx 1/\langle r \rangle \approx me^2/\hbar^2$ in this state.

Exercise 9.4.3. Ignore the fact that the hydrogen atom is a three-dimensional system and pretend that

$$H = \frac{P^2}{2m} - \frac{e^2}{(R^2)^{1/2}} \quad (P^2 = P_x^2 + P_y^2 + P_z^2, R^2 = X^2 + Y^2 + Z^2)$$

corresponds to a one-dimensional problem. Assuming

$$\Delta P \cdot \Delta R \geq \hbar/2$$

estimate the ground-state energy.

*Exercise 9.4.4.** Compute $\Delta T \cdot \Delta X$, where $T = P^2/2m$. Why is this relation not so famous?

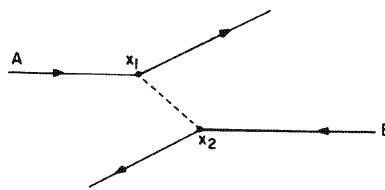


Figure 9.1. At the point x_1 , skater A throws the snowball towards skater B , who catches it at the point x_2 .

9.5. The Energy–Time Uncertainty Relation

There exists an uncertainty relation

$$\Delta E \cdot \Delta t \geq \hbar/2 \quad (9.5.1)$$

which does not follow from Eq. (9.2.12), since time t is not a dynamical variable but a parameter. The content of this equation is quite different from the others involving just dynamical variables. The rough meaning of this inequality is that the energy of a system that has been in existence only for a finite time Δt has a spread (or uncertainty) of at least ΔE , where ΔE and Δt are related by (9.5.1). To see how this comes about, recall that eigenstates of energy have a time dependence $e^{-iEt/\hbar}$, i.e., a definite energy is associated with a definite frequency, $\omega = E/\hbar$. Now, only a wave train that is infinitely long in time (that is to say, a system that has been in existence for infinite time) has a well-defined frequency. Thus a system that has been in existence only for a finite time, even if its time dependence goes as $e^{-iEt/\hbar}$ during this period, is not associated with a pure frequency $\omega = E/\hbar$ or definite energy E .

Consider the following example. At time $t=0$, we turn on light of frequency ω on an ensemble of hydrogen atoms all in their ground state. Since the light is supposed to consist of photons of energy $\hbar\omega$, we expect transitions to take place only to a level (if it exists) $\hbar\omega$ above the ground state. It will however be seen that initially the atoms make transitions to several levels not obeying this constraint. However, as t increases, the deviation ΔE from the expected final-state energy will decrease according to $\Delta E \simeq \hbar/t$. Only as $t \rightarrow \infty$ do we have a rigid law of conservation of energy in the classical sense. We interpret this result by saying that the light source is not associated with a definite frequency (i.e., does not emit photons of definite energy) if it has been in operation only for a finite time, even if the dial is set at a definite frequency ω during this time. [The output of the source is not just $e^{-i\omega t}$ but rather $\theta(t) e^{-i\omega t}$, whose transform is not a delta function peaked at ω .] Similarly when the excited atoms get deexcited and drop to the ground state, they do not emit photons of a definite energy $E = E_e - E_g$ (the subscripts e and g stand for “excited” and “ground”) but rather with a spread $\Delta E \simeq \hbar/\Delta t$, Δt being the duration for which they were in the excited state. [The time dependence of the atomic wave function is not $e^{-iE_e t/\hbar}$ but rather $\theta(t)\theta(T-t) e^{-iE_e t/\hbar}$ assuming it abruptly got excited to this state at $t=0$ and abruptly got deexcited at $t=T$.] We shall return to this point when we discuss the interaction of atoms with radiation in a later chapter.

Another way to describe this uncertainty relation is to say that violations in the classical energy conservation law by ΔE are possible over times $\Delta t \sim \hbar/\Delta E$. The following example should clarify the meaning of this statement.

Example 9.5.1. (Range of the Nuclear Force.) Imagine two ice skaters each equipped with several snowballs, and skating toward each other on trajectories that are parallel but separated by some perpendicular distance (Fig. 9.1). When skater A reaches some point x_1

let him throw a snowball toward B . He (A) will then recoil away from B and start moving along a new straight line. Let B now catch the snowball. He too will recoil as a result, as shown in the figure. If this whole process were seen by someone who could not see the snow balls, he would conclude that there is a repulsive force between A and B . If A (or B) can throw the ball at most 10 ft, the observer would conclude that the *range of the force* is 10 ft, meaning A and B will not affect each other if the perpendicular distance between them exceeds 10 ft.

This is roughly how elementary particles interact with each other: if they throw photons at each other the force is called the electromagnetic force and the ability to throw and catch photons is called “electric charge.” If the projectiles are pions the force is called the nuclear force. We would like to estimate the range of the nuclear force using the uncertainty principle. Now, unlike the two skaters endowed with snowballs, the protons and neutrons (i.e., nucleons) in the nucleus do not have a ready supply of pions, which have a mass μ and energy μc^2 . A nucleon can, however, produce a pion from nowhere (violating the classical law of energy conservation by $\simeq \mu c^2$) provided it is caught by the other nucleon within a time Δt such that $\Delta t \simeq \hbar / \Delta E = \hbar / \mu c^2$. Even if the pion travels toward the receiver at the speed of light, it can only cover a distance $r = c \Delta t = \hbar / \mu c$, which is called the *Compton wavelength* of the pion and is a measure of the range of nuclear force. The value of r is approximately 1 Fermi = 10^{-13} cm.

The picture of nuclear force given here is rather simpleminded and should be taken with a grain of salt. For example, neither is the pion the only particle that can be “exchanged” between nucleons nor is the number of exchanges limited to one per encounter. (The pion is, however, the lightest object that can be exchanged and hence responsible for the nuclear force of the longest range.) Also our analogy with snowballs does not explain any attractive interaction between particles. \square

Systems with N Degrees of Freedom

10.1. N Particles in One Dimension

So far, we have restricted our attention (apart from minor digressions) to a system with one degree of freedom, namely, a single particle in one dimension. We now consider the quantum mechanics of systems with N degrees of freedom. The increase in degrees of freedom may be due to an increase in the number of particles, number of spatial dimensions, or both. In this section we consider N particles in one dimension, and start with the case $N=2$.

The Two-Particle Hilbert Space

Consider two particles described classically by (x_1, p_1) and (x_2, p_2) . The rule for quantizing this system [Postulate II, Eq. (7.4.39)] is to promote these variables to quantum operators (X_1, P_1) and (X_2, P_2) obeying the canonical commutation relations:

$$[X_i, P_j] = i\hbar\{x_i, p_j\} = i\hbar\delta_{ij} \quad (i=1, 2) \quad (10.1.1a)$$

$$[X_i, X_j] = i\hbar\{x_i, x_j\} = 0 \quad (10.1.1b)$$

$$[P_i, P_j] = i\hbar\{p_i, p_j\} = 0 \quad (10.1.1c)$$

It might be occasionally possible (as it was in the case of the oscillator) to extract all the physics given just the canonical commutators. In practice one works in a basis, usually the coordinate basis. This basis consists of the kets $|x_1 x_2\rangle$ which are

simultaneous eigenkets of the commuting operators X_1 and X_2 :

$$\begin{aligned} X_1|x_1x_2\rangle &= x_1|x_1x_2\rangle \\ X_2|x_1x_2\rangle &= x_2|x_1x_2\rangle \end{aligned} \quad (10.1.2)$$

and are normalized as‡

$$\langle x'_1x'_2|x_1x_2\rangle = \delta(x'_1 - x_1)\delta(x'_2 - x_2) \quad (10.1.3)$$

In this basis

$$\begin{aligned} |\psi\rangle &\rightarrow \langle x_1x_2|\psi\rangle = \psi(x_1, x_2) \\ X_i &\rightarrow x_i \\ P_i &\rightarrow -i\hbar \frac{\partial}{\partial x_i} \end{aligned} \quad (10.1.4)$$

We may interpret

$$P(x_1, x_2) = |\langle x_1x_2|\psi\rangle|^2 \quad (10.1.5)$$

as the absolute probability density for catching particle 1 near x_1 and particle 2 near x_2 , provided we normalize $|\psi\rangle$ to unity

$$1 = \langle\psi|\psi\rangle = \int |\langle x_1x_2|\psi\rangle|^2 dx_1 dx_2 = \int P(x_1, x_2) dx_1 dx_2 \quad (10.1.6)$$

There are other bases possible besides $|x_1x_2\rangle$. There is, for example, the momentum basis, consisting of the simultaneous eigenkets $|p_1p_2\rangle$ of P_1 and P_2 . More generally, we can use the simultaneous eigenkets $|\omega_1\omega_2\rangle$ of two commuting operators§ $\Omega_1(X_1, P_1)$ and $\Omega_2(X_2, P_2)$ to define the Ω basis. We denote by $\mathbb{V}_{1\otimes 2}$ the two-particle Hilbert space spanned by any of these bases.

$\mathbb{V}_{1\otimes 2}$ As a Direct Product Space

There is another way to arrive at the space $\mathbb{V}_{1\otimes 2}$, and that is to build it out of two one-particle spaces. Consider a system of two particles described classically by (x_1, p_1) and (x_2, p_2) . If we want the quantum theory of just particle 1, we define operators X_1 and P_1 obeying

$$[X_1, P_1] = i\hbar I \quad (10.1.7)$$

The eigenvectors $|x_1\rangle$ of X_1 form a complete (coordinate) basis for the Hilbert space

‡ Note that we denote the bra corresponding to $|x'_1x'_2\rangle$ as $\langle x'_1x'_2|$.

§ Note that any function of X_1 and P_1 commutes with any function of X_2 and P_2 .

\mathbb{V}_1 of particle 1. Other bases, such as $|p_1\rangle$ of P_1 or in general, $|\omega_1\rangle$ of $\Omega_1(X_1, P_1)$ are also possible. Since the operators X_1 , P_1 , Ω_1 , etc., act on \mathbb{V}_1 , let us append a superscript (1) to all of them. Thus Eq. (10.1.7) reads

$$[X_1^{(1)}, P_1^{(1)}] = i\hbar I^{(1)} \quad (10.1.8a)$$

where $I^{(1)}$ is the identity operator on \mathbb{V}_1 . A similar picture holds for particle 2, and in particular,

$$[X_2^{(2)}, P_2^{(2)}] = i\hbar I^{(2)} \quad (10.1.8b)$$

Let us now turn our attention to the two-particle system. What will be the coordinate basis for this system? Previously we assigned to every possible outcome x_1 of a position measurement a vector $|x_1\rangle$ in \mathbb{V}_1 and likewise for particle 2. Now a position measurement will yield a pair of numbers (x_1, x_2) . Since after the measurement particle 1 will be in state $|x_1\rangle$ and particle 2 in $|x_2\rangle$, let us denote the corresponding ket by $|x_1\rangle \otimes |x_2\rangle$:

$$|x_1\rangle \otimes |x_2\rangle \leftrightarrow \begin{cases} \text{particle 1 at } x_1 \\ \text{particle 2 at } x_2 \end{cases} \quad (10.1.9)$$

Note that $|x_1\rangle \otimes |x_2\rangle$ is a new object, quite unlike the inner product $\langle \psi_1 | \psi_2 \rangle$ or the outer product $|\psi_1\rangle \langle \psi_2|$ both of which involve *two vectors from the same space*. The product $|x_1\rangle \otimes |x_2\rangle$, called the *direct product*, is the product of *vectors from two different spaces*. The direct product is a linear operation:

$$(\alpha|x_1\rangle + \alpha'|x'_1\rangle) \otimes (\beta|x_2\rangle) = \alpha\beta|x_1\rangle \otimes |x_2\rangle + \alpha'\beta|x'_1\rangle \otimes |x_2\rangle \quad (10.1.10)$$

The set of all vectors of the form $|x_1\rangle \otimes |x_2\rangle$ forms the basis for a space which we call $\mathbb{V}_1 \otimes \mathbb{V}_2$, and refer to as the *direct product of the spaces \mathbb{V}_1 and \mathbb{V}_2* . The dimensionality (number of possible basis vectors) of $\mathbb{V}_1 \otimes \mathbb{V}_2$ is the product of the dimensionality of \mathbb{V}_1 and the dimensionality of \mathbb{V}_2 . Although all the dimensionalities are infinite here, the statement makes heuristic sense: to each basis vector $|x_1\rangle$ of \mathbb{V}_1 and $|x_2\rangle$ of \mathbb{V}_2 , there is one and only one basis vector $|x_1\rangle \otimes |x_2\rangle$ of $\mathbb{V}_1 \otimes \mathbb{V}_2$. This should be compared to the direct sum (Section 1.4):

$$\mathbb{V}_{1+2} = \mathbb{V}_1 \oplus \mathbb{V}_2$$

in which case the dimensionalities of \mathbb{V}_1 and \mathbb{V}_2 add (assuming the vectors of \mathbb{V}_1 are linearly independent of those of \mathbb{V}_2).

The coordinate basis, $|x_1\rangle \otimes |x_2\rangle$, is just one possibility; we can use the momentum basis $|p_1\rangle \otimes |p_2\rangle$, or, more generally, $|\omega_1\rangle \otimes |\omega_2\rangle$. *Although these vectors span $\mathbb{V}_1 \otimes \mathbb{V}_2$, not every element of $\mathbb{V}_1 \otimes \mathbb{V}_2$ is a direct product.* For instance

$$|\psi\rangle = |x'_1\rangle \otimes |x'_2\rangle + |x''_1\rangle \otimes |x''_2\rangle$$

cannot be written as

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$$

where $|\psi_1\rangle$ and $|\psi_2\rangle$ are elements of \mathbb{V}_1 and \mathbb{V}_2 , respectively.

The inner product of $|x_1\rangle \otimes |x_2\rangle$ and $|x'_1\rangle \otimes |x'_2\rangle$ is

$$\begin{aligned} (\langle x'_1| \otimes \langle x'_2|)(|x_1\rangle \otimes |x_2\rangle) &= \langle x'_1|x_1\rangle \langle x'_2|x_2\rangle \\ &= \delta(x'_1 - x_1)\delta(x'_2 - x_2) \end{aligned} \quad (10.1.11)$$

Since any vector in $\mathbb{V}_1 \otimes \mathbb{V}_2$ can be expressed in terms of the $|x_1\rangle \otimes |x_2\rangle$ basis, this defines the inner product between any two vectors in $\mathbb{V}_1 \otimes \mathbb{V}_2$.

It is intuitively clear that when two particles are amalgamated to form a single system, the position and momentum operators of each particle, $X_1^{(1)}$, $P_1^{(1)}$ and $X_2^{(2)}$, $P_2^{(2)}$, which acted on \mathbb{V}_1 and \mathbb{V}_2 , respectively, must have counterparts in $\mathbb{V}_1 \otimes \mathbb{V}_2$ and have the same interpretation. Let us denote by $X_1^{(1) \otimes (2)}$ the counterpart of $X_1^{(1)}$, and refer to it also as the “ X operator of particle 1.” Let us define its action on $\mathbb{V}_1 \otimes \mathbb{V}_2$. Since the vectors $|x_1\rangle \otimes |x_2\rangle$ span the space, it suffices to define its action on these. Now the ket $|x_1\rangle \otimes |x_2\rangle$ denotes a state in which particle 1 is at x_1 . Thus it must be an eigenket of $X_1^{(1) \otimes (2)}$ with eigenvalue x_1 :

$$X_1^{(1) \otimes (2)}|x_1\rangle \otimes |x_2\rangle = x_1|x_1\rangle \otimes |x_2\rangle \quad (10.1.12)$$

Note that $X_1^{(1) \otimes (2)}$ does not really care about the second ket $|x_2\rangle$, i.e., it acts trivially (as the identity) on $|x_2\rangle$ and acts on $|x_1\rangle$ just as $X_1^{(1)}$ did. In other words

$$X_1^{(1) \otimes (2)}|x_1\rangle \otimes |x_2\rangle = |X_1^{(1)}x_1\rangle \otimes |I^{(2)}x_2\rangle \quad (10.1.13)$$

Let us define a *direct product of two operators*, $\Gamma_1^{(1)}$ and $\Lambda_2^{(2)}$ (denoted by $\Gamma_1^{(1)} \otimes \Lambda_2^{(2)}$), whose action on a direct product ket $|\omega_1\rangle \otimes |\omega_2\rangle$ is

$$(\Gamma_1^{(1)} \otimes \Lambda_2^{(2)})|\omega_1\rangle \otimes |\omega_2\rangle = |\Gamma_1^{(1)}\omega_1\rangle \otimes |\Lambda_2^{(2)}\omega_2\rangle \quad (10.1.14)$$

In this notation, we may write $X_1^{(1) \otimes (2)}$, in view of Eq. (10.1.13), as

$$X_1^{(1) \otimes (2)} = X_1^{(1)} \otimes I^{(2)} \quad (10.1.15)$$

We can similarly promote $P_2^{(2)}$, say, from \mathbb{V}_2 to $\mathbb{V}_1 \otimes \mathbb{V}_2$ by defining the momentum operator for particle 2, $P_2^{(1) \otimes (2)}$, as

$$P_2^{(1) \otimes (2)} = I^{(1)} \otimes P_2^{(2)} \quad (10.1.16)$$

The following properties of direct products of operators may be verified (say by acting on the basis vectors $|x_1\rangle \otimes |x_2\rangle$):

*Exercise 10.1.1.** Show the following:

$$(1) \quad [\Omega_1^{(1)} \otimes I^{(2)}, I^{(1)} \otimes \Lambda_2^{(2)}] = 0 \text{ for any } \Omega_1^{(1)} \text{ and } \Lambda_2^{(2)} \quad (10.1.17a)$$

(operators of particle 1 commute with those of particle 2).

$$(2) \quad (\Omega_1^{(1)} \otimes \Gamma_2^{(2)}) (\theta_1^{(1)} \otimes \Lambda_2^{(2)}) = (\Omega \theta)_1^{(1)} \otimes (\Gamma \Lambda)_2^{(2)} \quad (10.1.17b)$$

(3) If

$$[\Omega_1^{(1)}, \Lambda_1^{(1)}] = \Gamma_1^{(1)}$$

then

$$[\Omega_1^{(1) \otimes (2)}, \Lambda_1^{(1) \otimes (2)}] = \Gamma_1^{(1)} \otimes I^{(2)} \quad (10.1.17c)$$

and similarly with $1 \rightarrow 2$.

$$(4) \quad (\Omega_1^{(1) \otimes (2)} + \Omega_2^{(1) \otimes (2)})^2 = (\Omega_1^{(2)})^{(1)} \otimes I^{(2)} + I^{(1)} \otimes (\Omega_2^{(2)})^{(2)} + 2\Omega_1^{(1)} \otimes \Omega_2^{(2)} \quad (10.1.17d)$$

The notion of direct products of vectors and operators is no doubt a difficult one, with no simple analogs in elementary vector analysis. The following exercise should give you some valuable experience. It is recommended that you reread the preceding discussion after working on the exercise.

*Exercise 10.1.2.** Imagine a fictitious world in which the single-particle Hilbert space is two-dimensional. Let us denote the basis vectors by $|+\rangle$ and $|-\rangle$. Let

$$\sigma_1^{(1)} = \begin{pmatrix} + & - \\ a & b \\ - & c \end{pmatrix} \quad \text{and} \quad \sigma_2^{(2)} = \begin{pmatrix} + & - \\ e & f \\ - & g \end{pmatrix}$$

be operators in \mathbb{V}_1 and \mathbb{V}_2 , respectively (the \pm signs label the basis vectors. Thus $b = \langle + | \sigma_1^{(1)} | - \rangle$ etc.) The space $\mathbb{V}_1 \otimes \mathbb{V}_2$ is spanned by four vectors $|+\rangle \otimes |+\rangle$, $|+\rangle \otimes |-\rangle$, $|-\rangle \otimes |+\rangle$, $|-\rangle \otimes |-\rangle$. Show (using the method of images or otherwise) that

$$(1) \quad \sigma_1^{(1) \otimes (2)} = \sigma_1^{(1)} \otimes I^{(2)} = \begin{pmatrix} ++ & +- & -+ & -- \\ + & \begin{bmatrix} a & 0 & b & 0 \end{bmatrix} \\ - & \begin{bmatrix} 0 & a & 0 & b \\ c & 0 & d & 0 \end{bmatrix} \\ - & \begin{bmatrix} 0 & c & 0 & d \end{bmatrix} \end{pmatrix}$$

(Recall that $\langle \alpha | \otimes \langle \beta |$ is the bra corresponding to $|\alpha\rangle \otimes |\beta\rangle$.)

$$(2) \quad \sigma_2^{(1)\otimes(2)} = \begin{bmatrix} e & f & 0 & 0 \\ g & h & 0 & 0 \\ 0 & 0 & e & f \\ 0 & 0 & g & h \end{bmatrix}$$

$$(3) \quad (\sigma_1\sigma_2)^{(1)\otimes(2)} = \sigma_1^{(1)} \otimes \sigma_2^{(2)} = \begin{bmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{bmatrix}$$

Do part (3) in two ways, by taking the matrix product of $\sigma_1^{(1)\otimes(2)}$ and $\sigma_2^{(1)\otimes(2)}$ and by directly computing the matrix elements of $\sigma_1^{(1)} \otimes \sigma_2^{(2)}$.

From Eqs. (10.1.17a) and (10.1.17c) it follows that the commutation relations between the position and momentum operators on $\mathbb{V}_1 \otimes \mathbb{V}_2$ are

$$\begin{aligned} [X_i^{(1)\otimes(2)}, P_j^{(1)\otimes(2)}] &= i\hbar\delta_{ij}I^{(1)} \otimes I^{(2)} = i\hbar\delta_{ij}I^{(1)\otimes(2)} \\ [X_i^{(1)\otimes(2)}, X_j^{(1)\otimes(2)}] &= [P_i^{(1)\otimes(2)}, P_j^{(1)\otimes(2)}] = 0 \quad i, j = 1, 2 \end{aligned} \quad (10.1.18)$$

Now we are ready to assert something that may have been apparent all along: the space $\mathbb{V}_1 \otimes \mathbb{V}_2$ is just $\mathbb{V}_{1\otimes 2}$, $|x_1\rangle \otimes |x_2\rangle$ is just $|x_1x_2\rangle$, and $X_1^{(1)\otimes(2)}$ is just X_1 , etc. Notice first that both spaces have the same dimensionality: the vectors $|x_1x_2\rangle$ and $|x_1\rangle \otimes |x_2\rangle$ are both in one-to-one correspondence with points in the $x_1 - x_2$ plane. Notice next that the two sets of operators X_1, \dots, P_2 and $X_2^{(1)\otimes(2)}, \dots, P_2^{(1)\otimes(2)}$ have the same connotation and commutation rules [Eqs. (10.1.1) and (10.1.18)]. Since X and P are *defined* by their commutators we can make the identification

$$\begin{aligned} X_i^{(1)\otimes(2)} &= X_i \\ P_i^{(1)\otimes(2)} &= P_i \end{aligned} \quad (10.1.19a)$$

We can also identify the simultaneous eigenkets of the position operators (since they are nondegenerate):

$$|x_1\rangle \otimes |x_2\rangle = |x_1x_2\rangle \quad (10.1.19b)$$

In the future, we shall use the more compact symbols occurring on the right-hand side of Eqs. (10.1.19). We will, however, return to the concept of direct products of vectors and operators on and off and occasionally use the symbols on the left-hand side. Although the succinct notation suppresses the label $(1\otimes 2)$ of the space on

which the operators act, it should be clear from the context. Consider, for example, the CM kinetic energy operator of the two-particle system:

$$T_{\text{CM}} = \frac{P_{\text{CM}}^2}{2(m_1 + m_2)} = \frac{P_{\text{CM}}^2}{2M} = \frac{(P_1 + P_2)^2}{2M} = \frac{P_1^2 + P_2^2 + 2P_1P_2}{2M}$$

which really means

$$\begin{aligned} 2MT_{\text{CM}}^{(1)\otimes(2)} &= (P_1^2)^{(1)\otimes(2)} + (P_2^2)^{(1)\otimes(2)} + 2P_1^{(1)\otimes(2)} \cdot P_2^{(1)\otimes(2)} \\ &= (P_1^{(1)} \otimes I^{(2)})^2 + (I^{(1)} \otimes P_2^{(2)})^2 + 2P_1^{(1)} \otimes P_2^{(2)} \end{aligned}$$

The Direct Product Revisited

Since the notion of a direct product space is so important, we revisit the formation of $\mathbb{V}_{1\otimes 2}$ as a direct product of \mathbb{V}_1 and \mathbb{V}_2 , but this time in the coordinate basis instead of in the abstract. Let $\Omega_i^{(1)}$ be an operator on \mathbb{V}_1 whose nondegenerate eigenfunctions $\psi_{\omega_1}(x_1) \equiv \omega_1(x_1)$ form a complete basis. Similarly let $\omega_2(x_2)$ form a basis for \mathbb{V}_2 . Consider now a function $\psi(x_1, x_2)$, which represents the abstract ket $|\psi\rangle$ from $\mathbb{V}_{1\otimes 2}$. If we keep x_1 fixed at some value, say \bar{x}_1 , then ψ becomes a function of x_2 alone and may be expanded as

$$\psi(\bar{x}_1, x_2) = \sum_{\omega_2} C_{\omega_2}(\bar{x}_1) \omega_2(x_2) \quad (10.1.20)$$

Notice that the coefficients of the expansion depend on the value of \bar{x}_1 . We now expand the function $C_{\omega_2}(\bar{x}_1)$ in the basis $\omega_1(\bar{x}_1)$:

$$C_{\omega_2}(\bar{x}_1) = \sum_{\omega_1} C_{\omega_1, \omega_2} \omega_1(\bar{x}_1) \quad (10.1.21)$$

Feeding this back to the first expansion and dropping the bar on \bar{x}_1 we get

$$\psi(x_1, x_2) = \sum_{\omega_1} \sum_{\omega_2} C_{\omega_1, \omega_2} \omega_1(x_1) \omega_2(x_2) \quad (10.1.22a)$$

What does this expansion of an arbitrary $\psi(x_1, x_2)$ in terms of $\omega_1(x_1) \times \omega_2(x_2)$ imply? Equation (10.1.22a) is the coordinate space version of the abstract result

$$|\psi\rangle = \sum_{\omega_1} \sum_{\omega_2} C_{\omega_1, \omega_2} |\omega_1\rangle \otimes |\omega_2\rangle \quad (10.1.22b)$$

which means $\mathbb{V}_{1\otimes 2} = \mathbb{V}_1 \otimes \mathbb{V}_2$, for $|\psi\rangle$ belongs to $\mathbb{V}_{1\otimes 2}$ and $|\omega_1\rangle \otimes |\omega_2\rangle$ spans $\mathbb{V}_1 \otimes \mathbb{V}_2$. If we choose $\Omega = X$, we get the familiar basis $|x_1\rangle \otimes |x_2\rangle$. By dotting both sides of Eq. (10.1.22b) with these basis vectors we regain Eq. (10.1.22a). (In the coordinate basis, the direct product of the kets $|\omega_1\rangle$ and $|\omega_2\rangle$ becomes just the ordinary product of the corresponding wave functions.)

Consider next the operators. The momentum operator on \mathbb{V}_1 , which used to be $-i\hbar d/dx_1$ now becomes $-i\hbar \partial/\partial x_1$, where the partial derivative symbol tells us it

operates on x_1 as before and leaves x_2 alone. This is the coordinate space version of $P^{(1)} \otimes P^{(2)} = P_1^{(1)} \otimes I^{(2)}$. You are encouraged to pursue this analysis further.

Evolution of the Two-Particle State Vector

The state vector of the system is an element of $\mathbb{V}_{1 \otimes 2}$. It evolves in time according to the equation

$$i\hbar |\dot{\psi}\rangle = \left[\frac{P_1^2}{2m_1} + \frac{P_2^2}{2m_2} + V(X_1, X_2) \right] |\psi\rangle = H|\psi\rangle \quad (10.1.23)$$

There are two classes of problems.

Class A: H is separable, i.e.,

$$H = \frac{P_1^2}{2m_1} + V_1(X_1) + \frac{P_2^2}{2m_2} + V_2(X_2) = H_1 + H_2 \quad (10.1.24)$$

Class B: H is not separable, i.e.,

$$V(X_1, X_2) \neq V_1(X_1) + V_2(X_2)$$

and

$$H \neq H_1 + H_2 \quad (10.1.25)$$

Class A corresponds to two particles interacting with external potentials V_1 and V_2 but not with each other, while in class B there is no such restriction. We now examine these two classes.

Class A: Separable Hamiltonians. Classically, the decomposition

$$\mathcal{H} = \mathcal{H}_1(x_1, p_1) + \mathcal{H}_2(x_2, p_2)$$

means that the two particles evolve independently of each other. In particular, their energies are *separately* conserved and the total energy E is $E_1 + E_2$. Let us see these results reappear in quantum theory. For a stationary state,

$$|\psi(t)\rangle = |E\rangle e^{-iEt/\hbar} \quad (10.1.26)$$

Eq. (10.1.23) becomes

$$[H_1(X_1, P_1) + H_2(X_2, P_2)]|E\rangle = E|E\rangle \quad (10.1.27)$$

Since $[H_1, H_2] = 0$ [Eq. (10.1.17a)] we can find their simultaneous eigenstates, which are none other than $|E_1\rangle \otimes |E_2\rangle = |E_1 E_2\rangle$, where $|E_1\rangle$ and $|E_2\rangle$ are solutions to

$$H_1^{(1)}|E_1\rangle = E_1|E_1\rangle \quad (10.1.28a)$$

and

$$H_2^{(2)}|E_2\rangle = E_2|E_2\rangle \quad (10.1.28b)$$

It should be clear that the state $|E_1\rangle \otimes |E_2\rangle$ corresponds to particle 1 being in the energy eigenstate $|E_1\rangle$ and particle 2 being in the energy eigenstate $|E_2\rangle$. Clearly

$$H|E\rangle = (H_1 + H_2)|E_1\rangle \otimes |E_2\rangle = (E_1 + E_2)|E_1\rangle \otimes |E_2\rangle = (E_1 + E_2)|E\rangle$$

so that

$$E = E_1 + E_2 \quad (10.1.28c)$$

(The basis $|E_1\rangle \otimes |E_2\rangle$ is what we would get if in forming basis vectors of the direct product $\mathbb{V}_1 \otimes \mathbb{V}_2$, we took the energy eigenvalues from each space, instead of, say, the position eigenvectors.) Finally, feeding $|E\rangle = |E_1\rangle \otimes |E_2\rangle$, $E = E_1 + E_2$ into Eq. (10.1.26) we get

$$|\psi(t)\rangle = |E_1\rangle e^{-iE_1 t/\hbar} \otimes |E_2\rangle e^{-iE_2 t/\hbar} \quad (10.1.29)$$

It is worth rederiving Eqs. (10.1.28) and (10.1.29) in the coordinate basis to illustrate a useful technique that you will find in other textbooks. By projecting the eigenvalue Eq. (10.1.27) on this basis, and making the usual operator substitutions, Eq. (10.1.4), we obtain

$$\left[\frac{-\hbar^2}{2m_1} \frac{\partial^2}{\partial x_1^2} + V_1(x_1) - \frac{\hbar^2}{2m_2} \frac{\partial^2}{\partial x_2^2} + V_2(x_2) \right] \psi_E(x_1, x_2) = E \psi_E(x_1, x_2)$$

where

$$\psi_E(x_1, x_2) = \langle x_1 x_2 | E \rangle \quad (10.1.30)$$

We solve the equation by the method of *separation of variables*. We assume

$$\psi_E(x_1, x_2) = \psi_{E_1}(x_1) \psi_{E_2}(x_2) \quad (10.1.31)$$

The subscripts E_1 and E_2 have no specific interpretation yet and merely serve as labels. Feeding this *ansatz* into Eq. (10.1.30) and *then* dividing both sides by

$$\frac{1}{\psi_{E_1}(x_1)} \left[\frac{-\hbar^2}{2m_1} \frac{\partial^2}{\partial x_1^2} + V_1(x_1) \right] \psi_{E_1}(x_1) + \frac{1}{\psi_{E_2}(x_2)} \left[\frac{-\hbar^2}{2m_2} \frac{\partial^2}{\partial x_2^2} + V_2(x_2) \right] \psi_{E_2}(x_2) = E \quad (10.1.32)$$

This equation says that a function of x_1 alone, plus one of x_2 alone, equals a constant E . Since x_1 and x_2 , and hence the two functions, may be varied independently, it follows that each function separately equals a constant. We will call these constants E_1 and E_2 . Thus Eq. (10.1.32) breaks down into three equations:

$$\begin{aligned} \frac{1}{\psi_{E_1}(x_1)} \left[\frac{-\hbar^2}{2m_1} \frac{\partial^2}{\partial x_1^2} + V_1(x_1) \right] \psi_{E_1}(x_1) &= E_1 \\ \frac{1}{\psi_{E_2}(x_2)} \left[\frac{-\hbar^2}{2m_2} \frac{\partial^2}{\partial x_2^2} + V_2(x_2) \right] \psi_{E_2}(x_2) &= E_2 \\ E_1 + E_2 &= E \end{aligned} \quad (10.1.33)$$

Consequently

$$\begin{aligned} \psi_E(x_1, x_2, t) &= \psi_E(x_1, x_2) e^{-iEt/\hbar} \\ &= \psi_{E_1}(x_1) e^{-iE_1t/\hbar} \psi_{E_2}(x_2) e^{-iE_2t/\hbar} \end{aligned} \quad (10.1.34)$$

where ψ_{E_1} and ψ_{E_2} are eigenfunctions of the one-particle Schrödinger equation with eigenvalues E_1 and E_2 , respectively. We recognize Eqs. (10.1.33) and (10.1.34) to be the projections of Eqs. (10.1.28) and (10.1.29) on $|x_1x_2\rangle = |x_1\rangle \otimes |x_2\rangle$.

Case B: Two Interacting Particles. Consider next the more general problem of two interacting particles with

$$\mathcal{H} = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + V(x_1, x_2) \quad (10.1.35)$$

where

$$V(x_1, x_2) \neq V_1(x_1) + V_2(x_2)$$

Generally this cannot be reduced to two independent single-particle problems. If, however,

$$V(x_1, x_2) = V(x_1 - x_2) \quad (10.1.36)$$

which describes two particles responding to each other but nothing external, one can always, by employing the CM coordinate

$$x_{\text{CM}} = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2} \quad (10.1.37a)$$

and the relative coordinate

$$x = x_1 - x_2 \quad (10.1.37b)$$

reduce the problem to that of two independent fictitious particles: one, the CM, which is free, has mass $M = m_1 + m_2$ and momentum

$$p_{\text{CM}} = M \dot{x}_{\text{CM}} = m_1 \dot{x}_1 + m_2 \dot{x}_2$$

and another, with the reduced mass $\mu = m_1 m_2 / (m_1 + m_2)$, momentum $p = \mu \dot{x}$, moving under the influence of $V(x)$:

$$\begin{aligned} \mathcal{H}(x_1, p_1; x_2, p_2) &\rightarrow \mathcal{H}(x_{\text{CM}}, p_{\text{CM}}; x, p) \\ &= \mathcal{H}_{\text{CM}} + \mathcal{H}_{\text{relative}} = \frac{p_{\text{CM}}^2}{2M} + \frac{p^2}{2\mu} + V(x) \end{aligned} \quad (10.1.38)$$

which is just the result from Exercise 2.5.4 modified to one dimension. Since the new variables are also canonical (Exercise 2.7.6) and Cartesian, the quantization condition is just

$$[X_{\text{CM}}, P_{\text{CM}}] = i\hbar \quad (10.1.39a)$$

$$[X, P] = i\hbar \quad (10.1.39b)$$

and all other commutators zero. In the quantum theory,

$$H = \frac{P_{\text{CM}}^2}{2M} + \frac{P^2}{2\mu} + V(X) \quad (10.1.40)$$

and the eigenfunctions of H factorize:

$$\begin{aligned} \psi_E(x_{\text{CM}}, x) &= \frac{e^{ip_{\text{CM}} \cdot x_{\text{CM}}/\hbar}}{(2\pi\hbar)^{1/2}} \cdot \psi_{E_{\text{rel}}}(x) \\ E &= \frac{p_{\text{CM}}^2}{2M} + E_{\text{rel}} \end{aligned} \quad (10.1.41)$$

The real dynamics is contained in $\psi_{E_{\text{rel}}}(x)$ which is the energy eigenfunction for a particle of mass μ in a potential $V(x)$. Since the CM drifts along as a free particle, one usually chooses to study the problem in the CM frame. In this case $E_{\text{CM}} =$

$p_{\text{CM}}^2/2M$ drops out of the energy, and the plane wave factor in ψ representing CM motion becomes a constant. In short, one can forget all about the CM in the quantum theory just as in the classical theory.

N Particles in One Dimension

All the results but one generalize from $N=2$ to arbitrary N . The only exception is the result from the last subsection: for $N>2$, one generally cannot, by using CM and relative coordinates (or other sets of coordinates) reduce the problem to N independent one-particle problems. There are a few exceptions, the most familiar ones being Hamiltonians quadratic in the coordinates and momenta which may be reduced to a sum over oscillator Hamiltonians by the use of normal coordinates. In such cases the oscillators become independent and their energies add both in the classical and quantum cases. This result (with respect to the quantum oscillators) was assumed in the discussion on specific heats in Chapter 7.

*Exercise 10.1.3.** Consider the Hamiltonian of the coupled mass system:

$$\mathcal{H} = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \frac{1}{2} m\omega^2 [x_1^2 + x_2^2 + (x_1 - x_2)^2]$$

We know from Example 1.8.6 that \mathcal{H} can be decoupled if we use normal coordinates

$$x_{\text{I,II}} = \frac{x_1 \pm x_2}{2^{1/2}}$$

and the corresponding momenta

$$p_{\text{I,II}} = \frac{p_1 \pm p_2}{2^{1/2}}$$

(1) Rewrite \mathcal{H} in terms of normal coordinates. Verify that the normal coordinates are also canonical, i.e., that

$$\{x_i, p_j\} = \delta_{ij} \text{ etc.}; \quad i, j = \text{I, II}$$

Now quantize the system, promoting these variables to operators obeying

$$[X_i, P_j] = i\hbar\delta_{ij} \text{ etc.}; \quad i, j = \text{I, II}$$

Write the eigenvalue equation for H in the simultaneous eigenbasis of X_1 and X_{II} .

(2) Quantize the system directly, by promoting x_1 , x_2 , p_1 , and p_2 to quantum operators. Write the eigenvalue equation for H in the simultaneous eigenbasis of X_1 and X_2 . Now change from x_1 , x_2 (and of course $\partial/\partial x_1$, $\partial/\partial x_2$) to x_{I} , x_{II} (and $\partial/\partial x_{\text{I}}$, $\partial/\partial x_{\text{II}}$) in the differential equation. You should end up with the result from part (1).

In general, one can change coordinates and then quantize or first quantize and then change variables in the differential equation, if the change of coordinates is canonical. (We are assuming that all the variables are Cartesian. As mentioned earlier in the book, if one wants

to employ non-Cartesian coordinates, it is best to first quantize the Cartesian coordinates and then change variables in the differential equation.)

10.2. More Particles in More Dimensions

Mathematically, the problem of a single particle in two dimensions (in terms of Cartesian coordinates) is equivalent to that of two particles in one dimension. It is, however, convenient to use a different notation in the two cases. We will denote the two Cartesian coordinates of the single particle by x and y rather than x_1 and x_2 . Likewise the momenta will be denoted by p_x and p_y . The quantum operators will be called X and Y ; and P_x , and P_y , their common eigenkets $|xy\rangle$, $|p_x p_y\rangle$, respectively, and so on. The generalization to three dimensions is obvious. We will also write a position eigenket as $|\mathbf{r}\rangle$ and the orthonormality relation $\langle xyz|x'y'z'\rangle = \delta(x-x')\delta(y-y')\delta(z-z')$ as $\langle \mathbf{r}|\mathbf{r}'\rangle = \delta^3(\mathbf{r}-\mathbf{r}')$. The same goes for the momentum eigenkets $|\mathbf{p}\rangle$ also. When several particles labeled by numbers $1, \dots, N$ are involved, this extra label will also be used. Thus $|\mathbf{p}_1 \mathbf{p}_2\rangle$ will represent a two-particle state in which particle 1 has momentum \mathbf{p}_1 and particle 2 has momentum \mathbf{p}_2 and so on.

*Exercise 10.2.1** (Particle in a Three-Dimensional Box). Recall that a particle in a one-dimensional box extending from $x=0$ to L is confined to the region $0 \leq x \leq L$; its wave function vanishes at the edges $x=0$ and L and beyond (Exercise 5.2.5). Consider now a particle confined in a three-dimensional cubic box of volume L^3 . Choosing as the origin one of its corners, and the x , y , and z axes along the three edges meeting there, show that the normalized energy eigenfunctions are

$$\psi_E(x, y, z) = \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{n_x \pi x}{L}\right) \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{n_y \pi y}{L}\right) \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{n_z \pi z}{L}\right)$$

where

$$E = \frac{\hbar^2 \pi^2}{2ML^2} (n_x^2 + n_y^2 + n_z^2)$$

and n_i are positive integers.

*Exercise 10.2.2.** Quantize the two-dimensional oscillator for which

$$\mathcal{H} = \frac{p_x^2 + p_y^2}{2m} + \frac{1}{2} m \omega_x^2 x^2 + \frac{1}{2} m \omega_y^2 y^2$$

(1) Show that the allowed energies are

$$E = (n_x + 1/2)\hbar\omega_x + (n_y + 1/2)\hbar\omega_y, \quad n_x, n_y = 0, 1, 2, \dots$$

(2) Write down the corresponding wave functions in terms of single oscillator wave functions. Verify that they have definite parity (even/odd) number $x \rightarrow -x$, $y \rightarrow -y$ and that the parity depends only on $n = n_x + n_y$.

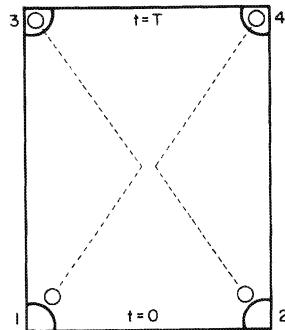


Figure 10.1. Two identical billiard balls start near holes 1 and 2 and end up in holes 3 and 4, respectively, as predicted by P_1 . The prediction of P_2 , that they would end up in holes 4 and 3, respectively, is wrong, even though the two final configurations would be indistinguishable to an observer who walks in at $t = T$.

(3) Consider next the *isotropic oscillator* ($\omega_x = \omega_y$). Write *explicit*, normalized eigenfunctions of the first three *states* (that is, for the cases $n=0$ and 1). Reexpress your results in terms of polar coordinates ρ and ϕ (for later use). Show that the degeneracy of a level with $E=(n+1)\hbar\omega$ is $(n+1)$.

*Exercise 10.2.3.** Quantize the three-dimensional *isotropic oscillator* for which

$$\mathcal{H} = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + \frac{1}{2} m\omega^2(x^2 + y^2 + z^2)$$

(1) Show that $E=(n+3/2)\hbar\omega$; $n=n_x+n_y+n_z$; $n_x, n_y, n_z=0, 1, 2, \dots$

(2) Write the corresponding eigenfunctions in terms of single-oscillator wave functions and verify that the parity of the level with a given n is $(-1)^n$. Reexpress the first four states in terms of spherical coordinates. Show that the degeneracy of a level with energy $E=(n+3/2)\hbar\omega$ is $(n+1)(n+2)/2$.

10.3. Identical Particles

The formalism developed above, when properly applied to a system containing identical particles, leads to some very surprising results. We shall say two particles are identical if they are exact replicas of each other in every respect—there should be no experiment that detects any intrinsic‡ difference between them. Although the definition of identical particles is the same classically and quantum mechanically, the implications are different in the two cases.

The Classical Case

Let us first orient ourselves by recapitulating the situation in classical physics. Imagine a billiard table with four holes, numbered 1 through 4 (Fig. 10.1). Near holes 1 and 2 rest two identical billiard balls. Let us call these balls 1 and 2. The difference between the labels reflects not any intrinsic difference in the balls (for they are identical) but rather a difference in their environments, namely, the holes near which they find themselves.

‡ By intrinsic I mean properties inherent to the particle, such as its charge or mass and not its location or momentum.

Now it follows from the definition of identity, that if these two balls are exchanged, the resulting configuration would appear exactly the same. Nonetheless these two configurations are treated as distinct in classical physics. In order for this distinction to be meaningful, there must exist some experiments in which these two configurations are inequivalent. We will now discuss one such experiment.

Imagine that at time $t=0$, two players propel the balls toward the center of the table. At once two physicists P_1 and P_2 take the initial-value data and make the following predictions:

$$\begin{aligned} P_1: & \left. \begin{array}{l} \text{ball 1 goes to hole 3} \\ \text{ball 2 goes to hole 4} \end{array} \right\} \quad \text{at } t=T \\ P_2: & \left. \begin{array}{l} \text{ball 1 goes to hole 4} \\ \text{ball 2 goes to hole 3} \end{array} \right\} \quad \text{at } t=T \end{aligned}$$

Say at time T we find that ball 1 ends up in hole 3 and ball 2 in hole 4. We declare that P_1 is correct and P_2 is wrong. Now, the configurations predicted by them for $t=T$ differ only by the exchange of two identical particles. If seen in isolation they would appear identical: an observer who walks in just at $t=T$ and is given the predictions of P_1 and P_2 will conclude that both are right. What do we know about the balls (that allows us to make a distinction between them and hence the two outcomes), that the newcomer does not? The answer of course is—their histories. Although both balls appear identical to the newcomer, we are able to trace the ball in hole 3 back to the vicinity of hole 1 and the one in hole 4 back to hole 2. Similarly at $t=0$, the two balls which seemed identical to us would be distinguishable to someone who had been following them from an earlier period. Now of course it is not really necessary that either we or any other observer be actually present in order for this distinction to exist. One imagines in classical physics the fictitious observer who sees everything and disturbs nothing; if he can make the distinction, the distinction exists.

To summarize, it is possible in classical mechanics to distinguish between identical particles by following their nonidentical trajectories (without disturbing them in any way). Consequently two configurations related by exchanging the identical particles are physically nonequivalent.

An immediate consequence of the above reasoning, and one that will play a dominant role in what follows, is that in quantum theory, which completely outlaws the notion of continuous trajectories for the particles, there exists no physical basis for distinguishing between identical particles. Consequently two configurations related by the exchange of identical particles must be treated as one and the same configuration and described by the same state vector. We now proceed to deduce the consequences of this restriction.

Two-Particle Systems—Symmetric and Antisymmetric States

Suppose we have a system of two *distinguishable* particles 1 and 2 and a position measurement on the system shows particle 1 to be at $x=a$ and particle 2 to be at

$x=b$. We write the state just after measurement as

$$|\psi\rangle = |x_1=a, x_2=b\rangle = |ab\rangle \quad (10.3.1)$$

where we are adopting the convention that the state of particle 1 is described by the first label (a) and that of particle 2 by the second label (b). Since the particles are distinguishable, the state obtained by exchanging them is distinguishable from the above. It is given by

$$|\psi\rangle = |ba\rangle$$

and corresponds to having found particle 1 at b and particle 2 at a .

Suppose we repeat the experiment with two *identical* particles and catch one at $x=a$ and the other at $x=b$. Is the state vector just after measurement $|ab\rangle$ or $|ba\rangle$? The answer is, neither. We have seen that in quantum theory two configurations related by the exchange of identical particles must be viewed as one and the same and be described by the same state vector. Since $|\psi\rangle$ and $\alpha|\psi\rangle$ are physically equivalent, we require that $|\psi(a, b)\rangle$, the state vector just after the measurement, satisfy the constraint

$$|\psi(a, b)\rangle = \alpha|\psi(b, a)\rangle \quad (10.3.2)$$

where α is any complex number. Since under the exchange

$$|ab\rangle \leftrightarrow |ba\rangle$$

and the two vectors are not multiples of each other‡ (i.e., are physically distinct) neither is acceptable. The problem is that our position measurement yields not an ordered pair of numbers (as in the distinguishable particle case) but just a pair of numbers: to assign them to the particles in a definite way is to go beyond what is physically meaningful in quantum theory. What our measurement *does* permit us to conclude is that the state vector is an eigenstate of $X_1 + X_2$ with eigenvalue $a+b$, the sum of the eigenvalue being insensitive to how the values a and b are assigned to the particles. In other words, given an *unordered* pair of numbers a and b we can still define a unique sum (but not difference). Now, there are just two product vectors, $|ab\rangle$ and $|ba\rangle$ with this eigenvalue, and the state vector lies somewhere in the two-dimensional degenerate (with respect to $X_1 + X_2$) eigenspace spanned by them. Let $|\psi(a, b)\rangle = \beta|ab\rangle + \gamma|ba\rangle$ be the allowed vector. If we impose the constraint Eq. (10.3.2):

$$\beta|ab\rangle + \gamma|ba\rangle = \alpha[\beta|ba\rangle + \gamma|ab\rangle]$$

we find, upon equating the coefficients of $|ab\rangle$ and $|ba\rangle$ that

$$\beta = \alpha\gamma, \quad \gamma = \alpha\beta$$

‡ We are assuming $a \neq b$. If $a = b$, the state is acceptable, but the choice we are agonizing over does not arise.

so that

$$\alpha = \pm 1 \quad (10.3.3)$$

It is now easy to construct the allowed state vectors. They are

$$|ab, S\rangle = |ab\rangle + |ba\rangle \quad (10.3.4)$$

called the *symmetric* state vector ($\alpha = 1$) and

$$|ab, A\rangle = |ab\rangle - |ba\rangle \quad (10.3.5)$$

called the *antisymmetric* state vector ($\alpha = -1$). (These are *unnormalized* vectors. Their normalization will be taken up shortly.)

More generally, if some variable Ω is measured and the values ω_1 and ω_2 are obtained, the state vector immediately following the measurement is either $|\omega_1\omega_2, S\rangle$ or $|\omega_1\omega_2, A\rangle$.[‡] Although we have made a lot of progress in nailing down the state vector corresponding to the measurement, we have still to find a way to choose between these two alternatives.

Bosons and Fermions

Although both S and A states seem physically acceptable (in that they respect the indistinguishability of the particles) we can go a step further and make the following assertion:

A given species of particles must choose once and for all between S and A states.

Suppose the contrary were true, and the Hilbert space of two identical particles contained both S and A vectors. Then the space also contains linear combinations such as

$$|\psi\rangle = \alpha|\omega_1\omega_2, S\rangle + \beta|\omega'_1\omega'_2, A\rangle$$

which are neither symmetric nor antisymmetric. So we rule out this possibility.

Nature seems to respect the constraints we have deduced. Particles such as the pion, photon, and graviton are *always* found in symmetric states and are called *bosons*, and particles such as the electron, proton, and neutron are *always* found in antisymmetric states and are called *fermions*.

Thus if we catch two identical bosons, one at $x=a$ and the other at $x=b$, the state vector immediately following the measurement is

$$\begin{aligned} |\psi\rangle &= |x_1=a, x_2=b\rangle + |x_1=b, x_2=a\rangle \\ &= |ab\rangle + |ba\rangle = |ab, S\rangle \end{aligned}$$

[‡] We are assuming Ω is nondegenerate. If not, let ω represent the eigenvalues of a complete set of commuting operators.

Had the particles been fermions, the state vector after the measurement would have been

$$\begin{aligned} |\psi\rangle &= |x_1=a, x_2=b\rangle - |x_1=b, x_2=a\rangle = |ab\rangle - |ba\rangle \\ &= |ab, A\rangle \end{aligned}$$

Note that although we still use the labels x_1 and x_2 , we do not attach them to the particles in any particular way. Thus having caught the bosons at $x=a$ and $x=b$, we need not agonize over whether $x_1=a$ and $x_2=b$ or vice versa. Either choice leads to the same $|\psi\rangle$ for bosons, and to state vectors differing only by an overall sign for fermions.

We are now in a position to deduce a fundamental property of fermions, which results from the antisymmetry of their state vectors. Consider a two-fermion state

$$|\omega_1\omega_2, A\rangle = |\omega_1\omega_2\rangle - |\omega_2\omega_1\rangle$$

Let us now set $\omega_1 = \omega_2 = \omega$. We find

$$|\omega\omega, A\rangle = |\omega\omega\rangle - |\omega\omega\rangle = 0 \quad (10.3.6)$$

This is the celebrated *Pauli exclusion principle*: *Two identical fermions cannot be in the same quantum state*. This principle has profound consequences—in statistical mechanics, in understanding chemical properties of atoms, in nuclear theory, astrophysics, etc. We will have occasion to return to it often.

With this important derivation out of our way, let us address a question that may have plagued you: our analysis has only told us that a given type of particle, say a pion, has to be either a boson or a fermion, but does not say which one. There are two ways to the answer. The first is by further cerebration, to be specific, within the framework of quantum field theory, which relates the spin of the particle to its “statistics”—which is the term physicists use to refer to its bosonic or fermionic nature. Since the relevant arguments are beyond the scope of this text I merely quote the results here. Recall that the spin of the particle is its internal angular momentum. The *magnitude* of spin happens to be an invariant for a particle (and thus serves as a label, like its mass or charge) and can have only one of the following values: 0, $\hbar/2$, \hbar , $3\hbar/2$, $2\hbar$, The spin statistics theorem, provable in quantum field theory, asserts that particles with (magnitude of spin) equal to an even multiple of $\hbar/2$ are bosons, and those with spin equal to an odd multiple of $\hbar/2$ are fermions. However, this connection, proven in three dimensions, does not apply to one dimension, where it is not possible to define spin or any form of angular momentum. (This should be clear classically.) Thus the only way to find if a particle in one dimension is a boson or fermion is to determine the symmetry of the wave function experimentally. This is the second method, to be discussed in a moment.

Before going on to this second method, let us note that the requirement that the state vector of two identical particles be symmetric or antisymmetric (under the exchange of the quantum numbers labeling them) applies in three dimensions as well, as will be clear by going through the arguments in one dimension. The only difference will be the increase in the number of labels. For example, the position

eigenket of a spin-zero boson will be labeled by three numbers x , y , and z . For fermions, which have spin at least equal to $\hbar/2$, the states will be labeled by the orientation of the spin as well as the orbital labels that describe spinless bosons.[‡] We shall consider just spin- $\frac{1}{2}$ particles, for which this label can take only two values, call them + and - or spin up and down (the meaning of these terms will be clear later). If we denote by ω all the orbital labels and by s the spin label, the state vector of the fermion that is antisymmetric under the exchange of the particles, i.e., under the exchange of all the labels, will be of the form

$$|\omega_1 s_1, \omega_2 s_2, A\rangle = |\omega_1 s_1, \omega_2 s_2\rangle - |\omega_2 s_2, \omega_1 s_1\rangle \quad (10.3.7)$$

We see that the state vector vanishes if

$$\omega_1 = \omega_2 \quad \text{and} \quad s_1 = s_2 \quad (10.3.8)$$

Thus we find once again that two fermions cannot be in the same quantum state, but we mean by a quantum state a state of definite ω and s . Thus two electrons can be in the same orbital state if their spin orientations are different.

We now turn to the second way of finding the statistics of a given species of particles, the method that works in one or three dimensions, because it appeals to a simple experiment which determines whether the two-particle state vector is symmetric or antisymmetric for the given species. As a prelude to the discussion of such an experiment, let us study in some detail the Hilbert space of bosons and fermions.

Bosonic and Fermionic Hilbert Spaces

We have seen that two identical bosons will always have symmetric state vectors and two identical fermions will always have antisymmetric state vectors. Let us call the Hilbert space of symmetric bosonic vectors \mathbb{V}_S and the Hilbert space of the antisymmetric fermionic vectors \mathbb{V}_A . We first examine the relation between these two spaces on the one hand and the direct product space $\mathbb{V}_{1\otimes 2}$ on the other.

The space $\mathbb{V}_{1\otimes 2}$ consists of all vectors of the form $|\omega_1 \omega_2\rangle = |\omega_1\rangle \otimes |\omega_2\rangle$. To each pair of vectors $|\omega_1 = a, \omega_2 = b\rangle$ and $|\omega_1 = b, \omega_2 = a\rangle$ there is one (unnormalized) bosonic vector $|\omega_1 = a, \omega_2 = b\rangle + |\omega_1 = b, \omega_2 = a\rangle$ and one fermionic vector $|\omega_1 = a, \omega_2 = b\rangle - |\omega_1 = a, \omega_2 = b\rangle$. If $a = b$; the vector $|\omega_1 = a, \omega_2 = a\rangle$ is already symmetric and we may take it to be the bosonic vector. There is no corresponding fermionic vector (the Pauli principle). Thus $\mathbb{V}_{1\otimes 2}$ has just enough basis vectors to form one bosonic Hilbert space and one fermionic Hilbert space. We express this relation as

$$\mathbb{V}_{1\otimes 2} = \mathbb{V}_S \oplus \mathbb{V}_A \quad (10.3.9)$$

[‡] Since spin has no classical counterpart, the operator representing it is not a function of the coordinate and momentum operators and it commutes with any orbital operator Ω . Thus spin may be specified simultaneously with the orbital variables.

with \mathbb{V}_S getting slightly more than half the dimensionality of $\mathbb{V}_{1\otimes 2}$.[‡] Our analysis has shown that at any given time, the state of two bosons is an element of \mathbb{V}_S and that of two fermions an element of \mathbb{V}_A . It can also be shown that a system that starts out in $\mathbb{V}_S(\mathbb{V}_A)$ remains in $\mathbb{V}_S(\mathbb{V}_A)$ (see Exercise 10.3.5). *Thus in studying two identical particles we need only consider \mathbb{V}_S or \mathbb{V}_A .* It is however convenient, for bookkeeping purposes, to view \mathbb{V}_S and \mathbb{V}_A as subspaces of $\mathbb{V}_{1\otimes 2}$ and the elements of \mathbb{V}_S or \mathbb{V}_A as elements also of $\mathbb{V}_{1\otimes 2}$.

Let us now consider the normalization of the vectors in \mathbb{V}_S . Consider first the eigenkets $|\omega_1\omega_2, S\rangle$ corresponding to a variable Ω with discrete eigenvalues. The unnormalized state vector is

$$|\omega_1\omega_2, S\rangle = |\omega_1\omega_2\rangle + |\omega_2\omega_1\rangle$$

Since $|\omega_1\omega_2\rangle$ and $|\omega_2\omega_1\rangle$ are orthonormal states in $\mathbb{V}_{1\otimes 2}$, the normalization factor is just $2^{-1/2}$, i.e.,

$$|\omega_1\omega_2, S\rangle = 2^{-1/2}[|\omega_1\omega_2\rangle + |\omega_2\omega_1\rangle] \quad (10.3.10a)$$

is a normalized eigenvector. You may readily check that $\langle\omega_1\omega_2, S|\omega_1\omega_2, S\rangle = 1$. The preceding discussion assumes $\omega_1 \neq \omega_2$. If $\omega_1 = \omega_2 = \omega$ the product ket $|\omega\omega\rangle$ is itself both symmetric and normalized and we choose

$$|\omega\omega, S\rangle = |\omega\omega\rangle \quad (10.3.10b)$$

Any vector $|\psi_S\rangle$ in \mathbb{V}_S may be expanded in terms of this Ω basis. As usual we identify

$$P_S(\omega_1, \omega_2) = |\langle\omega_1\omega_2, S|\psi_S\rangle|^2 \quad (10.3.11)$$

as the *absolute* probability of finding the particles in state $|\omega_1\omega_2, S\rangle$ when an Ω measurement is made on a system in state $|\psi_S\rangle$. The normalization condition of $|\psi_S\rangle$ and $P_S(\omega_1, \omega_2)$ may be written as

$$\begin{aligned} 1 &= \langle\psi_S|\psi_S\rangle = \sum_{\text{dist}} |\langle\omega_1\omega_2, S|\psi_S\rangle|^2 \\ &= \sum_{\text{dist}} P_S(\omega_1, \omega_2) \end{aligned} \quad (10.3.12a)$$

where \sum_{dist} denotes a sum over all physically *distinct* states. If ω_1 and ω_2 take values between ω_{\min} and ω_{\max} , then

$$\sum_{\text{dist}} = \sum_{\omega_2 = \omega_{\min}}^{\omega_{\max}} \sum_{\omega_1 = \omega_{\min}}^{\omega_2} \quad (10.3.12b)$$

In this manner we avoid counting *both* $|\omega_1\omega_2, S\rangle$ and $|\omega_2\omega_1, S\rangle$, which are physically equivalent. Another way is to count them both and then divide by 2.

