CHAPTER 3

Theory of Angular Momentum

This chapter is concerned with a systematic treatment of angular momentum and related topics. The importance of angular momentum in modern physics can hardly be overemphasized. A thorough understanding of angular momentum is essential in molecular, atomic, and nuclear spectroscopy; angular-momentum considerations play an important role in scattering and collision problems as well as in bound-state problems. Furthermore, angular-momentum concepts have important generalizations—isospin in nuclear physics, SU(3), SU(2)\(\otimes\)U(1) in particle physics, and so forth.

3.1. ROTATIONS AND ANGULAR MOMENTUM COMMUTATION RELATIONS

Finite Versus Infinitesimal Rotations

We recall from elementary physics that rotations about the same axis commute, whereas rotations about different axes do not. For instance, a 30° rotation about the z-axis followed by a 60° rotation about the same z-axis is obviously equivalent to a 60° rotation followed by a 30° rotation, both about the same axis. However, let us consider a 90° rotation about the z-axis, denoted by \(R_z(\pi/2)\), followed by a 90° rotation about the x-axis, denoted by \(R_x(\pi/2)\); compare this with a 90° rotation about the x-axis followed by a 90° rotation about the z-axis. The net results are different, as we can see from Figure 3.1.

Our first basic task is to work out quantitatively the manner in which rotations about different axes fail to commute. To this end, we first recall how to represent rotations in three dimensions by \(3 \times 3\) real, orthogonal matrices. Consider a vector \(\mathbf{V}\) with components \(V_x, V_y,\) and \(V_z\). When we rotate, the three components become some other set of numbers, \(V'_x, V'_y,\) and \(V'_z\). The old and new components are related via a \(3 \times 3\) orthogonal matrix \(R\):

\[
\begin{pmatrix}
V'_x \\
V'_y \\
V'_z
\end{pmatrix} =
\begin{pmatrix}
R
\end{pmatrix}
\begin{pmatrix}
V_x \\
V_y \\
V_z
\end{pmatrix},
\]

where the superscript \(T\) stands for a transpose of a matrix. It is a property of orthogonal matrices that

\[
\sqrt{V'^2_x + V'^2_y + V'^2_z} = \sqrt{V^2_x + V^2_y + V^2_z}
\]

is automatically satisfied.
To be definite, we consider a rotation about the $z$-axis by angle $\phi$. The convention we follow throughout this book is that a rotation operation affects a physical system itself, as in Figure 3.1, while the coordinate axes remain unchanged. The angle $\phi$ is taken to be positive when the rotation in question is counterclockwise in the $xy$-plane, as viewed from the positive $z$-side. If we associate a right-handed screw with such a rotation, a positive $\phi$ rotation around the $z$-axis means that the screw is advancing in the positive $z$-direction. With this convention, we easily verify that

$$R_z(\phi) = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3.1.3)$$

Had we adopted a different convention, in which a physical system remained fixed but the coordinate axes rotated, this same matrix with a positive $\phi$ would have represented a \textit{clockwise} rotation of the $x$- and $y$-axes, when viewed from the positive $z$-side. It is obviously important not to mix the two conventions! Some authors distinguish the two approaches by using “active rotations” for physical systems rotated and “passive rotations” for the coordinate axes rotated.

We are particularly interested in an infinitesimal form of $R_z$:

$$R_z(\epsilon) = \begin{pmatrix} 1 - \frac{\epsilon^2}{2} & -\epsilon & 0 \\ \epsilon & 1 - \frac{\epsilon^2}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (3.1.4)$$

where terms of order $\epsilon^3$ and higher are ignored. Likewise, we have

$$R_x(\epsilon) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \frac{\epsilon^2}{2} & -\epsilon \\ 0 & \epsilon & 1 - \frac{\epsilon^2}{2} \end{pmatrix} \quad (3.1.5a)$$

and

$$R_y(\epsilon) = \begin{pmatrix} 1 - \frac{\epsilon^2}{2} & 0 & \epsilon \\ 0 & 1 & 0 \\ -\epsilon & 0 & 1 - \frac{\epsilon^2}{2} \end{pmatrix}, \quad (3.1.5b)$$

which may be read from (3.1.4) by cyclic permutations of $x$, $y$, $z$—that is, $x \to y$, $y \to z$, $z \to x$. Compare now the effect of a $y$-axis rotation followed by an $x$-axis rotation with that of an $x$-axis rotation followed by a $y$-axis rotation. Elementary matrix manipulations lead to

$$R_y(\epsilon)R_x(\epsilon) = \begin{pmatrix} 1 - \frac{\epsilon^2}{2} & 0 & \epsilon \\ \epsilon & 1 - \frac{\epsilon^2}{2} & -\epsilon \\ -\epsilon & \epsilon & 1 - \epsilon^2 \end{pmatrix} \quad (3.1.6a)$$

and

$$R_x(\epsilon)R_y(\epsilon) = \begin{pmatrix} 1 - \frac{\epsilon^2}{2} & \epsilon^2 & 0 \\ 0 & 1 - \frac{\epsilon^2}{2} & -\epsilon \\ -\epsilon & \epsilon & 1 - \epsilon^2 \end{pmatrix}. \quad (3.1.6b)$$

From (3.1.6a) and (3.1.6b) we have the first important result: Infinitesimal rotations about different axes do commute if terms of order $\epsilon^2$ and higher are ignored.* The second and even more important result concerns the manner in which rotations about different axes \textit{fail} to commute when terms of order $\epsilon^3$ are kept:

$$R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon) = \begin{pmatrix} 0 & -\epsilon^2 & 0 \\ \epsilon^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = R_z(\epsilon^3) - 1, \quad (3.1.7)$$

where all terms of order higher than $\epsilon^2$ have been ignored throughout this derivation. We also have

$$1 = R_{\text{any}}(0) \quad (3.1.8)$$

where \textit{any} stands for any rotation axis. Thus the final result can be written as

$$R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon) = R_z(\epsilon^3) - R_{\text{any}}(0). \quad (3.1.9)$$

This is an example of the commutation relations between rotation operations about different axes, which we will use later in deducing the angular-momentum commutation relations in quantum mechanics.

*Actually there is a familiar example of this in elementary mechanics. The angular velocity vector $\omega$ that characterizes an infinitesimal change in rotation angle during an infinitesimal time interval follows the usual rule of vector addition, including commutativity of vector addition. However, we cannot ascribe a vectorial property to a \textit{finite} angular change.
Infinitesimal Rotations in Quantum Mechanics

So far we have not used quantum-mechanical concepts. The matrix $R$ is just a $3 \times 3$ orthogonal matrix acting on a vector $\mathbf{V}$ written in column matrix form. We must now understand how to characterize rotations in quantum mechanics.

Because rotations affect physical systems, the state ket corresponding to a rotated system is expected to look different from the state ket corresponding to the original unrotated system. Given a rotation operation $R$, characterized by a $3 \times 3$ orthogonal matrix $R$, we associate an operator $\mathcal{D}(R)$ in the appropriate ket space such that

$$|\alpha\rangle_R = \mathcal{D}(R)|\alpha\rangle,$$

(3.1.10)

where $|\alpha\rangle_R$ and $|\alpha\rangle$ stand for the kets of the rotated and original system, respectively. Note that the $3 \times 3$ orthogonal matrix $R$ acts on a column matrix made up of the three components of a classical vector, while the operator $\mathcal{D}(R)$ acts on state vectors in ket space. The matrix representation of $\mathcal{D}(R)$, which we will study in great detail in the subsequent sections, depends on the dimensionality $N$ of the particular ket space in question. For $N = 2$, appropriate for describing a spin $\frac{1}{2}$ system with no other degrees of freedom, $\mathcal{D}(R)$ is represented by a $2 \times 2$ matrix; for a spin 1 system, the appropriate representation is a $3 \times 3$ unitary matrix, and so on.

To construct the rotation operator $\mathcal{D}(R)$, it is again fruitful to examine first its properties under an infinitesimal rotation. We can almost guess how we must proceed by analogy. In both translations and time evolution, which we studied in Sections 1.6 and 2.1, respectively, the appropriate infinitesimal operators could be written as

$$U_\epsilon = 1 - iG\epsilon$$

(3.1.11)

with a Hermitian operator $G$. Specifically,

$$G \rightarrow \frac{P_x}{\hbar}, \quad \epsilon \rightarrow dx'$$

(3.1.12)

for an infinitesimal translation by a displacement $dx'$ in the $x$-direction, and

$$G \rightarrow \frac{H}{\hbar}, \quad \epsilon \rightarrow dt$$

(3.1.13)

for an infinitesimal time evolution with time displacement $dt$. We know from classical mechanics that angular momentum is the generator of rotation in much the same way as momentum and Hamiltonian are the generators of translation and time evolution, respectively. We therefore define the angular momentum operator $J_k$ in such a way that the operator for an infinitesimal rotation around the $k$th axis by angle $d\phi$ can be obtained by letting

$$G \rightarrow \frac{J_k}{\hbar}, \quad \epsilon \rightarrow d\phi$$

(3.1.14)

in (3.1.11). With $J_k$ taken to be Hermitian, the infinitesimal rotation operator is guaranteed to be unitary and reduces to the identity operator in the limit $d\phi \rightarrow 0$. More generally, we have

$$\mathcal{D}(\hbar, d\phi) = 1 - i\left[\frac{J}{\hbar}\right]d\phi$$

(3.1.15)

for a rotation about the direction characterized by a unit vector $\mathbf{n}$ by an infinitesimal angle $d\phi$.

We stress that in this book we do not define the angular-momentum operator to be $\mathbf{x} \times \mathbf{p}$. This is important because spin angular momentum, to which our general formalism also applies, has nothing to do with $x_i$ and $p_j$. Put another way, in classical mechanics one can prove that the angular momentum defined to be $\mathbf{x} \times \mathbf{p}$ is the generator of a rotation; in contrast, in quantum mechanics we define $\mathbf{J}$ so that the operator for an infinitesimal rotation takes form (3.1.15).

A finite rotation can be obtained by compounding successively infinitesimal rotations about the same axis. For instance, if we are interested in a finite rotation about the $z$-axis by angle $\phi$, we consider

$$\mathcal{D}_z(\phi) = \lim_{N \rightarrow \infty} \left[1 - i\frac{J_z}{\hbar}\left(\frac{\phi}{N}\right)\right]^N$$

$$= \exp\left(-\frac{iJ\phi}{\hbar}\right)$$

$$= 1 - \frac{iJ\phi}{\hbar} - \frac{J^2\phi^2}{2\hbar^2} + \cdots$$

(3.1.16)

In order to obtain the angular-momentum commutation relations, we need one more concept. As we remarked earlier, for every rotation $R$ represented by a $3 \times 3$ orthogonal matrix $R$ there exists a rotation operator $\mathcal{D}(R)$ in the appropriate ket space. We further postulate that $\mathcal{D}(R)$ has the same group properties as $R$:

- **Identity:** $R \cdot 1 = 1$ implies $\mathcal{D}(R) \cdot 1 = \mathcal{D}(R)$

(3.1.17a)

- **Closure:** $R_1 R_2 = R_3$ implies $\mathcal{D}(R_1) \mathcal{D}(R_2) = \mathcal{D}(R_3)$

(3.1.17b)

- **Inverses:** $R^{-1} R = 1$ implies $\mathcal{D}(R) \mathcal{D}^{-1}(R) = 1$

(3.1.17c)

- **Associativity:** $R_1 (R_2 R_3) = (R_1 R_2) R_3$ implies $\mathcal{D}(R_1) [\mathcal{D}(R_2) \mathcal{D}(R_3)]$.

(3.1.17d)
3.2. Spin ½ Systems and Finite Rotations

Rotation Operator for Spin ½

The lowest number, \( N \), of dimensions in which the angular-momentum commutation relations (3.1.20) are realized, is \( N = 2 \). The reader has already checked in Problem 8 of Chapter 1 that the operators defined by

\[
S_x = \left( \frac{\hbar}{2} \right) \{(+\hat{\chi}-\hat{\chi}) + (-\hat{\chi}+\hat{\chi}) \}, \\
S_y = \left( \frac{i\hbar}{2} \right) \{-(-\hat{\chi}+\hat{\chi}) + (+\hat{\chi}-\hat{\chi}) \}, \\
S_z = \left( \frac{\hbar}{2} \right) \{(+\hat{\chi}+\hat{\chi}) - (-\hat{\chi}-\hat{\chi}) \}
\]

(3.2.1)
satisfy commutation relations (3.1.20) with \( J_k \) replaced by \( S_k \). It is not a priori obvious that nature takes advantage of the lowest dimensional realization of (3.1.20), but numerous experiments—from atomic spectroscopy to nuclear magnetic resonance—suffice to convince us that this is in fact the case.

Consider a rotation by a finite angle \( \phi \) about the z-axis. If the ket of a spin \( \frac{1}{2} \) system before rotation is given by \( |\alpha\rangle \), the ket after rotation is given by

\[
|\alpha\rangle_R = \mathcal{D}_z(\phi)|\alpha\rangle
\]

(3.2.2)

with

\[
\mathcal{D}_z(\phi) = \exp \left( -\frac{iS\phi}{\hbar} \right).
\]

(3.2.3)

To see that this operator really rotates the physical system, let us look at its effect on \( \langle S_z \rangle \). Under rotation this expectation value changes as follows:

\[
\langle S_z \rangle \rightarrow \langle \alpha | S_z | \alpha \rangle_R = \langle \alpha | \mathcal{D}_z(\phi) S_z \mathcal{D}_z(\phi)^\dagger | \alpha \rangle.
\]

(3.2.4)

We must therefore compute

\[
\exp \left( \frac{iS\phi}{\hbar} \right) S_z \exp \left( -\frac{iS\phi}{\hbar} \right).
\]

(3.2.5)

For pedagogical reasons we evaluate this in two different ways.

**Derivation 1:** Here we use the specific form of \( S_z \) given by (3.2.1). We then obtain for (3.2.5)

\[
\left( \frac{\hbar}{2} \right) \exp \left( \frac{iS\phi}{\hbar} \right) \{(+\hat{\chi}-\hat{\chi}) + (-\hat{\chi}+\hat{\chi}) \} \exp \left( -\frac{iS\phi}{\hbar} \right) = \exp \left( \frac{i\phi}{\hbar} \right) \left( \frac{\hbar}{2} \right) \{ (e^{i\phi/2}) + (e^{-i\phi/2}) \} = \exp \left( \frac{i\phi}{\hbar} \right) \frac{\hbar}{2} \{ (1 + \cos \phi) + i(\sin \phi) \} = S_z \cos \phi - S_z \sin \phi.
\]

(3.2.6)

**Derivation 2:** Alternatively we may use formula (2.3.47) to evaluate (3.2.5):

\[
\exp \left( \frac{iS\phi}{\hbar} \right) S_z \exp \left( -\frac{iS\phi}{\hbar} \right) = S_z + \left( \frac{i\phi}{\hbar} \right) \left[ S_z, S_z \right] \frac{i\hbar}{h^2S_y} + \left( \frac{1}{2!} \right) \left( \frac{i\phi}{\hbar} \right)^2 \left[ S_z, [S_z, S_z] \right] \frac{1}{hS_y} + \left( \frac{1}{3!} \right) \left( \frac{i\phi}{\hbar} \right)^3 \left[ S_z, [S_z, [S_z, S_z]] \right] \frac{h^2S_y}{i\hbar^3} + \ldots
\]

(3.2.7)

\[
= S_z \left( 1 - \frac{\phi^2}{2!} + \ldots \right) - S_y \left( \phi - \frac{\phi^3}{3!} + \ldots \right)
\]

(3.2.8)

\[
= S_z \cos \phi - S_y \sin \phi.
\]
Notice that in derivation 2 we used only the commutation relations for \( S_z \), so this method can be generalized to rotations of systems with angular momentum higher than \( \frac{1}{2} \).

For spin \( \frac{1}{2} \), both methods give
\[
\langle S_x \rangle \rightarrow_R \langle a|S_x|a \rangle_R = \langle S_x \rangle \cos \phi - \langle S_y \rangle \sin \phi,
\]
(3.2.8)
where the expectation value without subscripts is understood to be taken with respect to the (old) unrotated system. Similarly,
\[
\langle S_y \rangle \rightarrow \langle S_y \rangle \cos \phi + \langle S_y \rangle \sin \phi.
\]
(3.2.9)
As for the expectation value of \( S_z \), there is no change because \( S_z \) commutes with \( \mathcal{B}_z(\phi) \):
\[
\langle S_z \rangle \rightarrow \langle S_z \rangle.
\]
(3.2.10)
Relations (3.2.8), (3.2.9), and (3.2.10) are quite reasonable. They show that rotation operator (3.2.3), when applied to the state ket, does rotate the expectation value of \( S \) around the \( z \)-axis by angle \( \phi \). In other words, the expectation value of the spin operator behaves as though it were a classical vector under rotation:
\[
\langle S_k \rangle \rightarrow \sum_i R_{ki} \langle S_i \rangle,
\]
(3.2.11)
where \( R_{ki} \) are the elements of the \( 3 \times 3 \) orthogonal matrix \( R \) that specifies the rotation in question. It should be clear from our derivation 2 that this property is not restricted to the spin operator of spin \( \frac{1}{2} \) systems. In general, we have
\[
\langle J_k \rangle \rightarrow \sum_i R_{ki} \langle J_i \rangle
\]
(3.2.12)
under rotation, where \( J_k \) are the generators of rotations satisfying the angular-momentum commutation relations (3.1.20). Later we will show that relations of this kind can be further generalized to any vector operator.

So far everything has been as expected. But now, be prepared for a surprise! We examine the effect of rotation operator (3.2.3) on a general ket,
\[
|\alpha \rangle = |+\rangle|+|a\rangle + |-\rangle|+|a\rangle,
\]
(3.2.13)
a little more closely. We see that
\[
\exp \left( -\frac{iS_\phi}{\hbar} \right)|\alpha\rangle = e^{-iφ/2}|+\rangle|+|a\rangle + e^{iφ/2}|-\rangle|-|a\rangle.
\]
(3.2.14)
The appearance of the half-angle \( φ/2 \) here has an extremely interesting consequence.

Let us consider a rotation by \( 2\pi \). We then have
\[
|\alpha\rangle \rightarrow (2\pi) \rightarrow -|\alpha\rangle.
\]
(3.2.15)

1.2. Spin \( \frac{1}{2} \) Systems and Finite Rotations

So the ket for the \( 360° \) rotated state differs from the original ket by a minus sign. We would need a \( 720° \) (\( φ = 4\pi \)) rotation to get back to the same ket with a plus sign. Notice that this minus sign disappears for the expectation value of \( S \) because \( S \) is sandwiched by \( |\alpha \rangle \) and \( \langle a| \), both of which change sign. Will this minus sign ever be observable? We will give the answer to this interesting question after we discuss spin precession once again.

**Spin Precession Revisited**

We now treat the problem of spin precession, already discussed in Section 2.1, from a new point of view. We recall that the basic Hamiltonian of the problem is given by
\[
H = -\left( \frac{e}{m_c} \right) \mathbf{S} \cdot \mathbf{B} = \omega S_z,
\]
(3.2.16)
where
\[
\omega = \frac{|e|B}{m_c}.
\]
(3.2.17)
The time-evolution operator based on this Hamiltonian is given by
\[
\mathcal{U}(t,0) = \exp \left( -iHt \right) = \exp \left( -\frac{i\mathbf{S} \cdot \mathbf{B} t}{\hbar} \right).
\]
(3.2.18)
Comparing this equation with (3.2.3), we see that the time-evolution operator here is precisely the same as the rotation operator in (3.2.3) with \( φ \) set equal to \( \omega t \). In this manner we see immediately why this Hamiltonian causes spin precession. Paraphrasing (3.2.8), (3.2.9), and (3.2.10), we obtain
\[
\langle S_x \rangle = \langle S_x \rangle_{t=0} \cos \omega t - \langle S_y \rangle_{t=0} \sin \omega t,
\]
(3.2.19a)
\[
\langle S_y \rangle = \langle S_y \rangle_{t=0} \cos \omega t + \langle S_x \rangle_{t=0} \sin \omega t,
\]
(3.2.19b)
\[
\langle S_z \rangle = \langle S_z \rangle_{t=0}.
\]
(3.2.19c)
After \( t = 2\pi/\omega \), the spin returns to its original direction.

This set of equations can be used to discuss the spin precession of a **muon**, an electronlike particle which, however, is 210 times as heavy. The muon magnetic moment can be determined from other experiments—for example, the hyperfine splitting in muonium, a bound state of a positive muon and an electron—to be \( e\hbar/2m_c \), just as expected from Dirac's relativistic theory of spin \( \frac{1}{2} \) particles. (We will here neglect very small corrections that arise from quantum field theory effects). Knowing the magnetic moment we can predict the angular frequency of precession. So (3.2.19) can be and, in fact, has been checked experimentally. In practice, as the external magnetic field causes spin precession, the spin direction is analyzed by taking advantage of the fact that electrons from muon decay tend to be emitted preferentially in the direction opposite to the muon spin.
Let us now look at the time evolution of the state ket itself. Assuming that the initial \((t=0)\) ket is given by (3.2.13), we obtain after time \(t\)
\[
|\alpha, t_0 = 0; t\rangle = e^{-i\omega t/2}|+\rangle + |\alpha\rangle + e^{+i\omega t/2}|-\rangle = |\alpha\rangle.
\] (3.2.20)
Expression (3.2.20) acquires a minus sign at \(t = 2\pi/\omega\), and we must wait until \(t = 4\pi/\omega\) to get back to the original state ket with the same sign. To sum up, the period for the state ket is twice as long as the period for spin precession
\[
\tau_{\text{precession}} = \frac{2\pi}{\omega},
\] (3.2.21a)
\[
\tau_{\text{state ket}} = \frac{4\pi}{\omega}.
\] (3.2.21b)

**Neutron Interferometry Experiment to Study \(2\pi \) Rotations**

We now describe an experiment performed to detect the minus sign in (3.2.15). Quite clearly, if every state ket in the universe is multiplied by a minus sign, there will be no observable effect. The only way to detect the predicted minus sign is to make a comparison between an unrotated state and a rotated state. As in gravity-induced quantum interference, discussed in Section 2.6, we rely on the art of neutron interferometry to verify this extraordinary prediction of quantum mechanics.

A nearly monoenergetic beam of thermal neutrons is split into two parts—path \(A\) and path \(B\); see Figure 3.2. Path \(A\) always goes through a magnetic-field-free region; in contrast, path \(B\) enters a small region where a static magnetic field is present. As a result, the neutron state ket going via path \(B\) suffers a phase change \(e^{+i\omega T/2}\), where \(T\) is the time spent in the \(B \neq 0\) region and \(\omega\) is the spin-precession frequency
\[
\omega = \frac{g_n e B}{m_p c}, \quad (g_n \approx -1.91)
\] (3.2.22)
for the neutron with a magnetic moment of \(g_n e h/2m_p c\), as we can see if we compare this with (3.2.17), which is appropriate for the electron with magnetic moment \(e\hbar/2m_e c\). When path \(A\) and path \(B\) meet again in the interference region of Figure 3.2, the amplitude of the neutron arriving via path \(B\) is
\[
c_2 = c_1 (B = 0) e^{-i\omega T/2},
\] (3.2.23)
while the amplitude of the neutron arriving via path \(A\) is \(c_1\), independent of \(B\). So the intensity observable in the interference region must exhibit a sinusoidal variation
\[
\cos \left(\frac{\pm \omega T}{2} + \delta\right),
\] (3.2.24)
where \(\delta\) is the phase difference between \(c_1\) and \(c_2\) \((B = 0)\). In practice, \(T\), the time spent in the \(B \neq 0\) region, is fixed but the precession frequency \(\omega\) is varied by changing the strength of the magnetic field. The intensity in the interference region as a function of \(B\) is predicted to have a sinusoidal variation. If we call \(\Delta B\) the difference in \(B\) needed to produce successive maxima, we can easily show that
\[
\Delta B = \frac{4\pi \hbar c}{eg_n \lambda l},
\] (3.2.25)
where \(l\) is the path length.

In deriving this formula we used the fact that a \(4\pi\) rotation is needed for the state ket to return to the original ket with the same sign, as required by our formalism. If, on the other hand, our description of spin \(1/2\) systems were incorrect and the ket were to return to its original ket with the same sign under a \(2\pi\) rotation, the predicted value for \(\Delta B\) would be just one-half of (3.2.25).

Two different groups have conclusively demonstrated experimentally that prediction (3.2.25) is correct to an accuracy of a fraction of a percent.* This is another triumph of quantum mechanics. The nontrivial prediction (3.2.15) has been experimentally established in a direct manner.

**Pauli Two-Component Formalism**

Manipulations with the state kets of spin \(1/2\) systems can be conveniently carried out using the two-component spinor formalism introduced by W. Pauli in 1926. In Section 1.3 we learned how a ket (bra) can be represented by a column (row) matrix; all we have to do is arrange the expansion coefficients in terms of a certain specified set of base kets into a

---

column (row) matrix. In the spin ½ case we have

\[
| + \rangle \doteq (1, 0) = x_+ \quad | - \rangle \doteq (0, 1) = x_-
\]

(3.2.27a)

for the base kets and bras and

\[
| \alpha \rangle = | + \rangle \langle + | \alpha \rangle + | - \rangle \langle - | \alpha \rangle = \begin{pmatrix} + | \alpha \rangle \\ - | \alpha \rangle \end{pmatrix}
\]

(3.2.27b)

and

\[
\langle \alpha | = \langle \alpha | + \rangle \langle + | \alpha \rangle + \langle \alpha | - \rangle \langle - | \alpha \rangle = \begin{pmatrix} \langle + | \alpha \rangle \\ \langle - | \alpha \rangle \end{pmatrix}
\]

(3.2.27b)

for an arbitrary state ket and the corresponding state bra. Column matrix (3.2.27a) is referred to as a two-component spinor and is written as

\[
\chi = \begin{pmatrix} c_+ \\ c_- \end{pmatrix} = c_+ x_+ + c_- x_-,
\]

(3.2.28)

where \(c_+\) and \(c_-\) are, in general, complex numbers. For \(\chi^\dagger\) we have

\[
\chi^\dagger = \langle \alpha | \alpha \rangle = (c_+^*, c_-^*).
\]

(3.2.29)

The matrix elements \(\langle \pm | S_k | \pm \rangle\) and \(\langle \pm | S_k | - \rangle\), apart from \(\hbar/2\), are to be set equal to those of \(2 \times 2\) matrices \(\sigma_k\), known as the Pauli matrices. We identify

\[
\langle \pm | S_k | \pm \rangle \equiv \begin{pmatrix} \hbar/2 \\ \sigma \end{pmatrix}_{\pm, \pm}, \quad \langle \pm | S_k | - \rangle \equiv \begin{pmatrix} \hbar/2 \\ \sigma \end{pmatrix}_{\pm, -}.
\]

(3.2.30)

We can now write the expectation value \(\langle S_k \rangle\) in terms of \(\chi\) and \(\sigma_k\):

\[
\langle S_k \rangle = \langle \alpha | S_k | \alpha \rangle = \sum_{a', a''} \sum_a \langle \alpha | a' \rangle \langle a' | S_k | a'' \rangle \langle a'' | a \rangle = \left( \frac{\hbar}{2} \right) \chi^\dagger \sigma_k \chi,
\]

(3.2.31)

where the usual rule of matrix multiplication is used in the last line. Explicitly, we see from (3.2.1) together with (3.2.30) that

\[
\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\]

(3.2.32)

where the subscripts 1, 2, and 3 refer to \(x, y, z\), respectively.

We record some properties of the Pauli matrices. First,

\[
\sigma_i^2 = 1
\]

(3.2.33a)

\[
\sigma_i \sigma_j + \sigma_j \sigma_i = 0, \quad \text{for} \ i \neq j,
\]

(3.2.33b)

where the right-hand side of (3.2.33a) is to be understood as the \(2 \times 2\) identity matrix. These two relations are, of course, equivalent to the anticommutation relations

\[
\{ \sigma_i, \sigma_j \} = 2 \delta_{ij}.
\]

(3.2.34)

We also have the commutation relations

\[
[ \sigma_i, \sigma_j ] = 2 i \epsilon_{ijk} \sigma_k.
\]

(3.2.35)

which we see to be the explicit \(2 \times 2\) matrix realizations of the angular-momentum commutation relations (3.1.20). Combining (3.2.34) and (3.2.35), we can obtain

\[
\sigma_i \sigma_2 = - \sigma_2 \sigma_i = i \sigma_3, \ldots
\]

(3.2.36)

Notice also that

\[
\sigma_i^\dagger = \sigma_i, \quad \det(\sigma) = -1, \quad \text{Tr}(\sigma) = 0.
\]

(3.2.37)

We now consider \(\sigma \cdot a\), where \(a\) is a vector in three dimensions. This is actually to be understood as a \(2 \times 2\) matrix. Thus

\[
\sigma \cdot a = \sum_k a_k \sigma_k
\]

\[
= \begin{pmatrix} + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & -a_3 \end{pmatrix}.
\]

(3.2.38)

There is also a very important identity,

\[
(\sigma \cdot a)(\sigma \cdot b) = a \cdot b + i \sigma \cdot (a \times b).
\]

(3.2.39)

To prove this all we need are the anticommutation and commutation relations, (3.2.34) and (3.2.35), respectively:

\[
\sum_j \sigma_j a_j \sum_k \sigma_k b_k = \sum_k \frac{1}{2} \left\{ \sigma_j, \sigma_k \right\} a_j b_k + \frac{1}{2} \left[ \sigma_j, \sigma_k \right] a_j b_k
\]

\[
= \sum_k \left( \delta_{jk} + i \epsilon_{jkl} a_l \right) a_j b_k
\]

\[
= a \cdot b + i \sigma \cdot (a \times b).
\]

(3.2.40)

If the components of \(a\) are real, we have

\[
(\sigma \cdot a)^2 = |a|^2,
\]

(3.2.41)

where \(|a|\) is the magnitude of the vector \(a\).

