14

Spin

14.1. Introduction

In this chapter we consider a class of quantum phenomena that cannot be handled by a straightforward application of the four postulates. The reason is that these phenomena involve a quantum degree of freedom called *spin*, which has no classical counterpart. Consequently, neither can we obtain the spin operator by turning to Postulate II, nor can we immediately write down the quantum Hamiltonian that governs its time evolution. The problem is very important, for most particles—electrons, protons, neutrons, photons—have the spin degree of freedom. Fortunately the problem can be solved by a shrewd mixture of classical intuition and reasoning by analogy. In this chapter we study just electron spin. The treatment of the spins of other particles is quite similar, with the exception of the photon, which moves at speed c and can't be treated nonrelativistically. Photon spin will be discussed in Chapter 18.

In the next three sections we address the following questions:

(1) What is the nature of this new spin degree of freedom?

(2) How is the Hilbert space modified to take this new degree of freedom into account? What do the spin operators look like in this space (kinematics of spin)?

(3) How does spin evolve with time, i.e., how does it enter the Hamiltonian (dynamics of spin)?

14.2. What is the Nature of Spin?

The best way to characterize spin is as a form of angular momentum. It is, however, not the angular momentum associated with the operator $\mathbf{L}$, as the following experiment shows. An electron is prepared in a state of zero linear momentum, i.e., in a state with a constant (space-independent) wave function. As the operators L_x, L_y, and L_z will give zero when acting on it, our existing formalism predicts that if the angular momentum along, say the z direction, is measured, a result of zero will obtain. The actual experiment, however, shows that this is wrong, that the result is

$\pm\hbar/2$.‡ It follows that the electron has "intrinsic" angular momentum, not associated with its orbital motion. This angular momentum is called *spin*, for it was imagined in the early days that if the electron has angular momentum without moving through space, then it must be spinning like a top. We adopt this nomenclature, but not the mechanical model that goes with it, for a consistent mechanical model doesn't exist. Fortunately one can describe spin and its dynamics without appealing to any model, starting with just the observed fact that it is a form of angular momentum. Let us now develop the formalism that deals with spin and, in particular, allows us to understand the above experiment.

14.3. Kinematics of Spin

The discussion following the general solution to the eigenvalue problem of angular momentum (Section 12.5) suggests the way for treating particles with intrinsic angular momentum or spin. Recall that if a particle is described by a wave function with many (n) components, the generator of infinitesimal rotation is not just $\mathbf{L}$ but something more. The reason is that under an infinitesimal rotation *two* things happen to the wave function: (1) the values at each spatial point are reassigned to the rotated point, and (2) the components of the wave function get transformed into linear combinations of each other.

The differential operator $\mathbf{L}$ does part (1), while an $n\times n$ matrix $\mathbf{S}$ is responsible for part (2).

By generalizing our findings from Exercise 12.5.1 to an n component wave function in three dimensions, we can say that under an infinitesimal rotation around the z axis, the wave function is transformed as follows:

$$\begin{bmatrix}\psi_1'\\ \vdots\\ \psi_n'\end{bmatrix}=\left(\begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}-\frac{i\varepsilon}{\hbar}\begin{bmatrix}-i\hbar\,\partial/\partial\phi & & 0\\ & \ddots & \\ 0 & & -i\hbar\,\partial/\partial\phi\end{bmatrix}-\frac{i\varepsilon}{\hbar}S_z\right)\begin{bmatrix}\psi_1\\ \vdots\\ \psi_n\end{bmatrix} \tag{14.3.1}$$

where S_z is an $n\times n$ matrix. In abstract form, this equation reads§

$$\begin{aligned}|\psi'\rangle&=\left[I-\frac{i\varepsilon}{\hbar}(L_z+S_z)\right]|\psi\rangle\\ &=\left[I-\frac{i\varepsilon}{\hbar}J_z\right]|\psi\rangle\end{aligned} \tag{14.3.2}$$

We identify J_z, the generator of infinitesimal rotations about the z axis, as the z component of angular momentum. We see it has two parts:

$$J_z=L_z+S_z$$

‡ In practice one measures not the angular momentum, but a related quantity called magnetic moment. More on this later. Also spin was first discovered on the basis of spectroscopic evidence and not from an experiment of the above type.