[‡] Since every element of \mathbb{V}_S is perpendicular to every element of \mathbb{V}_A (you should check this) the dimensionality of $\mathbb{V}_{1\otimes 2}$ equals the sum of the dimensionalities of \mathbb{V}_S and \mathbb{V}_A .

What if we want the absolute probability density for some continuous variable such as X ? In this case we must take the projection of $|\psi_S\rangle$ on the normalized position eigenket:

$$|x_1x_2, S\rangle = 2^{-1/2} [|x_1x_2\rangle + |x_2x_1\rangle] \quad (10.3.13)$$

to obtain

$$P_S(x_1, x_2) = |\langle x_1x_2, S | \psi_S \rangle|^2 \quad (10.3.14)$$

The normalization condition for $P_S(x_1, x_2)$ and $|\psi_S\rangle$ is

$$1 = \iint P_S(x_1, x_2) \frac{dx_1 dx_2}{2} = \iint |\langle x_1x_2, S | \psi_S \rangle|^2 \frac{dx_1 dx_2}{2} \quad (10.3.15)$$

where the factor $1/2$ makes up for the double counting done by the $dx_1 dx_2$ integration.[‡] In this case it is convenient to define the wave function as

$$\psi_S(x_1, x_2) = 2^{-1/2} \langle x_1x_2, S | \psi_S \rangle \quad (10.3.16)$$

so that the normalization of ψ_S is

$$1 = \iint |\psi_S(x_1, x_2)|^2 dx_1 dx_2 \quad (10.3.17)$$

However, in this case

$$P_S(x_1, x_2) = 2 |\psi_S(x_1, x_2)|^2 \quad (10.3.18)$$

due to the rescaling. Now, note that

$$\begin{aligned} \psi_S(x_1, x_2) &= \frac{1}{2^{1/2}} \langle x_1x_2, S | \psi_S \rangle = \frac{1}{2} [\langle x_1x_2 | \psi_S \rangle + \langle x_2x_1 | \psi_S \rangle] \\ &= \langle x_1x_2 | \psi_S \rangle \end{aligned} \quad (10.3.19)$$

where we have exploited the fact that $|\psi_S\rangle$ is symmetrized between the particles and has the same inner product with $\langle x_1x_2 |$ and $\langle x_2x_1 |$. Consequently, the normalization

[‡] The points $x_1 = x_2 = x$ pose some subtle questions both with respect to the factor $1/2$ and the normalization of the kets $|xx, S\rangle$. We do not get into these since the points on the line $x_1 = x_2 = x$ make only an infinitesimal contribution to the integration in the $x_1 - x_2$ plane (of any smooth function). In the following discussion you may assume that quantities such as $P_S(x, x)$, $\psi_S(x, x)$ are all given by the limits $x_1 \rightarrow x_2 \rightarrow x$ of $P_S(x_1, x_2)$, $\psi_S(x_1, x_2)$, etc.

condition Eq. (10.3.17) becomes

$$1 = \langle \psi_S | \psi_S \rangle = \iint |\psi_S|^2 dx_1 dx_2 = \iint \langle \psi_S | x_1 x_2 \rangle \langle x_1 x_2 | \psi_S \rangle dx_1 dx_2$$

which makes sense, as $|\psi_S\rangle$ is an element of $\mathbb{V}_{1\otimes 2}$ as well. Note, however, that the kets $|x_1 x_2\rangle$ enter the definition of the wave function Eq. (10.3.19), and the normalization integral above, only as bookkeeping devices. They are not elements of \mathbb{V}_S and the inner product $\langle x_1 x_2 | \psi \rangle$ would be of no interest to us, were it not for the fact that the quantity *that is* of physical interest $\langle x_1 x_2, S | \psi_S \rangle$, is related to it by just a scale factor of $2^{1/2}$. Let us now consider a concrete example. We measure the energy of two noninteracting bosons in a box extending from $x=0$ to $x=L$ and find them to be in the quantum states $n=3$ and $n=4$. The normalized state vector just after measurement is then

$$|\psi_S\rangle = \frac{|3, 4\rangle + |4, 3\rangle}{2^{1/2}} \quad (10.3.20)$$

in obvious notation. The wave function is

$$\begin{aligned} \psi_S(x_1, x_2) &= 2^{-1/2} \langle x_1 x_2, S | \psi_S \rangle \\ &= \frac{1}{2} (\langle x_1 x_2 | + \langle x_2 x_1 |) \left(\frac{|3, 4\rangle + |4, 3\rangle}{2^{1/2}} \right) \\ &= \frac{1}{2(2^{1/2})} [\langle x_1 x_2 | 3, 4 \rangle + \langle x_1 x_2 | 4, 3 \rangle + \langle x_2 x_1 | 3, 4 \rangle + \langle x_2 x_1 | 4, 3 \rangle] \\ &= \frac{1}{2(2^{1/2})} [\psi_3(x_1)\psi_4(x_2) + \psi_4(x_1)\psi_3(x_2) + \psi_3(x_2)\psi_4(x_1) \\ &\quad + \psi_4(x_2)\psi_3(x_1)] \\ &= 2^{-1/2} [\psi_3(x_1)\psi_4(x_2) + \psi_4(x_1)\psi_3(x_2)] \\ &= \langle x_1 x_2 | \psi_S \rangle \end{aligned} \quad (10.3.21a)$$

where in all of the above,

$$\psi_n(x) = \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{n\pi x}{L}\right) \quad (10.3.21b)$$

These considerations apply with obvious modifications to the fermionic space \mathbb{V}_A . The basis vectors are of the form

$$|\omega_1 \omega_2, A\rangle = 2^{-1/2} [|\omega_1 \omega_2\rangle - |\omega_2 \omega_1\rangle] \quad (10.3.22)$$

(The case $\omega_1 = \omega_2$ does not arise here.) The wave function is once again

$$\begin{aligned}\psi_A(x_1, x_2) &= 2^{-1/2} \langle x_1 x_2, A | \psi_A \rangle \\ &= \langle x_1 x_2 | \psi_A \rangle\end{aligned}\quad (10.3.23)$$

and as in the bosonic case

$$P_A(x_1, x_2) = 2|\psi_A(x_1, x_2)|^2 \quad (10.3.24)$$

The normalization condition is

$$1 = \iint P_A(x_1, x_2) \frac{dx_1 dx_2}{2} = \iint |\psi_A(x_1, x_2)|^2 dx_1 dx_2 \quad (10.3.25)$$

Returning to our example of two particles in a box, if we had obtained the values $n=3$ and $n=4$, then the state just after measurement would have been

$$|\psi_A\rangle = \frac{|3, 4\rangle - |4, 3\rangle}{2^{1/2}} \quad (10.3.26)$$

(We may equally well choose

$$|\psi_A\rangle = \frac{|4, 3\rangle - |3, 4\rangle}{2^{1/2}}$$

which makes no physical difference). The corresponding wave function may be written in the form of a determinant:

$$\begin{aligned}\psi_A(x_1, x_2) &= \langle x_1 x_2 | \psi_A \rangle = 2^{-1/2} [\psi_3(x_1)\psi_4(x_2) - \psi_4(x_1)\psi_3(x_2)] \\ &= 2^{-1/2} \begin{vmatrix} \psi_3(x_1) & \psi_4(x_1) \\ \psi_3(x_2) & \psi_4(x_2) \end{vmatrix}\end{aligned}\quad (10.3.27)$$

Had we been considering the state $|\omega_1 \omega_2, A\rangle$ [Eq. (10.3.22)],‡

$$\psi_A(x_1, x_2) = 2^{-1/2} \begin{vmatrix} \psi_{\omega_1}(x_1) & \psi_{\omega_2}(x_1) \\ \psi_{\omega_1}(x_2) & \psi_{\omega_2}(x_2) \end{vmatrix} \quad (10.3.28)$$

Determination of Particle Statistics

We are finally ready to answer the old question: how does one determine empirically the statistics of a given species, i.e., whether it is a boson or fermion, without turning to the spin statistics theorem? For concreteness, let us say we have two identical noninteracting pions and wish to find out if they are bosons or fermions.

‡ The determinantal form of ψ_A makes it clear that ψ_A vanishes if $x_1 = x_2$ or $\omega_1 = \omega_2$.

We proceed as follows. We put them in a one-dimensional box[‡] and make an energy measurement. Say we find one in the state $n=3$ and the other in the state $n=4$. The probability distribution in x space would be, depending on their statistics,

$$\begin{aligned} P_{S/A}(x_1, x_2) &= 2|\psi_{S/A}(x_1, x_2)|^2 \\ &= 2|2^{-1/2}[\psi_3(x_1)\psi_4(x_2) \pm \psi_4(x_1)\psi_3(x_2)]|^2 \\ &= |\psi_3(x_1)|^2|\psi_4(x_2)|^2 + |\psi_4(x_1)|^2|\psi_3(x_2)|^2 \\ &\quad \pm [\psi_3^*(x_1)\psi_4(x_1)\psi_4^*(x_2)\psi_3(x_2) + \psi_4^*(x_1)\psi_3(x_1)\psi_3^*(x_2)\psi_4(x_2)] \end{aligned} \quad (10.3.29)$$

Compare this situation with two particles carrying labels 1 and 2, but otherwise identical,[§] with particle 1 in state 3 and described by a probability distribution $|\psi_3(x)|^2$, and particle 2 in state 4 and described by the probability distribution $|\psi_4(x)|^2$. In this case, the first term represents the probability that particle 1 is at x_1 and particle 2 is at x_2 , while the second gives the probability for the exchanged event. The sum of these two terms then gives $P_D(x_1, x_2)$, the probability for finding one at x_1 and the other at x_2 , with no regard paid to their labels. (The subscript D denotes distinguishable.) The next two terms, called interference terms, remind us that there is more to identical particles in quantum theory than just their identical characteristics: they have no separate identities. Had they separate identities (as in the classical case) and we were just indifferent to which one arrives at x_1 and which one at x_2 , we would get just the first two terms. There is a parallel between this situation and the double-slit experiment, where the probabilities for finding a particle at a given point x on the screen with both slits open was not the sum of the probabilities with either slit open. In both cases, the interference terms arise, because in quantum theory, when an event can take place in two (or more) indistinguishable ways, we add the corresponding amplitudes and not the corresponding probabilities.

Just as we were not allowed then to assign a definite trajectory to the particle (through slits 1 or 2), we are not allowed now to assign definite labels to the two particles.

The interference terms tell us if the pions are bosons or fermions. The difference between the two cases is most dramatic as $x_1 \rightarrow x_2 \rightarrow x$:

$$P_A(x_1 \rightarrow x, x_2 \rightarrow x) \rightarrow 0 \text{ (Pauli principle applied to state } |x\rangle) \quad (10.3.30)$$

whereas

$$P_S(x_1 \rightarrow x, x_2 \rightarrow x) = 2[|\psi_3(x)|^2|\psi_4(x)|^2 + |\psi_4(x)|^2|\psi_3(x)|^2] \quad (10.3.31)$$

which is twice as big as $P_D(x_1 \rightarrow x, x_2 \rightarrow x)$, the probability density for two distinct label carrying (but otherwise identical) particles, whose labels are disregarded in the position measurement.

One refers to the tendency of fermions to avoid each other (i.e., avoid the state $x_1 = x_2 = x$) as obeying “Fermi–Dirac statistics” and the tendency of bosons to

[‡] We do this to simplify the argument. The basic idea works just as well in three dimensions.

[§] The label can, for example, be the electric charge.

conglomerate as “obeying Bose–Einstein statistics,” after the physicists who first explored the consequences of the antisymmetrization and symmetrization requirements on the statistical mechanics of an ensemble of fermions and bosons, respectively. (This is the reason for referring to the bosonic/fermionic nature of a particle as its statistics.)

Given the striking difference in the two distributions, we can readily imagine deciding (once and for all) whether pions are bosons or fermions by preparing an ensemble of systems (with particles in $n=3$ and 4) and measuring $P(x_1, x_2)$.

Note that $P(x_1, x_2)$ helps us decide not only whether the particles are bosons or fermions, but also whether they are identical in the first place. In other words, if particles that we think are identical differ with respect to some label that we are not aware of, the nature of the interference term will betray this fact. Imagine, for example, two bosons, call them K and \bar{K} , which are identical with respect to mass and charge, but different with respect to a quantum number called “hypercharge.” Let us assume we are ignorant of hypercharge. In preparing an ensemble that we think contains N identical pairs, we will actually be including some (K, K) pairs, some (\bar{K}, \bar{K}) pairs. If we now make measurements on the ensemble and extract the distribution $P(x_1, x_2)$ (once again ignoring the hypercharge), we will find the interference term has the + sign but is not as big as it should be. If the ensemble contained only identical bosons, $P(x, x)$ should be twice as big as $P_D(x, x)$, which describes label-carrying particles; if we get a ratio less than 2, we know the ensemble is contaminated by label-carrying particles which produce no interference terms.

From the above discussions, it is also clear that one cannot hastily conclude, upon catching two electrons in the same orbital state in three dimensions that they are not fermions. In this case, the label we are ignoring is the spin orientation s . As mentioned earlier on, s can have only two values, call them + and −. If we assume that s never changes (during the course of the experiment) it can serve as a particle label that survives with time. If $s=+$ for one electron and − for the other, they are like two distinct particles and *can* be in the same orbital state. The safe thing to do here is once again to work with an ensemble rather than an isolated measurement. Since we are ignorant of spin, our ensemble will contain $(+, +)$ pairs, $(-, -)$ pairs, and $(+, -)$ pairs. The $(+, +)$ and $(-, -)$ pairs are identical fermions and will produce a negative interference term, while the $(+, -)$ pairs will not. Thus we will find $P(\mathbf{r}, \mathbf{r})$ is smaller than $P_D(\mathbf{r}, \mathbf{r})$ describing labeled particles, but not zero. This will tell us that our ensemble has identical fermion pairs contaminated by pairs of distinguishable particles. It will then be up to us to find the nature of the hidden degree of freedom which provides the distinction.

Systems of N Identical Particles

The case $N=2$ lacks one feature that is found at larger N . We illustrate it by considering the case of three identical particles in a box. Let us say that an energy measurement shows the quantum numbers of the particles to be n_1 , n_2 , and n_3 . Since the particles are identical, all we can conclude from this observation is that the total energy is

$$E = \left(\frac{\hbar^2 \pi^2}{2mL^2} \right) (n_1^2 + n_2^2 + n_3^2)$$

Now there are $3! =$ six product states with this energy: $|n_1 n_2 n_3\rangle$, $|n_1 n_3 n_2\rangle$, $|n_2 n_3 n_1\rangle$, $|n_2 n_1 n_3\rangle$, $|n_3 n_2 n_1\rangle$, and $|n_3 n_1 n_2\rangle$. The physical states are elements of the six-dimensional eigenspace spanned by these vectors and distinguished by the property that under the exchange of *any* two particle labels, the state vector changes only by a factor α . Since double exchange of the same two labels is equivalent to no exchange, we conclude as before that $\alpha = \pm 1$. There are only two states with this property:

$$\begin{aligned} |n_1 n_2 n_3, S\rangle &= \frac{1}{(3!)^{1/2}} [|n_1 n_2 n_3\rangle + |n_1 n_3 n_2\rangle + |n_2 n_3 n_1\rangle \\ &\quad + |n_2 n_1 n_3\rangle + |n_3 n_2 n_1\rangle + |n_3 n_1 n_2\rangle] \end{aligned} \quad (10.3.32)$$

called the totally symmetric state,[‡] for which $\alpha = +1$ for all three possible exchanges ($1 \leftrightarrow 2$, $2 \leftrightarrow 3$, $1 \leftrightarrow 3$); and

$$\begin{aligned} |n_1 n_2 n_3, A\rangle &= \frac{1}{(3!)^{1/2}} [|n_1 n_2 n_3\rangle - |n_1 n_3 n_2\rangle + |n_2 n_3 n_1\rangle \\ &\quad - |n_2 n_1 n_3\rangle + |n_3 n_1 n_2\rangle - |n_3 n_2 n_1\rangle] \end{aligned} \quad (10.3.33)$$

called the totally antisymmetric state, for which $\alpha = -1$ for all three possible exchanges.

Bosons will always pick the S states and fermions, the A states. It follows that no two fermions can be in the same state.

As in the $N=2$ case, the wave function in the X basis is

$$\psi_{S/A}(x_1, x_2, x_3) = (3!)^{-1/2} \langle x_1 x_2 x_3, S/A | \psi_{S/A} \rangle = \langle x_1 x_2 x_3 | \psi_{S/A} \rangle \quad (10.3.34)$$

and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi_{S/A}|^2 dx_1 dx_2 dx_3 = 1$$

For instance, the wave function associated with $|n_1 n_2 n_3, S/A\rangle$, Eqs. (10.3.33) and (10.3.34), is

$$\begin{aligned} \psi_{n_1 n_2 n_3}(x_1, x_2, x_3, S/A) &= (3!)^{-1/2} [\psi_{n_1}(x_1) \psi_{n_2}(x_2) \psi_{n_3}(x_3) \pm \psi_{n_1}(x_1) \psi_{n_3}(x_2) \psi_{n_2}(x_3) \\ &\quad + \psi_{n_2}(x_1) \psi_{n_3}(x_2) \psi_{n_1}(x_3) \pm \psi_{n_2}(x_1) \psi_{n_1}(x_2) \psi_{n_3}(x_3) \\ &\quad + \psi_{n_3}(x_1) \psi_{n_1}(x_2) \psi_{n_2}(x_3) \pm \psi_{n_3}(x_1) \psi_{n_2}(x_2) \psi_{n_1}(x_3)] \end{aligned} \quad (10.3.35)$$

[‡] The normalization factor $(3!)^{-1/2}$ is correct only if all three n 's are different. If, for example, $n_1 = n_2 = n_3 = n$, then the product state $|nnn\rangle$ is normalized and symmetric and can be used as the S state. A similar question does not arise for the fermion state due to the Pauli principle.

The fermion wave function may once again be written as a determinant:

$$\psi_{n_1 n_2 n_3}(x_1, x_2, x_3, A) = \frac{1}{(3!)^{1/2}} \begin{vmatrix} \psi_{n_1}(x_1) & \psi_{n_2}(x_1) & \psi_{n_3}(x_1) \\ \psi_{n_1}(x_2) & \psi_{n_2}(x_2) & \psi_{n_3}(x_2) \\ \psi_{n_1}(x_3) & \psi_{n_2}(x_3) & \psi_{n_3}(x_3) \end{vmatrix} \quad (10.3.36)$$

Using the properties of the determinant, one easily sees that ψ vanishes if two of the x 's or n 's coincide. All these results generalize directly to any higher N .

Two questions may bother you at this point.

Question I. Consider the case $N=3$. There are three possible exchanges here: $(1 \leftrightarrow 2)$, $(1 \leftrightarrow 3)$, and $(2 \leftrightarrow 3)$. The S states pick up a factor $\alpha=+1$ for all three exchanges, while the A states pick up $\alpha=-1$ for all three exchanges. What about states for which some of the α 's are $+1$ and the others -1 ? Such states do not exist. You may verify this by exhaustion: take the $3!$ product vectors and try to form such a linear combination. Since a general proof for this case and all N involves group theory, we will not discuss it here. Note that since we get only two acceptable vectors for every $N!$ product vectors, the direct product space for $N \geq 3$ is bigger (in dimensionality) than $\mathbb{V}_S \oplus \mathbb{V}_A$.

Question II. We have tacitly assumed that if *two* identical particles of a given species always pick the S (or A) state, so will three or more, i.e., we have extended our definition of bosons and fermions from $N=2$ to all N . What if two pions always pick the S state while *three* always pick the A state? While intuition revolts at such a possibility, it still needs to be formally ruled out. We do so at the end of the next subsection.

When Can We Ignore Symmetrization and Antisymmetrization?

A basic assumption physicists make before they can make any headway is that they can single out some part of the universe (the system) and study it in isolation from the rest. While no system is truly isolated, one can often get close to this ideal. For instance, when we study the oscillations of a mass coupled to a spring, we ignore the gravitational pull of Pluto.

Classically, the isolation of the system is expressed by the separability of the Hamiltonian of the universe:

$$\mathcal{H}_{\text{universe}} = \mathcal{H}_{\text{sys}} + \mathcal{H}_{\text{rest}} \quad (10.3.37)$$

where \mathcal{H}_{sys} is a function of the system coordinates and momenta alone. It follows that the time evolution of the system's p 's and q 's are independent of what is going on in the rest of the universe. In our example, this separability is ruined (to give just one example) by the gravitational interaction between the mass and Pluto, which depends on their relative separation. If we neglect this absurdly small effect (and other such effects) we obtain separability to an excellent approximation.

Quantum mechanically, separability of H leads to the factorization of the wave function of the universe:

$$\psi_{\text{universe}} = \psi_{\text{sys}} \cdot \psi_{\text{rest}} \quad (10.3.38)$$

where ψ_{sys} is a function only of system coordinates, collectively referred to as x_s . Thus if we want the probability that the system has a certain coordinate x_s , and do not care about the rest, we find (symbolically)

$$\begin{aligned} P(x_s) &= \int |\psi_{\text{universe}}(x_s, x_{\text{rest}})|^2 dx_{\text{rest}} \\ &= |\psi_{\text{sys}}(x_s)|^2 \int |\psi(x_{\text{rest}})|^2 dx_{\text{rest}} \\ &= |\psi_{\text{sys}}(x_s)|^2 \end{aligned} \quad (10.3.39)$$

We could have obtained this result by simply ignoring ψ_{rest} from the outset.

Things get complicated when the system and the “rest” contain identical particles. Even if there is no interaction between the system and the rest, i.e., the Hamiltonian is separable, product states are not allowed and only S or A states must be used. Once the state vector fails to factorize, we will no longer have

$$P(x_s, x_{\text{rest}}) = P(x_s)P(x_{\text{rest}}) \quad (10.3.40)$$

(i.e., the systems will not be statistically independent), and we can not integrate out $P(x_{\text{rest}})$ and regain $P(x_s)$.

Now it seems reasonable that at least in certain cases it should be possible to get away with the product state and ignore the symmetrization or antisymmetrization conditions.

Suppose, for example, that at $t=0$, we find one pion in the ground state of an oscillator potential centered around a point on earth and another pion in the same state, but on the moon. It seems reasonable that we can give the particles the labels “earth pion” and “moon pion,” which will survive with time. Although we cannot follow their trajectories, we can follow their wave functions: we know the first wave function is a Gaussian $G_E(x_E)$ centered at a point in the lab on earth and that the second is a Gaussian $G_M(x_M)$ centered at a point on the moon. If we catch a pion somewhere on earth at time t , the theory tells us that it is almost certainly the “earth pion” and that the chances of its being the “moon pion” are absurdly small. Thus the uncertainty in the position of each pion is compensated by a separation that is much larger. (Even in classical mechanics, it is not necessary to know the trajectories exactly to follow the particles; the band of uncertainty about each trajectory has merely to be much thinner than the minimum separation between the particles during their encounter.) We therefore believe that if we assumed

$$\psi(x_E, x_M) = G_E(x_E)G_M(x_M) \quad (10.3.41)$$

we should be making an error that is as negligible as is the chance of finding the earth pion on the moon and vice versa. Given this product form, the person on earth can compute the probability for finding the earth pion at some x by integrating out the moon pion:

$$\begin{aligned} P(x_E) &= |G_E(x_E)|^2 \int |G_M(x_M)|^2 dx_M \\ &= |G_E(x_E)|^2 \end{aligned} \quad (10.3.42)$$

Likewise the person on the moon, who does not care about (i.e., sums over) the earth pion will obtain

$$P(x_M) = |G_M(x_M)|^2 \quad (10.3.43)$$

Let us now verify that if we took a properly symmetrized wave function it leads to essentially the same predictions (with negligible differences).

Let us start with

$$\psi_S(x_1, x_2) = 2^{-1/2} [G_E(x_1)G_M(x_2) + G_M(x_1)G_E(x_2)] \quad (10.3.44)$$

We use the labels x_1 and x_2 rather than x_E and x_M to emphasize that the pions are indeed being treated as indistinguishable. Now, the probability (density) of finding one particle near x_1 and one near x_2 is

$$\begin{aligned} P(x_1, x_2) &= 2|\psi|^2 = |G_E(x_1)|^2|G_M(x_2)|^2 + |G_M(x_1)|^2|G_E(x_2)|^2 \\ &\quad + G_E^*(x_1)G_M(x_1)G_M^*(x_2)G_E(x_2) \\ &\quad + G_M^*(x_1)G_E(x_1)G_E^*(x_2)G_M(x_2) \end{aligned} \quad (10.3.45)$$

Let us ask for the probability of finding one particle near some point x_E on the earth, with no regard to the other. This is given by setting *either* one of the variables (say x_1) equal to x_E and integrating out the other [since $P(x_1, x_2) = P(x_2, x_1)$]. There is no need to divide by 2 in doing this integration (why?). We get

$$\begin{aligned} P(x_E) &= |G_E(x_E)|^2 \int |G_M(x_2)|^2 dx_2 + |G_M(x_E)|^2 \int |G_E(x_2)|^2 dx_2 \\ &\quad + G_E^*(x_E)G_M(x_E) \int G_M^*(x_2)G_E(x_2) dx_2 \\ &\quad + G_M^*(x_E)G_E(x_E) \int G_E^*(x_2)G_M(x_2) dx_2 \end{aligned} \quad (10.3.46)$$

The first term is what we would get if we begin with a product wave function Eq. (10.3.41) and integrate out x_M . The other three terms are negligible since G_M is peaked on the moon and is utterly negligible at a point x_E on the earth. Similarly if we asked for $P(x_M)$, where x_M is a point on the moon, we will again get $|G_M(x_M)|^2$.

The labels “earth pion” and “moon pion” were useful only because the two Gaussians remained well separated for all times (being stationary states). If the two Gaussians had not been bound by the oscillating wells, and were wave packets drifting toward each other, the labeling (and the factorized wave function) would have become invalid when the Gaussians begin to have a significant overlap. The point is that at the start of any experiment, one can always assign the particles some labels. These labels acquire a physical significance only if they survive for some time. Labels like “a particle of mass m and charge +1” survive forever, while the longevity of a label like “earth pion” is controlled by whether or not some other pion is in the vicinity.

A dramatic illustration of this point is provided by the following example. At $t=0$ we catch two pions, one at $x=a$ and the other at $x=b$. We can give them the labels a and b since the two delta functions do not overlap even if a and b are in the same room. We may describe the initial state by a product wave function. But this labeling is quite useless, since after the passage of an infinitesimal period of time, the delta functions spread out completely: the probability distributions become constants. You may verify this by examining $|U(x, t; a, 0)|^2$ (the “fate” of the delta function)[‡] or by noting that $\Delta P = \infty$ for a delta function (the particle has all possible velocities from 0 to ∞) and which, therefore, spreads out in no time.

All these considerations apply with no modification to two fermions: the two cases differ in the sign of the interference term, which is irrelevant to these considerations.

What if there are three pions, two on earth and one on the moon? Since the two on the earth (assuming that their wave functions appreciably overlap) *can* be confused with each other, we must symmetrize between them, and the total wave function will be, in obvious notation,

$$\psi(x_{E_1}, x_{E_2}, x_M) = \psi_S(x_{E_1}, x_{E_2}) \cdot \psi(x_M) \quad (10.3.47)$$

The extension of this result to more particles and to fermions is obvious.

At this point the answer to Question II raised at the end of the last subsection becomes apparent. Suppose three-pion systems picked the A state while two-pion systems picked the S state. Let two of the three pions be on earth and the third one on the moon. Then, by assumption, the following function should provide an excellent approximation:

$$\psi(x_{E_1}, x_{E_2}, x_M) = \psi_A(x_{E_1}, x_{E_2})\psi(x_M) \quad (10.3.48)$$

If we integrate over the moon pion we get

$$P(x_{E_1}, x_{E_2}) = 2|\psi_A(x_{E_1}, x_{E_2})|^2 \quad (10.3.49)$$

We are thus led to conclude that two pions on earth will have a probability distribution corresponding to two fermions if there is a third pion on the moon and a distribution expected to two bosons if there is not a third one on the moon. Such

[‡] It is being assumed that the particles are free.

absurd conclusions are averted only if the statistics depend on the species and not the number of particles.

A word of caution before we conclude this long discussion. If two particles have nonoverlapping wave functions in x space, then it is only in x space that a product wave function provides a good approximation to the exact symmetrized wave function, which in our example was

$$\psi_s(x_1, x_2) = 2^{-1/2} [G_E(x_1)G_M(x_2) + G_M(x_1)G_E(x_2)] \quad (10.3.50)$$

The formal reason is that for any choice of the arguments x_1 and x_2 , only one or the other of the two terms in the right-hand side is important. (For example, if x_1 is on the earth and x_2 is on the moon, only the first piece is important.) Physically it is because the chance of finding one pion in the territory of the other is negligible and interference effects can be ignored.

If, however, we wish to switch to another basis, say the P basis, we must consider the Fourier transform of the symmetric function ψ_s and not the product, so that we end up with a symmetrized wave function in p space. The physical reason for this is that the two pions have the same momentum distributions—with $\langle P \rangle = 0$ and identical Gaussian fluctuations about this mean—since the momentum content of the oscillator is independent of its location. Consequently, there are no grounds in P space for distinguishing between them. Thus when a momentum measurement (which says nothing about the positions) yields two numbers, we cannot assign them to the pions in a unique way. Formally, symmetrization is important because the p -space wave functions of the pions overlap strongly and there exist values for the two momenta (both ≈ 0) for which both terms in the symmetric wave function are significant.

By the same token, if there are two particles with nonoverlapping wave functions in p space, we may describe the system by a product wave function in this space (using labels like “fast” and “slow” instead of “earth” and “moon” to distinguish between them), but not in another space where the distinction between them is absent. It should be clear that these arguments apply not just to X or P but to any arbitrary variable Ω .

*Exercise 10.3.1.** Two identical bosons are found to be in states $|\phi\rangle$ and $|\psi\rangle$. Write down the normalized state vector describing the system when $\langle\phi|\psi\rangle \neq 0$.

*Exercise 10.3.2.** When an energy measurement is made on a system of three bosons in a box, the n values obtained were 3, 3, and 4. Write down a symmetrized, normalized state vector.

*Exercise 10.3.3.** Imagine a situation in which there are three particles and only three states a , b , and c available to them. Show that the total number of allowed, distinct configurations for this system is

- (1) 27 if they are labeled
- (2) 10 if they are bosons
- (3) 1 if they are fermions

*Exercise 10.3.4.** Two identical particles of mass m are in a one-dimensional box of length L . Energy measurement of the system yields the value $E_{\text{sys}} = \hbar^2 \pi^2 / mL^2$. Write down the state vector of the system. Repeat for $E_{\text{sys}} = 5\hbar^2 \pi^2 / 2mL^2$. (There are two possible vectors in this case.) You are not told if they are bosons or fermions. You may assume that the only degrees of freedom are orbital.

*Exercise 10.3.5.** Consider the *exchange operator* P_{12} whose action on the X basis is

$$P_{12}|x_1, x_2\rangle = |x_2, x_1\rangle$$

- (1) Show that P_{12} has eigenvalues ± 1 . (It is Hermitian and unitary.)
- (2) Show that its action on the basis ket $|\omega_1, \omega_2\rangle$ is also to exchange the labels 1 and 2, and hence that $\mathbb{V}_{S,A}$ are its eigenspaces with eigenvalues ± 1 .
- (3) Show that $P_{12}X_1P_{12} = X_2$, $P_{12}X_2P_{12} = X_1$ and similarly for P_1 and P_2 . Then show that $P_{12}\Omega(X_1, P_1; X_2, P_2)P_{12} = \Omega(X_2, P_2; X_1, P_1)$. [Consider the action on $|x_1, x_2\rangle$ or $|p_1, p_2\rangle$. As for the functions of X and P , assume they are given by power series and consider any term in the series. If you need help, peek into the discussion leading to Eq. (11.2.22).]
- (4) Show that the Hamiltonian and propagator for two *identical* particles are left unaffected under $H \rightarrow P_{12}HP_{12}$ and $U \rightarrow P_{12}UP_{12}$. Given this, show that any eigenstate of P_{12} continues to remain an eigenstate with the same eigenvalue as time passes, i.e., elements of $\mathbb{V}_{S,A}$ never leave the symmetric or antisymmetric subspaces they start in.

*Exercise 10.3.6.** Consider a composite object such as the hydrogen atom. Will it behave as a boson or fermion? Argue in general that objects containing an even/odd number of fermions will behave as bosons/fermions.

Symmetries and Their Consequences

11.1. Overview

In Chapter 2, we explored the consequences of the symmetries of the Hamiltonian. We saw the following:

- (1) If \mathcal{H} is invariant under the infinitesimal canonical transformation generated by a variable $g(q, p)$, then g is conserved.
- (2) Any canonical transformation that leaves \mathcal{H} invariant maps solutions to the equations of motion into other solutions. Equivalently, an experiment and its transformed version will give the same result if the transformation is canonical and leaves \mathcal{H} invariant.

Here we address the corresponding results in quantum mechanics.[‡]

11.2. Translational Invariance in Quantum Theory

Consider a single particle in one dimension. How shall we define translational invariance? Since a particle in an arbitrary state has neither a well-defined position nor a well-defined energy, we cannot define translational invariance to be the invariance of the energy under an infinitesimal shift in the particle position. Our previous experience, however, suggests that in the quantum formulation the expectation values should play the role of the classical variables. We therefore make the correspondence shown in Table 11.1.

Having agreed to formulate the problem in terms of expectation values, we still have two equivalent ways to interpret the transformations:

$$\langle X \rangle \rightarrow \langle X \rangle + \varepsilon \quad (11.2.1a)$$

$$\langle P \rangle \rightarrow \langle P \rangle \quad (11.2.1b)$$

[‡] It may be worth refreshing your memory by going through Sections 2.7 and 2.8.

Table 11.1. Correspondence between Classical and Quantum Mechanical Concepts Related to Translational Invariance

Concept	Classical mechanics	Quantum mechanics
Translation	$x \rightarrow x + \varepsilon$	$\langle X \rangle \rightarrow \langle X \rangle + \varepsilon$
Translational invariance	$p \rightarrow p$	$\langle P \rangle \rightarrow \langle P \rangle$
Conservation law	$\mathcal{H} \rightarrow \mathcal{H}$	$\langle H \rangle \rightarrow \langle H \rangle$
	$\dot{p} = 0$	$\langle \dot{P} \rangle = 0$ (anticipated)

The first is to say that under the infinitesimal translation, each state $|\psi\rangle$ gets modified into a *translated state*, $|\psi_\varepsilon\rangle$ such that

$$\langle \psi_\varepsilon | X | \psi_\varepsilon \rangle = \langle \psi | X | \psi \rangle + \varepsilon \quad (11.2.2a)$$

$$\langle \psi_\varepsilon | P | \psi_\varepsilon \rangle = \langle \psi | P | \psi \rangle \quad (11.2.2b)$$

In terms of $T(\varepsilon)$, the *translation operator*, which translates the state (and which will be constructed explicitly in a while)

$$T(\varepsilon) |\psi\rangle = |\psi_\varepsilon\rangle \quad (11.2.3)$$

Eq. (11.2.2) becomes

$$\langle \psi | T^\dagger(\varepsilon) X T(\varepsilon) | \psi \rangle = \langle \psi | X | \psi \rangle + \varepsilon \quad (11.2.4a)$$

$$\langle \psi | T^\dagger(\varepsilon) P T(\varepsilon) | \psi \rangle = \langle \psi | P | \psi \rangle \quad (11.2.4b)$$

This point of view is called the active transformation picture (in the terminology of Section 1.7) and corresponds to physically displacing the particle to the *right* by ε .

The second point of view is to say that nothing happens to the state vectors; it is the operators X and P that get modified by $T(\varepsilon)$ as follows:

$$X \rightarrow T^\dagger(\varepsilon) X T(\varepsilon)$$

$$P \rightarrow T^\dagger(\varepsilon) P T(\varepsilon)$$

such that

$$T^\dagger(\varepsilon) X T(\varepsilon) = X + \varepsilon I \quad (11.2.5a)$$

$$T^\dagger(\varepsilon) P T(\varepsilon) = P \quad (11.2.5b)$$

This is called the passive transformation picture. Physically it corresponds to moving the environment (the coordinate system, sources of external field if any, etc.) to the *left* by ε .

Physically, the equivalence of the active and passive pictures is due to the fact that moving the particle one way is equivalent to moving the environment the other way by an equal amount.

Mathematically, we show the equivalence as follows. If we sandwich the operator equation (11.2.5) between $\langle \psi |$ and $|\psi \rangle$, we get Eq. (11.2.4). To go the other way, we first rewrite Eq. (11.2.4) as

$$\begin{aligned}\langle \psi | T^\dagger(\varepsilon) X T(\varepsilon) - X - \varepsilon I | \psi \rangle &= 0 \\ \langle \psi | T^\dagger(\varepsilon) P T(\varepsilon) - P | \psi \rangle &= 0\end{aligned}$$

We now reason as follows:

- (1) The operators being sandwiched are Hermitian (verify).
- (2) Since $|\psi\rangle$ is arbitrary, we can choose it to be any of the eigenvectors of these operators. It follows that all the eigenvalues vanish.
- (3) The operators themselves vanish, implying Eq. (11.2.5).

In what follows, we will examine both pictures. We will find that it is possible to construct $T(\varepsilon)$ given either of Eqs. (11.2.4) or (11.2.5), and of course that the two yield the same result. The active transformation picture is nice in that we work with the quantum state $|\psi\rangle$, which now plays the role of the classical state (x, p) . The passive transformation picture is nice because the response of the quantum operators X and P to a translation is formally similar to that of their classical counterparts.[‡]

We begin by discussing translations in terms of active transformations. Let us examine how the ket $|\psi_\varepsilon\rangle$ is related to $|\psi\rangle$ or, equivalently, the action of the Hilbert space operator $T(\varepsilon)$. The answer appears obvious if we work with kets of definite position, $|x\rangle$. In this case it is clear that

$$T(\varepsilon)|x\rangle = |x + \varepsilon\rangle \quad (11.2.6)$$

In other words, if the particle is originally at x , it must end up at $x + \varepsilon$. Notice that $T(\varepsilon)$ is unitary: it acts on an orthonormal basis $|x\rangle$, $-\infty \leq x \leq \infty$, and gives another, $|x + \varepsilon\rangle$, $-\infty \leq x + \varepsilon \leq \infty$. Once the action of $T(\varepsilon)$ on a complete basis is known, its action on any ket $|\psi\rangle$ follows:

$$\begin{aligned}|\psi_\varepsilon\rangle &= T(\varepsilon)|\psi\rangle = T(\varepsilon) \int_{-\infty}^{\infty} |x\rangle \langle x| \psi \rangle dx = \int_{-\infty}^{\infty} |x + \varepsilon\rangle \langle x| \psi \rangle dx \\ &= \int_{-\infty}^{\infty} |x'\rangle \langle x' - \varepsilon| \psi \rangle dx' \quad (x' = x + \varepsilon)\end{aligned} \quad (11.2.7)$$

In other words if

$$\langle x | \psi \rangle = \psi(x)$$

[‡] As we shall see, it is this point of view that best exposes many formal relations between classical and quantum mechanics.

then

$$\langle x | T(\varepsilon) | \psi \rangle = \psi(x - \varepsilon) \quad (11.2.8)$$

For example, if $\psi(x) \sim e^{-x^2}$ is a Gaussian peaked at the origin, $\psi(x - \varepsilon) \sim e^{-(x-\varepsilon)^2}$ is an identical Gaussian peaked at $x = \varepsilon$. Thus the wave function $\psi_\varepsilon(x)$ is obtained by translating (without distortion) the wave function $\psi(x)$ by an amount ε to the right. You may verify that the action of $T(\varepsilon)$ defined by Eq. (11.2.8) satisfies the condition Eq. (11.2.1a). How about the condition Eq. (11.2.1b)? It is *automatically* satisfied:

$$\begin{aligned} \langle \psi_\varepsilon | P | \psi_\varepsilon \rangle &= \int_{-\infty}^{\infty} \psi_\varepsilon^*(x) \left(-i\hbar \frac{d}{dx} \right) \psi_\varepsilon(x) dx \\ &= \int_{-\infty}^{\infty} \psi^*(x - \varepsilon) \left(-i\hbar \frac{d}{dx} \right) \psi(x - \varepsilon) dx \\ &= \int_{-\infty}^{\infty} \psi^*(x') \left(-i\hbar \frac{d}{dx'} \right) \psi(x') dx' \quad (x' = x - \varepsilon) \\ &= \langle \psi | P | \psi \rangle \end{aligned} \quad (11.2.9)$$

Now there is something odd here. Classically, translation is specified by two *independent* relations

$$x \rightarrow x + \varepsilon$$

$$p \rightarrow p$$

while in the quantum version we seem to find that in enforcing the former (on position eigenkets), the latter automatically follows. The reason is that in our derivation we have assumed more than what was explicitly stated. We reasoned earlier, on physical grounds, that since a particle initially located at x must end up at $x + \varepsilon$, it follows that

$$T(\varepsilon) |x\rangle = |x + \varepsilon\rangle$$

While our intuition was correct, our implementation was not. As seen in chapter 7, the X basis is not unique, and the general result consistent with our intuition is not Eq. (11.2.6) but rather

$$T(\varepsilon) |x\rangle = e^{i\epsilon g(x)/\hbar} |x + \varepsilon\rangle \quad (11.2.10)$$

(Note that as $\varepsilon \rightarrow 0$, $T(\varepsilon) |x\rangle \rightarrow |x\rangle$ as it should.) In ignoring $g(x)$, we had essentially assumed the quantum analog of $p \rightarrow p$. Let us see how. If we start with Eq. (11.2.10)

instead of Eq. (11.2.6), we find that

$$\langle X \rangle \xrightarrow{T(\varepsilon)} \langle X \rangle + \varepsilon \quad (11.2.11a)$$

$$\langle P \rangle \rightarrow \langle P \rangle + \varepsilon \langle f(X) \rangle \quad (11.2.11b)$$

where $f = g'$. Demanding now that $\langle P \rangle \rightarrow \langle P \rangle$, we eliminate f and reduce g to a harmless constant (which can be chosen to be 0).

Exercise 11.2.1. Verify Eq. (11.2.11b)

Note that there was nothing wrong with our initial choice $T|x\rangle = |x + \varepsilon\rangle$ —it was too restrictive given just the requirement $\langle X \rangle \rightarrow \langle X \rangle + \varepsilon$, but not so if we also considered $\langle P \rangle \rightarrow \langle P \rangle$. This situation reappears when we go to two or three dimensions and when we consider rotations. In all those cases we will make the analog of the naive choice $T(\varepsilon)|x\rangle = |x + \varepsilon\rangle$ to shorten the derivations.

Having defined translations, let us now define *translational invariance* in the same spirit. We define it by the requirement

$$\langle \psi | H | \psi \rangle = \langle \psi_e | H | \psi_e \rangle \quad (11.2.12)$$

To derive the conservation law that goes with the above equation, we must first construct the operator $T(\varepsilon)$ explicitly. Since $\varepsilon = 0$ corresponds to no translation, we may expand $T(\varepsilon)$ to order ε as

$$T(\varepsilon) = I - \frac{i\varepsilon}{\hbar} G \quad (11.2.13)$$

The operator G , called the *generator of translations*, is Hermitian (see Exercise 11.2.2 for the proof) and is to be determined. The constant $(-i/\hbar)$ is introduced in anticipation of what is to follow.

*Exercise 11.2.2.** Using $T^\dagger(\varepsilon)T(\varepsilon) = I$ to order ε , deduce that $G^\dagger = G$.

We find G by turning to Eq. (11.2.8):

$$\langle x | T(\varepsilon) | \psi \rangle = \psi(x - \varepsilon)$$

Expanding both sides to order ε , we find

$$\langle x | I | \psi \rangle - \frac{i\varepsilon}{\hbar} \langle x | G | \psi \rangle = \psi(x) - \frac{d\psi}{dx} \varepsilon$$

so that

$$\langle x|G|\psi\rangle = -i\hbar \frac{d\psi}{dx}$$

Clearly G is the momentum operator,

$$G = P$$

and

$$T(\varepsilon) = I - \frac{i\varepsilon}{\hbar} P \quad (11.2.14)$$

We see that exactly as in classical mechanics, the momentum is the generator of (infinitesimal) translations.

The momentum conservation law now follows from translational invariance, Eq. (11.2.12), if we combine it with Eq. (11.2.14):

$$\begin{aligned} \langle \psi | H | \psi \rangle &= \langle \psi_\varepsilon | H | \psi_\varepsilon \rangle \\ &= \langle T(\varepsilon)\psi | H | T(\varepsilon)\psi \rangle = \langle \psi | T^\dagger(\varepsilon)HT(\varepsilon)| \psi \rangle \\ &= \langle \psi | \left(I + \frac{i\varepsilon}{\hbar} P \right) H \left(I - \frac{i\varepsilon}{\hbar} P \right) | \psi \rangle \\ &= \langle \psi | H | \psi \rangle + \frac{i\varepsilon}{\hbar} \langle \psi | [P, H] | \psi \rangle + O(\varepsilon^2) \end{aligned}$$

so that, we get, upon equating the coefficient of ε to zero,

$$\langle \psi | [P, H] | \psi \rangle = 0 \quad (11.2.15)$$

It now follows from Ehrenfest's theorem that

$$\langle [P, H] \rangle = 0 \rightarrow \langle \dot{P} \rangle = 0 \quad (11.2.16)$$

Translation in Terms of Passive Transformations

Let us rederive $T(\varepsilon)$, given that it acts as follows on X and P :

$$T^\dagger(\varepsilon)XT(\varepsilon) = X + \varepsilon I \quad (11.2.17a)$$

$$T^\dagger(\varepsilon)PT(\varepsilon) = P \quad (11.2.17b)$$

The operator $T^\dagger(\varepsilon)XT(\varepsilon)$ is also a position operator, but it measures position from a new origin, shifted to the *left* by ε : This is the meaning of Eq. (11.2.17a).

Equation (11.2.17b) states that under the shift in the origin, the momentum is unaffected.

Writing once again

$$T(\varepsilon) = I - \frac{i\varepsilon G}{\hbar}$$

we find from Eq. (11.2.17a) (using the fact that $G^\dagger = G$)

$$\left(I + \frac{i\varepsilon G}{\hbar}\right)X\left(I - \frac{i\varepsilon G}{\hbar}\right) = X + \varepsilon I$$

or

$$-\frac{i\varepsilon}{\hbar} [X, G] = \varepsilon I \quad (11.2.18a)$$

$$[X, G] = i\hbar I \quad (11.2.18b)$$

This allows us to conclude that

$$G = P + f(X) \quad (11.2.19)$$

If we now turn to Eq. (11.2.17b) we find

$$-\frac{i\varepsilon}{\hbar} [P, G] = 0 \quad (11.2.20a)$$

or

$$[P, G] = 0 \quad (11.2.20b)$$

which eliminates $f(X)$.[‡] So once again

$$T(\varepsilon) = I - \frac{i\varepsilon P}{\hbar}$$

Having derived the translation operator in the passive transformation picture, let us reexamine the notion of translational invariance.

We define translational invariance by the requirement

$$T^\dagger(\varepsilon)HT(\varepsilon) = H \quad (11.2.21)$$

[‡] For the purists, it reduces f to a c number which commutes with X and P , which we choose to be zero.

We can rewrite Eq. (11.2.21) in a form that is closer to the classical definition of translational invariance. But first we need the following result: for any $\Omega(X, P)$ that can be expanded in a power series, and for any unitary operator U ,

$$U^\dagger \Omega(X, P) U = \Omega(U^\dagger X U, U^\dagger P U)$$

For the proof, consider a typical term in the series such as PX^2P . We have, using $UU^\dagger = I$,

$$U^\dagger PX^2PU = U^\dagger PUU^\dagger XUU^\dagger XUU^\dagger PU \quad \text{Q.E.D.}$$

Applying this result to the case $U = T(\varepsilon)$ we get the response of any dynamical variable to a translation:

$$\Omega(X, P) \rightarrow T^\dagger \Omega(X, P) T = \Omega(T^\dagger XT, T^\dagger PT) = \Omega(X + \varepsilon I, P) \quad (11.2.22)$$

Thus the transformed Ω is found by replacing X by $X + \varepsilon I$ and P by P . If we now apply this to Eq. (11.2.21) we get the following definition of translation invariance:

$$H(X + \varepsilon I, P) = H(X, P) \quad (11.2.23)$$

Not only does this condition have the same form as its classical counterpart

$$\mathcal{H}(x + \varepsilon, p) = \mathcal{H}(x, p)$$

but it is also satisfied whenever the classical counterpart is. The reason is simply that H is the same function of X and P as \mathcal{H} is of x and p , and both sets of variables undergo identical changes in a translation.

The conservation of momentum follows if we write $T(\varepsilon)$ in Eq. (11.2.21) in terms of P and expand things out to first order in ε :

$$\begin{aligned} 0 &= T^\dagger(\varepsilon) HT(\varepsilon) - H = (I + i\varepsilon P/\hbar) H (I - i\varepsilon P/\hbar) - H \\ &= \frac{-i\varepsilon}{\hbar} [H, P] \end{aligned} \quad (11.2.24)$$

which implies that $\langle \dot{P} \rangle = 0$, because of the Ehrenfest's theorem.

A Digression on the Analogy with Classical Mechanics[‡]

The passive transformation picture has the virtue that it bears a close formal resemblance to classical mechanics, with operators Ω in place of the classical variables

[‡] In a less advanced course, the reader may skip this digression.

ω [Eqs. (11.2.17), (11.2.22), (11.2.23)]. In fact, the infinitesimal unitary transformation $T(\varepsilon)$ generated by P is the quantum image of the infinitesimal canonical transformation generated by p : if we *define* the changes δX and δP by

$$\begin{aligned}\delta X &= T^\dagger(\varepsilon)XT(\varepsilon) - X \\ \delta P &= T^\dagger(\varepsilon)PT(\varepsilon) - P\end{aligned}$$

we get, on the one hand, from Eq. (11.2.17),

$$\begin{aligned}\delta X &= X + \varepsilon I - X = \varepsilon I \\ \delta P &= P - P = 0\end{aligned}$$

and on the other, from $T = I - i\varepsilon P/\hbar$ (working to first order in ε),

$$\begin{aligned}\delta X &= (I + i\varepsilon P/\hbar)X(I - i\varepsilon P/\hbar) - X = \frac{-i\varepsilon}{\hbar} [X, P] \\ \delta P &= (I + i\varepsilon P/\hbar)P(I - i\varepsilon P/\hbar) - P = \frac{-i\varepsilon}{\hbar} [P, P]\end{aligned}$$

combining which we obtain

$$\begin{aligned}\delta X &= \frac{-i\varepsilon}{\hbar} [X, P] = \varepsilon I \\ \delta P &= \frac{-i\varepsilon}{\hbar} [P, P] = 0\end{aligned}$$

More generally, upon combining, Eq. (11.2.22) and $T = I - i\varepsilon P/\hbar$, we obtain

$$\delta \Omega = \frac{-i\varepsilon}{\hbar} [\Omega, P] = \Omega(X + \varepsilon I, P) - \Omega(X, P)$$

These are the analogs of the canonical transformation generated by p :

$$\begin{aligned}\delta x &= \varepsilon \{x, p\} = \varepsilon \\ \delta p &= \varepsilon \{p, p\} = 0 \\ \delta \omega &= \varepsilon \{\omega, p\} = \omega(x + \varepsilon, p) - \omega(x, p)\end{aligned}$$

If the problem is translationally invariant, we have

$$\delta H = \frac{-i\varepsilon}{\hbar} [H, P] = 0 \rightarrow \langle \dot{P} \rangle = 0 \quad \text{by Ehrenfest's theorem}$$

while classically

$$\delta \mathcal{H} = \varepsilon \{ \mathcal{H}, p \} = 0 \rightarrow \dot{p} = 0 \quad \text{by } \dot{p} = \{ p, \mathcal{H} \}$$

The correspondence is achieved through the substitution rules already familiar to us:

$$\Omega \leftrightarrow \omega$$

$$\frac{-i}{\hbar} [\Omega, \Lambda] \leftrightarrow \{\omega, \lambda\}$$

In general, the infinitesimal canonical transformation generated by $g(x, p)$,

$$\delta \omega = \varepsilon \{ \omega, g \}$$

has as its image in quantum theory the infinitesimal unitary transformation $U_G(\varepsilon) = I - i\varepsilon G/\hbar$ in response to which

$$\delta \Omega = \frac{-i\varepsilon}{\hbar} [\Omega, G]$$

Now, we have seen that the transformation generated by any $g(x, p)$ is canonical, i.e., it preserves the PB between the x 's and the p 's. In the quantum theory, the quantities preserved are the commutation relations between the X 's and the P 's, for if

$$[X_i, P_j] = i\hbar \delta_{ij} I$$

then upon premultiplying by the unitary operator $U_G^\dagger(\varepsilon)$ and postmultiplying by $U_G(\varepsilon)$, we find that the transformed operators obey‡

$$[U^\dagger X_i U, U^\dagger P_j U] = i\hbar \delta_{ij} I$$

This completes the proof of the correspondence

$$\left\{ \begin{array}{l} \text{infinitesimal canonical} \\ \text{transformation generated} \\ \text{by } g(x, p) \end{array} \right\} \leftrightarrow \left\{ \begin{array}{l} \text{infinitesimal unitary} \\ \text{transformation generated} \\ \text{by } G(X, P) \end{array} \right\}$$

The correspondence holds for finite transformations as well, for these may be viewed as a sequence of infinitesimal transformations.

‡ More generally if $[\Omega, \theta] = \Gamma$, then a similar relation holds between the transformed operators $U^\dagger \Omega U$, $U^\dagger \theta U$, $U^\dagger \Gamma U$. This is the quantum version of the result that PB are invariant under canonical transformation.

The correspondence with unitary transformations also holds for *regular* canonical transformations which have no infinitesimal versions. For instance, in the coupled oscillator problem, Exercise 10.1.3, we performed a canonical transformation from x_1, x_2, p_1, p_2 to x_I, x_{II}, p_I , and p_{II} , where, for example, $x_I = (x_1 + x_2)/2$. In the quantum theory there will exist a unitary operator such that, for example, $U^\dagger X_I U = (X_1 + X_2)/2 = X_I$ and so on.[‡]

We can see why we can either perform the canonical transformation at the classical level and then quantize, or first quantize and then perform the unitary transformation—since the quantum operators respond to the unitary transformation as do their classical counterparts to the canonical transformation, the end effect will be the same.[§]

Let us now return to the problem of translational invariance. Notice that in a problem with translational invariance, Eq. (11.2.24) tells us that we can find the simultaneous eigenbasis of P and H . (This agrees with our result from Chapter 5, that the energy eigenstates of a free particle could be chosen to be momentum eigenstates as well.)^{||} If a system starts out in such an eigenstate, its momentum eigenvalue remains constant. To prove this, first note that

$$[P, H] = 0 \rightarrow [P, U(t)] = 0 \quad (11.2.25)$$

since the propagator is a function of just H .*

Suppose at $t=0$ we have a system in an eigenstate of P :

$$P|p\rangle = p|p\rangle \quad (11.2.26)$$

After time t , the state is $U(t)|p\rangle$ and we find

$$P U(t)|p\rangle = U(t)P|p\rangle = U(t)p|p\rangle = p U(t)|p\rangle \quad (11.2.27)$$

In other words, the state at time t is also an eigenstate of P with the same eigenvalue. For such states with well-defined momentum, the conservation law $\langle \dot{P} \rangle = 0$ reduces to the classical form $\dot{p} = 0$.

Finite Translations

What is the operator $T(a)$ corresponding to a finite translation a ? We find it by the following trick. We divide the interval a into N parts of size a/N . As $N \rightarrow \infty$,

[‡] If the transformation is not regular, we cannot find a unitary transformation in the quantum theory, since unitary transformations preserve the eigenvalue spectrum.

[§] End of digression.

^{||} Note that a single particle whose H is translationally invariant is necessarily free.

^{*} When H is time independent, we know $U(t) = \exp(-iHt/\hbar)$. If $H = H(t)$, the result is true if P commutes with $H(t)$ for all t . (Why?)

a/N becomes infinitesimal and we know

$$T(a/N) = I - \frac{ia}{\hbar N} P \quad (11.2.28)$$

Since a translation by a equals N translations by a/N ,

$$T(a) = \lim_{N \rightarrow \infty} [T(a/N)]^N = e^{-iaP/\hbar} \quad (11.2.29)$$

by virtue of the formula

$$e^{-ax} = \lim_{N \rightarrow \infty} \left(1 - \frac{ax}{N}\right)^N$$

We may apply this formula, true for c numbers, to the present problem, since P is the only operator in the picture and commutes with everything in sight, i.e., behaves like a c number. Since

$$T(a) \xrightarrow[X \text{ basis}]{} e^{-ad/dx} \quad (11.2.30)$$

we find

$$\langle x | T(a) | \psi \rangle = \psi(x) - \frac{d\psi}{dx} a + \frac{d^2\psi}{dx^2} \frac{a^2}{2!} + \dots \quad (11.2.31)$$

which is the full Taylor series for $\psi(x-a)$ about the point x .

A Consistency Check. A translation by a followed by a translation by b equals a translation by $a+b$. This result has nothing to do with quantum mechanics and is true whether you are talking about a quantum system or a sack of potatoes. It is merely a statement about how translations combine in space. Now, we have just built operators T , which are supposed to translate quantum states. For this interpretation to be consistent, it is necessary that the law of combination of the translation operators coincide with the law of combination of the translations they represent. Now, although we presumed this [see Eq. (11.2.29), and the line above it] in the very act of deriving the formula for $T(a)$, let us verify that our result $T(a) = \exp(-iaP/\hbar)$ satisfies

$$T(a)T(b) = T(a+b)? \quad (11.2.32)$$

We find that this is indeed so:

$$T(a)T(b) = e^{-iaP/\hbar} \cdot e^{-ibP/\hbar} = e^{-i(a+b)P/\hbar} = T(a+b) \quad (11.2.33)$$

A Digression on Finite Canonical and Unitary Transformations‡

Though it is clear that the correspondence between canonical and unitary transformations, established for the infinitesimal case in earlier discussions, must carry

‡ Optional.

over to the finite case, let us nonetheless go through the details. Consider, for definiteness, the case of translations. In the quantum theory we have

$$\Omega \rightarrow T^\dagger(a)\Omega T(a) = e^{iaP/\hbar}\Omega e^{-iaP/\hbar}$$

Using the identity

$$e^{-A}B e^{+A} = B + [B, A] + \frac{1}{2!} [[B, A], A] + \frac{1}{3!} \dots$$

we find

$$\Omega \rightarrow \Omega + a\left(\frac{-i}{\hbar}\right)[\Omega, P] + \frac{1}{2!} a^2\left(\frac{-i}{\hbar}\right)^2 [[\Omega, P], P] + \dots \quad (11.2.34)$$

For example, if we set $\Omega = X^2$ we get $X^2 \rightarrow (X + aI)^2$.

In the classical case, under an infinitesimal displacement δa ,

$$\delta\omega = \delta a \{\omega, p\}$$

or

$$\frac{d\omega}{da} = \{\omega, p\}$$

Applying the above result to the variable $d\omega/da$, we get

$$\frac{d}{da}(d\omega/da) = d^2\omega/da^2 = \{d\omega/da, p\} = \{\{\omega, p\}, p\}$$

and so on. The response of ω to the finite translation is given by the Taylor series about the point $a=0$:

$$\omega \rightarrow \omega + a\{\omega, p\} + \frac{a^2}{2!} \{\{\omega, p\}, p\} + \dots \quad (11.2.35)$$

which we see is in correspondence with Eq. (11.2.34) if we make the usual substitutions.

*Exercise 11.2.3.** Recall that we found the finite rotation transformation from the infinitesimal one, by solving differential equations (Section 2.8). Verify that if, instead, you relate the transformed coordinates \bar{x} and \bar{y} to x and y by the infinite string of Poisson brackets, you get the same result, $\bar{x} = x \cos \theta - y \sin \theta$, etc. (Recall the series for $\sin \theta$, etc.)