**Rotations in the Two-Component Formalism**

Let us now study the \(2 \times 2\) matrix representation of the rotation operator \(\mathcal{D}(\mathbf{h}, \phi)\). We have

\[
\exp \left( \frac{-i \mathbf{S} \cdot \mathbf{b} \phi}{\hbar} \right) = \exp \left( -i \sigma \cdot \mathbf{h} \phi \right).
\]

(3.2.42)
Using
\[
(\sigma \cdot \hat{n})^n = \begin{cases} 
1 & \text{for } n \text{ even,} \\
\sigma \cdot \hat{n} & \text{for } n \text{ odd,}
\end{cases}
\] (3.2.43)

which follows from (3.2.41), we can write
\[
\exp\left(-\frac{i \sigma \cdot \hat{n} \phi}{2}\right) = 1 - \frac{(\sigma \cdot \hat{n})^2}{2!} \left(\frac{\phi}{2}\right)^2 + \frac{(\sigma \cdot \hat{n})^4}{4!} \left(\frac{\phi}{2}\right)^4 - \ldots
\]
\[
- i \left[\frac{\sigma \cdot \hat{n}}{3} \left(\frac{\phi}{2}\right)^3 + \ldots\right] \exp\left(-\frac{i \sigma \cdot \hat{n} \phi}{2}\right) = 1 \cos\left(\frac{\phi}{2}\right) - i \sigma \cdot \hat{n} \sin\left(\frac{\phi}{2}\right).
\] (3.2.44)

Explicitly, in 2×2 form have we
\[
\exp\left(-\frac{i \sigma \cdot \hat{n} \phi}{2}\right) = \begin{pmatrix}
\cos\left(\frac{\phi}{2}\right) - i n_x \sin\left(\frac{\phi}{2}\right) & (-i n_x - n_y) \sin\left(\frac{\phi}{2}\right) \\
(-i n_x + n_y) \sin\left(\frac{\phi}{2}\right) & \cos\left(\frac{\phi}{2}\right) + i n_x \sin\left(\frac{\phi}{2}\right)
\end{pmatrix}
\] (3.2.45)

Just as the operator \(\exp(-i S \cdot \hat{n} \phi / \hbar)\) acts on a state ket \(|\alpha\rangle\), the 2×2 matrix \(\exp(-i \sigma \cdot \hat{n} \phi / 2)\) acts on a two-component spinor \(\chi\). Under rotations we change \(\chi\) as follows:
\[
\chi \rightarrow \exp\left(-\frac{i \sigma \cdot \hat{n} \phi}{2}\right) \chi.
\] (3.2.46)

On the other hand, the \(\sigma_k\)'s themselves are to remain \textit{unchanged} under rotations. So strictly speaking, despite its appearance, \(\sigma\) is not to be regarded as a vector; rather, it is \(\chi^\dagger \sigma \chi\) which obeys the transformation property of a vector:
\[
\chi^\dagger \sigma_k \chi \rightarrow \sum_i R_{ki} \chi^\dagger \sigma_i \chi.
\] (3.2.47)

An explicit proof of this may be given using
\[
\exp\left(-\frac{i \sigma \cdot \phi}{2}\right) \sigma_i \exp\left(-\frac{i \sigma \cdot \phi}{2}\right) = \sigma_1 \cos \phi - \sigma_2 \sin \phi
\] (3.2.48)

and so on, which is the 2×2 matrix analogue of (3.2.6).
pin state is then obtained; see Figure 3.3. In the Pauli spinor language this sequence of operations is equivalent to applying \( \exp(-i\alpha \beta /2) \) to \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) followed by an application of \( \exp(-i\alpha \alpha /2) \). The net result is

\[
\begin{pmatrix}
\cos(\frac{\alpha}{2}) - i\sin(\frac{\alpha}{2}) & 0 \\
0 & \cos(\frac{\alpha}{2}) + i\sin(\frac{\alpha}{2})
\end{pmatrix}
\begin{pmatrix}
\cos(\frac{\beta}{2}) & -\sin(\frac{\beta}{2}) \\
\sin(\frac{\beta}{2}) & \cos(\frac{\beta}{2})
\end{pmatrix}
\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix}
\cos(\frac{\beta}{2}) e^{-i\alpha/2} \\
\sin(\frac{\beta}{2}) e^{i\alpha/2}
\end{pmatrix}
\]

(3.2.52)
in complete agreement with Problem 9 of Chapter 1 if we realize that a phase common to both the upper and lower components is devoid of physical significance.

### 3.3. SO(3), SU(2), and Euler Rotations

**Orthogonal Group**

We will now study a little more systematically the group properties of the operations with which we have been concerned in the previous two sections.

The most elementary approach to rotations is based on specifying the axis of rotation and the angle of rotation. It is clear that we need three real numbers to characterize a general rotation: the polar and the azimuthal angles of the unit vector \( \hat{n} \) taken in the direction of the rotation axis and the rotation angle \( \phi \) itself. Equivalently, the same rotation can be specified by the three Cartesian components of the vector \( \hat{n}\phi \). However, these ways of characterizing rotation are not so convenient from the point of view of studying the group properties of rotations. For one thing, unless \( \phi \) is infinitesimal or \( \hat{n} \) is always in the same direction, we cannot add vectors of the form \( \hat{n}\phi \) to characterize a succession of rotations. It is much easier to work with a \( 3 \times 3 \) orthogonal matrix \( R \) because the effect of successive rotations can be obtained just by multiplying the appropriate orthogonal matrices.

How many independent parameters are there in a \( 3 \times 3 \) orthogonal matrix? A real \( 3 \times 3 \) matrix has 9 entries, but we have the orthogonality constraint

\[
RR^T = 1.
\]

(3.3.1)

which corresponds to 6 independent equations because the product \( RR^T \), being the same as \( R^T R \), is a symmetrical matrix with 6 independent entries. As a result, there are 3 (that is, \( 9-6 \)) independent numbers in \( R \), the same number previously obtained by a more elementary method.

The set of all multiplication operations with orthogonal matrices forms a group. By this we mean that the following four requirements are satisfied:

1. The product of any two orthogonal matrices is another orthogonal matrix, which is satisfied because

\[
(R_1 R_2)(R_1 R_2)^T = R_1 R_2 R_2^T R_1^T = 1.
\]

(3.3.2)

2. The associative law holds:

\[
R_1 (R_2 R_3) = (R_1 R_2) R_3.
\]

(3.3.3)

3. The identity matrix \( 1 \)—physically corresponding to no rotation—defined by

\[
R I = I R = R
\]

(3.3.4)
is a member of the class of all orthogonal matrices.

4. The inverse matrix \( R^{-1} \)—physically corresponding to rotation in the opposite sense—defined by

\[
R R^{-1} = R^{-1} R = 1
\]

(3.3.5)
is also a member. This group has the name SO(3), where S stands for special, O stands for orthogonal, 3 for three dimensions. Note only rotational operations are considered here, hence we have SO(3) rather than O(3) (which can include the inversion operation of Chapter 4 later).

**Unitary unimodular group**

In the previous section we learned yet another way to characterize an arbitrary rotation—that is, to look at the \( 2 \times 2 \) matrix (3.2.45) that acts on the two-component spinor \( \chi \). Clearly, (3.2.45) is unitary. As a result, for the \( c_+ \) and \( c_- \), defined in (3.2.28),

\[
|c_+|^2 + |c_-|^2 = 1
\]

(3.3.6)
is left invariant. Furthermore, matrix (3.3.45) is unimodular; that is, its determinant is 1, as will be shown explicitly below.

We can write the most general unitary unimodular matrix as

\[
U(a, b) = \begin{pmatrix}
a & b \\
-b^* & a^*
\end{pmatrix}
\]

(3.3.7)

where \( a \) and \( b \) are complex numbers satisfying the unimodular condition

\[
|a|^2 + |b|^2 = 1.
\]

(3.3.8)
We can easily establish the unitary property of (3.3.7) as follows:

\[
U(a, b)^T U(a, b) = \left( \begin{array}{cc}
a^* & -b \\
-b^* & a \\
\end{array} \right) \left( \begin{array}{cc}
a & b \\
b^* & a^* \\
\end{array} \right) = 1,
\]

where we have used (3.3.8). Notice that the number of independent real parameters in (3.3.7) is again three.

We can readily see that the 2x2 matrix (3.2.45) that characterizes a rotation of a spin \( \frac{1}{2} \) system can be written as \( U(a, b) \). Comparing (3.2.45) with (3.3.7), we identify

\[
\begin{align*}
\text{Re}(a) &= \cos \left( \frac{\phi}{2} \right), \\
\text{Im}(a) &= -n\sin \left( \frac{\phi}{2} \right), \\
\text{Re}(b) &= -n\sin \left( \frac{\phi}{2} \right), \\
\text{Im}(b) &= -n\sin \left( \frac{\phi}{2} \right),
\end{align*}
\]

(3.3.10)

from which the unimodular property of (3.3.8) is immediate. Conversely, it is clear that the most general unitary unimodular matrix of form (3.3.7) can be interpreted as representing a rotation.

The two complex numbers \( a \) and \( b \) are known as Cayley-Klein parameters. Historically the connection between a unitary unimodular matrix and a rotation was known long before the birth of quantum mechanics. In fact, the Cayley-Klein parameters were used to characterize complicated motions of gyroscopes in rigid-body kinematics.

Without appealing to the interpretations of unitary unimodular matrices in terms of rotations, we can directly check the group properties of multiplication operations with unitary unimodular matrices. Note in particular that

\[
U(a_1, b_1)U(a_2, b_2) = U(a_1a_2 - b_1b_2^*, a_1b_2 + a_2^*b_1),
\]

(3.3.11)

where the unimodular condition for the product matrix is

\[
|a_1a_2 - b_1b_2^*|^2 + |a_1b_2 + a_2^*b_1|^2 = 1.
\]

(3.3.12)

For the inverse of \( U \) we have

\[
U^{-1}(a, b) = U(a^*, -b).
\]

(3.3.13)

This group is known as SU(2), where \( S \) stands for special, \( U \) for unitary, and 2 for dimensionality 2. In contrast, the group defined by multiplication operations with general 2x2 unitary matrices (not necessarily constrained to be unimodular) is known as U(2). The most general unitary matrix in two dimensions has four independent parameters and can be written as \( e^{i\gamma} \) (with \( \gamma \) real) times a unitary unimodular matrix:

\[
U = e^{i\gamma} \left( \begin{array}{cc}
a & b \\
-b^* & a^* \\
\end{array} \right),
\]

(3.3.14)

\( SU(2) \) is called a subgroup of \( U(2) \).

Because we can characterize rotations using both the \( SO(3) \) language and the \( SU(2) \) language, we may be tempted to conclude that the groups \( SO(3) \) and \( SU(2) \) are isomorphic—that is, that there is a one-to-one correspondence between an element of \( SO(3) \) and an element of \( SU(2) \). This inference is not correct. Consider a rotation by \( 2\pi \) and another one by \( 4\pi \). In the \( SO(3) \) language, the matrices representing a \( 2\pi \) rotation and a \( 4\pi \) rotation are both \( 3x3 \) identity matrices; however, in the \( SU(2) \) language the corresponding matrices are \(-1\) times the \( 2x2 \) identity matrix and the identity matrix itself, respectively. More generally, \( U(a, b) \) and \( U(-a, -b) \) both correspond to a single \( 3x3 \) matrix in the \( SO(3) \) language. The correspondence therefore is two-to-one; for a given \( R \), the corresponding \( U \) is double-valued. One can say, however, that the two groups are locally isomorphic.

**Euler Rotations**

From classical mechanics the reader may be familiar with the fact that an arbitrary rotation of a rigid body can be accomplished in three steps, known as Euler rotations. The Euler rotation language, specified by three Euler angles, provides yet another way to characterize the most general rotation in three dimensions.

The three steps of Euler rotations are as follows. First, rotate the rigid body counterclockwise (as seen from the positive \( z \)-side) about the \( z \)-axis by angle \( \alpha \). Imagine now that there is a body \( y \)-axis embedded, so to speak, in the rigid body such that before the \( z \)-axis rotation is carried out, the body \( y \)-axis coincides with the usual \( y \)-axis, referred to as the space-fixed \( y \)-axis. Obviously, after the rotation about the \( z \)-axis, the body \( y \)-axis no longer coincides with the space-fixed \( y \)-axis; let us call the former the \( y' \)-axis. To see how all this may appear for a thin disk, refer to Figure 3.4a. We now perform a second rotation, this time about the \( y' \)-axis by angle \( \beta \). As a result, the body \( z \)-axis no longer points in the space-fixed \( z \)-axis direction. We call the body-fixed \( z \)-axis after the second rotation the \( z' \)-axis; see Figure 3.4b. The third and final rotation is about the \( z' \)-axis by angle \( \gamma \). The body \( y \)-axis now becomes the \( y'' \)-axis of Figure 3.4c. In terms of \( 3 \times 3 \) orthogonal matrices the product of the three operations can be written as

\[
R(\alpha, \beta, \gamma) = R_{y''}(\gamma)R_{y'}(\beta)R_{y}(\alpha).
\]

(3.3.15)

A cautionary remark is in order here. Most textbooks in classical mechanics prefer to perform the second rotation (the middle rotation) about the body \( x \)-axis rather than about the body \( y \)-axis (see, for example, Goldstein 1980). This convention is to be avoided in quantum mechanics for a reason that will become apparent in a moment.

In (3.3.15) there appear \( R_{y''} \) and \( R_{y'} \), which are matrices for rotations about body axes. This approach to Euler rotations is rather inconvenient in
To prove this assertion, let us look more closely at the effect of both sides of (3.3.16) on the circular disc of Figure 3.4a. Clearly, the orientation of the body $y$-axis is unchanged in both cases, namely, in the $y'$-direction. Furthermore, the orientation of the final body $z$-axis is the same whether we apply $R_y(\beta)$ or $R_z(\alpha)R_y(\beta)R_z^{-1}(\alpha)$. In both cases the final body $z$-axis makes a polar angle $\beta$ with the fixed $z$-axis (the same as the initial $z$-axis), and its azimuthal angle, as measured in the fixed-coordinate system, is just $\alpha$. In other words, the final body $z$-axis is the same as the $z'$-axis of Figure 3.4b. Similarly, we can prove

$$R_z(\gamma) = R_y(\beta)R_z(\gamma)R_y^{-1}(\beta). \quad (3.3.17)$$

Using (3.3.16) and (3.3.17), we can now rewrite (3.3.15). We obtain

$$R_z(\gamma)R_y(\beta)R_z(\alpha) = R_y(\beta)R_z(\gamma)R_y^{-1}(\beta)R_y(\beta)R_z(\alpha)$$

$$= R_z(\alpha)R_y(\beta)R_z^{-1}(\alpha)R_y(\gamma)R_z(\alpha)$$

$$= R_z(\alpha)R_y(\beta)R_z(\gamma), \quad (3.3.18)$$

where in the final step we used the fact that $R_z(\gamma)$ and $R_z(\alpha)$ commute. To summarize,

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma), \quad (3.3.19)$$

where all three matrices on the right-hand side refer to fixed-axis rotations.

Now let us apply this set of operations to spin $\frac{1}{2}$ systems in quantum mechanics. Corresponding to the product of orthogonal matrices in (3.3.19) there exists a product of rotation operators in the ket space of the spin $\frac{1}{2}$ system under consideration:

$$\mathcal{D}(\alpha, \beta, \gamma) = \mathcal{D}_z(\alpha)\mathcal{D}_y(\beta)\mathcal{D}_z(\gamma). \quad (3.3.20)$$

The $2 \times 2$ matrix representation of this product is

$$\exp\left(-\frac{i\alpha}{2}\right)\exp\left(-\frac{i\beta}{2}\right)\exp\left(-\frac{i\gamma}{2}\right)$$

$$= \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{pmatrix} \begin{pmatrix} \cos(\beta/2) & -\sin(\beta/2) \\ \sin(\beta/2) & \cos(\beta/2) \end{pmatrix} \begin{pmatrix} e^{-i\gamma/2} & 0 \\ 0 & e^{i\gamma/2} \end{pmatrix}$$

$$= \begin{pmatrix} e^{-i(a+\gamma)/2}\cos(\beta/2) - e^{-i(a-\gamma)/2}\sin(\beta/2) \\ e^{i(a+\gamma)/2}\sin(\beta/2) + e^{i(a-\gamma)/2}\cos(\beta/2) \end{pmatrix}, \quad (3.3.21)$$

where (3.2.44) was used. This matrix is clearly of the unitary unimodular form. Conversely, the most general $2 \times 2$ unitary unimodular matrix can be written in this Euler angle form.

Notice that the matrix elements of the second (middle) rotation $\exp(-i\phi/2)$ are purely real. This would not have been the case had we chosen to rotate about the $x$-axis rather than the $y$-axis, as done in most

![Figure 3.4. Euler rotations.](image-url)
textbooks in classical mechanics. In quantum mechanics it pays to stick to our convention because we prefer the matrix elements of the second rotation, which is the only rotation matrix containing off-diagonal elements, to be purely real."

The 2 × 2 matrix in (3.3.21) is called the \( j = \frac{1}{2} \) irreducible representation of the rotation operator \( D(\alpha, \beta, \gamma) \). Its matrix elements are denoted by \( D^{(1/2)}(\alpha, \beta, \gamma) \). In terms of the angular-momentum operators we have

\[
D^{(1/2)}(\alpha, \beta, \gamma) = \left( \begin{array}{cc} 1 & \frac{1}{2} \sin \beta \exp \left( -i \frac{1}{2} \alpha \right) \exp \left( -i \frac{1}{2} \gamma \right) \end{array} \right) \times \left( \begin{array}{c} \frac{1}{2} \sin \beta \exp \left( -i \frac{1}{2} \alpha \right) \exp \left( -i \frac{1}{2} \gamma \right) \end{array} \right) \text{ for } j = \frac{1}{2}, m. \tag{3.3.22}
\]

In Section 3.5 we will extensively study higher \( j \)-analogues of (3.3.21).

3.4. DENSITY OPERATORS AND PURE VERSUS MIXED ENSEMBLES

Polarized Versus Unpolarized Beams

The formalism of quantum mechanics developed so far makes statistical predictions on an ensemble, that is, a collection, of identically prepared physical systems. More precisely, in such an ensemble all members are supposed to be characterized by the same state ket \( |\alpha\rangle \). A good example of this is a beam of silver atoms coming out of an SG filtering apparatus. Every atom in the beam has its spin pointing in the same direction, namely, the direction determined by the inhomogeneity of the magnetic field of the filtering apparatus. We have not yet discussed how to describe quantum mechanically an ensemble of physical systems for which some, say 60%, are characterized by \( |\alpha\rangle \), and the remaining 40% are characterized by some other ket \( |\beta\rangle \).

To illustrate vividly the incompleteness of the formalism developed so far, let us consider silver atoms coming directly out of a hot oven, yet to be subjected to a filtering apparatus of the Stern-Gerlach type. On symmetry grounds we expect that such atoms have random spin orientations; in other words, there should be no preferred direction associated with such an ensemble of atoms. According to the formalism developed so far, the most general state ket of a spin \( \frac{1}{2} \) system is given by

\[
|\alpha\rangle = c_+ |+\rangle + c_- |-\rangle. \tag{3.4.1}
\]

\( \text{This, of course, depends on our convention that the matrix elements of } S_i \text{ (or, more generally, } J_i \text{) are taken to be purely imaginary.} \)

Is this equation capable of describing a collection of atoms with random spin orientations? The answer is clearly no; (3.4.1) characterizes a state ket whose spin is pointing in some definite direction, namely, in the direction of \( \mathbf{h} \), whose polar and azimuthal angles, \( \beta \) and \( \alpha \), respectively, are obtained by solving

\[
\frac{c_+}{c_-} = \frac{\cos(\beta/2)}{e^{i\alpha}\sin(\beta/2)}; \tag{3.4.2}
\]

and see (3.2.52).

To cope with a situation of this kind we introduce the concept of fractional population, or probability weight. An ensemble of silver atoms with completely random spin orientation can be viewed as a collection of silver atoms in which 50% of the members of the ensemble are characterized by \(|+\rangle\) and the remaining 50% by \(|-\rangle\). We specify such an ensemble by assigning

\[
w_+ = 0.5, \quad w_- = 0.5, \tag{3.4.3}
\]

where \( w_+ \) and \( w_- \) are the fractional population for spin-up and -down, respectively. Because there is no preferred direction for such a beam, it is reasonable to expect that this same ensemble can be regarded equally well as a 50-50 mixture of \(|S_z; +\rangle\) and \(|S_z; -\rangle\). The mathematical formalism needed to accomplish this will appear shortly.

It is very important to note that we are simply introducing here two real numbers \( w_+ \) and \( w_- \). There is no information on the relative phase between the spin-up and the spin-down ket. Quite often we refer to such a situation as an incoherent mixture of spin-up and spin-down states. What we are doing here is to be clearly distinguished from what we did with a coherent linear superposition, for example,

\[
\left( \frac{1}{\sqrt{2}} \right) |+\rangle + \left( \frac{1}{\sqrt{2}} \right) |-\rangle, \tag{3.4.4}
\]

where the phase relation between \(|+\rangle\) and \(|-\rangle\) contains vital information on the spin orientation in the xy-plane, in this case in the positive x-direction. In general, we should not confuse \( w_+ \) and \( w_- \) with \( |c_+|^2 \) and \( |c_-|^2 \). The probability concept associated with \( w_+ \) and \( w_- \) is much closer to that encountered in classical probability theory. The situation encountered in dealing with silver atoms directly from the hot oven may be compared with that of a graduating class in which 50% of the graduating seniors are male, the remaining 50% female. When we pick a student at random, the probability that the particular student is male (or female) is 0.5. Whoever heard of a student referred to as a coherent linear superposition of male and female with a particular phase relation?

The beam of silver atoms coming directly out of the oven is an example of a completely random ensemble; the beam is said to be un-
polarized because there is no preferred direction for spin orientation. In contrast, the beam that has gone through a selective Stern-Gerlach–type measurement is an example of a pure ensemble; the beam is said to be polarized because all members of the ensemble are characterized by a single common ket that describes a state with spin pointing in some definite direction. To appreciate the difference between a completely random ensemble and a pure ensemble, let us consider a rotatable SG apparatus where we can vary the direction of the inhomogeneous B just by rotating the apparatus. When a completely unpolarized beam directly out of the oven is subjected to such an apparatus, we always obtain two emerging beams of equal intensity no matter what the orientation of the apparatus may be. In contrast, if a polarized beam is subjected to such an apparatus, the relative intensities of the two emerging beams vary as the apparatus is rotated. For some particular orientation the ratio of the intensities actually becomes one to zero. In fact, the formalism we developed in Chapter 1 tells us that the relative intensities are simply \( \cos^2(\beta/2) \) and \( \sin^2(\beta/2) \), where \( \beta \) is the angle between the spin direction of the atoms and the direction of the inhomogeneous magnetic field in the SG apparatus.

A complete random ensemble and a pure ensemble can be regarded as the extremes of what is known as a mixed ensemble. In a mixed ensemble a certain fraction—for example, 70%—of the members are characterized by a state ket \( |\alpha\rangle \), the remaining 30% by \( |\beta\rangle \). In such a case the beam is said to be partially polarized. Here \( |\alpha\rangle \) and \( |\beta\rangle \) need not even be orthogonal; we can, for example, have 70% with spin in the positive x-direction and 30% with spin in the negative z-direction.*

**Ensemble Averages and Density Operator**

We now present the density operator formalism, pioneered by J. von Neumann in 1927, that quantitatively describes physical situations with mixed as well as pure ensembles. Our general discussion here is not restricted to spin \( \frac{1}{2} \) systems, but for illustrative purposes we return repeatedly to spin \( \frac{1}{2} \) systems.

A pure ensemble by definition is a collection of physical systems such that every member is characterized by the same ket \( |\alpha\rangle \). In contrast, in a mixed ensemble, a fraction of the members with relative population \( w_1 \) are characterized by \( |\alpha^{(1)}\rangle \), some other fraction with relative population \( w_2 \), by \( |\alpha^{(2)}\rangle \), and so on. Roughly speaking, a mixed ensemble can be viewed as a mixture of pure ensembles, just as the name suggests. The fractional populations are constrained to satisfy the normalization condition

\[
\sum_i w_i = 1. \tag{3.4.5}
\]

As we mentioned previously, \( |\alpha^{(1)}\rangle \) and \( |\alpha^{(2)}\rangle \) need not be orthogonal. Furthermore, the number of terms in the \( i \) sum of (3.4.5) need not coincide with the dimensionality \( N \) of the ket space; it can easily exceed \( N \). For example, for spin \( \frac{1}{2} \) systems with \( N = 2 \), we may consider 40% with spin in the positive \( z \)-direction, 30% with spin in the positive \( x \)-direction, and the remaining 30% with spin in the negative \( y \)-direction.

Suppose we make a measurement on a mixed ensemble of some observable \( A \). We may ask what is the average measured value of \( A \) when a large number of measurements are carried out. The answer is given by the ensemble average of \( A \), which is defined by

\[
[A] = \sum_i w_i \langle \alpha^{(i)} | A | \alpha^{(i)} \rangle = \sum_i \sum_{\alpha'} w_i \langle \alpha' | \alpha^{(i)} \rangle^2 \langle \alpha' \rangle, \tag{3.4.6}
\]

where \( |\alpha'\rangle \) is an eigenket of \( A \). Recall that \( \langle \alpha^{(i)} | A | \alpha^{(i)} \rangle \) is the usual quantum mechanical expectation value of \( A \) taken with respect to state \( |\alpha^{(i)}\rangle \). Equation (3.4.6) tells us that these expectation values must further be weighted by the corresponding fractional populations \( w_i \). Notice how probabilistic concepts enter twice: first in \( \langle \alpha' | \alpha^{(i)} \rangle^2 \) for the quantum-mechanical probability for state \( |\alpha^{(i)}\rangle \) to be found in an \( A \) eigenstate \( |\alpha'\rangle \); second, in the probability factor \( w_i \) for finding in the ensemble a quantum-mechanical state characterized by \( |\alpha^{(i)}\rangle \).*

We can now rewrite ensemble average (3.4.6) using a more general basis, \( \{|b'\rangle\} \):

\[
[A] = \sum_i w_i \sum_{b', b''} \langle \alpha^{(i)} | b' \rangle \langle b' | A | b'' \rangle \langle b'' | \alpha^{(i)} \rangle
= \sum_i \sum_{b', b''} \left( \sum_{\alpha'} w_i \langle \alpha' \rangle^2 \langle \alpha' | b'' \rangle \right) \langle b' \rangle \langle b'' \rangle. \tag{3.4.7}
\]

The number of terms in the sum of the \( b' \langle b'' \rangle \) is just the dimensionality of the ket space, while the number of terms in the sum of the \( i \) depends on how the mixed ensemble is viewed as a mixture of pure ensembles. Notice that in this form the basic property of the ensemble which does not depend on the particular observable \( A \) is factored out. This motivates us to define the density operator \( \rho \) as follows:

\[
\rho = \sum_i w_i |\alpha^{(i)}\rangle \langle \alpha^{(i)}|. \tag{3.4.8}
\]

*Quite often in the literature the ensemble average is also called the expectation value. However, in this book, the term expectation value is reserved for the average measured value when measurements are carried on a pure ensemble.
The elements of the corresponding density matrix have the following form:

\[ \langle b''|\rho|b'\rangle = \sum_i w_i \langle b''|\alpha^{(i)}\rangle \langle \alpha^{(i)}|b'\rangle. \]  

(3.4.9)

The density operator contains all the physically significant information we can possibly obtain about the ensemble in question. Returning to (3.4.7), we see that the ensemble average can be written as

\[ [A] = \sum_{b''} \sum_{b'} \langle b''|\rho|b'\rangle \langle b'|A|b''\rangle \]

\[ = \text{tr}(\rho A). \]  

(3.4.10)

Because the trace is independent of representations, \( \text{tr}(\rho A) \) can be evaluated using any convenient basis. As a result, (3.4.10) is an extremely powerful relation.