§ The spin operators will be denoted by the same symbol (S) whether they are referred to in the abstract or as matrices in some basis.

and more generally

$$\mathbf{J}=\mathbf{L}+\mathbf{S} \tag{14.3.3}$$

Our problem is to find the number (n) of components appropriate to the electron and the three spin matrices that rotate its components. We proceed as follows.

Since J_i are generators of rotations, they must obey the consistency condition

$$[J_i, J_j]=i\hbar\sum_k \varepsilon_{ijk}J_k \tag{14.3.4}$$

Since $\mathbf{L}$ and $\mathbf{S}$ act on different parts of the wave function (the former on x, y, z, the latter on the indices $i=1, \ldots, n$) they commute, and we may infer from Eq. (14.3.4) that

$$[L_i, L_j]+[S_i, S_j]=i\hbar\left[\sum_k \varepsilon_{ijk}L_k+\sum_k \varepsilon_{ijk}S_k\right] \tag{14.3.5}$$

Using the known commutation rules of the L_i, we deduce

$$[S_i, S_j]=i\hbar\sum_k \varepsilon_{ijk}S_k \tag{14.3.6}$$

Now recall that in Chapter 12 we found matrices J_x, J_y, and J_z [Eqs. (12.5.22)–(12.5.24)] that obey precisely these commutation relations. But these matrices were infinite dimensional. However, the infinite-dimensional matrices were built out of $(2j+1)\times(2j+1)$ blocks, with $j=0, 1/2, 1, 3/2, \ldots$, and the commutation relations were satisfied block by block. So which block shall we pick for the electron spin operators? The answer is given by the empirical fact that S_z has only the eigenvalues $\pm\hbar/2$. This singles out the 2×2 blocks in Eqs. (12.5.22)–(12.5.24):

$$S_x=\frac{\hbar}{2}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \qquad S_y=\frac{\hbar}{2}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}, \qquad S_z=\frac{\hbar}{2}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \tag{14.3.7}$$

Thus, the way to describe the electron is through a two-component wave function called a *spinor*:

$$\psi=\begin{bmatrix}\psi_+(x, y, z)\\ \psi_-(x, y, z)\end{bmatrix} \tag{14.3.8a}$$

$$\equiv\psi_+\begin{bmatrix}1\\ 0\end{bmatrix}+\psi_-\begin{bmatrix}0\\ 1\end{bmatrix} \tag{14.3.8b}$$

If $\psi_-=0$, $\psi_+\neq 0$, we have an eigenstate of S_z with eigenvalue $\hbar/2$; if $\psi_-\neq 0$, $\psi_+=0$, the S_z eigenvalue is $(-\hbar/2)$.

Let us now proceed to interpret the experiment mentioned earlier. Since we prepared a state of zero momentum, we want the operator **P** to give zero when acting on ψ. The operator **P** simply differentiates both components of ψ:

$$\mathbf{P} \rightarrow \begin{bmatrix} -i\hbar\nabla & 0 \\ 0 & -i\hbar\nabla \end{bmatrix} \tag{14.3.9}$$

We deduce from $\mathbf{P}|\psi\rangle = 0$, i.e.,

$$\begin{bmatrix} -i\hbar\nabla\psi_+ \\ -i\hbar\nabla\psi_- \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{14.3.10}$$

that ψ_+ and ψ_- are independent of x, y, and z. It follows that L_z acting on ψ gives zero. However, S_z doesn't: there is an amplitude $\psi_\pm$ for obtaining $\pm\hbar/2$.