System of Particles

We will not belabor the extension of the preceding ideas to a system of N particles. Starting with the analog of Eq. (11.2.8),

$$\langle x_1, \dots, x_N | T(\varepsilon) | \psi \rangle = \psi(x_1 - \varepsilon, \dots, x_N - \varepsilon) \quad (11.2.36)$$

we find, on expanding both sides to order ε , that

$$\langle x_1, \dots, x_N | I - \frac{i\varepsilon}{\hbar} P | \psi \rangle = \psi(x_1, \dots, x_N) - \sum_{i=1}^N \varepsilon \frac{\partial \psi}{\partial x_i} \quad (11.2.37)$$

from which it follows that

$$T(\varepsilon) = I - \frac{i\varepsilon}{\hbar} \sum_{i=1}^N P_i = I - \frac{i\varepsilon}{\hbar} P \quad (11.2.38)$$

where P is the *total momentum operator*. You may verify that

$$\begin{aligned} T^\dagger(\varepsilon) X_i T(\varepsilon) &= X_i + \varepsilon I \\ T^\dagger(\varepsilon) P_i T(\varepsilon) &= P_i, \quad i = 1, \dots, N \end{aligned} \quad (11.2.39)$$

Translational invariance means in this case (suppressing indices),

$$H(X, P) = T^\dagger(\varepsilon) H(X, P) T(\varepsilon) = H(X + \varepsilon I, P) \quad (11.2.40)$$

Whereas in the single-particle cases this implied the particle was free, here it merely requires that H (or rather V) be a function of the coordinate *differences*. Any system whose parts interact with each other, but nothing external, will have this property.

There are some profound consequences of translational invariance besides momentum conservation. We take these up next.

Implications of Translational Invariances

Consider a system with translational invariance. Premultiplying both sides of Eq. (11.2.21) with T and using its unitarity, we get

$$[T(a), H] = 0$$

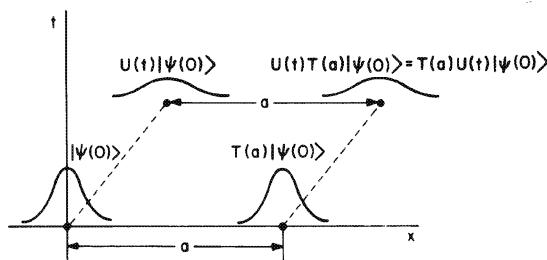


Figure 11.1. A symbolic depiction of translational invariance. The states are represented schematically by wave functions.

It follows that

$$[T(a), U(t)] = 0 \quad \text{or} \quad T(a)U(t) = U(t)T(a) \quad (11.2.41)$$

The consequence of this relation is illustrated by the following example (Fig. 11.1). At $t=0$ two observers *A* and *B* prepare identical systems at $x=0$ and $x=a$, respectively. If $|\psi(0)\rangle$ is the state vector of the system prepared by *A*, then $T(a)|\psi(0)\rangle$ is the state vector of the system prepared by *B*. The two systems look identical to the observers who prepared them. After time t , the state vectors evolve into $U(t)|\psi(0)\rangle$ and $U(t)T(a)|\psi(0)\rangle$. Using Eq. (11.2.41) the latter may be rewritten as $T(a)U(t)|\psi(0)\rangle$, which is just the translated version of *A*'s system at time t . Therefore the two systems, which differed only by a translation at $t=0$, differ only by the same translation at future times. In other words, the time evolution of each system appears the same to the observer who prepared it. Translational invariance of H implies that the same experiment repeated at two different places will give the same result (as seen by the local observers). We have already seen this result in the classical framework. We pursue it further now.

Now it turns out that every known interaction—gravitational, weak, electromagnetic, and strong (e.g., nuclear)—is translationally invariant, in that every experiment, if repeated at a new site, will give the same result. Consider the following illustrative example, which clarifies the meaning of this remark. A hydrogen atom is placed between the plates of a charged condenser. The Hamiltonian is

$$H = \frac{|\mathbf{P}_1|^2}{2m_1} + \frac{|\mathbf{P}_2|^2}{2m_2} + \frac{e_1 e_2}{|\mathbf{R}_1 - \mathbf{R}_2|} + e_1 V(\mathbf{R}_1) + e_2 V(\mathbf{R}_2) \quad (11.2.42)$$

where the subscripts 1 and 2 refer to the electron and the proton and $V(\mathbf{R})$ [‡] to the potential due to the plates. Now this problem has no translation invariance, i.e.,

$$H(\mathbf{R}_1 + \boldsymbol{\varepsilon}, \mathbf{P}_1; \mathbf{R}_2 + \boldsymbol{\varepsilon}, \mathbf{P}_2) \neq H(\mathbf{R}_1, \mathbf{P}_1; \mathbf{R}_2, \mathbf{P}_2)$$

which in turn means that if the atom alone is translated (away from the condenser) it will behave differently. But this does not correspond to repeating the *same* experiment and getting a different result, since the condenser, which materially affects the

[‡] Remember that \mathbf{R} is the operator corresponding to the classical variable \mathbf{r} .

dynamics, is left behind. To incorporate it in what is translated, we redefine our system to include the (N) charges on the condenser and write

$$H = \sum_{i=1}^{N+2} \frac{|\mathbf{P}_i|^2}{2m_i} + \frac{1}{2} \sum_{i=1}^{N+2} \sum_{j \neq i}^{N+2} \frac{e_i e_j}{|\mathbf{R}_i - \mathbf{R}_j|} \quad (11.2.43)$$

Now the charges on the condenser enter H , not via the external field which breaks translational invariance, but through the Coulomb interaction, which does not. Now it is true that (dropping indices),

$$H(\mathbf{R} + \boldsymbol{\varepsilon}, \mathbf{P}) = H(\mathbf{R}, \mathbf{P})$$

which implies that if the atom *and* the condenser are moved to a new site, the behavior of the composite system will be unaffected. This result should be viewed not as obvious or self-evident, but rather as a profound statement about the Coulomb interaction.

The content of the assertion made above is that every known interaction has translational invariance at the fundamental level—if we expand our system to include all degrees of freedom that affect the outcome of an experiment (so that there are not external fields, only interactions between parts of the system) the total H is translationally invariant. This is why we apply momentum conservation to every problem whatever be the underlying interaction. The translational invariance of natural laws reflects the uniformity or homogeneity of space. The fact that the dynamics of an isolated[‡] system (the condenser plus atom in our example) depends only on where the parts of the system are relative to each other and not on where the system is as a whole, represents the fact that one part of free space is as good as another.

It is translational invariance that allows experimentalists in different parts of the earth to claim they all did the “same” experiment, and to corroborate, correct, and complement each other. It is the invariance of the natural laws under translations that allows us to describe a hydrogen atom in some distant star as we do one on earth and to apply to its dynamics the quantum mechanical laws deduced on earth. We will examine further consequences of translational invariance toward the end of the next section.

11.3. Time Translational Invariance

Just as the homogeneity of space ensures that the same experiment performed at two different places gives the same result, homogeneity in time ensures that the

[‡] To be exact, no system is truly “isolated” except the whole universe (and only its momentum is exactly conserved). But in practice one draws a line somewhere, between what constitutes the system and what is irrelevant (for practical purposes) to its evolution. I use the term “isolated” in this practical sense. The real utility of the concepts of translational invariance and momentum conservation lies in these approximate situations. Who cares if the universe as a whole is translationally invariant and its momentum is conserved? What matters to me is that I can take my equipment to another town and get the same results and that the momentum of my system is conserved (to a good accuracy).

same experiment repeated at two different times gives the same result. Let us see what feature of the Hamiltonian ensures this and what conservation law follows.

Let us prepare at time t_1 a system in state $|\psi_0\rangle$ and let it evolve for an infinitesimal time ε . The state at time $t_1 + \varepsilon$, to first order in ε , will be

$$|\psi(t_1 + \varepsilon)\rangle = \left[I - \frac{i\varepsilon}{\hbar} H(t_1) \right] |\psi_0\rangle \quad (11.3.1)$$

If we repeat the experiment at time t_2 , beginning with the same initial state, the state at time $t_2 + \varepsilon$ will be

$$|\psi(t_2 + \varepsilon)\rangle = \left[I - \frac{i\varepsilon}{\hbar} H(t_2) \right] |\psi_0\rangle \quad (11.3.2)$$

The outcome will be the same in both cases if

$$\begin{aligned} 0 &= |\psi(t_2 + \varepsilon)\rangle - |\psi(t_1 + \varepsilon)\rangle \\ &= \left(-\frac{i\varepsilon}{\hbar} \right) [H(t_2) - H(t_1)] |\psi_0\rangle \end{aligned} \quad (11.3.3)$$

Since $|\psi_0\rangle$ is arbitrary, it follows that

$$H(t_2) = H(t_1) \quad (11.3.4)$$

Since t_2 and t_1 are arbitrary, it follows that H is time-independent:

$$\frac{dH}{dt} = 0 \quad (11.3.5)$$

Thus time translational invariance requires that H have no t dependence. Now Ehrenfest's theorem for an operator Ω that has no time dependence[‡] is

$$i\hbar \langle \dot{\Omega} \rangle = \langle [\Omega, H] \rangle$$

Applying it to $\Omega = H$ in a problem with time translational invariance, we find

$$\langle \dot{H} \rangle = 0 \quad (11.3.6)$$

which is the law of conservation of energy.

[‡] If $d\Omega/dt \neq 0$ there will be an extra piece $i\hbar \langle d\Omega/dt \rangle$ on the right-hand side.

An important simplification that arises if $dH/dt=0$ is one we have repeatedly exploited in the past: Schrödinger's equation

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = H|\psi\rangle \quad (11.3.7)$$

admits solutions of the form

$$|\psi(t)\rangle = |E\rangle e^{-iEt/\hbar} \quad (11.3.8)$$

where the time-independent ket $|E\rangle$ satisfies

$$H|E\rangle = E|E\rangle \quad (11.3.9)$$

The entire dynamics, i.e., the determination of the propagator $U(t)$, boils down to the solution of the time-independent Schrödinger equation (11.3.9).

The considerations that applied to space translation invariance apply here as well. In particular, all known interactions—from gravitational to strong—are time translational invariant. Consequently, if we suitably define the system (to include the sources of external fields that affect the experiment) the total H will be independent of t . Consider, for example, a hydrogen atom between the plates of a *discharging* condenser. If the system includes just the electron and the proton, H will depend on time—it will have the form of Eq. (11.2.42), with $V=V(\mathbf{R}, t)$. This simply means that repeating the experiment *without recharging the condenser*, will lead to a different result. If, however, we enlarge the system to include the N charges on the condenser, we end up with the H in Eq. (11.2.43), which has no t dependence.

The space-time invariance of natural laws has a profound impact on our quest for understanding nature. The very cycle of physics—of deducing laws from some phenomena studied at some time and place and then applying them to other phenomena at a different time and place—rests on the assumption that natural laws are space-time invariant. If nature were not to follow the same rules over space-time, there would be no rules to find, just a sequence of haphazard events with no rhyme or reason. By repeating the natural laws over and over through all of space-time, nature gives tiny earthlings, who probe just a minuscule region of space for a fleeting moment (in the cosmic scale), a chance of comprehending the universe at large. Should we at times be despondent over the fact that we know so few of nature's laws, let us find solace in these symmetry principles, which tell us that what little we know is universal and eternal.[‡]

[‡] The invariance of the laws of nature is not to be confused with our awareness of them, which does not change with time. For example, Einstein showed that Newtonian mechanics and gravitation are approximations to relativistic mechanics and gravitation. But this is not to say that the Newtonian scheme worked till Einstein came along. In other words, the relation of Newton's scheme to Einstein's (as a good approximation in a certain limit) has always been the same, before and after we learned of it.

11.4. Parity Invariance

Unlike space-time translations, and rotations, (which we will study in the next chapter), parity is a discrete transformation. Classically, the parity operation corresponds to reflecting the state of the particle through the origin

$$\begin{array}{ccc} x & \xrightarrow{\text{parity}} & -x \\ p & \xrightarrow{\text{parity}} & -p \end{array} \quad (11.4.1)$$

In quantum theory, we define the action of the parity operator on the X basis as follows

$$\Pi|x\rangle = |-x\rangle \quad (11.4.2)$$

in analogy with the classical case. Given this,

$$\Pi|p\rangle = |-p\rangle \quad (11.4.3)$$

follows, as you will see in a moment.

Given the action of Π on a complete (X) basis, its action on an arbitrary ket follows:

$$\begin{aligned} \Pi|\psi\rangle &= \Pi \int_{-\infty}^{\infty} |x\rangle\langle x| \psi \rangle dx \\ &= \int_{-\infty}^{\infty} |-x\rangle\langle x| \psi \rangle dx \\ &= \int_{-\infty}^{\infty} |x'\rangle\langle -x'| \psi \rangle dx' \quad (\text{where } x' = -x) \end{aligned} \quad (11.4.4)$$

It follows that if

$$\begin{aligned} \langle x| \psi \rangle &= \psi(x) \\ \langle x| \Pi| \psi \rangle &= \psi(-x) \end{aligned} \quad (11.4.5)$$

The function $\psi(-x)$ is the mirror image of $\psi(x)$ about the origin. Applying Eq. (11.4.5) to a momentum eigenstate, it will be readily found that $\Pi|p\rangle = |-p\rangle$.

The eigenvalues of Π are just ± 1 . A moment's "reflection" will prove this. Since

$$\Pi|x\rangle = |-x\rangle$$

$$\Pi^2|x\rangle = |-(-x)\rangle = |x\rangle$$

Since this is true for an entire basis,

$$\Pi^2 = I \quad (11.4.6)$$

Please note that

- (1) $\Pi = \Pi^{-1}$
- (2) The eigenvalues of Π are ± 1 .
- (3) Π is Hermitian and unitary.
- (4) Or $\Pi^{-1} = \Pi^\dagger = \Pi$.

The eigenvectors with eigenvalue ± 1 are said to have even/odd parity. In the X basis, where

$$\psi(x) \xrightarrow{\Pi} \psi(-x)$$

even-parity vectors have even wave functions and odd-parity vectors have odd wave functions. The same goes for the P basis since

$$\psi(p) \xrightarrow{\Pi} \psi(-p)$$

In an arbitrary Ω basis, $\psi(\omega)$ need not be even or odd even if $|\psi\rangle$ is a parity eigenstate (check this).

Rather than define Π in terms of its action on the kets, we may also define it through its action on the operators:

$$\begin{aligned} \Pi^\dagger X \Pi &= -X \\ \Pi^\dagger P \Pi &= -P \end{aligned} \quad (11.4.7)$$

We say $H(X, P)$ is parity invariant if

$$\Pi^\dagger H(X, P) \Pi = H(-X, -P) = H(X, P) \quad (11.4.8)$$

In this case

$$[\Pi, H] = 0$$

and a common eigenbasis of Π and H can be found. In particular, if we consider just bound states in one dimension (which we saw are nondegenerate), every eigenvector of H is necessarily an eigenvector of Π . For example, the oscillator Hamiltonian

satisfies Eq. (11.4.8) and its eigenfunctions have definite parity equal to $(-1)^n$, n being the quantum number of the state. The particle in a box has a parity-invariant Hamiltonian if the box extends from $-L/2$ to $L/2$. In this case the eigenfunctions have parity $(-1)^{n+1}$, n being the quantum number. If the box extends from 0 to L , $V(x)$ is not parity invariant and the eigenfunctions

$$\psi_n(x) = \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{n\pi x}{L}\right)$$

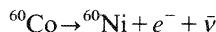
have no definite parity. (When $x \rightarrow -x$ they vanish, since ψ_n is given by the sine function *only between 0 and L, and vanishes outside.*)

If H is parity invariant, then

$$\Pi U(t) = U(t)\Pi \quad (11.4.9)$$

This means that if at $t=0$ I start with a system in a state $|\psi(0)\rangle$, and someone else starts with a system in the parity operated state $\Pi|\psi(0)\rangle$, then at a later time the state of his system will be related to mine by the parity transformation.

Whereas all natural laws are invariant under space-time translations (and rotations) some are not invariant under parity. These are the laws of weak interactions, which are responsible for nuclear β decay (among other things). This means formally that the Hamiltonian cannot be made parity invariant by any redefinition of the system if weak interactions are involved. Physically this means that if two observers prepare initial states $|\psi(0)\rangle$ and $\Pi|\psi(0)\rangle$ which are mirror images of each other, the final states $U(t)|\psi(0)\rangle$ and $U(t)\Pi|\psi(0)\rangle$ will not be mirror images of each other (since $\Pi U \neq U\Pi$).[‡] Consider the following concrete example of a β decay:



where e^- is an electron and $\bar{\nu}$ is an antineutrino. Now it turns out that the electron likes to come flying out in a direction opposite to the spin of ^{60}Co —and this implies parity noninvariance. Let us see how. At $t=0$ I prepare a system that consists of a ^{60}Co nucleus with its spin up along the z axis (Fig. 11.2) (experiment *A*). Although you are not yet familiar with spin, you may pretend here that ^{60}Co is spinning in the sense shown. Let another observer set up another system which is just the mirror image of mine (experiment *B*). Let M denote a *fictitious* experiment, which is what I see in a mirror in front of me. Notice how the spin \mathbf{S} gets reversed under a mirror reflection. Let the β decay take place. My electron comes out *down* the z axis. Of course the mirror also shows an electron coming down the z axis. In the other *real* experiment (*B*), the dynamics forces the electron to come *up* the z axis, since the initial \mathbf{S} was *down*. Thus *B* starts out as the mirror image of *A* but ends up different. Consequently, what I see in the mirror (experiment *M*) does not correspond to what can happen in real life, i.e., is not a solution to the equations of motion.

[‡] See Exercise 11.4.4 for a discussion of why the parity transformation is essentially a mirror reflection in three dimensions.

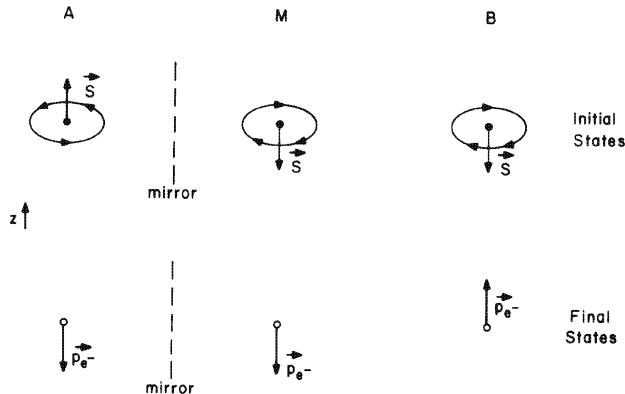


Figure 11.2. An example of parity noninvariance. In experiment *A*, which I perform, the spin of the nucleus points up the *z* axis. In its actual mirror image, it points down (experiment *M*). In experiment *B*, which is a real experiment, the spin is chosen to be down, i.e., *B* starts out as the mirror image of *A*. After the decay, the momentum of my electron, p_e^- , is down the *z* axis. The mirror image of course also shows the electron coming down. But in the actual experiment *B*, the dynamics forces the electron to come up the *z* axis, i.e., antiparallel to the initial nuclear spin \mathbf{S} .

This then is the big difference between parity and other transformations such as space-time translations and rotations. If a certain phenomenon can happen, its translated or rotated version can also happen, but not its mirror-reflected version, if the phenomenon involves weak interactions. In terms of conservation laws, if an isolated system starts out in a state of definite parity, it need not end in a state of same parity if weak interactions are at work. The possibility that weak interactions could be parity noninvariant was discussed in detail by Lee and Yang in 1956 and confirmed shortly thereafter by the experiments of C. S. Wu and collaborators.[‡]

*Exercise 11.4.1.** Prove that if $[\Pi, H] = 0$, a system that starts out in a state of even/odd parity maintains its parity. (Note that since parity is a discrete operation, it has no associated conservation law in classical mechanics.)

*Exercise 11.4.2.** A particle is in a potential

$$V(x) = V_0 \sin(2\pi x/a)$$

which is invariant under the translations $x \rightarrow x + ma$, where m is an integer. Is momentum conserved? Why not?

*Exercise 11.4.3.** You are told that in a certain reaction, the electron comes out with its spin always parallel to its momentum. Argue that parity is violated.

*Exercise 11.4.4.** We have treated parity as a mirror reflection. This is certainly true in one dimension, where $x \rightarrow -x$ may be viewed as the effect of reflecting through a (point) mirror at the origin. In higher dimensions when we use a plane mirror (say lying on the $x-y$

[‡] T. D. Lee and C. N. Yang, *Phys. Rev.*, **104**, 254 (1956); C. S. Wu, E. Ambler, R. W. Hayward, and R. P. Hudson, *Phys. Rev.*, **105**, 1413 (1957).

plane), only one (z) coordinate gets reversed, whereas the parity transformation reverses all three coordinates.

Verify that reflection on a mirror in the $x-y$ plane is the same as parity followed by 180° rotation about the z axis. Since rotational invariance holds for weak interactions, noninvariance under mirror reflection implies noninvariance under parity.

11.5. Time-Reversal Symmetry

This is a discrete symmetry like parity. Let us first understand what it means in classical physics. Consider a planet that is on a circular orbit around the sun. At $t=0$ it starts at $\theta=0$ and has a velocity in the direction of increasing θ . In other words, the orbit is running counterclockwise. Let us call the initial position and momentum $x(0)$, $p(0)$. (We should really be using vectors, but ignore this fact for this discussion.)

We now define the *time-reversed state* as one in which the position is the same but the momentum is reversed:

$$x_r(t) = x(t) \quad p_r(t) = -p(t).$$

In general, any quantity like position or kinetic energy, which involves an even power of t in its definition is left invariant and any quantity like momentum or angular momentum is reversed in sign under the time-reversal operation.

Say that after time T the planet has come to a final state $x(T)$, $p(T)$ at $\theta=\pi/2$ after doing a quarter of a revolution. Now Superman (for reasons best known to him) stops it dead in its tracks, reverses its speed, and lets it go. What will it do? We know it will retrace its path and at time $2T$ end up in the time-reversed state of the initial state:

$$x(2T) = x(0) \quad p(2T) = -p(0) \quad (11.5.1)$$

The above equation defines time-reversal invariance (TRI).

We can describe TRI more graphically as follows. Suppose we take a movie of the planet from $t=0$ to $t=T$. At $t=T$, we start playing the film backward. The backward motion of the planet will bring it back to the time-reversal initial state at $t=2T$. What we see in the movie can really happen, indeed, it was shown how Superman could make it happen even as you are watching the movie. More generally, if you see a movie of some planetary motion you will have no way of knowing if the projector is running forwards or backward. In some movies they get a big laugh out of the audience by showing cars and people zooming in reverse. As a serious physics student you should not laugh when you see this since these motions obey Newton's laws. In other words, it is perfectly possible for a set of people and cars to execute this motion. On the other hand, when a cartoon character falling under gravity suddenly starts clawing his way upwards in thin air using sheer will power, you may laugh since this is a gross violation of Newton's laws.

While the correctness of Eq.(11.5.1) is intuitively clear, we will now prove it with the help of Newton's Second Law using the fact that it is invariant under $t \rightarrow -t$: the acceleration is even in time and the potential or force has no reference to t . Here are the details. Just for this discussion let us use a new clock that has its zero at the

point of time-reversal, so that $t=0$ defines the point when the motion is time-reversed. When the movie is run backward we see the trajectory

$$x_r(t) = x(-t)$$

In other words, 5 seconds after the reversal, the object is where it was 5 seconds before the reversal. The reversal of velocities follows from this:

$$\dot{x}_r(t) = \frac{dx(-t)}{dt} = -\frac{dx(-t)}{d(-t)} = -\dot{x}(-t)$$

and does not have to be additionally encoded. The question is this: Does this orbit $x_r(t)$ obey Newton's Second Law

$$m \frac{d^2 x_r(t)}{dt^2} = F(x_r)$$

given that $x(t)$ does? We find it does:

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

Not all problems are time-reversal invariant. Consider a positively charged particle in the x - y plane moving under a magnetic field *down* the z -axis. Let us say it is released at $t=0$ just like the planet, with its velocity in the direction of increasing θ . Due to the $\mathbf{v} \times \mathbf{B}$ force it will go in a counterclockwise orbit. Let us wait till it has gone around by $\pi/2$ and at this time, $t=T$, time-reverse its state. Will it return to the time-reversed initial state at $t=2T$? No, it is readily seen that starting from $t=T$ it will once again go on a counterclockwise circular orbit tangential to the first at the point of reversal. We blame the magnetic interaction for this failure of TRI: the force now involves the velocity which is odd under time-reversal.

We now ask how all this appears in quantum mechanics. The ideas will be illustrated in the simplest context. Let us consider a particle in one dimension with a time-independent Hamiltonian H . In the x -representation the wave equation is

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = H(x)\psi(x, t)$$

Let us first note that

$$\psi \rightarrow \psi^*$$

performs time-reversal. This is clear from the fact that the detailed probability distribution in x is unaffected by this change. On the other hand, it is clear from looking at plane waves (or the momentum operator $-i\hbar(\hat{e}/\partial x)$) that $p \rightarrow -p$ under complex conjugation.

If the system has TRI, we must find the analog of Eq. (11.5.1). So let us prepare a state $\psi(x, 0)$, let it evolve for time T , complex conjugate it, let *that* evolve for another time T and see if we end up with the complex conjugate of the initial state. We find the following happens at each stage:

$$\psi(x, 0) \rightarrow e^{-iH(x)T/\hbar} \psi(x, 0) \rightarrow e^{iH^*(x)T/\hbar} \psi^*(x, 0) \rightarrow e^{-iH(x)T/\hbar} e^{iH^*(x)T/\hbar} \psi^*(x, 0)$$

It is clear that in order for the end result, which is $\psi(x, 2T)$, to obey

$$\psi(x, 2T) = \psi^*(x, 0)$$

we require that

$$H(x) = H^*(x) \quad (11.5.2)$$

i.e., that the Hamiltonian be real. For $H = P^2/2m + V(x)$ this is the case, even in higher dimensions. On the other hand, if we have a magnetic field, P enters linearly and $H(x) \neq H^*(x)$.

If H has TRI, i.e., is real, we have seen at the end of Chapter 6 that every eigenfunction implies a degenerate one which is its complex conjugate.

Notice that the failure of TRI in the presence of a magnetic field does not represent any *fundamental* asymmetry under time-reversal in electrodynamics. The laws of electrodynamics are invariant under $t \rightarrow -t$. The asymmetry in our example arose due to our treating the magnetic field as external to the system and hence not to be time-reversed. If we had included in our system the currents producing the magnetic field, and reversed them also, the entire system would have followed the time-reversed trajectory. Indeed, if you had taken a movie of the experiment and played it back, and you could have seen the charges in the wire, you would have found them running backward, the field would have been reversed at $t = T$, and the charge we chose to focus on would have followed the time-reversed trajectory.

On the other hand, certain experiments together with general arguments from quantum field theory suggest that there exist interactions in this universe which do not have this symmetry at the fundamental level.

There are ways to formulate TRI in a basis-independent way but we will not do so here. For most problems where the coordinate basis is the natural choice the above discussion will do. There will be a minor twist when the problem involves spin which has no classical counterpart. This can be handled by treating spin as we would treat orbital angular momentum.

12

Rotational Invariance and Angular Momentum

In the last chapter on symmetries, rotational invariance was not discussed, not because it is unimportant, but because it is all too important and deserves a chapter on its own. The reason is that most of the problems we discuss involve a single particle (which may be the reduced mass) in an external potential, and whereas translational invariance of H implies that the particle is free, rotational invariance of H leaves enough room for interesting dynamics. We first consider two dimensions and then move on to three.

12.1. Translations in Two Dimensions

Although we are concerned mainly with rotations, let us quickly review translations in two dimensions. By a straightforward extension of the arguments that led to Eq. (11.2.14) from Eq. (11.2.13), we may deduce that the generators of infinitesimal translations along the x and y directions are, respectively,

$$P_x \xrightarrow[\text{coordinate basis}]{} -i\hbar \frac{\partial}{\partial x} \quad (12.1.1)$$

$$P_y \xrightarrow[\text{coordinate basis}]{} -i\hbar \frac{\partial}{\partial y} \quad (12.1.2)$$

In terms of the *vector operator* \mathbf{P} , which represents momentum,

$$\mathbf{P} = P_x \mathbf{i} + P_y \mathbf{j} \quad (12.1.3)$$

P_x and P_y are the dot products of \mathbf{P} with the unit vector (\mathbf{i} or \mathbf{j}) in the direction of the translation. Since there is nothing special about these two directions, we conclude

that in general,

$$\hat{n} \cdot \mathbf{P} \equiv P_{\hat{n}} \quad (12.1.4)$$

is the generator of translations in the direction of the unit vector \hat{n} . Finite translation operators are found by exponentiation. Thus $T(\mathbf{a})$, which translates by \mathbf{a} , is given by

$$T(\mathbf{a}) = e^{-iaP_a/\hbar} = e^{-ia\hat{a} \cdot \mathbf{P}/\hbar} = e^{-i\mathbf{a} \cdot \mathbf{P}/\hbar} \quad (12.1.5)$$

where $\hat{a} = \mathbf{a}/a$.

The Consistency Test. Let us now ask if the translation operators we have constructed have the right laws of combination, i.e., if

$$T(\mathbf{b})T(\mathbf{a}) = T(\mathbf{a} + \mathbf{b}) \quad (12.1.6)$$

or equivalently if

$$e^{-i\mathbf{b} \cdot \mathbf{P}/\hbar} e^{-i\mathbf{a} \cdot \mathbf{P}/\hbar} = e^{-i(\mathbf{a} + \mathbf{b}) \cdot \mathbf{P}/\hbar} \quad (12.1.7)$$

This amounts to asking if P_x and P_y may be treated as c numbers in manipulating the exponentials. The answer is yes, since in view of Eqs. (12.1.1) and (12.1.2), the operators commute

$$[P_x, P_y] = 0 \quad (12.1.8)$$

and their q number nature does not surface here. The commutativity of P_x and P_y reflects the commutativity of translations in the x and y directions.

*Exercise 12.1.1.** Verify that $\hat{\mathbf{a}} \cdot \mathbf{P}$ is the generator of infinitesimal translations along \mathbf{a} by considering the relation

$$\langle x, y | I - \frac{i}{\hbar} \delta \mathbf{a} \cdot \mathbf{P} | \psi \rangle = \psi(x - \delta a_x, y - \delta a_y)$$

12.2. Rotations in Two Dimensions

Classically, the effect of a rotation $\phi_0 \mathbf{k}$, i.e., by an angle ϕ_0 about the z axis (counterclockwise in the x - y plane) has the following effect on the state of a particle:

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} \cos \phi_0 & -\sin \phi_0 \\ \sin \phi_0 & \cos \phi_0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (12.2.1)$$

$$\begin{bmatrix} p_x \\ p_y \end{bmatrix} \rightarrow \begin{bmatrix} \bar{p}_x \\ \bar{p}_y \end{bmatrix} = \begin{bmatrix} \cos \phi_0 & -\sin \phi_0 \\ \sin \phi_0 & \cos \phi_0 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \end{bmatrix} \quad (12.2.2)$$

Let us denote the operator that rotates these two-dimensional vectors by $R(\phi_0 \mathbf{k})$. It is represented by the 2×2 matrix in Eqs. (12.2.1) and (12.2.2). Just as $T(\mathbf{a})$ is the operator in Hilbert space associated with the translation \mathbf{a} , let $U[R(\phi_0 \mathbf{k})]$ be the operator associated with the rotation $R(\phi_0 \mathbf{k})$. In the active transformation picture‡

$$|\psi\rangle \xrightarrow{U[R]} |\psi_R\rangle = U[R]|\psi\rangle \quad (12.2.3)$$

The rotated state $|\psi_R\rangle$ must be such that

$$\langle X \rangle_R = \langle X \rangle \cos \phi_0 - \langle Y \rangle \sin \phi_0 \quad (12.2.4a)$$

$$\langle Y \rangle_R = \langle X \rangle \sin \phi_0 + \langle Y \rangle \cos \phi_0 \quad (12.2.4b)$$

$$\langle P_x \rangle_R = \langle P_x \rangle \cos \phi_0 - \langle P_y \rangle \sin \phi_0 \quad (12.2.5a)$$

$$\langle P_y \rangle_R = \langle P_x \rangle \sin \phi_0 + \langle P_y \rangle \cos \phi_0 \quad (12.2.5b)$$

where

$$\langle X \rangle_R = \langle \psi_R | X | \psi_R \rangle$$

and

$$\langle X \rangle = \langle \psi | X | \psi \rangle, \text{ etc.}$$

In analogy with the translation problem, we define the action of $U[R]$ on position eigenkets:

$$U[R]|x, y\rangle = |x \cos \phi_0 - y \sin \phi_0, x \sin \phi_0 + y \cos \phi_0\rangle \quad (12.2.6)$$

As in the case of translations, this equation is guided by more than just Eq. (12.2.4), which specifies how $\langle X \rangle$ and $\langle Y \rangle$ transform: in omitting a possible phase factor $g(x, y)$, we are also ensuring that $\langle P_x \rangle$ and $\langle P_y \rangle$ transform as in Eq. (12.2.5).

One way to show this is to keep the phase factor and use Eqs. (12.2.5a) and (12.2.5b) to eliminate it. We will take the simpler route of dropping it from the outset and proving at the end that $\langle P_x \rangle$ and $\langle P_y \rangle$ transform according to Eq. (12.2.5).

Explicit Construction of $U[R]$

Let us now construct $U[R]$. Consider first an infinitesimal rotation $\varepsilon_z \mathbf{k}$. In this case we set

$$U[R(\varepsilon_z \mathbf{k})] = I - \frac{i\varepsilon_z L_z}{\hbar} \quad (12.2.7)$$

‡ We will suppress the rotation angle when it is either irrelevant or obvious.

where L_z , the generator of infinitesimal rotations, is to be determined. Starting with Eq. (12.2.6), which becomes to first order in ε_z

$$U[R]|x, y\rangle = |x - y\varepsilon_z, x\varepsilon_z + y\rangle \quad (12.2.8)$$

it can be shown that

$$\langle x, y|I - \frac{i\varepsilon_z L_z}{\hbar}|\psi\rangle = \psi(x + y\varepsilon_z, y - x\varepsilon_z) \quad (12.2.9)$$

*Exercise 12.2.1.** Provide the steps linking Eq. (12.2.8) to Eq. (12.2.9). [Hint: Recall the derivation of Eq. (11.2.8) from Eq. (11.2.6).]

Expanding both sides to order ε_z

$$\begin{aligned} \langle x, y|I|\psi\rangle - \frac{i\varepsilon_z}{\hbar} \langle x, y|L_z|\psi\rangle &= \psi(x, y) + \frac{\partial\psi}{\partial x}(y\varepsilon_z) + \frac{\partial\psi}{\partial y}(-x\varepsilon_z) \\ \langle x, y|L_z|\psi\rangle &= \left[x \left(-i\hbar \frac{\partial}{\partial y} \right) - y \left(-i\hbar \frac{\partial}{\partial x} \right) \right] \psi(x, y) \end{aligned}$$

So

$$L_z \xrightarrow[\text{coordinate basis}]{} x \left(-i\hbar \frac{\partial}{\partial y} \right) - y \left(-i\hbar \frac{\partial}{\partial x} \right) \quad (12.2.10)$$

or in the abstract

$$L_z = X P_y - Y P_x \quad (12.2.11)$$

Let us verify that $\langle P_x \rangle$ and $\langle P_y \rangle$ transform according to Eq. (12.2.5). Since

$$L_z \xrightarrow[\text{momentum basis}]{} \left(i\hbar \frac{\partial}{\partial p_x} p_y - i\hbar \frac{\partial}{\partial p_y} p_x \right) \quad (12.2.12)$$

it is clear that

$$\frac{-i\varepsilon_z}{\hbar} \langle p_x, p_y|L_z|\psi\rangle = \frac{\partial\psi}{\partial p_x}(p_y\varepsilon_z) + \frac{\partial\psi}{\partial p_y}(-p_x\varepsilon_z) \quad (12.2.13)$$

Thus $I - i\varepsilon_z L_z/\hbar$ rotates the momentum space wave function $\psi(p_x, p_y)$ by ε_z in momentum space, and as a result $\langle P_x \rangle$ and $\langle P_y \rangle$ transform just as $\langle X \rangle$ and $\langle Y \rangle$ do, i.e., in accordance with Eq. (12.2.5).

We could have also derived Eq. (12.2.11) for L_z by starting with the passive transformation equations for an infinitesimal rotation:

$$U^\dagger[R]XU[R]=X-Y\varepsilon_z \quad (12.2.14a)$$

$$U^\dagger[R]YU[R]=X\varepsilon_z+Y \quad (12.2.14b)$$

$$U^\dagger[R]P_xU[R]=P_x-P_y\varepsilon_z \quad (12.2.15a)$$

$$U^\dagger[R]P_yU[R]=P_x\varepsilon_z+P_y \quad (12.2.15b)$$

By feeding Eq. (12.2.7) into the above we can deduce that

$$[X, L_z]=-i\hbar Y \quad (12.2.16a)$$

$$[Y, L_z]=i\hbar X \quad (12.2.16b)$$

$$[P_x, L_z]=-i\hbar P_y \quad (12.2.17a)$$

$$[P_y, L_z]=i\hbar P_x \quad (12.2.17b)$$

These commutation relations suffice to fix L_z as XP_y-YP_x .

Exercise 12.2.2. Using these commutation relations (and your keen hindsight) derive $L_z=XP_y-YP_x$. At least show that Eqs. (12.2.16) and (12.2.17) are consistent with $L_z=XP_y-YP_x$.

The finite rotation operator $U[R(\phi_0\mathbf{k})]$ is

$$U[R(\phi_0\mathbf{k})]=\lim_{N\rightarrow\infty}\left(I-\frac{i}{\hbar}\frac{\phi_0}{N}L_z\right)^N=\exp(-i\phi_0L_z/\hbar) \quad (12.2.18)$$

Given

$$L_z \xrightarrow[\text{coordinate basis}]{} x\left(-i\hbar\frac{\partial}{\partial y}\right)-y\left(-i\hbar\frac{\partial}{\partial x}\right)$$

it is hard to see that $e^{-i\phi_0L_z/\hbar}$ indeed rotates the state by the angle ϕ_0 . For one thing, expanding the exponential is complicated by the fact that $x(-i\hbar\partial/\partial y)$ and $y(-i\hbar\partial/\partial x)$ do not commute. So let us consider an alternative form for L_z . It can be shown, by changing to polar coordinates, that

$$L_z \xrightarrow[\text{coordinate basis}]{} -i\hbar\frac{\partial}{\partial\phi} \quad (12.2.19)$$

This result can also be derived more directly by starting with the requirement that under an infinitesimal rotation $\varepsilon_z \mathbf{k}$, $\psi(x, y) = \psi(\rho, \phi)$ becomes $\psi(\rho, \phi - \varepsilon_z)$.

*Exercise 12.2.3.** Derive Eq. (12.2.19) by doing a coordinate transformation on Eq. (12.2.10), and also by the direct method mentioned above.

Now it is obvious that

$$\exp(-i\phi_0 L_z/\hbar) \xrightarrow[\text{coordinate basis}]{} \exp\left(-\phi_0 \frac{\partial}{\partial \phi}\right) \quad (12.2.20)$$

rotates the state by an angle ϕ_0 about the z axis, for

$$\exp(-\phi_0 \partial/\partial \phi) \psi(\rho, \phi) = \psi(\rho, \phi - \phi_0)$$

by Taylor's theorem. It is also obvious that $U[R(\phi_0 \mathbf{k})] U[R(\phi_0 \mathbf{k})] = U[R((\phi_0 + \phi_0') \mathbf{k})]$. Thus the rotation operators have the right law of combination.

Physical Interpretation of L_z . We identify L_z as the angular momentum operator, since (i) it is obtained from $l_z = xp_y - yp_x$ by the usual substitution rule (Postulate II), and (ii) it is the generator of infinitesimal rotations about the z axis. L_z is conserved in a problem with rotational invariance: if

$$U^\dagger[R] H(X, P_x; Y, P_y) U[R] = H(X, P_x; Y, P_y) \quad (12.2.21)$$

it follows (by choosing an infinitesimal rotation) that

$$[L_z, H] = 0 \quad (12.2.22)$$

Since X , P_x , Y , and P_y respond to the rotation as do their classical counterparts [Eqs. (12.2.14) and (12.2.15)] and H is the same function of these operators as \mathcal{H} is of the corresponding classical variables, H is rotationally invariant whenever \mathcal{H} is.

Besides the conservation of $\langle L_z \rangle$, Eq. (12.2.22) also implies the following:

- (1) An experiment and its rotated version will give the same result if H is rotationally invariant.
- (2) There exists a common basis for L_z and H . (We will spend a lot of time discussing this basis as we go along.)

The Consistency Check. Let us now verify that our rotation and translation operators combine as they should. In contrast to pure translations or rotations, which have a simple law of composition, the combined effect of translations and rotations is nothing very simple. We seem to be facing the prospect of considering every possible combination of rotations and translations, finding their net effect, and then verifying that the product of the corresponding quantum operators equals the

operator corresponding to the result of all the transformations. Let us take one small step in this direction, which will prove to be a giant step toward our goal.

Consider the following product of four infinitesimal operations:

$$U[R(-\varepsilon_z \mathbf{k})]T(-\varepsilon)U[R(\varepsilon_z \mathbf{k})]T(\varepsilon)$$

where $\varepsilon = \varepsilon_x \mathbf{i} + \varepsilon_y \mathbf{j}$. By subjecting a point in the x - y plane to these four operations we find

$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix} &\xrightarrow{-\varepsilon} \begin{bmatrix} x + \varepsilon_x \\ y + \varepsilon_y \end{bmatrix} \xrightarrow{R(\varepsilon_z \mathbf{k})} \begin{bmatrix} (x + \varepsilon_x) - (y + \varepsilon_y) \varepsilon_z \\ (x + \varepsilon_x) \varepsilon_z + (y + \varepsilon_y) \end{bmatrix} \\ &\xrightarrow{-\varepsilon} \begin{bmatrix} x - (y + \varepsilon_y) \varepsilon_z \\ (x + \varepsilon_x) \varepsilon_z + y \end{bmatrix} \xrightarrow{R(-\varepsilon_z \mathbf{k})} \begin{bmatrix} x - \varepsilon_y \varepsilon_z \\ y + \varepsilon_x \varepsilon_z \end{bmatrix} \end{aligned} \quad (12.2.23)$$

i.e., that the net effect is a translation by $-\varepsilon_y \varepsilon_z \mathbf{i} + \varepsilon_x \varepsilon_z \mathbf{j}$.[‡] In the above, we have ignored terms involving ε_x^2 , ε_y^2 , ε_z^2 , and beyond. We do, however, retain the $\varepsilon_x \varepsilon_z$ and $\varepsilon_y \varepsilon_z$ terms since they contain the first germ of noncommutativity. Note that although these are second-order terms, they are fully determined in our approximation, i.e. unaffected by the second-order terms that we have ignored. Equation (12.2.23) imposes the following restriction on the quantum operators:

$$U[R(-\varepsilon_z \mathbf{k})]T(-\varepsilon)U[R(\varepsilon_z \mathbf{k})]T(\varepsilon) = T(-\varepsilon_y \varepsilon_z \mathbf{i} + \varepsilon_x \varepsilon_z \mathbf{j}) \quad (12.2.24)$$

or

$$\begin{aligned} &\left(I + \frac{i}{\hbar} \varepsilon_z L_z \right) \left[I + \frac{i}{\hbar} (\varepsilon_x P_x + \varepsilon_y P_y) \right] \left(I - \frac{i}{\hbar} \varepsilon_z L_z \right) \left[I - \frac{i}{\hbar} (\varepsilon_x P_x + \varepsilon_y P_y) \right] \\ &= I + \frac{i}{\hbar} \varepsilon_y \varepsilon_z P_x - \frac{i}{\hbar} \varepsilon_x \varepsilon_z P_y \end{aligned} \quad (12.2.25)$$

By matching coefficients (you should do this) we can deduce the following constraints:

$$[P_x, L_z] = -i\hbar P_y$$

$$[P_y, L_z] = i\hbar P_x$$

which are indeed satisfied by the generators [Eq. (12.2.17)].

So our operators have passed this test. But many other tests are possible. How about the coefficients of terms such as $\varepsilon_x \varepsilon_z^2$, or more generally, how about finite

[‡] Note that if rotations and translations commuted, the fourfold product would equal I , as can be seen by rearranging the factors so that the two opposite rotations and the two opposite translations cancel each other. The deviation from this result of I is a measure of noncommutativity. Given two symmetry operations that do not commute, the fourfold product provides a nice characterization of their noncommutativity. As we shall see, this characterization is complete.

rotations? How about tests other than the fourfold product, such as one involving 14 translations and six rotations interlaced?

There is a single answer to all these equations: there is no need to conduct any further tests. Although it is beyond the scope of this book to explain why this is so, it is not hard to explain when it is time to stop testing. We can stop the tests when all possible commutators between the generators have been considered. In the present case, given the generators P_x , P_y , and L_z , the possible commutators are $[P_x, L_z]$, $[P_y, L_z]$, and $[P_x, P_y]$. We have just finished testing the first two. Although the third was tested implicitly in the past, let us do it explicitly again. If we convert the law of combination

$$\begin{bmatrix} x \\ y \end{bmatrix} \xrightarrow{\varepsilon_x i} \begin{bmatrix} x + \varepsilon_x \\ y \end{bmatrix} \xrightarrow{\varepsilon_y i} \begin{bmatrix} x + \varepsilon_x \\ y + \varepsilon_y \end{bmatrix} \xrightarrow{-\varepsilon_x i} \begin{bmatrix} x \\ y + \varepsilon_y \end{bmatrix} \xrightarrow{-\varepsilon_y i} \begin{bmatrix} x \\ y \end{bmatrix} \quad (12.2.26)$$

into the operator constraint

$$T(-\varepsilon_y \mathbf{j}) T(-\varepsilon_x \mathbf{i}) T(\varepsilon_y \mathbf{j}) T(\varepsilon_x \mathbf{i}) = I \quad (12.2.27)$$

we deduce that

$$[P_x, P_y] = 0$$

which of course is satisfied by the generators P_x and P_y . [Although earlier on, we did not consider the fourfold product, Eq. (12.2.27), we did verify that the arguments of the T operators combined according to the laws of vector analysis. Equation (12.2.26) is just a special case which brings out the commutativity of P_x and P_y .]

When I say that there are no further tests to be conducted, I mean the following:

(1) Every consistency test will reduce to just another relation between the *commutators* of the generators.

(2) This relation will be automatically satisfied if the generators pass the tests we have finished conducting. The following exercise should illustrate this point.

*Exercise 12.2.4.** Rederive the equivalent of Eq. (12.2.23) keeping terms of order $\varepsilon_x \varepsilon_z^2$. (You may assume $\varepsilon_y = 0$.) Use this information to rewrite Eq. (12.2.24) to order $\varepsilon_x \varepsilon_z^2$. By equating coefficients of this term deduce the constraint

$$-2L_z P_x L_z + P_x L_z^2 + L_z^2 P_x = \hbar^2 P_x$$

This seems to conflict with statement (1) made above, but not really, in view of the identity

$$-\Lambda \Omega \Lambda + \Omega \Lambda^2 + \Lambda^2 \Omega \equiv [\Lambda, [\Lambda, \Omega]]$$

Using the identify, verify that the new constraint coming from the $\varepsilon_x \varepsilon_z^2$ term is satisfied given the commutation relations between P_x , P_y , and L_z .

Vector Operators

We call $\mathbf{V} = V_x \mathbf{i} + V_y \mathbf{j}$ a vector operator if V_x and V_y transform as components of a vector under a passive transformation generated by $U[R]$:

$$U^\dagger[R] V_i U[R] = \sum_j R_{ij} V_j$$

where R_{ij} is the 2×2 rotation matrix appearing in Eq. (12.2.1). Examples of \mathbf{V} are $\mathbf{P} = P_x \mathbf{i} + P_y \mathbf{j}$ and $\mathbf{R} = X \mathbf{i} + Y \mathbf{j}$ [see Eqs. (12.2.14) and (12.2.15)]. Note the twofold character of a vector operator such as \mathbf{P} : on the one hand, its components are operators in Hilbert space, and on the other, it transforms as a vector in $\mathbb{V}^2(R)$.

The same definition of a vector operator holds in three dimensions as well, with the obvious difference that R_{ij} is a 3×3 matrix.

12.3. The Eigenvalue Problem of L_z

We have seen that in a rotationally invariant problem, H and L_z share a common basis. In order to exploit this fact we must first find the eigenfunctions of L_z . We begin by writing

$$L_z |l_z\rangle = l_z |l_z\rangle \quad (12.3.1)$$

in the coordinate basis:

$$-i\hbar \frac{\partial \psi_{l_z}(\rho, \phi)}{\partial \phi} = l_z \psi_{l_z}(\rho, \phi) \quad (12.3.2)$$

The solution to this equation is

$$\psi_{l_z}(\rho, \phi) = R(\rho) e^{il_z \phi / \hbar} \quad (12.3.3)$$

where $R(\rho)$ is an arbitrary function normalizable with respect to $\int_0^\infty \rho d\rho$.[‡] We shall have more to say about $R(\rho)$ in a moment. But first note that l_z seems to be arbitrary: it can even be complex since ϕ goes only from 0 to 2π . (Compare this to the eigenfunctions $e^{ipx/\hbar}$ of linear momentum, where we could argue that p had to be real to keep $|\psi|$ bounded as $|x| \rightarrow \infty$.) The fact that complex eigenvalues enter the answer, signals that we are overlooking the Hermiticity constraint. Let us impose it. The condition

$$\langle \psi_1 | L_z | \psi_2 \rangle = \langle \psi_2 | L_z | \psi_1 \rangle^* \quad (12.3.4)$$

[‡] This will ensure that ψ is normalizable with respect to

$$\iint dx dy = \int_0^\infty \int_0^{2\pi} \rho d\rho d\phi$$

becomes in the coordinate basis

$$\int_0^\infty \int_0^{2\pi} \psi_1^* \left(-i\hbar \frac{\partial}{\partial \phi} \right) \psi_2 \rho \, d\rho \, d\phi = \left[\int_0^\infty \int_0^{2\pi} \psi_2^* \left(-i\hbar \frac{\partial}{\partial \phi} \right) \psi_1 \rho \, d\rho \, d\phi \right]^* \quad (12.3.5)$$

If this requirement is to be satisfied for all ψ_1 and ψ_2 , one can show (upon integrating by parts) that it is enough if each ψ obeys

$$\psi(\rho, 0) = \psi(\rho, 2\pi) \quad (12.3.6)$$

If we impose this constraint on the L_z eigenfunctions, Eq. (12.3.3), we find

$$1 = e^{2\pi i l_z / \hbar} \quad (12.3.7)$$

This forces l_z not merely to be real, but also to be an integral multiple of \hbar :

$$l_z = m\hbar, \quad m = 0, \pm 1, \pm 2, \dots \quad (12.3.8)$$

One calls m the *magnetic quantum number*. Notice that $l_z = m\hbar$ implies that ψ is a single-valued function of ϕ . (However, see Exercise 12.3.2.)

Exercise 12.3.1. Provide the steps linking Eq. (12.3.5) to Eq. (12.3.6).

Exercise 12.3.2. Let us try to deduce the restriction on l_z from another angle. Consider a superposition of two allowed l_z eigenstates:

$$\psi(\rho, \phi) = A(\rho) e^{i\phi l_z / \hbar} + B(\rho) e^{i\phi l_z' / \hbar}$$

By demanding that upon a 2π rotation we get the same physical state (not necessarily the same state vector), show that $l_z - l_z' = m\hbar$, where m is an integer. By arguing on the grounds of symmetry that the allowed values of l_z must be symmetric about zero, show that these values are either $\dots, 3\hbar/2, \hbar/2, -\hbar/2, -3\hbar/2, \dots$ or $\dots, 2\hbar, \hbar, 0, -\hbar, -2\hbar, \dots$. It is not possible to restrict l_z any further this way. \square

Let us now return to the arbitrary function $R(\rho)$ that accompanies the eigenfunctions of L_z . Its presence implies that the eigenvalue $l_z = m\hbar$ does not nail down a unique state in Hilbert space but only a subspace \mathbb{V}_m . The dimensionality of this space is clearly infinite, for the space of all normalizable functions R is infinite dimensional. The natural thing to do at this point is to introduce some operator that commutes with L_z and whose simultaneous eigenfunctions with L_z pick out a unique basis in each \mathbb{V}_m . We shall see in a moment that the Hamiltonian in a rotationally invariant problem does just this. Physically this means that a state is not uniquely specified by just its angular momentum (which only fixes the angular part of the wave function), but it can be specified by its energy and angular momentum in a rotationally invariant problem.

It proves convenient to introduce the functions

$$\Phi_m(\phi) = (2\pi)^{-1/2} e^{im\phi} \quad (12.3.9)$$

which would have been nondegenerate eigenfunctions of L_z if the ρ coordinate had not existed. These obey the orthonormality condition

$$\int_0^{2\pi} \Phi_m^*(\phi) \Phi_{m'}(\phi) d\phi = \delta_{mm'} \quad (12.3.10)$$

It will be seen that these functions play an important role in problems with rotational invariance.

*Exercise 12.3.3.** A particle is described by a wave function

$$\psi(\rho, \phi) = A e^{-\rho^2/2\Delta^2} \cos^2 \phi$$

Show (by expressing $\cos^2 \phi$ in terms of Φ_m) that

$$P(l_z=0) = 2/3$$

$$P(l_z=2\hbar) = 1/6$$

$$P(l_z=-2\hbar) = 1/6$$

(Hint: Argue that the radial part $e^{-\rho^2/2\Delta^2}$ is irrelevant here.)

*Exercise 12.3.4.** A particle is described by a wave function

$$\psi(\rho, \phi) = A e^{-\rho^2/2\Delta^2} \left(\frac{\rho}{\Delta} \cos \phi + \sin \phi \right)$$

Show that

$$P(l_z=\hbar) = P(l_z=-\hbar) = \frac{1}{2}$$

Solutions to Rotationally Invariant Problems

Consider a problem where $V(\rho, \phi) = V(\rho)$. The eigenvalue equation for H is

$$\left[\frac{-\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \right) + V(\rho) \right] \psi_E(\rho, \phi) = E \psi_E(\rho, \phi) \quad (12.3.11)$$

(We shall use μ to denote the mass, since m will denote the angular momentum quantum number.) Since $[H, L_z] = 0$ in this problem, we seek simultaneous eigenfunctions of H and L_z . We have seen that the most general eigenfunction of L_z with

eigenvalue $m\hbar$ is of the form

$$\psi_m(\rho, \phi) = R(\rho)(2\pi)^{-1/2}e^{im\phi} = R(\rho)\Phi_m(\phi)$$

where $R(\rho)$ is undetermined. In the present case R is determined by the requirement that

$$\psi_{Em}(\rho, \phi) = R_{Em}(\rho)\Phi_m(\phi) \quad (12.3.12)$$

be an eigenfunction of H as well, with eigenvalue E , i.e., that ψ_{Em} satisfy Eq. (12.3.11). Feeding the above form into Eq. (12.3.11), we get the *radial equation* that determines $R_{Em}(\rho)$ and the allowed values for E :

$$\left[\frac{-\hbar^2}{2\mu} \left(\frac{d^2}{d\rho^2} + \frac{1}{\rho} \frac{d}{d\rho} - \frac{m^2}{\rho^2} \right) + V(\rho) \right] R_{Em}(\rho) = ER_{Em}(\rho) \quad (12.3.13)$$

As we change the potential, only the *radial part* of the wave function, R , changes; the angular part Φ_m is unchanged. Thus the functions $\Phi_m(\phi)$, which were obtained by pretending ρ does not exist, provide the angular part of the wave function in the eigenvalue problem of any rotationally invariant Hamiltonian.

*Exercise 12.3.5**. Note that the angular momentum seems to generate a repulsive potential in Eq. (12.3.13). Calculate its gradient and identify it as the centrifugal force.

Exercise 12.3.6. Consider a particle of mass μ constrained to move on a circle of radius a . Show that $H = L_z^2/2\mu a^2$. Solve the eigenvalue problem of H and interpret the degeneracy.

Exercise 12.3.7. (The Isotropic Oscillator)*. Consider the Hamiltonian

$$H = \frac{P_x^2 + P_y^2}{2\mu} + \frac{1}{2} \mu \omega^2 (X^2 + Y^2)$$

(1) Convince yourself $[H, L_z] = 0$ and reduce the eigenvalue problem of H to the radial differential equation for $R_{Em}(\rho)$.

(2) Examine the equation as $\rho \rightarrow 0$ and show that

$$R_{Em}(\rho) \xrightarrow[\rho \rightarrow 0]{} \rho^{|m|}$$

(3) Show likewise that up to powers of ρ

$$R_{Em}(\rho) \xrightarrow[\rho \rightarrow \infty]{} e^{-\mu \omega \rho^2/2\hbar}$$

So assume that $R_{Em}(\rho) = \rho^{|m|} e^{-\mu \omega \rho^2/2\hbar} U_{Em}(\rho)$.

(4) Switch to dimensionless variables $\varepsilon = E/\hbar\omega$, $y = (\mu\omega/\hbar)^{1/2}\rho$.

(5) Convert the equation for R into an equation for U . (I suggest proceeding in two stages: $R = y^{|m|}f$, $f = e^{-y^2/2}U$.) You should end up with

$$U'' + \left[\left(\frac{2|m|+1}{y} \right) - 2y \right] U' + (2\varepsilon - 2|m| - 2)U = 0$$

(6) Argue that a power series for U of the form

$$U(y) = \sum_{r=0}^{\infty} C_r y^r$$

will lead to a *two-term* recursion relation.

(7) Find the relation between C_{r+2} and C_r . Argue that the series must terminate at some finite r if the $y \rightarrow \infty$ behavior of the solution is to be acceptable. Show $\varepsilon = r + |m| + 1$ leads to termination after r terms. Now argue that r is necessarily even—i.e., $r = 2k$. (Show that if r is odd, the behavior of R as $\rho \rightarrow 0$ is not $\rho^{|m|}$.) So finally you must end up with

$$E = (2k + |m| + 1)\hbar\omega, \quad k = 0, 1, 2, \dots$$

Define $n = 2k + |m|$, so that

$$E_n = (n + 1)\hbar\omega$$

(8) For a given n , what are the allowed values of $|m|$? Given this information show that for a given n , the degeneracy is $n + 1$. Compare this to what you found in Cartesian coordinates (Exercise 10.2.2).

(9) Write down all the normalized eigenfunctions corresponding to $n = 0, 1$.

(10) Argue that the $n = 0$ function *must* equal the corresponding one found in Cartesian coordinates. Show that the two $n = 2$ solutions are linear combinations of their counterparts in Cartesian coordinates. Verify that the parity of the states is $(-1)^n$ as you found in Cartesian coordinates.

*Exercise 12.3.8.** Consider a particle of charge q in a vector potential

$$\mathbf{A} = \frac{B}{2}(-y\mathbf{i} + x\mathbf{j})$$

(1) Show that the magnetic field is $\mathbf{B} = B\mathbf{k}$.

(2) Show that a classical particle in this potential will move in circles at an angular frequency $\omega_0 = qB/\mu c$.

(3) Consider the Hamiltonian for the corresponding quantum problem:

$$H = \frac{[P_x + qYB/2c]^2}{2\mu} + \frac{[P_y - qXB/2c]^2}{2\mu}$$

Show that $Q = (cP_x + qYB/2)/qB$ and $P = (P_y - qXB/2c)$ are canonical. Write H in terms of P and Q and show that allowed levels are $E = (n + 1/2)\hbar\omega_0$.

(4) Expand H out in terms of the original variables and show

$$H = H\left(\frac{\omega_0}{2}, \mu\right) - \frac{\omega_0}{2} L_z$$

where $H(\omega_0/2, \mu)$ is the Hamiltonian for an isotropic two-dimensional harmonic oscillator of mass μ and frequency $\omega_0/2$. Argue that the same basis that diagonalized $H(\omega_0/2, \mu)$ will diagonalize H . By thinking in terms of this basis, show that the allowed levels for H are $E = (k + \frac{1}{2}|m| - \frac{1}{2}m + \frac{1}{2})\hbar\omega_0$, where k is any integer and m is the angular momentum. Convince yourself that you get the same levels from this formula as from the earlier one [$E = (n + 1/2)\hbar\omega_0$]. We shall return to this problem in Chapter 21.