There are two properties of the density operator worth recording. First, the density operator is Hermitian, as is evident from (3.4.8). Second, the density operator satisfies the normalization condition

\[ \text{tr}(\rho) = \sum_i \sum_{b'} w_i \langle b'|\alpha^{(i)}\rangle \langle \alpha^{(i)}|b'\rangle \]

\[ = \sum_i w_i \langle \alpha^{(i)}|\alpha^{(i)}\rangle \]

\[ = 1. \]  

(3.4.11)

Because of the Hermiticity and the normalization condition, for spin \( \frac{1}{2} \) systems with dimensionality 2 the density operator, or the corresponding density matrix, is characterized by three independent real parameters. Four real numbers characterize a \( 2 \times 2 \) Hermitian matrix. However, only three are independent because of the normalization condition. The three numbers needed are \([S_z], [S_\uparrow], \text{and } [S_z] \); the reader may verify that knowledge of these three ensemble averages is sufficient to reconstruct the density operator. The manner in which a mixed ensemble is formed can be rather involved. We may mix pure ensembles characterized by all kinds of \( |\alpha^{(i)}\rangle \)'s with appropriate \( w_i \)'s; yet for spin \( \frac{1}{2} \) systems three real numbers completely characterize the ensemble in question. This strongly suggests that a mixed ensemble can be decomposed into pure ensembles in many different ways. A problem to illustrate this point appears at the end of this chapter.

A pure ensemble is specified by \( w_i = 1 \) for some \( |\alpha^{(i)}\rangle \) — with \( i = n \), for instance—and \( w_i = 0 \) for all other conceivable state kets, so the corresponding density operator is written as

\[ \rho = |\alpha^{(n)}\rangle \langle \alpha^{(n)}|. \]  

(3.4.12)

with no summation. Clearly, the density operator for a pure ensemble is idempotent, that is,

\[ \rho^2 = \rho \]  

(3.4.13)

or, equivalently,

\[ \rho(\rho - 1) = 0. \]  

(3.4.14)

Thus, for a pure ensemble only, we have

\[ \text{tr}(\rho^2) = 1. \]  

(3.4.15)

in addition to (3.4.11). The eigenvalues of the density operator for a pure ensemble are zero or one, as can be seen by inserting a complete set of base kets that diagonalize the Hermitian operator \( \rho \) between \( \rho \) and \( (\rho - 1) \) of (3.4.14). When diagonalized, the density matrix for a pure ensemble must therefore look like

\[ \rho \cong \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \]  

(diagonal form)

(3.4.16)

It can be shown that \( \text{tr}(\rho^2) \) is maximal when the ensemble is pure; for a mixed ensemble \( \text{tr}(\rho^2) \) is a positive number less than one.

Given a density operator, let us see how we can construct the corresponding density matrix in some specified basis. To this end we first recall that

\[ |\alpha\rangle \langle \alpha| = \sum_{b''} \sum_{b'} \langle b'|\alpha\rangle \langle \alpha|b''\rangle \langle b''|b'\rangle \]  

(3.4.17)

this shows that we can form the square matrix corresponding to \( |\alpha^{(i)}\rangle \langle \alpha^{(i)}| \) by combining, in the sense of outer product, the column matrix formed by \( \langle b'|\alpha^{(i)}\rangle \) with the row matrix formed by \( \langle \alpha^{(i)}|b''\rangle \), which, of course, is equal to \( \langle b''|\alpha^{(i)}\rangle^* \). The final step is to sum such square matrices with weighting factors \( w_i \), as indicated in (3.4.8). The final form agrees with (3.4.9), as expected.

It is instructive to study several examples, all referring to spin \( \frac{1}{2} \) systems.

**Example 1.** A completely polarized beam with \( S_z + \).

\[ \rho = |+\rangle \langle +| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \]  

(3.4.18)
3.4. Density Operators and Pure Versus Mixed Ensembles

The corresponding $\rho$ can be represented by

$$\rho = \frac{1}{4} \left( \begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right) + \frac{1}{4} \left( \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right)$$

from which follows

$$[S_x] = \frac{\hbar}{8}, \quad [S_y] = 0, \quad [S_z] = \frac{3\hbar}{8}. \quad (3.4.26)$$

We leave as an exercise for the reader the task of showing that this ensemble can be decomposed in ways other than (3.4.24).

Time Evolution of Ensembles

How does the density operator $\rho$ change as a function of time? Let us suppose that at some time $t_0$ the density operator is given by

$$\rho(t_0) = \sum_i w_i |\alpha^{(i)}\rangle \langle \alpha^{(i)}|.$$

If the ensemble is to be left undisturbed, we cannot change the fractional population $w_i$. So the change in $\rho$ is governed solely by the time evolution of state ket $|\alpha^{(i)}\rangle$:

$$|\alpha^{(i)}\rangle \quad \text{at} \quad t_0 \rightarrow |\alpha^{(i)}, t_0; t\rangle. \quad (3.4.28)$$

From the fact that $|\alpha^{(i)}, t_0; t\rangle$ satisfies the Schrödinger equation we obtain

$$i\hbar \frac{\partial}{\partial t} \rho = \sum_i w_i \left( H |\alpha^{(i)}, t_0; t\rangle \langle \alpha^{(i)}, t_0; t| - |\alpha^{(i)}, t_0; t\rangle \langle \alpha^{(i)}, t_0; t| H \right)$$

$$= -[\rho, H]. \quad (3.4.29)$$

This looks like the Heisenberg equation of motion except that the sign is wrong! This is not disturbing because $\rho$ is not a dynamic observable in the Heisenberg picture. On the contrary, $\rho$ is built up of Schrödinger-picture state kets and state bras which evolve in time according to the Schrödinger equation.

It is amusing that (3.4.29) can be regarded as the quantum-mechanical analogue of Liouville's theorem in classical statistical mechanics,

$$\frac{\partial \rho_{\text{classical}}}{\partial t} = -[\rho_{\text{classical}}, H]_{\text{classical}}, \quad (3.4.30)$$

where $\rho_{\text{classical}}$ stands for the density of representative points in phase space.* Thus the name density operator for the $\rho$ appearing in (3.4.29) is

Example 2. A completely polarized beam with $S_z = \pm$:

$$\rho = |S_z; \pm \rangle \langle S_z; \pm| = \frac{1}{\sqrt{2}} (|+\rangle \langle +| + |\pm\rangle \langle \pm|) = \frac{1}{\sqrt{2}} \pm \frac{1}{\sqrt{2}} \pm.$$  \hspace{1cm} (3.4.19)

The ensembles of Examples 1 and 2 are both pure.

Example 3. An unpolarized beam. This can be regarded as an incoherent mixture of a spin-up ensemble and a spin-down ensemble with equal weights (50% each):

$$\rho = \left( \begin{array}{c}
\frac{1}{2} \\
0 \\
\frac{1}{2}
\end{array} \right) = \left( \begin{array}{c}
\frac{1}{2} \\
0 \\
\frac{1}{2}
\end{array} \right)$$

which is just the identity matrix divided by 2. As we remarked earlier, the same ensemble can also be regarded as an incoherent mixture of an $S_z +$ ensemble and an $S_z -$ ensemble with equal weights. It is gratifying that our formalism automatically satisfies the expectation

$$\left( \begin{array}{c}
\frac{1}{2} \\
0 \\
\frac{1}{2}
\end{array} \right) = \frac{1}{2} \left( \left( \begin{array}{c}
\frac{1}{2} \\
\frac{1}{2}
\end{array} \right) + \left( \begin{array}{c}
\frac{1}{2} \\
\frac{1}{2}
\end{array} \right) \right) = \frac{1}{2} \left( \left( \begin{array}{c}
\frac{1}{2} \\
\frac{1}{2}
\end{array} \right) + \left( \begin{array}{c}
\frac{1}{2} \\
\frac{1}{2}
\end{array} \right) \right),$$

where we see from Example 2 that the two terms on the right-hand side are the density matrices for pure ensemble with $S_z +$ and $S_z -$. Because $\rho$ in this case is just the identity operator divided by 2 (the dimensionality), we have

$$\text{tr}(\rho S_x) = \text{tr}(\rho S_y) = \text{tr}(\rho S_z) = 0, \quad (3.4.22)$$

where we used the fact that $S_k$ is traceless. Thus for the ensemble average of $S$ we have

$$[S] = 0. \quad (3.4.23)$$

This is reasonable because there should be no preferred spin direction in a completely random ensemble of spin $\frac{1}{2}$ systems.

Example 4. As an example of a partially polarized beam, let us consider a 75-25 mixture of two pure ensembles, one with $S_z +$ and the other with $S_z +$:

$$w(S_z +) = 0.75, \quad w(S_z +) = 0.25. \quad (3.4.24)$$

*Remember, a pure classical state is one represented by a single moving point in phase space $(q_1, \ldots, q_r, p_1, \ldots, p_r)$ at each instant of time. A classical statistical state, on the other hand, is described by our nonnegative density function $\rho_{\text{classical}}(q_1, \ldots, q_r, p_1, \ldots, p_r; t)$ such that the probability that a system is found in the interval $dq_1, \ldots, dp_r$ at time $t$ is $\rho_{\text{classical}}dq_1, \ldots, dp_r$.}
indeed appropriate. The classical analogue of (3.4.10) for the ensemble average of some observable \( A \) is given by

\[
\langle A \rangle_{\text{average}} = \frac{\int \rho_{\text{classical}} A(q, p) d\Gamma_{q, p}}{\int \rho_{\text{classical}} d\Gamma_{q, p}},
\]

(3.4.31)

where \( d\Gamma_{q, p} \) stands for a volume element in phase space.

**Continuum Generalizations**

So far we have considered density operators in ket space where the base kets are labeled by the discrete-eigenvalues of some observable. The concept of density matrix can be generalized to cases where the base kets used are labeled by continuous eigenvalues. In particular, let us consider the ket space spanned by the position eigenkets \( |x\rangle \). The analogue of (3.4.10) is given by

\[
[A] = \int d^3 x' \int d^3 x'' \langle x''|\rho|x'\rangle\langle x'|A|x''\rangle.
\]

(3.4.32)

The density matrix here is actually a function of \( x' \) and \( x'' \), namely,

\[
\langle x''|\rho|x'\rangle = \langle x''|\left( \sum_i w_i |\alpha^{(i)}\rangle \langle \alpha^{(i)}| \right)|x'\rangle
\]

\[
= \sum_i w_i \psi_i(x'') \psi_i^*(x'),
\]

(3.4.33)

where \( \psi_i \) is the wave function corresponding to the state ket \( |\alpha^{(i)}\rangle \). Notice that the diagonal element (that is, \( x' = x'' \)) of this is just the weighted sum of the probability densities. Once again, the term *density matrix* is indeed appropriate.

In continuum cases, too, it is important to keep in mind that the same mixed ensemble can be decomposed in different ways into pure ensembles. For instance, it is possible to regard a "realistic" beam of particles either as a mixture of plane-wave states (monoenergetic free-particle states) or as a mixture of wave-packet states.

**Quantum Statistical Mechanics**

We conclude this section with a brief discussion on the connection between the density operator formalism and statistical mechanics. Let us first record some properties of completely random and of pure ensembles.

The density matrix of a completely random ensemble looks like

\[
\rho = \frac{1}{N} \begin{bmatrix}
1 & 1 & \cdots & 1 \\
1 & 1 & & \\
\vdots & & \ddots & \vdots \\
1 & & & 1
\end{bmatrix}
\]

(3.4.34)

in any representation [compare Example 3 with (3.4.20)]. This follows from the fact that all states corresponding to the base kets with respect to which the density matrix is written are equally populated. In contrast, in the basis where \( \rho \) is diagonalized, we have (3.4.16) for the matrix representation of the density operator for a pure ensemble. The two diagonal matrices (3.4.34) and (3.4.16), both satisfying the normalization requirement (3.4.11), cannot look more different. It would be desirable if we could somehow construct a quantity that characterizes this dramatic difference.

Thus we define a quantity called \( \sigma \) by

\[
\sigma = -\text{tr}(\rho \ln \rho).
\]

(3.4.35)

The logarithm of the operator \( \rho \) may appear rather formidable, but the meaning of (3.4.35) is quite unambiguous if we use the basis in which \( \rho \) is diagonal:

\[
\sigma = -\sum_k \rho_k^{(\text{diag})} \ln \rho_k^{(\text{diag})}.
\]

(3.4.36)

Because each element \( \rho_k^{(\text{diag})} \) is a real number between 0 and 1, \( \sigma \) is necessarily positive semidefinite. For a completely random ensemble (3.4.34), we have

\[
\sigma = -\sum_{k=1}^N \frac{1}{N} \ln \left( \frac{1}{N} \right) = \ln N.
\]

(3.4.37)

In contrast, for a pure ensemble (3.4.16) we have

\[
\sigma = 0
\]

(3.4.38)

where we have used

\[
\rho_k^{(\text{diag})} = 0 \quad \text{or} \quad \ln \rho_k^{(\text{diag})} = 0
\]

(3.4.39)

for each term in (3.4.36).

We now argue that physically \( \sigma \) can be regarded as a quantitative measure of disorder. A pure ensemble is an ensemble with a maximum amount of order because all members are characterized by the same quantum-mechanical state ket; it may be likened to marching soldiers in a
well-regimented army. According to (3.4.38), $\sigma$ vanishes for such an ensemble. On the other extreme, a completely random ensemble, in which all quantum-mechanical states are equally likely, may be likened to drunken soldiers wandering around in random directions. According to (3.4.37), $\sigma$ is large; indeed, we will show later that $\ln N$ is the maximum possible value for $\sigma$ subject to the normalization condition

$$\sum_k \rho_{kk} = 1.$$  \tag{3.4.40}

In thermodynamics we learn that a quantity called entropy measures disorder. It turns out that our $\sigma$ is related to the entropy per constituent member, denoted by $S$, of the ensemble via

$$S = k\sigma,$$  \tag{3.4.41}

where $k$ is a universal constant identifiable with the Boltzmann constant. In fact, (3.4.41) may be taken as the definition of entropy in quantum statistical mechanics.

We now show how the density operator $\rho$ can be obtained for an ensemble in thermal equilibrium. The basic assumption we make is that nature tends to maximize $\sigma$ subject to the constraint that the ensemble average of the Hamiltonian has a certain prescribed value. To justify this assumption would involve us in a delicate discussion of how equilibrium is established as a result of interactions with the environment, which is beyond the scope of this book. In any case, once thermal equilibrium is established, we expect

$$\frac{\partial \rho}{\partial t} = 0.$$  \tag{3.4.42}

Because of (3.4.29), this means that $\rho$ and $H$ can be simultaneously diagonalized. So the kets used in writing (3.4.36) may be taken to be energy eigenkets. With this choice $\rho_{kk}$ stands for the fractional population for an energy eigenstate with energy eigenvalue $E_k$.

Let us maximize $\sigma$ by requiring that

$$\delta \sigma = 0.$$  \tag{3.4.43}

However, we must take into account the constraint that the ensemble average of $H$ has a certain prescribed value. In the language of statistical mechanics, $[H]$ is identified with the internal energy per constituent denoted by $U:

$$[H] = \text{tr}(\rho H) = U.$$  \tag{3.4.44}

In addition, we should not forget the normalization constraint (3.4.40). So our basic task is to require (3.4.43) subject to the constraints

$$\delta [H] = \sum_k \delta \rho_{kk} E_k = 0$$  \tag{3.4.45a}

and

$$\delta (\text{tr} \rho) = \sum_k \delta \rho_{kk} = 0.$$  \tag{3.4.45b}

We can most readily accomplish this by using Lagrange multipliers. We obtain

$$\sum_k \delta \rho_{kk} \left[ (\ln \rho_{kk} + 1) + \beta E_k + \gamma \right] = 0,$$  \tag{3.4.46}

which for an arbitrary variation is possibly only if

$$\rho_{kk} = \exp(-\beta E_k - \gamma - 1).$$  \tag{3.4.47}

The constant $\gamma$ can be eliminated using the normalization condition (3.4.40), and our final result is

$$\rho_{kk} = \frac{\exp(-\beta E_k)}{\sum \exp(-\beta E_i)},$$  \tag{3.4.48}

which directly gives the fractional population for an energy eigenstate with eigenvalue $E_k$. It is to be understood throughout that the sum is over distinct energy eigenstates; if there is degeneracy we must sum over states with the same energy eigenvalue.

The density matrix element (3.4.48) is appropriate for what is known in statistical mechanics as a canonical ensemble. Had we attempted to maximize $\sigma$ without the internal-energy constraint (3.4.45a), we would have obtained instead

$$\rho_{kk} = \frac{1}{N}, \quad \text{(independent of } k),$$  \tag{3.4.49}

which is the density matrix element appropriate for a completely random ensemble. Comparing (3.4.48) with (3.4.49), we infer that a completely random ensemble can be regarded as the $\beta \to 0$ limit (physically the high-temperature limit) of a canonical ensemble.

We recognize the denominator of (3.4.48) as the partition function

$$Z = \sum_k^N \exp(-\beta E_k)$$  \tag{3.4.50}

in statistical mechanics. It can also be written as

$$Z = \text{tr}(e^{-\beta H}).$$  \tag{3.4.51}

Knowing $\rho_{kk}$ given in the energy basis, we can write the density operator as

$$\rho = \frac{e^{-\beta H}}{Z}.$$  \tag{3.4.52}

This is the most basic equation from which everything follows. We can
immediately evaluate the ensemble average of any observable $A$:

$$ [A] = \frac{\text{tr}(e^{-\beta H_A})}{Z} = \frac{\sum_k^N \langle A \rangle_k \exp(-\beta E_k)}{\sum_k^N \exp(-\beta E_k)} .$$

In particular, for the internal energy per constituent we obtain

$$ U = \frac{\sum_k^N \langle E \rangle_k \exp(-\beta E_k)}{\sum_k^N \exp(-\beta E_k)} = -\frac{\partial}{\partial \beta} \ln(Z) ,$$

a formula well known to every student of statistical mechanics. The parameter $\beta$ is related to the temperature $T$ as follows:

$$ \beta = \frac{1}{kT} .$$

where $k$ is the Boltzmann constant. It is instructive to convince ourselves of this identification by comparing the ensemble average $[H]$ of simple harmonic oscillators with the $kT$ expected for the internal energy in the classical limit, which is left as an exercise. We have already commented that in the high-temperature limit, a canonical ensemble becomes a completely random ensemble in which all energy eigenstates are equally populated. In the opposite low-temperature limit ($\beta \rightarrow \infty$), (3.4.18) tells us that a canonical ensemble becomes a pure ensemble where only the ground state is populated.

As a simple illustrative example, consider a canonical ensemble made up of spin $\frac{1}{2}$ systems, each with a magnetic moment $e\hbar/2m_e c$ subjected to a uniform magnetic field in the $z$-direction. The Hamiltonian relevant to this problem has already been given [see (3.2.16)]. Because $H$ and $S_z$ commute, the density matrix for this canonical ensemble is diagonal in the $S_z$ basis. Thus

$$ \rho \equiv \begin{pmatrix} e^{-\beta \hbar\omega/2} & 0 \\ 0 & e^{\beta \hbar\omega/2} \end{pmatrix} ,$$

where the partition function $Z = e^{-\beta \hbar\omega/2} + e^{\beta \hbar\omega/2}$.


3.5. Eigenvalues and Eigenstates of Angular Momentum

From this we compute

$$ [S_x] = [S_y] = 0, \quad [S_z] = -\frac{\hbar}{2} \tanh\left(\frac{\beta \hbar \omega}{2}\right) .$$

The ensemble average of the magnetic moment component is just $e/m_e c$ times $[S_z]$. The paramagnetic susceptibility $\chi$ may be computed from

$$ \left(\frac{e}{m_e c}\right) [S_z] = \chi B .$$

In this way we arrive at Brillouin's formula for $\chi$:

$$ \chi = \left(\frac{e\hbar}{2m_e c B}\right) \tanh\left(\frac{\beta \hbar \omega}{2}\right) .$$

3.5. EIGENVALUES AND EIGENSTATES OF ANGULAR MOMENTUM

Up to now our discussion of angular momentum has been confined exclusively to spin $\frac{1}{2}$ systems with dimensionality $N = 2$. In this and following sections we study more-general angular momentum states. To this end we first work out the eigenvalues and eigenkets of $J^2$ and $J_z$ and derive the expressions for matrix elements of angular momentum operators, first obtained in a 1926 paper by M. Born, W. Heisenberg, and P. Jordan.

Commutation Relations and the Ladder Operators

Everything we will do follows from the angular-momentum commutation relations (3.1.20), where we may recall that $J_z$ is defined as the generator of infinitesimal rotation. The first important property we derive from the basic commutation relations is the existence of a new operator $\mathbf{J}^2$, defined by

$$ \mathbf{J}^2 \equiv J_x J_x + J_y J_y + J_z J_z ,$$

that commutes with every one of $J_k$:

$$ [\mathbf{J}^2, J_k] = 0, \quad (k = 1, 2, 3) .$$

To prove this let us look at the $k = 3$ case:

$$ [ J_x J_x + J_y J_y + J_z J_z, J_z ] = J_x [J_x, J_z] + [J_x, J_z] J_x + J_y [J_y, J_z] + [J_y, J_z] J_y + [J_z, J_z] J_z = J_z ( -i \hbar J_y ) + (-i \hbar J_y ) J_z + J_z ( i \hbar J_x ) + ( i \hbar J_x ) J_z = 0 .$$

The proofs for the cases where $k = 1$ and 2 follow by cyclic permutation ($1 \rightarrow 2 \rightarrow 3 \rightarrow 1$) of the indices. Because $J_x$, $J_y$, and $J_z$ do not commute with
each other, we can choose only one of them to be the observable to be
diagonalized simultaneously with \( J^2 \). By convention we choose \( J_\pm \) for this
purpose.

We now look for the simultaneous eigenkets of \( J^2 \) and \( J_\pm \). We denote
the eigenvalues of \( J^2 \) and \( J_\pm \) by \( a \) and \( b \), respectively:

\[
J^2 |a,b\rangle = a |a,b\rangle \quad (3.5.4a)
J_\pm |a,b\rangle = b |a,b\rangle. \quad (3.5.4b)
\]

To determine the allowed values for \( a \) and \( b \), it is convenient to work with
the non-Hermitian operators

\[
J_\pm \equiv J_x \pm i J_y, \quad (3.5.5)
\]
called the \textit{ladder operators}, rather than with \( J_x \) and \( J_y \). They satisfy the
commutation relations

\[
[J_+, J_-] = 2h J_z \quad (3.5.6a)
\]
and

\[
[J_\pm, J_\pm] = \pm h J_\pm, \quad (3.5.6b)
\]
which can easily be obtained from (3.1.20). Note also that

\[
[J^2, J_\pm] = 0, \quad (3.5.7)
\]
which is an obvious consequence of (3.5.2).

What is the physical meaning of \( J_\pm \)? To answer this we examine how
\( J_\pm \) acts on \( J_\pm |a,b\rangle \):

\[
J_\pm (J_\pm |a,b\rangle) = ([J_\pm, J_\pm] + J_\pm J_\pm) |a,b\rangle = (b \pm h)(J_\pm |a,b\rangle) \quad (3.5.8)
\]
where we have used (3.5.6b). In other words, if we apply \( J_\pm(J_-) \) to a \( J_z \)
eigenket, the resulting ket is still a \( J_z \) eigenket except that its eigenvalue is
now increased (decreased) by one unit of \( h \). So now we see why \( J_\pm \), which
step one step up (down) on the "ladder" of \( J_z \) eigenvalues, are known as the
ladder operators.

We now digress to recall that the commutation relations in (3.5.6b)
are reminiscent of some commutation relations we encountered in the earlier
chapters. In discussing the translation operator \( \mathcal{T} \) (l) we had

\[
[x_i, \mathcal{T}(l)] = l_i \mathcal{T}(l); \quad (3.5.9)
\]
also, in discussing the simple harmonic oscillator we had

\[
[N, a^\dagger] = a^\dagger, \quad [N, a] = -a. \quad (3.5.10)
\]
We see that both (3.5.9) and (3.5.10) have a structure similar to (3.5.6b). The
physical interpretation of the translation operator is that it changes the
eigenvalue of the position operator \( x \) by \( 1 \) in much the same way as the

ladder operator \( J_\pm \) changes the eigenvalue of \( J_z \) by one unit of \( h \). Likewise,
the oscillator creation operator \( a^\dagger \) increases the eigenvalue of the number
operator \( N \) by unity.

Even though \( J_\pm \) changes the eigenvalue of \( J_z \) by one unit of \( h \), it does not change the
eigenvalue of \( J^2 \):

\[
J^2 (J_\pm |a,b\rangle) = J_\pm J^2 |a,b\rangle = a(J_\pm |a,b\rangle), \quad (3.5.11)
\]
where we have used (3.5.7). To summarize, \( J_\pm |a,b\rangle \) are simultaneous
eigenkets of \( J^2 \) and \( J_\pm \) with eigenvalues \( a \) and \( b \pm h \).
We may write

\[
J_\pm |a,b\rangle = c_{\pm} |a,b \pm h\rangle, \quad (3.5.12)
\]
where the proportionality constant \( c_{\pm} \) will be determined later from the
normalization requirement of the angular-momentum eigenkets.

\section*{Eigenvalues of \( J^2 \) and \( J_\pm \)}

We now have the machinery needed to construct angular-momentum
eigenkets and to study their eigenvalue spectrum. Suppose we apply \( J_\pm \)
successively, say \( n \) times, to a simultaneous eigenket of \( J^2 \) and \( J_\pm \). We then
obtain another eigenket of \( J^2 \) and \( J_\pm \) with the \( J_z \) eigenvalue increased by \( nh \),
while its \( J^2 \) eigenvalue is unchanged. However, this process cannot go on
indefinately. It turns out that there exists an upper limit to \( b \) (the \( J_z \)
eigenvalue) for a given \( a \) (the \( J^2 \) eigenvalue):

\[
a \geq b^2. \quad (3.5.13)
\]
To prove this assertion we first note that

\[
J^2 - J_z^2 = \frac{1}{2}(J_+ J_- + J_- J_+) = \frac{1}{2}(J_+ J_+^\dagger + J_-^\dagger J_-), \quad (3.5.14)
\]
Now \( J_+ J_+^\dagger \) and \( J_-^\dagger J_- \) must have nonnegative expectation values because

\[
J_+^\dagger |a,b\rangle \leftrightarrow \langle a,b|J_+, \quad J_- |a,b\rangle \leftrightarrow \langle a,b|J_-^\dagger ; \quad (3.5.15)
\]
thus

\[
\langle a,b|(J^2 - J_z^2)|a,b\rangle \geq 0, \quad (3.5.16)
\]
which, in turn, implies (3.5.13). It therefore follows that there must be a \( b_{\text{max}} \)
such that

\[
J_+ |a,b_{\text{max}}\rangle = 0, \quad (3.5.17)
\]
Stated another way, the eigenvalue of \( b \) cannot be increased beyond \( b_{\text{max}} \).
Now (3.5.17) also implies

\[
J_- J_+ |a,b_{\text{max}}\rangle = 0. \quad (3.5.18)
\]
But
\[ J_- J_+ = J_z^2 + J_y^2 - i(J_x J_y - J_y J_x) = J_z^2 - J_y^2 - hJ_z. \]  

So
\[ (J_z^2 - J_y^2 - hJ_z)|a, b_{\text{max}}\rangle = 0. \]  

Because \(|a, b_{\text{max}}\rangle\) itself is not a null ket, this relationship is possible only if
\[ a - b_{\text{max}}^2 - b_{\text{max}} h = 0 \]  
or
\[ a = b_{\text{max}}(b_{\text{max}} + h). \]  

In a similar manner we argue from (3.5.13) that there must also exist a \(b_{\text{min}}\) such that
\[ J_- |a, b_{\text{min}}\rangle = 0. \]  

By writing \(J_+ J_-\) as
\[ J_+ J_- = J_z^2 - J_y^2 + hJ_z \]  
in analogy with (3.5.19), we conclude that
\[ a = b_{\text{min}}(b_{\text{min}} - h). \]  

By comparing (3.5.22) with (3.5.25) we infer that
\[ b_{\text{max}} = -b_{\text{min}}, \]  
with \(b_{\text{max}}\) positive, and that the allowed values of \(b\) lie within
\[ -b_{\text{max}} \leq b \leq b_{\text{max}}. \]  

Clearly, we must be able to reach \(|a, b_{\text{max}}\rangle\) by applying \(J_+\) successively to \(|a, b_{\text{min}}\rangle\) a finite number of times. We must therefore have
\[ b_{\text{max}} = b_{\text{min}} + nh, \]  
where \(n\) is some integer. As a result, we get
\[ b_{\text{max}} = \frac{nh}{2}. \]  

It is more conventional to work with \(j\), defined to be \(b_{\text{max}}/h\), instead of with \(b_{\text{max}}\) so that
\[ j = \frac{n}{2}. \]  

The maximum value of the \(J_z\) eigenvalue is \(jh\), where \(j\) is either an integer or a half-integer. Equation (3.5.22) implies that the eigenvalue of \(J_z^2\) is given by
\[ a = h^2 j(j + 1). \]  

Let us also define \(m\) so that
\[ b = mh. \]  

If \(j\) is an integer, all \(m\) values are integers; if \(j\) is a half-integer, all \(m\) values are half-integers. The allowed \(m\)-values for a given \(j\) are
\[ m = -j, -j + 1, \ldots, j - 1, j. \]  

Instead of \(|a, b\rangle\) it is more convenient to denote a simultaneous eigenket of \(J_z^2\) and \(J_z\) by \(|j, m\rangle\). The basic eigenvalue equations now read
\[ J_z^2 |j, m\rangle = j(j + 1)h^2 |j, m\rangle \]  
and
\[ J_\pm |j, m\rangle = mh |j, m\rangle, \]  
with \(j\) either an integer or a half-integer and \(m\) given by (3.5.33). It is very important to recall here that we have used only the commutation relations (3.1.20) to obtain these results. The quantization of angular momentum, manifested in (3.34), is a direct consequence of the angular-momentum commutation relations, which, in turn, follow from the properties of rotations together with the definition of \(J_z\) as the generator of rotation.