The electron spinor is a *two*-component object, which puts it between a scalar, which has one component, and a vector, which has three. However, the components of the spinor are complex.

A significant difference between spin and orbital angular momentum is this: we can change the magnitude of orbital angular momentum of a particle (by applying external fields) but not the magnitude of its spin. The S^2 operator is

$$S^2 = \hbar^2 \begin{bmatrix} (\frac{1}{2})(\frac{1}{2}+1) & 0 \\ 0 & (\frac{1}{2})(\frac{1}{2}+1) \end{bmatrix} = \tfrac{3}{4}\hbar^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{14.3.11}$$

and yields a value $\frac{3}{4}\hbar^2$ on any state ψ. [For any particle, the magnitude of spin is decided by the number of components in the wave function and is an invariant. Thus the spin of the electron is always 1/2 (in units of $\hbar$) and serves as an invariant label of the particle, like its charge or rest mass.]

We have deduced that the electron is to be described by a two-component wave function in the coordinate basis.‡ Let us restate this result in Hilbert space. First, it is easy to see that the introduction of spin has doubled the size of Hilbert space; if it was ∞ dimensional before, now it is 2∞ dimensional, if you know what I mean. The basis vectors $|xyzs_z\rangle$ diagonalize the mutually commuting operators X, Y, Z, and S_z (one can also think of other bases such as $|\mathbf{p}s_z\rangle$ or $|\mathbf{p}s_x\rangle$ etc.). The state vector

‡ We made the deduction given the empirical input from experiment. When we come to the Dirac equation, we will see that incorporating relativistic kinematics will automatically lead to a multicomponent wave function, i.e., lead to spin, if we demand that the equation be first order in time and space.

$|\psi\rangle$ is a 2∞-dimensional column vector in this basis:

$$|\psi\rangle \xrightarrow[\mathbf{R},S_z \text{ basis}]{} \langle xyzs_z|\psi\rangle = \begin{bmatrix} \psi(x=-\infty, y=-\infty, z=-\infty, s_z=+\hbar/2) \\ \vdots \\ \psi(x, y, z, s_z=+\hbar/2) \\ \vdots \\ \psi(x=\infty, y=\infty, z=\infty, s_z=+\hbar/2) \\ \hline \psi(x=-\infty, y=-\infty, z=-\infty, s_z=-\hbar/2) \\ \vdots \\ \psi(x, y, z, s_z=-\hbar/2) \\ \vdots \\ \psi(x=\infty, y=\infty, z=\infty, s_z=-\hbar/2) \end{bmatrix} \tag{14.3.12}$$

Clearly $\psi(\mathbf{r}, \pm\hbar/2)$ gives the amplitude to find the electron at $\mathbf{r}$ with $s_z = \pm\hbar/2$. The horizontal dashed line separates the components with $s_z = \hbar/2$ from those with $s_z = -\hbar/2$. Now if s_z is fixed at $\hbar/2$ and we vary x, y, z from $-\infty$ to ∞, the component of $|\psi\rangle$ will vary smoothly, i.e., define a continuous function $\psi_+(x, y, z)$. Likewise the components below the dotted line define a function $\psi_-(x, y, z)$. In terms of these functions, we may compactify Eq. (14.3.12) to the form

$$|\psi\rangle \xrightarrow[\mathbf{R},S_z \text{ basis}]{} \begin{bmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{bmatrix} \tag{14.3.13}$$

This notation blends two notations we have used so far: if the vector has components labeled by a discrete index i ($i = 1, \ldots, n$) we denote it as a column vector, while if it is labeled by a continuous index such as x, we denote it by a function $\psi(x)$; but here, since it depends on discrete (s_z) as well as continuous (x, y, z) indices, we write it as a column vector whose components are functions. The normalization condition is

$$\begin{aligned} 1 = \langle\psi|\psi\rangle &= \sum_{s_z} \int \langle\psi|xyzs_z\rangle\langle xyzs_z|\psi\rangle \, dx\, dy\, dz \\ &= \int (|\psi_+|^2 + |\psi_-|^2)\, dx\, dy\, dz \end{aligned} \tag{14.3.14}$$