12.4. Angular Momentum in Three Dimensions

It is evident that as we pass from two to three dimensions, the operator L_z picks up two companions L_x and L_y which generate infinitesimal rotations about the x and y axes, respectively. So we have

$$L_x = YP_z - ZP_y \quad (12.4.1a)$$

$$L_y = ZP_x - XP_z \quad (12.4.1b)$$

$$L_z = XP_y - YP_x \quad (12.4.1c)$$

As usual, we subject these to the consistency test. It may be verified, (Exercise 12.4.2), that if we take a point in three-dimensional space and subject it to the following rotations: $R(\varepsilon_x \mathbf{i})$, $R(\varepsilon_y \mathbf{j})$, $R(-\varepsilon_x \mathbf{i})$ and lastly $R(-\varepsilon_y \mathbf{j})$, it ends up rotated by $-\varepsilon_x \varepsilon_y \mathbf{k}$. In other words

$$R(-\varepsilon_y \mathbf{j})R(-\varepsilon_x \mathbf{i})R(\varepsilon_y \mathbf{j})R(\varepsilon_x \mathbf{i}) = R(-\varepsilon_x \varepsilon_y \mathbf{k}) \quad (12.4.2)$$

It follows that the quantum operators $U[R]$ must satisfy

$$U[R(-\varepsilon_y \mathbf{j})]U[R(-\varepsilon_x \mathbf{i})]U[R(\varepsilon_y \mathbf{j})]U[R(\varepsilon_x \mathbf{i})] = U[R(-\varepsilon_x \varepsilon_y \mathbf{k})] \quad (12.4.3)$$

If we write each U to order ε and match coefficients of $\varepsilon_x \varepsilon_y$, we will find

$$[L_x, L_y] = i\hbar L_z \quad (12.4.4a)$$

By considering two similar tests involving $\varepsilon_y \varepsilon_z$ and $\varepsilon_z \varepsilon_x$, we can deduce the constraints

$$[L_y, L_z] = i\hbar L_x \quad (12.4.4b)$$

$$[L_z, L_x] = i\hbar L_y \quad (12.4.4c)$$

You may verify that the operators in Eq. (12.4.1) satisfy these constraints. So they are guaranteed to generate finite rotation operators that obey the right laws of combination.

The three relations above may be expressed compactly as one vector equation

$$\mathbf{L} \times \mathbf{L} = i\hbar \mathbf{L} \quad (12.4.5)$$

Yet another way to write the commutation relations is

$$[L_i, L_j] = i\hbar \sum_{k=1}^3 \varepsilon_{ijk} L_k \quad (12.4.6)$$

In this equation, i and j run from 1 to 3, L_1 , L_2 , and L_3 stand for L_x , L_y , and L_z , respectively,[†] and ε_{ijk} are the components of an antisymmetric tensor of rank 3, with the following properties:

- (1) They change sign when any two indices are exchanged. Consequently no two indices can be equal.
- (2) $\varepsilon_{123} = 1$.

This fixes all other components. For example,

$$\varepsilon_{132} = -1, \quad \varepsilon_{312} = (-1)(-1) = +1 \quad (12.4.7)$$

and so on. In short, ε_{ijk} is +1 for any cyclic permutation of the indices in ε_{123} and -1 for the others. (The relation

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} \quad (12.4.8)$$

between three vectors from $\mathbb{V}^3(R)$ may be written in component form as

$$c_i = \sum_{j=1}^3 \sum_{k=1}^3 \varepsilon_{ijk} a_j b_k \quad (12.4.9)$$

Of course $\mathbf{a} \times \mathbf{a}$ is zero if \mathbf{a} is a vector whose components are numbers, but not zero if it is an operator such as \mathbf{L} .)

*Exercise 12.4.1.** (1) Verify that Eqs. (12.4.9) and Eq. (12.4.8) are equivalent, given the definition of ε_{ijk} .

(2) Let U_1 , U_2 , and U_3 be three energy eigenfunctions of a single particle in some potential. Construct the wave function $\psi_A(x_1, x_2, x_3)$ of three fermions in this potential, one of which is in U_1 , one in U_2 , and one in U_3 , using the ε_{ijk} tensor.

*Exercise 12.4.2.** (1) Verify Eq. (12.4.2) by first constructing the 3×3 matrices corresponding to $R(\varepsilon_x \mathbf{i})$ and $R(\varepsilon_y \mathbf{j})$, to order ε .

(2) Provide the steps connecting Eqs. (12.4.3) and (12.4.4a).

[†] We will frequently let the indices run over 1, 2, and 3 instead of x , y , and z .

(3) Verify that L_x and L_y defined in Eq. (12.4.1) satisfy Eq. (12.4.4a). The proof for other commutators follows by cyclic permutation.

We next define the total angular momentum operator squared

$$L^2 = L_x^2 + L_y^2 + L_z^2 \quad (12.4.10)$$

It may be verified (by you) that

$$[L^2, L_i] = 0, \quad i = x, y, \text{ or } z \quad (12.4.11)$$

Finite Rotation Operators. Rotations about a given axis commute. So a finite rotation may be viewed as a sequence of infinitesimal rotations about the same axis. What is the operator that rotates by angle θ , i.e., by an amount θ about an axis parallel to $\hat{\theta}$? If $\theta = \theta_x \mathbf{i}$, then clearly

$$U[R(\theta_x \mathbf{i})] = e^{-i\theta_x L_x / \hbar}$$

The same goes for θ along the unit vectors \mathbf{j} and \mathbf{k} . What if θ has some arbitrary direction? We conjecture that $L_{\hat{\theta}} \equiv \hat{\theta} \cdot \mathbf{L}$ (where $\hat{\theta} = \theta / \theta$) is the generator of infinitesimal rotations about that axis and that

$$\begin{aligned} U[R(\theta)] &= \lim_{N \rightarrow \infty} \left(I - \frac{i}{\hbar} \frac{\theta}{N} \hat{\theta} \cdot \mathbf{L} \right)^N = e^{-i\theta \hat{\theta} \cdot \mathbf{L} / \hbar} \\ &= e^{-\theta \cdot \mathbf{L} / \hbar} \end{aligned} \quad (12.4.12)$$

Our conjecture is verified in the following exercise.

*Exercise 12.4.3.** We would like to show that $\hat{\theta} \cdot \mathbf{L}$ generates rotations about the axis parallel to $\hat{\theta}$. Let $\delta\theta$ be an infinitesimal rotation parallel to θ .

(1) Show that when a vector \mathbf{r} is rotated by an angle $\delta\theta$, it changes to $\mathbf{r} + \delta\theta \times \mathbf{r}$. (It might help to start with $\mathbf{r} \perp \delta\theta$ and then generalize.)

(2) We therefore demand that (to first order, as usual)

$$\psi(\mathbf{r}) \xrightarrow{U[R(\delta\theta)]} \psi(\mathbf{r} - (\delta\theta \times \mathbf{r})) = \psi(\mathbf{r}) - (\delta\theta \times \mathbf{r}) \cdot \nabla \psi$$

Comparing to $U[R(\delta\theta)] = I - (i \delta\theta / \hbar) L_{\hat{\theta}}$, show that $L_{\hat{\theta}} = \hat{\theta} \cdot \mathbf{L}$.

*Exercise 12.4.4.** Recall that \mathbf{V} is a vector operator if its components V_i transform as

$$U^\dagger[R] V_i U[R] = \sum_j R_{ij} V_j \quad (12.4.13)$$

(1) For an infinitesimal rotation $\delta\theta$, show, on the basis of the previous exercise, that

$$\sum_j R_{ij} V_j = V_i + (\delta\theta \times \mathbf{V})_i = V_i + \sum_j \sum_k \epsilon_{ijk} (\delta\theta)_j V_k$$

(2) Feed in $U[R] = 1 - (i/\hbar)\delta\theta \cdot \mathbf{L}$ into the left-hand side of Eq. (12.4.13) and deduce that

$$[V_i, L_j] = i\hbar \sum_k \epsilon_{ijk} V_k \quad (12.4.14)$$

This is as good a definition of a vector operator as Eq. (12.4.13). By setting $\mathbf{V} = \mathbf{L}$, we can obtain the commutation rules among the L 's.

If the Hamiltonian is invariant under arbitrary rotations,

$$U^\dagger[R] H U[R] = H \quad (12.4.15)$$

it follows (upon considering infinitesimal rotations around the x , y , and z axes) that

$$[H, L_i] = 0 \quad (12.4.16)$$

and from it

$$[H, L^2] = 0 \quad (12.4.17)$$

Thus L^2 and all three components of \mathbf{L} are conserved. It does not, however, follow that there exists a basis common to H and all three L 's. This is because the L 's do not commute with each other. So the best one can do is find a basis common to H , L^2 , and one of the L 's, usually chosen to be L_z .

We now examine the eigenvalue problem of the commuting operators L^2 and L_z . When this is solved, we will turn to the eigenvalue problem of H , L^2 , and L_z .

12.5. The Eigenvalue Problem of L^2 and L_z

There is a close parallel between our approach to this problem and that of the harmonic oscillator. Recall that in that case we (1) solved the eigenvalue problem of H in the coordinate basis; (2) solved the problem in the energy basis directly, using the a and a^\dagger operators, the commutation rules, and the positivity of H ; (3) obtained the coordinate wave function $\psi_n(y)$ given the results of part (2), by the following trick. We wrote

$$a|0\rangle = 0$$

in the coordinate basis as

$$\left(y + \frac{\partial}{\partial y} \right) \psi_0(y) = 0$$

which immediately gave us $\psi_0(y) \sim e^{-y^2/2}$, up to a normalization that could be easily determined.

Given the normalized eigenfunction $\psi_0(y)$, we got $\psi_n(y)$ by the application of the (differential) operator $(a^\dagger)^n/(n!)^{1/2} \rightarrow (y - \hat{c}/\hat{c}y)^n/(2^n n!)^{1/2}$.

In the present case we omit part (1), which involves just one more bout with differential equations and is not particularly enlightening.

Let us now consider part (2). It too has many similarities with part (2) of the oscillator problem.[‡] We begin by assuming that there exists a basis $|\alpha, \beta\rangle$ common to L^2 and L_z :

$$L^2|\alpha\beta\rangle = \alpha|\alpha\beta\rangle \quad (12.5.1)$$

$$L_z|\alpha\beta\rangle = \beta|\alpha\beta\rangle \quad (12.5.2)$$

We now define *raising and lower operators*

$$L_\pm = L_x \pm iL_y \quad (12.5.3)$$

which satisfy

$$[L_z, L_\pm] = \pm \hbar L_\pm \quad (12.5.4)$$

and of course (since L^2 commutes with L_x and L_y)

$$[L^2, L_\pm] = 0 \quad (12.5.5)$$

Equations (12.5.4) and (12.5.5) imply that L_\pm raise/lower the eigenvalue of L_z by \hbar , while leaving the eigenvalue of L^2 alone. For example,

$$\begin{aligned} L_z(L_+|\alpha\beta\rangle) &= (L_+L_z + \hbar L_+)|\alpha\beta\rangle \\ &= (L_+\beta + \hbar L_+)|\alpha\beta\rangle \\ &= (\beta + \hbar)(L_+|\alpha\beta\rangle) \end{aligned} \quad (12.5.6)$$

and

$$L^2 L_+ |\alpha\beta\rangle = L_+ L^2 |\alpha\beta\rangle = \alpha L_+ |\alpha\beta\rangle \quad (12.5.7)$$

From Eqs. (12.5.6) and (12.5.7) it is clear that $L_+|\alpha\beta\rangle$ is proportional to the normalized eigenket $|\alpha, \beta + \hbar\rangle$:

$$L_+|\alpha\beta\rangle = C_+(\alpha, \beta)|\alpha, \beta + \hbar\rangle \quad (12.5.8a)$$

[‡] If you have forgotten the latter, you are urged to refresh your memory at this point.

It can similarly be shown that

$$L_-|\alpha\beta\rangle = C_-(\alpha, \beta)|\alpha, \beta - \hbar\rangle \quad (12.5.8b)$$

The existence of L_{\pm} implies that given an eigenstate $|\alpha\beta\rangle$ there also exist eigenstates $|\alpha, \beta + \hbar\rangle, |\alpha, \beta + 2\hbar\rangle, \dots$; and $|\alpha, \beta - \hbar\rangle, |\alpha, \beta - 2\hbar\rangle, \dots$. This clearly signals trouble, for classical intuition tells us that the z component of angular momentum cannot take arbitrarily large positive or negative values for a given value of the square of the total angular momentum; in fact classically $|l_z| \leq (l^2)^{1/2}$.

Quantum mechanically we have

$$\langle\alpha\beta|L^2 - L_z^2|\alpha\beta\rangle = \langle\alpha\beta|L_x^2 + L_y^2|\alpha\beta\rangle \quad (12.5.9)$$

which implies

$$\alpha - \beta^2 \geq 0$$

(since $L_x^2 + L_y^2$ is positive definite) or

$$\alpha \geq \beta^2 \quad (12.5.10)$$

Since β^2 is bounded by α , it follows that there must exist a state $|\alpha\beta_{\max}\rangle$ such that it cannot be raised:

$$L_+|\alpha\beta_{\max}\rangle = 0 \quad (12.5.11)$$

Operating with L_- and using $L_-L_+ = L^2 - L_z^2 - \hbar L_z$, we get

$$\begin{aligned} (L^2 - L_z^2 - \hbar L_z)|\alpha\beta_{\max}\rangle &= 0 \\ (\alpha - \beta_{\max}^2 - \hbar\beta_{\max})|\alpha\beta_{\max}\rangle &= 0 \\ \alpha &= \beta_{\max}(\beta_{\max} + \hbar) \end{aligned} \quad (12.5.12)$$

Starting with $|\alpha\beta_{\max}\rangle$ let us operate k times with L_- , till we reach a state $|\alpha\beta_{\min}\rangle$ that cannot be lowered further without violating the inequality (12.5.10):

$$\begin{aligned} L_-|\alpha\beta_{\min}\rangle &= 0 \\ L_+L_-|\alpha\beta_{\min}\rangle &= 0 \\ (L^2 - L_z^2 + \hbar L_z)|\alpha\beta_{\min}\rangle &= 0 \\ \alpha &= \beta_{\min}(\beta_{\min} - \hbar) \end{aligned} \quad (12.5.13)$$

A comparison of Eqs. (12.5.12) and (12.5.13) shows (as is to be expected)

$$\beta_{\min} = -\beta_{\max} \quad (12.5.14)$$

Table 12.1. Some Low-Angular-Momentum States

(Angular momentum)			α	$ \alpha\beta\rangle$
$k/2$	β_{\max}			
0	0	0	0	$ 0, 0\rangle$
1/2	$\hbar/2$	$(1/2)(3/2)\hbar^2$	$ (3/4)\hbar^2, \hbar/2\rangle$ $ (3/4)\hbar^2, -\hbar/2\rangle$	
1	\hbar	$(1)(2)\hbar^2$	$ 2\hbar^2, \hbar\rangle$ $ 2\hbar^2, 0\rangle$ $ 2\hbar^2, -\hbar\rangle$	
3/2	⋮	⋮	⋮	⋮

Since we got to $|\alpha\beta_{\min}\rangle$ from $|\alpha\beta_{\max}\rangle$ in k steps of \hbar each, it follows that

$$\begin{aligned}\beta_{\max} - \beta_{\min} &= 2\beta_{\max} = \hbar k \\ \beta_{\max} &= \frac{\hbar k}{2}, \quad k = 0, 1, 2, \dots\end{aligned}\tag{12.5.15a}$$

$$\alpha = (\beta_{\max})(\beta_{\max} + \hbar) = \hbar^2 \left(\frac{k}{2}\right) \left(\frac{k}{2} + 1\right)\tag{12.5.15b}$$

We shall refer to $(k/2) = (\beta_{\max}/\hbar)$ as the *angular momentum of the state*. Notice that unlike in classical physics, β_{\max}^2 is less than α , the square of the magnitude of angular momentum, except when $\alpha = \beta_{\max} = 0$, i.e., in a state of zero angular momentum.

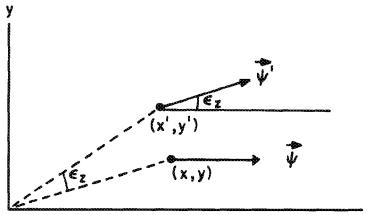
Let us now take a look at a few of the low-angular-momentum states listed in Table 12.1.

At this point the astute reader raises the following objection.

A.R.: I am disturbed by your results for odd k . You seem to find that L_z can have half-integral eigenvalues (in units of \hbar). But you just convinced us in Section 12.3 that L_z has only integral eigenvalues m (in units of \hbar). Where did you go wrong?

R.S.: Nowhere, but your point is well taken. The extra (half-integral) eigenvalues arise because we have solved a more general problem than that of L_x , L_y , L_z , and L^2 (although we didn't intend to). Notice that nowhere in the derivation did we use the explicit expressions for the L 's [Eq. (12.4.1)] and in particular $L_z \rightarrow -i\hbar\partial/\partial\phi$. (Had we done so, we would have gotten only integral eigenvalues as you expect.) We relied instead on just the commutation relations, $\mathbf{L} \times \mathbf{L} = i\hbar\mathbf{L}$. Now, these commutation relations reflect the law of combinations of infinitesimal rotations in three dimensions and must be satisfied by the three generators of rotations whatever the nature of the wave functions they rotate. We have so far considered just scalar wave functions $\psi(x, y, z)$, which assign a complex number (scalar) to each point. Now, there can be particles in nature for which the wave function is more complicated, say a vector field $\Psi(x, y, z) = \psi_x(x, y, z)\mathbf{i} + \psi_y(x, y, z)\mathbf{j} + \psi_z(x, y, z)\mathbf{k}$. The response of such a wave function to rotations is more involved. Whereas in the scalar case the effect of rotation by $\delta\theta$ is to take the number assigned to each point (x, y, z)

Figure 12.1. The effect of the infinitesimal rotations by ϵ_z on a vector ψ in two dimensions is to (1) first reassign it to the rotated point (x', y') (2) and then rotate the vector itself by the infinitesimal angle. The differential operator L_z does the first part while a 2×2 spin matrix S_z does the second.



and reassign it to the rotated point (x', y') , in the vector case the vector at (x, y, z) (i) must itself be rotated by $\delta\theta$ and (ii) then reassigned to (x', y', z') . (A simple example from two dimensions is given in Fig. 12.1.) The differential operators L_x , L_y , and L_z will only do part (ii) but not part (i), which has to be done by 3×3 matrices S_x , S_y , and S_z which shuffle the components ψ_x , ψ_y , ψ_z of Ψ . In such cases, the generators of infinitesimal rotations will be of the form

$$J_i = L_i + S_i$$

where L_i does part (2) and S_i does part (1) (see Exercise 12.5.1 for a concrete example). One refers to L_i as the *orbital angular momentum*, S_i as the *spin angular momentum* (or simply spin), and J_i as the *total angular momentum*. We do not yet know what J_i or S_i look like in these general cases, but we do know this: the J_i 's must obey the same commutation rules as the L_i 's, for the commutation rules reflect the law of combination of rotations and must be obeyed by any triplet of generators (the consistency condition), whatever be the nature of wave function they rotate. So in general we have

$$\mathbf{J} \times \mathbf{J} = i\hbar \mathbf{J} \quad (12.5.16)$$

with \mathbf{L} as a special case when the wave function is a scalar. So our result, which followed from just the commutation relations, applies to the problem of arbitrary \mathbf{J} and not just \mathbf{L} . Thus the answer to the question raised earlier is that unlike L_z , J_z is not restricted to have integral eigenvalues. But our analysis tells us, who know very little about spin, that S_z can have only integral or half-integral eigenvalues if the commutation relations are to be satisfied. Of course, our analysis doesn't imply that there *must* exist particles with spin integral or half integral—but merely reveals the possible variety in wave functions. But the old maxim—if something can happen, it will—is true here and nature does provide us with particles that possess spin—i.e., particles whose wave functions are more complicated than scalars. We will study them in Chapter 14 on spin.

*Exercise 12.5.1.** Consider a vector field $\Psi(x, y)$ in two dimensions. From Fig. 12.1 it follows that under an infinitesimal rotation $\epsilon_z \mathbf{k}$,

$$\psi_x \rightarrow \psi'_x(x, y) = \psi_x(x + y\epsilon_z, y - x\epsilon_z) - \psi_y(x + y\epsilon_z, y - x\epsilon_z)\epsilon_z$$

$$\psi_y \rightarrow \psi'_y(x, y) = \psi_x(x + y\epsilon_z, y - x\epsilon_z)\epsilon_z + \psi_y(x + y\epsilon_z, y - x\epsilon_z)$$

Show that (to order ε_z)

$$\begin{bmatrix} \psi'_x \\ \psi'_y \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{i\varepsilon_z}{\hbar} \begin{bmatrix} L_z & 0 \\ 0 & L_z \end{bmatrix} - \frac{i\varepsilon_z}{\hbar} \begin{bmatrix} 0 & -i\hbar \\ i\hbar & 0 \end{bmatrix} \right) \begin{bmatrix} \psi_x \\ \psi_y \end{bmatrix}$$

so that

$$\begin{aligned} J_z &= L_z^{(1)} \otimes I^{(2)} + I^{(1)} \otimes S_z^{(2)} \\ &= L_z + S_z \end{aligned}$$

where $I^{(2)}$ is a 2×2 identity matrix with respect to the vector components, $I^{(1)}$ is the identity operator with respect to the argument (x, y) of $\Psi(x, y)$. This example only illustrates the fact that $J_z = L_z + S_z$ if the wave function is not a scalar. An example of half-integral eigenvalues will be provided when we consider spin in a later chapter. (In the present example, S_z has eigenvalues $\pm \hbar$.)

Let us return to our main discussion. To emphasize the generality of the results we have found, we will express them in terms of J 's rather than L 's and also switch to a more common notation. Here is a summary of what we have found. The eigenvectors of the operators J^2 and J_z are given by

$$J^2|jm\rangle = j(j+1)\hbar^2|jm\rangle, \quad j=0, 1/2, 1, 3/2, \dots \quad (12.5.17a)$$

$$J_z|jm\rangle = m\hbar|jm\rangle, \quad m=j, j-1, j-2, \dots, -j \quad (12.5.17b)$$

We shall call j the angular momentum of the state. Note that in the above m can be an integer or half-integer depending on j .

The results for the restricted problem $\mathbf{J}=\mathbf{L}$ that we originally set out to solve are contained in Eq. (12.5.17): we simply ignore the states with half-integral m and j . To remind us in these cases that we are dealing with $\mathbf{J}=\mathbf{L}$, we will denote these states by $|lm\rangle$. They obey

$$L^2|lm\rangle = l(l+1)\hbar^2|lm\rangle, \quad l=0, 1, 2, \dots \quad (12.5.18a)$$

$$L_z|lm\rangle = m\hbar|lm\rangle, \quad m=l, l-1, \dots, -l \quad (12.5.18b)$$

Our problem has not been fully solved: we have only found the eigenvalues—the eigenvectors aren't fully determined yet. (As in the oscillator problem, finding the eigenvectors means finding the matrices corresponding to the basic operators whose commutation relations are given.) Let us continue our analysis in terms of the J 's. If we rewrite Eq. (12.5.8) in terms of J_{\pm} , j , and m (instead of L_{\pm} , α , and β), we get

$$J_{\pm}|jm\rangle = C_{\pm}(j, m)|j, m \pm 1\rangle \quad (12.5.19)$$

where $C_{\pm}(j, m)$ are yet to be determined. We will determine them now.

If we take the adjoint of

$$J_+|jm\rangle = C_+(j, m)|j, m+1\rangle$$

we get

$$\langle jm| J_- = C_+^*(j, m)\langle j, m+1|$$

Equating the inner product of the objects on the left-hand side to the product of the objects on the right-hand side, we obtain

$$\begin{aligned} \langle jm| J_- J_+ |jm\rangle &= |C_+(j, m)|^2 \langle j, m+1| j, m+1\rangle \\ &= |C_+(j, m)|^2 \\ \langle jm| J^2 - J_z^2 - \hbar J_z |jm\rangle &= |C_+(j, m)|^2 \end{aligned}$$

or

$$\begin{aligned} |C_+(j, m)|^2 &= j(j+1)\hbar^2 - m^2\hbar^2 - m\hbar^2 \\ &= \hbar^2(j-m)(j+m+1) \end{aligned}$$

or[‡]

$$C_+(j, m) = \hbar[(j-m)(j+m+1)]^{1/2}$$

It can likewise be shown that

$$C_-(j, m) = \hbar[(j+m)(j-m+1)]^{1/2}$$

so that finally

$$J_{\pm}|jm\rangle = \hbar[(j \mp m)(j \pm m+1)]^{1/2}|j, m \pm 1\rangle \quad (12.5.20)$$

Notice that when J_{\pm} act on $|j, \pm j\rangle$ they kill the state, so that each family with a given angular momentum j has only $2j+1$ states with eigenvalues $j\hbar, (j-1)\hbar, \dots, -(j\hbar)$ for J_z .

Equation (12.5.20) brings us to the end of our calculation, for we can write down the matrix elements of J_x and J_y in this basis:

$$\begin{aligned} \langle j'm'| J_x |jm\rangle &= \langle j'm' | \frac{J_+ + J_-}{2} |jm\rangle \\ &= \frac{\hbar}{2} \{ \delta_{jj'} \delta_{m'm+1} [(j-m)(j+m+1)]^{1/2} + \delta_{jj'} \delta_{m'm-1} \\ &\quad \times [(j+m)(j-m+1)]^{1/2} \} \end{aligned} \quad (12.5.21a)$$

[‡] There can be an overall phase factor in front of C_+ . We choose it to be unity according to standard convention.

$$\langle j'm' | J_y | jm \rangle = \langle j'm' | \frac{J_+ - J_-}{2i} | jm \rangle$$

$$= \frac{\hbar}{2i} \left\{ \delta_{jj'} \delta_{m', m+1} [(j-m)(j+m+1)]^{1/2} - \delta_{jj'} \delta_{m', m-1} \right. \\ \left. \times [(j+m)(j-m+1)]^{1/2} \right\} \quad (12.5.21b)$$

Using these (or our mnemonic based on images) we can write down the matrices corresponding to J^2 , J_z , J_x , and J_y in the $|jm\rangle$ basis[†]:

$$J^2 \rightarrow \begin{array}{c} \diagdown jm \\ j'm' \end{array} \begin{matrix} (0, 0) & (\frac{1}{2}, \frac{1}{2}) & (\frac{1}{2}, -\frac{1}{2}) & (1, 1) & (1, 0) & (1, -1) & \cdots \\ \left[\begin{array}{cccccc} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & \frac{3}{4}\hbar^2 & 0 & 0 & 0 & 0 & \\ 0 & 0 & \frac{3}{4}\hbar^2 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 2\hbar^2 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 2\hbar^2 & 0 & \\ 0 & 0 & 0 & 0 & 0 & 2\hbar^2 & \\ \vdots & & & & & & \ddots \end{array} \right] \end{matrix} \quad (12.5.22)$$

J_z is also diagonal with elements $m\hbar$.

$$J_z \rightarrow \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \hbar/2 & 0 & 0 & 0 & \\ 0 & \hbar/2 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \hbar/2^{1/2} & 0 & \\ 0 & 0 & 0 & \hbar/2^{1/2} & 0 & \hbar/2^{1/2} & \\ 0 & 0 & 0 & 0 & \hbar/2^{1/2} & 0 & \\ \vdots & & & & & & \ddots \end{bmatrix} \quad (12.5.23)$$

$$J_y \rightarrow \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & -i\hbar/2 & 0 & 0 & 0 & \\ 0 & i\hbar/2 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & -i\hbar/2^{1/2} & 0 & \\ 0 & 0 & 0 & i\hbar/2^{1/2} & 0 & -i\hbar/2^{1/2} & \\ 0 & 0 & 0 & 0 & i\hbar/2^{1/2} & 0 & \\ \vdots & & & & & & \ddots \end{bmatrix} \quad (12.5.24)$$

Notice that although J_x and J_y are not diagonal in the $|jm\rangle$ basis, they are *block diagonal*: they have no matrix elements between one value of j and another. This is

[†] The quantum numbers j and m do not fully label a state; a state is labeled by $|ajm\rangle$, where a represents the remaining labels. In what follows, we suppress a but assume it is the same throughout.

because J_{\pm} (out of which they are built) do not change j when they act on $|jm\rangle$. Since the J 's are all block diagonal, the blocks do not mix when we multiply them. In particular when we consider a commutation relation such as $[J_x, J_y] = i\hbar J_z$, it will be satisfied within each block. If we denote the $(2j+1) \times (2j+1)$ block in J_i , corresponding to a certain j , by $J_i^{(j)}$, then we have

$$[J_x^{(j)}, J_y^{(j)}] = i\hbar J_z^{(j)}, \quad j=0, \frac{1}{2}, 1, \dots \quad (12.5.25)$$

Exercise 12.5.2. (1) Verify that the 2×2 matrices $J_x^{(1/2)}$, $J_y^{(1/2)}$, and $J_z^{(1/2)}$ obey the commutation rule $[J_x^{(1/2)}, J_y^{(1/2)}] = i\hbar J_z^{(1/2)}$.

(2) Do the same for the 3×3 matrices $J_i^{(1)}$.

(3) Construct the 4×4 matrices and verify that

$$[J_x^{(3/2)}, J_y^{(3/2)}] = i\hbar J_z^{(3/2)}$$

*Exercise 12.5.3.** (1) Show that $\langle J_x \rangle = \langle J_y \rangle = 0$ in a state $|jm\rangle$.

(2) Show that in these states

$$\langle J_x^2 \rangle = \langle J_y^2 \rangle = \frac{1}{2} \hbar^2 [j(j+1) - m^2]$$

(use symmetry arguments to relate $\langle J_x^2 \rangle$ to $\langle J_y^2 \rangle$).

(3) Check that $\Delta J_x \cdot \Delta J_y$ from part (2) satisfies the inequality imposed by the uncertainty principle [Eq. (9.2.9)].

(4) Show that the uncertainty bound is saturated in the state $|j, \pm j\rangle$.

Finite Rotations‡

Now that we have explicit matrices for the generators of rotations, J_x , J_y , and J_z , we can construct the matrices representing $U[R]$ by exponentiating $(-\imath\theta \cdot \mathbf{J}/\hbar)$. But this is easier said than done. The matrices J_i are infinite dimensional and exponentiating them is not practically possible. But the situation is not as bleak as it sounds for the following reason. First note that since J_i are block diagonal, so is the linear combination $\theta \cdot \mathbf{J}$, and so is its exponential. Consequently, *all* rotation operators $U[R]$ will be represented by block diagonal matrices. The $(2j+1)$ -dimensional block at a given j is denoted by $D^{(j)}[R]$. The block diagonal form of the rotation matrices implies (recall the mnemonic of images) that any vector $|\psi_j\rangle$ in the subspace \mathbb{V}_j spanned by the $(2j+1)$ vectors $|jj\rangle, \dots, |j-j\rangle$ goes into another element $|\psi'\rangle$ of \mathbb{V}_j . Thus to rotate $|\psi_j\rangle$, we just need the matrix $D^{(j)}$. More generally, if $|\psi\rangle$ has components only in $\mathbb{V}_0, \mathbb{V}_1, \mathbb{V}_2, \dots, \mathbb{V}_j$, we need just the first $(j+1)$ matrices $D^{(j)}$. What makes the situation hopeful is that it is possible, in practice, to evaluate these if j is small. Let us see why. Consider the series representing $D^{(j)}$:

$$D^{(j)}[R(\theta)] = \exp \left[-\frac{i\theta \cdot \mathbf{J}^{(j)}}{\hbar} \right] = \sum_0^{\infty} \left(\frac{-i\theta}{\hbar} \right)^n (\hat{\theta} \cdot \mathbf{J}^{(j)})^n \frac{1}{n!}$$

‡ The material from here to the end of Exercise 12.5.7 may be skimmed over in a less advanced course.

It can be shown (Exercise 12.5.4) that $(\hat{\theta} \cdot \mathbf{J}^{(j)})^n$ for $n > 2j$ can be written as a linear combination of the first $2j$ powers of $\hat{\theta} \cdot \mathbf{J}^{(j)}$. Consequently the series representing $D^{(j)}$ may be reduced to

$$D^{(j)} = \sum_0^{2j} f_n(\theta) (\hat{\theta} \cdot \mathbf{J}^{(j)})^n$$

It is possible, in practice, to find closed expressions for $f_n(\theta)$ in terms of trigonometric functions, for modest values of j (see Exercise 12.5.5). For example,

$$D^{(1/2)}[R] = \cos\left(\frac{\theta}{2}\right) - \frac{2i}{\hbar} \hat{\theta} \cdot \mathbf{J}^{(1/2)} \sin\left(\frac{\theta}{2}\right)$$

Let us return to the subspaces \mathbb{V}_j . Since they go into themselves under arbitrary rotations, they are called *invariant subspaces*. The physics behind the invariance is simple: each subspace contains states of a definite magnitude of angular momentum squared $j(j+1)\hbar^2$, and a rotation cannot change this. Formally it is because $[J^2, U[R]] = 0$ and so $U[R]$ cannot change the eigenvalue of J^2 .

The invariant subspaces have another feature: they are *irreducible*. This means that \mathbb{V}_j itself does not contain *invariant subspaces*. We prove this by showing that any invariant subspace $\bar{\mathbb{V}}_j$ of \mathbb{V}_j is as big as the latter. Let $|\psi\rangle$ be an element of $\bar{\mathbb{V}}_j$. Since we haven't chosen a basis yet, let us choose one such that $|\psi\rangle$ is one of the basis vectors, and furthermore, such that it is the basis vector $|jj\rangle$, up to a normalization factor, which is irrelevant in what follows. (What if we had already chosen a basis $|jj\rangle, \dots, |j, -j\rangle$ generated by the operators J_i ? Consider any unitary transformation U which converts $|jj\rangle$ into $|\psi\rangle$ and a different triplet of operators J'_i defined by $J'_i = U J_i U^\dagger$. The primed operators have the same commutation rules and hence eigenvalues as the J_i . The eigenvectors are just $|jm'\rangle = U|jm\rangle$, with $|jj'\rangle = |\psi\rangle$. In the following analysis we drop all primes.)

Let us apply an infinitesimal rotation $\delta\theta$ to $|\psi\rangle$. This gives

$$\begin{aligned} |\psi'\rangle &= U[R(\delta\theta)]|jj\rangle \\ &= [I - (i/\hbar)(\delta\theta \cdot \mathbf{J})]|jj\rangle \\ &= [I - (i/2\hbar)(\delta\theta_+ J_- + \delta\theta_- J_+ + 2\delta\theta_z J_z)]|jj\rangle \end{aligned}$$

where

$$\delta\theta_\pm = (\delta\theta_x \pm i\delta\theta_y)$$

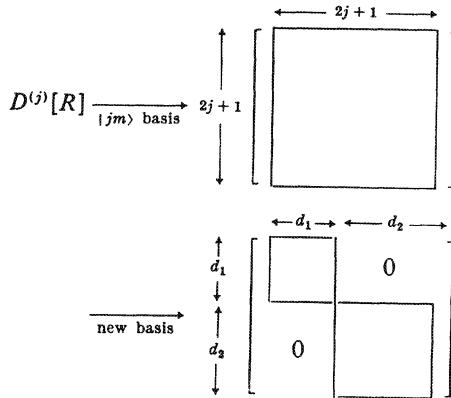
Since $J_+|jj\rangle = 0$, $J_z|jj\rangle = j\hbar|jj\rangle$, and $J_-|jj\rangle = \hbar(2j)^{1/2}|j, j-1\rangle$, we get

$$|\psi'\rangle = (1 - ij\delta\theta_z)|jj\rangle - \frac{1}{2}i(2j)^{1/2}\delta\theta_+|j, j-1\rangle$$

Since $\bar{\mathbb{V}}_j$ is assumed to be invariant under any rotation, $|\psi'\rangle$ also belongs to $\bar{\mathbb{V}}_j$. Subtracting $(1 - ij\delta\theta_z)|jj\rangle$, which also belongs to $\bar{\mathbb{V}}_j$, from $|\psi'\rangle$, we find that $|j, j-1\rangle$ also belongs to $\bar{\mathbb{V}}_j$. By considering more of such rotations, we can easily establish that the $(2j+1)$ orthonormal vectors, $|jj\rangle, |j, j-1\rangle, \dots, |j, -j\rangle$ all belong to $\bar{\mathbb{V}}_j$.

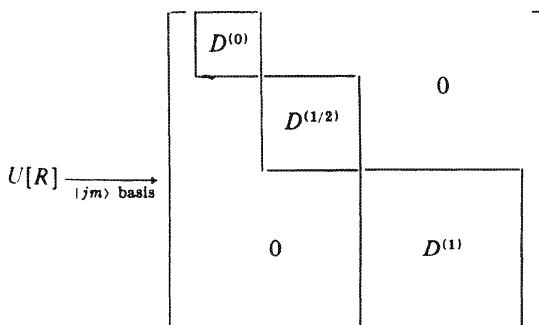
Thus $\bar{\mathbb{V}}_j$ has the same dimensionality as \mathbb{V}_j . Thus \mathbb{V}_j has no invariant subspaces. (In a technical sense, \mathbb{V}_j is its own subspace and is invariant. We are concerned here with subspaces of smaller dimensionality.)

The irreducibility of \mathbb{V}_j means that we cannot, by a change of basis within \mathbb{V}_j , further block diagonalize all the $D^{(j)}$. We show that if this were not true, then a contradiction would arise. Let it be possible to block diagonalize *all* the $D^{(j)}$, say, as follows:



(The boxed regions are generally nonzero). It follows that \mathbb{V}_j contains two invariant subspaces of dimensionalities d_1 and d_2 , respectively. (For example, any vector with just the first d_1 components nonzero will get rotated into another such vector. Such vectors form a d_1 -dimensional subspace.) We have seen this is impossible.

The block diagonal matrices representing the rotation operators $U[R]$ are said to provide an *irreducible (matrix) representation* of these operators. For the set of all rotation operators, the elements of which do not generally commute with each other, this irreducible form is the closest one can come to simultaneous diagonalization. All this is summarized schematically in the sketch below, where the boxed regions represent the blocks, $D^{(0)}$, $D^{(1)}$, ... etc. The unboxed regions contain zeros.



Consider next the matrix representing a rotationally invariant Hamiltonian in this basis. Since $[H, \mathbf{J}] = 0$, H has the same form as J^2 , which also commutes with

all the generators, namely,

- (1) H is diagonal, since $[H, J^2] = 0$, $[H, J_z] = 0$.
- (2) Within each block, H has the same eigenvalue E_j , since $[H, J_{\pm}] = 0$.

It follows from (2) that \mathbb{V}_j is an eigenspace of H with eigenvalue E_j , i.e., all states of a given j are degenerate in a rotationally invariant problem. Although the same result is true classically, the relation between degeneracy and rotational invariance is different in the two cases. Classically, if we are given two states with the same magnitude of angular momentum but different orientation, we argue that they are degenerate because

- (1) One may be rotated into the other.
- (2) This rotation does not change the energy.

Quantum mechanically, given two elements of \mathbb{V}_j , it is not always true that they may be rotated into each other (Exercise 12.5.6). However, we argue as follows:

- (1) One may be reached from the other (in general) by the combined action of J_{\pm} and $U[R]$.
- (2) These operators commute with H .

In short, rotational invariance is the cause of degeneracy in both cases, but the degenerate states are not always rotated versions of each other in the quantum case (Exercises 12.5.6 and 12.5.7).

*Exercise 12.5.4.** (1) Argue that the eigenvalues of $J_x^{(j)}$ and $J_y^{(j)}$ are the same as those of $J_z^{(j)}$, namely, $j\hbar, (j-1)\hbar, \dots, (-j)\hbar$. Generalize the result to $\hat{\theta} \cdot \mathbf{J}^{(j)}$.

(2) Show that

$$(J - j\hbar)(J - (j-1)\hbar)(J - (j-2)\hbar) \cdots (J + j\hbar) = 0$$

where $J \equiv \hat{\theta} \cdot \mathbf{J}^{(j)}$. (Hint: In the case $J = J_z$ what happens when both sides are applied to an arbitrary eigenket $|jm\rangle$? What about an arbitrary superpositions of such kets?)

(3) It follows from (2) that J^{2j+1} is a linear combination of J^0, J^1, \dots, J^{2j} . Argue that the same goes for J^{2j+k} , $k = 1, 2, \dots$.

Exercise 12.5.5. (Hard). Using results from the previous exercise and Eq. (12.5.23), show that

- (1) $D^{(1/2)}[R] = \exp(-i\hat{\theta} \cdot \mathbf{J}^{(1/2)}/\hbar) = \cos(\theta/2)I^{(1/2)} - (2i/\hbar)\sin(\theta/2)\hat{\theta} \cdot \mathbf{J}^{(1/2)}$
- (2) $D^{(1)}[R] = \exp(-i\theta_x J_x^{(1)}/\hbar) = (\cos \theta_x - 1) \left(\frac{J_x^{(1)}}{\hbar} \right)^2 - i \sin \theta_x \left(\frac{J_x^{(1)}}{\hbar} \right) + I^{(1)}$

Exercise 12.5.6. Consider the family of states $|jj\rangle, \dots, |jm\rangle, \dots, |j, -j\rangle$. One refers to them as states of the same magnitude but different orientation of angular momentum. If one takes this remark literally, i.e., in the classical sense, one is led to believe that one may rotate these into each other, as is the case for classical states with these properties. Consider, for

instance, the family $|1, 1\rangle, |1, 0\rangle, |1, -1\rangle$. It may seem, for example, that the state with zero angular momentum along the z axis, $|1, 0\rangle$, may be obtained by rotating $|1, 1\rangle$ by some suitable ($\frac{1}{2}\pi$?) angle about the x axis. Using $D^{(1)}[R(\theta_x \mathbf{i})]$ from part (2) in the last exercise show that

$$|1, 0\rangle \neq D^{(1)}[R(\theta_x \mathbf{i})]|1, 1\rangle \quad \text{for any } \theta_x$$

The error stems from the fact that classical reasoning should be applied to $\langle \mathbf{J} \rangle$, which responds to rotations like an ordinary vector, and not directly to $|jm\rangle$, which is a vector in Hilbert space. Verify that $\langle \mathbf{J} \rangle$ responds to rotations like its classical counterpart, by showing that $\langle \mathbf{J} \rangle$ in the state $D^{(1)}[R(\theta_x \mathbf{i})]|1, 1\rangle$ is $\hbar[-\sin \theta_x \mathbf{j} + \cos \theta_x \mathbf{k}]$.

It is not too hard to see why we can't always satisfy

$$|jm'\rangle = D^{(j)}[R]|jm\rangle$$

or more generally, for two normalized kets $|\psi'\rangle$ and $|\psi_j\rangle$, satisfy

$$|\psi_j'\rangle = D^{(j)}[R]|\psi_j\rangle$$

by any choice of R . These abstract equations imply $(2j+1)$ linear, complex relations between the components of $|\psi_j'\rangle$ and $|\psi_j\rangle$ that can't be satisfied by varying R , which depends on only three parameters, θ_x , θ_y , and θ_z . (Of course one can find a unitary matrix in \mathbb{V}_j that takes $|jm\rangle$ into $|jm'\rangle$ or $|\psi_j\rangle$ into $|\psi_j'\rangle$, but it will not be a *rotation* matrix corresponding to $U[R]$.)

Exercise 12.5.7: Euler Angles. Rather than parametrize an arbitrary rotation by the angle θ , which describes a *single* rotation by θ about an axis parallel to θ , we may parametrize it by three angles, γ , β , and α called *Euler angles*, which define three successive rotations:

$$U[R(\alpha, \beta, \gamma)] = e^{-i\alpha J_z/\hbar} e^{-i\beta J_y/\hbar} e^{-i\gamma J_z/\hbar}$$

(1) Construct $D^{(1)}[R(\alpha, \beta, \gamma)]$ explicitly as a product of three 3×3 matrices. (Use the result from Exercise 12.5.5 with $J_x \rightarrow J_y$.)

(2) Let it act on $|1, 1\rangle$ and show that $\langle \mathbf{J} \rangle$ in the resulting state is

$$\langle \mathbf{J} \rangle = \hbar(\sin \beta \cos \alpha \mathbf{i} + \sin \beta \sin \alpha \mathbf{j} + \cos \beta \mathbf{k})$$

(3) Show that for no value of α , β , and γ can one rotate $|1, 1\rangle$ into just $|1, 0\rangle$.

(4) Show that one can always rotate any $|1, m\rangle$ into a linear combination that involves $|1, m'\rangle$, i.e.,

$$\langle 1, m' | D^{(1)}[R(\alpha, \beta, \gamma)] | 1, m \rangle \neq 0$$

for some α , β , γ and any m, m' .

(5) To see that one can occasionally rotate $|jm\rangle$ into $|jm'\rangle$, verify that a 180° rotation about the y axis applied to $|1, 1\rangle$ turns it into $|1, -1\rangle$.

Angular Momentum Eigenfunctions in the Coordinate Basis

We now turn to step (3) outlined at the beginning of this section, namely, the construction of the eigenfunctions of L^2 and L_z in the coordinate basis, given the information on the kets $|lm\rangle$.

Consider the states corresponding to a given l . The “topmost” state $|ll\rangle$ satisfies

$$L_+|ll\rangle = 0 \quad (12.5.26)$$

If we write the operator $L_{\pm} = L_x \pm iL_y$ in spherical coordinates we find

$$L_{\pm} \xrightarrow[\text{coordinate basis}]{} \pm \hbar e^{\pm i\phi} \left(\frac{\partial}{\partial \theta} \pm i \cot \theta \frac{\partial}{\partial \phi} \right) \quad (12.5.27)$$

Exercise 12.5.8 (Optional). Verify that

$$L_x \xrightarrow[\text{coordinate basis}]{} i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cos \phi \cot \theta \frac{\partial}{\partial \phi} \right)$$

$$L_y \xrightarrow[\text{coordinate basis}]{} i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \sin \phi \cot \theta \frac{\partial}{\partial \phi} \right)$$

If we denote by $\psi'_l(r, \theta, \phi)$ the eigenfunction corresponding to $|ll\rangle$, we find that it satisfies

$$\left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) \psi'_l(r, \theta, \phi) = 0 \quad (12.5.28)$$

Since ψ'_l is an eigenfunction of L_z with eigenvalue $l\hbar$, we let

$$\psi'_l(r, \theta, \phi) = U'_l(r, \theta) e^{il\phi} \quad (12.5.29)$$

and find that

$$\left(\frac{\partial}{\partial \theta} - l \cot \theta \right) U'_l = 0 \quad (12.5.30)$$

$$\frac{dU'_l}{U'_l} = l \frac{d(\sin \theta)}{\sin \theta}$$

or

$$U'_l(r, \theta) = R(r)(\sin \theta)^l \quad (12.5.31)$$

where $R(r)$ is an arbitrary (normalizable) function of r . When we address the eigenvalue problem of rotationally invariant Hamiltonians, we will see that H will nail down R if we seek simultaneous eigenfunctions of H , L^2 , and L_z . But first let us introduce, as we did in the study of L_z in two dimensions, the function that would

have been the unique, nondegenerate solution in the absence of the radial coordinate:

$$Y_l^l(\theta, \phi) = (-1)^l \left[\frac{(2l+1)!}{4\pi} \right]^{1/2} \frac{1}{2^l l!} (\sin \theta)^l e^{il\phi} \quad (12.5.32)$$

Whereas the phase factor $(-1)^l$ reflects our convention, the others ensure that

$$\int |Y_l^l|^2 d\Omega \equiv \int_{-1}^1 \int_0^{2\pi} |Y_l^l|^2 d(\cos \theta) d\phi = 1 \quad (12.5.33)$$

We may obtain Y_l^{l-1} by using the lowering operator. Since

$$\begin{aligned} L_-|ll\rangle &= \hbar[(l+l)(1)]^{1/2}|l, l-1\rangle = \hbar(2l)^{1/2}|l, l-1\rangle \\ Y_l^{l-1}(\theta, \phi) &= \frac{1}{(2l)^{1/2}} \frac{(-1)}{\hbar} \left[\hbar e^{-i\phi} \left(\frac{\partial}{\partial \theta} - i \cot \theta \frac{\partial}{\partial \phi} \right) \right] Y_l^l \end{aligned} \quad (12.5.34)$$

We can keep going in this manner until we reach Y_l^{-1} . The result is, for $m \geq 0$,

$$\begin{aligned} Y_l^m(\theta, \phi) &= (-1)^l \left[\frac{(2l+1)!}{4\pi} \right]^{1/2} \frac{1}{2^l l!} \left[\frac{(l+m)!}{(2l)!(l-m)!} \right]^{1/2} e^{im\phi} (\sin \theta)^{-m} \\ &\times \frac{d^{l-m}}{d(\cos \theta)^{l-m}} (\sin \theta)^{2l} \end{aligned} \quad (12.5.35)$$

For $m < 0$, see Eq. (12.5.40). These functions are called *spherical harmonics* and satisfy the orthonormality condition

$$\int Y_l^m(\theta, \phi) Y_l^{m'}(\theta, \phi) d\Omega = \delta_{ll'} \delta_{mm'}$$

Another route to the Y_l^m is the direct solution of the L^2, L_z eigenvalue problem in the coordinate basis where

$$L^2 \rightarrow (-\hbar^2) \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) \quad (12.5.36)$$

and of course

$$L_z \rightarrow -i\hbar \frac{\partial}{\partial \phi}$$

If we seek common eigenfunctions of the form $\dagger f(\theta) e^{im\phi}$, which are regular between $\theta = 0$ and π , we will find that L^2 has eigenvalues of the form $l(l+1)\hbar^2$, $l = 0, 1, 2, \dots$,

\ddagger We neglect the function $R(r)$ that can tag along as a spectator.

where $l \geq |m|$. The Y_l^m functions are mutually orthogonal because they are *nondegenerate* eigenfunctions of L^2 and L_z , which are Hermitian on single-valued functions of θ and ϕ .

Exercise 12.5.9. Show that L^2 above is Hermitian in the sense

$$\int \psi_1^*(L^2\psi_2) d\Omega = \left[\int \psi_2^*(L^2\psi_1) d\Omega \right]^*$$

The same goes for L_z , which is insensitive to θ and is Hermitian with respect to the ϕ integration.

We may expand any $\psi(r, \theta, \phi)$ in terms of $Y_l^m(\theta, \phi)$ using *r-dependent coefficients* [consult Eq. (10.1.20) for a similar expansion]:

$$\psi(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l C_l^m(r) Y_l^m(\theta, \phi) \quad (12.5.37a)$$

where

$$C_l^m(r) = \int Y_l^m(\theta, \phi) \psi(r, \theta, \phi) d\Omega \quad (12.5.37b)$$

If we compute $\langle \psi | L^2 | \psi \rangle$ and interpret the result as a weighted average, we can readily see (assuming ψ is normalized to unity) that

$$P(L^2 = l(l+1)\hbar^2, L_z = m\hbar) = \int_0^r |C_l^m(r)|^2 r^2 dr \quad (12.5.38)$$

It is clear from the above that C_l^m is the amplitude to find the particle at a radial distance r with angular momentum (l, m) .[‡] The expansion Eq. (12.5.37a) tells us how to rotate any $\psi(r, \theta, \phi)$ by an angle θ (in principle):

- (1) We construct the block diagonal matrices, $\exp(-i\theta \cdot \mathbf{L}^{(l)}/\hbar)$.
- (2) Each block will rotate the C_l^m into linear combination of each other, i.e., under the action of $U[R]$, the coefficients $C_l^m(r)$, $m = 1, l-1, \dots, -l$; will get mixed with each other by $D_{m'm}^{(l)}$.

In practice, one can explicitly carry out these steps only if ψ contains only Y_l^m 's with small l . A concrete example will be provided in one of the exercises.

[‡] Note that r is just the eigenvalue of the operator $(X^2 + Y^2 + Z^2)^{1/2}$ which commutes with L^2 and L_z .

Here are the first few Y_l^m functions:

$$\begin{aligned}
 Y_0^0 &= (4\pi)^{-1/2} \\
 Y_1^{\pm 1} &= \mp (3/8\pi)^{1/2} \sin \theta e^{\pm i\phi} \\
 Y_1^0 &= (3/4\pi)^{1/2} \cos \theta \\
 Y_2^{\pm 2} &= (15/32\pi)^{1/2} \sin^2 \theta e^{\pm 2i\phi} \\
 Y_2^{\pm 1} &= \mp (15/8\pi)^{1/2} \sin \theta \cos \theta e^{\pm i\phi} \\
 Y_2^0 &= (5/16\pi)^{1/2} (3 \cos^2 \theta - 1)
 \end{aligned} \tag{12.5.39}$$

Note that

$$Y_l^{-m} = (-1)^m (Y_l^m)^* \tag{12.5.40}$$

Closely related to the spherical harmonics are the *associated Legendre polynomials* P_l^m (with $0 \leq m \leq l$) defined by

$$Y_l^m(\theta, \phi) = \left[\frac{(2l+1)(l-m)!}{4\pi(l+m)!} \right]^{1/2} (-1)^m e^{im\phi} P_l^m(\cos \theta) \tag{12.5.41}$$

If $m=0$, $P_l^0(\cos \theta) \equiv P_l(\cos \theta)$ is called a Legendre polynomial.

The Shape of the Y_l^m Functions. For large l , the functions $|Y_l^m|$ exhibit many classical features. For example, $|Y_l^l| \propto |\sin^l \theta|$, is almost entirely confined to the $x-y$ plane, as one would expect of a classical particle with all its angular momentum pointing along the z axis. Likewise, $|Y_l^0|$ is, for large l , almost entirely confined to the z axis. Polar plots of these functions may be found in many textbooks.

Exercise 12.5.10. Write the differential equation corresponding to

$$L^2|\alpha\beta\rangle = \alpha|\alpha\beta\rangle$$

in the coordinate basis, using the L^2 operator given in Eq. (12.5.36). We already know $\beta = m\hbar$ from the analysis of $-i\hbar(\partial/\partial\phi)$. So assume that the simultaneous eigenfunctions have the form

$$\psi_{am}(\theta, \phi) = P_a^m(\theta) e^{im\phi}$$

and show that P_a^m satisfies the equation

$$\left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{\alpha}{\hbar^2} - \frac{m^2}{\sin^2 \theta} \right) P_a^m(\theta) = 0$$

We need to show that

$$(1) \frac{\alpha}{\hbar^2} = l(l+1), \quad l=0, 1, 2, \dots$$

$$(2) |m| \leq l$$

We will consider only part (1) and that too for the case $m=0$. By rewriting the equation in terms of $u=\cos \theta$, show that P_α^0 satisfies

$$(1-u^2) \frac{d^2 P_\alpha^0}{du^2} - 2u \frac{dP_\alpha^0}{du} + \left(\frac{\alpha}{\hbar^2} \right) P_\alpha^0 = 0$$

Convince yourself that a power series solution

$$P_\alpha^0 = \sum_{n=0}^{\infty} C_n u^n$$

will lead to a two-term recursion relation. Show that $(C_{n+2}/C_n) \rightarrow 1$ as $n \rightarrow \infty$. Thus the series diverges when $|u| \rightarrow 1$ ($\theta \rightarrow 0$ or π). Show that if $\alpha/\hbar^2 = l(l+1)$; $l=0, 1, 2, \dots$, the series will terminate and be either an even or odd function of u . The functions $P_\alpha^0(u) = P_{l(l+1)\hbar^2}^0(u) \equiv P_l^0(u) \equiv P_l(u)$ are just the Legendre polynomials up to a scale factor. Determine P_0 , P_1 , and P_2 and compare (ignoring overall scales) with the Y_l^0 functions.

Exercise 12.5.11. Derive Y_1^1 starting from Eq.(12.5.28) and normalize it yourself. [Remember the $(-1)^l$ factor from Eq. (12.5.32).] Lower it to get Y_1^0 and Y_1^{-1} and compare it with Eq. (12.5.39).

*Exercise 12.5.12.** Since L_z^2 and L_z commute with Π , they should share a basis with it. Verify that under parity $Y_l^m \rightarrow (-1)^l Y_l^m$. (First show that $\theta \rightarrow \pi - \theta$, $\phi \rightarrow \phi + \pi$ under parity. Prove the result for Y_l^l . Verify that L_z does not alter the parity, thereby proving the result for all Y_l^m .)

*Exercise 12.5.13.** Consider a particle in a state described by

$$\psi = N(x+y+2z) e^{-ar}$$

where N is a normalization factor.

(1) Show, by rewriting the $Y_l^{\pm 1,0}$ functions in terms of x , y , z , and r , that

$$\begin{aligned} Y_1^{\pm 1} &= \mp \left(\frac{3}{4\pi} \right)^{1/2} \frac{x \pm iy}{2^{1/2}r} \\ Y_1^0 &= \left(\frac{3}{4\pi} \right)^{1/2} \frac{z}{r} \end{aligned} \tag{12.5.42}$$

(2) Using this result, show that for a particle described by ψ above, $P(l_z=0)=2/3$; $P(l_z=+\hbar)=1/6=P(l_z=-\hbar)$.

Exercise 12.5.14. Consider a rotation $\theta_x \mathbf{i}$. Under this

339

$$\begin{aligned}x &\rightarrow x \\y &\rightarrow y \cos \theta_x - z \sin \theta_x \\z &\rightarrow z \cos \theta_x + y \sin \theta_x\end{aligned}$$

ROTATION
INVARIANCE
AND ANGULAR
MOMENTUM

Therefore we must have

$$\psi(x, y, z) \xrightarrow{U[R(\theta_x \mathbf{i})]} \psi_R = \psi(x, y \cos \theta_x - z \sin \theta_x, z \cos \theta_x + y \sin \theta_x)$$

Let us verify this prediction for a special case

$$\psi = Az e^{-r^2/a^2}$$

which must go into

$$\psi_R = A(z \cos \theta_x - y \sin \theta_x) e^{-r^2/a^2}$$

- (1) Expand ψ in terms of Y_1^1 , Y_1^0 , Y_1^{-1} .
- (2) Use the matrix $e^{-i\theta_x L_z/\hbar}$ to find the fate of ψ under this rotation.[‡] Check your result against that anticipated above. [Hint: (1) $\psi \sim Y_1^0$, which corresponds to

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

- (2) Use Eq. (12.5.42).]

12.6. Solution of Rotationally Invariant Problems

We now consider a class of problems of great practical interest: problems where $V(r, \theta, \phi) = V(r)$. The Schrödinger equation in spherical coordinates becomes

$$\left[\frac{-\hbar^2}{2\mu} \left(\frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) + V(r) \right] \times \psi_E(r, \theta, \phi) = E \psi_E(r, \theta, \phi) \quad (12.6.1)$$

Since $[H, \mathbf{L}] = 0$ for a spherically symmetric potential, we seek simultaneous eigenfunctions of H , L^2 , and L_z :

$$\psi_{Elm}(r, \theta, \phi) = R_{Elm}(r) Y_l^m(\theta, \phi) \quad (12.6.2)$$

Feeding in this form, and bearing in mind that the angular part of ∇^2 is just the L^2 operator in the coordinate basis [up to a factor $(-\hbar^2 r^2)^{-1}$, see Eq. (12.5.36)], we get

[‡] See Exercise 12.5.5.

$$\left\{ -\frac{\hbar^2}{2\mu} \left[\frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} - \frac{l(l+1)}{r^2} \right] + V(r) \right\} R_{El} = E R_{El} \quad (12.6.3)$$

Notice that the subscript m has been dropped: neither the energy nor the radial function depends on it. We find, as anticipated earlier, the $(2l+1)$ -fold degeneracy of H .

*Exercise 12.6.1.** A particle is described by the wave function

$$\psi_E(r, \theta, \phi) = A e^{-r/a_0} \quad (a_0 = \text{const})$$

- (1) What is the angular momentum content of the state?
- (2) Assuming ψ_E is an eigenstate in a potential that vanishes as $r \rightarrow \infty$, find E . (Match leading terms in Schrödinger's equation.)
- (3) Having found E , consider finite r and find $V(r)$.

At this point it becomes fruitful to introduce an auxiliary function U_{El} defined as follows:

$$R_{El} = U_{El}/r \quad (12.6.4)$$

and which obeys the equation

$$\left\{ \frac{d^2}{dr^2} + \frac{2\mu}{\hbar^2} \left[E - V(r) - \frac{l(l+1)\hbar^2}{2\mu r^2} \right] \right\} U_{El} = 0 \quad (12.6.5)$$

*Exercise 12.6.2.** Provide the steps connecting Eq. (12.6.3) and Eq. (12.6.5).