### Matrix Elements of Angular-Momentum Operators

Let us work out the matrix elements of the various angular-momentum operators. Assuming \(|j, m\rangle\) to be normalized, we obviously have from (3.5.34)
\[ \langle j', m' | J_z^2 | j, m \rangle = j(j + 1)h^2 \delta_{j'j} \delta_{m'm}. \]  

To obtain the matrix elements of \(J_\pm\), we first consider
\[ \langle j, m | J_\pm | j, m \rangle = \langle j, m | (J_z^2 - J_y^2 - hJ_z) | j, m \rangle = h^2 [j(j + 1) - m^2 - m]. \]  

Now \(J_\pm |j, m\rangle\) must be the same as \(|j, m + 1\rangle\) (normalized) up to a multiplicative constant [see (3.5.12)]. Thus
\[ J_\pm |j, m\rangle = c_{jm}^\pm |j, m + 1\rangle. \]  

Comparison with (3.5.36) leads to
\[ |c_{jm}^\pm|^2 = h^2 [j(j + 1) - m(m + 1)]. \]  

Therefore, we have the same result for \(J_-\) and \(J_+\).

### Angular-Momentum Operators

Let us now consider the angular-momentum operators. Assuming \(|j, m\rangle\) to be normalized, we obviously have from (3.5.34)
\[ \langle j', m' | J_z^2 | j, m \rangle = j(j + 1)h^2 \delta_{j'j} \delta_{m'm}. \]  

To obtain the matrix elements of \(J_\pm\), we first consider
\[ \langle j, m | J_\pm | j, m \rangle = \langle j, m | (J_z^2 - J_y^2 - hJ_z) | j, m \rangle = h^2 [j(j + 1) - m^2 - m]. \]  

Now \(J_\pm |j, m\rangle\) must be the same as \(|j, m + 1\rangle\) (normalized) up to a multiplicative constant [see (3.5.12)]. Thus
\[ J_\pm |j, m\rangle = c_{jm}^\pm |j, m + 1\rangle. \]  

Comparison with (3.5.36) leads to
\[ |c_{jm}^\pm|^2 = h^2 [j(j + 1) - m(m + 1)]. \]  

Therefore, we have the same result for \(J_-\) and \(J_+\).
we have determined \( c_{jm}^+ \) up to an arbitrary phase factor. It is customary to choose \( c_{jm}^+ \) to be real and positive by convention. So

\[
J_+ |j, m\rangle = \sqrt{(j - m)(j + m + 1)} \hbar |j, m + 1\rangle. \tag{3.5.39}
\]

Similarly, we can derive

\[
J_- |j, m\rangle = \sqrt{(j + m)(j - m + 1)} \hbar |j, m - 1\rangle. \tag{3.5.40}
\]

Finally, we determine the matrix elements of \( J_\pm \) to be

\[
\langle j', m' | J_\pm | j, m \rangle = \sqrt{(j \pm m)(j \pm m + 1)} \hbar \delta_{j,j'} \delta_{m',m \pm 1}. \tag{3.5.41}
\]

**Representations of the Rotation Operator**

Having obtained the matrix elements of \( J_+ \) and \( J_- \), we are now in a position to study the matrix elements of the rotation operator \( \mathcal{D}(R) \). If a rotation \( R \) is specified by \( \phi \) and \( \theta \), we can define its matrix elements by

\[
\mathcal{D}_{m'm}^{(j)}(R) = \langle j, m' | \exp\left( -\frac{i\mathbf{J} \cdot \mathbf{R}}{\hbar} \right) | j, m \rangle. \tag{3.5.42}
\]

These matrix elements are sometimes called Wigner functions after E. P. Wigner, who made pioneering contributions to the group-theoretical properties of rotations in quantum mechanics. Notice here that the same \( j \)-value appears in the ket and bra of (3.5.42); we need not consider matrix elements \( \mathcal{D}(R) \) between states with different \( j \)-values because they all vanish trivially. This is because \( \mathcal{D}(R) | j, m \rangle \) is still an eigenket of \( \mathbf{J}^2 \) with the same eigenvalue \( j(j+1)\hbar^2 \):

\[
\mathbf{J}^2 \mathcal{D}(R) | j, m \rangle = \mathcal{D}(R) \mathbf{J}^2 | j, m \rangle
\]

\[
= j(j+1)\hbar^2 [\mathcal{D}(R) | j, m \rangle]. \tag{3.5.43}
\]

which follows directly from the fact that \( \mathbf{J}^2 \) commutes with \( \mathbf{J}_k \) (hence with any function of \( \mathbf{J}_k \)). Simply stated, rotations cannot change the \( j \)-value, which is an eminently sensible result.

Often in the literature the \((2j+1) \times (2j+1)\) matrix formed by \( \mathcal{D}_{m'm}^{(j)}(R) \) is referred to as the \((2j+1)\)-dimensional irreducible representation of the rotation operator \( \mathcal{D}(R) \). This means that the matrix which corresponds to an arbitrary rotation operator in ket space not necessarily characterized by a single \( j \)-value can, with a suitable choice of basis, be

brought to block-diagonal form:

where each shaded square is a \((2j+1) \times (2j+1)\) square matrix formed by \( \mathcal{D}_{m'm}^{(j)} \) with some definite value of \( j \). Furthermore, each square matrix itself cannot be broken into smaller blocks

with any choice of basis.

The rotation matrices characterized by definite \( j \) form a group. First, the identity is a member because the rotation matrix corresponding to no rotation \((\phi = 0)\) is the \((2j+1) \times (2j+1)\) identity matrix. Second, the inverse is also a member; we simply reverse the rotation angle \((\phi \rightarrow -\phi)\) without changing the rotation axis \( \hat{n} \). Third, the product of any two
members is also a member; explicitly we have
\[ \sum_{m'} \mathcal{D}_{m'm}^{(j)}(R_1) \mathcal{D}_{m'm}^{(j)}(R_2) = \mathcal{D}_{m'm}^{(j)}(R_1 R_2), \]
where the product \( R_1 R_2 \) represents a single rotation. We also note that the rotation matrix is unitary because the corresponding rotation operator is unitary; explicitly we have
\[ \mathcal{D}_{m'm}^{(j)}(R^{-1}) = \mathcal{D}_{m'm}^{(j)}(R). \]
To appreciate the physical significance of the rotation matrix let us start with a state represented by \(|j, m\rangle\). We now rotate it:
\[ |j, m\rangle \rightarrow \mathcal{D}(R)|j, m\rangle. \]
Even though this rotation operation does not change \( j \), we generally obtain states with \( m \)-values other than the original \( m \). To find the amplitude for being found in \(|j, m'\rangle\), we simply expand the rotated state as follows:
\[ \mathcal{D}(R)|j, m\rangle = \sum_{m'} |j, m'\rangle \langle j, m'| \mathcal{D}(R)|j, m\rangle = \sum_{m'} |j, m'\rangle \mathcal{D}_{m'm}^{(j)}(R), \]
where, in using the completeness relation, we took advantage of the fact that \( \mathcal{D}(R) \) connects only states with the same \( j \). So the matrix element \( \mathcal{D}_{m'm}^{(j)}(R) \) is simply the amplitude for the rotated state to be found in \(|j, m'\rangle\) when the original unrotated state is given by \(|j, m\rangle\).

In Section 3.3 we saw how Euler angles may be used to characterize the most general rotation. We now consider the matrix realization of (3.3.20) for an arbitrary \( j \) (not necessarily \( \frac{1}{2} \)):
\[ \mathcal{D}_{m'm}^{(j)}(\alpha, \beta, \gamma) = \langle j, m'| \exp \left( -\frac{iJ_\beta}{\hbar} \right) \exp \left( -\frac{iJ_\gamma}{\hbar} \right) \exp \left( -\frac{iJ_\alpha}{\hbar} \right) |j, m\rangle = e^{-i(m'a + m'\gamma)} \langle j, m'| \exp \left( -\frac{iJ_\beta}{\hbar} \right) |j, m\rangle. \]
Notice that the only nontrivial part is the middle rotation about the \( y \)-axis, which mixes different \( m \)-values. It is convenient to define a new matrix \( d^{(j)}(\beta) \) as
\[ d_{m'm}^{(j)}(\beta) = \langle j, m'| \exp \left( -\frac{iJ_\beta}{\hbar} \right) |j, m\rangle. \]
Finally, let us turn to some examples. The \( j = \frac{1}{2} \) case has already been worked out in Section 3.3. See the middle matrix of (3.3.21),
\[ d^{(1/2)} = \begin{pmatrix} \cos \left( \frac{\beta}{2} \right) & -\sin \left( \frac{\beta}{2} \right) \\ \sin \left( \frac{\beta}{2} \right) & \cos \left( \frac{\beta}{2} \right) \end{pmatrix}. \]

The next simplest case is \( j = 1 \), which we consider in some detail. Clearly, we must first obtain the \( 3 \times 3 \) matrix representation of \( J_y \). Because
\[ J_y = \frac{(J_+ - J_-)}{2i}, \]
from the defining equation (3.5.5) for \( J_+ \), we can use (3.5.41) to obtain
\[ m = 1 \quad m = 0 \quad m = -1 \]
\[ J_y^{(j=1)} = \left( \frac{\hbar}{2} \right) \begin{pmatrix} 0 & -\sqrt{2}i & 0 \\ \sqrt{2}i & 0 & -\sqrt{2}i \\ 0 & \sqrt{2}i & 0 \end{pmatrix} \]

Our next task is to work out the Taylor expansion of \( \exp(-iJ_y\beta/\hbar) \). Unlike the case \( j = \frac{1}{2} \), \( [J_y^{(j=1)}, J_y^{(j=1)}] \) is independent of 1 and \( J_y^{(j=1)} \). However, it is easy to work out
\[ J_y^{(j=1)} = \frac{1}{\hbar} \]

Consequently, for \( j = 1 \) only, it is legitimate to replace
\[ \exp \left( -\frac{iJ_y\beta}{\hbar} \right) \rightarrow 1 - \left( \frac{J_y}{\hbar} \right) (1 - \cos \beta) - i \left( \frac{J_y}{\hbar} \right) \sin \beta, \]
and the reader may verify in detail. Explicitly we have
\[ d^{(1)}(\beta) = \begin{pmatrix} \left( \frac{1}{2} \right)(1 + \cos \beta) & -\left( \frac{1}{\sqrt{2}} \right) \sin \beta & \left( \frac{1}{2} \right)(1 - \cos \beta) \\ \left( \frac{1}{\sqrt{2}} \right) \sin \beta & \cos \beta & -\left( \frac{1}{\sqrt{2}} \right) \sin \beta \\ \left( \frac{1}{2} \right)(1 - \cos \beta) & \left( \frac{1}{\sqrt{2}} \right) \sin \beta & \left( \frac{1}{2} \right)(1 + \cos \beta) \end{pmatrix}. \]

Clearly, this method becomes time-consuming for large \( j \). In Section 3.6 we will learn a much easier method for obtaining \( d_{m'm}^{(j)}(\beta) \) for any \( j \).

### 3.6. Orbital Angular Momentum

We introduced the concept of angular momentum by defining it to be the generator of an infinitesimal rotation. There is another way to approach the subject of angular momentum when spin-angular momentum is zero or can be ignored. The angular momentum \( J \) for a single particle is then the same as orbital angular momentum, which is defined as
\[ \mathbf{L} = \mathbf{x} \times \mathbf{p}. \]
Orbital Angular Momentum as Rotation Generator

We first note that the orbital angular-momentum operator defined as

\[ \{ L_i, L_j \} = i \epsilon_{ijk} \hbar L_k \]  

(3.6.2)

by virtue of the commutation relations among the components of \( x \) and \( p \). This can easily be proved as follows:

\[ \{ L_x, L_y \} = \left[ y p_z - z p_y, z p_x - x p_z \right] = \left[ y p_z, z p_x \right] + \left[ z p_y, x p_z \right] = y p_z [z, p_x] + p_x [z, p_y] = \hbar \left( x p_y - y p_x \right) = \hbar L_z \]

(3.6.3)

Next we let

\[ 1 - i \left( \frac{\delta \phi}{\hbar} \right) L_z = 1 - i \left( \frac{p_y}{\hbar} \right) \left( x p_y - y p_x \right) \]

(3.6.4)

act on an arbitrary position eigenket \( |x', y', z'\rangle \) to examine whether it can be interpreted as the infinitesimal rotation operator about the \( z \)-axis by angle \( \delta \phi \). Using the fact that momentum is the generator of translation, we obtain [see (1.6.32)]

\[ \left[ 1 - i \left( \frac{\delta \phi}{\hbar} \right) L_z \right] |x', y', z\rangle = \left[ 1 - i \left( \frac{p_y}{\hbar} \right) \left( \delta \phi x \right) + i \left( \frac{p_x}{\hbar} \right) \left( \delta \phi y \right) \right] |x', y', z\rangle = |x' - y \delta \phi, y' + x \delta \phi, z'\rangle. \]

(3.6.5)

This is precisely what we expect if \( L_z \) generates an infinitesimal rotation about the \( z \)-axis. So we have demonstrated that if \( p \) generates translation, then \( L \) generates rotation.

Suppose the wave function for an arbitrary physical state of a spinless particle is given by \( \langle x', y', z | \alpha \rangle \). After an infinitesimal rotation about the \( z \)-axis is performed, the wave function for the rotated state is

\[ \langle x', y', z | L_z | x, y, z \rangle = \langle x + y \delta \phi, y - x \delta \phi, z \rangle | \alpha \rangle. \]

(3.6.6) 

It is actually more transparent to change the coordinate basis

\[ \langle x', y', z | \alpha \rangle \rightarrow \langle r, \theta, \phi | \alpha \rangle. \]

(3.6.7)

For the rotated state we have, according to (3.6.6),

\[ \langle r, \theta, \phi | L_z | r, \theta, \phi \rangle = \langle r, \theta, \phi | \delta \phi | \alpha \rangle \]

(3.6.8)

Because \( \langle r, \theta, \phi | \alpha \rangle \) is an arbitrary position eigenket, we can identify

\[ \langle x | L_z | \alpha \rangle = -i \hbar \frac{\partial}{\partial \phi} \langle x | \alpha \rangle, \]

(3.6.9)

which is a well-known result from wave mechanics. Even though this relation can also be obtained just as easily using the position representation of the momentum operator, the derivation given here emphasizes the role of \( L_z \) as the generator of rotation.

We next consider a rotation about the \( x \)-axis by angle \( \delta \phi \). In analogy with (3.6.6) we have

\[ \langle x', y', z | L_x \rangle | \alpha \rangle = \langle x + \delta \phi, y, z - y \delta \phi | \alpha \rangle = \langle x', y, z' \rangle | \alpha \rangle. \]

(3.6.10)

By expressing \( x' \), \( y' \), and \( z' \) in spherical coordinates, we can show that

\[ \langle x | L_x | \alpha \rangle = -i \hbar \left( -\sin \phi \frac{\partial}{\partial \theta} - \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \langle x | \alpha \rangle \]

(3.6.11)

Likewise,

\[ \langle x | L_y | \alpha \rangle = -i \hbar \left( \cos \phi \frac{\partial}{\partial \theta} - \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \langle x | \alpha \rangle \]

(3.6.12)

Using (3.6.11) and (3.6.12), for the ladder operator \( L_\pm \) defined as in (3.5.5), we have

\[ \langle x | L_\pm | \alpha \rangle = -i \hbar e^{\pm i \phi} \left( \pm i \frac{\partial}{\partial \theta} - \cot \theta \frac{\partial}{\partial \phi} \right) \langle x | \alpha \rangle. \]

(3.6.13)

Finally, it is possible to write \( \langle x | L^2 | \alpha \rangle \) using

\[ L^2 = L_z^2 + \left( \frac{1}{2} \right) (L_+ L_- + L_- L_+) \]

(3.6.9), and (3.6.13), as follows:

\[ \langle x | L^2 | \alpha \rangle = -\hbar^2 \left[ \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \phi} \right) \right] \langle x | \alpha \rangle. \]

(3.6.15)

Apart from \( 1/r^2 \), we recognize the differential operator that appears here to be just the angular part of the Laplacian in spherical coordinates.

It is instructive to establish this connection between the \( L^2 \) operator and the angular part of the Laplacian in another way by looking directly at the kinetic-energy operator. We first record an important operator identity,

\[ L^2 = x^2 \mathbf{p}^2 - (\mathbf{x} \cdot \mathbf{p})^2 + i\hbar \mathbf{x} \cdot \mathbf{p}, \]

(3.6.16)

where \( x^2 \) is understood to be the operator \( x \cdot x \), just as \( p^2 \) stands for the
operator \( \mathbf{p} \cdot \mathbf{p} \). The proof of this is straightforward:

\[
\mathbf{L}^2 = \sum_{ijklm} \epsilon_{ijk} x_i \epsilon_{lmk} x_l p_m
\]

\[
= \sum_{ijlm} (\delta_{ij} \delta_{jm} - \delta_{im} \delta_{jl}) x_i p_j x_l p_m
\]

\[
= \sum_{ijlm} [\delta_{ij} \delta_{jm} x_i (x_l p_j) p_m - \delta_{im} \delta_{jl} x_l (p_m x_l) p_j + i \hbar \delta_{im} x_l p_j]
\]

\[
= x^2 \mathbf{p}^2 - i \hbar \mathbf{x} \cdot \mathbf{p} - \sum_{ijlm} \delta_{im} \delta_{jl} [x_i p_m (x_l p_j) - \delta_{jm} \delta_{jl} x_i p_m + i \hbar \delta_{im} x_l p_j]
\]

\[
= x^2 \mathbf{p}^2 - (\mathbf{x} \cdot \mathbf{p})^2 + i \hbar \mathbf{x} \cdot \mathbf{p}.
\]

Before taking the preceding expression between \( \langle x \ket \) and \( |\alpha\rangle \), first note that

\[
\langle x \ket \mathbf{x} \cdot \mathbf{p} |\alpha\rangle = \left( \begin{array}{c} 1 \end{array} \right)_{\alpha} = -i \hbar r \frac{\partial}{\partial r} \langle x \ket |\alpha\rangle.
\]

Likewise,

\[
\langle x \ket (\mathbf{x} \cdot \mathbf{p})^2 |\alpha\rangle = - \hbar^2 r \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} \langle x \ket |\alpha\rangle \right)
\]

\[
= - \hbar^2 \left( r^2 \frac{\partial^2}{\partial r^2} \langle x \ket |\alpha\rangle + r \frac{\partial}{\partial r} \langle x \ket |\alpha\rangle \right).
\]

Thus

\[
\langle x \ket \mathbf{L}^2 |\alpha\rangle = r^2 \langle x \ket |\mathbf{p}^2\rangle + \hbar^2 \left( r^2 \frac{\partial^2}{\partial r^2} \langle x \ket |\alpha\rangle + 2 r \frac{\partial}{\partial r} \langle x \ket |\alpha\rangle \right).
\]

In terms of the kinetic energy \( \mathbf{p}^2/2m \), we have

\[
\frac{1}{2m} \langle x \ket |\mathbf{p}^2\rangle |\alpha\rangle = - \left( \frac{\hbar^2}{2m} \right) \nabla^2 \langle x \ket |\alpha\rangle
\]

\[
= - \left( \frac{\hbar^2}{2m} \right) \left( \frac{\partial^2}{\partial r^2} \langle x \ket |\alpha\rangle + \frac{2}{r} \frac{\partial}{\partial r} \langle x \ket |\alpha\rangle - \frac{1}{r^2} \frac{\partial}{\partial r} \langle x \ket |\mathbf{L}^2|\alpha\rangle \right).
\]

The first two terms in the last line are just the radial part of the Laplacian acting on \( \langle x \ket |\alpha\rangle \). The last term must then be the angular part of the Laplacian acting on \( \langle x \ket |\alpha\rangle \), in complete agreement with (3.6.15).

**Spherical Harmonics**

Consider a spinless particle subjected to a spherical symmetrical potential. The wave equation is known to be separable in spherical coordinates and the energy eigenfunctions can be written as

\[
\langle x \ket |n, l, m\rangle = R_n(r) Y_l^m(\theta, \phi),
\]

where the position vector \( x' \) is specified by the spherical coordinates \( r, \theta, \) and \( \phi \), and \( n \) stands for some quantum number other than \( l \) and \( m \), for example, the radial quantum number for bound-state problems or the energy for a free-particle spherical wave. As will be made clearer in Section 3.10, this form can be regarded as a direct consequence of the rotational invariance of the problem. When the Hamiltonian is spherically symmetrical, \( H \) commutes with \( L_z \) and \( \mathbf{L}^2 \), and the energy eigenkets are expected to be eigenkets of \( \mathbf{L}^2 \) and \( L_z \) also. Because \( L_z \) with \( k = 1, 2, 3 \) satisfy the angular-momentum commutation relations, the eigenvalues of \( \mathbf{L}^2 \) and \( L_z \) are expected to be \( l(l + 1) \hbar^2 \), and \( m \hbar = [-\hbar, (-l + 1) \hbar, \ldots, (l - 1) \hbar, l \hbar] \).

Because the angular dependence is common to all problems with spherical symmetry, we can isolate it and consider

\[
\langle \mathbf{n} \ket |l, m\rangle = Y_l^m(\theta, \phi) = Y_l^m(\mathbf{n}),
\]

where we have defined a direction eigenket \( \ket{\mathbf{n}} \). From this point of view, \( Y_l^m(\theta, \phi) \) is the amplitude for a state characterized by \( l, m \) to be found in the direction \( \mathbf{n} \) specified by \( \theta \) and \( \phi \).

Suppose we have relations involving orbital angular-momentum eigenkets. We can immediately write the corresponding relations involving the spherical harmonics. For example, take the eigenvalue equation

\[
L_z |l, m\rangle = m \hbar |l, m\rangle.
\]

Multiplying \( \langle \mathbf{n} \ket \) on the left and using (3.6.9), we obtain

\[
- \hbar \frac{\partial}{\partial \phi} \langle \mathbf{n} \ket |l, m\rangle = m \hbar \langle \mathbf{n} \ket |l, m\rangle.
\]

We recognize this equation to be

\[
- \hbar \frac{\partial}{\partial \phi} Y_l^m(\theta, \phi) = m \hbar Y_l^m(\theta, \phi),
\]

which implies that the \( \phi \)-dependence \( Y_l^m(\theta, \phi) \) must behave like \( e^{im\phi} \). Likewise, corresponding to

\[
L_z |l, m\rangle = l(l + 1) \hbar^2 |l, m\rangle,
\]

we have [see (3.6.15)]

\[
\left[ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + l(l + 1) \right] Y_l^m = 0,
\]

which is simply the partial differential equation satisfied by \( Y_l^m \) itself. The orthogonality relation

\[
\langle l', m'|l, m\rangle = \delta_{l'l} \delta_{m'm},
\]

(3.6.22)
leads to
\[ \int_{0}^{2\pi} d\phi \int_{-1}^{1} d(\cos \theta) Y_{m}^{*}(\theta, \phi) Y_{m}^{m}(\theta, \phi) = \delta_{m,m'}. \] (3.6.30)
where we have used the completeness relation for the direction eigenkets,
\[ \int d\Omega \langle \hat{n} | \hat{n} \rangle = 1. \] (3.6.31)

To obtain the \( Y_{m}^{m} \) themselves, we may start with the \( m = l \) case. We have
\[ L_{+}[l,l] = 0, \] (3.6.32)
which, because of (3.6.13), leads to
\[ -i\hbar e^{i\phi} \left[ i \frac{\partial}{\partial \theta} - \cot \theta \frac{\partial}{\partial \phi} \right] \langle \hat{n} | l, l \rangle = 0. \] (3.6.33)
Remembering that the \( \phi \)-dependence must behave like \( e^{i\phi} \), we can easily show that this partial differential equation is satisfied by
\[ \langle \hat{n} | l, l \rangle = Y_{l}^{l}(\theta, \phi) = c_{l} e^{i\phi} \sin l \theta, \] (3.6.34)
where \( c_{l} \) is the normalization constant determined from (3.6.30) to be
\[ c_{l} = \left( -1 \right)^{l} \frac{l!}{2^{l}!} \sqrt{\frac{(2l+1)(2l)!}{4\pi}}. \] (3.6.35)

Starting with (3.6.34) we can use
\[ \langle \hat{n} | l, m \rangle = \frac{\langle \hat{n} | L_{-} | l, m \rangle}{\sqrt{(l+m)(l-m+1)\hbar}} \]
\[ = \frac{1}{\sqrt{(l+m)(l-m+1)}} e^{-i\phi} \left( - \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) \langle \hat{n} | l, m \rangle \] (3.6.36)
successively to obtain all \( Y_{m}^{m} \) with \( l \) fixed. Because this is done in many textbooks on elementary quantum mechanics, we will not work out the details here. The result for \( m \geq 0 \) is
\[ Y_{m}^{m}(\theta, \phi) = \left( -1 \right)^{l} \frac{(l+m)!}{2^{l}! (l-m)!} \sqrt{\frac{2l+1}{4\pi}} \frac{e^{im\phi}}{\sin m \theta} \frac{d^{l-m}}{d(\cos \theta)^{l-m}}(\sin \theta)^{2l} \] (3.6.37)

and we define \( Y_{m}^{-m} \) by
\[ Y_{m}^{-m}(\theta, \phi) = (-1)^{m} \left[ Y_{m}^{m}(\theta, \phi) \right]^*. \] (3.6.38)
Regardless of whether \( m \) is positive or negative, the \( \theta \)-dependent part of \( Y_{m}^{m}(\theta, \phi) \) is \([\sin \theta]^{m}\) times a polynomial in \( \cos \theta \) with a highest power of \( l - |m| \). For \( m = 0 \), we obtain
\[ Y_{0}^{0}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}} P_{l}(\cos \theta). \] (3.6.39)

From the point of view of the angular-momentum commutation relations alone, it might not appear obvious why \( l \) cannot be a half-integer. It turns out that several arguments can be advanced against half-integer \( l \)-values. First, for half-integer \( l \), and hence for half-integer \( m \), the wave function would acquire a minus sign,
\[ e^{-i\phi(2\pi)} = -1, \] (3.6.40)
under a \( 2\pi \) rotation. As a result, the wave function would not be single-valued; we pointed out in Section 2.4 that the wave function must be single-valued because of the requirement that the expansion of a state ket in terms of position eigenkets be unique. We can prove that if \( L \), defined to be \( \bf{\times} \bf{p} \), is to be identified as the generator of rotation, then the wave function must acquire a plus sign under a \( 2\pi \) rotation. This follows from the fact that the wave function for a \( 2\pi \)-rotated state is the original wave function itself with no sign change:
\[ \langle x'| \exp \left( -iL_{2\pi \hbar} \right) |\alpha \rangle = \langle x' | \cos 2\pi + y'sin 2\pi, y'cos 2\pi - x'sin 2\pi, z' |\alpha \rangle \]
\[ = \langle x' |\alpha \rangle, \] (3.6.41)
where we have used the finite-angle version of (3.6.6). Next, let us suppose \( Y_{m}^{m}(\theta, \phi) \) with a half-integer \( l \) were possible. To be specific, we choose the simplest case, \( l = m = \frac{1}{2} \). According to (3.6.34) we would have
\[ Y_{l=1/2}^{1/2}(\theta, \phi) = c_{1/2} e^{i\phi/2 \sqrt{\sin \theta}}. \] (3.6.42)
From the property of \( L_{-} \) [see (3.6.36)] we would then obtain
\[ Y_{1/2}^{1/2}(\theta, \phi) = e^{-i\phi} \left( - \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) \left( c_{1/2} e^{i\phi/2 \sqrt{\sin \theta}} \right) \]
\[ = - c_{1/2} e^{-i\phi/2 \cot \theta \sqrt{\sin \theta}}. \] (3.6.43)
This expression is not permissible because it is singular at \( \theta = 0, \pi \). What is worse, from the partial differential equation
\[ \langle \hat{n} | L_{-} | \frac{1}{2}, -\frac{1}{2} \rangle = -i\hbar e^{-i\phi} \left( - i \frac{\partial}{\partial \theta} - \cot \theta \frac{\partial}{\partial \phi} \right) \langle \hat{n} | \frac{1}{2}, -\frac{1}{2} \rangle \]
\[ = 0 \] (3.6.44)
we directly obtain
\[ Y_{l/2} = c_{l/2} e^{-i\theta / 2\sqrt{\sin \theta}}, \]
(3.6.45)
in sharp contradiction with (3.6.43). Finally, we know from the Sturm-Liouville theory of differential equations that the solutions of (3.6.28) with integer form a complete set. An arbitrary function of \( \theta \) and \( \phi \) can be expanded in terms of \( Y_{l}^{m} \) with integer \( l \) and \( m \) only. For all these reasons it is futile to contemplate orbital angular momentum with half-integer \( l \)-values.