In the compact notation, S_z is a 2×2 matrix:

$$S_z|\psi\rangle \xrightarrow[\mathbf{R},S_z \text{ basis}]{} \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{bmatrix} \tag{14.3.15a}$$

whereas in its full glory, it is a 2∞-dimensional matrix:

$$S_z|\psi\rangle \xrightarrow[\mathbf{R},S_z \text{ basis}]{} \frac{\hbar}{2}\begin{bmatrix} 1 & & & & & & 0 \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & -1 & & \\ & & & & & -1 & \\ & & & & & & \ddots \\ 0 & & & & & & -1 \end{bmatrix}\begin{bmatrix} \psi(-\infty, \hbar/2) \\ \vdots \\ \\ \vdots \\ \psi(\infty, -\hbar/2) \end{bmatrix} \tag{14.3.15b}$$

What about the familiar operators $\Omega(\mathbf{R}, \mathbf{P})$? Equation (14.3.9) gives $\mathbf{P}$ in the compact notation. Likewise, L_z becomes

$$L_z|\psi\rangle \xrightarrow[\mathbf{R},S_z \text{ basis}]{} \begin{bmatrix} -i\hbar\,\partial/\partial\phi & 0 \\ 0 & -i\hbar\,\partial/\partial\phi \end{bmatrix}\begin{bmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{bmatrix} \tag{14.3.16}$$

The forms of these operators are consistent with the requirement that operators built out of $\mathbf{R}$ and $\mathbf{P}$ commute with the spin operators. Observe that the Hilbert space $\mathbb{V}_e$ of the electron may be viewed as a direct product of an infinite-dimensional space $\mathbb{V}_0$, which describes a particle with just orbital degrees of freedom, and a two-dimensional space $\mathbb{V}_s$, which describes a particle with just spin degrees of freedom:

$$\mathbb{V}_e = \mathbb{V}_0 \otimes \mathbb{V}_s \tag{14.3.17}$$

The basis vector $|x, y, z, s_z\rangle$ of $\mathbb{V}_e$ is just a direct product

$$|x, y, z, s_z\rangle = |xyz\rangle \otimes |s = 1/2, s_z\rangle \tag{14.3.18}$$

Of course $\mathbb{V}_0$ and $\mathbb{V}_s$ do not describe two particles which are amalgamated into a single system, but, rather, two independent degrees of freedom of the electron.

Since we already know how to handle the orbital degrees of freedom, let us pretend from now on that only the spin degree of freedom exists. Or, to be more precise, let us assume the orbital degree of freedom exists but evolves independently. Formally this means that the Hamiltonian is separable:

$$H = H_0 + H_s \tag{14.3.19}$$

where H_0 and H_s depend on just the orbital and spin operators, respectively. Consequently the state vector factorizes into‡

$$|\psi(t)\rangle = |\psi_0(t)\rangle \otimes |\chi_s(t)\rangle \tag{14.3.20}$$

where $|\psi_0\rangle$ and $|\chi_s\rangle$ are elements of $\mathbb{V}_0$ and $\mathbb{V}_s$, respectively. Now $|\psi_0(t)\rangle$ evolves in response to H_0, while the evolution of $|\chi_s(t)\rangle$ is dictated by H_s. We will follow just the evolution of $|\chi_s\rangle$. The product form of $|\psi\rangle$ ensures that the spin and orbital degrees of freedom are statistically independent. Of course, there are many interesting cases in which H is not separable, and the orbital and spin degrees are coupled in their evolution. We will tackle them in a later chapter.