The equation is the same as the one-dimensional Schrödinger equation except for the following differences:

- (1) The independent variable (r) goes from 0 to ∞ and not from $-\infty$ to ∞ .
 - (2) In addition to the actual potential $V(r)$, there is the *repulsive centrifugal barrier*, $l(l+1)\hbar^2/2\mu r^2$, in all but the $l=0$ states.
 - (3) The boundary conditions on U are different from the one-dimensional case.
- We find these by rewriting Eq. (12.6.5) as an eigenvalue equation

$$\left[-\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + V(r) + \frac{l(l+1)\hbar^2}{2\mu r^2} \right] U_{El} \equiv D_l(r) U_{El} = E U_{El} \quad (12.6.6)$$

and demanding that the functions U_{El} be such that D_l is Hermitian with respect to them. In other words, if U_1 and U_2 are two such functions, then we demand that

$$\int_0^\infty U_1^*(D_l U_2) dr = \left[\int_0^\infty U_2^*(D_l U_1) dr \right]^* \equiv \int_0^\infty (D_l U_1)^* U_2 dr \quad (12.6.7a)$$

This reduces to the requirement

$$\left(U_1^* \frac{dU_2}{dr} - U_2 \frac{dU_1^*}{dr} \right) \Big|_0^\infty = 0 \quad (12.6.7b)$$

→

Exercise 12.6.3. Show that Eq. (12.6.7b) follows from Eq. (12.6.7a).

Now, a necessary condition for

$$\int_0^\infty |R_{El}|^2 r^2 dr = \int_0^\infty |U_{El}|^2 dr$$

to be normalizable to unity or the Dirac delta function is that

$$U_{El} \xrightarrow[r \rightarrow \infty]{} 0 \quad (12.6.8a)$$

or

$$U_{El} \xrightarrow[r \rightarrow \infty]{} e^{ikr} \quad (12.6.8b)$$

the first corresponding to bound states and the second to unbound states. In either case, the expression in the brackets in Eq. (12.6.7b) vanishes at the upper limit‡ and the Hermiticity of D_l hinges on whether or not

$$\left[U_1^* \frac{dU_2}{dr} - U_2 \frac{dU_1^*}{dr} \right]_0 = 0 \quad (12.6.9)$$

Now this condition is satisfied if

$$U \xrightarrow[r \rightarrow 0]{} c, \quad c = \text{const} \quad (12.6.10)$$

‡ For the oscillating case, we must use the limiting scheme described in Section 1.10.

If c is nonzero, then

$$R \sim \frac{U}{r} \sim \frac{c}{r}$$

diverges at the origin. This in itself is not a disqualification, for R is still square integrable. The problem with $c \neq 0$ is that the corresponding total wave function[‡]

$$\psi \sim \frac{c}{r} Y_0^0$$

does not satisfy Schrödinger's equation at the origin. This is because of the relation

$$\nabla^2(1/r) = -4\pi\delta^3(\mathbf{r}) \quad (12.6.11)$$

the proof of which is taken up in Exercise 12.6.4. Thus unless $V(r)$ contains a delta function at the origin (which we assume it does not) the choice $c \neq 0$ is untenable. Thus we deduce that

$$U_{El} \xrightarrow[r \rightarrow 0]{} 0 \quad (12.6.12)$$

*Exercise 12.6.4.** (1) Show that

$$\delta^3(\mathbf{r} - \mathbf{r}') \equiv \delta(x - x')\delta(y - y')\delta(z - z') = \frac{1}{r^2 \sin \theta} \delta(r - r')\delta(\theta - \theta')\delta(\phi - \phi')$$

(consider a test function).

(2) Show that

$$\nabla^2(1/r) = -4\pi\delta^3(\mathbf{r})$$

(Hint: First show that $\nabla^2(1/r) = 0$ if $r \neq 0$. To see what happens at $r = 0$, consider a small sphere centered at the origin and use Gauss's law and the identity $\nabla^2\phi = \mathbf{V} \cdot \nabla\phi$).[§]

General Properties of U_{El}

We have already discussed some of the properties of U_{El} as $r \rightarrow 0$ or ∞ . We shall try to extract further information on U_{El} by analyzing the equation governing it in these limits, without making detailed assumptions about $V(r)$. Consider first the limit $r \rightarrow 0$. Assuming $V(r)$ is less singular than r^{-2} , the equation is dominated by the

[‡] As we will see in a moment, $l \neq 0$ is incompatible with the requirement that $\psi(\mathbf{r}) \rightarrow r^{-1}$ as $r \rightarrow 0$. Thus the angular part of ψ has to be $Y_0^0 = (4\pi)^{-1/2}$.

[§] Or compare this equation to Poisson's equation in electrostatics $\nabla^2\phi = -4\pi\rho$. Here $\rho = \delta^3(\mathbf{r})$, which represents a unit point charge at the origin. In this case we know from Coulomb's law that $\phi = 1/r$.

centrifugal barrier:

$$U_l'' \approx \frac{l(l+1)}{r^2} U_l \quad (12.6.13)$$

We have dropped the subscript E , since E becomes inconsequential in this limit. If we try a solution of the form

$$U_l \sim r^\alpha$$

we find

$$\alpha(\alpha - 1) = l(l+1)$$

or

$$\alpha = l+1 \quad \text{or } (-l)$$

and

$$U_l \sim \begin{cases} r^{l+1} & (\text{regular}) \\ r^{-l} & (\text{irregular}) \end{cases} \quad (12.6.14)$$

We reject the irregular solution since it does not meet the boundary condition $U(0) = 0$. The behavior of the regular solutions near the origin is in accord with our expectation that as the angular momentum increases the particle should avoid the origin more and more.

The above arguments are clearly true only if $l \neq 0$. If $l=0$, the centrifugal barrier is absent, and the answer may be sensitive to the potential. In the problems we will consider, $U_{l=0}$ will also behave as r^{l+1} with $l=0$. Although $U_0(r) \rightarrow 0$ as $r \rightarrow 0$, note that a particle in the $l=0$ state has a nonzero amplitude to be at the origin, since $R_0(r) = U_0(r)/r \neq 0$ at $r=0$.

Consider now the behavior of U_E as $r \rightarrow \infty$. If $V(r)$ does not vanish as $r \rightarrow \infty$, it will dominate the result (as in the case of the isotropic oscillator, for which $V(r) \propto r^2$) and we cannot say anything in general. So let us consider the case where $rV(r) \rightarrow 0$ as $r \rightarrow \infty$. At large r the equation becomes

$$\frac{d^2 U_E}{dr^2} = -\frac{2\mu E}{\hbar^2} U_E \quad (12.6.15)$$

(We have dropped the subscript l since the answer doesn't depend on l .) There are now two cases:

1. $E > 0$: the particle is allowed to escape to infinity classically. We expect U_E to oscillate as $r \rightarrow \infty$.
2. $E < 0$: The particle is bound. The region $r \rightarrow \infty$ is classically forbidden and we expect U_E to fall exponentially there.

Consider the first case. The solutions to Eq. (12.6.15) are of the form

$$U_E = A e^{ikr} + B e^{-ikr}, \quad k = (2\mu E/\hbar^2)^{1/2}$$

that is to say, the particle behaves as a free particle far from the origin.[‡] Now, you might wonder why we demanded that $rV(r) \rightarrow 0$ and not simply $V(r) \rightarrow 0$ as $r \rightarrow \infty$. To answer this question, let us write

$$U_E = f(r) e^{\pm ikr}$$

and see if $f(r)$ tends to a constant as $r \rightarrow \infty$. Feeding in this form of U_E into Eq. (12.6.5) we find (ignoring the centrifugal barrier)

$$f'' \pm (2ik)f' - \frac{2\mu V(r)}{\hbar^2} f = 0$$

Since we expect $f(r)$ to be slowly varying as $r \rightarrow \infty$, we ignore f'' and find

$$\begin{aligned} \frac{df}{f} &= \mp \frac{i}{k} \frac{\mu}{\hbar^2} V(r) dr \\ f(r) &= f(r_0) \cdot \exp \mp \left[\frac{i\mu}{k\hbar^2} \int_{r_0}^r V(r') dr' \right] \end{aligned} \quad (12.6.16)$$

where r_0 is some constant. If $V(r)$ falls faster than r^{-1} , i.e., $rV(r) \rightarrow 0$ as $r \rightarrow \infty$, we can take the limit as $r \rightarrow \infty$ in the integral and $f(r)$ approaches a constant as $r \rightarrow \infty$. If instead

$$V(r) = -\frac{e^2}{r}$$

as in the Coulomb problem,[§] then

$$f(r) = f(r_0) \exp \pm \left[\frac{i\mu e^2}{k\hbar^2} \ln \left(\frac{r}{r_0} \right) \right]$$

and

$$U_E(r) \sim \exp \pm \left[i \left(kr + \frac{\mu e^2}{k\hbar^2} \ln r \right) \right] \quad (12.6.17)$$

This means that no matter how far away the particle is from the origin, it is never completely free of the Coulomb potential. If $V(r)$ falls even slower than a Coulomb potential, this problem only gets worse.

[‡] Although A and B are arbitrary in this asymptotic form, their ratio is determined by the requirement that if U_E is continued inward to $r=0$, it must vanish. That there is just one free parameter in the solution (the overall scale), and not two, is because D_i is nondegenerate even for $E>0$, which in turn is due to the constraint $U_E(r=0)=0$; see Exercise 12.6.5.

[§] We are considering the case of equal and opposite charges with an eye on the next chapter.

Consider now the case $E < 0$. All the results from the $E > 0$ case carry over with the change

$$k \rightarrow i\kappa, \quad \kappa = (2\mu|E|/\hbar^2)^{1/2}$$

Thus

$$U_E \xrightarrow[r \rightarrow \infty]{} A e^{-\kappa r} + B e^{+\kappa r} \quad (12.6.18)$$

Again B/A is not arbitrary if we demand that U_E continued inward vanish at $r=0$. Now, the growing exponential is disallowed. For arbitrary $E < 0$, both $e^{\kappa r}$ and $e^{-\kappa r}$ will be present in U_E . Only for certain discrete values of E will the $e^{\kappa r}$ piece be absent; these will be the allowed bound state levels. (If A/B were arbitrary, we could choose $B=0$ and get a normalizable bound state for every $E < 0$.)

As before, Eq. (12.6.18) is true only if $rV(r) \rightarrow 0$. In the Coulomb case we expect [from Eq. (12.6.17) with $k \rightarrow i\kappa$]

$$\begin{aligned} U_E &\sim \exp\left(\pm \frac{\mu e^2}{\kappa \hbar^2} \ln r\right) e^{\mp \kappa r} \\ &= (r)^{\pm \mu e^2 / \kappa \hbar^2} e^{\mp \kappa r} \end{aligned} \quad (12.6.19)$$

When we solve the problem of the hydrogen atom, we will find that this is indeed the case.

When $E < 0$, the energy eigenfunctions are normalizable to unity. As the operator $D_l(r)$ is nondegenerate (Exercise 12.6.5), we have

$$\int_0^\infty U_{E'l}(r) U_{El}(r) dr = \delta_{EE'}$$

and

$$\psi_{Elm}(r, \theta, \phi) = R_{El}(r) Y_l^m(\theta, \phi)$$

obeys

$$\iiint \psi_{Elm}^*(r, \theta, \phi) \psi_{E'l'm'}(r, \theta, \phi) r^2 dr d\Omega = \delta_{EE'} \delta_{ll'} \delta_{mm'}$$

We will consider the case $E > 0$ in a moment.

Exercise 12.6.5. Show that D_l is nondegenerate in the space of functions U that vanish as $r \rightarrow 0$. (Recall the proof of Theorem 15, Section 5.6.) Note that U_{El} is nondegenerate even for $E > 0$. This means that E , l , and m , label a state fully in three dimensions.

The Free Particle in Spherical Coordinates[‡]

If we begin as usual with

$$\psi_{Elm}(r, \theta, \phi) = R_{El}(r) Y_l^m(\theta, \phi)$$

and switch to U_{El} , we end up with

$$\left[\frac{d^2}{dr^2} + k^2 - \frac{l(l+1)}{r^2} \right] U_{El} = 0, \quad k^2 = \frac{2\mu E}{\hbar^2}$$

Dividing both sides by k^2 , and changing to $\rho = kr$, we obtain

$$\left[-\frac{d^2}{d\rho^2} + \frac{l(l+1)}{\rho^2} \right] U_l = U_l \quad (12.6.20)$$

The variable k , which has disappeared, will reappear when we rewrite the answer in terms of $r = \rho/k$. This problem looks a lot like the harmonic oscillator except for the fact that we have a potential $1/\rho^2$ instead of ρ^2 . So we define operators analogous to the raising and lowering operators. These are

$$d_l = \frac{d}{d\rho} + \frac{l+1}{\rho} \quad (12.6.21a)$$

and its adjoint

$$d_l^\dagger = -\frac{d}{d\rho} + \frac{l+1}{\rho} \quad (12.6.21b)$$

(Note that $d/d\rho$ is anti-Hermitian.) In terms of these, Eq. (12.6.20) becomes

$$(d_l d_l^\dagger) U_l = U_l \quad (12.6.22)$$

Now we premultiply both sides by d_l^\dagger to get

$$d_l^\dagger d_l (d_l^\dagger U_l) = d_l^\dagger U_l \quad (12.6.23)$$

You may verify that

$$d_l^\dagger d_l = d_{l+1} d_{l+1}^\dagger \quad (12.6.24)$$

so that

$$d_{l+1} d_{l+1}^\dagger (d_l^\dagger U_l) = d_l^\dagger U_l \quad (12.6.25)$$

[‡] The present analysis is a simplified version of the work of L. Infeld, *Phys. Rev.*, **59**, 737 (1941).

It follows that

$$d_l^\dagger U_l = c_l U_{l+1} \quad (12.6.26)$$

where c_l is a constant. We choose it to be unity, for it can always be absorbed in the normalization. We see that d_l^\dagger serves as a “raising operator” in the index l . Given U_0 , we can find the others.[‡] From Eq. (12.6.20) it is clear that if $l=0$ there are two independent solutions:

$$U_0^A(\rho) = \sin \rho, \quad U_0^B = -\cos \rho \quad (12.6.27)$$

The constants in front are chosen according to a popular convention. Now U_0^B is unacceptable at $\rho=0$ since it violates Eq. (12.6.12). If, however, one is considering the equation in a region that excludes the origin, U_0^B must be included. Consider now the tower of solutions built out of U_0^A and U_0^B . Let us begin with the equation

$$U_{l+1} = d_l^\dagger U_l \quad (12.6.28)$$

Now, we are really interested in the functions $R_l = U_l/\rho$.[§] These obey (from the above)

$$\begin{aligned} \rho R_{l+1} &= d_l^\dagger (\rho R_l) \\ &= \left(-\frac{d}{d\rho} + \frac{l+1}{\rho} \right) (\rho R_l) \\ R_{l+1} &= \left(-\frac{d}{d\rho} + \frac{l}{\rho} \right) R_l \\ &= \rho' \left(-\frac{d}{d\rho} \right) \frac{R_l}{\rho'} \end{aligned}$$

or

$$\begin{aligned} \frac{R_{l+1}}{\rho'^{l+1}} &= \left(-\frac{1}{\rho} \frac{d}{d\rho} \right) \frac{R_l}{\rho'} \\ &= \left(-\frac{1}{\rho} \frac{d}{d\rho} \right)^2 \frac{R_{l-1}}{\rho'^{l-1}} \\ &= \left(-\frac{1}{\rho} \frac{d}{d\rho} \right)^{l+1} \frac{R_0}{\rho^0} \end{aligned}$$

[‡] In Chapter 15, we will gain some insight into the origin of such a ladder of solutions.

[§] Actually we want $R_l = U_l/r = k U_l/\rho$. But the factor k may be absorbed in the normalization factors of U and R .

so that finally we have

$$R_l = (-\rho)^l \left(\frac{1}{\rho} \frac{\partial}{\partial \rho} \right)^l R_0 \quad (12.6.29)$$

Now there are two possibilities for R_0 :

$$R_0^A = \frac{\sin \rho}{\rho}$$

$$R_0^B = \frac{-\cos \rho}{\rho}$$

These generate the functions

$$R_l^A \equiv j_l = (-\rho)^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \left(\frac{\sin \rho}{\rho} \right) \quad (12.6.30a)$$

called the *spherical Bessel functions* of order l , and

$$R_l^B \equiv n_l = (-\rho)^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \left(\frac{-\cos \rho}{\rho} \right) \quad (12.6.30b)$$

called *spherical Neumann functions* of order l .[‡] Here are a few of these functions:

$$\begin{aligned} j_0(\rho) &= \frac{\sin \rho}{\rho}, & n_0(\rho) &= \frac{-\cos \rho}{\rho} \\ j_1(\rho) &= \frac{\sin \rho}{\rho^2} - \frac{\cos \rho}{\rho}, & n_1(\rho) &= \frac{-\cos \rho}{\rho^2} - \frac{\sin \rho}{\rho} \\ j_2(\rho) &= \left(\frac{3}{\rho^3} - \frac{1}{\rho} \right) \sin \rho - \frac{3 \cos \rho}{\rho^2}, & n_2(\rho) &= -\left(\frac{3}{\rho^3} - \frac{1}{\rho} \right) \cos \rho - \frac{3 \sin \rho}{\rho^2} \end{aligned} \quad (12.6.31)$$

As $\rho \rightarrow \infty$, these functions behave as

$$j_l \xrightarrow[\rho \rightarrow \infty]{} \frac{1}{\rho} \sin \left(\rho - \frac{l\pi}{2} \right) \quad (12.6.32)$$

$$n_l \xrightarrow[\rho \rightarrow \infty]{} -\frac{1}{\rho} \cos \left(\rho - \frac{l\pi}{2} \right) \quad (12.6.32)$$

Despite the apparent singularities as $\rho \rightarrow 0$, the $j_l(l)$ functions are finite and in fact

$$j_l(\rho) \xrightarrow[\rho \rightarrow 0]{} \frac{\rho^l}{(2l+1)!!} \quad (12.6.33)$$

[‡] One also encounters *spherical Hankel functions* $h_l = j_l + in_l$ in some problems.

where $(2l+1)!! = (2l+1)(2l-1)(2l-3)\dots(5)(3)(1)$. These are just the regular solutions listed in Eq. (12.6.14). The Neumann functions, on the other hand, are singular

$$n_l(\rho) \xrightarrow[\rho \rightarrow 0]{} -\frac{(2l-1)!!}{\rho^{l+1}} \quad (12.6.34)$$

and correspond to the irregular solutions listed in Eq. (12.6.14).

Free-particle solutions that are *regular in all space* are then

$$\psi_{Elm}(r, \theta, \phi) = j_l(kr) Y_l^m(\theta, \phi), \quad E = \frac{\hbar^2 k^2}{2\mu} \quad (12.6.35)$$

These satisfy

$$\iiint \psi_{Elm}^* \psi_{E'l'm'} r^2 dr d\Omega = \frac{2}{\pi k^2} \delta(k - k') \delta_{ll'} \delta_{mm'} \quad (12.6.36)$$

We are using here the fact that

$$\int_0^\infty j_l(kr) j_l(k'r) r^2 dr = \frac{\pi}{2k^2} \delta(k - k') \quad (12.6.37)$$

*Exercise 12.6.6.** (1) Verify that Eqs. (12.6.21) and (12.6.22) are equivalent to Eq. (12.6.20)

(2) Verify Eq. (12.6.24).

Exercise 12.6.7. Verify that j_0 and j_1 have the limits given by Eq. (12.6.33).

*Exercise 12.6.8.** Find the energy levels of a particle in a spherical box of radius r_0 in the $l=0$ sector.

*Exercise 12.6.9.** Show that the quantization condition for $l=0$ bound states in a spherical well of depth $-V_0$ and radius r_0 is

$$k'/\kappa = -\tan k'r_0$$

where k' is the wave number inside the well and $i\kappa$ is the complex wave number for the exponential tail outside. Show that there are no bound states for $V_0 < \pi^2 \hbar^2 / 8\mu r_0^2$. (Recall Exercise 5.2.6.)

Connection with the Solution in Cartesian Coordinates

If we had attacked the free-particle problem in Cartesian coordinates, we would have readily obtained

$$\psi_E(x, y, z) = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p} \cdot \mathbf{r}/\hbar}, \quad E = \frac{p^2}{2\mu} = \frac{\hbar^2 k^2}{2\mu} \quad (12.6.38)$$

Consider now the case which corresponds to a particle moving along the z axis with momentum p . As

$$\mathbf{p} \cdot \mathbf{r}/\hbar = (pr \cos \theta)\hbar = kr \cos \theta$$

we get

$$\psi_E(r, \theta, \phi) = \frac{e^{ikr \cos \theta}}{(2\pi\hbar)^{3/2}}, \quad E = \frac{\hbar^2 k^2}{2\mu} \quad (12.6.39)$$

It should be possible to express this solution, describing a particle moving in the z direction with energy $E = \hbar^2 k^2 / 2\mu$, as a linear combination of the functions ψ_{Elm} which have the same energy, or equivalently, the same k :

$$e^{ikr \cos \theta} = \sum_{l=0}^{\infty} \sum_{m=-l}^l C_l^m j_l(kr) Y_l^m(\theta, \phi) \quad (12.6.40)$$

Now, only terms with $m=0$ are relevant since the left-hand side is independent of ϕ . Physically this means that a particle moving along the z axis has no angular momentum in that direction. Since we have

$$Y_l^0(\theta) = \left(\frac{2l+1}{4\pi} \right)^{1/2} P_l(\cos \theta)$$

$$e^{ikr \cos \theta} = \sum_{l=0}^{\infty} C_l j_l(kr) P_l(\cos \theta), \quad C_l = C_l^0 \cdot \left(\frac{2l+1}{4\pi} \right)^{1/2}$$

It can be shown that

$$C_l = i^l (2l+1)$$

so that

$$e^{ikr \cos \theta} = \sum_{l=0}^{\infty} i^l (2l+1) j_l(kr) P_l(\cos \theta) \quad (12.6.41)$$

This relation will come in handy when we study scattering. This concludes our study of the free particle.

Exercise 12.6.10. (Optional). Verify Eq. (12.6.41) given that

$$(1) \int_{-1}^1 P_l(\cos \theta) P_l(\cos \vartheta) d(\cos \theta) = [2/(2l+1)] \delta_{ll'}$$

$$(2) P_l(x) = \frac{1}{2^l l!} \frac{d^l (x^2 - 1)^l}{dx^l}$$

$$(3) \int_0^1 (1 - x^2)^m dx = \frac{(2m)!!}{(2m+1)!!}$$

Hint: Consider the limit $kr \rightarrow 0$ after projecting out C_l .

We close this section on rotationally invariant problems with a brief study of the isotropic oscillator. The most celebrated member of this class, the hydrogen atom, will be discussed in detail in the next chapter.

The Isotropic Oscillator

The isotropic oscillator is described by the Hamiltonian

$$H = \frac{P_x^2 + P_y^2 + P_z^2}{2\mu} + \frac{1}{2} \mu \omega^2 (X^2 + Y^2 + Z^2) \quad (12.6.42)$$

If we write as usual

$$\psi_{Elm} = \frac{U_{El}(r)}{r} Y_l^m(\theta, \phi) \quad (12.6.43)$$

we obtain the radial equation

$$\left\{ \frac{d^2}{dr^2} + \frac{2\mu}{\hbar^2} \left[E - \frac{1}{2} \mu \omega^2 r^2 - \frac{l(l+1)\hbar^2}{2\mu r^2} \right] \right\} U_{El} = 0 \quad (12.6.44)$$

As $r \rightarrow \infty$, we find

$$U \sim e^{-y^2/2} \quad (12.6.45)$$

where

$$y = \left(\frac{\mu \omega}{\hbar} \right)^{1/2} r \quad (12.6.46)$$

is dimensionless. So we let

$$U(y) = e^{-y^2/2} v(y) \quad (12.6.47)$$

and obtain the following equation for $v(y)$:

$$v'' - 2yv' + \left[2\lambda - 1 - \frac{l(l+1)}{y^2} \right] v = 0, \quad \lambda = \frac{E}{\hbar \omega} \quad (12.6.48)$$

It is clear upon inspection that a two-term recursion relation will obtain if a power-series solution is plugged in. We set

$$v(y) = y^{l+1} \sum_{n=0}^{\infty} C_n y^n \quad (12.6.49)$$

where we have incorporated the known behavior [Eq. (12.6.14)] near the origin.

By going through the usual steps (left as an exercise) we can arrive at the following quantization condition:

$$E = (2k + l + 3/2)\hbar\omega, \quad k = 0, 1, 2, \dots \quad (12.6.50)$$

If we define the principal quantum number (which controls the energy)

$$n = 2k + l \quad (12.6.51)$$

we get

$$E = (n + 3/2)\hbar\omega \quad (12.6.52)$$

At each n , the allowed l values are

$$l = n - 2k = n, n - 2, \dots, 1 \text{ or } 0 \quad (12.6.53)$$

Here are the first few eigenstates:

$$\begin{array}{lll} n=0 & l=0 & m=0 \\ n=1 & l=1 & m=\pm 1, 0 \\ n=2 & l=0, 2 & m=0; \pm 2, \pm 1, 0 \\ n=3 & l=1, 3 & m=\pm 1, 0; \pm 3, \pm 2, \pm 1, 0 \\ \vdots & \vdots & \vdots \end{array}$$

Of particular interest to us is the fact that states of different l are degenerate. The degeneracy in m at each l we understand in terms of rotational invariance. The degeneracy of the different l states (which are not related by rotation operators or the generators) appears mysterious. For this reason it is occasionally termed *accidental degeneracy*. This is, however, a misnomer, for the degeneracy in l can be attributed to additional invariance properties of H . Exactly what these extra invariances or symmetries of H are, and how they explain the degeneracy in l , we will see in Chapter 15.

*Exercise 12.6.11.** (1) By combining Eqs. (12.6.48) and (12.6.49) derive the two-term recursion relation. Argue that $C_0 \neq 0$ if U is to have the right properties near $y=0$. Derive the quantizations condition, Eq. (12.6.50).

(2) Calculate the degeneracy and parity at each n and compare with Exercise 10.2.3, where the problem was solved in Cartesian coordinates.

(3) Construct the normalized eigenfunction ψ_{nlm} for $n=0$ and 1. Write them as linear combinations of the $n=0$ and $n=1$ eigenfunctions obtained in Cartesian coordinates.

The Hydrogen Atom

13.1. The Eigenvalue Problem

We have here a two-body problem, of an electron of charge $-e$ and mass m , and a proton of charge $+e$ and mass M . By using CM and relative coordinates and working in the CM frame, we can reduce the problem to the dynamics of a single particle whose mass $\mu = mM/(m+M)$ is the reduced mass and whose coordinate \mathbf{r} is the relative coordinate of the two particles. However, since $m/M \approx 1/2000$, as a result of which the relative coordinate is essentially the electron's coordinate and the reduced mass is essentially m , let us first solve the problem in the limit $M \rightarrow \infty$. In this case we have just the electron moving in the field of the immobile proton. At a later stage, when we compare the theory with experiment, we will see how we can easily take into account the finiteness of the proton mass.

Since the potential energy of the electron in the Coulomb potential

$$\phi = e/r \quad (13.1.1)$$

due to the proton is $V = -e^2/r$, the Schrödinger equation

$$\left\{ \frac{d^2}{dr^2} + \frac{2m}{\hbar^2} \left[E + \frac{e^2}{r} - \frac{l(l+1)\hbar^2}{2mr^2} \right] \right\} U_{El} = 0 \quad (13.1.2)$$

determines the energy levels in the rest frame of the atom, as well as the wave functions‡

$$\psi_{Elm}(r, \theta, \phi) = R_{El}(r) Y_l^m(\theta, \phi) = \frac{U_{El}(r)}{r} Y_l^m(\theta, \phi) \quad (13.1.3)$$

It is clear upon inspection of Eq. (13.1.2) that a power series ansatz will lead to a three-term recursion relation. So we try to factor out the asymptotic behavior.

‡ It should be clear from the context whether m stands for the electron mass or the z component of angular momentum.

We already know from Section 12.6 that up to (possibly fractional) powers of r [Eq. (12.6.19)],

$$U_{El} \underset{r \rightarrow \infty}{\sim} \exp[-(2mW/\hbar^2)^{1/2}r] \quad (13.1.4)$$

where

$$W = -E$$

is the *binding energy* (which is the energy it would take to liberate the electron) and that

$$U_{El} \underset{r \rightarrow 0}{\sim} r^{l+1} \quad (13.1.5)$$

Equation (13.1.4) suggests the introduction of the dimensionless variable

$$\rho = (2mW/\hbar^2)^{1/2}r \quad (13.1.6)$$

and the auxiliary function v_{El} defined by

$$U_{El} = e^{-\rho} v_{El} \quad (13.1.7)$$

The equation for v is then

$$\frac{d^2v}{d\rho^2} - 2 \frac{dv}{d\rho} + \left[\frac{e^2\lambda}{\rho} - \frac{l(l+1)}{\rho^2} \right] v = 0 \quad (13.1.8)$$

where

$$\lambda = (2m/\hbar^2 W)^{1/2} \quad (13.1.9)$$

and the subscripts on v are suppressed. You may verify that if we feed in a series into Eq. (13.1.8), a two-term recursion relation will obtain. Taking into account the behavior near $\rho = 0$ [Eq. (13.1.5)] we try

$$v_{El} = \rho^{l+1} \sum_{k=0}^{\infty} C_k \rho^k \quad (13.1.10)$$

and obtain the following recursion relation between *successive* coefficients:

$$\frac{C_{k+1}}{C_k} = \frac{-e^2\lambda + 2(k+l+1)}{(k+l+2)(k+l+1) - l(l+1)} \quad (13.1.11)$$

The Energy Levels

Since

$$\frac{C_{k+1}}{C_k} \xrightarrow{k \rightarrow \infty} \frac{2}{k} \quad (13.1.12)$$

is the behavior of the series $\rho^m e^{2\rho}$, and would lead to $U \sim e^{-\rho} v \sim \rho^m e^{-\rho} e^{2\rho} \sim \rho^m e^\rho$ as $\rho \rightarrow \infty$, we demand that the series terminate at some k . This will happen if

$$e^2 \lambda = 2(k + l + 1) \quad (13.1.13)$$

or [from Eq. (13.1.9)]

$$E = -W = \frac{-me^4}{2\hbar^2(k+l+1)^2}, \quad k=0, 1, 2, \dots : \quad l=0, 1, 2, \dots \quad (13.1.14)$$

In terms of the *principal quantum number*

$$n = k + l + 1 \quad (13.1.15)$$

the allowed energies are

$$E_n = \frac{-me^4}{2\hbar^2 n^2}, \quad n=1, 2, 3, \dots \quad (13.1.16)$$

and at each n the allowed values of l are, according to Eq. (13.1.15),

$$l = n - k - 1 = n - 1, n - 2, \dots, 1, 0 \quad (13.1.17)$$

That states of different l should be degenerate indicates that H contains more symmetries besides rotational invariance. We discuss these later. For the present, let us note that the degeneracy at each n is

$$\sum_{l=0}^{n-1} (2l+1) = n^2 \quad (13.1.18)$$

It is common to refer to the states with $l=0, 1, 2, 3, 4, \dots$ as s, p, d, f, g, h, \dots states. In this *spectroscopic notation*, $1s$ denotes the state ($n=1, l=0$); $2s$ and $2p$ the $l=0$ and $l=1$ states at $n=2$; $3s, 3p$, and $3d$ the $l=0, 1$, and 2 states at $n=3$, and so on. No attempt is made to keep track of m .

It is convenient to employ a natural unit of energy, called a *Rydberg* (Ry), for measuring the energy levels of hydrogen:

$$Ry = \frac{me^4}{2\hbar^2} \quad (13.1.19)$$

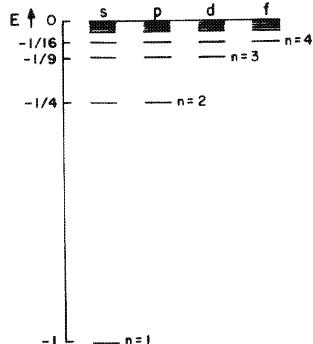


Figure 13.1. The first few eigenstates of hydrogen. The energy is measured in Rydbergs and the states are labelled in the spectroscopic notation.

in terms of which

$$E_n = \frac{-Ry}{n^2} \quad (13.1.20)$$

Figure 13.1 shows some of the lowest-energy states of hydrogen.

The Wave Functions

Given the recursion relations, it is a straightforward matter to determine the wave functions and to normalize them. Consider a given n and l . Since the series in Eq. (13.1.10) terminates at

$$k = n - l - 1 \quad (13.1.21)$$

the corresponding function v_l is ρ^{l+1} times a polynomial of degree $n-l-1$. This polynomial is called the *associated Laguerre polynomial*, $L_{n-l-1}^{2l+1}(2\rho)$.[‡] The corresponding radial function is

$$R_{nl}(\rho) \sim e^{-\rho} \rho^l L_{n-l-1}^{2l+1}(2\rho) \quad (13.1.22)$$

Recall that

$$\begin{aligned} \rho &= \left(\frac{2mW}{\hbar^2} \right)^{1/2} r = \left[\frac{2m}{\hbar^2} \left(\frac{me^4}{2\hbar^2 n^2} \right) \right]^{1/2} r \\ &= \frac{me^2}{\hbar^2 n} r \end{aligned} \quad (13.1.23)$$

[‡] $L_p^k(x) = (-1)^k (d^k/dx^k) L_p^0(x)$, $L_p^0 = e^x (d^p/dx^p) (e^{-x} x^p)$.

$$a_0 = \frac{\hbar^2}{me^2} \quad (13.1.24)$$

called the *Bohr radius*, which provides the natural distance scale for the hydrogen atom,

$$R_{nl}(r) \sim e^{-r/a_0} \left(\frac{r}{na_0} \right)^l L_{n-l-1}^{2l+1} \left(\frac{2r}{na_0} \right) \quad (13.1.25)$$

As $r \rightarrow \infty$, L will be dominated by the highest power, r^{n-l-1} , and

$$R_{nl} \underset{r \rightarrow \infty}{\sim} (r)^{n-1} e^{-r/a_0} \quad (\text{independent of } l) \quad (13.1.26)$$

(If $l=n-1$, this form is valid at all r since L_0^{2l+1} is a constant.) Equation (13.1.26) was anticipated in the last chapter when we considered the behavior of U_{El} as $r \rightarrow \infty$, in a Coulomb potential (see Exercise 13.1.4).

The following are the first few normalized eigenfunctions, $\psi_{Elm} \equiv \psi_{nlm}$:

$$\begin{aligned} \psi_{1,0,0} &= \left(\frac{1}{\pi a_0^3} \right)^{1/2} e^{-r/a_0} \\ \psi_{2,0,0} &= \left(\frac{1}{32\pi a_0^3} \right)^{1/2} \left(2 - \frac{r}{a_0} \right) e^{-r/2a_0} \\ \psi_{2,1,0} &= \left(\frac{1}{32\pi a_0^3} \right)^{1/2} \frac{r}{a_0} e^{-r/2a_0} \cos \theta \\ \psi_{2,1,\pm 1} &= \mp \left(\frac{1}{64\pi a_0^3} \right)^{1/2} \frac{r}{a_0} e^{-r/2a_0} \sin \theta e^{\pm i\phi} \end{aligned} \quad (13.1.27)$$

Exercise 13.1.1. Derive Eqs. (13.1.11) and (13.1.14) starting from Eqs. (13.1.8)–(13.1.10).

Exercise 13.1.2. Derive the degeneracy formula, Eq. (13.1.18).

Exercise 13.1.3. Starting from the recursion relation, obtain ψ_{210} (normalized).

Exercise 13.1.4. Recall from the last chapter [Eq. (12.6.19)] that as $r \rightarrow \infty$, $U_E \sim (r)^{me^2/\kappa\hbar^2} e^{-\kappa r}$ in a Coulomb potential $V = -e^2/r$ [$\kappa = (2mW/\hbar^2)^{1/2}$]. Show that this agrees with Eq. (13.1.26).

Let us explore the statement that a_0 provides a natural length scale for the hydrogen atom. Consider the state described by

$$\psi_{n,n-1,m} \propto e^{-r/a_0} r^{n-1} Y_{n-1}^m(\theta, \phi) \quad (13.1.28)$$

Let us ask for the probability of finding the electron in a spherical shell of radius r and thickness dr :

$$\int_{\Omega} P(\mathbf{r}) r^2 dr d\Omega \propto e^{-2r/a_0} r^{2n} dr \quad (13.1.29)$$

The probability density in r reaches a maximum when

$$\frac{d}{dr} (e^{-2r/a_0} r^{2n}) = 0$$

or

$$r = n^2 a_0 \quad (13.1.30)$$

When $n=1$, this equals a_0 . Thus the Bohr radius gives the most probable value of r in the ground state and this defines the “size” of the atom (to the extent one may speak of it in quantum theory). If $n>1$ we see that the size grows as n^2 , at least in the state of $l=n-1$. If $l \neq n-1$, the radial function has $n-l-1$ zeros and the density in r has several bumps. In this case, we may define the size by $\langle r \rangle$.‡ It can be shown, by using properties of L_{n-l-1}^{2l+1} that

$$\langle r \rangle_{nlm} = \frac{a_0}{2} [3n^2 - l(l+1)] \quad (13.1.31)$$

Rather than go through the lengthy derivation of this formula let us consider the following argument, which indicates that the size grows as $n^2 a_0$. In any eigenstate

$$\langle H \rangle = E = \langle T \rangle + \langle V \rangle = \langle P^2/2m \rangle - \langle e^2/r \rangle \quad (13.1.32)$$

It can be shown (Exercise 13.1.5) that

$$\langle T \rangle = -\frac{1}{2} \langle V \rangle \quad (13.1.33)$$

which is just the quantum version of the classical virial theorem, which states that if $V=cr^k$, then the averages \bar{T} and \bar{U} are related by

$$\bar{T} = \frac{k}{2} \bar{V}$$

‡ Even though r represents the abstract operator $(X^2 + Y^2 + Z^2)^{1/2}$ only in the coordinate basis, we shall use the same symbol to refer to it in the abstract, so as to keep the notation simple.

It follows that

$$E = \frac{1}{2} \langle V \rangle = -\frac{1}{2} \langle e^2/r \rangle \quad (13.1.34)$$

Now, in the state labeled by n ,

$$E_n = \frac{-me^4}{2\hbar^2 n^2} = \frac{-e^2}{2a_0 n^2} \quad (13.1.35)$$

from which it follows that

$$\left\langle \frac{1}{r} \right\rangle_n = \frac{1}{a_0 n^2} \quad (13.1.36)$$

Although

$$\frac{1}{\langle r \rangle} \neq \left\langle \frac{1}{r} \right\rangle$$

the two are of the same order of magnitude (see Exercise 9.4.2) and we infer that

$$\langle r \rangle_n \sim n^2 a_0 \quad (13.1.37)$$

which agrees with the result Eq. (13.1.31). (One must be somewhat cautious with statements like $\langle 1/r \rangle \approx 1/\langle r \rangle$. For example, it is not true in an s state that $\langle 1/r^4 \rangle \approx 1/\langle r^4 \rangle$, since $\langle 1/r^4 \rangle$ is divergent while $1/\langle r^4 \rangle$ is not. In the present case, however, $\langle 1/r \rangle$ is well defined in all states and indeed $\langle 1/r \rangle$ and $1/\langle r \rangle$ are of the same order of magnitude.)

This completes our analysis of the hydrogen spectrum and wave functions. Several questions need to be answered, such as (1) What are the numerical values of E_n , a_0 , etc.? (2) How does one compare the energy levels and wave functions deduced here with experiment?

These questions will be taken up in Section 13.3. But first let us address a question raised earlier: what is the source of the degeneracy in l at each n ?

Exercise 13.1.5. (Virial Theorem).* Since $|n, l, m\rangle$ is a stationary state, $\langle \dot{\Omega} \rangle = 0$ for any Ω . Consider $\Omega = \mathbf{R} \cdot \mathbf{P}$ and use Ehrenfest's theorem to show that $\langle T \rangle = (-1/2) \langle V \rangle$ in the state $|n, l, m\rangle$.

13.2. The Degeneracy of the Hydrogen Spectrum

The hydrogen atom, like the oscillator, exhibits “accidental degeneracy.” Quotation marks are used once again, because, as in the case of the oscillator, the degeneracy can be explained in terms of other symmetries the Hamiltonian has besides rotational invariance. Now, we have seen that the symmetries of H imply

the conservation of the generators of the symmetries. Consequently, if there is an extra symmetry (besides rotational invariance) there must be some extra conserved quantities (besides angular momentum). Now it is well known classically that the Coulomb[†] potential is special (among rotationally invariant potentials) in that it conserves the *Runge-Lenz vector*

$$\mathbf{n} = \frac{\mathbf{p} \times \mathbf{l}}{m} - \frac{e^2}{r} \mathbf{r} \quad (13.2.1)$$

The conservation of \mathbf{n} implies that not only is the orbit confined to a plane perpendicular to \mathbf{l} (as in any rotationally invariant problem) it is also *closed* (Exercise 13.2.1).

In quantum theory then, there will be an operator \mathbf{N} which commutes with H :

$$[\mathbf{N}, H] = 0 \quad (13.2.2)$$

and is given by[§]

$$\mathbf{N} = \frac{1}{2m} [\mathbf{P} \times \mathbf{L} - \mathbf{L} \times \mathbf{P}] - \frac{e^2 \mathbf{R}}{(X^2 + Y^2 + Z^2)^{1/2}} \quad (13.2.3)$$

We have seen that the conservation of \mathbf{L} implies that $[L_z, H] = 0$, which means that we can raise and lower the m values at a given l without changing the energy. This is how the degeneracy in m is “explained” by rotational invariance.

So it must be that since $[\mathbf{N}, H] = 0$, we must be able to build some operator out of the components of \mathbf{N} , which commutes with H and which raises l by one unit. This would then explain the degeneracy in l at each n . Precisely what this operator is and how it manages to raise l by one unit will be explained in Section 15.4, devoted to the study of “accidental” degeneracy. You will also find therein the explanation of the degeneracy of the oscillator.

Exercise 13.2.1. Let us see why the conservation of the Runge-Lenz vector \mathbf{n} implies closed orbits.

(1) Express \mathbf{n} in terms of \mathbf{r} and \mathbf{p} alone (get rid of \mathbf{l}).

(2) Since the particle is bound, it cannot escape to infinity. So, as we follow it from some arbitrary time onward, it must reach a point \mathbf{r}_{\max} where its distance from the origin stops growing. Show that

$$\mathbf{n} = \mathbf{r}_{\max} \left(2E + \frac{e^2}{r_{\max}} \right)$$

[†] Or generally any $1/r$ potential, say, gravitational.

[§] Since $[\mathbf{P}, \mathbf{L}] \neq 0$, we have used the symmetrization rule to construct \mathbf{N} from \mathbf{n} , i.e., $\mathbf{p} \times \mathbf{l} \rightarrow \frac{1}{2}[(\mathbf{P} \times \mathbf{L}) + (\mathbf{P} \times \mathbf{L})^\dagger] = \frac{1}{2}[\mathbf{P} \times \mathbf{L} - \mathbf{L} \times \mathbf{P}]$ (verify this).

$$\mathbf{n} = \mathbf{r}_{\min} \left(2E + \frac{e^2}{r_{\min}} \right)$$

at this point. (Use the law of conservation of energy to eliminate p^2 .) Show that, for similar reasons, if we wait some more, it will come to \mathbf{r}_{\max} , where

Thus \mathbf{r}_{\max} and \mathbf{r}_{\min} are parallel to each other and to \mathbf{n} . The conservation or constancy of \mathbf{n} implies that the maximum (minimum) separation is always reached at the same point $\mathbf{r}_{\max}(\mathbf{r}_{\min})$, i.e., the orbit is closed. In fact, all three vectors \mathbf{r}_{\max} , \mathbf{r}_{\min} , and \mathbf{n} are aligned with the major axis of the ellipse along which the particle moves; \mathbf{n} and \mathbf{r}_{\min} are parallel, while \mathbf{n} and \mathbf{r}_{\max} are antiparallel. (Why?) Convince yourself that for a circular orbit, \mathbf{n} must and does vanish.

13.3. Numerical Estimates and Comparison with Experiment

In this section we (1) obtain numerical estimates for various quantities such as the Bohr radius, energy levels, etc.; (2) ask how the predictions of the theory are actually compared with experiment.

Numerical Estimates

Consider first the particle masses. We will express the rest energies of the particles in million-electron volts or MeV:

$$mc^2 \approx 0.5 \text{ Mev} \quad (0.511 \text{ is a more exact value}) \quad (13.3.1)$$

$$Mc^2 = 1000 \text{ MeV} \quad (938.3) \dagger \quad (13.3.2)$$

$$m/M \approx 1/2000 \quad (1/1836) \dagger \quad (13.3.3)$$

Consequently the reduced mass μ and electron mass m are almost equal:

$$\mu = \frac{mM}{m+M} \approx \frac{mM}{M} = m \quad (13.3.4)$$

as are the relative coordinate and the electron coordinate.

Consider now an estimate of the Bohr radius

$$a_0 = \hbar^2 / me^2 \quad (13.3.5)$$

\dagger A more exact value.

To find this we need the values of \hbar and e . It was mentioned earlier that

$$\hbar = 1.054 \times 10^{-27} \text{ erg sec}$$

A more useful thing to remember for performing quick estimates is†

$$\hbar c \approx 2000 \text{ eV Å} \quad (1973.3) \quad (13.3.6)$$

where 1 angstrom (\AA) = 10^{-8} cm. The best way to remember e^2 is through the *fine-structure constant*:

$$\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137} \left(\frac{1}{137.04} \right) \quad (13.3.7)$$

This constant plays a fundamental role in quantum mechanical problems involving electrodynamics. Since it is dimensionless, its numerical value has an absolute significance: no matter what units we use for length, mass, and time, α will be 1/137. Thus, although no one tries to explain why $c = 3 \times 10^{10}$ cm/sec, several attempts have been made to arrive at the magic figure of 1/137. Since it is a God-given number (independent of mortal choice of units) one tries to relate it to fundamental numbers such as π , e , e^π , π^e , the number of space-time dimensions, etc.

Anyway, returning to our main problem, we can now estimate a_0 :

$$a_0 \approx \frac{\hbar^2}{me^2} = \frac{\hbar c}{mc^2} \left(\frac{\hbar c}{e^2} \right) = \frac{(2000)(137)}{0.5 \times 10^6} \text{ Å} \approx 0.55 \text{ Å} \quad (0.53)$$

Consider next the energy levels

$$E_n = -Ry/n^2$$

We estimate

$$\begin{aligned} Ry &= \frac{me^4}{2\hbar^2} = \frac{mc^2}{2} \left(\frac{e^2}{\hbar c} \right)^2 \\ &\approx \frac{0.25 \times 10^6}{(137)^2} \text{ eV} \approx 13.3 \text{ eV} \end{aligned} \quad (13.6)$$

So, using the more accurate value of Ry,

$$E_n = \frac{-13.6}{n^2} \text{ eV}$$

† Many of the tricks used here were learned from Professor A. Rosenfeld at the University of California, Berkeley.

The electron in the ground state needs 13.6 eV to be liberated or ionized. One may imagine that it is 13.6 eV down the infinitely deep Coulomb potential.

Let us digress to consider two length scales related to a_0 . The first

$$a_0\alpha = \frac{\hbar^2}{me^2} \cdot \frac{e^2}{\hbar c} = \frac{\hbar}{mc} \equiv \lambda_e \quad (13.3.8)$$

is called the *Compton wavelength* of the electron and is 137 times smaller than the Bohr radius. What does λ_e represent? In discussing the nuclear force, it was pointed out that the Compton wavelength of the pion was the distance over which it could be exchanged. It can also be defined as the lower limit on how well a particle can be localized. In the nonrelativistic theory we are considering, the lower limit is zero, since we admit position eigenkets $|x\rangle$. But in reality, as we try to locate the particle better and better, we use more and more energetic probes, say photons to be specific. To locate it to some ΔX , we need a photon of momentum

$$\Delta P \sim \frac{\hbar}{\Delta X}$$

Since the photon is massless, the corresponding energy is

$$\Delta E \sim \frac{\hbar c}{\Delta X}$$

in view of Einstein's formula $E^2 = c^2 p^2 + m^2 c^4$.

If this energy exceeds twice the rest energy of the particle, relativity allows the production of a particle–antiparticle pair in the measurement process. So we demand

$$\Delta E \lesssim 2mc^2$$

$$\frac{\hbar c}{\Delta X} \lesssim 2mc^2$$

or

$$\Delta X \gtrsim \frac{\hbar}{2mc} \sim \frac{\hbar}{mc}$$

If we attempt to localize the particle any better, we will see pair creation and we will have three (or more) particles instead of the one we started to locate.

In our analysis of the hydrogen atom, we treated the electron as a localized point particle. The preceding analysis shows that this is not strictly correct, but it

also shows that it is a fair approximation, since the “fuzziness” or “size” of the electron is α times smaller than the size of the atom, a_0

$$\frac{\hbar/mc}{a_0} = \alpha \approx \frac{1}{137}$$

Had the electric charge been 10 times as big, α would have been of order unity, and the size of the electron and the size of its orbit would have been of the same order and the point particle approximation would have been untenable. Let us note that

$$\lambda_e = a \cdot a_0 \approx 0.5 \times \frac{1}{137} \text{ Å} \approx \frac{1}{250} \text{ Å} \approx 4 \times 10^{-3} \text{ Å}$$

If we multiply λ_e by α we get another length, called the *classical radius of the electron*:

$$r_e = \alpha \lambda_e = \frac{\hbar}{mc} \cdot \frac{e^2}{\hbar c} = \frac{e^2}{mc^2} \approx 3 \times 10^{-5} \text{ Å} \quad (13.3.9)$$

If we imagine the electron to be a spherical charge distribution, the Coulomb energy of the distribution (the energy it takes to assemble it) will be of the order e^2/r_e , where r_e is the radius of the sphere. If we attribute the rest energy of the electron to this Coulomb energy, we arrive at the classical radius. In summary,

$$\begin{array}{ccccc} a_0 & \xrightarrow{\alpha} & \lambda_e & \xrightarrow{\alpha} & r_0 \\ \left(\frac{1}{2} \text{ Å}\right) & & \left(\frac{\alpha}{2} \text{ Å}\right) & & \left(\frac{\alpha^2}{2} \text{ Å}\right) \end{array}$$

Let us now return to the hydrogen atom. The mnemonics discussed so far are concerned only with the numbers. Let us now consider mnemonics that help us remember the dynamics. These must be used with caution, for they are phrased in terms not allowed in quantum theory.

The source of these mnemonics is the *Bohr model* of the hydrogen atom. About a decade or so prior to the formulation of quantum mechanics as described in this text, Bohr proposed a model of the atom along the following lines. Consider a particle of mass m in $V(r) = -e^2/r$, moving in a circular orbit of radius r . The dynamical equation is

$$\frac{mv^2}{r} = \frac{e^2}{r^2} \quad (13.3.10)$$

or

$$mv^2 = \frac{e^2}{r} \quad (13.3.11)$$

Thus any radius is allowed if r satisfies this equation. It also follows that any energy is allowed since

$$E = \frac{1}{2}mv^2 - \frac{e^2}{r} = -\frac{e^2}{2r} = -\frac{1}{2}mv^2 \quad (13.3.12)$$

Bohr conjectured that the only allowed orbits were those that had integral angular momentum in units of \hbar :

$$mv r = n\hbar \quad (13.3.13)$$

Feeding this into Eq. (13.3.11) we get

$$m \cdot \frac{n^2 \hbar^2}{m^2 r^2} = \frac{e^2}{r}$$

or

$$r = n^2 \frac{\hbar^2}{me^2} = n^2 a_0 \quad (13.3.14)$$

and

$$E_n = -\frac{e^2}{2r} = -\frac{e^2}{2a_0} \cdot \left(\frac{1}{n^2}\right) \quad (13.3.15)$$

Thus, if you ever forget the formula for a_0 or E_n , you can go back to this model for the formulas (though not for the physics, since it is perched on the fence between classical and quantum mechanics; it speaks of orbits, but quantizes angular momentum and so on). The most succinct way to remember the Bohr atom (i.e., a mnemonic for the mnemonic) is the equation

$$\alpha = \beta \quad (13.3.16)$$

where β is the velocity of the electron in the ground state of hydrogen measured in units of velocity of light ($\beta = v/c$). Given this, we get the ground state energy as

$$\begin{aligned} E_1 &= -\frac{1}{2}mv^2 = -\frac{1}{2}mc^2(v/c)^2 = -\frac{1}{2}mc^2\beta^2 = -\frac{1}{2}mc^2\alpha^2 \\ &= -\frac{1}{2}mc^2\left(\frac{e^2}{\hbar c}\right)^2 = -\frac{me^4}{2\hbar^2} \end{aligned} \quad (13.3.17)$$

Given this, how could one forget that the levels go as n^{-2} , i.e.,

$$E_n = -\frac{E_1}{n^2}$$

If we rewrite E_1 as $-e^2/2a_0$, we can get the formula for a_0 . The equation $\alpha=\beta$ also justifies the use of nonrelativistic quantum mechanics. An equivalent way (which avoids the use of velocity) is Eq. (13.3.17), which states that the binding energy is $\simeq(1/137)^2$ times the rest energy of the electron.

*Exercise 13.3.1.** The pion has a range of 1 Fermi = 10^{-5} Å as a mediator of nuclear force. Estimate its rest energy.

*Exercise 13.3.2.** Estimate the de Broglie wavelength of an electron of kinetic energy 200 eV. (Recall $\lambda=2\pi\hbar/p$.)

Comparison with Experiment

Quantum theory makes very detailed predictions for the hydrogen atom. Let us ask how these are to be compared with experiment. Let us consider first the energy levels and then the wave functions. In principle, one can measure the energy levels by simply weighing the atom. In practice, one measures the *differences in energy levels* as follows. If we start with the atom in an eigenstate $|nlm\rangle$, it will stay that way forever. However, if we perturb it for a time T , by turning on some external field (i.e., change the Hamiltonian from H^0 , the Coulomb Hamiltonian, to $H^0 + H^1$) its state vector can start moving around in Hilbert space, since $|nlm\rangle$ is not a stationary state of $H^0 + H^1$. If we measure the energy at time $t > T$, we may find it corresponds to another state with $n' \neq n$. One measures the energy by detecting the photon emitted by the atom. The frequency of the detected photon will be

$$\omega_{nn'} = \frac{E_n - E_{n'}}{\hbar} \quad (13.3.18)$$

Thus the frequency of light coming out of hydrogen will be

$$\begin{aligned} \omega_{nn'} &= \frac{\text{Ry}}{\hbar} \left(-\frac{1}{n^2} + \frac{1}{n'^2} \right) \\ &= \frac{\text{Ry}}{\hbar} \left(\frac{1}{n'^2} - \frac{1}{n^2} \right) \end{aligned} \quad (13.3.19)$$

For a fixed value $n' = 1, 2, 3, \dots$, we obtain a family of lines as we vary n . These families have in fact been seen, at least for several values of n' . The $n' = 1$ family is

called the *Lyman series* (it corresponds to transitions to the ground state from the upper ones):

$$\omega_{n1} = \frac{Ry}{\hbar} \left(\frac{1}{1} - \frac{1}{n^2} \right) \quad (13.3.20)$$

The $n'=2$ family is called the *Balmer series* and corresponds to transitions to the states $|2lm\rangle$ from $n=3, 4, \dots$, etc. The $n'=3$ family called the *Paschen series*, etc. Let us estimate the wavelength of a typical line in the Lyman series, say the one corresponding to the transition $n=2 \rightarrow n'=1$:

$$\begin{aligned} \omega_{21} &= \frac{13.5 \text{ eV}}{\hbar} \left(1 - \frac{1}{4} \right) \\ &\simeq \frac{10}{\hbar} \text{ eV} \end{aligned}$$

The wavelength is estimated to be

$$\lambda = \frac{2\pi c}{\omega} = \frac{2\pi}{10} (\hbar c) \simeq 1200 \text{ \AA}$$

A more refined estimate gives a value of 1216 Å, in very good agreement with experiment. Equally good is the agreement for all other observed lines. However, there are, in all cases, small discrepancies. Much of these may be explained by corrections that are calculable in theory. First we must correct for the fact that the proton is not really immobile; that we have here a two-body problem. As explained in Chapter 10, this is done by writing Schrödinger's equation for the relative (and not electron) coordinate and working in the CM frame. This equation would differ from Eq. (13.1.2) only in that m would be replaced by μ . This in fact would be the only change in all the formulas that follow, in particular Eq. (13.1.16) for the energy levels. This would simply rescale the entire spectrum by a factor $\mu/m = M/(M+m)$, which differs from 1 by less than a tenth of a percent. This difference is, however, observable in practice: one sees it in the difference between the levels of hydrogen and deuterium (whose nucleus has a proton and a neutron).

Then there is the correction due to the fact that the kinetic energy of the electron is not $\frac{1}{2}mv^2 = p^2/2m$ in Einstein's theory, but instead $mc^2[(1-v^2/c^2)^{-1/2} - 1]$, which is the difference between the energy at velocity v and the energy at rest. The $\frac{1}{2}mv^2$ term is just the first in the power series expansion of the above, in the variable v^2/c^2 . In Chapter 17 we will take into account the effect of the next term, which is $-3mv^4/8c^2$, or in terms of the momentum, $-3p^4/8m^3c^2$. This is a correction of order v^2/c^2 relative to the $p^2/2m$ piece we included, or since $v/c \approx \alpha$, a correction of order α^2 relative to main piece. There are other corrections of the same order, and these go by the name of *fine-structure corrections*. They will be included (in some approximation) in Chapter 17. The Dirac equation, which we will not solve in this book, takes into account the relativistic corrections to all orders in v/c . However, it too doesn't give the full story; there are tiny corrections due to quantum fluctuations of

the electromagnetic field (which we have treated classically so far). These corrections are calculable in theory and measurable experimentally. The agreement between theory and experiment is spectacular. It is, however, important to bear in mind that all these corrections are icing on the cake; that the simple nonrelativistic Schrödinger equation by itself provides an excellent description of the hydrogen spectrum. (Much of the present speculation on what the correct theory of elementary particles is will be put to rest if one can come up with a description of these particles that is half as good as the description of the hydrogen atom by Schrödinger's equation.)

Consider next the wave functions. To test the predictions, one once again relies on perturbing the system. The following example should give you a feeling for how this is done. Suppose we apply an external perturbation H^1 for a short time ε . During this time, the system goes from $|nlm\rangle$ to

$$\begin{aligned} |\psi(\varepsilon)\rangle &= \left[I - \frac{i\varepsilon}{\hbar} (H^0 + H^1) \right] |nlm\rangle \\ &= |nlm\rangle - \left(\frac{i\varepsilon E_n}{\hbar} + \frac{i\varepsilon H^1}{\hbar} \right) |nlm\rangle \end{aligned}$$

The probability of it being in a state $|n'l'm'\rangle$ (assuming $|n'l'm'\rangle$ is different from $|nlm\rangle$) is

$$|\langle n'l'm' | \psi(\varepsilon) \rangle|^2 = \left| -\frac{i\varepsilon}{\hbar} \langle n'l'm' | H^1 | nlm \rangle \right|^2$$

Thus quantum theory can also determine for us the rate of transition to the state $|n'l'm'\rangle$. This rate is controlled by the matrix element $\langle n'l'm' | H^1 | nlm \rangle$, which in coordinate space, will be some integral over $\psi_{n'l'm}'$ and ψ_{nlm} with H^1 sandwiched between them. The evaluation of the integrals entails detailed knowledge of the wave functions, and conversely, agreement of the calculated rates with experiment is a check on the predicted wave functions. We shall see a concrete example of this when we discuss the interaction of radiation with matter in Chapter 18.