Spherical Harmonics as Rotation Matrices

We conclude this section on orbital angular momentum by discussing the spherical harmonics from the point of view of the rotation matrices introduced in the last section. We can readily establish the desired connection between the two approaches by constructing the most general direction eigenket |\( \mathbf{h} \rangle \) by applying appropriate rotation operators to |\( \mathbf{h} \rangle \), the direction eigenket in the positive \( z \)-direction. We wish to find \( \mathcal{D}(R) \) such that
\[ |\mathbf{h} \rangle = \mathcal{D}(R)|\mathbf{h} \rangle. \]
(3.6.46)

We can rely on the technique used in constructing the eigenspinor of \( \mathbf{\sigma} \cdot \mathbf{\mathbf{h}} \) in Section 3.2. We first rotate about the \( y \)-axis by angle \( \theta \), then around the \( z \)-axis by angle \( \phi \); see Figure 3.3 with \( \beta \rightarrow \theta \), \( \alpha \rightarrow \phi \). In the notation of Euler angles we have
\[ \mathcal{D}(R) = \mathcal{D}(\alpha, \beta, \gamma = 0). \]
(3.6.47)
Writing (3.6.46) as
\[ |\mathbf{h} \rangle = \sum_{l, m} \mathcal{D}(R)|l, m \rangle \langle l, m \mathbf{h} \rangle, \]
(3.6.48)
we see that |\( \mathbf{h} \rangle \), when expanded in terms of \( |l, m \rangle \), contains all possible \( l \)-values. However, when this equation is multiplied by \( \langle l, m \rangle \) on the left, only one term in the \( l \)-sum contributes, namely,
\[ \langle l, m |\mathbf{h} \rangle = \sum_{m} \mathcal{D}^{(l)}|l, m \rangle \langle l, m \mathbf{h} \rangle \]
(3.6.49)
Now \( \langle l, m \mathbf{h} \rangle \) is just a number; in fact, it is precisely \( Y_{l}^{m}(\theta, \phi) \) evaluated at \( \theta = 0 \) with \( \phi \) undetermined. At \( \theta = 0 \), \( Y_{l}^{m} \) is known to vanish for \( m \neq 0 \), which can also be seen directly from the fact that \( |\mathbf{h} \rangle \) is an eigenket of \( L_{z} \) (which equals \( xp_{y} - yp_{x} \)) with eigenvalue zero. So we can write
\[ \langle l, m |\mathbf{h} \rangle = Y_{l}^{m}(\theta = 0, \phi \text{ undetermined}) \delta_{m,0}. \]
(3.6.50)

1.7. Addition of Angular Momentum

Returning to (3.6.49), we have
\[ Y_{l}^{m}(\theta, \phi) = \sqrt{(2l+1) \over 4\pi} \mathcal{D}^{(l)}|l, m \rangle \langle l, m \mathbf{h} \rangle \]
(3.6.51)
or
\[ \mathcal{D}^{(l)}|l, m \rangle \langle l, m \mathbf{h} \rangle \]
(3.6.52)
Notice the \( m = 0 \) case, which is of particular importance:
\[ d^{(l)}_{m,0} = P_{l}(\cos \theta). \]
(3.6.53)

3.7. Addition of Angular Momentum

Angular-momentum addition has important applications in all areas of modern physics—from atomic spectroscopy to nuclear and particle collisions. Furthermore, a study of angular-momentum addition provides an excellent opportunity to illustrate the concept of change of basis, which we discussed extensively in Chapter 1.

Simple Examples of Angular-Momentum Addition

Before studying a formal theory of angular-momentum addition, it is worth looking at two simple examples with which the reader may be familiar: (1) how to add orbital angular momentum and spin-angular momentum and (2) how to add the spin-angular momenta of two spin \( \tfrac{1}{2} \) particles.

Previously we studied both spin \( \tfrac{1}{2} \) systems with all quantum-mechanical degrees of freedom other than spin—such as position and momentum—ignored and quantum-mechanical particles with the space degrees of freedom (such as position or momentum) taken into account but the internal degrees of freedom (such as spin) ignored. A realistic description of a particle with spin must of course take into account both the space degree of freedom and the internal degrees of freedom. The base ket for a spin \( \tfrac{1}{2} \) particle may be visualized to be in the direct-product space of the infinite-dimensional ket space spanned by the position eigenkets \( \{|X_{\alpha}\rangle\} \) and the two-dimensional spin space spanned by \(|+\rangle\) and \(|-\rangle\). Explicitly, we have for the base ket
\[ |X_{\alpha}, \pm\rangle = |X_{\alpha}\rangle \otimes |\pm\rangle, \]
(3.7.1)
where any operator in the space spanned by \( \{|X_{\alpha}\rangle\} \) commutes with any operator in the two-dimensional space spanned by \(|\pm\rangle\).
The rotation operator still takes the form $\exp(-iJ\cdot\hat{\mathbf{\phi}}/\hbar)$ but $J$ is a generator of rotations, is now made up of two parts, namely,

$$J = L + S.$$  

It is actually more obvious to write (3.7.2) as

$$J = L \otimes 1 + 1 \otimes S,$$  

where the 1 in $L \otimes 1$ stands for the identity operator in the spin space, while the 1 in $1 \otimes S$ stands for the identity operator in the infinite-dimensional spin space spanned by the position eigenkets. Because $L$ and $S$ commute, we can write

$$\mathcal{D}(R) = \mathcal{D}^{(\text{orb})}(R) \otimes \mathcal{D}^{(\text{spin})}(R)$$

$$= \exp\left(-iL \cdot \hat{\mathbf{\phi}}/\hbar\right) \otimes \exp\left(-iS \cdot \hat{\mathbf{\phi}}/\hbar\right).$$  

The wave function for a particle with spin is written as

$$\langle \mathbf{x}', \pm |a\rangle = \psi_{\pm}(\mathbf{x}').$$  

The two components $\psi_{\pm}$ are often arranged in column matrix form as follows:

$$\begin{pmatrix}
\psi_{+}(\mathbf{x}') \\
\psi_{-}(\mathbf{x}')
\end{pmatrix}.$$  

where $|\psi_{\pm}(\mathbf{x}')|^2$ stands for the probability density for the particle to be found at $\mathbf{x}'$ with spin up and down, respectively. Instead of $|\mathbf{x}'\rangle$ as the base kets for the space part, we may use $|n, l, m\rangle$, which are eigenkets of $L^2$ and $L_z$, with eigenvalues $\hbar^2(l + 1)$ and $m\hbar$, respectively. For the spin part, $|\pm\rangle$ are eigenkets of $S^2$ and $S_z$ with eigenvalues $\hbar^2/4$ and $\pm \hbar/2$, respectively. However, as we will show later, we can also use base kets which are the eigenkets of $J^2$, $J_z$, $L^2$, and $S^2$. In other words, we can expand a state ket of a particle with spin in terms of simultaneous eigenkets of $L^2$, $S^2$, $L_z$, and $S_z$ or in terms of simultaneous eigenkets of $J^2$, $J_z$, $L^2$, and $S^2$. We will study in detail how the two descriptions are related.

As a second example, we study two spin $\frac{1}{2}$ particles—say two electrons—with the orbital degree of freedom suppressed. The total spin operator is usually written as

$$S = S_1 + S_2,$$  

but again it is to be understood as

$$S_1 \otimes 1 + 1 \otimes S_2,$$  

where the 1 in the first (second) term stands for the identity operator in the spin space of electron 2 (1). We, of course, have

$$[S_{1x}, S_{2y}] = 0$$  

and so forth. Within the space of electron 1(2) we have the usual commutation relations

$$[S_{1x}, S_{1y}] = i\hbar S_{1z}, [S_{2x}, S_{2y}] = i\hbar S_{2z}, \ldots.$$  

A direct consequence of (3.7.9) and (3.7.10), we have

$$[S_x, S_y] = i\hbar S_z.$$  

Again, we can expand the ket corresponding to an arbitrary spin state of two electrons in terms of either the eigenkets of $S^2$ and $S_z$ or the eigenkets of $S_{1z}$ and $S_{2z}$. The two possibilities are as follows:

1. The $(m_1, m_2)$ representation based on the eigenkets of $S_{1z}$ and $S_{2z}$:

$$|+ +\rangle, |+ -\rangle, |- +\rangle, \text{ and } | - -\rangle,$$  

where $|+ -\rangle$ stands for $m_1 = \frac{1}{2}, m_2 = -\frac{1}{2}$, and so forth.

2. The $(s, m)$ representation (or the triplet-singlet representation) based on the eigenkets of $S^2$ and $S_z$:

$$|s = 1, m = \pm 1, 0\rangle, |s = 0, m = 0\rangle,$$  

where $s = 1$ $(s = 0)$ is referred to as spin triplet (spin singlet).

Notice that in each set there are four base kets. The relationship between the two sets of base kets is as follows:

$$|s = 1, m = 1\rangle = |+ +\rangle,$$  

$$|s = 1, m = 0\rangle = \left(\frac{1}{\sqrt{2}}\right)(|+ -\rangle + |- +\rangle),$$  

$$|s = 1, m = -1\rangle = |- -\rangle,$$  

$$|s = 0, m = 0\rangle = \left(\frac{1}{\sqrt{2}}\right)(|+ -\rangle - |- +\rangle).$$  

The right-hand side of (3.7.15a) tells us that we have both electrons with spin up; this situation can correspond only to $s = 1, m = 1$. We can obtain (3.7.15b) from (3.7.15a) by applying the ladder operator

$$S_+ \equiv S_{1+} + S_{2+}$$

$$= (S_{1x} - iS_{1y}) + (S_{2x} - iS_{2y})$$  

and so forth. Within the space of electron 1(2) we have the usual commutation relations

$$[S_{1x}, S_{1y}] = i\hbar S_{1z}, [S_{2x}, S_{2y}] = i\hbar S_{2z}, \ldots.$$  

A direct consequence of (3.7.9) and (3.7.10), we have

$$[S_x, S_y] = i\hbar S_z.$$  

Again, we can expand the ket corresponding to an arbitrary spin state of two electrons in terms of either the eigenkets of $S^2$ and $S_z$ or the eigenkets of $S_{1z}$ and $S_{2z}$. The two possibilities are as follows:

1. The $(m_1, m_2)$ representation based on the eigenkets of $S_{1z}$ and $S_{2z}$:

$$|+ +\rangle, |+ -\rangle, |- +\rangle, \text{ and } | - -\rangle,$$  

where $|+ -\rangle$ stands for $m_1 = \frac{1}{2}, m_2 = -\frac{1}{2}$, and so forth.

2. The $(s, m)$ representation (or the triplet-singlet representation) based on the eigenkets of $S^2$ and $S_z$:

$$|s = 1, m = \pm 1, 0\rangle, |s = 0, m = 0\rangle,$$  

where $s = 1$ $(s = 0)$ is referred to as spin triplet (spin singlet).
To both sides of (3.7.15a). In doing so we must remember that an electron operator like \( S_{1-} \) affects just the first entry of \(|+\rangle \), and so on. We can write
\[
S_- |s = 1, m = 1\rangle = (S_1 + S_2) |+\rangle
\]
and
\[
\sqrt{(1+1)(1-1+1)} |s = 1, m = 0\rangle = \sqrt{(1+1)(1-1+1) + |+\rangle}
\]
which immediately leads to (3.7.15b). Likewise, we can obtain \(|s = 1, m = -1\rangle\) by applying (3.7.16) once again to (3.7.15b). Finally, we can obtain (3.7.15d) by requiring it to be orthogonal to the other three kets, in particular to (3.7.15b).

The coefficients that appear on the right-hand side of (3.7.15) are the simplest example of Clebsch-Gordan coefficients to be discussed further at a later time. They are simply the elements of the transformation matrix that connects the \( \{m_1, m_2\} \) basis to the \( \{s, m\} \) basis. It is instructive to derive these coefficients in another way. Suppose we write the \( 4 \times 4 \) matrix corresponding to
\[
S^2 = S_1^2 + S_2^2 + 2S_1 S_2
\]
using the \( \{m_1, m_2\} \) basis. The square matrix is obviously not diagonal because an operator like \( S_{1+} \) connects \(|+\rangle\) with \(|+\rangle\). The unitary matrix that diagonalizes this matrix carries the \( \{|m_1, m_2\}\) base kets into the \( \{|s, m\}\) base kets. The elements of this unitary matrix are precisely the Clebsch-Gordan coefficients for this problem. The reader is encouraged to work out all this in detail.

### Formal Theory of Angular-Momentum Addition

Having gained some physical insight by considering simple examples, we are now in a position to study more systematically the formal theory of angular-momentum addition. Consider two angular-momentum operators \( J_1 \) and \( J_2 \) in different subspaces. The components of \( J_1(J_2) \) satisfy the usual angular-momentum commutation relations:
\[
[J_1^i, J_1^j] = i\hbar \epsilon_{ijk} J_{1k}
\]
and
\[
[J_2^i, J_2^j] = i\hbar \epsilon_{ijk} J_{2k}.
\]

However, we have
\[
[J_{1k}, J_{2l}] = 0
\]
between any pair of operators from different subspaces.

The infinitesimal rotation operator that affects both subspace 1 and subspace 2 is written as
\[
(1 - i \frac{\mathbf{J}_1 \cdot \hat{n} \delta \phi}{\hbar}) \otimes (1 - i \frac{\mathbf{J}_2 \cdot \hat{n} \delta \phi}{\hbar}) = 1 - i(\mathbf{J}_1 \otimes 1 + 1 \otimes \mathbf{J}_2) \cdot \hat{n} \delta \phi.
\]

We define the total angular momentum by
\[
\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2,
\]
which is more commonly written as
\[
\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2.
\]

The finite-angle version of (3.7.22) is
\[
\mathcal{D}_1(R) \otimes \mathcal{D}_2(R) = \exp\left(-i \frac{\mathbf{J}_1 \cdot \mathbf{h}}{\hbar}\phi \right) \otimes \exp\left(-i \frac{\mathbf{J}_2 \cdot \mathbf{h}}{\hbar}\phi \right).
\]

Notice the appearance of the same axis of rotation and the same angle of rotation.

It is very important to note that the total \( \mathbf{J} \) satisfies the angular-momentum commutation relations
\[
[J_i, J_j] = i\hbar \epsilon_{ijk} J_k
\]
as a direct consequence of (3.7.20) and (3.7.21). In other words, \( \mathbf{J} \) is an angular momentum in the sense of Section 3.1. Physically this is reasonable because \( \mathbf{J} \) is the generator for the entire system. Everything we learned in Section 3.5—for example, the eigenvalue spectrum of \( \mathbf{J}^2 \) and \( \mathbf{J} \), and the matrix elements of the ladder operators—also holds for the total \( \mathbf{J} \).

As for the choice of base kets we have two options.

**Option A:** Simultaneous eigenkets of \( J_{1z}, J_{1z}, J_{1z}, \text{ and } J_{2z} \). Denoted by \( |J_{1z}; m_1 m_2\rangle \). Obviously the four operators commute with each other. The defining equations are
\[
J_{1z} |J_{1z}; m_1 m_2\rangle = m_1 |J_{1z}; m_1 m_2\rangle,
\]
\[
J_{2z} |J_{2z}; m_1 m_2\rangle = m_2 |J_{2z}; m_1 m_2\rangle.
\]

**Option B:** Simultaneous eigenkets of \( J^2, J_{1z}^2, J_{2z}^2, \text{ and } J_z \). First, note that this set of operators mutually commute. In particular, we have
\[
[J^2, J_z^2] = 0.
\]
which can readily be seen by writing \( J^2 \) as

\[
J^2 = J_x^2 + J_y^2 + 2J_z; J_x; J_y + J_1; J_2 + J_{1-} J_{2+}.
\]

(3.7.29)

We use \( |j_1, j_2; jm\rangle \) to denote the base kets of option B:

\[
\begin{align*}
J_x |j_1, j_2; jm\rangle &= j_x (j_1 + 1)h_x |j_1, j_2; jm\rangle, \\
J_y |j_1, j_2; jm\rangle &= j_y (j_2 + 1)h_y |j_1, j_2; jm\rangle, \\
J_z |j_1, j_2; jm\rangle &= j_z (j + 1)h_z |j_1, j_2; jm\rangle, \\
J_{1\pm} |j_1, j_2; jm\rangle &= m \pm |j_1, j_2; jm\rangle.
\end{align*}
\]

(3.7.30)

Quite often \( j_1, j_2 \) are understood, and the base kets are written simply as \( |j, m\rangle \).

It is very important to note that even though

\[
[J^2, J_z] = 0,
\]

(3.7.31)

we have

\[
[J^2, J_{1\pm}] \neq 0, \quad [J^2, J_{2\pm}] \neq 0,
\]

(3.7.32)

as the reader may easily verify using (3.7.29). This means that we cannot add \( J^2 \) to the set of operators of option A. Likewise, we cannot add \( J_{1\pm} \) and/or \( J_{2\pm} \) to the set of operators of option B. We have two possible sets of base kets corresponding to the two maximal sets of mutually compatible observables we have constructed.

Let us consider the unitary transformation in the sense of Section 1.5 that connects the two bases:

\[
|j_1, j_2; jm\rangle = \sum_{m_1} \sum_{m_2} |j_1, j_2; m_1m_2\rangle |j_1, j_2; m_1m_2; jm\rangle,
\]

(3.7.33)

where we have used

\[
\sum_{m_1} \sum_{m_2} |j_1, j_2; m_1m_2\rangle |j_1, j_2; m_1m_2\rangle = 1
\]

(3.7.34)

and where the right-hand side is the identity operator in the ket space of given \( j_1 \) and \( j_2 \). The elements of this transformation matrix \( \langle j_1, j_2; m_1m_2|j_1, j_2; jm\rangle \) are Clebsch-Gordan coefficients.

There are many important properties of Clebsch-Gordan coefficients that we are now ready to study. First, the coefficients vanish unless

\[
m = m_1 + m_2.
\]

(3.7.35)

To prove this, first note that

\[
(J_z - J_{1z} - J_{2z}) |j_1, j_2; jm\rangle = 0.
\]

(3.7.36)

Multiplying \( \langle j_1, j_2; m_1m_2|j_1, j_2; jm\rangle \) on the left, we obtain

\[
(m - m_2) \langle j_1, j_2; m_1m_2|j_1, j_2; jm\rangle = 0,
\]

(3.7.37)

which proves our assertion. Admire the power of the Dirac notation! It really pays to write the Clebsch-Gordan coefficients in Dirac's bracket form, as we have done.

Second, the coefficients vanish unless

\[
|j_1 - j_2| \leq j \leq j_1 + j_2.
\]

(3.7.38)

This property may appear obvious from the vector model of angular-momentum addition, where we visualize \( J \) to be the vectorial sum of \( J_1 \) and \( J_2 \). However, it is worth checking this point by showing that if (3.7.38) holds, then the dimensionality of the space spanned by \( \{j_1, j_2; m_1m_2\}\) is the same as that of the space spanned by \( \{j_1, j_2; jm\}\). For the \((m_1, m_2)\) way of counting we obtain

\[
N = (2j_1 + 1)(2j_2 + 1)
\]

(3.7.39)

because for given \( j_1 \) there are \( 2j_1 + 1 \) possible values of \( m_1 \); a similar statement is true for the other angular momentum \( j_2 \). As for the \((j, m)\) way of counting, we note that for each \( j \), there are \( 2j + 1 \) states, and according to (3.7.38), \( j \) itself runs from \( j_1 - j_2 \) to \( j_1 + j_2 \), where we have assumed, without loss of generality, that \( j_1 \geq j_2 \). We therefore obtain

\[
N = \sum_{j = j_1 - j_2}^{j_1 + j_2} (2j + 1)
\]

\[
= \frac{1}{2} \left[ (2(j_1 - j_2) + 1) + (2(j_1 + j_2) + 1) \right] (2j_2 + 1)
\]

\[
= (2j_1 + 1)(2j_2 + 1).
\]

(3.7.40)

Because both ways of counting give the same \( N\)-value, we see that (3.7.38) is quite consistent.*

The Clebsch-Gordan coefficients form a unitary matrix. Furthermore, the matrix elements are taken to be real by convention. An immediate consequence of this is that the inverse coefficient \( \langle j_1, j_2; jm|j_1, j_2; m_1m_2\rangle \) is the same as \( \langle j_1, j_2; m_1m_2|j_1, j_2; jm\rangle \) itself. A real unitary matrix is orthogonal, so we have the orthogonality condition

\[
\sum_{j} \sum_{m} \langle j_1, j_2; m_1m_2|j_1, j_2; jm\rangle \langle j_1, j_2; m_1m_2|j_1, j_2; jm\rangle = \delta_{m_1m_1'} \delta_{m_2m_2'}
\]

(3.7.41)

which is obvious from the orthonormality of \( \{j_1, j_2; m_1m_2\}\) together with the reality of the Clebsch-Gordan coefficients. Likewise, we also have

\[
\sum_{j} \sum_{m} \langle j_1, j_2; m_1m_2|j_1, j_2; jm\rangle \langle j_1, j_2; m_1m_2|j_1, j_2; jm\rangle = \delta_{j_1j_1'} \delta_{m_1m_1'}
\]

(3.7.42)

*A complete proof of (3.7.38) is given in Gottfried 1966, 215, and also in Appendix B of this book.
As a special case of this we may set \( j' = j \), \( m' = m = m_1 + m_2 \). We then obtain
\[
\sum_{m_1} \sum_{m_2} |j_1 j_2; m_1 m_2; j_1 j_2; jm\rangle^2 = 1, \tag{3.7.43}
\]
which is just the normalization condition for \( |j_1 j_2; jm\rangle \).

Some authors use somewhat different notations for the Clebsch-Gordan coefficients. Instead of \( |j_1 j_2; m_1 m_2; j_1 j_2; jm\rangle \) we sometimes see \( |j_1 m_1; j_2 m_2; j_1 j_2; jm\rangle \), \( C(j_1 j_2; j_1 m_1; j_2 m_2 \mid jm) \), \( C_{j_1 j_2} \), and so on. They can also be written in terms of Wigner's 3-\( j \) symbol, which is occasionally found in the literature:
\[
\langle j_1 j_2; m_1 m_2; j_1 j_2; jm \rangle = (-1)^{j_1 + j_2 + m} \sqrt{2j_1 + 1} \binom{j_1 j_2}{m_1 j_2; m_2 - m}. \tag{3.7.44}
\]

### Recursion Relations for the Clebsch-Gordan Coefficients

With \( j_1 \), \( j_2 \), and \( j \) fixed, the coefficients with different \( m_1 \) and \( m_2 \) are related to each other by recursion relations. We start with
\[
J_+ |j_1 j_2; jm\rangle = (J_1 + J_2) \sum_{m_1 m_2} |j_1 j_2; m_1 m_2 \rangle \langle j_1 j_2; m_1 m_2; j_1 j_2; jm \rangle. \tag{3.7.45}
\]

Using (3.5.39) and (3.5.40) we obtain (with \( m_1 \to m'_1 \), \( m_2 \to m'_2 \))
\[
\sqrt{(j + m)(j + m + 1)} |j_1 j_2; j, m + 1\rangle
= \sum_{m'_1 m'_2} \left( \sqrt{(j_1 + m'_1)(j_1 + m'_1 + 1)} |j_1 j_2; m'_1 \pm 1, m'_2\rangle
+ \sqrt{(j_2 + m'_2)(j_2 + m'_2 + 1)} |j_1 j_2; m'_1, m'_2 \pm 1\rangle \right)
\times |j_1 j_2; m'_1 m'_2; j_1 j_2; jm\rangle. \tag{3.7.46}
\]

Our next step is to multiply by \( |j_1 j_2; m_1 m_2\rangle \) on the left and use orthonormality, which means that nonvanishing contributions from the right-hand side are possible only with
\[
m_1 = m'_1 \pm 1, \quad m_2 = m'_2 \tag{3.7.47}
\]
for the first term and
\[
m_1 = m'_1, \quad m_2 = m'_2 \pm 1 \tag{3.7.48}
\]
for the second term. In this manner we obtain the desired recursion relations:
\[
\sqrt{(j + m)(j + m + 1)} \langle j_1 j_2; m_1 m_2; j_1 j_2; jm \rangle
= \sqrt{(j_1 + m_1)(j_1 \pm m_1)} \langle j_1 j_2; m_1 \pm 1, m_2 j_1 j_2; jm \rangle
+ \sqrt{(j_2 + m_2)(j_2 \pm m_2)} \langle j_1 j_2; m_1, m_2 \pm 1; j_1 j_2; jm \rangle. \tag{3.7.49}
\]

It is important to note that because the \( J_+ \) operators have shifted the \( m \)-values, the nonvanishing condition (3.7.35) for the Clebsch-Gordan coefficients has now become \( [\text{when applied to (3.7.49)}] \)
\[
m_1 + m_2 = m \pm 1. \tag{3.7.50}
\]

We can appreciate the significance of the recursion relations by looking at (3.7.49) in an \( m_1, m_2 \)-plane. The \( J_+ \) recursion relation (upper sign) tells us that the coefficient at \( (m_1, m_2) \) is related to the coefficients at \( (m_1 - 1, m_2) \) and \( (m_1, m_2 - 1) \), as shown in Figure 3.5a. Likewise, the \( J_- \) recursion relation (lower sign) relates the three coefficients whose \( m_1, m_2 \) values are given in Figure 3.5b.

Recursion relations (3.7.49), together with normalization condition (3.7.43), almost uniquely determine all Clebsch-Gordan coefficients. (We say "almost uniquely" because certain sign conventions have yet to be specified.) Our strategy is as follows. We go back to the \( m_1 m_2 \)-plane, again for fixed \( j_1 \), \( j_2 \), and \( j \), and plot the boundary of the allowed region determined by
\[
|m_1| \leq j_1, \quad |m_2| \leq j_2, \quad -j \leq m_1 + m_2 \leq j; \tag{3.7.51}
\]
see Figure 3.6a. We may start with the upper right-hand corner, denoted by \( A \). Because we work near \( A \) at the start, a more detailed "map" is in order;

**FIGURE 3.5.** \( m_1 m_2 \)-plane showing the Clebsch-Gordan coefficients related by the recursion relations (3.7.49).
see Figure 3.6b. We apply the $J_-$ recursion relation (3.7.49) (lower sign) with $(m_1, m_2 + 1)$ corresponding to $A$. Observe now that the recursion relation connects $A$ with only $B$ because the site corresponding to $(m_1, m_2)$ is forbidden by $m_1 \leq j_1$. As a result, we can obtain the Clebsch-Gordan coefficient of $B$ in terms of the coefficient of $A$. Next, we form a triangle made up of $A$, $B$, and $D$. This enables us to obtain the coefficient of $D$ once the coefficient of $A$ is specified. We can continue in this fashion. Knowing $B$ and $D$, we can get to $E$; knowing $B$ and $E$ we can get to $F$, and so on. With enough patience we can obtain the Clebsch-Gordan coefficient of every site in terms of the coefficient of starting site, $A$. For overall normalization we use (3.7.43). The final overall sign is fixed by the convention. (See the following example.)