With this assumption, we have just a (complex) two-dimensional Hilbert space $\mathbb{V}_s$ to work with. A complete basis is provided by the vectors $|s, s_z\rangle = |s, m\hbar\rangle \equiv |s, m\rangle$. They are

$$|s, m\rangle = |1/2, 1/2\rangle \xrightarrow[S_z \text{ basis}]{} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{14.3.21a}$$

$$|s, m\rangle = |1/2, -1/2\rangle \xrightarrow[S_z \text{ basis}]{} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{14.3.21b}$$

Any ket $|\chi\rangle$ in $\mathbb{V}_s$ may be expanded as

$$|\chi\rangle = \alpha|1/2, 1/2\rangle + \beta|1/2, -1/2\rangle \xrightarrow[S_z \text{ basis}]{} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{14.3.22}$$

The normalization condition is

$$1 = \langle\chi|\chi\rangle \xrightarrow[S_z \text{ basis}]{} 1 = [\alpha^*, \beta^*]\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = |\alpha|^2 + |\beta|^2 \tag{14.3.23}$$

If one calculates $\langle \mathbf{S} \rangle$ in the eigenstates of S_z, one finds

$$\langle 1/2, \pm 1/2|\mathbf{S}|1/2, \pm 1/2\rangle = \pm(\hbar/2)\mathbf{k} \tag{14.3.24}$$

One refers to these as states with spin pointing up/down the z axis. More generally, the eigenstates $|\hat{n}, \pm\rangle$ of $\hat{n}\cdot\mathbf{S}$ with eigenvalues $\pm\hbar/2$, in which

$$\langle \hat{n}, \pm|\mathbf{S}|\hat{n}, \pm\rangle = \pm(\hbar/2)\hat{n} \tag{14.3.25}$$

are said to be states with spin up/down the direction of the unit vector $\hat{n}$. Let us address the determination (in the S_z basis) of the components of $|\hat{n}, \pm\rangle$ and the verification of Eq. (14.3.25).

‡ In the $\mathbf{R}$, S_z basis, this means $\psi(x, y, z, s_z, t) = \psi_0(x, y, z, t)\chi(t)$ where χ is a two-component spinor independent of x, y, and z.

Let us say $\hat{n}$ points in the direction (θ, ϕ), i.e., that

$$\begin{aligned} \hat{n}_z &= \cos\theta \\ \hat{n}_x &= \sin\theta\cos\phi \\ \hat{n}_y &= \sin\theta\sin\phi \end{aligned} \tag{14.3.26}$$

The kets $|\hat{n}, \pm\rangle$ are eigenvectors of

$$\begin{aligned} \hat{n}\cdot\mathbf{S} &= n_x S_x + n_y S_y + n_z S_z \\ &= \frac{\hbar}{2}\begin{bmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{bmatrix} \\ &= \frac{\hbar}{2}\begin{bmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{bmatrix} \end{aligned} \tag{14.3.27}$$

It is a simple matter to solve the eigenvalue problem (Exercise 14.3.2) and to find

$$|\hat{n}\ \text{up}\rangle \equiv |\hat{n}+\rangle = \begin{bmatrix} \cos(\theta/2)\, e^{-i\phi/2} \\ \sin(\theta/2)\, e^{i\phi/2} \end{bmatrix} \tag{14.3.28a}$$

$$|\hat{n}\ \text{down}\rangle \equiv |\hat{n}-\rangle = \begin{bmatrix} -\sin(\theta/2)\, e^{-i\phi/2} \\ \cos(\theta/2)\, e^{i\phi/2} \end{bmatrix} \tag{14.3.28b}$$

You may verify that as claimed

$$\begin{aligned} \langle \hat{n}\pm|\mathbf{S}|\hat{n}\pm\rangle &= \pm(\hbar/2)(\mathbf{i}\sin\theta\cos\phi + \mathbf{j}\sin\theta\sin\phi + \mathbf{k}\cos\theta) \\ &= \pm(\hbar/2)\hat{n} \end{aligned} \tag{14.3.29}$$