Exercise 13.3.3. Instead of looking at the emission spectrum, we can also look at the *absorption* spectrum of hydrogen. Say some hydrogen atoms are sitting at the surface of the sun. From the interior of the sun, white light tries to come out and the atoms at the surface absorb what they can. The atoms in the ground state will now *absorb* the Lyman series and this will lead to dark lines if we analyze the light coming from the sun. The presence of these lines will tell us that there is hydrogen at the surface of the sun. We can also estimate the surface temperature as follows. Let T be the surface temperature. The probabilities $P(n=1)$ and $P(n=2)$ of an atom being at $n=1$ and $n=2$, respectively, are related by Boltzmann's formula

$$\frac{P(n=2)}{P(n=1)} = 4 e^{-(E_2 - E_1)/kT}$$

where the factor 4 is due to the degeneracy of the $n=2$ level. Now only atoms in $n=2$ can produce the Balmer lines in the absorption spectrum. The relative strength of the Balmer and

Lyman lines will tell us $P(n=2)/P(n=1)$, from which we may infer T . Show that for $T = 6000$ K, $P(n=2)/P(n=1)$ is negligible and that it becomes significant only for $T \approx 10^5$ K. (The Boltzmann constant is $k \approx 9 \times 10^{-5}$ eV/K. A mnemonic is $kT \approx \frac{1}{40}$ eV at room temperature, $T = 300$ K.)

13.4. Multielectron Atoms and the Periodic Table

It is not possible to treat multielectron atoms analytically even if we treat the nucleus as immobile. Although it is possible, in principle, to treat an arbitrarily complex atom by solving the exact Schrödinger equation numerically, a more practical method is to follow some approximation scheme. Consider the one due to Hartree. Here one assumes that each electron obeys a one-particle Schrödinger equation wherein the potential energy $V = -e\phi(r)$ is due to the nucleus and the other electrons. In computing the electronic contribution to $\phi(r)$, each electron is assigned a charge distribution which is $(-e)$ times the probability density associated with its wave function. And what are the wave functions? They are the eigenstates in the potential $\phi(r)$! To break the vicious circle, one begins with a reasonable guess for the potential, call it $\phi_0(r)$, and computes the allowed energy eigenstates. One then fills them up in the order of increasing energy, putting in just two electrons in each orbital state, with opposite spins (the Pauli principle will not allow any more)[‡] until all the electrons have been used up. One then computes the potential $\phi_1(r)$ due to this electronic configuration.[§] If it coincides with $\phi_0(r)$ (to some desired accuracy) one stops here and takes the configuration one got to be the ground state of the atom. If not, one goes through one more round, this time starting with $\phi_1(r)$. The fact that, in practice, one soon finds a potential that reproduces itself, signals the soundness of this scheme.

What do the eigenstates look like? They are still labeled by (nlm) as in hydrogen, with states of different m degenerate at a given n and l . [This is because $\phi(r)$ is rotationally invariant.] The degeneracy in l is, however, lost. Formally this is because the potential is no longer $1/r$ and physically this is because states with lower angular momentum have a larger amplitude to be near the origin and hence sample more of the nuclear charge, while states of high angular momentum, which are suppressed at the origin, see the nuclear charge shielded by the electrons in the inner orbits. As a result, at each n the energy goes up with l . The “radius” of each state grows with n , with a slight dependence on l . States of a given n are thus said to form a *shell* (for, in a semiclassical sense, they may be viewed as moving on a sphere of radius equal to the most probable value of r). States of a given l and n are said to form a *subshell*.

Let us now consider the electronic configurations of some low Z (Z is the nuclear charge) atoms. Hydrogen (${}^1\text{H}$) has just one electron, which is in the $1s$ state. This configuration is denoted by $1s^1$. Helium (${}^2\text{He}$) has two electrons in the $1s$ state with opposite spins, a configuration denoted by $1s^2$. ${}^2\text{He}$ has its $n=1$ shell filled. Lithium (${}^3\text{Li}$) has its third electron in the $2s$ state, i.e., it is in the configuration $1s^2 2s^1$. (Recall

[‡] In this discussion electron spin is viewed as a spectator variable whose only role is to double the states. This is a fairly good approximation.

[§] If necessary, one averages over angles to get a spherically symmetric ϕ .

that the s state is lower than the p state.) We keep going this way through beryllium (${}^4\text{Be}$), boron (${}^5\text{B}$), carbon (${}^6\text{C}$), nitrogen (${}^7\text{N}$), oxygen (${}^8\text{O}$), and fluorine (${}^9\text{F}$), till neon (${}^{10}\text{Ne}$). Neon is in the configuration $1s^2 2s^2 2p^6$, i.e., has its $n=2$ shell filled. The next element, sodium (${}^{11}\text{Na}$), has a solitary electron in the $3s$ state. The $3s$ and $3p$ subshells are filled when we get to argon (${}^{18}\text{Ar}$). The next one, potassium (${}^{19}\text{K}$) has its 19th electron in the $4s$ and not $3d$ state. This is because the growth in energy due to a change in n from 3 to 4 is less than the growth due to change in l from 1 to 2 at $n = 3$. This phenomenon occurs often as we move up in Z . For example, in the “rare earth” elements, the $6s$ shell is filled before the $4f$ shell.

Given the electronic configurations, one can anticipate many of the chemical properties of the elements. Consider an element such as ${}^{10}\text{Ne}$, which has a closed outer shell. Since the total electronic charge is spherically symmetric ($|R_{nl}|^2 \sum_{m=-l}^l |Y_l^m|^2$ is independent of θ and ϕ), it shields the nuclear charge very effectively and the atom has no significant electrostatic affinity for electrons in other atoms. If one of the electrons in the outer shell could be excited to a higher level, this would change, but there is a large gap in energy to cross. Thus the atom is rarely excited and is chemically inert. On the other hand, consider an element like ${}^{11}\text{Na}$, which has one more electron, which occupies the $3s$ state. This electron sees a charge of $+e$ when it looks inward (the nuclear charge of 11 shielded by the 10 electrons in the $n=1$ and 2 shells) and is thus very loosely bound. Its binding energy is 5.1 eV compared to an $n=2$ electron in Ne, which has a binding energy of 21.6 eV. If ${}^{11}\text{Na}$ could get rid of this electron, it could reach a stable configuration with a closed $n=2$ shell. If we look one place to the left (in Z) of ${}^{10}\text{Ne}$, we see a perfect acceptor for this electron: we have here ${}^9\text{F}$, whose $n=2$ shell is all full except for one electron. So when ${}^{11}\text{Na}$ and ${}^9\text{F}$ get together, Na passes on its electron to F and the system as a whole lowers its energy, since the binding energy in F is 17.4 eV. Having carried out the transfer, the atoms cannot part company, for they have now become charged ions, Na^+ and F^- , which are held together by electrostatic attraction, called the *ionic bond* and form the NaF molecule.

Once we grasp that the chemical behavior is dictated by what is happening in the outermost shell, we can see that several elements will have similar chemical properties because they have similar outer shells. For example, we expect all elements with filled outer shells to be chemically inert. This is true. It is also true that some elements with filled subshells are also inert, such as ${}^{18}\text{Ar}$, in which just the $3s$ and $3p$ subshells are filled. The origin of this inertness is the same as in the case with filled shells: a spherically symmetric electronic charge distribution and a large excitation energy. If we move one place to the right of the inert elements, we meet those that behave like Na, i.e., eager to give up an electron, while if we move one place to the left, we meet the likes of F, eager to accept an electron. If we move two places to the left, we see the likes of oxygen, which want two electrons, while two places to the right we have elements like magnesium, which want to get rid of two electrons. It follows that as we move in Z , we see a certain chemical tendency over and over again. This quasiperiodic behavior was emphasized in 1869 by Mendeleev, who organized the elements into a *periodic table*, in which the elements are arranged into a matrix, with all similar elements in the same column. As we go down the first column, for example, we see H, Li, Na, etc., i.e., elements with one electron to spare. In the last column we see the inert elements, He, Ne, etc. Given the maxim that happiness is a filled outer shell, we can guess who will interact with whom. For

instance, not only can Na give its electron to F, it can give to Cl, which is one shy of a filled $3p$ subshell. Likewise F can get its electron from K as well, which has a lone electron in the $4s$ state. More involved things can happen, such as the formation of H_2O when two H atoms get together with an oxygen atom, forming the *covalent bond*, in which each hydrogen atom shares an electron with the oxygen atom. This way all three atoms get to fill their outer shells at least part of the time.

There are many more properties of elements that follow from the configuration of the outer electrons. Consider the rare earth elements, ^{58}Ce through ^{71}Lu , which have very similar chemical properties. Why doesn't the chemical behavior change with Z in this range? The answer is that in these elements the $6s$ subshell is filled and the $4f$ subshell, deep in the interior (but of a higher energy), is *being* filled. Since what happens in the interior does not affect the chemical properties, they all behave alike. The same goes for the *actinides*, ^{90}Th to ^{103}Lw , which have a filled $7s$ subshell and a $5f$ subshell that is getting filled up.

Since we must stop somewhere, let us stop here. If you want to know more, you must consult books devoted to the subject.‡

*Exercise 13.4.1.** Show that if we ignore interelectron interactions, the energy levels of a multielectron atom go as Z^2 . Since the Coulomb potential is Ze/r , why is the energy $\propto Z^2$?

*Exercise 13.4.2.** Compare (roughly) the sizes of the uranium atom and the hydrogen atom. Assume levels fill in the order of increasing n , and that the nonrelativistic description holds. Ignore interelectron effects.

*Exercise 13.4.3.** Visible light has a wavelength of approximately 5000 Å. Which of the series—Lyman, Balmer, Paschen—do you think was discovered first?

‡ See, for a nice trip through the periodic table, U. Fano and L. Fano, *Basic Physics of Atoms and Molecules*, Chapter 18, Wiley, New York (1959).

14

Spin

14.1. Introduction

In this chapter we consider a class of quantum phenomena that cannot be handled by a straightforward application of the four postulates. The reason is that these phenomena involve a quantum degree of freedom called *spin*, which has no classical counterpart. Consequently, neither can we obtain the spin operator by turning to Postulate II, nor can we immediately write down the quantum Hamiltonian that governs its time evolution. The problem is very important, for most particles—electrons, protons, neutrons, photons—have the spin degree of freedom. Fortunately the problem can be solved by a shrewd mixture of classical intuition and reasoning by analogy. In this chapter we study just electron spin. The treatment of the spins of other particles is quite similar, with the exception of the photon, which moves at speed c and can't be treated nonrelativistically. Photon spin will be discussed in Chapter 18.

In the next three sections we address the following questions:

- (1) What is the nature of this new spin degree of freedom?
- (2) How is the Hilbert space modified to take this new degree of freedom into account? What do the spin operators look like in this space (kinematics of spin)?
- (3) How does spin evolve with time, i.e., how does it enter the Hamiltonian (dynamics of spin)?

14.2. What is the Nature of Spin?

The best way to characterize spin is as a form of angular momentum. It is, however, not the angular momentum associated with the operator \mathbf{L} , as the following experiment shows. An electron is prepared in a state of zero linear momentum, i.e., in a state with a constant (space-independent) wave function. As the operators L_x , L_y , and L_z will give zero when acting on it, our existing formalism predicts that if the angular momentum along, say the z direction, is measured, a result of zero will obtain. The actual experiment, however, shows that this is wrong, that the result is

$\pm\hbar/2$.[†] It follows that the electron has “intrinsic” angular momentum, not associated with its orbital motion. This angular momentum is called *spin*, for it was imagined in the early days that if the electron has angular momentum without moving through space, then it must be spinning like a top. We adopt this nomenclature, but not the mechanical model that goes with it, for a consistent mechanical model doesn’t exist. Fortunately one can describe spin and its dynamics without appealing to any model, starting with just the observed fact that it is a form of angular momentum. Let us now develop the formalism that deals with spin and, in particular, allows us to understand the above experiment.

14.3. Kinematics of Spin

The discussion following the general solution to the eigenvalue problem of angular momentum (Section 12.5) suggests the way for treating particles with intrinsic angular momentum or spin. Recall that if a particle is described by a wave function with many (n) components, the generator of infinitesimal rotation is not just \mathbf{L} but something more. The reason is that under an infinitesimal rotation *two* things happen to the wave function: (1) the values at each spatial point are reassigned to the rotated point, and (2) the components of the wave function get transformed into linear combinations of each other.

The differential operator \mathbf{L} does part (1), while an $n \times n$ matrix \mathbf{S} is responsible for part (2).

By generalizing our findings from Exercise 12.5.1 to an n component wave function in three dimensions, we can say that under an infinitesimal rotation around the z axis, the wave function is transformed as follows:

$$\begin{bmatrix} \psi'_1 \\ \vdots \\ \psi'_n \end{bmatrix} = \left(\begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} - \frac{i\varepsilon}{\hbar} \begin{bmatrix} -i\hbar \partial/\partial\phi & & 0 \\ & \ddots & \\ 0 & & -i\hbar \partial/\partial\phi \end{bmatrix} - \frac{i\varepsilon}{\hbar} S_z \right) \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_n \end{bmatrix} \quad (14.3.1)$$

where S_z is an $n \times n$ matrix. In abstract form, this equation reads[§]

$$\begin{aligned} |\psi'\rangle &= \left[I - \frac{i\varepsilon}{\hbar} (L_z + S_z) \right] |\psi\rangle \\ &= \left[I - \frac{i\varepsilon}{\hbar} J_z \right] |\psi\rangle \end{aligned} \quad (14.3.2)$$

We identify J_z , the generator of infinitesimal rotations about the z axis, as the z component of angular momentum. We see it has two parts:

$$J_z = L_z + S_z$$

[†] In practice one measures not the angular momentum, but a related quantity called magnetic moment. More on this later. Also spin was first discovered on the basis of spectroscopic evidence and not from an experiment of the above type.

[§] The spin operators will be denoted by the same symbol (S) whether they are referred to in the abstract or as matrices in some basis.

and more generally

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (14.3.3)$$

Our problem is to find the number (n) of components appropriate to the electron and the three spin matrices that rotate its components. We proceed as follows.

Since J_i are generators of rotations, they must obey the consistency condition

$$[J_i, J_j] = i\hbar \sum_k \epsilon_{ijk} J_k \quad (14.3.4)$$

Since \mathbf{L} and \mathbf{S} act on different parts of the wave function (the former on x, y, z , the latter on the indices $i=1, \dots, n$) they commute, and we may infer from Eq. (14.3.4) that

$$[L_i, L_j] + [S_i, S_j] = i\hbar \left[\sum_k \epsilon_{ijk} L_k + \sum_k \epsilon_{ijk} S_k \right] \quad (14.3.5)$$

Using the known commutation rules of the L_i , we deduce

$$[S_i, S_j] = i\hbar \sum_k \epsilon_{ijk} S_k \quad (14.3.6)$$

Now recall that in Chapter 12 we found matrices J_x, J_y , and J_z [Eqs. (12.5.22)–(12.5.24)] that obey precisely these commutation relations. But these matrices were infinite dimensional. However, the infinite-dimensional matrices were built out of $(2j+1) \times (2j+1)$ blocks, with $j=0, 1/2, 1, 3/2, \dots$, and the commutation relations were satisfied block by block. So which block shall we pick for the electron spin operators? The answer is given by the empirical fact that S_z has only the eigenvalues $\pm\hbar/2$. This singles out the 2×2 blocks in Eqs. (12.5.22)–(12.5.24):

$$S_x = \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad S_y = \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad S_z = \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (14.3.7)$$

Thus, the way to describe the electron is through a two-component wave function called a *spinor*:

$$\psi = \begin{bmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{bmatrix} \quad (14.3.8a)$$

$$\equiv \psi_+ \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \psi_- \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (14.3.8b)$$

If $\psi_- = 0$, $\psi_+ \neq 0$, we have an eigenstate of S_z with eigenvalue $\hbar/2$; if $\psi_- \neq 0$, $\psi_+ = 0$, the S_z eigenvalue is $(-\hbar/2)$.

Let us now proceed to interpret the experiment mentioned earlier. Since we prepared a state of zero momentum, we want the operator \mathbf{P} to give zero when acting on ψ . The operator \mathbf{P} simply differentiates both components of ψ :

$$\mathbf{P} \rightarrow \begin{bmatrix} -i\hbar\nabla & 0 \\ 0 & -i\hbar\nabla \end{bmatrix} \quad (14.3.9)$$

We deduce from $\mathbf{P}|\psi\rangle=0$, i.e.,

$$\begin{bmatrix} -i\hbar\nabla\psi_+ \\ -i\hbar\nabla\psi_- \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (14.3.10)$$

that ψ_+ and ψ_- are independent of x , y , and z . It follows that L_z acting on ψ gives zero. However, S_z doesn't: there is an amplitude ψ_{\pm} for obtaining $\pm\hbar/2$.

The electron spinor is a *two*-component object, which puts it between a scalar, which has one component, and a vector, which has three. However, the components of the spinor are complex.

A significant difference between spin and orbital angular momentum is this: we can change the magnitude of orbital angular momentum of a particle (by applying external fields) but not the magnitude of its spin. The S^2 operator is

$$S^2 = \hbar^2 \begin{bmatrix} (\frac{1}{2})(\frac{1}{2}+1) & 0 \\ 0 & (\frac{1}{2})(\frac{1}{2}+1) \end{bmatrix} = \frac{3}{4}\hbar^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (14.3.11)$$

and yields a value $\frac{3}{4}\hbar^2$ on any state ψ . [For any particle, the magnitude of spin is decided by the number of components in the wave function and is an invariant. Thus the spin of the electron is always $1/2$ (in units of \hbar) and serves as an invariant label of the particle, like its charge or rest mass.]

We have deduced that the electron is to be described by a two-component wave function in the coordinate basis.[‡] Let us restate this result in Hilbert space. First, it is easy to see that the introduction of spin has doubled the size of Hilbert space; if it was ∞ dimensional before, now it is 2∞ dimensional, if you know what I mean. The basis vectors $|xyzs_z\rangle$ diagonalize the mutually commuting operators X , Y , Z , and S_z (one can also think of other bases such as $|\mathbf{p}s_z\rangle$ or $|\mathbf{p}s_x\rangle$ etc.). The state vector

[‡] We made the deduction given the empirical input from experiment. When we come to the Dirac equation, we will see that incorporating relativistic kinematics will automatically lead to a multicomponent wave function, i.e., lead to spin, if we demand that the equation be first order in time and space.

$|\psi\rangle$ is a 2∞ -dimensional column vector in this basis:

$$|\psi\rangle \xrightarrow[\mathbf{R}, S_z \text{ basis}]{} \langle xyzs_z | \psi \rangle = \begin{bmatrix} \psi(x = -\infty, y = -\infty, z = -\infty, s_z = +\hbar/2) \\ \vdots \\ \psi(x, y, z, s_z = +\hbar/2) \\ \vdots \\ \psi(x = \infty, y = \infty, z = \infty, s_z = +\hbar/2) \\ \hline \psi(x = -\infty, y = -\infty, z = -\infty, s_z = -\hbar/2) \\ \vdots \\ \psi(x, y, z, s_z = -\hbar/2) \\ \vdots \\ \psi(x = \infty, y = \infty, z = \infty, s_z = -\hbar/2) \end{bmatrix} \quad (14.3.12)$$

Clearly $\psi(\mathbf{r}, \pm\hbar/2)$ gives the amplitude to find the electron at \mathbf{r} with $s_z = \pm\hbar/2$. The horizontal dashed line separates the components with $s_z = \hbar/2$ from those with $s_z = -\hbar/2$. Now if s_z is fixed at $\hbar/2$ and we vary x, y, z from $-\infty$ to ∞ , the component of $|\psi\rangle$ will vary smoothly, i.e., define a continuous function $\psi_+(x, y, z)$. Likewise the components below the dotted line define a function $\psi_-(x, y, z)$. In terms of these functions, we may compactify Eq. (14.3.12) to the form

$$|\psi\rangle \xrightarrow[\mathbf{R}, S_z \text{ basis}]{} \begin{bmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{bmatrix} \quad (14.3.13)$$

This notation blends two notations we have used so far: if the vector has components labeled by a discrete index i ($i = 1, \dots, n$) we denote it as a column vector, while if it is labeled by a continuous index such as x , we denote it by a function $\psi(x)$; but here, since it depends on discrete (s_z) as well as continuous (x, y, z) indices, we write it as a column vector whose components are functions. The normalization condition is

$$\begin{aligned} 1 &= \langle \psi | \psi \rangle = \sum_{s_z} \int \langle \psi | xyzs_z \rangle \langle xyzs_z | \psi \rangle dx dy dz \\ &= \int (|\psi_+|^2 + |\psi_-|^2) dx dy dz \end{aligned} \quad (14.3.14)$$

In the compact notation, S_z is a 2×2 matrix:

$$S_z |\psi\rangle \xrightarrow[\mathbf{R}, S_z \text{ basis}]{} \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{bmatrix} \quad (14.3.15a)$$

whereas in its full glory, it is a 2 ∞ -dimensional matrix:

$$S_z |\psi\rangle \xrightarrow[\text{R}, S_z \text{ basis}]{\frac{\hbar}{2}} \begin{bmatrix} 1 & & & & 0 \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -1 \\ & & & & & -1 \\ & & & & & & \ddots \\ & & & & & & & -1 \\ 0 & & & & & & & & \end{bmatrix} \begin{bmatrix} \psi(-\infty, \hbar/2) \\ \vdots \\ \psi(\infty, -\hbar/2) \end{bmatrix} \quad (14.3.15b)$$

What about the familiar operators $\Omega(\mathbf{R}, \mathbf{P})$? Equation (14.3.9) gives \mathbf{P} in the compact notation. Likewise, L_z becomes

$$L_z |\psi\rangle \xrightarrow[\text{R}, S_z \text{ basis}]{\frac{\hbar}{2}} \begin{bmatrix} -i\hbar \partial/\partial\phi & 0 \\ 0 & -i\hbar \partial/\partial\phi \end{bmatrix} \begin{bmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{bmatrix} \quad (14.3.16)$$

The forms of these operators are consistent with the requirement that operators built out of \mathbf{R} and \mathbf{P} commute with the spin operators. Observe that the Hilbert space \mathbb{V}_e of the electron may be viewed as a direct product of an infinite-dimensional space \mathbb{V}_0 , which describes a particle with just orbital degrees of freedom, and a two-dimensional space \mathbb{V}_s , which describes a particle with just spin degrees of freedom:

$$\mathbb{V}_e = \mathbb{V}_0 \otimes \mathbb{V}_s \quad (14.3.17)$$

The basis vector $|x, y, z, s_z\rangle$ of \mathbb{V}_e is just a direct product

$$|x, y, z, s_z\rangle = |xyz\rangle \otimes |s=1/2, s_z\rangle \quad (14.3.18)$$

Of course \mathbb{V}_0 and \mathbb{V}_s do not describe two particles which are amalgamated into a single system, but, rather, two independent degrees of freedom of the electron.

Since we already know how to handle the orbital degrees of freedom, let us pretend from now on that only the spin degree of freedom exists. Or, to be more precise, let us assume the orbital degree of freedom exists but evolves independently. Formally this means that the Hamiltonian is separable:

$$H = H_0 + H_s \quad (14.3.19)$$

where H_0 and H_s depend on just the orbital and spin operators, respectively. Consequently the state vector factorizes into[‡]

$$|\psi(t)\rangle = |\psi_0(t)\rangle \otimes |\chi_s(t)\rangle \quad (14.3.20)$$

where $|\psi_0\rangle$ and $|\chi_s\rangle$ are elements of \mathbb{V}_0 and \mathbb{V}_s , respectively. Now $|\psi_0(t)\rangle$ evolves in response to H_0 , while the evolution of $|\chi_s(t)\rangle$ is dictated by H_s . We will follow just the evolution of $|\chi_s\rangle$. The product form of $|\psi\rangle$ ensures that the spin and orbital degrees of freedom are statistically independent. Of course, there are many interesting cases in which H is not separable, and the orbital and spin degrees are coupled in their evolution. We will tackle them in a later chapter.

With this assumption, we have just a (complex) two-dimensional Hilbert space \mathbb{V}_s to work with. A complete basis is provided by the vectors $|s, s_z\rangle = |s, m\hbar\rangle \equiv |s, m\rangle$. They are

$$|s, m\rangle = |1/2, 1/2\rangle \xrightarrow[S_z \text{ basis}]{ } \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (14.3.21a)$$

$$|s, m\rangle = |1/2, -1/2\rangle \xrightarrow[S_z \text{ basis}]{ } \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (14.3.21b)$$

Any ket $|\chi\rangle$ in \mathbb{V}_s may be expanded as

$$|\chi\rangle = \alpha|1/2, 1/2\rangle + \beta|1/2, -1/2\rangle \xrightarrow[S_z \text{ basis}]{ } \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (14.3.22)$$

The normalization condition is

$$1 = \langle \chi | \chi \rangle \xrightarrow[S_z \text{ basis}]{ } 1 = [\alpha^*, \beta^*] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = |\alpha|^2 + |\beta|^2 \quad (14.3.23)$$

If one calculates $\langle \mathbf{S} \rangle$ in the eigenstates of S_z , one finds

$$\langle 1/2, \pm 1/2 | \mathbf{S} | 1/2, \pm 1/2 \rangle = \pm(\hbar/2)\mathbf{k} \quad (14.3.24)$$

One refers to these as states with spin pointing up/down the z axis. More generally, the eigenstates $|\hat{n}, \pm\rangle$ of $\hat{n} \cdot \mathbf{S}$ with eigenvalues $\pm\hbar/2$, in which

$$\langle \hat{n}, \pm | \mathbf{S} | \hat{n}, \pm \rangle = \pm(\hbar/2)\hat{n} \quad (14.3.25)$$

are said to be states with spin up/down the direction of the unit vector \hat{n} . Let us address the determination (in the S_z basis) of the components of $|\hat{n}, \pm\rangle$ and the verification of Eq. (14.3.25).

[‡] In the \mathbf{R}, S_z basis, this means $\psi(x, y, z, s_z, t) = \psi_0(x, y, z, t)\chi(t)$ where χ is a two-component spinor independent of x, y , and z .

Let us say \hat{n} points in the direction (θ, ϕ) , i.e., that

$$\begin{aligned}\hat{n}_z &= \cos \theta \\ \hat{n}_x &= \sin \theta \cos \phi \\ \hat{n}_y &= \sin \theta \sin \phi\end{aligned}\tag{14.3.26}$$

The kets $|\hat{n}, \pm\rangle$ are eigenvectors of

$$\begin{aligned}\hat{n} \cdot \mathbf{S} &= n_x S_x + n_y S_y + n_z S_z \\ &= \frac{\hbar}{2} \begin{bmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{bmatrix} \\ &= \frac{\hbar}{2} \begin{bmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{bmatrix}\end{aligned}\tag{14.3.27}$$

It is a simple matter to solve the eigenvalue problem (Exercise 14.3.2) and to find

$$|\hat{n} \text{ up}\rangle \equiv |\hat{n}+\rangle = \begin{bmatrix} \cos(\theta/2) e^{-i\phi/2} \\ \sin(\theta/2) e^{i\phi/2} \end{bmatrix}\tag{14.3.28a}$$

$$|\hat{n} \text{ down}\rangle \equiv |\hat{n}-\rangle = \begin{bmatrix} -\sin(\theta/2) e^{-i\phi/2} \\ \cos(\theta/2) e^{i\phi/2} \end{bmatrix}\tag{14.3.28b}$$

You may verify that as claimed

$$\begin{aligned}\langle \hat{n} \pm | \mathbf{S} | \hat{n} \pm \rangle &= \pm(\hbar/2)(\mathbf{i} \sin \theta \cos \phi + \mathbf{j} \sin \theta \sin \phi + \mathbf{k} \cos \theta) \\ &= \pm(\hbar/2)\hat{n}\end{aligned}\tag{14.3.29}$$

An interesting feature of \mathbb{V}_s is that not only can we calculate $\langle \mathbf{S} \rangle$ given a state, but we can also go the other way, i.e., deduce the state vector given $\langle \mathbf{S} \rangle$. This has to do with the fact that any element of \mathbb{V}_s has only two (complex) components α and β , constrained by the normalization requirement $|\alpha|^2 + |\beta|^2 = 1$, i.e., three real degrees of freedom, and $\langle \mathbf{S} \rangle$ contains exactly three pieces of information. If we write $\langle \mathbf{S} \rangle$ as $(\hbar/2)\hat{n}$, then the corresponding ket is $|\hat{n}, +\rangle$ or if you want $|-\hat{n}, -\rangle$. Another way to state this result is as follows. Instead of specifying a state by α and β , we can give the operator $\hat{n} \cdot \mathbf{S}$ of which it is an eigenvector with eigenvalue $\hbar/2$. An interesting corollary is that every spinor in \mathbb{V}_s is an eigenket of some spin operator $\hat{n} \cdot \mathbf{S}$ with eigenvalue $\hbar/2$.

Exercise 14.3.1. Let us verify the above corollary explicitly. Take some spinor with components $\alpha = \rho_1 e^{i\phi_1}$ and $\beta = \rho_2 e^{i\phi_2}$. From $\langle \chi | \chi \rangle = 1$, deduce that we can write $\rho_1 = \cos(\theta/2)$ and $\rho_2 = \sin(\theta/2)$ for some θ . Next pull out a common phase factor so that the spinor takes the form in Eq. (14.3.28a). This verifies the corollary and also fixes \hat{n} .

So much for the state vectors in \mathbb{V}_s . How about the operators on this space? Let us commence with S_x , S_y , and S_z . It is convenient to introduce the *Pauli matrices* σ , defined by

$$\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\sigma} \quad (14.3.30)$$

so that

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (14.3.31)$$

It is worth memorizing these matrices. Here are some of their important properties.

(1) They *anticommute* with each other:

$$[\sigma_i, \sigma_j]_+ = 0 \quad \text{or} \quad \sigma_i \sigma_j = -\sigma_j \sigma_i \quad (i \neq j) \quad (14.3.32)$$

(2) From the commutation rules for the spin operators \mathbf{S} , we get, upon using the anticommutativity of the Pauli matrices,

$$\sigma_x \sigma_y = i \sigma_z \quad \text{and cyclic permutations} \quad (14.3.33)$$

(3) They are traceless

$$\text{Tr } \sigma_i = 0, \quad i = x, y, z \quad (14.3.34)$$

(See Exercise 14.3.3 for the proof.)

(4) The square of any Pauli matrix equals I :

$$\sigma_i^2 = I \quad (14.3.35)$$

or more generally,

$$(\hat{n} \cdot \boldsymbol{\sigma})^2 = I \quad (14.3.36)$$

Proof. Since S_z has eigenvalues $\pm \hbar/2$, it follows that

$$\left(S_z + \frac{I\hbar}{2} \right) \left(S_z - \frac{I\hbar}{2} \right) = 0$$

in this Hilbert space.[‡] But since what we call the z axis is arbitrary, it must be true that

$$\left(\hat{n} \cdot \mathbf{S} + \frac{I\hbar}{2}\right) \left(\hat{n} \cdot \mathbf{S} - \frac{I\hbar}{2}\right) = 0$$

or

$$(\hat{n} \cdot \mathbf{S})^2 = \frac{\hbar^2}{4} I$$

or

$$(\hat{n} \cdot \boldsymbol{\sigma})^2 = I$$

Q.E.D.

(5) We can combine Eqs. (14.3.32) and (14.3.35) into

$$[\sigma_i, \sigma_j]_+ = 2\delta_{ij}I \quad (14.3.37)$$

(6) Combining this relation with the commutation rules

$$[\sigma_x, \sigma_y] = 2i\sigma_z \quad \text{and cyclic permutations} \quad (14.3.38)$$

we may establish a very useful identity (Exercise 14.3.4):

$$(\mathbf{A} \cdot \boldsymbol{\sigma})(\mathbf{B} \cdot \boldsymbol{\sigma}) = \mathbf{A} \cdot \mathbf{B} I + i(\mathbf{A} \times \mathbf{B}) \cdot \boldsymbol{\sigma} \quad (14.3.39)$$

where \mathbf{A} and \mathbf{B} are vectors or vector operators that commute with $\boldsymbol{\sigma}$.

(7) Combining Eqs. (14.3.33), (14.3.34), and (14.3.35) we find that

$$\text{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}, \quad i, j = x, y, z \quad (14.3.40a)$$

Let us view the identity, I , as the fourth Pauli matrix. If we call it σ_0 , then

$$\text{Tr}(\sigma_\alpha \sigma_\beta) = 2\delta_{\alpha\beta} \quad (\alpha, \beta = x, y, z, 0) \quad (14.3.40b)$$

This equation implies that the σ_α matrices are *linearly independent*. By this I mean as usual that

$$\sum_\alpha c_\alpha \sigma_\alpha = 0 \rightarrow c_\alpha = 0 \quad \text{for all } \alpha \quad (14.3.41)$$

To prove this for say c_β , multiply both sides by σ_β and take the trace.

[‡] See Exercise 12.5.4.

[§] From now on α, β will run over four values $x, y, z, 0$; while i, j will run over just x, y , and z .

Since any 2×2 matrix M has only four independent (complex) degrees of freedom, it may be written as

$$M = \sum m_\alpha \sigma_\alpha \quad (14.3.42)$$

To find m_β , we multiply by σ_β and take the trace, to find

$$m_\beta = \frac{1}{2} \text{Tr}(M\sigma_\beta) \quad (14.3.43)$$

(The coefficients m_α will be complex in general, and real if M is Hermitian.)

Thus, any operator in \mathbb{V}_s may be expressed in terms of the σ_α , which form a basis that is orthonormal with respect to the inner product $\frac{1}{2} \text{Tr}(\sigma_\alpha \sigma_\rho)$.[‡]

Explicit Forms of Rotation Operators

The fact that $(\hat{n} \cdot \boldsymbol{\sigma})^2 = I$ greatly simplifies many computations and allows us to compute in closed form several operators such as $U(t) = \exp(-iHt/\hbar)$, $U[R(\theta)] = \exp(-i\theta \cdot \mathbf{S}/\hbar)$, which are intractable in infinite-dimensional spaces. In this section we consider the rotation operators, and in the next, the propagator.

Consider

$$\begin{aligned} U[R(\theta)] &= \exp(-i\theta \cdot \mathbf{S}/\hbar) = \exp(-i\theta \cdot \boldsymbol{\sigma}/2) \\ &= \exp\left[-i\left(\frac{\theta}{2}\right)\hat{\theta} \cdot \boldsymbol{\sigma}\right] \\ &= \sum_{n=0}^{\infty} \left(-\frac{i\theta}{2}\right)^n \frac{1}{n!} (\hat{\theta} \cdot \boldsymbol{\sigma})^n \\ &= I + \left(-\frac{i\theta}{2}\right)\hat{\theta} \cdot \boldsymbol{\sigma} + \frac{1}{2!} \left(-\frac{i\theta}{2}\right)^2 I + \frac{1}{3!} \left(-\frac{i\theta}{2}\right)^3 (\hat{\theta} \cdot \boldsymbol{\sigma}) + \dots \end{aligned}$$

Grouping together the coefficients of I and $\hat{\theta} \cdot \boldsymbol{\sigma}$, we get

$$U[R(\theta)] = \cos(\theta/2)I - i \sin(\theta/2)\hat{\theta} \cdot \boldsymbol{\sigma} \quad (14.3.44)$$

Let us put this operator to a test. Suppose we have a particle with spin up along the z direction, i.e., in the state $|_0^1\rangle$. If we want to get from this a particle in the state $|\hat{n}, +\rangle$, it is clear that we must rotate $|_0^1\rangle$ by an angle θ about an axis perpendicular to the z axis and the \hat{n} axis. Thus the rotation angle is

$$\theta = \theta \hat{\theta} = \theta \frac{\mathbf{k} \times \hat{n}}{|\mathbf{k} \times \hat{n}|} \quad (14.3.45)$$

[‡] The inner product between two matrices M and M' acting on \mathbb{V}_s is actually $\text{Tr}(MM'^\dagger)$. However, the dagger is irrelevant for the Hermitian σ 's. It is an interesting exercise to check that this inner product obeys the three axioms.

where \mathbf{k} is the unit vector along the z axis. Since $\hat{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$, it follows that

$$\hat{\theta} = \frac{1}{\sin \theta} (-\sin \theta \sin \phi, \sin \theta \cos \phi, 0) = (-\sin \phi, \cos \phi, 0) \quad (14.3.46)$$

The rotation matrix is, from Eq. (14.3.44),

$$\exp\left(-\frac{i\theta}{2} \hat{\theta} \cdot \boldsymbol{\sigma}\right) = \begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) e^{-i\phi} \\ \sin(\theta/2) e^{i\phi} & \cos(\theta/2) \end{bmatrix} \quad (14.3.47)$$

According to our mnemonic, the first column gives the rotated version of $|\downarrow\rangle$. We see that it agrees with $|\hat{n}, +\rangle$ given in Eq. (14.3.28) up to an overall phase.

Here is a summary of useful formulas that were derived or simply stated:

$$\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\sigma}$$

$$[\sigma_i, \sigma_j]_+ = 2I\delta_{ij}$$

$$[\sigma_i, \sigma_j] = 2i \sum_k \epsilon_{ijk} \sigma_k$$

$$(\hat{n} \cdot \boldsymbol{\sigma})^2 = I$$

$$\text{Tr } \sigma_i = 0$$

$$\text{Tr}(\sigma_\alpha \sigma_\beta) = 2\delta_{\alpha\beta} \quad (\alpha, \beta = x, y, z, 0)$$

$$\exp\left(-i \frac{\theta}{2} \hat{\theta} \cdot \boldsymbol{\sigma}\right) = \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) \hat{\theta} \cdot \boldsymbol{\sigma}$$

$$(\mathbf{A} \cdot \boldsymbol{\sigma})(\mathbf{B} \cdot \boldsymbol{\sigma}) = (\mathbf{A} \cdot \mathbf{B})I + i(\mathbf{A} \times \mathbf{B}) \cdot \boldsymbol{\sigma}$$

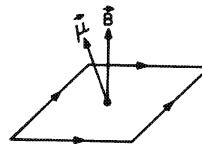
*Exercise 14.3.2.** (1) Show that the eigenvectors of $\boldsymbol{\sigma} \cdot \hat{n}$ are given by Eq. (14.3.28).
(2) Verify Eq. (14.3.29).

*Exercise 14.3.3.** Using Eqs. (14.3.32) and (14.3.33) show that the Pauli matrices are traceless.

*Exercise 14.3.4.** Derive Eq. (14.3.39) in two different ways.

- (1) Write $\sigma_i \sigma_j$ in terms of $[\sigma_i, \sigma_j]_+$ and $[\sigma_i, \sigma_j]$.
- (2) Use Eqs. (14.3.42) and (14.3.43).

Figure 14.1. In the figure, \mathbf{B} is the magnetic field and μ is the magnetic moment of the loop. The direction of the arrows in the loop is that of the current.



Exercise 14.3.5. Express the following matrix M in terms of the Pauli matrices:

$$M = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

Exercise 14.3.6. (1) Argue that $|\hat{n}, +\rangle = U[R(\phi\mathbf{k})]U[R(\theta\mathbf{j})]|s_z = \hbar/2\rangle$. (2) Verify by explicit calculation.

Exercise 14.3.7. Express the following as linear combinations of the Pauli matrices and I :

- (1) $(I + i\sigma_x)^{1/2}$. (Relate it to half a certain rotation.)
- (2) $(2I + \sigma_x)^{-1}$.
- (3) σ_x^{-1} .

*Exercise 14.3.8.** (1) Show that any matrix that commutes with σ is a multiple of the unit matrix.

(2) Show that we cannot find a matrix that anticommutes with all three Pauli matrices. (If such a matrix exists, it must equal zero.)

14.4. Spin Dynamics

Since the quest for the spin Hamiltonian is based on classical analogy, let us recall some basic ideas from classical magnetostatics. Consider a square loop (Fig. 14.1) carrying a current I , in a magnetic field \mathbf{B} . From standard magnetostatics (force per unit length on a current-carrying conductor etc.) one can show that the torque on the loop is

$$\mathbf{T} = \mu \times \mathbf{B} \quad (14.4.1)$$

where μ , the *magnetic moment*, is given by

$$\mu = \frac{I \cdot A}{c} \mathbf{e}_\perp \quad (14.4.2)$$

where A is the area of the loop, c is the velocity of light, and \mathbf{e}_\perp is a unit vector perpendicular to the plane of the loop.[‡] The effect of \mathbf{T} will be to rotate the loop until μ and \mathbf{B} are parallel.

Since we finally wish to address a quantum mechanical problem, it is preferable to summarize the interaction between the loop and the magnetic field in terms of

[‡] The sense of \mathbf{e}_\perp is related to the current flow by the right-hand rule.

the potential energy associated with the torque: If θ is the angle between μ and \mathbf{B} , the interaction energy is[‡]

$$\mathcal{H}_{\text{int}} = \int T(\theta) d\theta = \int \mu B \sin \theta d\theta = -\mu B \cos \theta = -\mu \cdot \mathbf{B} \quad (14.4.3)$$

As we would expect, this energy is minimized, i.e., a stable configuration obtains, when μ and \mathbf{B} are parallel.

Although we derived the above equations for a square loop, they are true for any tiny planar loop, over whose extent \mathbf{B} is constant. So we may apply it to the following problem. Imagine a particle of mass m , charge q , moving in a circular orbit of radius r . The current associated with this charge is

I = charge flow past any point in the circle per second

$$= \frac{qv}{2\pi r} \quad (14.4.4)$$

and the magnetic moment has a magnitude

$$\mu = \frac{qv}{2\pi r} \cdot \frac{\pi r^2}{c} = \frac{qvr}{2c} = \left(\frac{q}{2mc} \right) mvr = \frac{q}{2mc} \cdot l \quad (14.4.5)$$

where l is the magnitude of the angular momentum. Since μ and \mathbf{l} are parallel,

$$\mu = \left(\frac{q}{2mc} \right) \mathbf{l} \quad (14.4.6)$$

The ratio of μ to \mathbf{l} is called the *gyromagnetic ratio* γ . For the particle considered above,

$$\gamma = \frac{q}{2mc} \quad (14.4.7)$$

In the case of the current loop, it was stated that the effect of the torque \mathbf{T} is to cause μ to align with \mathbf{B} . This picture changes when μ has its origin in angular momentum, as is the case for the particle in question. In this case, \mathbf{T} causes a

[‡] This is not the full Hamiltonian (for it does not include the kinetic energy of the loop) but just the potential energy of interaction with the magnetic field.

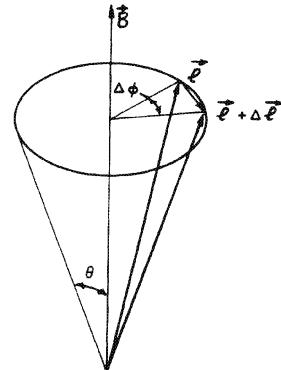


Figure 14.2. In a small time Δt , the tip of the \mathbf{l} vector precesses by an angle $\Delta\phi$ around the magnetic field vector.

precession of μ around \mathbf{B} . We may see this as follows (see Fig. 14.2). The equation of motion is

$$\mathbf{T} = \frac{d\mathbf{l}}{dt} = \mu \times \mathbf{B} = \gamma(\mathbf{l} \times \mathbf{B}) \quad (14.4.8)$$

So in a small time Δt ,

$$\Delta\mathbf{l} = \gamma(\mathbf{l} \times \mathbf{B})\Delta t$$

or

$$\Delta l = \gamma l B \sin \theta \Delta t$$

Since $\Delta\mathbf{l}$ is perpendicular to \mathbf{l} , the tip of the \mathbf{l} vector moves by an angle

$$\Delta\phi = \left(\frac{-\Delta l}{l \sin \theta} \right) = (-\gamma B) \Delta t \quad (14.4.9)$$

i.e., precesses at a frequency

$$\omega_0 = -\gamma \mathbf{B} \quad (14.4.10)$$

Orbital Magnetic Moment in Quantum Theory

These ideas reemerge in the quantum theory. The Hamiltonian for a particle of mass m and charge q in a magnetic field is

$$H = \frac{(\mathbf{P} - q\mathbf{A}/c)^2}{2m} = \frac{|\mathbf{P}|^2}{2m} - \frac{q}{2mc} (\mathbf{P} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{P}) + \frac{q^2 |\mathbf{A}|^2}{2mc^2} \quad (14.4.11)$$

Let

$$\mathbf{A} = \frac{B}{2}(-y\mathbf{i} + x\mathbf{j}) \quad (14.4.12)$$

so that

$$\nabla \times \mathbf{A} = \mathbf{B} = B\mathbf{k} \quad (14.4.13)$$

is constant and along the z axis. We will assume B is small and drop the last term in H , quadratic in B . When the middle term acts on any $|\psi\rangle$,

$$\begin{aligned} (\mathbf{P} \cdot \mathbf{A})|\psi\rangle &\rightarrow -i\hbar \nabla \cdot (\mathbf{A}\psi) \\ &= -i\hbar[(\nabla \cdot \mathbf{A})\psi + \mathbf{A} \cdot \nabla \psi] \\ &= (-i\hbar \mathbf{A} \cdot \nabla)\psi \rightarrow (\mathbf{A} \cdot \mathbf{P})|\psi\rangle \end{aligned}$$

since $\nabla \cdot \mathbf{A} = 0$ here.[‡] Thus the interaction Hamiltonian is

$$\begin{aligned} H_{\text{int}} &= -\frac{q}{2mc}(2\mathbf{A} \cdot \mathbf{P}) \\ &= -\frac{q}{mc}\frac{B}{2}(-YP_x + XP_y) \\ &= -\frac{q}{2mc}\mathbf{L} \cdot \mathbf{B} \equiv -\boldsymbol{\mu} \cdot \mathbf{B} \end{aligned} \quad (14.4.14)$$

so that

$$\boldsymbol{\mu} = \frac{q}{2mc}\mathbf{L} \quad (14.4.15)$$

exactly as in the classical case. (We use the same symbol $\boldsymbol{\mu}$ to denote the classical variable and the quantum operator. We will occasionally violate our convention in this manner, so as to follow other widely used conventions.)

If we project this relation along the z axis, we get

$$\mu_z = \frac{q}{2mc}L_z = \frac{q\hbar}{2mc}(0, \pm 1, \pm 2, \dots)$$

[‡] It is shown in Section 18.4 that \mathbf{A} corresponding to a given \mathbf{B} can always be chosen divergenceless.

The quantity $q\hbar/2mc$ is called the *Bohr magneton* of the particle. The *electron Bohr magneton*, simply called the *Bohr magneton*, has a magnitude

$$\frac{e\hbar}{2mc} \approx 0.6 \times 10^{-8} \text{ eV/G} \quad (14.4.16)$$

where m is the mass of the electron and G stands for gauss. The *nucleon Bohr magneton* is about 2000 times smaller:

$$\frac{e\hbar}{2Mc} \approx 0.3 \times 10^{-11} \text{ eV/G} \quad (14.4.17)$$

where M is the nucleon (proton or neutron)[‡] mass. (The nucleon Bohr magneton is also called the *nuclear Bohr magneton*.)

It may be verified, by the use of Ehrenfest's theorem, that $\langle \mathbf{L} \rangle$ precesses around the constant field \mathbf{B} just as \mathbf{l} would (Exercise 14.4.1).

Spin Magnetic Moment

Armed with all this knowledge, we now address the problem of how the electron interacts with an external magnetic field. We assume once again that there is a magnetic moment operator μ associated with the spin angular momentum. Since any operator on \mathbb{V}_s is a linear combination of the identity and the spin operators, and since μ is a vector operator, we conclude that

$$\mu = \gamma \mathbf{S} \quad (14.4.18a)$$

where γ is a constant. Since $\gamma = -e/2mc$ for the orbital case, let us write

$$\mu = g(-e/2mc)\mathbf{S} \quad (14.4.18b)$$

where g is a constant. We also assume that

$$\begin{aligned} H_{\text{int}} &= -\mu \cdot \mathbf{B} = \frac{ge}{2mc} \mathbf{S} \cdot \mathbf{B} \\ &= \frac{ge\hbar}{4mc} \boldsymbol{\sigma} \cdot \mathbf{B} \end{aligned} \quad (14.4.19)$$

The intrinsic magnetic moment due to spin is $g/2$ magnetons. Our present formalism does not tell us what g is; to find it we must confront the above H with experiment and hope that for some value of g it gives the right physics. This happens to be the case, and the experimental value for g is very close to 2. We assume

[‡] Recall that these two are nearly equal: $M_p c^2 = 938.28 \text{ MeV}$, while $M_n c^2 = 939.57 \text{ MeV}$.

hereafter that

$$g = 2 \quad (14.4.20)$$

Thus the gyromagnetic ratio for spin is *twice as big* as for orbital angular momentum.

Why is $g \approx 2$? And why isn't it exactly equal to 2, which would be much prettier? Our formalism doesn't tell us. But it is irresistible to digress and mention that the Dirac equation, which we will discuss in Chapter 20, predicts that $g=2$ exactly. Quantum electrodynamics, which we will not discuss in this book, predicts that the Dirac result will receive corrections that can be calculated in a power series in α , the fine-structure constant. The physics behind the corrections is the following. Recall that the interaction between the electron and other charged particles is mediated by the exchange of photons. Occasionally, an electron will recapture the photon it emitted. Between the emission and reabsorption, the system that originally contained just the electron will contain an electron and the photon. If the magnetic moment of the system is probed at this time, we can get a result that corresponds to $g \neq 2$, since the electron in the two-particle system has both spin and orbital angular momentum. In fact, quantum electrodynamics predicts that what we call the electron is really a superposition of states that contain one Dirac electron, a Dirac electron and a photon, a Dirac electron, several photons, and several electron-positron pairs, etc.[‡] The reason the observed value of g is so close to the Dirac value of 2 is that configurations of increasing complexity are suppressed by increasing powers of the fine-structure constant in the superposition. Thus the simplest configuration, with just the Dirac electron, will dominate the picture and the complicated states will provide smaller and smaller corrections to the result $g=2$. The corrections may be calculated in a power series in α :

$$g = 2 \left[1 + \frac{1}{2\pi} \cdot \alpha + O(\alpha^2) + \dots \right]$$

which has been evaluated to order α^3 . The result is[§]

$$g_{\text{theory}} = 2[1.001159652140(\pm 28)]$$

where the error ± 28 in the last two digits is mostly due to uncertainties in the value of α itself and in the numerical evaluation of some of the integrals in the calculation.

In addition to higher-order corrections, this result also receives corrections due to other interactions of the electron, i.e., due to its ability to exchange other quanta such as the graviton. But these effects are negligible to the accuracy considered above. The experimental value of g is^{||}

$$g_{\text{exp}} = 2[1.0011596521884(\pm 43)]$$

[‡] The time-energy uncertainty relation allows the production of these particles for short times.

[§] T. Kinoshita and W. B. Lindquist, *Phys. Rev.* **D42**, 636, 1990.

^{||} R. S. Van Dyck, P. B. Schwinberg, and H. G. Dehmelt, *Phys. Rev. Lett.* **59**, 26, 1987.

in splendid agreement with theory. Feynman has pointed out that this is equivalent to predicting and measuring the distance between New York and Los Angeles to within the width of a human hair!

The theoretical situation is bad for the nucleons. The reason is that these participate in strong interactions as well, i.e., can emit and absorb pions etc., and the counterpart of α is large (≈ 15). In other words, the state with just the Dirac particle no longer dominates, and the corrections are no longer tiny. We can of course measure g experimentally, and the result is (to two places)

$$\gamma_{\text{proton}} = 5.6 (e/2Mc)$$

$$\gamma_{\text{neutron}} = -3.8 (e/2Mc)$$

Dirac theory predicts $\gamma = e/Mc$ or $g = 2$ for the proton and $\gamma = 0$ for the neutral neutron. The nonzero γ of the neutron reflects the fact that the neutron can be in a state that has particles with compensating electrical charges but not necessarily compensating magnetic moments.

Because of their large masses, the magnetic moments of the nucleons are negligible compared to that of the electron.[‡]

Let us now return to the dynamics of spin in a magnetic field \mathbf{B} . All we need from now on is the Hamiltonian

$$H = -\boldsymbol{\mu} \cdot \mathbf{B} = -\gamma \mathbf{S} \cdot \mathbf{B} \quad (14.4.21)$$

where

$$\gamma = \frac{-e \cdot 2}{2mc} = \frac{-e}{mc} \quad (14.4.22)$$

Let $|\psi(0)\rangle$ be the initial state of the electron. The state at a later time is

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle$$

where

$$U(t) = e^{-iHt/\hbar} = e^{+i\gamma t(\mathbf{S} \cdot \mathbf{B})/\hbar} \quad (14.4.23)$$

Since $\exp(-i\boldsymbol{\theta} \cdot \mathbf{S}/\hbar)$ is the operator that rotates by $\boldsymbol{\theta}$, the effect of $U(t)$ is clearly to rotate the state by an angle

$$\boldsymbol{\theta}(t) = -\gamma \mathbf{B}t \quad (14.4.24)$$

[‡] End of digression.

It follows that $\langle \mathbf{S} \rangle$ will precess around \mathbf{B} at a frequency $\omega_0 = -\gamma \mathbf{B}$. If this seems too abstract, let us consider a concrete example. Let \mathbf{B} be along the z axis: $\mathbf{B} = B \mathbf{k}$. In this case

$$\begin{aligned} U(t) &= \exp(i\gamma t S_z B / \hbar) \\ &= \exp(i\omega_0 t \sigma_z / 2) \quad (\omega_0 = \gamma B) \end{aligned}$$

Since σ_z is diagonal,

$$U(t) \rightarrow \begin{bmatrix} e^{i\omega_0 t / 2} & 0 \\ 0 & e^{-i\omega_0 t / 2} \end{bmatrix}$$

Consider an electron that starts out in the state $|\hat{n}, +\rangle$:

$$|\psi(0)\rangle = |\hat{n}, +\rangle \rightarrow \begin{bmatrix} \cos(\theta/2) e^{-i\phi/2} \\ \sin(\theta/2) e^{i\phi/2} \end{bmatrix}$$

in which case

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \rightarrow \begin{bmatrix} \cos(\theta/2) e^{-i(\phi - \omega_0 t)/2} \\ \sin(\theta/2) e^{i(\phi - \omega_0 t)/2} \end{bmatrix}$$

i.e., ϕ decreases at a rate ω_0 .

Paramagnetic Resonance

Consider a classical magnetic moment μ in a field $\mathbf{B}_0 = B_0 \mathbf{k}$. It will precess around \mathbf{B}_0 at a frequency

$$\omega_0 = -\gamma \mathbf{B}_0$$

Suppose we view this process in a frame that is rotating at a frequency ω parallel to \mathbf{B}_0 . In this rotating frame, the precession frequency will be

$$\omega_r = \omega_0 - \omega = -\gamma \mathbf{B}_0 - \omega = -\gamma (\mathbf{B}_0 + \omega/\gamma) \quad (14.4.25)$$

Thus the effective field in this rotating frame will be

$$\mathbf{B}_r = \mathbf{B}_0 + \omega/\gamma \quad (14.4.26)$$

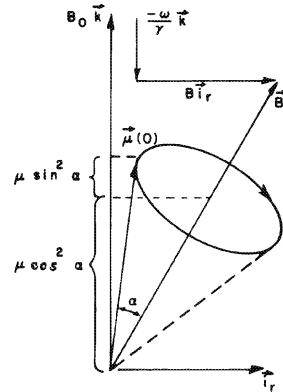


Figure 14.3. The situation in the rotating frame. The effective magnetic field is \mathbf{B}_r . The magnetic moment starts out along the z axis (but is slightly displaced in the figure for clarity) and precesses around \mathbf{B}_r . The z component of the moment oscillates with an amplitude $\mu \sin^2 \alpha$, where α is the opening angle of the cone. At resonance, \mathbf{B}_r lies along the x axis and μ precesses in the plane normal to it. The amplitude of the μ_z oscillation is then at its maximum value of μ .

This result is valid even if ω and \mathbf{B}_0 are not parallel (Exercise 14.4.5). Consider now the problem of interest, where, in a nonrotating (lab) frame

$$\mathbf{B} = B \cos \omega t \mathbf{i} - B \sin \omega t \mathbf{j} + B_0 \mathbf{k} \quad (B \ll B_0) \quad (14.4.27)$$

and at $t=0$,

$$\mu(0) = \mu \mathbf{k} \quad (14.4.28)$$

We would like to find out the evolution of $\mu(t)$. Since \mathbf{B} depends on time, it proves advantageous to first consider the problem in a frame that rotates at the same frequency $\omega = -\omega \mathbf{k}$ as the tiny clockwise rotating field B . In this frame, the rotating component of \mathbf{B} gets frozen (say along the x axis) and the constant component $B_0 \mathbf{k}$ gets reduced as per Eq. (14.4.26) so that the effective, time-independent field is

$$\mathbf{B}_r = Bi_r + (B_0 - \omega/\gamma) \mathbf{k} \quad (14.4.29)$$

where i_r is the unit vector in the x direction in the rotating frame. ($\mathbf{k} = \mathbf{k}_r$, of course.) In this frame, μ will precess around \mathbf{B}_r at a frequency

$$\omega_r = -\gamma \mathbf{B}_r \quad (14.4.30a)$$

where

$$|\omega_r| = \omega_r = \gamma [B^2 + (B_0 - \omega/\gamma)^2]^{1/2} \quad (14.4.30b)$$

It is a simple matter to deduce from Fig. 14.3 that μ_z oscillates as follows:

$$\begin{aligned} \mu_z(t) &= \mu \cos^2 \alpha + \mu \sin^2 \alpha \cos \omega_r t \\ &= \mu_z(0) \left[\frac{(\omega_0 - \omega)^2}{(\omega_0 - \omega)^2 + \gamma^2 B^2} + \frac{\gamma^2 B^2 \cos \omega_r t}{(\omega_0 - \omega)^2 + \gamma^2 B^2} \right] \end{aligned} \quad (14.4.31)$$

This formula for $\mu_z(t)$ applies in the lab frame as well, since μ_z is invariant under z rotations. As ω increases from 0, the z component of \mathbf{B}_r steadily decreases; α , the opening angle of the cone, increases, and the amplitude of oscillation, $\mu \sin^2 \alpha$, grows. At *paramagnetic resonance*, $\omega = \omega_0$, $\mathbf{B}_r = \mathbf{B}_l$, $\alpha = \pi/2$, the cone becomes a circle in the y - z plane, and μ_z oscillates with the largest amplitude μ at a frequency γB . The behavior for $\omega > \omega_0$ is obvious.

What if we apply the rotating field at the resonance frequency, but for a time τ such that

$$\gamma B \tau = \pi/2$$

Such a pulse, called a 90° *pulse*, will swing the magnetic moment into the x - y plane (in either frame). Thereafter μ will precess around $B_0 \mathbf{k}$ at the frequency ω_0 in the lab frame. If we apply a 180° pulse, i.e., choose τ such that

$$\gamma B \tau = \pi$$

the pulse will reverse the sign of μ and leave it pointing down the z axis, where it will stay (in either frame).

These results for the classical moment μ apply to the expectation value $\langle \mu \rangle$ in the quantum problem, as you may verify by doing Exercise 14.4.1, where it is proved in general, and Exercise 14.4.3, where the explicit verification in this case is discussed.

Negative Absolute Temperature (Optional Digression)

The absolute zero of temperature, 0 K, ($\simeq -273^\circ\text{C}$) is defined so that nothing can be colder, yet here we speak of negative absolute temperatures! There is no conflict, however, since we will see that negative temperatures are hotter than positive temperatures! Before you give up all faith, let us quickly sort this thing out.

The absolute temperature T is defined as follows:

$$\beta = \frac{1}{kT} = \frac{1}{k} \frac{\partial S}{\partial E} = \frac{\partial \ln \Omega(E)}{\partial E} \quad (14.4.32)$$

where β is the *thermodynamic temperature*, k is Boltzmann's constant, $S = k \ln \Omega$ is the entropy and $\Omega(E)$ is the number of states available to the system as a function of its energy. (Ω depends on other variables, assumed to be fixed.) In most systems, β is positive because adding energy only opens up more states and increases Ω . For instance, if we have a box of gas molecules, they all stay in the ground state at $T = 0$. So, $S = k \ln \Omega = k \ln 1 = 0$. As we pump in energy, they can occupy higher states, and S and Ω can increase without limit.