As an important practical example we consider the problem of adding the orbital and spin-angular momenta of a single spin $\frac{1}{2}$ particle. We have

\[
\begin{align*}
    j_1 &= l \quad \text{(integer)}, \\
    j_2 &= s = \frac{1}{2}, \\
    m_1 &= m_2 = \pm \frac{1}{2}. \\
\end{align*}
\]

The allowed values of $j$ are given by

\[
    j = l \pm \frac{1}{2}, \quad l > 0; \quad j = \frac{1}{2}, \quad l = 0
\]

So for each $l$ there are two possible $j$-values; for example, for $l = 1$ (p state) we get, in spectroscopic notation, $p_{3/2}$ and $p_{1/2}$, where the subscript refers to $j$. The $m_1 m_2$-plane, or better the $m_1 m_s$-plane, of this problem is particularly simple. The allowed sites form only two rows: the upper row for $m_1 = +\frac{1}{2}$ and the lower row for $m_1 = -\frac{1}{2}$; see Figure 3.7. Specifically, we work out:

\[
\begin{align*}
    \langle m_1 = m_j = m - \frac{1}{2}, & \quad m_2 = m_j = \frac{1}{2} \rangle \\
& \quad \text{the case } j = l + \frac{1}{2}. \quad \text{Because } m_j \text{ cannot exceed } \frac{1}{2}, \text{ we can use the } J_- \text{ recursion in such a way that we always stay in the upper row } (m_2 = m_j = \frac{1}{2}) \text{, while the } m_j \text{-value changes by one unit each time we consider a new } J_- \text{ triangle.}
\end{align*}
\]

suppressing $j_1 = l, \quad j_2 = \frac{1}{2}$ in writing the Clebsch-Gordan coefficients, we obtain from (3.7.49) (lower sign)

\[
\begin{align*}
    \sqrt{(l + \frac{1}{2} + m + 1)(l + \frac{1}{2} - m)} \langle m - \frac{1}{2}, \frac{1}{2}| l + \frac{1}{2}, m \rangle \\
& = \sqrt{(l + m + \frac{1}{2})(l - m + \frac{1}{2})} \langle m + \frac{1}{2}, \frac{1}{2}| l + \frac{1}{2}, m + 1 \rangle,
\end{align*}
\]

where we have used

\[
\begin{align*}
    m_1 &= m_j = m - \frac{1}{2}, \\
    m_2 &= m_j = \frac{1}{2}.
\end{align*}
\]

In this way we can move horizontally by one unit:

\[
\langle m - \frac{1}{2}, \frac{1}{2}| l + \frac{1}{2}, m \rangle = \sqrt{\frac{l + m + \frac{1}{2}}{l + m + \frac{1}{2}}} \langle m + \frac{1}{2}, \frac{1}{2}| l + \frac{1}{2}, m + 1 \rangle.
\]

We can in turn express $\langle m + \frac{1}{2}, \frac{1}{2}| l + \frac{1}{2}, m + 1 \rangle$ in terms of $\langle m + \frac{1}{2}, \frac{1}{2}| l + \frac{1}{2}, m + 2 \rangle$, and so forth. Clearly, this procedure can be continued until $m_j$ reaches $l$, the maximum possible value:

\[
\begin{align*}
    \langle m - \frac{1}{2}, \frac{1}{2}| l + \frac{1}{2}, m \rangle &= \sqrt{\frac{l + m + \frac{1}{2}}{l + m + \frac{1}{2}}} \sqrt{\frac{l + m + \frac{1}{2}}{l + m + \frac{1}{2}}} \langle m + \frac{1}{2}, \frac{1}{2}| l + \frac{1}{2}, m + 2 \rangle \\
& = \sqrt{\frac{l + m + \frac{1}{2}}{l + m + \frac{1}{2}}} \sqrt{\frac{l + m + \frac{1}{2}}{l + m + \frac{1}{2}}} \sqrt{\frac{l + m + \frac{1}{2}}{l + m + \frac{1}{2}}} \langle m + \frac{1}{2}, \frac{1}{2}| l + \frac{1}{2}, m + 3 \rangle \\
& \cdots \\
& = \sqrt{\frac{l + m + \frac{1}{2}}{2l + 1}} \langle l, \frac{1}{2}| l + \frac{1}{2}, l + \frac{1}{2} \rangle.
\end{align*}
\]
Consider the angular-momentum configuration in which \( m_j \) and \( m_l \) are both maximal, that is, \( l \) and \( \frac{1}{2} \), respectively. The total \( m = m_j + m_l \) is \( + \frac{1}{2} \), which is possible only for \( j = l + \frac{1}{2} \) and not for \( j = l - \frac{1}{2} \). So \( |m_j = m_l = \frac{1}{2} \rangle \) must be equal to \( |j = l + \frac{1}{2}, m = l + \frac{1}{2} \rangle \) up to a phase factor. We take this phase factor to be real and positive by convention. With this choice we have

\[
\langle j, \frac{1}{2} | l + \frac{1}{2}, l + \frac{1}{2} \rangle = 1.
\] (3.7.58)

Returning to (3.7.57), we finally obtain

\[
\left\langle m - \frac{1}{2}, \frac{1}{2} | j + \frac{1}{2}, m \right\rangle = \sqrt{\frac{l + m + \frac{1}{2}}{2l + 1}}.
\] (3.7.59)

But this is only about one-fourth of the story. We must still determine the value of the question marks that appear in the following:

\[
| j = l + \frac{1}{2}, m \rangle = \sqrt{\frac{l + m + \frac{1}{2}}{2l + 1}} | m_j = m - \frac{1}{2}, m_s = \frac{1}{2} \rangle
\] (3.7.60)

\[
| j = l - \frac{1}{2}, m \rangle = \begin{cases} | m_j = m + \frac{1}{2}, m_s = -\frac{1}{2} \rangle & \text{if } m = m_j + m_l \text{ is maximal} \\ | m_j = m - \frac{1}{2}, m_s = \frac{1}{2} \rangle & \text{if } m = m_j + m_l \text{ is not maximal} \end{cases}
\]

We note that the transformation matrix with fixed \( m \) from the \((m, m)\) basis to the \((j, m)\) basis is, because of orthogonality, expected to have the form

\[
\begin{pmatrix}
-\cos \alpha & \sin \alpha \\
-\sin \alpha & \cos \alpha
\end{pmatrix}
\] (3.7.61)

Comparison with (3.7.60) shows that \( \cos \alpha \) is (3.7.59) itself; so we can readily determine \( \sin \alpha \) up to a sign ambiguity:

\[
\sin^2 \alpha = 1 - \frac{(l + m + \frac{1}{2})}{(2l + 1)} = \frac{(l - m + \frac{1}{2})}{(2l + 1)}.
\] (3.7.62)

We claim that \( \langle m_j = m + \frac{1}{2}, m_s = -\frac{1}{2} | j = l + \frac{1}{2}, m \rangle \) must be positive because all \( j = l + \frac{1}{2} \) states are reachable by applying the \( J^- \) operator successively to \( | j = l + \frac{1}{2}, m = l + \frac{1}{2} \rangle \), and the matrix elements of \( J^- \) are always positive by convention. So the \( 2 \times 2 \) transformation matrix (3.7.61) can be only

\[
\begin{pmatrix}
\sqrt{\frac{l + m + \frac{1}{2}}{2l + 1}} & i \sqrt{\frac{l - m + \frac{1}{2}}{2l + 1}} \\
-i \sqrt{\frac{l - m + \frac{1}{2}}{2l + 1}} & \sqrt{\frac{l + m + \frac{1}{2}}{2l + 1}}
\end{pmatrix}
\] (3.7.63)

We define spin-angular functions in two-component form as follows:

\[
\phi_j^{l \pm 1/2, m} = \pm \sqrt{\frac{l + m + \frac{1}{2}}{2l + 1}} Y_l^{m-1/2}(\theta, \phi) \chi_\pm
\] (3.7.64)

\[
= \frac{1}{\sqrt{2l + 1}} \left( \pm \sqrt{\frac{l + m + \frac{1}{2}}{2l + 1}} Y_l^{m+1/2}(\theta, \phi) \right).
\] (3.7.65)

They are, by construction, simultaneous eigenfunctions of \( L^2 \), \( S^2 \), \( J^2 \), and \( J_z \). They are also eigenfunctions of \( L \cdot S \) but \( L \cdot S \), being just

\[
L \cdot S = (\frac{1}{2}) (J^2 - L^2 - S^2),
\] (3.7.66)

is not independent. Indeed, its eigenvalue can easily be computed as follows:

\[
\left( \frac{h^2}{2} \right) \left[ j(j + 1) - l(l + 1) - \frac{3}{4} \right] = \left( \frac{(l + 1)h^2}{2} \right) \text{ for } j = l + \frac{1}{2},
\]

\[
\left( \frac{(l - 1)h^2}{2} \right) \text{ for } j = l - \frac{1}{2}.
\] (3.7.66)

Clebsch-Gordan Coefficients and Rotation Matrices

Angular-momentum addition may be discussed from the point of view of rotation matrices. Consider the rotation operator \( \mathcal{O}_{j(l)}(R) \) in the ket space spanned by the angular-momentum eigenkets with eigenvalue \( j \). Likewise, consider \( \mathcal{O}_{j(l)}(R) \). The product \( \mathcal{O}_{j(l)} \otimes \mathcal{O}_{j(l)} \) is reducible in the sense that after suitable choice of base kets, its matrix representation can take the following form:

\[
\begin{pmatrix}
\phi_{j+1}^{l(l)} & \phi_{j-1}^{l(l)} \\
\phi_{j+1}^{l(l)} & \phi_{j-1}^{l(l)}
\end{pmatrix}
\] (3.7.67)
In the notation of group theory this is written as
\[ \mathcal{D}^{(j_1)} \mathcal{D}^{(j_2)} = \mathcal{D}^{(j_1 + j_2)} \mathcal{D}^{(j_1 + j_2 - 1)} \cdots \mathcal{D}^{(j_1 - j_2)}. \] (3.7.66)

In terms of the elements of rotation matrices, we have an important expansion known as the Clebsch-Gordan series:
\[ \mathcal{D}_{m_1 m_2}^{(j_1)}(R) \mathcal{D}_{m'_2}^{(j_2)}(R) = \sum_j \sum_{m m'} \langle j_1 j_2; m_1 m_2 | j_1 j_2; jm \rangle \times \langle j_1 j_2; m'_1 m'_2 | j'_1 j'_2; jm' \rangle \mathcal{D}^{(j_1)}_{m m_1}(R), \] (3.7.69)
where the \( j \)-sum runs from \( |j_1 - j_2| \) to \( j_1 + j_2 \). The proof of this equation follows. First, note that the left-hand side of (3.69) is the same as
\[ \langle j_1 j_2; m_1 m_2 | \mathcal{D}(R) | j_1 j_2; m'_1 m'_2 \rangle = \langle j_1 j_2; m_1 m_2 | \mathcal{D}(R) | j_1 j_2; m'_2 | \mathcal{D}(R) | j_1 j_2; m'_1 \rangle \times \langle j_1 j_2; m'_2 | \mathcal{D}(R) | j_1 j_2; jm' \rangle \delta_{jj'}, \] (3.7.70)
But the same matrix element is also computable by inserting a complete set of states in the \((j, m)\) basis. Thus
\[ \langle j_1 j_2; m_1 m_2 | \mathcal{D}(R) | j_1 j_2; m'_2 \rangle = \sum_j \sum_{m} \sum_{j'} \sum_{m'} \langle j_1 j_2; m_1 m_2 | j_1 j_2; jm \rangle \langle j_1 j_2; jm | \mathcal{D}(R) | j_1 j_2; j'm' \rangle \times \langle j_1 j_2; j'm' | j_1 j_2; m'_2 \rangle \] (3.7.71)
which is just the right-hand side of (3.69).

As an interesting application of (3.69), we derive an important formula for an integral involving three spherical harmonics. First, recall the connection between \( Y_{l_1}^{m_1}(\theta, \phi) \) and \( Y_{l_2}^{m_2}(\theta, \phi) \) given by (3.6.52). Letting \( l_1 \to l_1', l_2 \to l_2', m_1 \to 0, m_2 \to 0 \) (hence \( m' \to 0 \)) in (3.6.59), we obtain, after some complex conjugation,
\[ Y_{l_1}^{m_1}(\theta, \phi) Y_{l_2}^{m_2}(\theta, \phi) = \frac{(2l_1 + 1)(2l_2 + 1)}{4\pi} \sum_{l' m'} \langle l_1 l_2; m_1 m_2 | l_1' l_2'; l' m' \rangle \times \langle l_1' l_2'; 00 | l_1 l_2; 00 \rangle \sqrt{\frac{4\pi}{2l' + 1}} Y_{l'}^{m'}(\theta, \phi). \] (3.7.72)

We multiply both sides by \( Y_{l_1}^{m_1}(\theta, \phi) \) and integrate over solid angles. The summations drop out because of the orthogonality of spherical harmonics, and we are left with
\[ \int d\Omega Y_{l_1}^{m_1}(\theta, \phi) Y_{l_2}^{m_2}(\theta, \phi) = \sqrt{\frac{(2l_1 + 1)(2l_2 + 1)}{4\pi(2l + 1)}} \langle l_1 l_2; 00 | l_1 l_2; 00 \rangle \langle l_1 l_2; m_1 m_2 | l_1 l_2; bn \rangle. \] (3.7.73)

The square root factor times the first Clebsch-Gordan coefficient is independent of orientations; that is, of \( m_1 \) and \( m_2 \). The second Clebsch-Gordan coefficient is the one appropriate for adding \( l_1 \) and \( l_2 \) to obtain total \( l \).

Equation (3.7.73) turns out to be a special case of the Wigner-Eckart theorem to be derived in Section 3.10. This formula is extremely useful in evaluating multipole matrix elements in atomic and nuclear spectroscopy.

3.8. Schwinger's Oscillator Model of Angular Momentum

Angular Momentum and Uncoupled Oscillators

There exists a very interesting connection between the algebra of angular momentum and the algebra of two independent (that is, uncoupled) oscillators, which was worked out in J. Schwinger's notes [see Quantum Theory of Angular Momentum, edited by L. C. Biedenharn and H. Van Dam, Academic Press (1965), p. 229]. Let us consider two simple harmonic oscillators, which we call the plus type and the minus type. We have the annihilation and creation operators, denoted by \( a_+ \) and \( a_+^{\dagger} \) for the plus-type oscillator; likewise, we have \( a_- \) and \( a_-^{\dagger} \) for the minus-type oscillators. We also define the number operators \( N_+ \) and \( N_- \) as follows:
\[ N_+ = a_+^{\dagger} a_+, \quad N_- = a_-^{\dagger} a_- \] (3.8.1)

We assume that the usual commutation relations among \( a, a^{\dagger} \), and \( N \) hold for oscillators of the same type (see Section 2.3).
\[ [a_+, a_+] = 1, \quad [a_-, a_-] = 1, \] (3.8.2a)
\[ [N_+, a_+] = -a_+, \quad [N_-, a_-] = -a_-. \] (3.8.2b)
\[ [N_+, a_+] = a_+^{\dagger}, \quad [N_-, a_-] = a_-^{\dagger}. \] (3.8.2c)

However, we assume that any pair of operators between different oscillators commute:
\[ [a_+, a_-^{\dagger}] = [a_-, a_+^{\dagger}] = 0 \] (3.8.3)
and so forth. So it is in this sense that we say the two oscillators are uncoupled.
Because \( N_+ \) and \( N_- \) commute by virtue of (3.8.3), we can build simultaneous eigenkets of \( N_+ \) and \( N_- \) with eigenvalues \( n_+ \) and \( n_- \) respectively. So we have the following eigenvalue equations for \( N_\pm \):

\[
N_+|n_+,n_-\rangle = n_+|n_+,n_-\rangle, \quad N_-|n_+,n_-\rangle = n_-|n_+,n_-\rangle.
\]

(3.8.5)

In complete analogy with (2.3.16) and (2.3.17), the creation and annihilation operators, \( a^\dagger_\pm \) and \( a_\pm \), act on \( |n_+,n_-\rangle \) as follows:

\[
a^\dagger_+|n_+,n_-\rangle = \sqrt{n_+ + 1}|n_+ + 1,n_-\rangle, \quad a^\dagger_-|n_+,n_-\rangle = \sqrt{n_- + 1}|n_+,n_- + 1\rangle,
\]

\[
a_+|n_+,n_-\rangle = \sqrt{n_+ - 1}|n_+,n_- - 1\rangle, \quad a_-|n_+,n_-\rangle = \sqrt{n_- - 1}|n_+,n_- - 1\rangle.
\]

(3.8.6)

We can obtain the most general eigenkets of \( N_+ \) and \( N_- \) by applying \( a_\pm \) and \( a^\dagger_\pm \) successively to the vacuum ket defined by

\[
a_+|0,0\rangle = 0, \quad a_-|0,0\rangle = 0.
\]

(3.8.7)

In this way we obtain

\[
|n_+,n_-\rangle = \left( \frac{(a^\dagger_+)^n a^n_-}{\sqrt{n_+!n_-!}} \right)|0,0\rangle.
\]

(3.8.8)

Next, we define

\[
J_+ = ha^\dagger_-, \quad J_- = ha^\dagger_+ a_+,
\]

(3.8.9a)

and

\[
J_z = \left( \frac{\hbar}{2} \right)(a^\dagger_+ a_+ - a^\dagger_- a_-) = \left( \frac{\hbar}{2} \right)(N_+ - N_-).
\]

(3.8.9b)

We can readily prove that these operators satisfy the angular-momentum commutation relations of the usual form

\[
[J_-,J_+] = \pm \hbar J_z,
\]

\[
[J_+,J_-] = 2\hbar J_z.
\]

(3.8.9)

For example, we prove (3.8.9b) as follows:

\[
\hbar^2[a^\dagger_+ a_- a^\dagger_- a_+] = \hbar^2 a^\dagger_+ a_- a^\dagger_- a_+ - \hbar^2 a^\dagger_- a_+ a^\dagger_+ a_-
\]

\[
= \hbar^2(a^\dagger_+ a_- a^\dagger_- a_+ - a^\dagger_- a_+ a^\dagger_+ a_-) = 2\hbar J_z.
\]

(3.8.10)

Defining the total \( N \) to be

\[
N = N_+ + N_- = a^\dagger_+ a_+ + a^\dagger_- a_-.
\]

(3.8.11)

We can also prove

\[
J^2 = J_+^2 + \left( \frac{1}{2} \right)(J_+ J_- + J_- J_+)
\]

\[
= \left( \frac{\hbar^2}{2} \right) N \left( \frac{N}{2} + 1 \right),
\]

(3.8.12)

which is left as an exercise.

What are the physical interpretations of all this? We associate spin up \((m = \frac{1}{2})\) with one quantum unit of the plus-type oscillator and spin down \((m = -\frac{1}{2})\) with one quantum unit of the minus-type oscillator. If you like, you may imagine one spin \(\frac{1}{2}\) "particle" with spin up (down) with each quantum unit of the plus- (minus-) type oscillator. The eigenvalues \(n_+ \) and \(n_- \) are just the numbers of spins up and spins down, respectively. The meaning of \(J_+ \) is that it destroys one unit of spin down with the \(z\)-component of spin-angular momentum \(-\hbar/2\) and creates one unit of spin up with the \(z\)-component of spin-angular momentum \(+\hbar/2\); the \(z\)-component of angular momentum is therefore increased by \(\hbar\). Likewise \(J_- \) destroys one unit of spin up and creates one unit of spin down; the \(z\)-component of angular momentum is therefore decreased by \(\hbar\). As for the \(J_z \) operator, it simply counts \(\hbar/2\) times the difference of \(n_+ \) and \(n_- \), just the \(z\)-component of the total angular momentum. With (3.8.5) at our disposal we can easily examine how \(J_\pm \) and \(J_z \) act on \( |n_+,n_-\rangle \) as follows:

\[
J_+|n_+,n_-\rangle = \hbar a^\dagger_-|n_+,n_-\rangle = \sqrt{n_-(n_+ + 1)} \hbar |n_+,n_- + 1\rangle,
\]

(3.8.13a)

\[
J_-|n_+,n_-\rangle = \hbar a^\dagger_+|n_+,n_-\rangle = \sqrt{n_-(n_+ + 1)} \hbar |n_+,n_- - 1\rangle,
\]

(3.8.13b)

\[
J_z|n_+,n_-\rangle = \left( \frac{\hbar}{2} \right)(N_+ - N_-)|n_+,n_-\rangle = \left( \frac{1}{2} \right)(n_+ - n_-) \hbar |n_+,n_-\rangle.
\]

(3.8.13c)

Notice that in all these operations, the sum \(n_+ + n_-\), which corresponds to the total number of spin \(\frac{1}{2}\) particles remains unchanged.

Observe now that (3.8.13a), (3.8.13b), and (3.8.13c) reduce to the familiar expressions for the \(J_\pm \) and \(J_z \) operators we derived in Section 3.5, provided we substitute

\[
n_+ \rightarrow j + m, \quad n_- \rightarrow j - m.
\]

(3.8.14)

The square root factors in (3.8.13a) and (3.8.13b) change to

\[
\sqrt{n_-(n_+ + 1)} \rightarrow \sqrt{(j - m)(j + m + 1)},
\]

\[
\sqrt{n_+(n_- + 1)} \rightarrow \sqrt{(j + m)(j - m + 1)},
\]

(3.8.15)

which are exactly the square root factors appearing in (3.5.39) and (3.5.41).
Notice also that the eigenvalue of the $J^2$ operator defined by (3.8.12) changes as follows:

$$\left(\frac{\hbar^2}{2}\right)(n_+ + n_-) \left[\frac{(n_+ + n_-)}{2} + 1\right] \rightarrow \hbar^2(j + 1). \quad (3.8.16)$$

All this may not be too surprising because we have already proved that $J_\pm$ and $J^2$ operators we constructed out of the oscillator operators satisfy the usual angular-momentum commutation relations. But it is instructive to see in an explicit manner the connection between the oscillator matrix elements and angular-momentum matrix elements. In any case, it is now natural to use

$$j = \frac{(n_+ + n_-)}{2}, \quad m = \frac{(n_+ - n_-)}{2} \quad (3.8.17),$$

in place of $n_+$ and $n_-$ to characterize simultaneous eigenkets of $J^2$ and $J_z$. According to (3.8.13a) the action of $J_-$ changes $n_-$ into $n_- + 1$, $n_-$ into $n_- - 1$, which means that $j$ is unchanged and $m$ goes into $m + 1$. Likewise, we see that the $J_+$ operator that changes $n_-$ into $n_- - 1$; $n_-$ into $n_- + 1$ lowers $m$ by one unit without changing $j$. We can now write as (3.8.7) for the most general $N_+$, $N_-$ eigenkets

$$|j, m\rangle = \frac{(a_+^*)^j(a^*_+)^m}{\sqrt{(j + m)! (j - m)!}} |0\rangle, \quad (3.8.18)$$

where we have used $|0\rangle$ for the vacuum ket, earlier denoted by $|0, 0\rangle$.

A special case of (3.8.18) is of interest. Let us set $m = j$, which physically means that the eigenvalue of $J_z$ is as large as possible for a given $j$. We have

$$|j, j\rangle = \frac{(a_+^*)^{2j}}{\sqrt{(2j)!}} |0\rangle. \quad (3.8.19)$$

We can imagine this state to be built up of $2j$ spin $\frac{1}{2}$ particles with their spins all pointing in the positive $z$-direction.

In general, we note that a complicated object of high $j$ can be visualized as being made up of primitive spin $\frac{1}{2}$ particles, $j + m$ of them with spin up and the remaining $j - m$ of them with spin down. This picture is extremely convenient even though we obviously cannot always regard an object of angular momentum $j$ literally as a composite system of spin $\frac{1}{2}$ particles. All we are saying is that, as far as the transformation properties under rotations are concerned, we can visualize any object of angular momentum $j$ as a composite system of $2j$ spin $\frac{1}{2}$ particles formed in the manner indicated by (3.8.18).

From the point of view of angular-momentum addition developed in the previous section, we can add the spins of $2j$ spin $\frac{1}{2}$ particles to obtain states with angular momentum $j$, $j - 1$, $j - 2$, ..., as a simple example, we can add the spin-angular momenta of two spin $\frac{1}{2}$ particles to obtain a total angular momentum of zero as well as one. In Schwinger's oscillator scheme, however, we obtain only states with angular momentum $j$ when we start with $2j$ spin $\frac{1}{2}$ particles. In the language of permutation symmetry to be developed in Chapter 6, only totally symmetrical states are constructed by this method. The primitive spin $\frac{1}{2}$ particles appearing here are actually bosons! This method is quite adequate if our purpose is to examine the properties under rotations of states characterized by $j$ and $m$ without asking how such states are built up initially.

The reader who is familiar with isospin in nuclear and particle physics may note that what we are doing here provides a new insight into the isospin (or isotopic spin) formalism. The operator $J_0$ that destroys one unit of the minus type and creates one unit of the plus type is completely analogous to the isospin ladder operator $I_0$ (sometimes denoted by $I_+^-$) that annihilates a neutron (isospin down) and creates a proton (isospin up), thus raising the $z$-component of isospin by one unit. In contrast, $J_+$ is analogous to $T_+$, which simply counts the difference between the number of protons and neutrons in nuclei.

### Explicit Formula for Rotation Matrices

Schwinger's scheme can be used to derive, in a very simple way, a closed formula for rotation matrices, first obtained by E. P. Wigner using a similar (but not identical) method. We apply the rotation operator $\mathcal{D}(R)$ to $|j, m\rangle$, written as (3.8.18). In the Euler angle notation the only nontrivial rotation is the second one about the $y$-axis, so we direct our attention to

$$\mathcal{D}(R) = \mathcal{D}(\alpha, \beta, \gamma)|_{\alpha = \gamma = 0} = \exp\left(\frac{-ij\beta}{\hbar}\right). \quad (3.8.20)$$

We have

$$\mathcal{D}(R)|j, m\rangle = \left[\mathcal{D}(R)a_+^\dag \mathcal{D}^{-1}(R)\right]^{j + m}\left[\mathcal{D}(R)a^\dag \mathcal{D}^{-1}(R)\right]^{-m}\mathcal{D}(R)|0\rangle. \quad (3.8.21)$$

Now, $\mathcal{D}(R)$ acting on $|0\rangle$ just reproduces $|0\rangle$ because, by virtue of (3.8.6), only the leading term, 1, in the expansion of exponential (3.8.20) contributes. So

$$\mathcal{D}(R)a_+^\dag \mathcal{D}^{-1}(R) = \exp\left(\frac{-ij\beta}{\hbar}\right) a_+^\dag \exp\left(\frac{ij\beta}{\hbar}\right). \quad (3.8.22)$$
Thus we may use formula (2.3.47). Letting
\[ G \rightarrow \frac{-J_y}{\hbar}, \quad \lambda \rightarrow \beta \]
(3.8.23)
in (2.3.47), we realize that we must look at various commutators, namely,
\[ \left[ \frac{-J_y}{\hbar}, a_+^{\dagger} \right] = \left( \frac{1}{2i} \right) \left[ a_+^{\dagger}, a_+ \right] = \left( \frac{1}{2i} \right) a_+^{\dagger}, \]
(3.8.24)
and so forth. Clearly, we always obtain either \( a_+^{\dagger} \) or \( a_+ \). Collecting terms, we get
\[ \mathcal{D}(R) a_+^{\dagger} \mathcal{D}^{-1}(R) = a_+^{\dagger} \cos \left( \frac{\beta}{2} \right) + a_+ \sin \left( \frac{\beta}{2} \right). \]
(3.8.25)
Likewise,
\[ \mathcal{D}(R) a_+ \mathcal{D}^{-1}(R) = a_+ \cos \left( \frac{\beta}{2} \right) - a_+^{\dagger} \sin \left( \frac{\beta}{2} \right). \]
(3.8.26)
Actually this result is not surprising. After all, the basic spin-up state is supposed to transform as
\[ a_+ \left| 0 \right> \rightarrow \cos \left( \frac{\beta}{2} \right) a_+ \left| 0 \right> + \sin \left( \frac{\beta}{2} \right) a_+^{\dagger} \left| 0 \right>, \]
(3.8.27)
under a rotation about the \( y \)-axis. Substituting (3.8.25) and (3.8.26) into (3.8.21) and recalling the binomial theorem
\[ (x+y)^N = \sum_k \binom{N}{k} x^{N-k} y^k, \]
(3.8.28)
we obtain
\[ \mathcal{D}(\alpha = 0, \beta, \gamma = 0 | j, m) = \sum_k \frac{(j+m)! (j-m)!}{(j+k-m+k!)(j-m-k)!} \times \left[ a_+ \cos \left( \frac{\beta}{2} \right) \right]^{j+m-k} \left[ a_+^{\dagger} \sin \left( \frac{\beta}{2} \right) \right]^k \]
\[ \times \left[ -a_+^{\dagger} \sin \left( \frac{\beta}{2} \right) \right]^{j-m-k} \left[ a_+ \cos \left( \frac{\beta}{2} \right) \right]^0, \]
(3.8.29)
We may compare (3.8.29) with
\[ \mathcal{D}(\alpha = 0, \beta, \gamma = 0 | j, m) = \sum_{m'} d_{m_m m}^{(j)}(\beta) \]
\[ = \sum_{m'} d_{m_m m}^{(j)}(\beta) \frac{a_+^{\dagger} \left[ a_+ \right]^{j+m'}}{\sqrt{(j+m')!(j-m')!}} \left| 0 \right>, \]
(3.8.30)
3.9. Spin Correlation Measurements and Bell's Inequality

We can obtain an explicit form for \( d_{m_m m}^{(j)}(\beta) \) by equating the coefficients of powers of \( a_+^{\dagger} \) in (3.8.29) and (3.8.30). Specifically, we want to compare \( a_+^{\dagger} \) raised to \( j + m' \) in (3.8.30) with \( a_+^{\dagger} \) raised to \( 2j - k - l \), so we identify
\[ l = j - k - m'. \]
(3.8.31)
We are seeking \( d_{m_m m}^{(j)}(\beta) \) with \( m' \) fixed. The \( k \)-sum and the \( l \)-sum in (3.8.29) are not independent of each other; we eliminate \( l \) in favor of \( k \) by taking advantage of (3.8.31). As for the powers of \( a_+^{\dagger} \), we note that \( a_+^{\dagger} \) raised to \( j - m' \) in (3.8.30) automatically matches with \( a_+^{\dagger} \) raised to \( k + l \) in (3.8.29) when (3.8.31) is imposed. The last step is to identify the exponents of \( \cos(\beta/2), \sin(\beta/2) \), and \(-1\), which are, respectively,
\[ j + m - k + l = 2j - 2k + m - m', \]
(3.8.32a)
\[ k + j - m = 2k - m + m', \]
(3.8.32b)
\[ j - m - l = k - m + m'. \]
(3.8.32c)
where we have used (3.8.31) to eliminate \( l \). In this way we obtain Wigner's formula for \( d_{m_m m}^{(j)}(\beta) \):
\[ d_{m_m m}^{(j)}(\beta) = \sum_k (-1)^{k - m - m'} \sqrt{(j+m)!(j-m)!(j+m')!(j-m')!} \]
\[ \times \left( \frac{\cos \left( \frac{\beta}{2} \right)}{\sin \left( \frac{\beta}{2} \right)} \right)^{j-k+m'-k}, \]
(3.8.33)
where we take the sum over \( k \) whenever none of the arguments of factorials in the denominator are negative.

3.9. SPIN CORRELATION MEASUREMENTS AND BELL'S INEQUALITY

Correlations in Spin-Singlet States

The simplest example of angular-momentum addition we encountered in Section 3.7 was concerned with a composite system made up of spin \( \frac{1}{2} \) particles. In this section we use such a system to illustrate one of the most astonishing consequences of quantum mechanics.