An interesting feature of $\mathbb{V}_s$ is that not only can we calculate $\langle\mathbf{S}\rangle$ given a state, but we can also go the other way, i.e, deduce the state vector given $\langle\mathbf{S}\rangle$. This has to do with the fact that any element of $\mathbb{V}_s$ has only two (complex) components α and β, constrained by the normalization requirement $|\alpha|^2 + |\beta|^2 = 1$, i.e., three real degrees of freedom, and $\langle\mathbf{S}\rangle$ contains exactly three pieces of information. If we write $\langle\mathbf{S}\rangle$ as $(\hbar/2)\hat{n}$, then the corresponding ket is $|\hat{n}, +\rangle$ or if you want $|-\hat{n}, -\rangle$. Another way to state this result is as follows. Instead of specifying a state by α and β, we can give the operator $\hat{n}\cdot\mathbf{S}$ of which it is an eigenvector with eigenvalue $\hbar/2$. An interesting corollary is that every spinor in $\mathbb{V}_s$ is an eigenket of some spin operator $\hat{n}\cdot\mathbf{S}$ with eigenvalue $\hbar/2$.

Exercise 14.3.1. Let us verify the above corollary explicitly. Take some spinor with components $\alpha = \rho_1 e^{i\phi_1}$ and $\beta = \rho_2 e^{i\phi_2}$. From $\langle\chi|\chi\rangle = 1$, deduce that we can write $\rho_1 = \cos(\theta/2)$ and $\rho_2 = \sin(\theta/2)$ for some θ. Next pull out a common phase factor so that the spinor takes the form in Eq. (14.3.28a). This verifies the corollary and also fixes $\hat{n}$.

So much for the state vectors in $\mathbb{V}_s$. How about the operators on this space? Let us commence with S_x, S_y, and S_z. It is convenient to introduce the *Pauli matrices* $\boldsymbol{\sigma}$, defined by

$$\mathbf{S}=\frac{\hbar}{2}\boldsymbol{\sigma} \tag{14.3.30}$$

so that

$$\sigma_x=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},\qquad \sigma_y=\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix},\qquad \sigma_z=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \tag{14.3.31}$$

It is worth memorizing these matrices. Here are some of their important properties.

(1) They *anti*commute with each other:

$$[\sigma_i, \sigma_j]_+=0 \quad \text{or} \quad \sigma_i\sigma_j=-\sigma_j\sigma_i \qquad (i\neq j) \tag{14.3.32}$$

(2) From the commutation rules for the spin operators $\mathbf{S}$, we get, upon using the anticommutativity of the Pauli matrices,

$$\sigma_x\sigma_y=i\sigma_z \quad \text{and cyclic permutations} \tag{14.3.33}$$

(3) They are traceless

$$\operatorname{Tr}\sigma_i=0, \qquad i=x, y, z \tag{14.3.34}$$

(See Exercise 14.3.3 for the proof.)

(4) The square of any Pauli matrix equals I:

$$\sigma_i^2=I \tag{14.3.35}$$

or more generally,

$$(\hat{n}\cdot\boldsymbol{\sigma})^2=I \tag{14.3.36}$$

Proof. Since S_z has eigenvalues $\pm\hbar/2$, it follows that

$$\left(S_z+\frac{I\hbar}{2}\right)\left(S_z-\frac{I\hbar}{2}\right)=0$$

in this Hilbert space.‡ But since what we call the z axis is arbitrary, it must be true that

$$\left(\hat{n}\cdot\mathbf{S}+\frac{I\hbar}{2}\right)\left(\hat{n}\cdot\mathbf{S}-\frac{I\hbar}{2}\right)=0$$