Consider now a collection of N spin-half particles sitting on some crystal lattice which is immersed in a field $\mathbf{B} = B_0 \mathbf{k}$. Each magnetic moment (or spin) has two states only, with energies $E = \pm \mu B_0$, where μ is the magnitude of the magnetic moment. At $T = 0$ K, all are in the ground state (μ parallel to \mathbf{B}); $\Omega = 1$, and $S = 0$. The system has a magnetic moment $\mathbf{M} = n\mu \mathbf{k}$. If we pump in energy $2\mu B_0$, one of the moments can move to the upper energy state; there are N ways to pick the one that moves

up, so that $\Omega = N$ and $S = k \ln N$. Clearly β and T are positive. As we pump in more and more energy, S keeps growing until half are up and half are down. At this point, S reaches a maximum, $\beta = \partial S / \partial E = 0$, and $T = +\infty$. The system has no mean magnetic moment along the z axis. Pumping in more energy only reduces S , with more and more particles piling up in the *upper* state. So β and T become negative. Finally, when $E = N\mu B_0$, all moments are in the upper energy state (antiparallel to \mathbf{B}), $\mathbf{M} = -N\mu\mathbf{k}$, there is only one such state; $\Omega = 1$ and $S = 0$. This corresponds to $\beta = -\infty$, $T = 0^-$. Thus the sequence of temperatures is $T = 0^+, \dots, 300, \dots, \infty, -\infty, \dots, -300, \dots, 0^-$. In terms of β , there is more continuity: $\beta = \infty, \dots, 0^+, 0^-, \dots, -\infty$. (We should have chosen $-\beta$ as the temperature, for it rises monotonically from $-\infty$ to $+\infty$ as we heat the system.) It should be clear that negative temperatures are hotter than positive temperatures since we go from the latter to the former by pumping in energy. We can also see this by imagining a system at $T = -300$ K brought in contact with identical system at $T = +300$ K. Since the populations of the two systems are identical, except for the change, parallel \leftrightarrow antiparallel, they can increase their entropies by moving toward the state with equal numbers up and down. In this process energy clearly flows from the negative temperature system to the positive temperature system, i.e., the former is hotter. Also note that the final equilibrium temperature is not 0 K but ∞ K.

How does one prepare negative temperatures in the lab? One takes a sample at room temperature, say at $T = 300$ K. It will have more moments parallel than antiparallel:

$$\frac{N(\text{parallel})}{N(\text{antiparallel})} = \frac{e^{-(\beta\mu B_0)}}{e^{-\beta\mu B_0}} = e^{2\beta\mu B_0} > 1 \quad (14.4.33)$$

and a net magnetic moment \mathbf{M} along the z axis. If one applies a 180° pulse, there will be *population inversion* (parallel \leftrightarrow antiparallel), which amounts to a change in the sign of β and T [see Eq. (14.4.33)]. The spin system cannot stay in this hot state ($T = -300$ K) forever, because it is in contact with the lattice, which will eventually cool it down to room temperature.

The return to thermal equilibrium is easier to observe if one applies a 90° pulse which swings \mathbf{M} into the x - y plane. The temperature now is $T = \infty$ K, since $M_z = 0 \rightarrow N(\text{parallel}) = N(\text{antiparallel}) \rightarrow T = \infty$. Thus \mathbf{M} , which will initially begin to precess around $\mathbf{B} = B_0\mathbf{k}$, will eventually realign itself with \mathbf{B} . The decay of its rotating components in the x - y plane may be observed as follows. Suppose the specimen is a long cylinder whose axis lies in the x - y plane. If one winds a coil around it, the transverse (x - y) components of \mathbf{M} , which simulate a bar magnet rotating in the x - y plane, will induce an oscillating voltage in the coil. The frequency of the (damped) oscillation will be ω_0 and the half-life will be a time τ , called the *transverse relaxation time*.†

*Exercise 14.4.1.** Show that if $H = -\gamma\mathbf{L} \cdot \mathbf{B}$, and \mathbf{B} is position independent,

$$\frac{d\langle \mathbf{L} \rangle}{dt} = \langle \boldsymbol{\mu} \times \mathbf{B} \rangle = \langle \boldsymbol{\mu} \rangle \times \mathbf{B}$$

† The transverse components of \mathbf{M} decay for other reasons, besides restoration of thermal equilibrium. See R. Schumacher, *Magnetic Resonance*, W. A. Benjamin, New York (1970).

Comparing this to Eq. (14.4.8), we see that $\langle \mu \rangle$ evolves exactly like μ . Notice that this conclusion is valid even if \mathbf{B} depends on time and also if we are talking about spin instead of orbital angular momentum. A more explicit verification follows in Exercise 14.4.3.

Exercise 14.4.2. Derive (14.4.31) by studying Fig. 14.3.

*Exercise 14.4.3.** We would like to study here the evolution of a state that starts out as $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and is subject to the \mathbf{B} field given in Eq. (14.4.27). This state obeys

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi\rangle \quad (14.4.34)$$

where $H = -\gamma \mathbf{S} \cdot \mathbf{B}$, and \mathbf{B} is time dependent. Since classical reasoning suggests that in a frame rotating at frequency $(-\omega \mathbf{k})$ the Hamiltonian should be time independent and governed by \mathbf{B}_r [Eq. (14.4.29)], consider the ket in the rotating frame, $|\psi_r(t)\rangle$, related to $|\psi(t)\rangle$ by a rotation angle ωt :

$$|\psi_r(t)\rangle = e^{-i\omega t S_z / \hbar} |\psi(t)\rangle \quad (14.4.35)$$

Combine Eqs. (14.4.34) and (14.4.35) to derive Schrödinger's equation for $|\psi_r(t)\rangle$ in the S_z basis and verify that the classical expectation is borne out. Solve for $|\psi_r(t)\rangle = U_r(t)|\psi_r(0)\rangle$ by computing $U_r(t)$, the propagator in the rotating frame. Rotate back to the lab and show that

$$|\psi(t)\rangle \xrightarrow[S_z \text{ basis}]{} \begin{bmatrix} \left[\cos\left(\frac{\omega_r t}{2}\right) + i \frac{\omega_0 - \omega}{\omega_r} \sin\left(\frac{\omega_r t}{2}\right) \right] e^{+i\omega t/2} \\ \frac{i\gamma B}{\omega_r} \sin\left(\frac{\omega_r t}{2}\right) e^{-i\omega t/2} \end{bmatrix} \quad (14.4.36)$$

Compare this to the state $|\hat{n}, +\rangle$ and see what is happening to the spin for the case $\omega_0 = \omega$. Calculate $\langle \mu_z(t) \rangle$ and verify that it agrees with Eq. (14.4.31).

Exercise 14.4.4. At $t=0$, an electron is in the state with $s_z = \hbar/2$. A steady field $\mathbf{B} = Bi$, $B = 100$ G, is turned on. How many seconds will it take for the spin to flip?

Exercise 14.4.5. We would like to establish the validity of Eq. (14.4.26) when $\boldsymbol{\omega}$ and \mathbf{B}_0 are not parallel.

(1) Consider a vector \mathbf{V} in the inertial (nonrotating) frame which changes by $\Delta \mathbf{V}$ in a time Δt . Argue, using the results from Exercise 12.4.3, that the change as seen in a frame rotating at an angular velocity $\boldsymbol{\omega}$, is $\Delta \mathbf{V} - \boldsymbol{\omega} \times \mathbf{V} \Delta t$. Obtain a relation between the time derivatives of \mathbf{V} in the two frames.

(2) Apply this result to the case of \mathbf{l} [Eq. (14.4.8)], and deduce the formula for the effective field in the rotating frame.

Exercise 14.4.6 (A Density Matrix Problem). (1) Show that the density matrix for an ensemble of spin-1/2 particles may be written as

$$\rho = \frac{1}{2}(I + \mathbf{a} \cdot \boldsymbol{\sigma})$$

where \mathbf{a} is a *c*-number vector.

- (2) Show that \mathbf{a} is the mean polarization, $\langle \bar{\sigma} \rangle$.
- (3) An ensemble of electrons in a magnetic field $\mathbf{B} = B\mathbf{k}$, is in thermal equilibrium at temperature T . Construct the density matrix for this ensemble. Calculate $\langle \bar{\mu} \rangle$.

14.5. Return of Orbital Degrees of Freedom

Let us now put back the orbital degrees of freedom. The simplest case is when H is separable:

$$H = H_0 + H_s \quad (14.5.1)$$

so that the energy eigenstates factorize

$$|\psi\rangle = |\psi_0\rangle \otimes |\chi_s\rangle$$

An example is provided by the hydrogen atom, where the Coulomb interaction is independent of spin:

$$H = H_0 \quad (14.5.2)$$

Here the spin is a constant in time, and all that happens is that we attach a constant spinor χ to the wave functions we found in Chapter 13. If we choose χ to be an eigenstate of S_z , we have‡

$$\begin{aligned} |nlmm_s=1/2\rangle &\rightarrow \psi_{nlm}(r, \theta, \phi)\chi_+ & \left[\chi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right] \\ |nlmm_s=-1/2\rangle &\rightarrow \psi_{nlm}(r, \theta, \phi)\chi_- & \left[\chi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right] \end{aligned} \quad (14.5.3)$$

The energy levels are of course unaffected. All we have is a doubling of states, with the electron spin being up or down (the z axis) in each of the orbital states (nlm).

Consider next the problem of the hydrogen atom in a weak magnetic field $\mathbf{B} = B\mathbf{k}$. Although both the proton and the electron couple to \mathbf{B} , the smallness of the ratio m/M allows us to ignore, in the first approximation, the coupling of the proton's intrinsic and orbital magnetic moments [these are of order m/M and $(m/M)^2$ relative

‡ We use the subscript s on m_s to remind us that it measures the *spin* projection: $s_z = m_s\hbar$. It will be dropped whenever it is obvious that we are dealing with spin.

to that of the electron; see Exercise 14.5.1]. Thus we have, from Eqs. (14.4.14) and (14.4.19),

$$H = H_{\text{Coulomb}} - \left(\frac{-eB}{2mc} \right) L_z - \left(\frac{-eB}{mc} \right) S_z \quad (14.5.4)$$

Since the additional terms in H commute with H_{Coulomb} , L^2 , L_z , and S_z , this H is diagonalized by the same states as before, namely, $|nlmm_s\rangle$. The eigenvalues are, however, different:

$$H|nlmm_s\rangle = \left[\frac{-Ry}{n^2} + \frac{eB\hbar}{2mc} (m + 2m_s) \right] |nlmm_s\rangle \quad (14.5.5)$$

The degeneracy is greatly reduced by the \mathbf{B} field. The ground state, which was twofold degenerate, splits into two levels:

$$E_{n=1} = -Ry \pm \frac{e\hbar B}{2mc} \quad (14.5.6)$$

The second, which was eightfold degenerate, splits into five levels:

$$E_{n=2} = -\frac{Ry}{4} + \frac{eB\hbar}{2mc} \times \begin{bmatrix} 2(m=1, m_s=1/2) \\ 1(m=0, m_s=1/2)(l=0 \text{ or } 1) \\ 0(m=1, m_s=-1/2, \text{ or } m=-1, m_s=1/2) \\ -1(m=0, m_s=-1/2)(l=0 \text{ or } 1) \\ -2(m=-1, m_s=-1/2) \end{bmatrix} \quad (14.5.7)$$

and so on. In a multielectron atom, one simply adds the contributions from all the electrons. The splitting of levels leads to an increase in the number of spectral lines; where there was one, there will now be several, and the spacing between them may be varied by varying B . This phenomenon is called the *Zeeman effect*.

Consider lastly the Hamiltonian

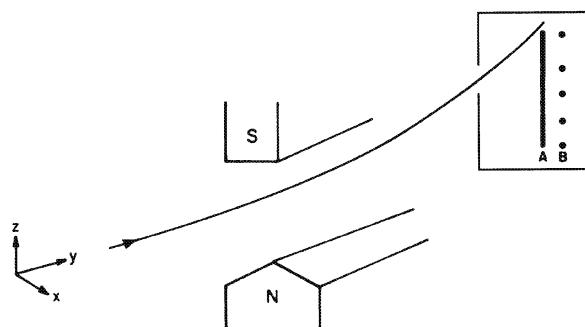
$$H = H_{\text{Coulomb}} + a\mathbf{L} \cdot \mathbf{S} \quad (14.5.8)$$

whose origin will be explained in a later chapter. For the present, we note that it is not separable, and consequently the spin and orbital degrees of freedom are coupled in their time evolution. The eigenstates of H will not be simply products of orbital and spin parts, but instead superpositions of such states that diagonalize $\mathbf{L} \cdot \mathbf{S}$. The details will be explained in the next chapter.

*Exercise 14.5.1.** (1) Why is the coupling of the proton's intrinsic moment to \mathbf{B} an order m/M correction to Eq. (14.5.4)?

(2) Why is the coupling of its orbital motion an order $(m/M)^2$ correction? (You may reason classically in both parts.)

Figure 14.4. The Stern-Gerlach experiment. A beam of particles endowed with magnetic moments enters the inhomogeneous field. Classically the beam is expected to fan out and produce a continuous trace (A) on the screen. What one observes is a set of discrete dots (B). This implies the quantization of magnetic moment and angular momentum.



*Exercise 14.5.2.** (1) Estimate the relative size of the level splitting in the $n=1$ state to the unperturbed energy of the $n=1$ state, when a field $\mathbf{B}=1000\text{kG}$ is applied.

(2) Recall that we have been neglecting the order B^2 term in H . Estimate its contribution in the $n=1$ state relative to the linear $(-\boldsymbol{\mu} \cdot \mathbf{B})$ term we have kept, by assuming the electron moves on a classical orbit of radius a_0 . Above what $|\mathbf{B}|$ does it begin to be a poor approximation?

The Stern-Gerlach (SG) Experiment

We now consider (in simplified form) the SG experiment, which clearly displays the quantization of angular momentum (along any direction). The apparatus (Fig. 14.4) consists of north and south pole pieces, between which is an *inhomogeneous* magnetic field. A beam of (particles with) magnetic moments, traveling along the y axis, enters the apparatus in a region where \mathbf{B} is predominantly along the z axis and $\partial B_z / \partial z < 0$. What do we expect will happen classically? If we pretend that the magnetic moment is due to a pair of equal and opposite (fictitious) magnetic charges, it is clear that any inhomogeneity in \mathbf{B} can lead to a net force on the dipole. This is confirmed if we calculate the force associated with the gradient of the interaction energy

$$\mathbf{F} = -\nabla \mathcal{H} = \nabla(\boldsymbol{\mu} \cdot \mathbf{B}) = (\boldsymbol{\mu} \cdot \nabla) \mathbf{B} = \mu_z \frac{\partial B_z}{\partial z} \mathbf{k} \quad (14.5.9)$$

[We have used the identity $\nabla(\boldsymbol{\mu} \cdot \mathbf{B}) = (\boldsymbol{\mu} \cdot \nabla) \mathbf{B} + (\mathbf{B} \cdot \nabla) \boldsymbol{\mu} + \boldsymbol{\mu} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \boldsymbol{\mu})$. In the present case, $\boldsymbol{\mu}$ is not a function of \mathbf{r} , and by Maxwell's equations, $\nabla \times \mathbf{B} = 0$. Both F_x and F_y vanish on average due to the precession of spin in the $x-y$ plane.] Classically, since μ_z is continuous, the beam is expected to fan out and produce a continuous trace (A in figure) on a screen placed behind the magnet. The actual experiment performed with atoms reveals a series of discrete dots (B in figure). We understand this in semiclassical terms, by saying that μ_z in Eq. (14.5.9) is discrete and therefore so is the angular momentum along the z axis.

This experiment can also be used to reveal the existence of electron spin. For example, if we send in a beam of hydrogen atoms in their ground state, the beam will split into two parts.

Let us describe the above-mentioned hydrogen atom experiment in quantum mechanical terms. Suppose the initial state of a hydrogen atom is

$$\psi_{\text{initial}} = \psi_y(\mathbf{r}_{\text{CM}}) \psi_{100}(\mathbf{r}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (14.5.10)$$

where ψ_y is a wave packet drifting along the y axis that describes the CM motion, ψ_{100} is the ground state wave function, and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is the electron spinor. (The proton spin is ignored, for the associated magnetic moment is too small to affect the dynamics.) Since the electron spin is up, its μ_z is down. Since $\partial B_z / \partial z < 0$, the classical force on the atom is up. So by Ehrenfest's theorem[‡] we expect the atom to emerge from the apparatus in a state (up to a phase factor)

$$\psi_{\text{out}} = \psi_{y,+z}(\mathbf{r}_{\text{CM}}) \psi_{100}(\mathbf{r}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (14.5.11)$$

where $\psi_{y,+z}$ describes a wave packet that is displaced (relative to the incoming one) along the positive z axis and has also a small velocity in the same direction. Likewise, if the electron spinor had initially been $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, the CM would have emerged in the state $\psi_{y,-z}$ (in the same notation). More generally, if

$$\psi_{\text{initial}} = \psi_y \psi_{100} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \psi_y \psi_{100} \begin{bmatrix} \alpha \\ 0 \end{bmatrix} + \psi_y \psi_{100} \begin{bmatrix} 0 \\ \beta \end{bmatrix} \quad (14.5.12)$$

then, by the linearity of Schrödinger's equation

$$\psi_{\text{out}} = \psi_{y,+z} \psi_{100} \begin{bmatrix} \alpha \\ 0 \end{bmatrix} + \psi_{y,-z} \psi_{100} \begin{bmatrix} 0 \\ \beta \end{bmatrix} \quad (14.5.13)$$

Assuming $\psi_{y,\pm z}$ are narrow packets with no overlap, we see that the SG apparatus has introduced a correlation between the spin and orbital coordinates: if we catch (by placing a screen) the outgoing atom above the original line of flight (i.e., in a region where $\psi_{y,+z}$ is peaked) it will have spin up, while if we catch it below, the spin is down.

The SG apparatus can be used to prepare a state of definite spin orientation: to get a pure spin up/down beam we simply block the lower/upper beam. But note that the filtering process changes the average z component of linear momentum. This can be undone and the particle restored its original momentum (but filtered with respect to spin) if we place some more magnets (with \mathbf{B} along the z axis) behind this apparatus. With this modification (which is assumed in the following exercises) the

[‡] Recall the warning at the end of Chapter 6. In the present case, the system follows the classical trajectory (approximately) thanks to the massive proton. If we send in just the electron, quantum fluctuations would wipe out the effect. See, for example, pages 324–330 of G. Baym, *Lectures on Quantum Mechanics*, Benjamin, New York (1969).

only effect of the SG apparatus with one or the other beams blocked is to filter the spin without affecting the orbital motion.

*Exercise 14.5.3.** A beam of spin-1/2 particles moving along the y axis goes through two collinear SG apparatuses, both with lower beams blocked. The first has its \mathbf{B} field along the z axis and the second has its \mathbf{B} field along the x axis (i.e., is obtained by rotating the first by an angle $\pi/2$ about the y axis). What fraction of particles leaving the first will exit the second? If a third filter that transmits only spin up along the z axis is introduced, what fraction of particles leaving the first will exit the third? If the middle filter transmits both spins up and down (no blocking) the x axis, but the last one transmits only spin down the z axis, what fraction of particles leaving the first will leave the last?

Exercise 14.5.4. A beam of spin-1 particles, moving along the y axis, is incident on two collinear SG apparatuses, the first with \mathbf{B} along the z axis and the second with \mathbf{B} along the z' axis, which lies in the $x-z$ plane at an angle θ relative to the z axis. Both apparatuses transmit only the uppermost beams. What fraction leaving the first will pass the second?

15

Addition of Angular Momenta

15.1. A Simple Example

Consider a system of two spin-1/2 particles (whose orbital degrees of freedom we ignore). If \mathbf{S}_1 and \mathbf{S}_2 [‡] are their spin operators, the two-particle Hilbert space $\mathbb{V}_{1\otimes 2}$ is spanned by the four vectors

$$|s_1 m_1\rangle \otimes |s_2 m_2\rangle \equiv |s_1 m_1, s_2 m_2\rangle \quad (15.1.1)$$

which obey

$$S_i^2 |s_1 m_1, s_2 m_2\rangle = \hbar^2 s_i(s_i + 1) |s_1 m_1, s_2 m_2\rangle \quad (15.1.2a)$$

$$S_{iz} |s_1 m_1, s_2 m_2\rangle = \hbar m_i |s_1 m_1, s_2 m_2\rangle \quad (i = 1, 2) \quad (15.1.2b)$$

Since $s_i = 1/2$, and $m_i = \pm 1/2$ has freedom only in sign, let us use the compact notation $|++\rangle, |+-\rangle, |-+\rangle, |--\rangle$ to denote the states. For instance,

$$|+-\rangle = |s_1 = \frac{1}{2}, m_1 = \frac{1}{2}, s_2 = \frac{1}{2}, m_2 = -\frac{1}{2}\rangle \quad (15.1.3)$$

and so on. These four vectors form the *product basis*. They represent states that have well-defined values for the magnitude and z component of the individual spins.

Suppose now that we choose not to look at the individual spins but the system as a whole. What are the possible values for the magnitude and z component of the system spin, and what are the states that go with these values? This is a problem in *addition of angular momenta*, which is the topic of this chapter.

[‡] In terms of the operators $\mathbf{S}_1^{(1)}$ and $\mathbf{S}_2^{(2)}$ which act on the one-particle spaces, $\mathbf{S}_1 = \mathbf{S}_1^{(1)} \otimes I^{(2)}$ and $\mathbf{S}_2 = I^{(1)} \otimes \mathbf{S}_2^{(2)}$.

Consider the operator

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 \quad (15.1.4)$$

which we call the *total angular momentum operator*. That \mathbf{S} is indeed the total angular momentum operator is supported by (1) our intuition; (2) the fact that it is the generator of rotation for the product kets, i.e., rotations of the whole system; (3) the fact that it obeys the commutation rules expected of a generator of rotations, namely,

$$[S_i, S_j] = \sum_k i\hbar \epsilon_{ijk} S_k \quad (15.1.5)$$

as may be readily verified. *Our problem is to find the eigenvalues and eigenvectors of S^2 and S_z .* Consider first

$$S_z = S_{1z} + S_{2z} \quad (15.1.6)$$

which commutes with S_1^2 , S_2^2 , S_{1z} , and S_{2z} . We expect it to be diagonal in the product basis. This is readily verified:

$$\begin{aligned} S_z |++\rangle &= (S_{1z} + S_{2z}) |++\rangle = \left(\frac{\hbar}{2} + \frac{\hbar}{2}\right) |++\rangle \\ S_z |+-\rangle &= 0 |+-\rangle \\ S_z |-+\rangle &= 0 |-+\rangle \\ S_z |--\rangle &= -\hbar |--\rangle \end{aligned} \quad (15.1.7)$$

Thus the allowed values for the total z component are \hbar , 0, and $-\hbar$.

By the method of images (or any other method)

$$S_z \xrightarrow[\text{product basis}]{} \hbar \begin{bmatrix} ++ & +- & -+ & -- \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (15.1.8)$$

Note that the eigenvalue $s_z=0$ is twofold degenerate, and the eigenspace is spanned by the vectors $|+-\rangle$ and $|-+\rangle$. If we form some linear combination, $\alpha|+-\rangle + \beta|-+\rangle$, we still get an eigenstate with $s_z=0$, but this state will not have definite values for S_{1z} and S_{2z} (unless α or $\beta=0$).

Consider next the operator

$$S^2 = (\mathbf{S}_1 + \mathbf{S}_2) \cdot (\mathbf{S}_1 + \mathbf{S}_2) = S_1^2 + S_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2 \quad (15.1.9)$$

Although S^2 commutes with S_1^2 and S_2^2 , it does not commute with S_{1z} and S_{2z} because of the $\mathbf{S}_1 \cdot \mathbf{S}_2$ term, which has S_{1x} , S_{1y} , etc. in it. By explicit computation,

$$S^2 \xrightarrow[\text{product basis}]{} \hbar^2 \begin{bmatrix} ++ & +- & -+ & -- \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad (15.1.10)$$

Thus we see that although $|++\rangle$ and $|--\rangle$ are eigenstates of $S^2[s(s+1)=2]$, the states of zero S_z , namely, $|+-\rangle$ and $|-+\rangle$, are not. However, the following linear combinations are:

$$\frac{|+-\rangle + |-\rangle}{2^{1/2}} \quad (s=1) \quad (15.1.11)$$

$$\frac{|+-\rangle - |-\rangle}{2^{1/2}} \quad (s=0)$$

*Exercise 15.1.1.** Derive Eqs. (15.1.10) and (15.1.11). It might help to use

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = S_{1z}S_{2z} + \frac{1}{2}(S_{1+}S_{2-} + S_{1-}S_{2+}) \quad (15.1.12)$$

This completes the solution to the problem we undertook. The allowed values for total spin are $s=1$ and 0 , while the allowed values of s_z are \hbar , 0 , and $-\hbar$. The corresponding eigenstates in the product basis are

$$\begin{aligned} |s=1 m=1, s_1=1/2 s_2=1/2\rangle &= |++\rangle \\ |s=1 m=0, s_1=1/2 s_2=1/2\rangle &= 2^{-1/2}[|+-\rangle + |-\rangle] \\ |s=1 m=-1, s_1=1/2 s_2=1/2\rangle &= |--\rangle \\ |s=0 m=0, s_1=1/2 s_2=1/2\rangle &= 2^{-1/2}[|+-\rangle - |-\rangle] \end{aligned} \quad (15.1.13)$$

These vectors represent states with well-defined total angular momentum; they form the *total-s basis*. The three spin-1 states are called *triplets* and the solitary spin-0 state is called the *singlet*. The problem of adding angular momenta is essentially a change of basis, from one that diagonalizes $(S_1^2, S_2^2, S_{1z}, S_{2z})$ to one that diagonalizes (S^2, S_z, S_1^2, S_2^2) . We can describe our findings symbolically as

$$1/2 \otimes 1/2 = 1 \oplus 0 \quad (15.1.14)$$

which means that the direct product of two spin-1/2 Hilbert spaces is a direct sum of a spin-1 space and a spin-0 space. The way the dimensionalities work out in

Eq. (15.1.14) is as follows:

$$\begin{aligned} \text{left-hand side: } & (2s_1+1)(2s_2+1) = (2 \times 1/2 + 1)(2 \times 1/2 + 1) = 4 \\ \text{right-hand side: } & \sum_{s=0}^1 (2s+1) = 1 + 3 = 4 \end{aligned} \quad (15.1.15)$$

The decomposition of the direct product space into a sum over spaces with well-defined total spin can also be viewed this way. The rotation operators for the entire system will be 4×4 matrices in the product basis. These matrices are, however, reducible: by changing to the total- s basis, they may be block diagonalized into a 3×3 block (spin-1 sector) and a 1×1 block (spin-0 sector). The total- s basis is, however, irreducible; we cannot further subdivide the spin-1 space into parts that do not mix under rotations.

The total- s states have another property: they have definite symmetry under the exchange of the two particles. The triplets are symmetric and the singlet is antisymmetric. Now, the state vector for two identical spin-1/2 particles must be antisymmetric under the exchange of particle labels, i.e., under the exchange of their spin *and* orbital degrees of freedom. We already know that if Ω is some orbital operator (built out of coordinates and momenta), then

$$|\omega_1\omega_2, S\rangle = 2^{-1/2}[|\omega_1\omega_2\rangle + |\omega_2\omega_1\rangle]$$

and

$$|\omega_1\omega_2, A\rangle = 2^{-1/2}[|\omega_1\omega_2\rangle - |\omega_2\omega_1\rangle]$$

are symmetric and antisymmetric, respectively, under the exchange of the orbital variable. To form the complete state vector, we simply multiply orbital and spin states of *opposite* symmetry:

$$|\omega_1m_1, \omega_2m_2, A\rangle = \begin{cases} |\omega_1\omega_2, S\rangle \otimes \frac{|+-\rangle - |--\rangle}{2^{1/2}} \\ |\omega_1\omega_2, A\rangle \otimes \begin{cases} |++\rangle \\ |+-\rangle + |-\rangle \\ |--\rangle \end{cases} \end{cases} \quad (15.1.16)$$

These vectors provide a complete basis for the Hilbert space of two identical spin-1/2 particles. As an example, consider the ground state of the He atom, which has two electrons. In connection with the periodic table it was said that in this state of lowest energy, both electrons are in the lowest orbital state $|n=1, l=0, m=0\rangle$ [‡] and

[‡] If we neglect interelectron forces, the states allowed to the electrons are hydrogenlike, in that they are labeled $|n, l, m\rangle$. But the energies and wave functions are obtained upon making the replacement $e^2 \rightarrow Ze^2 = 2e^2$.

have opposite spins. We can sharpen that statement now. The orbital part of the ground-state ket is just the direct product,

$$|\psi_o\rangle = |100\rangle \otimes |100\rangle \quad (15.1.17)$$

which is already symmetric. So the spin part must be

$$|\chi_s\rangle = 2^{-1/2}(|+-\rangle - |-+\rangle) \quad (15.1.18)$$

and so

$$|\psi_{\text{ground}}\rangle = |\psi_o\rangle \otimes |\chi_s\rangle \quad (15.1.19)$$

In this state, both the orbital and spin angular momenta are zero.

Let us now return to the problem of just the two spins (and no orbital coordinates). Now that we have two bases, which one should we use? The answer depends on the Hamiltonian. For instance, if the two spins only interact with an external field $\mathbf{B} = B_0 \mathbf{k}$,

$$H = -(\gamma_1 \mathbf{S}_1 + \gamma_2 \mathbf{S}_2) \cdot \mathbf{B} = -B_0(\gamma_1 S_{1z} + \gamma_2 S_{2z}) \quad (15.1.20)$$

the product basis, which diagonalizes S_{1z} and S_{2z} is the obvious choice. (If, however, $\gamma_1 = \gamma_2$, then $H \propto S_z$, and we can use the total- s basis as well.) On the other hand, if the spins are mutually interacting and, say,

$$H = A \mathbf{S}_1 \cdot \mathbf{S}_2 = \frac{1}{2}A(S^2 - S_1^2 - S_2^2) \quad (15.1.21)$$

the total- s basis diagonalizes H .

*Exercise 15.1.2.** In addition to the Coulomb interaction, there exists another, called the *hyperfine interaction*, between the electron and proton in the hydrogen atom. The Hamiltonian describing this interaction, which is due to the magnetic moments of the two particles is,

$$H_{hf} = A \mathbf{S}_1 \cdot \mathbf{S}_2 \quad (A > 0) \quad (15.1.22)$$

(This formula assumes the orbital state of the electron is $|1, 0, 0\rangle$.) The total Hamiltonian is thus the Coulomb Hamiltonian plus H_{hf} .

(1) Show that H_{hf} splits the ground state into two levels:

$$\begin{aligned} E_+ &= -Ry + \frac{\hbar^2 A}{4} \\ E_- &= -Ry - \frac{3\hbar^2 A}{4} \end{aligned} \quad (15.1.23)$$

and that corresponding states are triplets and singlet, respectively.

- (2) Try to estimate the frequency of the emitted radiation as the atom jumps from the triplet to the singlet. To do so, you may assume that the electron and proton are two dipoles μ_e and μ_p separated by a distance a_0 , with an interaction energy of the order[‡]

$$\mathcal{H}_{hf} \approx \frac{\mu_e \cdot \mu_p}{a_0^3}$$

Show that this implies that the constant in Eq. (15.1.22) is

$$A \sim \frac{2e}{2mc} \frac{(5.6)e}{2Mc} \frac{1}{a_0^3}$$

(where 5.6 is the g factor for the proton), and that

$$\Delta E = E_+ - E_- = A\hbar^2$$

is a correction of order $(m/M)\alpha^2$ relative to the ground-state energy. Estimate that the frequency of emitted radiation is a few tens of centimeters, using the mnemonics from Chapter 13. The measured value is 21.4 cm. This radiation, called the *21-cm line*, is a way to detect hydrogen in other parts of the universe.

- (3) Estimate the probability ratio $P(\text{triplet})/P(\text{singlet})$ of hydrogen atoms in thermal equilibrium at room temperature.

15.2. The General Problem

Consider now the general problem of adding two angular momenta \mathbf{J}_1 and \mathbf{J}_2 . What are the eigenvalues and eigenkets of J^2 and J_z , where $\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2$? One way to find out is to mimic the last section: construct the $(2j_1 + 1) \cdot (2j_2 + 1)$ -dimensional matrices J^2 and J_z and diagonalize them. Now, J_z will be diagonal in the product basis itself, for

$$J_z |j_1 m_1, j_2 m_2\rangle = \hbar(m_1 + m_2) |j_1 m_1, j_2 m_2\rangle \quad (15.2.1)$$

It will be a degenerate operator, for there are many ways to build up a total $m = m_1 + m_2$, except when $m = \pm(j_1 + j_2)$ when both angular momenta have maximal projections up/down the z axis. For instance, if $m = j_1 + j_2 - 2$, there are three product kets: $(m_1 = j_1, m_2 = j_2 - 2)$, $(m_1 = j_1 - 1, m_2 = j_2 - 1)$, and $(m_1 = j_1 - 2, m_2 = j_2)$. In each of the degenerate eigenspaces of J_z , we must choose a basis that diagonalizes J^2 (and undiagonalizes J_{1z} and J_{2z}). We can do this by constructing the matrix J^2 and then diagonalizing it. But this can be a tedious business. (If you have done Exercise 15.1.1 you will know that the construction of S^2 is quite tedious even in this four-dimensional case.) There is, however, a more efficient alternative to be described now.

As a first step, we need to know the allowed values for j . Our intuition and our experience from the last section suggest that j can take on values $j_1 + j_2$,

[‡] The description here is oversimplified; both \mathcal{H}_{hf} and H_{hf} are rather tricky to derive. Our aim is just to estimate $|A|$ and not to get into its precise origin.

$j_1 + j_2 - 1, \dots, j_1 - j_2$ (assuming $j_1 \geq j_2$).[‡] Let us check this. The number of product kets is $(2j_1 + 1) \cdot (2j_2 + 1)$. This must equal the number of total- j kets. According to our conjecture, this number is

$$\sum_{j=j_1-j_2}^{j_1+j_2} (2j+1) = \sum_{j=0}^{j_1+j_2} (2j+1) - \sum_{j=0}^{j_1-j_2-1} (2j+1) = (2j_1+1)(2j_2+1) \quad (15.2.2)$$

using the formula

$$\sum_{n=0}^N n = \frac{N(N+1)}{2}$$

We take this to be proof of our conjecture:

$$j_1 \otimes j_2 = (j_1 + j_2) \oplus (j_1 + j_2 - 1) \oplus \dots \oplus (j_1 - j_2) \quad (15.2.3)$$

In other words, the total- j kets are

$$|jm, j_1 j_2\rangle \quad \text{with} \quad j_1 + j_2 \geq j \geq j_1 - j_2, \quad j \geq m \geq -j \quad (15.2.4)$$

Let us write them in the form of an array:

	\nearrow^j				
m		$j_1 + j_2$	$j_1 + j_2 - 1$	\dots	$j_1 - j_2$
	\downarrow	$ j_1 + j_2, j_1 + j_2\rangle$			
		$ j_1 + j_2, j_1 + j_2 - 1\rangle$	$ j_1 + j_2 - 1, j_1 + j_2 - 1\rangle$		$ j_1 - j_2, j_1 - j_2\rangle$
		$ j_1 + j_2, j_1 + j_2 - 2\rangle$	$ j_1 + j_2 - 1, j_1 + j_2 - 2\rangle$	\dots	$ j_1 - j_2, -(j_1 - j_2)\rangle$
		\vdots	\vdots	\dots	\vdots
		$ j_1 + j_2, -(j_1 + j_2 - 2)\rangle$	$ j_1 + j_2 - 1, -(j_1 + j_2 - 2)\rangle$		$ j_1 - j_2, -(j_1 - j_2)\rangle$
		$ j_1 + j_2, -(j_1 + j_2 - 1)\rangle$	$ j_1 + j_2 - 1, -(j_1 + j_2 - 1)\rangle$		
		$ j_1 + j_2, -(j_1 + j_2)\rangle$			

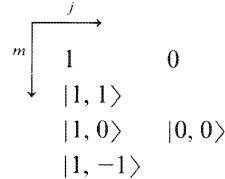
(15.2.5)

(Note that the labels $j_1 j_2$ are suppressed on the total- j kets. We shall do so frequently to simplify the notation.)

Our problem is to express each of these kets as a linear combination of product kets. To get an idea of how one goes about doing this, let us consider the problem

[‡] There is no loss of generality, for we can always call the larger one j_1 .

solved in the last section ($j_1=j_2=1/2$). In this case the states are



Consider the *top state* in the first column, $|1, 1\rangle$, which has the largest possible z component. There is only one product state with the right value of m , namely, both spins up. So by inspection,

$$|1, 1\rangle = |++\rangle$$

We can multiply the right-hand side by a phase factor, but we follow the convention, called the *Condon–Shortley convention*, in which the coefficient of this top state is chosen to be unity. Consider next the state below this one, namely, $|1, 0\rangle$. There are two product states with $m=0$, namely, $|+-\rangle$ and $|-+\rangle$; and $|1, 0\rangle$ must be a linear combination of these. We find the combination as follows. We know that[‡]

$$S_-|1, 1\rangle = 2^{1/2}\hbar|1, 0\rangle$$

so that

$$|1, 0\rangle = \frac{1}{2^{1/2}\hbar} S_-|1, 1\rangle$$

But we do not want $|1, 0\rangle$ in terms of $|1, 1\rangle$, we want it in terms of the product kets. So we rewrite the right-hand side as

$$= \frac{1}{2^{1/2}\hbar} (S_{1-} + S_{2-})|++\rangle = \frac{1}{2^{1/2}\hbar} (\hbar|+-\rangle + \hbar|-+\rangle)$$

so that

$$|1, 0\rangle = 2^{-1/2}(|+-\rangle + |-+\rangle)$$

in accordance with our earlier result.

The next state $|1, -1\rangle$ can be obtained by lowering this one more step in the above sense, or more simply by noting that there is only one ket with m maximally negative, namely, $|--\rangle$. So

$$|1, -1\rangle = |--\rangle$$

Our phase convention is such that this is what you would get if you lowered $|1, 0\rangle$.

[‡] Recall $J_z|j, m\rangle = \hbar[(j \mp m)(j \pm m + 1)]^{1/2}|j, m \pm 1\rangle$.

This takes care of the $j=1$ states. Consider next $j=0$. The state $|0, 0\rangle$ has $m=0$ and is also a linear combination of $|+-\rangle$ and $|-+\rangle$. We find the combination using two constraints: (1) The combination must be orthogonal to the one that forms the other state with $m=0$, namely, $|1, 0\rangle$ and have real coefficients.[†] (2) The combination is normalized to unity. If we call the combination $\alpha|+-\rangle + \beta|-+\rangle$, these constraints tell us that

$$\begin{aligned}\alpha + \beta &= 0 \\ \alpha^2 + \beta^2 &= 1\end{aligned}$$

It follows that

$$|0, 0\rangle = 2^{-1/2}(|+-\rangle - |-+\rangle)$$

Note that we could still have multiplied the state by (-1) . Our convention is as follows: in each column in Eq. (15.2.5) the top state is given the overall sign which makes the coefficient of the product ket with $m_1=j_1$ positive.

Let us now turn to the general problem, Eq. (15.2.5). Once again the top state in the first column, with m equal to its maximum value of $j_1 + j_2$, can be built out of only one product ket, the one in which both angular momenta take on maximum possible projections along the z axis:

$$|j_1 + j_2, j_1 + j_2\rangle = |j_1 j_1, j_2 j_2\rangle \quad (15.2.6)$$

The other m states at this value of j are obtained by lowering. Let us consider going down just one step. Since

$$J_- |j_1 + j_2, j_1 + j_2\rangle = \hbar[2(j_1 + j_2)]^{1/2} |j_1 + j_2, j_1 + j_2 - 1\rangle$$

we have, as in the spin-(1/2 \otimes 1/2) problem

$$\begin{aligned}& |j_1 + j_2, j_1 + j_2 - 1\rangle \\&= \frac{1}{[2(j_1 + j_2)]^{1/2}\hbar} \cdot (J_{1-} + J_{2-}) |j_1 j_1, j_2 j_2\rangle \\&= \frac{1}{[2(j_1 + j_2)]^{1/2}\hbar} [\hbar(2j_1)^{1/2} |j_1(j_1 - 1), j_2 j_2\rangle + \hbar(2j_2)^{1/2} |j_1 j_1, j_2(j_2 - 1)\rangle] \\&= \left(\frac{j_1}{j_1 + j_2}\right)^{1/2} |j_1(j_1 - 1), j_2 j_2\rangle + \left(\frac{j_2}{j_1 + j_2}\right)^{1/2} |j_1 j_1, j_2(j_2 - 1)\rangle \quad (15.2.7)\end{aligned}$$

Proceeding in this manner we can get to the bottom state in the first column.[§]

Now for the top state in the second column. Since it has $m=j_1 + j_2 - 1$, there are two product kets that are eligible to enter the linear combination; they are

[†] This is a matter of convention.

[§] In practice one goes only to $m=0$. The states of negative m can be found using special properties of the expansion, to be discussed shortly.

$|j_1 j_1, j_2(j_2 - 1)\rangle$ and $|j_1(j_1 - 1), j_2 j_2\rangle$. The combination must be normalized to unity, be orthogonal to the other state formed out of these kets, namely, $|j_1 + j_2, j_1 + j_2 - 1\rangle$ [see Eq. (15.2.7)], and by convention have real coefficients. The answer is, by inspection,

$$\begin{aligned} |j_1 + j_2 - 1, j_1 + j_2 - 1\rangle &= \left(\frac{j_1}{j_1 + j_2} \right)^{1/2} |j_1 j_1, j_2(j_2 - 1)\rangle \\ &\quad - \left(\frac{j_2}{j_1 + j_2} \right)^{1/2} |j_1(j_1 - 1), j_2 j_2\rangle \end{aligned} \quad (15.2.8)$$

The overall sign is fixed by requirement that the coefficient of the product ket with $m_1 = j_1$ be positive. Given the top state, the rest of the second column may be obtained by lowering. Let us go just one more column. The top state in the third column, $|j_1 + j_2 - 2, j_1 + j_2 - 2\rangle$, can be a superposition of three product kets. The three (real) coefficients are determined by these three requirements: orthogonality to the two preceding total- j kets of the same m , and unit normalization. It is clear that there are always enough constraints to determine the top states of each column, and once the top states are known, the rest follow by lowering.

Exercise 15.2.1. (1) Verify that $|j_1 j_1, j_2 j_2\rangle$ is indeed a state of $j = j_1 + j_2$ by letting $J^2 = J_1^2 + J_2^2 + 2J_{1z}J_{2z} + J_{1+}J_{2-} + J_{1-}J_{2+}$ act on it.

(2) (optional) Verify that the right-hand side of Eq. (15.2.8) indeed has angular momentum $j = j_1 + j_2 - 1$.

Clebsch–Gordan (CG) Coefficients

The completeness of the product kets allows us to write the total- j kets as

$$|jm, j_1 j_2\rangle = \sum_{m_1, m_2} \langle j_1 m_1, j_2 m_2 | \langle j_1 m_1, j_2 m_2 | jm, j_1 j_2 \rangle$$

The coefficients of the expansion

$$\langle j_1 m_1, j_2 m_2 | jm, j_1 j_2 \rangle \equiv \langle j_1 m_1, j_2 m_2 | jm \rangle$$

are called *Clebsch–Gordan coefficients* or *vector addition coefficients*. (Since the labels $j_1 j_2$ appear in the bra, we suppress them in the ket.) Here are some properties of these coefficients:

$$(1) \quad \langle j_1 m_1, j_2 m_2 | jm \rangle \neq 0 \quad \text{only if } j_1 - j_2 \leq j \leq j_1 + j_2 \quad (15.2.9)$$

(This is called the *triangle inequality*, for geometrically it means that we must be able to form a triangle with sides j_1 , j_2 , and j).

$$(2) \quad \langle j_1 m_1, j_2 m_2 | jm \rangle \neq 0 \quad \text{only if } m_1 + m_2 = m \quad (15.2.10)$$

(3) they are real (conventional)

(4) $\langle j_1 j_1, j_2 (j - j_1) | jj \rangle$ is positive (conventional)

(This condition fixes the overall sign in the expansion of each top state and was invoked in the preceding discussion.)

$$(5) \langle j_1 m_1, j_2 m_2 | jm \rangle = (-1)^{j_1 + j_2 - j} \langle j_1 (-m_1), j_2 (-m_2) | j(-m) \rangle \quad (15.2.11)$$

This relation halves the work we have to do: we start at the top state and work our way down to $m=0$ (or $1/2$ if j is half-integral). The coefficients for the negative m states are then determined by this relation.

*Exercise 15.2.2.** Find the CG coefficients of

$$(1) \frac{1}{2} \otimes 1 = \frac{3}{2} \oplus \frac{1}{2}$$

$$(2) 1 \otimes 1 = 2 \oplus 1 \oplus 0$$

Exercise 15.2.3. Argue that $\frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2} \oplus \frac{1}{2}$.

If we assemble the CG coefficients into a matrix, we find it is orthogonal (real and unitary). This follows from the fact that it relates one orthonormal basis to another. If we invert the matrix, we can write the product kets in terms of total- j kets. The coefficients in this expansion are also CG coefficients:

$$\langle jm | j_1 m_1, j_2 m_2 \rangle = \langle j_1 m_1, j_2 m_2 | jm \rangle^* = \langle j_1 m_1, j_2 m_2 | jm \rangle$$

because the CG coefficients are real. As an example, consider the $\frac{1}{2} \otimes \frac{1}{2}$ problem. There we have

$$\begin{bmatrix} |jm\rangle \\ |1, 1\rangle \\ |1, 0\rangle \\ |1, -1\rangle \\ |0, 0\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2^{1/2} & 1/2^{1/2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1/2^{1/2} & -1/2^{1/2} & 0 \end{bmatrix} \begin{bmatrix} |m_1 m_2\rangle \\ |++\rangle \\ |+-\rangle \\ |-+\rangle \\ |--\rangle \end{bmatrix}$$

(Notice that the columns contain not the components of vectors, but the basis vectors themselves.) We can invert this relation to get

$$\begin{bmatrix} |++\rangle \\ |+-\rangle \\ |-+\rangle \\ |--\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2^{1/2} & 0 & 1/2^{1/2} \\ 0 & 1/2^{1/2} & 0 & -1/2^{1/2} \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} |1, 1\rangle \\ |1, 0\rangle \\ |1, -1\rangle \\ |0, 0\rangle \end{bmatrix}$$

Thus we can write

$$|+-\rangle = 2^{-1/2}(|1, 0\rangle + |0, 0\rangle)$$

etc. In practice one uses CG coefficients to go both ways, from the product to the total- j basis and vice versa.

Addition of L and S

Consider an electron bound to a proton in a state of orbital angular momentum l . Since the electron has spin 1/2, its total angular momentum $\mathbf{J}=\mathbf{L}+\mathbf{S}$ can have values of $j=l\pm 1/2$. We wish to express the total- j states in terms of product states $|lm_o, sm_s\rangle$.[‡] Since $m_s=\pm 1/2$, at each m there will be at the most two eligible product kets.[§] Let

$$|j=l+1/2, m\rangle = \alpha |l, m-1/2; 1/2, 1/2\rangle + \beta |l, m+1/2; 1/2, -1/2\rangle \quad (15.2.12)$$

$$|j=l-1/2, m\rangle = \alpha' |l, m-1/2; 1/2, 1/2\rangle + \beta' |l, m+1/2; 1/2, -1/2\rangle \quad (15.2.13)$$

The requirement that these states be orthonormal tells us that

$$\alpha^2 + \beta^2 = 1 \quad (15.2.14)$$

$$\alpha'^2 + \beta'^2 = 1 \quad (15.2.15)$$

$$\alpha\alpha' + \beta\beta' = 0 \quad (15.2.16)$$

So we only need one more constraint, say the ratio α/β . We find it by demanding that

$$J^2 |j=l+1/2, m\rangle = \hbar^2(l+1/2)(l+3/2) |j=l+1/2, m\rangle \quad (15.2.17)$$

Writing

$$J^2 = L^2 + S^2 + 2L_zS_z + L_-S_+ + L_+S_- \quad (15.2.18)$$

we can deduce that

$$\frac{\beta}{\alpha} = \left(\frac{l+1/2-m}{l+1/2+m} \right)^{1/2} \quad (15.2.19)$$

[‡] Here, m_o , m_s , and m stand for orbital, spin, and total projections along the z axis.

[§] It might help to construct the table as in Eq. (15.2.5). It will contain just two columns, one for $j=l+1/2$ and one for $j=l-1/2$.

Given this, and our convention for the overall sign,

$$\begin{aligned} |j=l\pm 1/2, m\rangle &= \frac{1}{(2l+1)^{1/2}} [\pm(l+1/2\pm m)^{1/2}|l, m-1/2; 1/2, 1/2\rangle \\ &\quad + (l+1/2\mp m)^{1/2}|l, m+1/2; 1/2, -1/2\rangle] \end{aligned} \quad (15.2.20)$$

[Notice that if $j=l+1/2, m=\pm(l+1/2)$; only one term survives with unit coefficient.] If the Hamiltonian contains just the Coulomb interaction, or, in addition, an interaction with a weak constant magnetic field, the product basis is adequate. The total- j basis will come in handy when we study the spin-orbit interaction [which involves the operator $\mathbf{L}\cdot\mathbf{S}=\frac{1}{2}(J^2-L^2-S^2)$] in Chapter 17.

Exercise 15.2.4. Derive Eqs. (15.2.19) and (15.2.20).

*Exercise 15.2.5.** (1) Show that $\mathbb{P}_1=\frac{3}{4}I+(\mathbf{S}_1\cdot\mathbf{S}_2)/\hbar^2$ and $\mathbb{P}_0=\frac{1}{4}I-(\mathbf{S}_1\cdot\mathbf{S}_2)/\hbar^2$ are projection operators, i.e., obey $\mathbb{P}_i\mathbb{P}_j=\delta_{ij}\mathbb{P}_j$ [use Eq. (14.3.39)].
 (2) Show that these project into the spin-1 and spin-0 spaces in $\frac{1}{2}\otimes\frac{1}{2}=1\oplus 0$.

Exercise 15.2.6. Construct the project operators \mathbb{P}_{\pm} for the $j=l\pm 1/2$ subspaces in the addition $\mathbf{L}+\mathbf{S}=\mathbf{J}$.

Exercise 15.2.7. Show that when we add j_1 to j_1 , the states with $j=2j_1$ are symmetric. Show that the states with $j=2j_1=1$ are antisymmetric. (Argue for the symmetry of the top states and show that lowering does not change symmetry.) This pattern of alternating symmetry continues as j decreases, but is harder to prove.

The Modified Spectroscopic Notation

In the absence of spin, it is sufficient to use a single letter such as s, p, d, \dots to denote the (orbital) angular momentum of a particle. In the presence of spin one changes the notation as follows:

- (1) Use capital letters S, P, D, \dots (let us call a typical letter L), to indicate the value of the orbital angular momentum.
- (2) Append a *subscript J* to the right of L to indicate the j value.
- (3) Append a *superscript $2S+1$* to the left of L to indicate the multiplicity due to spin projections.

Thus, for example

$${}^{2S+1}L_J = {}^2P_{3/2}$$

denotes a state with $l=1, s=1/2, j=3/2$. For a single electron the $2S+1$ label is redundant and always equals 2. For a multielectron system, S and L stand

for total spin and total orbital angular momentum, and J for their sum. Thus in the ground state of He,

$$^{2S+1}L_J = ^1S_0$$

15.3. Irreducible Tensor Operators

We have already discussed scalar and vector operators. A scalar operator S transforms like a scalar under rotations, i.e., remains invariant:

$$S \rightarrow S' = U^\dagger[R] S U[R] = S \quad (15.3.1)$$

By considering arbitrary infinitesimal rotations we may deduce that

$$[J_i, S] = 0$$

or in a form that will be used later

$$\begin{aligned} [J_\pm, S] &= 0 \\ [J_z, S] &= 0 \end{aligned} \quad (15.3.2)$$

Examples of S are rotationally invariant Hamiltonians such as the Coulomb or isotropic oscillator Hamiltonian. A vector operator \mathbf{V} was defined as a collection of three operators (V_x , V_y , V_z) which transform as the components of a vector in $\mathbb{V}^3(R)$:

$$V_i \rightarrow V'_i = U^\dagger[R] V_i U[R] = \sum_j R_{ij} V_j \quad (15.3.3)$$

where R is the usual 3×3 rotation matrix. By considering infinitesimal rotations, we may deduce that [Eq. (12.4.14)]:

$$[V_i, J_j] = i\hbar \sum_k \epsilon_{ijk} V_k \quad (15.3.4)$$

Let us rewrite Eq. (15.3.3) in an equivalent form. Replace R by $R^{-1} = R^T$ everywhere to get

$$U[R] V_i U^\dagger[R] = \sum_j R_{ji} V_j \quad (15.3.5)$$

Notice that we are summing now over the first index of R . This seems peculiar, for we are accustomed to the likes of Eq. (15.3.3) where the sum is over the second index. The relation of Eq. (15.3.3) to Eq. (15.3.5) is the following. Let $|1\rangle$, $|2\rangle$, and $|3\rangle$ be basis kets in $\mathbb{V}^3(R)$ and R a rotation operator on it. If $|V\rangle$ is some vector

with components $v_i = \langle i | V \rangle$, its rotated version $|V'\rangle = R|V\rangle$ has components

$$v'_i = \langle i | R | V \rangle = \sum_j \langle i | R | j \rangle \langle j | V \rangle = \sum_j R_{ij} v_j \quad (15.3.6)$$

If instead we ask what R does to the basis, we find $|i\rangle \rightarrow |i'\rangle = R|i\rangle$ where

$$|i'\rangle = R|i\rangle = \sum_j |j\rangle \langle j | R | i \rangle = \sum_j R_{ji} |j\rangle \quad (15.3.7)$$

Since $R_{ji} = (R^{-1})_{ij}$, we see that vector components and the basis vectors transform in “opposite” ways. Equation (15.3.3) defines a vector operator as one whose components transform under $V_i \rightarrow U^\dagger V_i U$ as do *components* of a vector $|V\rangle$ under $|V\rangle \rightarrow R|V\rangle$, while Eq. (15.3.5) defines it as one whose components V_i transform under $V_i \rightarrow UV_i U^\dagger$ as do the *kets* $|i\rangle$ under $|i\rangle \rightarrow R|i\rangle$. Both definitions are of course equivalent. The first played a prominent role in the past and the second will play a prominent role in what follows.

Tensor Operators

We know that a vector $|V\rangle$ is an element of $\mathbb{V}^3(R)$, i.e., may be written as

$$|V\rangle = \sum_{i=1}^3 v_i |i\rangle \quad (15.3.8)$$

in terms of its components v_i and the basis kets $|i\rangle$. A *second-rank tensor* $|T^{(2)}\rangle$ is an element of the direct product space $\mathbb{V}^3(R) \otimes \mathbb{V}^3(R)$, spanned by the nine kets $|i\rangle \otimes |j\rangle$:

$$|T^{(2)}\rangle = \sum_{i=1}^3 \sum_{j=1}^3 t_{ij} |i\rangle \otimes |j\rangle \quad (15.3.9)$$

One refers to t_{ij} as the components of $|T^{(2)}\rangle$ in the basis $|i\rangle \otimes |j\rangle$.

As in the case of vectors, a tensor operator of rank 2 is a collection of nine operators T_{ij} which, under $T_{ij} \rightarrow U^\dagger T_{ij} U$, respond as do the tensor components t_{ij} , or, equivalently, under $T_{ij} \rightarrow U T_{ij} U^\dagger$, respond as do the basis kets $|i\rangle \otimes |j\rangle$. Tensors and tensor operators of rank $n > 2$ are defined in a similar way. (Note that a vector may be viewed as a tensor of rank 1.) We shall call these tensors *Cartesian tensors*.

Of greater interest to us are objects called *spherical tensor operators*. A spherical tensor operator of rank k has $2k+1$ components T_k^q , $q = +k, (k-1), \dots, -k$, which,

under $T_k^q \rightarrow UT_k^qU^\dagger$ respond like the angular momentum eigenkets $|j=k, m=q\rangle = |kq\rangle$ [‡]:

$$U[R]T_k^q U^\dagger[R] = \sum_{q'} D_{q'q}^{(k)} T_k^{q'} \quad (15.3.10)$$

Since the $2k+1$ kets $|kq\rangle$ transform irreducibly, so do the operators T_k^q . For this reason, they are also called *irreducible tensor operators*.

By considering infinitesimal rotations, we may deduce from Eq. (15.3.10) that (Exercise 15.3.1):

$$\begin{aligned} [J_\pm, T_k^q] &= \pm\hbar[(k \mp q)(k \pm q + 1)]^{1/2} T_k^{q \pm 1} \\ [J_z, T_k^q] &= \hbar q T_k^q \end{aligned} \quad (15.3.11)$$

Notice that commuting a J with T_k^q is like letting J act on the ket $|kq\rangle$.

Why are irreducible tensor operators interesting? Consider the effect of acting on a state $|\alpha lm\rangle$ with T_k^q . (Here α denotes labels besides angular momentum.) Let us rotate the resulting state and see what happens:

$$\begin{aligned} U[R]T_k^q|jm\rangle &= U[R]T_k^q U^\dagger[R]U[R]|jm\rangle \\ &= \sum_{q'} D_{q'q}^{(k)} T_k^{q'} \sum_{m'} D_{m'm}^{(j)}|jm'\rangle \\ &= \sum_{q'} \sum_{m'} D_{q'q}^{(k)} D_{m'm}^{(j)} T_k^{q'}|jm'\rangle \end{aligned} \quad (15.3.12)$$

We find that $T_k^q|jm\rangle$ responds to rotations like the product ket $|kq\rangle \otimes |jm\rangle$. Thus, when we act on a state with T_k^q , we add angular momentum (k, q) to the state. In other words, an irreducible tensor operator T_k^q imparts a definite amount of angular momentum (k, q) to the state it acts on. This allows us to say the following about matrix elements of T_k^q between angular momentum eigenstates:

$$\langle \alpha' j' m' | T_k^q | \alpha jm \rangle = 0 \quad \text{unless } k+j \geq j' \geq |k-j|, \quad m'=m+q \quad (15.3.13)$$

This is because $T_k^q|\alpha jm\rangle$ contains only those angular momenta that can be obtained by adding (k, q) and (j, m) ; so $|\alpha' j' m'\rangle$ is orthogonal to $T_k^q|jm\rangle$ unless (j', m') is one of the possible results of adding (k, q) and (j, m) . Equation (15.3.13) is an example of a *selection rule*.

Let us consider some examples, starting with the tensor operator of rank 0. It has only one component T_0^0 , which transforms like $|00\rangle$, i.e., remains invariant.

[‡] Recall that

$$\begin{aligned} |kq\rangle \rightarrow U[R]|kq\rangle &= \sum_{k'} \sum_{q'} |k'q'\rangle \langle k'q'| U[R] |kq\rangle \\ &= \sum_{q'} D_{q'q}^{(k)} |kq'\rangle \end{aligned}$$

Thus T_0^0 is just a scalar operator S , discussed earlier. Our selection rule tells us that

$$\langle \alpha' j' m' | T_0^0 | \alpha j m \rangle = 0 \quad \text{unless} \quad j=j', \quad m=m' \quad (15.3.14)$$

Consider next T_k^q ($q=1, 0, -1$). Here we have three objects that go into each other under rotations. Since a vector operator \mathbf{V} also has three components that transform irreducibly (why?) into each other, we conjecture that some linear combinations of the vector operator components should equal each T_k^q . In fact

$$\begin{aligned} T_1^{\pm 1} &= \mp \frac{V_x \pm i V_y}{2^{1/2}} \equiv V_1^{\pm 1} \\ T_1^0 &= V_z \equiv V_1^0 \end{aligned} \quad (15.3.15)^\ddagger$$

Given Eq. (15.3.4) and the above definitions, it may be readily verified that $V_1^{\pm 1}$ and V_1^0 obey Eq. (15.3.11) with $k=1$, $q=\pm 1, 0$. The selection rule for, say, V_x is

$$\begin{aligned} \langle \alpha' j' m' | V_x | \alpha j m \rangle &= \langle \alpha' j' m' | \frac{V_1^{-1} - V_1^1}{2^{1/2}} | \alpha j m \rangle \\ &= 0 \quad \text{unless} \quad j+1 \geq j' \geq |j-1|, \quad m'=m \pm 1 \end{aligned} \quad (15.3.16a)$$

and likewise

$$\begin{aligned} \langle \alpha' j' m' | V_z | \alpha j m \rangle &= \langle \alpha' j' m' | V_1^0 | \alpha j m \rangle \\ &= 0 \quad \text{unless} \quad j+1 \geq j' \geq |j-1|, \quad m'=m \end{aligned} \quad (15.3.16b)$$

Once we go beyond rank 1, it is no longer possible to express Cartesian and spherical tensors of the same rank in terms of each other. A Cartesian tensor of rank n has 3^n components, whereas a spherical tensor of rank k has $(2k+1)$ components. For $n=0$ and $n=1$, the Cartesian tensors happened to have the same number of components as spherical tensors of rank $k=0$ and 1, respectively, and also transformed irreducibly. But consider higher ranks, say rank 2. The tensor T_2^q has five components that transform irreducibly. The tensor T_{ij} has nine components which transform reducibly, i.e., it is possible to form combinations of T_{ij} such that some of them never mix with others under rotations. There is one combination that is invariant, i.e., transforms like T_0^0 ; there are three combinations that transform like a vector or in light of Eq. (15.3.15) like T_1^q ; and finally there are five that transform like T_2^q . We will see what these combinations are when we study the degeneracy of the isotropic oscillator of a few pages hence. Cartesian tensors of higher rank are likewise reducible. Let us now return to the selection rule, Eq. (15.3.13).