Consider a two-electron system in a spin-singlet state, that is, with a total spin of zero. We have already seen that the state ket can be written as [see (3.7.15d)]
\[ \left| \text{spin singlet} \right> = \left( \frac{1}{\sqrt{2}} \right) (|2+; 2-\rangle - |2-; 2+\rangle), \]
(3.9.1)
where we have explicitly indicated the quantization direction. Recall that
\(|\hat{\mathbf{z}}^+; \hat{\mathbf{z}}-\rangle\) means that electron 1 is in the spin-up state and electron 2 is in the spin-down state. The same is true for \(|\hat{\mathbf{z}}^-; \hat{\mathbf{z}}^+\rangle\).

Suppose we make a measurement on the spin component of one of the electrons. Clearly, there is a 50-50 chance of getting either up or down because the composite system may be in \(|\hat{\mathbf{z}}^+; \hat{\mathbf{z}}-\rangle\) or \(|\hat{\mathbf{z}}^-; \hat{\mathbf{z}}^+\rangle\) with equal probabilities. But if one of the components is shown to be in the spin-up state, the other is necessarily in the spin-down state, and vice versa. When the spin component of electron 1 is shown to be up, the measurement apparatus has selected the first term, \(|\hat{\mathbf{z}}^+; \hat{\mathbf{z}}-\rangle\) of (3.9.1); a subsequent measurement of the spin component of electron 2 must ascertain that the state ket of the composite system is given by \(|\hat{\mathbf{z}}^-; \hat{\mathbf{z}}^+\rangle\).

It is remarkable that this kind of correlation can persist even if the two particles are well separated and have ceased to interact provided that as they fly apart, there is no change in their spin states. This is certainly the case for a \(J = 0\) system disintegrating spontaneously into two spin \(\frac{1}{2}\) particles with no relative orbital angular momentum, because angular-momentum conservation must hold in the disintegration process. An example of this would be a rare decay of the \(\eta\) meson (mass 549 MeV/c\(^2\)) into a muon pair

\[\eta \rightarrow \mu^+ + \mu^-\]  
(3.9.2)

which, unfortunately, has a branching ratio of only approximately \(6 \times 10^{-6}\). More realistically, in proton-proton scattering at low kinetic energies, the Pauli principle to be discussed in Chapter 6 forces the interacting protons to be in \(^1S_0\) (orbital angular momentum 0, spin-singlet state), and the spin states of the scattered protons must be correlated in the manner indicated by (3.9.1) even after they get separated by a macroscopic distance.

To be more pictorial we consider a system of two spin \(\frac{1}{2}\) particles moving in opposite directions, as in Figure 38. Observer A specializes in measuring \(S_z\) of particle 1 (flying to the right), while observer B specializes in measuring \(S_z\) of particle 2 (flying to the left). To be specific, let us assume that observer A finds \(S_z\) to be positive for particle 1. Then he or she can predict, even before B performs any measurement, the outcome of B’s measurement with certainty: B must find \(S_z\) to be negative for particle 2. On the other hand, if A makes no measurement, B has a 50-50 chance of getting \(S_z^+\) or \(S_z^-\).

This by itself might not be so peculiar. One may say, “It is just like an urn known to contain one black ball and one white ball. When we blindly pick one of them, there is a 50-50 chance of getting black or white. But if the first ball we pick is black, then we can predict with certainty that the second ball will be white.”

It turns out that this analogy is too simple. The actual quantum-mechanical situation is far more sophisticated than that! This is because observers may choose to measure \(S_z\) in place of \(S_x\). The same pair of

![Figure 3.8. Spin correlation in a spin-singlet state.](image)

“quantum-mechanical balls” can be analyzed either in terms of black and white or in terms of blue and red!

Recall now that for a single spin \(\frac{1}{2}\) system the \(S_x\) eigenkets and \(S_z\) eigenkets are related as follows:

\[|\mathbf{S}_{\pm}\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\hat{\mathbf{z}}^+\rangle \pm |\hat{\mathbf{z}}^-\rangle), \quad |\hat{\mathbf{z}}\pm\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\mathbf{S}^+\rangle \pm |\mathbf{S}^-\rangle).\]  
(3.9.3)

Returning now to our composite system, we can rewrite spin-singlet ket (3.9.1) by choosing the \(x\)-direction as the axis of quantization:

\[|\text{spin singlet}\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\hat{\mathbf{z}}^-; \hat{\mathbf{z}}^+\rangle - |\hat{\mathbf{z}}^+; \hat{\mathbf{z}}^-\rangle).\]  
(3.9.4)

Apart from the overall sign, which in any case is a matter of convention, we could have guessed this form directly from (3.9.1) because spin-singlet states have no preferred direction in space. Let us now suppose that observer A can choose to measure \(S_x\) or \(S_z\) of particle 1 by changing the orientation of his or her spin analyzer, while observer B always specializes in measuring \(S_z\) of particle 2. If A determines \(S_x\) of particle 1 to be positive, B clearly has a 50-50 chance for getting \(S_z^+\) or \(S_z^-\); even though \(S_x\) of particle 2 is known to be negative with certainty, its \(S_z\) is completely undetermined. On the other hand, let us suppose that A also chooses to measure \(S_z\); if observer A determines \(S_z\) of particle 1 to be positive, then without fail, observer B will measure \(S_z\) of particle 2 to be negative. Finally, if A chooses to make no measurement, B, of course, will have a 50-50 chance of getting \(S_z^+\) or \(S_z^-\). To sum up:

1. If A measures \(S_x\) and B measures \(S_z\), there is a completely random correlation between the two measurements.
2. If A measures \(S_z\) and B measures \(S_z\), there is a 100% (opposite sign) correlation between the two measurements.
3. If A makes no measurement, B’s measurements show random results.

Table 3.1 shows all possible results of such measurements when B and A are allowed to choose to measure \(S_x\) or \(S_z\).
These considerations show that the outcome of B's measurement depends on what kind of measurement A decides to perform: an measurement, an Sz measurement, or no measurement. Notice again that A and B can be miles apart with no possibility of communications or physical interaction. Observer A can decide how to orient his or her analyser apparatus long after the two particles have separated. It is as though particle 2 "knows" which spin component of particle 1 is being measured.

The orthodox quantum-mechanical interpretation of this situation is as follows. A measurement is a selection (or filtration) process. When Sz of particle 1 is measured to be positive, then component \( |z^+; z^- \rangle \) is selected. Subsequent measurement of the other particle's Sz merely ascertains that the system is still in \( |z^+; z^- \rangle \). We must accept that a measurement on what appears to be a part of the system is to be regarded as a measurement on the whole system.

### Einstein's Locality Principle and Bell's Inequality

Many physicists have felt uncomfortable with the preceding orthodox interpretation of spin-correlation measurements. Their feelings are illustrated in the following frequently quoted remarks by A. Einstein, which call **Einstein's locality principle**: "But on one supposition we should, in opinion, absolutely hold fast: The real factual situation of the system Sz is independent of what is done with the system Sz, which is spatially separated from the former." Because this problem was first discussed in a 1935 paper of A. Einstein, B. Podolsky, and N. Rosen, it is sometimes known as the Einstein-Podolsky-Rosen paradox.*

Some have argued that the difficulties encountered here are inherent in the probabilistic interpretations of quantum mechanics and that the dynamic behavior at the microscopic level appears probabilistic only because some yet unknown parameters—so-called hidden variables—have not been specified. It is not our purpose here to discuss various alternatives to quantum mechanics based on hidden-variable or other considerations. Rather, let us ask, Do such theories make predictions different from those of quantum mechanics? Until 1964, it could be thought that the alternative theories could be concocted in such a way that they would give no predictions, other than the usual quantum-mechanical predictions, that could be verified experimentally. The whole debate would have belonged to the realm of metaphysics rather than physics. It was then pointed out by J. S. Bell that the alternative theories based on Einstein's locality principle actually predict a testable inequality relation among the observables of spin-correlation experiments that disagrees with the predictions of quantum mechanics.

We derive Bell's inequality within the framework of a simple model, conceived by E. P. Wigner, that incorporates the essential features of the various alternative theories. Proponents of this model agree that it is impossible to determine Sz and Sz simultaneously. However, when we have a large number of spin ½ particles, we assign a certain fraction of them to have the following property:

**If Sz is measured, we obtain a plus sign with certainty.**

**If Sz is measured, we obtain a minus sign with certainty.**

A particle satisfying this property is said to belong to type \( |z^+; z^- \rangle \). Notice that we are not asserting that we can simultaneously measure Sz and Sz to be \(+\) and \(-\), respectively. When we measure Sz, we do not measure Sz, and vice versa. We are assigning definite values of spin components in more than one direction with the understanding that only one or the other of the components can actually be measured. Even though this approach is fundamentally different from that of quantum mechanics, the quantum-mechanical predictions for Sz and Sz measurements performed on the spin-up (Sz+) state are reproduced provided there are as many particles belonging to type \( |z^+; z^+ \rangle \) as to type \( |z^+; z^- \rangle \).

Let us now examine how this model can account for the results of spin-correlation measurements made on composite spin-singlet systems.

---

*To be historically accurate, the original Einstein-Podolsky-Rosen paper dealt with measurements of \( x \) and \( y \). The use of composite spin ½ systems to illustrate the Einstein-Podolsky-Rosen paradox started with D. Bohm.
Clearly, for a particular pair, there must be a perfect matching between particle 1 and particle 2 to ensure zero total angular momentum: If particle 1 is of type $\ell +, \ell -$, then particle 2 must belong to type $\ell -, \ell +$, and so forth. The results of correlation measurements, such as in Table 3.1, can be reproduced if particle 1 and particle 2 are matched as follows:

\[
\text{particle 1} \quad \text{particle 2} \\
(\ell +, \ell -) \leftrightarrow (\ell -, \ell +), \quad (3.9.5a) \\
(\ell +, \ell +) \leftrightarrow (\ell -, \ell -), \quad (3.9.5b) \\
(\ell -, \ell +) \leftrightarrow (\ell +, \ell -), \quad (3.9.5c) \\
(\ell -, \ell -) \leftrightarrow (\ell +, \ell +). \quad (3.9.5d)
\]

with equal populations, that is, 25% each. A very important assumption is implied here. Suppose a particular pair belongs to type (3.9.5a) and observer A decides to measure $S_\ell$ of particle 1; then he or she necessarily obtains a plus sign regardless of whether B decides to measure $S_\ell$ or $S_\ell$. It is in this sense that Einstein's locality principle is incorporated in this model: A's result is predetermined independently of B's choice as to what to measure.

In the examples considered so far, this model has been successful in reproducing the predictions of quantum mechanics. We now consider more-complicated situations where the model leads to predictions different from the usual quantum-mechanical predictions. This time we start with three unit vectors $\hat{a}$, $\hat{b}$, and $\hat{c}$, which are, in general, not mutually orthogonal. We imagine that one of the particles belongs to some definite type, say $(\ell -, \ell +, \ell +)$, which means that if $S_\ell \hat{a}$ is measured, we obtain a minus sign with certainty; if $S_\ell \hat{b}$ is measured, we obtain a plus sign with certainty; if $S_\ell \hat{c}$ is measured, we obtain a plus with certainty. Again there must be a perfect matching in the sense that the other particle necessarily belongs to type $(\ell +, \ell -, \ell -)$ to ensure zero total angular momentum. In any given event, the particle pair in question must be a member of one of the eight types shown in Table 3.2. These eight possibilities are mutually exclusive and disjoint. The population of each type is indicated in the first column.

Let us suppose that observer A finds $S_\ell \hat{a}$ to be plus and observer B finds $S_\ell \hat{b}$ to be plus also. It is clear from Table 3.2 that the pair belong to either type 3 or type 4, so the number of particle pairs for which this situation is realized is $N_3 + N_4$. Because $N_i$ is positive semidefinite, we must have inequality relations like

\[
N_3 + N_4 \leq (N_2 + N_4) + (N_3 + N_7). \quad (3.9.6)
\]

Let $P(\hat{a}^+;\hat{b}^+)$ be the probability that, in a random selection, observer A measures $S_\ell \hat{a}$ to be $+$ and observer B measures $S_\ell \hat{b}$ to be $+$, and so on.

<table>
<thead>
<tr>
<th>Population</th>
<th>Particle 1</th>
<th>Particle 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_1$</td>
<td>$(\hat{a}^+, \hat{b}^+, \hat{c}^+)$</td>
<td>$(\hat{a}^-, \hat{b}^-, \hat{c}^-)$</td>
</tr>
<tr>
<td>$N_2$</td>
<td>$(\hat{a}^+, \hat{b}^+, \hat{c}^-)$</td>
<td>$(\hat{a}^-, \hat{b}^-, \hat{c}^+)$</td>
</tr>
<tr>
<td>$N_3$</td>
<td>$(\hat{a}^+, \hat{b}^-, \hat{c}^+)$</td>
<td>$(\hat{a}^-, \hat{b}^+, \hat{c}^-)$</td>
</tr>
<tr>
<td>$N_4$</td>
<td>$(\hat{a}^-, \hat{b}^+, \hat{c}^-)$</td>
<td>$(\hat{a}^+, \hat{b}^-, \hat{c}^+)$</td>
</tr>
<tr>
<td>$N_5$</td>
<td>$(\hat{a}^-, \hat{b}^-, \hat{c}^+)$</td>
<td>$(\hat{a}^+, \hat{b}^+, \hat{c}^-)$</td>
</tr>
<tr>
<td>$N_6$</td>
<td>$(\hat{a}^-, \hat{b}^-, \hat{c}^-)$</td>
<td>$(\hat{a}^+, \hat{b}^+, \hat{c}^+)$</td>
</tr>
<tr>
<td>$N_7$</td>
<td>$(\hat{a}^-, \hat{b}^+, \hat{c}^-)$</td>
<td>$(\hat{a}^+, \hat{b}^-, \hat{c}^+)$</td>
</tr>
</tbody>
</table>

Clearly, we have

\[
P(\hat{a}^+;\hat{b}^+) = \frac{(N_1 + N_4)}{\Sigma_i N_i}. \quad (3.9.7)
\]

In a similar manner, we obtain

\[
P(\hat{a}^+;\hat{c}^+) = \frac{(N_2 + N_4)}{\Sigma_i N_i} \quad \text{and} \quad P(\hat{c}^+;\hat{b}^+) = \frac{(N_3 + N_7)}{\Sigma_i N_i}. \quad (3.9.8)
\]

The positivity condition (3.9.6) now becomes

\[
P(\hat{a}^+;\hat{b}^+) \leq P(\hat{a}^+;\hat{c}^+) + P(\hat{c}^+;\hat{b}^+). \quad (3.9.9)
\]

This is Bell's inequality, which follows from Einstein's locality principle.

Quantum Mechanics and Bell's Inequality

We now return to the world of quantum mechanics. In quantum mechanics we do not talk about a certain fraction of particle pairs, say $N_i/\Sigma_i N_i$, belonging to type 3. Instead, we characterize all spin-singlet systems by the same ket (3.9.1); in the language of Section 3.4 we are concerned here with a pure ensemble. Using this ket and the rules of quantum mechanics we have developed, we can unambiguously calculate each of the three terms in inequality (3.9.9).

We first evaluate $P(\hat{a}^+;\hat{b}^+)$. Suppose observer A finds $S_\ell \hat{a}$ to be positive; because of the 100% (opposite sign) correlation we discussed earlier, B's measurement of $S_\ell \hat{a}$ will yield a minus sign with certainty. But to calculate $P(\hat{a}^+;\hat{b}^+)$ we must consider a new quantization axis $\hat{b}$ that makes an angle $\theta_{ab}$ with $\hat{a}$; see Figure 3.9. According to the formalism of Section 3.2, the probability that the $S_\ell \hat{b}$ measurement yields $+$ when particle 2 is known to be in an eigenket of $S_\ell \hat{a}$ with negative eigenvalue is
Several experiments have been performed to test Bell’s inequality. In one of the experiments spin correlations between the final protons in low-energy proton-proton scattering were measured. In all other experiments photon-polarization correlations between a pair of photons in a cascade transition of an excited atom (Ca, Hg, …),

\[(J = 0) \frac{\gamma_0}{(J = 1)} \frac{\gamma_0}{(J = 0)}, \tag{3.9.16}\]

or in the decay of a positronium (an \(e^+e^-\) bound state in \(1S_0\)) were measured; studying photon-polarization correlations should be just as good in view of the analogy developed in Section 1.1:*

\[S_x \rightarrow \hat{e} \quad \text{in} \quad x\text{-direction,} \]
\[S_y \rightarrow \hat{e} \quad \text{in} \quad y\text{-direction,} \tag{3.9.17}\]
\[S_z \rightarrow \hat{e} \quad \text{in} \quad 45^\circ \text{ diagonal direction,} \]
\[S_z \rightarrow \hat{e} \quad \text{in} \quad 135^\circ \text{ diagonal direction.} \]

The results of all recent precision experiments have conclusively established that Bell’s inequality was violated, in one case by more than nine standard deviations. Furthermore, in all these experiments the inequality relation was violated in such a way that the quantum-mechanical predictions were fulfilled within error limits. In this controversy, quantum mechanics has triumphed with flying colors.

The fact that the quantum-mechanical predictions have been verified does not mean that the whole subject is now a triviality. Despite the experimental verdict we may still feel psychologically uncomfortable about many aspects of measurements of this kind. Consider in particular the following point: Right after observer A performs a measurement on particle 1, how does particle 2—which may, in principle, be many light years away from particle 1—get to “know” how to orient its spin so that the remarkable correlations apparent in Table 3.1 are realized? In one of the experiments to test Bell’s inequality (performed by A. Aspect and collaborators) the analyzer settings were changed so rapidly that A’s decision as to what to measure could not be made until it was too late for any kind of influence, traveling slower than light, to reach B.

We conclude this section by showing that despite these peculiarities we cannot use spin-correlation measurements to transmit any useful information between two macroscopically separated points. In particular, superluminal (faster than light) communications are impossible.

Suppose A and B both agree in advance to measure \(S_z\); then, without asking A, B knows precisely what A is getting. But this does not mean that

*It should be kept in mind here that by working with photons we are going outside the realm of nonrelativistic quantum mechanics, which is the subject of this book.
A and B are communicating; B just observes a random sequence of positive and negative signs. There is obviously no useful information contained in it. B verifies the remarkable correlations predicted by quantum mechanics only after he or she gets together with A and compares the notes (or computer sheets).

It might be thought that A and B can communicate if one of them suddenly changes the orientation of his or her analyzing apparatus. Let us suppose that A agrees initially to measure $S_x$ and $B, S_z$. The results of A’s measurements are completely uncorrelated with the results of B’s measurements, so there is no information transferred. But then, suppose A suddenly breaks his or her promise and without telling B starts measuring $S_z$. There are now complete correlations between A’s results and B’s results. However, B has no way of inferring that A has changed the orientation of his or her analyzer. B continues to see just a random sequence of ‘+’s and ‘−’s by looking at his or her own notebook only. So again there is no information transferred.

### 3.10. Tensor Operators

**Vector Operator**

We have been using notations such as $x$, $p$, $S$, and $L$, but as yet we have not systematically discussed their rotational properties. They are vector operators, but what are their properties under rotations? In this section we give a precise quantum-mechanical definition of vector operators based on their commutation relations with the angular-momentum operator. We then generalize to tensor operators with more-complicated transformation properties and derive an important theorem on the matrix elements of vector and tensor operators.

We all know that a vector in classical physics is a quantity with three components that transforms by definition like $V_i \rightarrow \sum R_{ij} V_j$ under a rotation. It is reasonable to demand that the expectation value of a vector operator $V$ in quantum mechanics be transformed like a classical vector under rotation. Specifically, as the state ket is changed under rotation according to

$$|\alpha\rangle \rightarrow \mathcal{D}(R)|\alpha\rangle,$$

the expectation value of $V$ is assumed to change as follows:

$$\langle \alpha|V|\alpha\rangle \rightarrow \langle \alpha|\mathcal{D}^*(R)V_\mathcal{D}(R)|\alpha\rangle = \sum_j R_{ij}\langle \alpha|V_j|\alpha\rangle.$$

This must be true for an arbitrary ket $|\alpha\rangle$. Therefore,

$$\mathcal{D}^*(R)V_i\mathcal{D}(R) = \sum_j R_{ij}V_j \quad (3.10.3)$$

must hold as an **operator equation**, where $R_{ij}$ is the $3 \times 3$ matrix that corresponds to rotation $R$.

Let us now consider a specific case, an infinitesimal rotation. When the rotation is infinitesimal, we have

$$\mathcal{D}(R) = 1 - \frac{i\epsilon J \cdot \hbar}{\hbar}.$$  

(3.10.4)

We can now write (3.10.3) as

$$V_i + \frac{\epsilon}{i\hbar}[V_i, J \cdot \hbar] = \sum_j R_{ij}(\hbar; \epsilon)V_j.$$  

(3.10.5)

In particular, for $\hbar$ along the $z$-axis, we have

$$R(\xi; \epsilon) = \begin{pmatrix} 1 & -\epsilon & 0 \\ \epsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$  

(3.10.6)

so

1. $i = 1$: $V_x + \frac{\epsilon}{i\hbar}[V_x, J_z] = V_x - \epsilon V_y$

(3.10.7a)

2. $i = 2$: $V_y + \frac{\epsilon}{i\hbar}[V_y, J_z] = \epsilon V_x + V_y$

(3.10.7b)

3. $i = 3$: $V_z + \frac{\epsilon}{i\hbar}[V_z, J_z] = V_z$

(3.10.7c)

This means that $V$ must satisfy the commutation relations

$$[V_i, J_j] = i\epsilon_{ijk} \hbar V_k.$$  

(3.10.8)

Clearly, the behavior of $V$ under a **finite** rotation is completely determined by the preceding commutation relations; we just apply the by-now familiar formula (2.3.47) to

$$\exp\left(\frac{iJ \cdot \phi}{\hbar}\right)V_i\exp\left(-\frac{-iJ \cdot \phi}{\hbar}\right).$$  

(3.10.9)

We simply need to calculate

$$[J_i, [J_i, \cdots [J_i, V_i], \cdots]].$$  

(3.10.10)

Multiple commutators keep on giving back to us $V_i$ or $V_k$ ($k \neq i, j$) as in spin case (3.2.7).

We can use (3.10.8) as the defining property of a vector operator. Notice that the angular-momentum commutation relations are a special case of (3.10.8) in which we let $V_i \rightarrow J_i, V_k \rightarrow J_k$. Other special cases are $[y, L_z] = -i\hbar x, [x, L_z] = -i\hbar y, [p_x, L_z] = -i\hbar p_y, [p_y, L_z] = i\hbar p_x$; these can be proved explicitly.
Cartesian Tensors Versus Irreducible Tensors

In classical physics it is customary to define a tensor \( T_{ijk} \ldots \) by generalizing \( V_i \rightarrow \sum_j R_{ij} V_j \) as follows:

\[
T_{ijk} \ldots \rightarrow \sum_{i'} \sum_{j'} \sum_{k'} \cdots R_{ij} R_{j'k'} \cdots T_{ij'k'} \ldots \tag{3.10.11}
\]

under a rotation specified by the \( 3 \times 3 \) orthogonal matrix \( R \). The number of indices is called the rank of a tensor. Such a tensor is known as a Cartesian tensor.

The simplest example of a Cartesian tensor of rank 2 is a dyad formed out of two vectors \( U \) and \( V \). One simply takes a Cartesian component of \( U \) and a Cartesian component of \( V \) and puts them together:

\[
T_{ij} = U_i V_j. \tag{3.10.12}
\]

Notice that we have nine components altogether. They obviously transform like (3.10.11) under rotation.

The trouble with a Cartesian tensor like (3.10.12) is that it is reducible—that is, it can be decomposed into objects that transform differently under rotations. Specifically, for the dyadic in (3.10.12) we have

\[
U_i V_j = \frac{U \cdot V}{3} \delta_{ij} + \frac{1}{2} \left( U_i V_j - U_j V_i \right) - \frac{U \cdot V}{3} \delta_{ij}. \tag{3.10.13}
\]

The first term on the right-hand side, \( U \cdot V \), is a scalar product invariant under rotation. The second is an antisymmetric tensor which can be written as vector product \( \epsilon_{ijk} (U \times V)_k \). There are altogether 3 independent components. The last is a \( 3 \times 3 \) symmetric traceless tensor with 5 \((= 6 - 1)\), where 1 comes from the traceless condition) independent components. The number of independent components checks:

\[
3 \times 3 = 1 + 3 + 5. \tag{3.10.14}
\]

We note that the numbers appearing on the right-hand side of (3.10.14) are precisely the multiplicities of objects with angular momentum \( l = 0 \), \( l = 1 \), and \( l = 2 \), respectively. This suggests that the dyadic has been decomposed into tensors that can transform like spherical harmonics with \( l = 0, 1, \) and 2. In fact, (3.10.13) is the simplest nontrivial example to illustrate the reduction of a Cartesian tensor into irreducible spherical tensors.

Before presenting the precise definition of a spherical tensor, we first give an example of a spherical tensor of rank \( k \). Suppose we take a spherical harmonic \( Y^m_l(\theta, \phi) \). We have already seen that it can be written as \( Y^m_l(\hat{n}) \), where the orientation of \( \hat{n} \) is characterized by \( \theta \) and \( \phi \). We now replace \( \hat{n} \) by some vector \( V \). The result is that we have a spherical tensor of rank \( k \) (in place of \( l \)) with magnetic quantum number \( q \) (in place of \( m \)), namely,

\[
T_q^{(k)} = Y^m_{-k} a(V). \tag{3.10.15}
\]

Specifically, in the case \( k = 1 \) we take spherical harmonics with \( l = 1 \) and replace \((z/r) = (\hat{n})\) by \( V \), and so on.

\[
Y^0_1 = \sqrt{3 \over 4 \pi} \cos \theta = \sqrt{3 \over 4 \pi} {z \over r} \rightarrow T^0_1 = \sqrt{3 \over 4 \pi} V_z,
\]

\[
Y^\pm_1 = \mp \sqrt{3 \over 4 \pi} {x \pm iy \over \sqrt{2} r} \rightarrow T^{\pm}_1 = \sqrt{3 \over 4 \pi} \left( \mp {V_x \pm iV_y \over \sqrt{2}} \right).
\]

Obviously this can be generalized for higher \( k \), for example,

\[
Y^\pm_2 = \sqrt{15 \over 32 \pi} {x \pm iy \over r^2} \rightarrow T^{\pm}_2 = \sqrt{15 \over 32 \pi} \left( V_x \pm iV_y \right)^2.
\]

\( T_q^{(k)} \) are irreducible, just as \( Y^m_l \) are. For this reason, working with spherical tensors is more satisfactory than working with Cartesian tensors.

To see the transformation of spherical tensors constructed in this manner, let us first review how \( Y^m_l \) transform under rotations. First, we have for the direction eigenket;

\[
|\hat{n}\rangle \rightarrow \mathcal{D}(R) |\hat{n}\rangle = |\hat{n}'\rangle,
\]

which defines the rotated eigenket \( |\hat{n}'\rangle \). We wish to examine how \( Y^m_l(\hat{n}') = \langle \hat{n}' | l, m \rangle \) would look in terms of \( Y^m_l(\hat{n}) \). We can easily see this by starting with

\[
\mathcal{D}(R^{-1}) |l, m\rangle = \sum_{m'} |l, m\rangle \mathcal{D}_{m m'}^{(l)} (R^{-1}) \tag{3.10.19}
\]

and contracting with \( \langle \hat{n}| \) on the left, using (3.10.18).

\[
Y^m_l(\hat{n}') = \sum_{m'} Y^m_l(\hat{n}) \mathcal{D}_{m' m}^{(l)} (R^{-1}). \tag{3.10.20}
\]

If there is an operator that acts like \( Y^m_l(V) \), it is then reasonable to expect

\[
\mathcal{D}^*(R) Y^m_l(V) \mathcal{D}(R) = \sum_{m'} Y^m_{l'}(V) \mathcal{D}_{m' m}^{(l')} (R), \tag{3.10.21}
\]

where we have used the unitarity of the rotation operator to rewrite \( \mathcal{D}_{m m'}^{(l)} (R^{-1}) \).

All this work is just to motivate the definition of a spherical tensor. We now consider spherical tensors in quantum mechanics. Motivated by (3.10.21) we define a spherical tensor operator of rank \( k \) with \((2k + 1)\) components as

\[
\mathcal{D}^*(R) T_q^{(k)}(R) = \sum_{q'} \mathcal{D}_{q q'}^{(k)}(R) T_q^{(k)} \tag{3.10.22a}
\]
or, equivalently,
\[
\mathcal{D}(R)T_q^{(k)}\mathcal{D}(R) = \sum_{q' = -k}^{k} \mathcal{D}^{(k)}_{q'q}(R)T_q^{(k)}. \tag{3.10.22b}
\]

This definition holds regardless of whether \(T_q^{(k)}\) can be written as \(Y_{q-k}^{m}(V)\); for example, \((U_x + iU_y)V_x + iV_y\) is the \(q = +2\) component of a spherical tensor of rank 2 even though, unlike \((V_x + iV_y)^2\), it cannot be written as \(Y_k(V)\).