We can go a step further and relate the nonvanishing matrix elements. Consider the concrete example of R_1^q , the position operator in spherical form. We have

[†] In the special case $\mathbf{V} = \mathbf{J}$, $J_1^{\pm 1} = \mp(J_x \pm J_y)/2^{1/2} = \mp J_{\pm}/2^{1/2}$ and $J_1^0 = J_z$.

(assuming no spin, so $\mathbf{J}=\mathbf{L}$)

$$\begin{aligned}
\langle \alpha_2 l_2 m_2 | R_1^q | \alpha_1 l_1 m_1 \rangle &= \int R_{\alpha_2 l_2}^*(r) Y_{l_2}^{m_2}(\theta, \phi) r \left(\frac{4\pi}{3} \right)^{1/2} Y_1^q R_{\alpha_1 l_1}(r) Y_{l_1}^{m_1}(\theta, \phi) r^2 dr d\Omega \\
&= \left(\frac{4\pi}{3} \right)^{1/2} \int R_{\alpha_2 l_2}^* r R_{\alpha_1 l_1} r^2 dr \cdot \int Y_{l_2}^{m_2} Y_1^q Y_{l_1}^{m_1} d\Omega \\
&= \langle \alpha_2 l_2 | R_1 | \alpha_1 l_1 \rangle \cdot \langle l_2 m_2 | 1q, l_1 m_1 \rangle
\end{aligned} \tag{15.3.17}\dagger$$

where $\langle \alpha_2 l_2 | R_1 | \alpha_1 l_1 \rangle$, the *reduced matrix element*, is independent of m_1, m_2 , and q , which appear only in the CG coefficient, which is essentially the angular integral (up to a factor independent of m_1, m_2 , and q).

This example illustrates a general result (not proven here):

$$\langle \alpha_2 j_2 m_2 | T_k^q | \alpha_1 j_1 m_1 \rangle = \langle \alpha_2 j_2 | T_k | \alpha_1 j_1 \rangle \cdot \langle j_2 m_2 | kq, j_1 m_1 \rangle \tag{15.3.18}$$

This is called the *Wigner–Eckart theorem*. It separates the dependence of the matrix element on spatial orientation (on m_2, m_1 , and q) from the rest. The former is expressed entirely in terms of the CG coefficients.

Exercise 15.3.1. (1) Show that Eq. (15.3.11) follows from Eq. (15.3.10) when one considers infinitesimal rotations. (Hint: $D_{qq'}^{(k)} = \langle kq' | I - (i\delta\mathbf{\Theta} \cdot \mathbf{J}) / \hbar | kq \rangle$. Pick $\delta\theta$ along, say, the x direction and then generalize the result to the other directions.)

(2) Verify that the spherical tensor V_1^q constructed out of \mathbf{V} as in Eq. (15.3.15) obeys Eq. (15.3.11).

Exercise 15.3.2. It is claimed that $\sum_q (-1)^q S_k^q T_k^{(-q)}$ is a scalar operator.

(1) For $k=1$, verify that this is just $\mathbf{S} \cdot \mathbf{T}$.

(2) Prove it in general by considering its response to a rotation. [Hint: $D_{-m,-m'}^{(j)} = (-1)^{m-m'} (D_{m,m'}^{(j)})^*$.]

Exercise 15.3.3. (1) Using $\langle jj | jj, 10 \rangle = [j/(j+1)]^{1/2}$ show that

$$\langle \alpha j | J_1 | \alpha' j' \rangle = \delta_{\alpha\alpha'} \delta_{jj'} \hbar [j(j+1)]^{1/2}$$

(2) Using $\mathbf{J} \cdot \mathbf{A} = J_z A_z + \frac{1}{2} (J_- A_+ + J_+ A_-)$ (where $A_{\pm} = A_x \pm iA_y$) argue that

$$\langle \alpha' jm' | \mathbf{J} \cdot \mathbf{A} | \alpha jm \rangle = c \langle \alpha' j | A | \alpha j \rangle$$

where c is a constant independent of α, α' and \mathbf{A} . Show that $c = \hbar [j(j+1)]^{1/2} \delta_{mm'}$.

(3) Using the above, show that

$$\langle \alpha' jm' | A^q | \alpha jm \rangle = \frac{\langle \alpha' jm | \mathbf{J} \cdot \mathbf{A} | \alpha jm \rangle}{\hbar^2 j(j+1)} \langle jm' | J^q | jm \rangle \tag{15.3.19}$$

[†] Note that R_1^q is the tensor operator and $R_{\alpha l}(r)$ is the radial part of the wave function. We have also used Eq. (12.5.42) to obtain R_1^q .

*Exercise 15.3.4.** (1) Consider a system whose angular momentum consists of two parts \mathbf{J}_1 and \mathbf{J}_2 and whose magnetic moment is

$$\boldsymbol{\mu} = \gamma_1 \mathbf{J}_1 + \gamma_2 \mathbf{J}_2$$

In a state $|jm, j_1 j_2\rangle$ show, using Eq. (15.3.19), that

$$\langle \mu_x \rangle = \langle \mu_y \rangle = 0$$

$$\langle \mu_z \rangle = m\hbar \left[\frac{\gamma_1 + \gamma_2}{2} + \frac{(\gamma_1 - \gamma_2)}{2} \frac{j_1(j_1+1) - j_2(j_2+1)}{j(j+1)} \right]$$

(2) Apply this to the problem of a proton ($g = 5.6$) in a ${}^2P_{1/2}$ state and show that $\langle \mu_z \rangle = \pm 0.26$ nuclear magnetons.

(3) For an electron in a ${}^2P_{1/2}$ state show that $\langle \mu_z \rangle = \pm \frac{1}{3}$ Bohr magnetons.

*Exercise 15.3.5.** Show that $\langle jm | T_k^q | jm \rangle = 0$ if $k > 2j$.

15.4. Explanation of Some “Accidental” Degeneracies

In this section the degeneracy of states of different l at a given value of n in the hydrogen atom and the isotropic oscillator (see Section 12.6) will be explained. But first let us decide what it means to explain any degeneracy. Consider for example the $(2l+1)$ -fold degeneracy of the different m states at a given l in both these problems. We explain it in terms of the rotational invariance of the Hamiltonian as follows:

(1) For every rotation $R(\theta)$ on $\mathbb{V}^3(R)$ there exists a unitary operator $U[R]$ which rotates the vector operators

$$U^\dagger V_i U = \sum_j R_{ij} V_j \quad (15.4.1)$$

If the Hamiltonian depends only on the “lengths” of various vector operators like \mathbf{P} , \mathbf{R} , \mathbf{L} etc., then it is rotationally invariant:

$$U^\dagger H U = H \quad (15.4.2)$$

i.e., rotations are symmetries of H . This is the case for the two problems in question.

(2) If we write this relation in infinitesimal form, we find

$$[H, L_i] = 0, \quad i = 1, 2, 3 \quad (15.4.3)$$

where L_i are the generators of rotation. For every free parameter that defines a rotation (θ_x , θ_y , and θ_z) there is a corresponding generator. They are all conserved.

(3) From the three generators we construct the operator

$$L_- = L_x - iL_y \quad (15.4.4)$$

which lowers the m value:

$$L_-|l, m\rangle = c|l, m-1\rangle \quad (15.4.5)$$

Since $[L, H]=0$, the lowering operation does not change the energy.

This explains the degeneracy in m , for, starting with the state of highest m at a given l , we can go down all the way to the lowest m without changing the energy. (We can equally well work with L_+ .)

Let us try to do the same for the two problems in question. We follow these steps:

Step (1): Identify symmetries of H besides rotational invariance.

Step (2): Find the generators of the symmetry transformations.

Step (3): Construct an operator from these generators that can change l by one unit in the case of hydrogen and two units in the case of the oscillator.

Hydrogen

Steps (1) and (2). Unfortunately the only obvious symmetry of the Coulomb Hamiltonian is rotational invariance. The additional symmetry, the one we are after, is very subtle and clearest in momentum space. We will not discuss it. But how then do we go to step (2)? The answer lies in the fact that the generators of the symmetry are conserved quantities. Now we have seen that the Coulomb problem admits an extra conserved quantity, the Runge–Lenz vector. Thus the three components of

$$\mathbf{N} = \frac{1}{2m} (\mathbf{P} \times \mathbf{L} - \mathbf{L} \times \mathbf{P}) - \frac{e^2 \mathbf{R}}{(X^2 + Y^2 + Z^2)^{1/2}} \quad (15.4.6)$$

must be the generators of the additional symmetry transformations (or linear combinations thereof).

Step (3). Since we wish to talk about angular momentum let us write \mathbf{N} in spherical form:

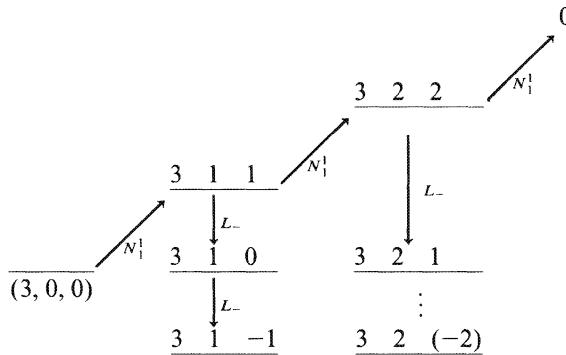
$$N_1^{\pm 1} = \mp \frac{N_x \pm iN_y}{2^{1/2}} \quad (15.4.7)$$

$$N_1^0 = N_z$$

Consider the state $|nl\ell\rangle$ of the H-atom. Acting on it with N_1^1 , we get another state of the same energy or same n (since $[H, N_1^1]=0$) but with higher angular momentum: $N_1^1|nl\ell\rangle$ behaves as $|11\rangle \otimes |l\ell\rangle = |l+1, l+1\rangle$. So

$$N_1^1|n, l, \ell\rangle = c|n, l+1, l+1\rangle \quad (15.4.8)$$

(It will turn out that c vanishes when $l=l_{\max}=n-1$.) Using N_+^l we can connect all the different l states at a given n , and using L_-^l we can connect all the m states at a given l . For example, at $n=3$ the network that connects degenerate states is as follows:



The Oscillator

Step (1). To find the extra symmetry of H , let us look at it again:

$$H = \frac{P_x^2 + P_y^2 + P_z^2}{2\mu} + \frac{1}{2}\mu\omega^2(X^2 + Y^2 + Z^2) \quad (15.4.9)$$

We say H is rotationally invariant because it depends only on the lengths squared of the (real) vectors \mathbf{P} and \mathbf{R} . Let us now rewrite H in a way that reveals the extra symmetry. Define a *complex vector* (operator) whose real and imaginary parts are proportional to \mathbf{R} and \mathbf{P} :

$$\mathbf{a} = \frac{1}{(2\mu\omega\hbar)^{1/2}} (\mu\omega\mathbf{R} + i\mathbf{P}) \quad (15.4.10)$$

and its adjoint, whose components are complex conjugates of those of \mathbf{a} :

$$\mathbf{a}^\dagger = \frac{1}{(2\mu\omega\hbar)^{1/2}} (\mu\omega\mathbf{R} - i\mathbf{P}) \quad (15.4.11)$$

The components of \mathbf{a} and \mathbf{a}^\dagger are just the lowering and raising operators for the x , y , and z oscillators. They obey

$$[a_i, a_j^\dagger] = \delta_{ij}$$

In terms of \mathbf{a} and \mathbf{a}^\dagger ,

$$H = \hbar\omega(\mathbf{a}^\dagger \cdot \mathbf{a} + 3/2) \quad (15.4.12)$$

Thus we find that H is a function of the length squared of a *complex* three-dimensional vector \mathbf{a} . So it is invariant under “rotations” in $\mathbb{V}^3(C)$, i.e., under unitary transformations in $\mathbb{V}^3(C)$. Just as we denoted the rotations in $\mathbb{V}^3(R)$ by R , let us call these C .[‡] For every “rotation” C (unitary transformation) in $\mathbb{V}^3(C)$, there will exist Hilbert space operators $U[C]$ which rotate the complex vector operator \mathbf{a} :

$$a_i \rightarrow a'_i = U^\dagger[C]a_i U[C] = \sum_j C_{ij}a_j \quad (15.4.13)$$

where C_{ij} are matrix elements of the unitary operator C in $\mathbb{V}^3(C)$. Since H depends only on the norm squared of \mathbf{a} ,

$$U^\dagger[C]HU[C] = H \quad (15.4.14)$$

Step (2). How many generators of $U[C]$ are there and what are they? The answer to the first part is the number of parameters that define a rotation in $\mathbb{V}^3(C)$, i.e., the number of independent parameters in a 3×3 unitary matrix C . Now any such matrix can be written as

$$C = e^{i\Omega} \quad (15.4.15)$$

where Ω is a 3×3 Hermitian matrix. It is easy to see that Ω has three real diagonal elements and three independent complex off-diagonal elements. Thus it depends on nine real parameters. So there are nine conserved generators. What are they? Rather than deduce them (as we did the L 's by considering the effect of infinitesimal rotations on ψ) we write down the nine conserved quantities by inspection. It is clear that in the oscillator case, the nine operators

$$T_{ij} = a_i^\dagger a_j \quad (i, j = x, y, \text{ or } z) \quad (15.4.16)$$

are conserved. The proof is simple: a_j destroys a j quantum and a_i^\dagger creates an i quantum and this leaves the energy invariant since the x , y , and z oscillators have the same ω (isotropy). To see what impact T_{ij} has on l degeneracy, we must decompose T_{ij} into its irreducible parts.

Consider first the combination

$$\text{Tr } T = T_{xx} + T_{yy} + T_{zz} = a_x^\dagger a_x + a_y^\dagger a_y + a_z^\dagger a_z = \mathbf{a}^\dagger \cdot \mathbf{a} \quad (15.4.17)$$

This is clearly a scalar, i.e., transforms like T_0^0 . The fact that it commutes with H does not explain the degeneracy in l because it “carries” no angular momentum. In fact $\mathbf{a}^\dagger \cdot \mathbf{a}$ is just H up to a scale factor and an additive constant.

[‡] We should really be calling these U . But that will complicate the notation.

Consider next the three antisymmetric combinations

$$\begin{aligned} T_{xy} - T_{yx} &= a_x^\dagger a_y - a_y^\dagger a_x = (\mathbf{a}^\dagger \times \mathbf{a})_z \\ T_{yz} - T_{zy} &= (\mathbf{a}^\dagger \times \mathbf{a})_x \\ T_{zx} - T_{xz} &= (\mathbf{a}^\dagger \times \mathbf{a})_y \end{aligned} \quad (15.4.18)$$

These clearly transform as a vector $\mathbf{V} = \mathbf{a}^\dagger \times \mathbf{a}$. There seems to be a problem here. Suppose we form the operator $V_1^l = -(V_x + iV_y)/2^{1/2}$. Then we expect

$$V_1^l |nll\rangle = c |n, l+1, l+1\rangle \quad (15.4.19)$$

as in Eq. (15.4.8). This would mean that states differing by *one* unit in l are degenerate. But we know from Section 12.6 that states differing by *two* units in l are degenerate. So how do we get out of the fix? To find out, you must work out any one of the components of the operator $\mathbf{V} = \mathbf{a}^\dagger \times \mathbf{a}$ in terms of \mathbf{R} and \mathbf{P} . If you do, you will see that c in Eq. (15.4.19) is really zero, and the paradox will be resolved.

We are now left with $9 - 1 - 3 = 5$ degrees of freedom out of the original nine T_{ij} 's. We argue that these must transform *irreducibly*. Why? Suppose the contrary is true. Then it must be possible to form irreducible tensors with fewer than five components out of these residual degrees of freedom. The only possibilities are tensors with 1 or 3 components, that is to say, scalars or vectors. But we know that given two vectors \mathbf{a}^\dagger and \mathbf{a} we can form only one scalar, $\mathbf{a}^\dagger \cdot \mathbf{a}$ and only one vector $\mathbf{a}^\dagger \times \mathbf{a}$, both of which we have already used up. So we are driven to the conclusion that the five residual degrees of freedom are linear combinations of some T_2^q . One usually refers to this object as the *quadrupole tensor* Q_2^q . All we need here is the component Q_2^2 , since

$$Q_2^2 |nll\rangle = c |n, l+2, l+2\rangle \quad (15.4.20)$$

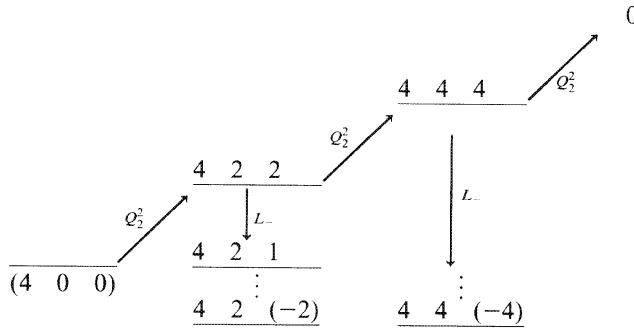
which explains the degeneracy in l at each n . (When $l=n=l_{\max}$, c vanishes.)

Let us explicitly construct the operator Q_2^2 in terms of $a_i^\dagger a_i$ to gain some experience. Now \mathbf{a} and \mathbf{a}^\dagger are vector operators from which we can form the tensor operators a_i^q and $(a^\dagger)^q$ which behave like $|1, q\rangle$. The product $a_i^\dagger a_i$ then behaves like the direct product of (linear combinations) of two spin-1 objects. Since Q_2^2 behaves like $|22\rangle$ and since $|22\rangle = |11\rangle \otimes |11\rangle$, we deduce that

$$\begin{aligned} Q_2^2 &= (a^\dagger)_1^q (a_1^q) \\ &= \left(\frac{a_x^\dagger + ia_y^\dagger}{2^{1/2}} \right) \left(\frac{a_x + ia_y}{2^{1/2}} \right) \\ &= \frac{1}{2} [a_x^\dagger a_x - a_y^\dagger a_y + i(a_x^\dagger a_y + a_y^\dagger a_x)] \end{aligned} \quad (15.4.21)$$

Other components of Q_2^q may be constructed by similar techniques. (It is just a matter of adding angular momenta $1 \otimes 1$ to get 2.) Starting with the smallest value

of l at each n (namely, 0 or 1), we can move up in steps of 2 until we reach $l=n$, at which point c in Eq. (15.4.20) will vanish. The network for $n=4$ is shown below:



This completes the explanation of the degeneracy of the oscillator.

The Free-Particle Solutions

We examine the free-particle solutions from Section 12.6 in the light of the preceding discussion. Here again we have a case where states with different l , in fact an infinite number of them, are degenerate at each energy $E = \hbar^2 k^2 / 2\mu$. This degeneracy is, however, not “accidental,” since the extra symmetry of the free-particle Hamiltonian, namely, translational invariance, is obvious. We therefore have a conserved vector operator \mathbf{P} from which we can form P_+ ,^f which can raise l and m by one unit. Thus, given the state with $l=m=0$, we can move up in l using

$$|kl\rangle = c(P_+)^l |k00\rangle \quad (15.4.22)$$

where c is some normalization constant.

Recall that in the coordinate basis it was easy to find

$$|k00\rangle \rightarrow \psi_{k00} = \frac{U_0(\rho)}{\rho} Y_0^0 \quad (15.4.23)$$

where $\rho = kr$, and $U_0(\rho)$ is $\sin \rho$ or $-\cos \rho$ (regular or irregular solutions). It is easy to verify that

$$\begin{aligned} P_+ |k00\rangle &\xrightarrow[\text{coordinate basis}]{} -i\hbar(x+iy) \frac{1}{r} \cdot \frac{d}{dr} \left[\frac{U_0(\rho)}{\rho} \right] Y_0^0 \\ &= C_1(x+iy) \frac{1}{\rho} \frac{d}{d\rho} \left[\frac{U_0(\rho)}{\rho} \right] \end{aligned} \quad (15.4.24)$$

^f $P_+ = P_x + iP_y$ is, up to a scale factor ($-2^{1/2}$) which does not change its rotational properties, just P_+^1 .

where C_1 has absorbed all the factors that have no ρ dependence. If we operate once again with P_+ and use $[P_+, R_+] = 0$ (where $R_+ = R_x + iR_y \propto R_1^1$), we get

$$(P_+)^2 |k00\rangle \rightarrow C_2(x + iy)^2 \left(\frac{1}{\rho} \frac{\partial}{\partial \rho} \right)^2 \frac{U_0(\rho)}{\rho} \quad (15.4.25)$$

and so finally

$$\begin{aligned} (P_+)^l |k00\rangle &\rightarrow \psi_{kl} = C_l (x + iy)^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \frac{U_0(\rho)}{\rho} \\ &= \tilde{C}_l (\sin \theta)^l e^{il\phi} \rho^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \frac{U_0(\rho)}{\rho} \\ &= \tilde{C}_l Y_l^l \rho^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \frac{U_0(\rho)}{\rho} \\ &= R_l Y_l^l \end{aligned} \quad (15.4.26)$$

where

$$R_l = \tilde{C}_l \rho^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \frac{U_0(\rho)}{\rho} = \tilde{C}_l \rho^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l R_0(\rho) \quad (15.4.27)$$

This agrees with Eq. (12.6.29) if we set $\tilde{C}_l = (-1)^l$.

16

The Variational and WKB Methods

16.1. The Variational Method

More often than not, it is impossible to find exact solutions to the eigenvalue problem of the Hamiltonian. One then turns to approximation methods, some of which will be described in this and the following chapters. In this section we consider a few examples that illustrate the *variational method*.

Our starting point is the inequality

$$E[\psi] \equiv \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \geq E_0 \quad (16.1.1)$$

where E_0 is the lowest eigenvalue of H , i.e., the ground-state energy. Although this result was proved earlier, let us recall the idea behind it. $E[\psi]$ is just the mean value of the energy in the state $|\psi\rangle$. The inequality states that the mean value cannot be less than the lowest value that enters the average. More formally, if $|\psi\rangle$ is expanded in terms of the eigenfunctions $|E_n\rangle$ of H ,

$$E[\psi] = \frac{\sum E_n |\langle E_n | \psi \rangle|^2}{\sum |\langle E_n | \psi \rangle|^2} \geq \frac{E_0 \sum |\langle E_n | \psi \rangle|^2}{\sum |\langle E_n | \psi \rangle|^2} = E_0 \quad (16.1.2)$$

This inequality suggests a way (at least in principle) of determining the ground-state energy and eigenket. We take all the kets in the Hilbert space one by one and make a table of the corresponding $E[\psi]$. At the end we read off the lowest entry and the ket that goes with it. Clearly this is not a practical algorithm. What one does in practice is to consider just a subset (not necessarily a subspace) of vectors which are parametrized by some variables ($\alpha, \beta, \gamma, \dots$) and which have the general features one expects of the true ground-state ket. In this limited search $E[\psi]$ reduces to a function of the parameters, $E(\alpha, \beta, \dots)$. We then find the values $(\alpha_0, \beta_0, \dots)$ which minimize E . This minimum $E(\alpha_0, \beta_0, \dots)$ provides an upper bound on E_0 .

The name of the game is finding the lowest upper bound for a given amount of work. If H happens to be positive definite, $E_0 \geq 0$, and we will be able to restrict E_0 to the range $E(\alpha_0, \beta_0, \dots) \geq E_0 \geq 0$.

As an example, consider the problem of a particle in a potential $V(x) = \lambda x^4$. Here are the features we expect of the ground state. It will have definite parity, and, since the ground-state function will have no nodes (more nodes \rightarrow more wiggles \rightarrow more kinetic energy), it will have even parity. It will be peaked at $x=0$ so as to minimize $\langle V \rangle$. And of course it will vanish as $|x| \rightarrow \infty$. A trial function that has all these features (and is also easy to differentiate and integrate) is

$$\psi(x, \alpha) = e^{-\alpha x^2/2} \quad (16.1.3)$$

where α is a free parameter that determines the width of the Gaussian. The energy as a function of α is

$$E(\alpha) = \int e^{-\alpha x^2/2} \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \lambda x^4 \right) e^{-\alpha x^2/2} dx / \int e^{-\alpha x^2} dx = \frac{\hbar^2 \alpha}{4m} + \frac{3\lambda}{4\alpha^2}$$

We see here the familiar struggle between the kinetic and potential energy terms. The former would like to see $\alpha \rightarrow 0$ so that the wave function is wide and has only large wavelength (small momentum) components, while the latter would like to see $\alpha \rightarrow \infty$, so that the wave function is a narrow spike near $x=0$, where the potential is a minimum. The optimal point, both effects included, is

$$\alpha_0 = \left(\frac{6m\lambda}{\hbar^2} \right)^{1/3} \quad (16.1.4)$$

The corresponding energy is

$$E(\alpha_0) = \frac{3}{8} \left(\frac{6\hbar^4 \lambda}{m^2} \right)^{1/3} \quad (16.1.5)$$

Since H is positive definite, we conclude

$$0 \leq E_0 \leq E(\alpha_0) \quad (16.1.6)$$

The best approximation to the wave function of the ground state (among all Gaussians) is $\psi(x, \alpha_0) = \exp(-\frac{1}{2}\alpha_0 x^2)$.

The inequality (16.1.6) is of course rigorous, but its utility depends on how close $E(\alpha_0)$ is to E_0 . Our calculation does not tell us this. All we know is that since we paid attention to parity, nodes, etc., our upper bound $E(\alpha_0)$ is lower than that obtained by someone whose test functions had odd parity and 15 nodes. For instance, if $V(x)$ had been $\frac{1}{2}m\omega^2 x^2$ instead of λx^4 , we would have found $\alpha_0 = (m\omega/\hbar)^{1/2}$ and $E(\alpha_0) = \hbar\omega/2$. Although this is the exact answer, our calculation would not tell us this. The way to estimate the quality of the bound obtained is to try to lower it further by considering a trial function with more parameters. If this produces sub-

stantial lowering, we keep going. On the other hand, if we begin to feel a “resistance” to the lowering of the bound as we try more elaborate test functions, we may suspect that E_0 is not too far below. In the case of $V(x) = \frac{1}{2}m\omega^2x^2$, it will eventually be found that there is no way of going below $E(\alpha_0) = \hbar\omega/2$.

Our faith in the variational method stems from its performance in cases where the exact answer is known either analytically or experimentally. Let us consider two examples. The first is that of the electron in a Coulomb potential $V = -e^2/r$. We expect the ground-state wave function to have no angular momentum, no nodes, behave like r^0 as $r \rightarrow 0$, and vanish as $r \rightarrow \infty$. So we choose $\psi(r, \theta, \phi, \alpha) = \exp(-ar^2)$.‡ We find (upon ignoring the irrelevant angular variables throughout)

$$E(\alpha) = \int \left[e^{-ar^2} \left(-\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} - \frac{e^2}{r} \right) e^{-ar^2} \right] r^2 dr / \int e^{-2ar^2} r^2 dr \quad (16.1.7)$$

$$= \frac{3\hbar^2\alpha}{2m} - \left(\frac{2}{\pi} \right)^{1/2} 2e^2 \alpha^{1/2} \quad (16.1.8)$$

which is minimized by

$$\alpha_0 = \left(\frac{me^2}{\hbar^2} \right)^2 \cdot \frac{8}{9\pi} \quad (16.1.9)$$

The upper bound is then

$$E(\alpha_0) = -\frac{me^4}{2\hbar^2} \frac{8}{3\pi} = -0.85 \text{ Ry} \quad (16.1.10)$$

which is slightly above§ the true energy. The true wave function is of course not a Gaussian, but the general features are the same. For example $\psi(r, \alpha_0) = e^{-\alpha_0 r^2}$ predicts an uncertainty $\Delta X = (9\pi/32)^{1/2}a_0 = 0.94a_0$, while the exact result is $\Delta X = a_0$ (the Bohr radius).||

The second example deals with the ground state of He. Ignoring nuclear motion ($m/M \rightarrow 0$), the Hamiltonian in the coordinate basis is

$$H \rightarrow -\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{r_{12}} \quad (16.1.11)$$

where r_1 and r_2 are the radial coordinates of the two electrons and r_{12} is the radial separation between them. We have already seen that if the mutual repulsion (e^2/r_{12}) is ignored, the ground-state wave function is just

$$\psi = \psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2) \quad (16.1.12)$$

‡ We could also choose e^{-ar} , which would give the exact answer. But let us not.

§ Remember that we are dealing with negative energies here.

|| This agreement is rather fortuitous. In general, the variational method provides much better approximations to the energies than to the wave functions. The reason follows shortly.

where the singlet spin wave function is suppressed, and ψ_{100} is the hydrogen-like wave function with $e^2 \rightarrow Ze^2$:

$$\psi_{100} = \left(\frac{Z^3}{\pi a_0^3} \right)^{1/2} e^{-Zr/a_0} \quad (Z=2) \quad (16.1.13)$$

Consequently

$$\psi = \frac{Z^3}{\pi a_0^3} e^{-Z(r_1+r_2)/a_0} \quad (Z=2) \quad (16.1.14)$$

The energy that goes with this simplified picture is

$$E = 2 \left(-\frac{m(2e^2)^2}{2\hbar^2} \right) = -8 \text{ Ry} \simeq -108.8 \text{ eV}$$

which is far below the measured value of -78.6 eV .‡ So we find that omitting the Coulomb repulsion between the electrons is a bad approximation. But if we include the e^2/r_{12} term, the problem cannot be solved analytically. So we apply to it the variational method for a trial wave function, we use just the product function in Eq. (16.1.14) but treat Z as a variable rather than setting it equal to 2. The idea is that as each electron shields the nuclear charge seen by the other, the effective Z is less than 2. This is borne out by the calculation of

$$E(Z) = \frac{\left[\int \psi(r_1 r_2 Z) \left[-\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) - 2e^2 \left(\frac{1}{r_1} + \frac{1}{r_2} \right) + \frac{e^2}{r_{12}} \right] \times \psi(r_1 r_2 Z) d^3 r_1 d^3 r_2 \right]}{\int |\psi(r_1 r_2 Z)|^2 d^3 r_1 d^3 r_2} \\ = -2 \text{ Ry} [4Z - Z^2 - \frac{5}{8}Z] \quad (16.1.15)$$

whose minimum lies at $Z=2$ but at $Z=2-5/16$. The corresponding energy is

$$E(2-5/16) = -2(2-5/16)^2 \text{ Ry} \simeq -77.5 \text{ eV} \quad (16.1.16)$$

which is much closer to the real answer. Notice also that it lies above it, as demanded by the inequality (16.1.1). By considering trial functions with more parameters, one can get closer to the exact answer, and one can also feel the “resistance” to further lowering.

‡ This is not in contradiction with Eq. (16.1.1) since we are using the wrong Hamiltonian when we neglect the Coulomb repulsion between the electrons.

A virtue of the variational method is that even a poor approximation to the actual wave function can yield an excellent approximation to the actual energy. The reason is the following. Suppose we had chosen a trial function

$$|\psi\rangle = |E_0\rangle + \frac{1}{10}|E_1\rangle$$

which contains a 10% contamination from the state $|E_1\rangle$. The estimated energy would have been

$$\begin{aligned} E(\psi) &= \frac{\langle E_0 | H | E_0 \rangle + \frac{1}{100} \langle E_1 | H | E_1 \rangle}{1 + \frac{1}{100}} = \frac{E_0 + 0.01 E_1}{1.01} \\ &\simeq 0.99 E_0 + 0.01 E_1 \end{aligned}$$

which is off by just 1%. (We are assuming that E_1 is not anomalously large.)

More generally, let

$$|\psi\rangle = |E_0\rangle + |\delta\psi\rangle \quad (16.1.17a)$$

be a trial ket. Let us decompose $|\delta\psi\rangle$ into parts parallel to and perpendicular to $|E_0\rangle$:

$$\begin{aligned} |\delta\psi\rangle &= |\delta\psi_{\parallel}\rangle + |\delta\psi_{\perp}\rangle \\ &= \alpha|E_0\rangle + |\delta\psi_{\perp}\rangle \end{aligned} \quad (16.1.17b)$$

In this state

$$\begin{aligned} E[\psi] &= \frac{E_0|1+\alpha|^2 + \langle \delta\psi_{\perp} | H | \delta\psi_{\perp} \rangle}{|1+\alpha|^2 + \langle \delta\psi_{\perp} | \delta\psi_{\perp} \rangle} \\ &= E_0 + O(\delta\psi_{\perp})^2 \end{aligned} \quad (16.1.18)$$

Thus the error in energy is of the second order in the error in the state vector. Notice that $|\delta\psi_{\parallel}\rangle$ produces no error in energy. This is because rescaling the normalized eigenket does not change the mean energy.

All these results are true for any eigenket of H . If

$$|\psi_n\rangle = |E_n\rangle + |\delta\psi_n\rangle$$

is an approximation to $|E_n\rangle$, then by similar reasoning

$$E[\psi_n] = E_n + O[(\delta\psi_n)^2] \quad (16.1.19)$$

Thus the eigenkets of H are characterized by the fact that when they are changed to first order, there is no energy change to first order: *the eigenkets of H are stationary points of $E[\psi]$.* (The ground state happens, in addition, to be an absolute minimum.) If we could carry out the impossible task of tabulating all the $E[\psi]$ we can then read off *all* the eigenstates by looking for the stationary points. This is of course not

a practical proposition. In practice, we use the following trick for finding the higher eigenvalues and eigenkets. Consider the case $V = \lambda x^4$. Since H is parity invariant, the states will occur with alternating parity. Suppose we take a trial state with odd parity. Then in the expansion $|\psi\rangle = \sum C_n |n\rangle$, $C_n = \langle n| \psi \rangle = 0$ for all even n , because the integral of an even and an odd function is zero. Consequently the lowest energy that enters the averaging is E_1 and we have the inequality

$$E[\psi] \geq E_1 \quad (16.1.20)$$

So we expect that if we take a trial state with odd parity, one node (in one dimension, there is one extra node for each upward step in energy), and the usual behavior as $|x| \rightarrow \infty$, we can get a good estimate E_1 and a rough picture of the corresponding wave function. What if we want to get a bound on E_2 ? The general idea is of course the same, to consider trial states in whose expansion $|E_0\rangle$ and $|E_1\rangle$ do not appear. But this cannot be done simply by choosing trial states of some definite parity. What we can do is the following. We have approximate wave functions for the first two levels from the variational energy estimates. We can choose our trial states to be orthogonal to these. The corresponding bounds are not rigorous, for we do not know $|E_0\rangle$ and $|E_1\rangle$ exactly, but they may still be useful.

This general idea is easier to implement in three dimensions if H is rotationally invariant. In this case the energy eigenstates have definite angular momentum. The ground state will have $l=0$. By varying spherically symmetric trial functions we can estimate the ground-state energy. If we next choose $l=1$ trial functions [$\psi = R(r) Y_1^m$], $E[\psi]$ will obey

$$E[\psi] \geq E_{l=1}$$

where $E_{l=1}$ is the lowest energy level with $l=1$. We can clearly keep going up in l . Suppose we do this for the Coulomb problem. We know that at each l , the lowest energy corresponds to $n=l+1$. The variational method applied to $l=0, 1, 2, \dots$ will yield energies close to those of the $n=1, 2, \dots$ levels. Of course we must pay attention to the radial part of ψ as well. For instance $R(r)$ must behave like r^l as $r \rightarrow 0$ in the angular momentum l sector. It must have the least number of nodes, namely zero, if it is to have the lowest energy for the given l . With these features built in, both the energy and wave function will come close to $\psi_{n,n-1,m}$.

We can also use any other operator that commutes with H in choosing trial functions. The angular momentum is especially convenient because its eigenfunctions are easy to write down and its eigenvalues are correlated (grow) with energy.

*Exercise 16.1.1.** Try $\psi = \exp(-\alpha x^2)$ for $V = \frac{1}{2} m\omega^2 x^2$ and find α_0 and $E(\alpha_0)$.

*Exercise 16.1.2.** For a particle in a box that extends from $-a$ to a , try (within the box) $\psi = (x-a)(x+a)$ and calculate E . There is no parameter to vary, but you still get an upper bound. Compare it to the true energy, E_0 . (Convince yourself that the singularities in ψ'' at $x = \pm a$ do not contribute to the energy.)

*Exercise 16.1.3.** For the attractive delta function potential $V = -aV_0\delta(x)$ use a Gaussian trial function. Calculate the upper bound on E_0 and compare it to the exact answer ($-ma^2 V_0^2 / 2\hbar^2$) (Exercise 5.2.3).

Exercise 16.1.4 (Optional). For the oscillator choose

$$\begin{aligned}\psi &= (x-a)^2(x+a)^2, & |x| \leq a \\ &= 0, & |x| > a\end{aligned}$$

Calculate $E(a)$, minimize it and compare to $\hbar\omega/2$.

*Exercise 16.1.5.** Solve the variational problem for the $l=1$ states of the electron in a potential $V = -e^2/r$. In your trial function incorporate (i) correct behavior as $r \rightarrow 0$, appropriate to $l=1$, (ii) correct number of nodes to minimize energy, (iii) correct behavior of wave function as $r \rightarrow \infty$ in a Coulomb potential (i.e., exponential instead of Gaussian damping). Does it matter what m you choose for Y_l^m ? Comment on the relation of the energy bound you obtain to the exact answer.

16.2. The Wentzel-Kramers-Brillouin Method

Consider a particle of energy E in one dimension moving in a constant potential V . The energy eigenfunctions are

$$\psi(x) = \psi(0) e^{\pm ipx/\hbar}, \quad p = [2m(E-V)]^{1/2} \quad (16.2.1)$$

where the \pm signs correspond to right- and left-moving plane waves. The general solution is a combination of both waves. The real and imaginary parts of ψ oscillate in space with a wavelength $\lambda = 2\pi\hbar/p$ or equivalently, the phase change per unit length is a constant, p/\hbar . Suppose now that V , instead of being a constant, varies very slowly. We then expect that over a small region [small compared to the distance over which $V(x)$ varies appreciably] ψ will still behave like a plane wave, with the local value of the wavelength:

$$\lambda(x) = \frac{2\pi\hbar}{p(x)} = \frac{2\pi\hbar}{\{2m[E-V(x)]\}^{1/2}} \quad (16.2.2)$$

Since λ varies with x , the accumulated phase shift between $x=0$ and $x=x$ is given by an integral, so that

$$\psi(x) = \psi(0) \exp\left[\pm(i/\hbar) \int_0^x p(x') dx'\right]$$

or more generally

$$\psi(x) = \psi(x_0) \exp\left[\pm(i/\hbar) \int_{x_0}^x p(x') dx'\right] \quad (16.2.3)$$

Once again the \pm stand for right- and left-moving waves, and the general solution is formed by taking an arbitrary linear combination of both. As mentioned above, we trust this formula only if the wavelength varies slowly. How slow is slow enough? Note that although there is a well defined function $\lambda(x)$ at each x , it makes no sense

to speak of the wavelength *at* a point. The wavelength is a characteristic of a repetitive phenomenon and consequently defined only over a region that contains many repetitions. Thus the statement “a position-dependent wavelength $\lambda(x)$ ” makes sense only if $\delta\lambda$ over a length λ is negligible compared to λ :

$$\left| \frac{\delta\lambda}{\lambda} \right| = \left| \frac{(d\lambda/dx) \cdot \lambda}{\lambda} \right| = \left| \frac{d\lambda}{dx} \right| \ll 1 \quad (16.2.4)$$

Let us now derive all of the above results more formally. The derivation will also provide corrections to this result and clarify the nature of the approximation. Our problem is to solve the equation

$$\left\{ \frac{d^2}{dx^2} + \frac{2m}{\hbar^2} [E - V(x)] \right\} \psi(x) = 0$$

or

$$\left[\frac{d^2}{dx^2} + \frac{1}{\hbar^2} p^2(x) \right] \psi(x) = 0$$

Let us write

$$\psi(x) = \exp[i\phi(x)/\hbar] \quad (16.2.5)$$

Since $\phi(x)$ is not assumed to be real, there is no loss of generality.[‡] Feeding this form into the equation, we get

$$-\left(\frac{\phi'}{\hbar}\right)^2 + \frac{i\phi''}{\hbar} + \frac{p^2(x)}{\hbar^2} = 0 \quad (16.2.6)$$

We now expand ϕ in a power series in \hbar :

$$\phi = \phi_0 + \hbar\phi_1 + \hbar^2\phi_2 + \dots \quad (16.2.7)$$

The logic is the following. If $\hbar \rightarrow 0$, the wavelength $\lambda = 2\pi\hbar/p$ tends to zero. Consequently any potential can be considered slowly varying in this limit and our approximation Eq. (16.2.3) should become increasingly reliable. Conversely, any corrections to this formula can be traced to the fact that \hbar isn't really zero. In situations where \hbar may be treated as a small number, we hope that corrections may be computed in powers of \hbar .

The WKB approximation (also called the *semiclassical approximation*) consists of keeping just the first two terms in Eq. (16.2.7). If we feed the truncated expansion

[‡] In other words, any complex number $\psi = \rho e^{i\tilde{\phi}} = e^{i\tilde{\phi} + i\ln\rho} = e^{i\phi}$, where $\phi = \tilde{\phi} - i\ln\rho$.

into Eq. (16.2.6) and group terms with the same \hbar dependence, we get

$$\frac{-(\phi'_0)^2 + p^2(x)}{\hbar^2} + \frac{i\phi''_0 - 2\phi'_1\phi'_0}{\hbar} + O(\hbar^0) = 0 \quad (16.2.8)$$

In the first approximation we concentrate on just the \hbar^{-2} term. This gives

$$\phi'_0 = \pm p(x)$$

or

$$\phi_0(x) = \pm \int_{-\infty}^x p(x') dx' \quad (16.2.9)$$

and

$$\begin{aligned} \psi(x) &= A \exp \left[\pm (i/\hbar) \int_{-\infty}^x p(x') dx' \right] \\ &= \psi(x_0) \exp \left[\pm (i/\hbar) \int_{x_0}^x p(x') dx' \right] \end{aligned} \quad (16.2.10)$$

where A was found by setting $x = x_0$ in the first equation. All this agrees with our previous result. But we can go a step further and include the \hbar^{-1} term in Eq. (16.2.8). We still choose ϕ'_0 so that the \hbar^{-2} term continues to vanish. For the \hbar^{-1} term to vanish, we need

$$\begin{aligned} i\phi''_0 &= 2\phi'_1\phi'_0 \\ \frac{\phi''_0}{\phi'_0} &= -2i\phi'_1 \\ \ln \phi'_0 &= -2i\phi_1 + c \\ \phi_1 &= +i \ln(\phi'_0)^{1/2} + c/2i = i \ln p^{1/2} + \tilde{c} \end{aligned} \quad (16.2.11)$$

To this order in \hbar ,

$$\begin{aligned} \psi(x) &= e^{i\phi(x)/\hbar} = A e^{-\ln[p(x)]^{1/2}} \exp \left[\pm \left(\frac{i}{\hbar} \right) \int_{-\infty}^x p(x') dx' \right] \\ &= \frac{A}{[p(x)]^{1/2}} \exp \left[\pm \left(\frac{i}{\hbar} \right) \int_{-\infty}^x p(x') dx' \right] \end{aligned} \quad (16.2.12)$$

or

$$\psi(x) = \psi(x_0) \left[\frac{p(x_0)}{p(x)} \right]^{1/2} \exp \left[\pm \frac{i}{\hbar} \int_{x_0}^x p(x') dx' \right] \quad (16.2.13)$$

The probability density associated with $\psi(x)$ behaves as $[p(x)]^{-1}$. This inverse dependence on the classical velocity is familiar to us from the classical probability distribution $P_{\text{cl}}(x)$ we studied in connection with the oscillator. Conversely, we could have written down Eq. (16.2.13) by combining classical and quantum reasoning, i.e., by semiclassical reasoning. We know classically that if a particle starts at x_0 with momentum $p(x_0) = \{2m[E - V(x_0)]\}^{1/2}$, it will have a momentum $p(x) = \{2m[E - V(x)]\}^{1/2}$ when it gets to x . Now we argue that since p/\hbar is the phase change per unit length of the quantum wave function, its phase must be $(1/\hbar) \int p(x') dx'$. As for its amplitude, we argue that since the probability $P_{\text{cl}}(x) \sim 1/v(x)$, $|\psi| \simeq 1/[v(x)]^{1/2} \simeq 1/[p(x)]^{1/2}$.

Whenever we do an approximate calculation, it behooves us to verify at the end that the solution we obtain is consistent with the assumptions that went into its derivation. Our fundamental assumption in the recursive approach to Eq. (16.2.8) has been that the part of the equation with less powers of \hbar is more important than the part with more powers, because \hbar is so small. This is fine as long as the coefficients of the various powers of \hbar are not anomalously big or small. For instance, if the coefficient of the \hbar^{-2} term $[-(\phi'_0)^2 + p^2(x)]$ is very small, of the order of, say, \hbar , then it makes no sense to ignore the \hbar^{-1} term in comparison. The same is true if the coefficient of the \hbar^{-1} is as large as \hbar^{-1} . So we demand that the absolute magnitude of the first term be much bigger than that of the second. Since in the solution, $\phi'_0 = p(x)$, we choose $(\phi'_0/\hbar)^2$ as a measure of the first term and for similar reasons ϕ''_0/\hbar as a measure of the second. The condition for the validity of the WKB approximation (to this order) is

$$\left| \frac{\phi''_0}{\hbar} \right| \ll \left| \frac{\phi'_0}{\hbar} \right|^2 \quad (16.2.14)$$

or

$$\hbar \left| \frac{d}{dx} \left(\frac{1}{\phi'_0} \right) \right| = \left| \frac{d}{dx} \left(\frac{\hbar}{p(x)} \right) \right| = \frac{1}{2\pi} \left| \frac{dp}{dx} \right| \ll 1 \quad (16.2.15)$$

which agrees with our heuristic expectation, Eq. (16.2.4).

Connection with the Path Integral Formalism

Let us now rederive the semiclassical wave functions of Eq. (16.2.10) in the path integral approach, in the semiclassical approximation, in which one writes

$$U_{\text{cl}}(xt, x'0) = A e^{(i/\hbar)S_{\text{cl}}[xt; x'0]} \quad (16.2.16)$$

where $S_{\text{cl}}[xt; x'0]$ is the action for the classical path connecting the end points $(x'0)$ and (xt) . We have chosen the initial time $t'=0$, assuming the problem has time translation invariance. The prefactor A has no dependence on x, x' , or t in our approximation in which the “area” under the functional integral is replaced by the value of the integrand at the stationary point times some constant. We will illustrate the procedure in the following situation.

- (1) The potential is always negative and goes to a constant, chosen for convenience to be zero, as $|x| \rightarrow \infty$.
- (2) The particle is in a quantum state with energy $E > 0$. This means that at the classical level there are no turning points, a fact which will prove significant.

Our strategy will be as follows. We will first show how to project out the exact wave functions from the exact propagator by performing some integrals. Then we will compute the propagator in the semiclassical approximation and project out the corresponding (approximate) wave functions of Eq. (16.2.10)).

How are the wave functions to be extracted from the propagator? In general, we have

$$U(xt; x') \equiv U(xt; x'0) = \sum \psi_n(x)\psi_n^*(X') e^{-iE_n t/\hbar}$$

where the sum may have to be an integral if the spectrum is continuous. Indeed, in the present problem this is the case. In the asymptotic region $|x| \rightarrow \infty$, the solutions must be plane waves with momentum

$$p_\infty = \pm \sqrt{2mE}$$

So we can use p as a label for the $E > 0$ eigenstates. (Hereafter the subscript on p_∞ will be dropped.) At each energy E , there will be two solutions just as in the free particle case, with $E = p^2/2m$. The wave functions will, however, not be plane waves in the interior and these are what we are after. So let us begin with

$$U(xt; x') = \int_{-\infty}^{\infty} \frac{dp}{2\pi\hbar} \psi_p(x)\psi_p^*(x') e^{-ip^2 t/2m\hbar} + BS \quad (16.2.17)$$

where BS stands for the sum over bound states. These will be dropped since we will soon be projecting out the states at some positive energy. Factors like π in the p -integration do not matter too much, we just want to get the right x -dependence of the wave functions and not their normalization.

Remember that $U(t)$ has been constructed to propagate the system forward in time, i.e., it is to be used for $t > 0$. Let us now define its transform

$$U(x, x', z) = \int_0^{\infty} dt U(xt; x') e^{izt/\hbar} \quad z = E + i\varepsilon \quad (16.2.18a)$$

where ε is a positive infinitesimal introduced to ensure convergence as $t \rightarrow \infty$. In what follows we will not distinguish between ε and a finite positive factor times ε and will ignore terms of higher order in ε .

It is readily seen by combining the two equations above that

$$U(x, x', z) = 2m \int_{-\infty}^{\infty} \frac{dp}{2\pi i} \frac{\psi_p(x)\psi_p^*(x')}{p^2 - 2mE - i\varepsilon} \quad (16.2.18b)$$

Writing (to order ε)

$$p^2 - 2mE - i\varepsilon = p^2 - (\sqrt{2mE} + i\varepsilon)^2$$

we factorize as follows:

$$\frac{1}{p^2 - 2mE - i\varepsilon} = \frac{1}{2\sqrt{2mE}} \left[\frac{1}{p - \sqrt{2mE} - i\varepsilon} - \frac{1}{p + \sqrt{2mE} + i\varepsilon} \right]$$

and use the formula (derived and explained in Appendix A4)

$$\frac{1}{(x-a) \mp i\varepsilon} = \mathcal{P} \frac{1}{x-a} \pm i\pi\delta(x-a)$$

to obtain

$$\begin{aligned} U(x, x', z) &= \sqrt{\frac{m}{2E}} \int_{-\infty}^{\infty} \frac{dp}{2\pi i} (\psi_p(x)\psi_p^*(x')) \\ &\times \left[\mathcal{P}\left(\frac{1}{p - \sqrt{2mE}}\right) + i\pi\delta(p - \sqrt{2mE}) \right. \\ &\quad \left. - \mathcal{P}\left(\frac{1}{p + \sqrt{2mE}}\right) + i\pi\delta(p + \sqrt{2mE}) \right] \end{aligned}$$

where \mathcal{P} means the principal value integral is to be evaluated. As it stands, U depends on the eigenfunctions not just at the energy E , but nearby energies also, due to the principal value integral. So we form the combination which singles out the eigenfunctions at just one energy:

$$\begin{aligned} U(x, x', z) + [U(x', x, z)]^* &= \sqrt{\frac{m}{2E}} \int_{-\infty}^{\infty} dp (\psi_p(x)\psi_p^*(x')) \left[\delta(p - \sqrt{2mE}) + \delta(p + \sqrt{2mE}) \right] \\ &= \sqrt{\frac{m}{2E}} [\psi_{\sqrt{2mE}}(x)\psi_{\sqrt{2mE}}^*(x') + \psi_{-\sqrt{2mE}}(x)\psi_{-\sqrt{2mE}}^*(x')] \quad (16.2.19) \end{aligned}$$

We now compare this with $U_{\text{cl}}(x, x', z)$:

$$U_{\text{cl}}(x, x', z) = \int_0^\infty dt U_{\text{cl}}(x, x', t) e^{(i/\hbar)(E + i\varepsilon)t} = \int_0^\infty dt e^{(i/\hbar)S_{\text{cl}}[x, x', t]} e^{(i/\hbar)(E + i\varepsilon)t}$$

Since U_{cl} was itself evaluated in the stationary point approximation of a functional integral (on the basis of the smallness of \hbar), to be consistent, we must evaluate the ordinary integral in t also by stationary phase, by setting the argument of the exponent to its value at the point t^* defined by

$$\frac{\partial S}{\partial t} + E = -E_{\text{cl}} + E = 0 \quad (16.2.20)$$

and equating the integral to the integrand at this point times some constant. For the stationary point, we have recalled Eq. (2.8.18) and dropped the convergence factor ε since it is not needed. What the stationary point does is to choose from all trajectories connecting the two given end points (with various travel times) the one whose classical energy equals the energy E of the quantum state [and whose travel time is $t^* = t^*(E)$].

Note that previously we were interested in trajectories that connected x' to x in a fixed time t . Given the equations of motion are second order in time, there will be just one trajectory that can meet these requirements. But now we are asking for a trajectory with a fixed energy connecting the two points x' and x with no restriction on travel time. This can generally have many solutions. For example, in a confining potential like the oscillator (or in any bound state) a particle leaving x' with some energy E will hit any other point x (within the turning points) an infinite number of times as it rattles back and forth. Each such orbit will have a different travel time and $U_{\text{cl}}(x, x', E)$ will receive contributions from an infinite number of stationary points obeying Eq. (16.2.20).

In the present problem with no turning points, there will be just two solutions—call them R and L —which are moving to the right or to the left. The right mover can go from x' to x if $x' < x$ and the left mover if $x' > x$.

Let us proceed with our calculation bearing all this in mind, and start with

$$U_{\text{cl}}(x, x', E) = A' \sum_{R, L} e^{(i/\hbar)[S_{\text{cl}}(x, x', t^*) + Et^*]}$$

where A' is a new constant. We now manipulate as follows bearing in mind that $E_{\text{cl}} = E$ is conserved on the classical path:

$$\begin{aligned} S_{\text{cl}}[x, x', t^*] &= \int_0^{t^*} (T - V) dt = \int_0^{t^*} 2T dt - Et^* \\ &= \int_0^{t^*} p(x(t)) \frac{dx}{dt} dt - Et^* = \int_{x'}^x p(x'') dx'' - Et^* \\ &\equiv W[x, x', E] - Et^* \end{aligned} \quad (16.2.21a)$$

It follows that

$$\begin{aligned} U_{\text{cl}}(x, x', E) &= A' \sum_{R, L} \exp \left[\frac{i}{\hbar} \int_{x'}^x p(x'') dx'' \right] \\ &= A' \sum_{R, L} \exp \left(\frac{i}{\hbar} W[x, x', E] \right) \end{aligned} \quad (16.2.21b)$$

If $x > x'$, the classical momentum has to be positive and we set $p(x) = \sqrt{2m(E - V(x))}$ whereas if $x < x'$ we make the other choice for the square root. Thus

$$\begin{aligned} U_{\text{cl}}(x, x', E) &= \theta(x - x') A' \exp \left(\frac{i}{\hbar} \left[\int_{x'}^x \sqrt{2m(E - V(x''))} dx'' \right] \right) \\ &\quad + \theta(x' - x) A' \exp \left(-\frac{i}{\hbar} \left[\int_{x'}^x \sqrt{2m(E - V(x''))} dx'' \right] \right) \end{aligned} \quad (16.2.22)$$

We now find

$$\begin{aligned} U_{\text{cl}}(x, x', E) + U_{\text{cl}}^*(x', x, E) &= A' \exp \left(\frac{i}{\hbar} \left[\int_{x'}^x \sqrt{2m(E - V(x''))} dx'' \right] \right) \\ &\quad + A' \exp \left(-\frac{i}{\hbar} \left[\int_{x'}^x \sqrt{2m(E - V(x''))} dx'' \right] \right) \end{aligned} \quad (16.2.23)$$

Comparing this to Eq. (16.2.19) we find

$$\psi_{\pm\sqrt{2mE}}(x) \psi_{\pm\sqrt{2mE}}^*(x') \simeq \exp \left(\pm \frac{i}{\hbar} \left[\int_{x'}^x \sqrt{2m(E - V(x''))} dx'' \right] \right)$$

It is easily seen that this may be written as

$$\psi_{\pm}(x) = \psi(x_0) \exp \left(\pm \frac{i}{\hbar} \left[\int_{x_0}^x \sqrt{2m(E - V(x''))} dx'' \right] \right)$$

Several comments are in order.

First, if we want to get Eq. (16.2.13), with the factor $p^{-1/2}$, we need to do a more accurate calculation. So far we have evaluated the functional integral and the ordinary t integral by setting the integral equal to the value of the integrand at the stationary point, times some constant to represent the “area” around this point. To get Eq. (16.2.13) we must approximate the integrands at the stationary point by Gaussians and do the Gaussian integrals. If you are interested in the details, consult one of the works in the Bibliography at the end of Chapter 21.

Second, note that in going from $U(t)$ to $U(E)$, (at fixed x', x), we shifted our attention from paths with a fixed time to paths with a fixed energy. Since

$E = -\partial S_{\text{cl}}/\partial t$, one is trading t for the derivative of S_{cl} with respect to t . This clearly requires a Legendre transformation, as explained in Section 2.5. It is clear from Eq. (16.2.21a) that $W(E)$ is the Legendre transform in question. This idea is pursued further in Exercise (16.2.1).

Finally, let us look at the combination $U(x, x', E) + U^*(x', x, E)$ we were led to in our attempt to filter out one energy. From our discussion of time reversal you should expect that the complex conjugated version of U (with initial and final points exchanged) stands for the time-reversed propagator. If so, the integral of this combination over positive times is the integral of U for all times. It is now clear that such an integration done on Eq. (16.2.17) will indeed produce a factor $\delta(E - p^2/2m)$ which projects out states at the energy E . This point is pursued in Exercise (16.2.3).

Exercise 16.2.1. Consider the function $W(E)$ introduced in Eq. (16.2.21a). Since $-E$ is the t derivative of $S(t)$, it follows that $W(E)$, must be the Legendre transform of S . In this case t must emerge as the derivative of $W(E)$ with respect to E . Verify that differentiation of the formula

$$W(E) = \int_{x'}^x \sqrt{2m(E - V(x''))} dx''$$

gives the time t taken to go from start to finish at this preassigned energy.

Exercise 16.2.2. Consider the free particle problem using the approach given above. Now that you know the wave functions explicitly, evaluate the integral in Eq. (16.2.18b) by contour integration to obtain

$$U(x, x', t) = \theta(x - x') e^{(i/\hbar)\sqrt{2mE}(x - x')} + \theta(x' - x) e^{(-i/\hbar)\sqrt{2mE}(x - x')}.$$

In doing the contour integrals, ask in which half-plane you can close the contour for a given sign of $x - x'$. Compare the above result to the semiclassical result and make sure it all works out. Note that there is no need to form the combination $U(x, x', t) + U^*(x, x', t)$; we can calculate both the principal part and the delta function contributions explicitly because we know the p -dependence of the integrand explicitly. To see this more clearly, avoid the contour integral approach and use instead the formula for $(x \pm i\varepsilon)^{-1}$ given above. Evaluate the principal value integral and the contribution from the delta function and see how they add up to the result of contour integration. Although both routes are possible here, in the problem with $V \neq 0$ contour integration is not possible since the p -dependence of the wave function in the complex p plane is not known. The advantage of the $U + U^*$ approach is that it only refers to quantities on the real axis and at just one energy.

Exercise 16.2.3. Let us take a second look at our derivation. We worked quite hard to isolate the eigenfunctions at one energy: we formed the combination $U(x, x', z) + U^*(x, x', z)$ to get rid of the principal part and filter out the delta function. Now it is clear that if in Eq. (16.2.18a) we could integrate in the range $-\infty \leq t \leq \infty$, we would get the δ function we want. What kept us from doing that was the fact the $U(t)$ was constructed to be used for $t > 0$. However, if we use the time evolution operator $e^{-(i/\hbar)Ht}$ for negative times, it will simply tell us what the system was doing at earlier times, assuming the same Hamiltonian. Thus we can make sense of the operator for negative times as well and define a transform that extends over all times. This is true in classical mechanics as well. For instance, if a stone is thrown straight up from a tall building and we ask when it will be at ground level, we get two answers,