A more convenient definition of a spherical tensor is obtained by considering the infinitesimal form of (3.10.22b), namely,
\[
\left(1 + \frac{i\mathbf{J} \cdot \mathbf{n}}{h}\right)T_q^{(k)} \left(1 - \frac{i\mathbf{J} \cdot \mathbf{n}}{h}\right) = \sum_{q' = -k}^{k} T_q^{(k)}(kq')\left(1 + \frac{i\mathbf{J} \cdot \mathbf{n}}{h}\right)(kq')
\]

or
\[
[\mathbf{J} \cdot \mathbf{n}, T_q^{(k)}] = \sum_{q'} T_q^{(k)}(kq')\left[\mathbf{J} \cdot \mathbf{n}\right](kq'). \tag{3.10.23}
\]

By taking \(\mathbf{n}\) in the \(\hat{z}\)- and in the \((\hat{x} + i\hat{y})\) directions and using the nonvanishing matrix elements of \(J_z\) and \(J_\pm\) [see (3.5.35b) and (3.5.41)], we obtain
\[
\left[J_z, T_q^{(k)}\right] = hqT_q^{(k)} \tag{3.10.25a}
\]

and
\[
\left[J_+, T_q^{(k)}\right] = h\left(k + 1\right)\left(k + q + 1\right)T_q^{(k+1)}. \tag{3.10.25b}
\]

These commutation relations can be considered as a definition of spherical tensors in place of (3.10.22).

**Product of Tensors**

We have seen how to form a scalar, vector (or antisymmetric tensor), and a traceless symmetric tensor out of two vectors using the Cartesian tensor language. Of course, spherical tensor language can also be used (Baym 1969, Chapter 17), for example,
\[
T_0^{(0)} = -\frac{U \cdot V}{3} = \frac{U_xV_{-1} + U_yV_{+1} - U_0V_0}{3},
\]
\[
T_1^{(1)} = \frac{(U \times V)_q}{\sqrt{2}},
\]
\[
T_1^{(2)} = U_{+1}V_{-1},
\]
\[
T_1^{(2)} = \frac{U_{-1}V_0 + U_0V_{+1}}{\sqrt{2}},
\]
\[
T_0^{(2)} = \frac{U_{+1}V_{-1} + 2U_0V_0 + U_{-1}V_{+1}}{\sqrt{6}},
\]
where \(U_q(V_q)\) is the \(q\)th component of a spherical tensor of rank 1, corresponding to vector \(U(V)\). The preceding transformation properties can be checked by comparing with \(Y_{q-m}\) and remembering that \(U_{+1} = -(U_x + iU_y)/\sqrt{2}, U_{-1} = (U_x - iU_y)/\sqrt{2}\), \(U_0 = U_z\). A similar check can be made for \(V_{\pm1,0}\). For instance,
\[
Y_2^0 = \sqrt{\frac{5}{16\pi}} \frac{3z^2 - r^2}{r^2},
\]
where \(3z^2 - r^2\) can be written as
\[
2z^2 + 2\left[-\frac{(x + iy)}{\sqrt{2}}\left(x - iy\right)\right],
\]
hence, \(Y_2^0\) is just a special case of \(T_0^{(2)}\) for \(U = V = r\).

A more systematic way of forming tensor products goes as follows. We start by stating a theorem:

**Theorem.** Let \(X_{q_1}^{(k_1)}\) and \(Z_{q_2}^{(k_2)}\) be irreducible spherical tensors of rank \(k_1\) and \(k_2\), respectively. Then
\[
T_q^{(k)} = \sum_{q_1, q_2} \langle k_1 k_2; q_1 q_2|k_1 k_2, qk \rangle X_{q_1}^{(k_1)}Z_{q_2}^{(k_2)} \tag{3.10.27}
\]
is a spherical (irreducible) tensor of rank \(k\).

**Proof.** We must show that under rotation \(T_q^{(k)}\) must transform according to (3.10.22)
\[
\mathcal{D}(R)T_q^{(k)}\mathcal{D}(R) = \sum_{q_1, q_2} \langle k_1 k_2; q_1 q_2|k_1 k_2, qk \rangle \times \mathcal{D}(R)X_{q_1}^{(k_1)}\mathcal{D}(R) \times \mathcal{D}(R)Z_{q_2}^{(k_2)}\mathcal{D}(R) = \sum_{q_1, q_2, q_1', q_2'} \langle k_1 k_2; q_1 q_2|k_1 k_2, qk \rangle \times X_{q_1}^{(k_1)}\mathcal{D}(R)(R^{-1})Z_{q_2}^{(k_2)}\mathcal{D}(R^{-1})(R^{-1})
\]
where we have used the Clebsch-Gordan series formula (3.7.69). The preceding expression becomes
\[
= \sum_{k'' q_1 q_2 q'''} \sum_{q_1 q_2 q_1'} \sum_{q_1 q_2} \delta_{k' k'' q'''} \langle k_1 k_2; q_1 q_2|k_1 k_2, k' q''; \mathcal{D}(R)(R^{-1})X_{q_1}^{(k_1)}Z_{q_2}^{(k_2)} \rangle,
where we have used the orthogonality of Clebsch-Gordan coefficients (3.7.42). Finally, this expression reduces to
\[
\sum_{q'} \left( \sum_{q_1} \sum_{q_2} \langle k_1 k_2; q_1' q_2'| k_1 k_2; q' \rangle X_{q_1'}^{(k_1)} Z_{q_2'}^{(k_2)} \right) D_{q_1 q_2'}^{(k)}(R^{-1})
\]
\[
= \sum_{q'} T_q^{(k)} D_{q q'}^{(k)}(R^{-1}) = \sum_{q'} D_{q q'}^{(k)}(R) T_q^{(k)}
\]

The foregoing shows how we can construct tensor operators of higher or lower ranks by multiplying two tensor operators. Furthermore, the manner in which we construct tensor products out of two tensors is completely analogous to the manner in which we construct an angular-momentum eigenstate by adding two angular momentums; exactly the same Clebsch-Gordan coefficients appear if we let \(k_1,2 \rightarrow j_{1,2}, q_1,2 \rightarrow m_{1,2}\).

**Matrix Elements of Tensor Operators; the Wigner-Eckart Theorem**

In considering the interactions of an electromagnetic field with atoms and nuclei, it is often necessary to evaluate matrix elements of tensor operators with respect to angular-momentum eigenstates. Examples of this will be given in Chapter 5. In general, it is a formidable dynamic task to calculate such matrix elements. However, there are some properties of these matrix elements that follow purely from kinematic or geometric considerations, which we now discuss.

First, there is a very simple \(m\)-selection rule:

- **\(m\)-selection Rule**
  \[
  \langle \alpha', j' m'| T_q^{(k)} | \alpha, j m \rangle = 0, \quad \text{unless } m' = q + m. \tag{3.10.28}
  \]

**Proof.** Using (3.10.25a), we have
\[
\langle \alpha', j' m'| \left[ J_z, T_q^{(k)} \right] | \alpha, j m \rangle = \left[ (m' - m) h - hq \right] \times \langle \alpha', j' m'| T_q^{(k)} | \alpha, j m \rangle = 0;
\]

hence,
\[
\langle \alpha', j' m'| T_q^{(k)} | \alpha, j m \rangle = 0 \quad \text{unless } m' = q + m. \tag{3.10.29}
\]

Another way to see this is to note that transformation property of \(T_q^{(k)} | \alpha, j m \rangle\) under rotation, namely,
\[
\mathcal{D} T_q^{(k)} | \alpha, j m \rangle = \mathcal{D} T_q^{(k)} \mathcal{D}^\dagger | \alpha, j m \rangle. \tag{3.10.29a}
\]

If we now let \(\mathcal{D}\) stand for a rotation operator around the \(z\)-axis, we get [see (3.10.22b) and (3.1.16)]
\[
\mathcal{D} (\hat{z}, \phi) T_q^{(k)} | \alpha, j m \rangle = e^{-i\phi q} e^{-i\phi q} T_q^{(k)} | \alpha, j m \rangle. \tag{3.10.30}
\]

which is orthogonal to \(| \alpha', j' m' \rangle\) unless \(q + m = m'\).

We are going to prove one of the most important theorems in quantum mechanics, the **Wigner-Eckart theorem.**

**The Wigner-Eckart Theorem.** The matrix elements of tensor operators with respect to angular-momentum eigenstates satisfy
\[
\langle \alpha', j' m'| T_q^{(k)} | \alpha, j m \rangle = \langle j k; m q| j k; j' m' \rangle \langle \alpha' j' || T_q^{(k)} || \alpha j \rangle \frac{\sqrt{2j + 1}}{\sqrt{2j' + 1}}. \tag{3.10.31}
\]

where the **double-bar matrix element** is independent of \(m\) and \(m'\), and \(q\).

Before we present a proof of this theorem, let us look at its significance. First, we see that the matrix element is written as the product of two factors. The first factor is a Clebsch-Gordan coefficient for adding \(j\) and \(k\) to get \(j'\). It depends only on the geometry, that is, the way the system is oriented with respect to the \(z\)-axis. There is no reference whatsoever to the particular nature of the tensor operator. The second factor does depend on the dynamics, for instance, \(\alpha\) may stand for the radial quantum number and its evaluation may involve, for example, evaluation of radial integrals. On the other hand, it is completely independent of the magnetic quantum numbers \(m, m'\), and \(q\), which specify the orientation of the physical system. To evaluate \(\langle \alpha', j' m'| T_q^{(k)} | \alpha, j m \rangle\) with various combinations of \(m, m'\), and \(q\) it is sufficient to know just one of them; all others can be related geometrically because they are proportional to Clebsch-Gordan coefficients, which are known. The common proportionality factor is \(\langle \alpha' j' || T_q^{(k)} || \alpha \rangle\), which makes no reference whatsoever to the geometric features.

The selection rules for the tensor operator matrix element can be immediately read off from the selection rules for adding angular momentum. Indeed, from the requirement that the Clebsch-Gordan coefficient be nonvanishing, we immediately obtain the \(m\)-selection rule (3.10.28) derived before and also the triangular relation
\[
| j - k | \leq j' \leq j + k. \tag{3.10.32}
\]

Now we prove the theorem.

**Proof.** Using (3.10.25b) we have
\[
\langle \alpha', j' m'| \left[ J_\pm, T_q^{(k)} \right] | \alpha, j m \rangle = h \sqrt{(k + q)(k \pm q + 1)} < \alpha', j' m'| T_q^{(k)} | \alpha, j m \rangle. \tag{3.10.33}
\]
or using (3.5.39) and (3.5.40) we have
\[
\sqrt{(j' + m')(j' + m' + 1)} \langle \alpha', j', m' + 1 | T_q^{(k)} | \alpha, jm \rangle \\
= \sqrt{(j + m)(j + m + 1)} \langle \alpha', j'm' | T_q^{(k)} | \alpha, j, m + 1 \rangle \\
+ \sqrt{(k + q)(k + q + 1)} \langle \alpha', j'm' | T_q^{(k)} | \alpha, jm \rangle.
\] (3.10.34)

Compare this with the recursion relation for the Clebsch-Gordan coefficient (3.7.49). Note the striking similarity if we substitute \( j' \rightarrow j, m' \rightarrow m, j \rightarrow j_1, m \rightarrow m_1, k \rightarrow j_2, \) and \( q \rightarrow m_2 \). Both recursion relations are of the form \( \sum a_{ij} x_j = 0, \) that is, first-order linear homogeneous equations with the same coefficients \( a_{ij} \). Whenever we have
\[
\sum_j a_{ij} x_j = 0, \quad \sum_j a_{ij} y_j = 0,
\] (3.10.35)
we cannot solve for the \( x_j \) (or \( y_j \)) individually but we can solve for the ratios; so
\[
\frac{x_j}{x_k} = \frac{y_j}{y_k} \quad \text{or} \quad x_j = cy_j,
\] (3.10.36)
where \( c \) is a universal proportionality factor. Noting that \( \langle j_1 j_2; m_1, m_2 | j_1 j_2; jm \rangle \) in the Clebsch-Gordan recursion relation (3.7.49) corresponds to \( \langle \alpha', j'm' | T_q^{(k)} | \alpha, jm \rangle \), we see that
\[
\langle \alpha', j'm' | T_q^{(k)} | \alpha, jm \rangle = (\text{universal proportionality constant independent of} \ m, q, \text{and} \ m') \langle jk; m q \pm 1 | jk; j'm' \rangle,
\] (3.10.37)
which proves the theorem.

Let us now look at two simple examples of the Wigner-Eckart theorem.

Example 1. Tensor of rank 0, that is, scalar \( T_0^{(0)} = S \). The matrix element of a scalar operator satisfies
\[
\langle \alpha', j'm' | S | \alpha, jm \rangle = \delta_{jj'} \delta_{mm'} \frac{\langle \alpha | \alpha' \rangle \langle j'm' | j m \rangle}{\sqrt{2j + 1}}
\] (3.10.38)
because \( S \) acting on \( | \alpha, jm \rangle \) is like adding an angular momentum of zero. Thus the scalar operator cannot change \( j, m \) values.

Example 2. Vector operator which in the spherical tensor language is a rank 1 tensor. The spherical component of \( V \) can be written as \( V_q = \pm 1,0 \), so we have the selection rule
\[
\Delta m = m' - m = \pm 1,0 \quad \Delta j = j' - j = \begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}.
\] (3.10.39)

In addition, the \( 0 \rightarrow 0 \) transition is forbidden. This selection rule is of fundamental importance in the theory of radiation; it is the dipole selection rule obtained in the long-wavelength limit of emitted photons.*

For \( j = j' \) the Wigner-Eckart theorem—when applied to the vector operator—takes a particularly simple form, often known as the projection theorem for obvious reasons.

The Projection Theorem
\[
\langle \alpha', jm | V_{q} | \alpha, jm \rangle = \frac{\langle \alpha', jm | J \cdot V | \alpha, jm \rangle}{\hbar^2 (j + 1)} \langle jm | J_q | jm \rangle,
\] (3.10.40)
where analogous to our discussion after (3.10.26) we choose
\[
J_{+1} = \mp \frac{1}{\sqrt{2}} (J_x \pm J_y) = \frac{1}{\sqrt{2}} J_z, \quad J_0 = J_z.
\] (3.10.41)

Proof. Noting (3.10.26) we have
\[
\langle \alpha', jm | J \cdot V | \alpha, jm \rangle = \langle \alpha', jm | (J_0 V_0 - J_{+1} V_{-1} - J_{-1} V_{+1}) | \alpha, jm \rangle
\]
\[
= m \hbar \langle \alpha', jm | V_0 | \alpha, jm \rangle + \frac{\hbar}{\sqrt{2}} \sqrt{(j + m)(j - m + 1)} \langle \alpha', j m - 1 | V_{-1} | \alpha, jm \rangle
\]
\[
- \frac{\hbar}{\sqrt{2}} \sqrt{(j - m)(j + m + 1)} \langle \alpha', j m + 1 | V_{+1} | \alpha, jm \rangle
\]
\[
= c_{jm} \langle \alpha | \alpha' \rangle \langle j m | j m \rangle
\] (3.10.42)
by the Wigner-Eckart theorem (3.10.31), where \( c_{jm} \) is independent of \( \alpha, \alpha' \), and \( V \), and the matrix elements of \( V_0, \pm 1 \) are all proportional to the double-bar matrix element (sometimes also called the reduced matrix element). Furthermore, \( c_{jm} \) is independent of \( m \) because \( J \cdot V \) is a scalar operator, so we may as well write it as \( c_j \). Because \( c_j \) does not depend on \( V \), (3.10.42) holds even if we let \( V \rightarrow J \) and \( \alpha' \rightarrow \alpha \), that is,
\[
\langle \alpha, jm | J^2 | \alpha, jm \rangle = c_j \langle \alpha | J^2 | \alpha \rangle.
\] (3.10.43)

Returning to the Wigner-Eckart theorem applied to \( V_q \) and \( J_q \), we have
\[
\frac{\langle \alpha', jm | V_{q} | \alpha, jm \rangle}{\langle \alpha, jm | J_{q} | \alpha, jm \rangle} = \frac{\langle \alpha' | J | \alpha \rangle}{\langle \alpha | J | \alpha \rangle},
\] (3.10.44)

*Additional parity selection rules are discussed in Chapter 4, Section 2. They lead to these I1 dipole selection rules.
1. Find the eigenvalues and eigenvectors of $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. Suppose an electron is in the spin state $\left( \alpha \right)_\beta$. If $s_y$ is measured, what is the probability of the result $\hbar / 2$?

2. Consider the $2 \times 2$ matrix defined by

$$U = \frac{a_0 + i\sigma \cdot a}{a_0 - i\sigma \cdot a},$$

where $a_0$ is a real number and $a$ is a three-dimensional vector with real components.

a. Prove that $U$ is unitary and unimodular.

b. In general, a $2 \times 2$ unimodular matrix represents a rotation in three dimensions. Find the axis and angle of rotation appropriate for $U$ in terms of $a_0$, $a_1$, $a_2$, and $a_3$.

3. The spin-dependent Hamiltonian of an electron-positron system in the presence of a uniform magnetic field in the $z$-direction can be written as

$$H = AS^{(e)}_z S^{(e)} + \left( \frac{eB}{mc} \right) \left( S^{(e)}_z - S^{(e')}_z \right).$$

Suppose the spin function of the system is given by $\chi^{(e)}_\alpha \chi^{(e')}_\beta$.

a. Is this an eigenfunction of $H$ in the limit $A \to 0, eB/mc \to 0$? If it is, what is the energy eigenvalue? If it is not, what is the expectation value of $H$?

b. Same problem when $eB/mc \to 0$, $A \neq 0$.

4. Consider a spin 1 particle. Evaluate the matrix elements of

$$S_z(S_z + \hbar)(S_z - \hbar) \quad \text{and} \quad S_z(S_z + \hbar)(S_z - \hbar).$$

5. Let the Hamiltonian of a rigid body be

$$H = \frac{1}{2} \left( \frac{K_1^2}{I_1} + \frac{K_2^2}{I_2} + \frac{K_3^2}{I_3} \right),$$

where $K$ is the angular momentum in the body frame. From this expression obtain the Heisenberg equation of motion for $K$ and then find Euler’s equation of motion in the correspondence limit.

6. Let $U = e^{iG_eG_3G_y}$, where $(\alpha, \beta, \gamma)$ are the Eulerian angles. In order that $U$ represent a rotation $(\alpha, \beta, \gamma)$, what are the commutation rules satisfied by the $G_e$? Relate $G$ to the angular momentum operators.

7. What is the meaning of the following equation:

$$U^{-1} A U = \sum R_k A_k,$$

where the three components of $A$ are matrices? From this equation show that matrix elements $\langle m | A_k | n \rangle$ transform like vectors.

8. Consider a sequence of Euler rotations represented by

$$\varphi^{(1/2)}(\alpha, \beta, \gamma) = \exp \left( -i \frac{\sigma_\alpha}{2} \right) \exp \left( -i \frac{\sigma_\beta}{2} \right) \exp \left( -i \frac{\sigma_\gamma}{2} \right)$$

$$\cdot \begin{pmatrix} e^{-i(\alpha + \gamma)/2} \cos \frac{\beta}{2} - e^{-i(\alpha - \gamma)/2} \sin \frac{\beta}{2} \\ e^{i(\alpha - \gamma)/2} \sin \frac{\beta}{2} e^{i(\alpha + \gamma)/2} \cos \frac{\beta}{2} \end{pmatrix}.$$
13. Show that the $3 \times 3$ matrices $G_i$ ($i = 1, 2, 3$) whose elements are given by
\[ (G_i)_{jk} = -i\hbar \delta_{ij} \delta_{jk}, \]
where $j$ and $k$ are the row and column indices, satisfy the angular momentum commutation relations. What is the physical (or geometric) significance of the transformation matrix that connects $G_i$ to the more usual $3 \times 3$ representations of the angular-momentum operator $J_i$ with $J_3$ taken to be diagonal? Relate your result to
\[ V \rightarrow V + \hbar \delta \phi \times V \]
under infinitesimal rotations. (Note: This problem may be helpful in understanding the photon spin.)

14. a. Let $\mathbf{J}$ be angular momentum. It may stand for orbital $\mathbf{L}$, spin $\mathbf{S}$, or $\mathbf{J}_{\text{total}}$. Using the fact that $J_x, J_y, J_z$ satisfy the usual angular-momentum commutation relations, prove
\[ J^2 = J_z^2 + J_+ J_- - \hbar J_z. \]
b. Using (a) (or otherwise), derive the "famous" expression for the coefficient $c_-$ that appears in
\[ J_- \psi_{jm} = c_- \psi_{j, m-1}. \]
15. The wave function of a particle subjected to a spherically symmetrical potential $V(r)$ is given by
\[ \psi(x) = (x + y + 3z)f(r). \]
a. Is $\psi$ an eigenfunction of $L^2$? If so, what is the $l$-value? If not, what are the possible values of $l$ we may obtain when $L^2$ is measured?
b. What are the probabilities for the particle to be found in various $m_l$ states?
c. Suppose it is known somehow that $\psi(x)$ is an energy eigenfunction with eigenvalue $E$. Indicate how we may find $V(r)$.
16. A particle in a spherically symmetrical potential is known to be in an eigenstate of $L^2$ and $L_z$, with eigenvalues $\hbar^2(l+1)$ and $m\hbar$, respectively. Prove that the expectation values between $|lm\rangle$ states satisfy
\[ \langle L_x \rangle = \langle L_y \rangle = 0, \quad \langle L_z^2 \rangle = \langle L_z \rangle = \frac{[l(l+1)\hbar^2 - m^2\hbar^2]}{2}. \]
Interpret this result semiclassically.
17. Suppose a half-integer $l$-value, say $\frac{1}{2}$, were allowed for orbital angular momentum. From
\[ L_z Y_{l\pm 1/2}(\theta, \phi) = 0, \]
we may deduce, as usual,
\[ Y_{l\pm 1/2}(\theta, \phi) \propto e^{i\phi/2} \sqrt{\sin \theta}. \]
18. Consider an orbital angular-momentum eigenstate $|l = 2, m = 0\rangle$. Suppose this state is rotated by an angle $\beta$ about the $y$-axis. Find the probability for the new state to be found in $m = 0$, $1$, and $2$. (The spherical harmonics for $l = 0, 1$, and 2 given in Appendix A may be useful.)
19. What is the physical significance of the operators
\[ K_+ = a^+ a_- \quad \text{and} \quad K_- = a_+ a^- \]
in Schwinger's scheme for angular momentum? Give the nonvanishing matrix elements of $K_\pm$.
20. We are to add angular momenta $j_1 = 1$ and $j_2 = 1$ to form $j = 2, 1, 0$ states. Using either the ladder operator method or the recursion relation, express all (nine) $(j, m)$ eigenkets in terms of $|j_1 j_2; m_1, m_2\rangle$. Write your answer as
\[ |j = 1, m = 1\rangle = \frac{1}{\sqrt{2}} |+, 0\rangle - \frac{1}{\sqrt{2}} |0, +\rangle, \ldots, \]
where $+$ and $0$ stand for $m_1, m_2 = 1, 0$, respectively.
21. a. Evaluate
\[ \sum_{m = -j}^{j} |d^{(j)}_{mm}(\beta)|^2 m \]
for any $j$ (integer or half-integer); then check your answer for $j = \frac{1}{2}$.
b. Prove, for any $j$,
\[ \sum_{m = -j}^{j} m^2 |d^{(j)}_{mm}(\beta)|^2 = \frac{1}{2} j(j + 1) \sin^2 \beta + m^2 \frac{1}{2} (3 \cos^2 \beta - 1). \]
[Hint: This can be proved in many ways. You may, for instance, examine the rotational properties of $J_z$ using the spherical (irreducible) tensor language.]
22. a. Consider a system with $j = 1$. Explicitly write
\[ \langle j = 1, m \mid J_z \mid j = 1, m\rangle \]
in $3 \times 3$ matrix form.
b. Show that for $j = 1$ only, it is legitimate to replace $e^{-i\beta J_z/\hbar}$ by
\[ 1 - i \left( \frac{J_z}{\hbar} \right) \sin \beta - \left( \frac{J_z}{\hbar} \right)^2 (1 - \cos \beta). \]
e. Using (b), prove
\[ d^{(j-1)}(\beta) = \]
\[
\begin{pmatrix}
\frac{1}{2}(1 + \cos \beta) & -\left(\frac{1}{\sqrt{2}}\right) \sin \beta & \frac{1}{2}(1 - \cos \beta)
\end{pmatrix}
\]
\[
\begin{pmatrix}
\frac{1}{\sqrt{2}} \sin \beta & \cos \beta & -\left(\frac{1}{\sqrt{2}}\right) \sin \beta
\end{pmatrix}
\]
\[
\begin{pmatrix}
\frac{1}{2}(1 - \cos \beta) & \frac{1}{\sqrt{2}} \sin \beta & \frac{1}{2}(1 + \cos \beta)
\end{pmatrix}
\]
Express the matrix element \( \langle \alpha_2|\beta_2\gamma_2|J^2|\alpha_1\beta_1\gamma_1 \rangle \) in terms of a series in \( \mathcal{D}_{mn}^{j}(a\beta\gamma) = \langle a\beta\gamma|jmn \rangle \).

Consider a system made up of two spin \( \frac{1}{2} \) particles. Observer A specializes in measuring the spin components of one of the particles \( S_{1z}, S_{1x} \) and so on, while observer B measures the spin components of the other particle. Suppose the system is known to be in a spin-singlet state, that is, \( S_{\text{total}} = 0 \).

a. What is the probability for observer A to obtain \( S_{1z} = \hbar/2 \) when observer B makes no measurement? Same problem for \( S_{1x} = \hbar/2 \).

b. Observer B determines the spin of particle 2 to be in the \( S_{2z} = \hbar/2 \) state with certainty. What can we then conclude about the outcome of observer A’s measurement if (i) A measures \( S_{1z} \) and (ii) A measures \( S_{1x} \)? Justify your answer.

Consider a spherical tensor of rank 1 (that is, a vector)
\[ V^{(1)} = \frac{V_x \pm iV_y}{\sqrt{2}} \quad V_0^{(1)} = V_z. \]

Using the expression for \( d^{(j-1)} \) given in Problem 22, evaluate
\[ \sum_q d^{(j-1)}(\beta) V_q^{(1)} \]
and show that your results are just what you expect from the transformation properties of \( V_{x,y,z} \) under rotations about the \( y \)-axis.

a. Construct a spherical tensor of rank 1 out of two different vectors \( U = (U_x, U_y, U_z) \) and \( V = (V_x, V_y, V_z) \). Explicitly write \( T_{\pm1,0}^{(1)} \) in terms of \( U_{x,y,z} \) and \( V_{x,y,z} \).

b. Construct a spherical tensor of rank 2 out of two different vectors \( U \) and \( V \). Write down explicitly \( T_{\pm1,0}^{(2)} \) in terms of \( U_{x,y,z} \) and \( V_{x,y,z} \).

Consider a spinless particle bound to a fixed center by a central force potential.

a. Relate, as much as possible, the matrix elements
\[ \langle n', l', m' | \pm \frac{1}{\sqrt{2}}(x \pm iy)|n, l, m \rangle \quad \text{and} \quad \langle n', l', m' | z|n, l, m \rangle \]
using only the Wigner-Eckart theorem. Make sure to state under what conditions the matrix elements are nonvanishing.

b. Do the same problem using wave functions \( \psi(x) = R_{n}(r)Y_{l}^{m}(\theta, \phi) \).

28. a. Write \( xy, xz, \) and \( (x^2 - y^2) \) as components of a spherical (irreducible) tensor of rank 2.

b. The expectation value
\[ Q = \langle \alpha, j, m = j(3z^2 - r^2)|\alpha, j, m = j \rangle \]
is known as the quadrupole moment. Evaluate
\[ \langle \alpha, j, m' = (x^2 - y^2)|\alpha, j, m = j \rangle \]
(\( m' = j, j - 1, j - 2, \ldots \)) in terms of \( Q \) and appropriate Clebsch-Gordan coefficients.

29. A spin \( \frac{1}{2} \) nucleus situated at the origin is subjected to an external inhomogeneous electric field. The basic electric quadrupole interaction may be taken to be
\[ H_{\text{int}} = \frac{eQ}{2s(s-1)\hbar^2} \left[ \left( \frac{\partial^2 \phi}{\partial x^2} \right)_0 S_x^2 + \left( \frac{\partial^2 \phi}{\partial y^2} \right)_0 S_y^2 + \left( \frac{\partial^2 \phi}{\partial z^2} \right)_0 S_z^2 \right], \]
where \( \phi \) is the electrostatic potential satisfying Laplace’s equation and the coordinate axes are so chosen that
\[ \left( \frac{\partial^2 \phi}{\partial x \partial y} \right)_0 = \left( \frac{\partial^2 \phi}{\partial y \partial z} \right)_0 = \left( \frac{\partial^2 \phi}{\partial z \partial x} \right)_0 = 0. \]

Show that the interaction energy can be written as
\[ A(3S_x^2 - S_y^2) + B(S_x^2 + S_y^2), \]
and express \( A \) and \( B \) in terms of \( (\partial^2 \phi / \partial x^2)_0 \) and so on. Determine the energy eigenkets (in terms of \( |m\rangle \), where \( m = \pm \frac{1}{2}, \pm \frac{3}{2} \)) and the corresponding energy eigenvalues. Is there any degeneracy?