or

$$(\hat{n}\cdot\mathbf{S})^2=\frac{\hbar^2}{4}I$$

or

$$(\hat{n}\cdot\boldsymbol{\sigma})^2=I \qquad \text{Q.E.D.}$$

(5) We can combine Eqs. (14.3.32) and (14.3.35) into

$$[\sigma_i,\sigma_j]_+=2\delta_{ij}I \tag{14.3.37}$$

(6) Combining this relation with the commutation rules

$$[\sigma_x,\sigma_y]=2i\sigma_z \quad \text{and cyclic permutations} \tag{14.3.38}$$

we may establish a very useful identity (Exercise 14.3.4):

$$(\mathbf{A}\cdot\boldsymbol{\sigma})(\mathbf{B}\cdot\boldsymbol{\sigma})=\mathbf{A}\cdot\mathbf{B}I+i(\mathbf{A}\times\mathbf{B})\cdot\boldsymbol{\sigma} \tag{14.3.39}$$

where $\mathbf{A}$ and $\mathbf{B}$ are vectors or vector operators that commute with $\boldsymbol{\sigma}$.

(7) Combining Eqs. (14.3.33), (14.3.34), and (14.3.35) we find that

$$\mathrm{Tr}(\sigma_i\sigma_j)=2\delta_{ij}, \qquad i,j=x,y,z \tag{14.3.40a}$$

Let us view the identity, I, as the fourth Pauli matrix. If we call it σ_0, then

$$\mathrm{Tr}(\sigma_\alpha\sigma_\beta)=2\delta_{\alpha\beta} \qquad (\alpha,\beta=x,y,z,0)\S \tag{14.3.40b}$$

This equation implies that the σ_α matrices are *linearly independent*. By this I mean as usual that

$$\sum_\alpha c_\alpha\sigma_\alpha=0\rightarrow c_\alpha=0 \qquad \text{for all } \alpha \tag{14.3.41}$$

To prove this for say c_β, multiply both sides by σ_β and take the trace.

‡ See Exercise 12.5.4.

§ From now on α, β will run over four values $x, y, z, 0$; while i, j will run over just x, y, and z.

Since any 2×2 matrix M has only four independent (complex) degrees of freedom, it may be written as

$$M=\sum m_\alpha \sigma_\alpha \tag{14.3.42}$$

To find m_β, we multiply by σ_β and take the trace, to find

$$m_\beta = \tfrac{1}{2}\,\mathrm{Tr}(M\sigma_\beta) \tag{14.3.43}$$

(The coefficients m_α will be complex in general, and real if M is Hermitian.)

Thus, any operator in $\mathbb{V}_s$ may be expressed in terms of the σ_α, which form a basis that is orthonormal with respect to the inner product $\frac{1}{2}\,\mathrm{Tr}(\sigma_\alpha\sigma_\rho)$.‡

Explicit Forms of Rotation Operators

The fact that $(\hat{n}\cdot\boldsymbol{\sigma})^2 = I$ greatly simplifies many computations and allows us to compute in closed form several operators such as $U(t)=\exp(-iHt/\hbar)$, $U[R(\boldsymbol{\theta})]=\exp(-i\boldsymbol{\theta}\cdot\mathbf{S}/\hbar)$, which are intractable in infinite-dimensional spaces. In this section we consider the rotation operators, and in the next, the propagator.

Consider

$$\begin{aligned} U[R(\boldsymbol{\theta})] &= \exp(-i\boldsymbol{\theta}\cdot\mathbf{S}/\hbar) = \exp(-i\boldsymbol{\theta}\cdot\boldsymbol{\sigma}/2) \\ &= \exp\left[-i\left(\frac{\theta}{2}\right)\hat{\theta}\cdot\boldsymbol{\sigma}\right] \\ &= \sum_{n=0}^{\infty}\left(-\frac{i\theta}{2}\right)^n \frac{1}{n!}(\hat{\theta}\cdot\boldsymbol{\sigma})^n \\ &= I+\left(-\frac{i\theta}{2}\right)\hat{\theta}\cdot\boldsymbol{\sigma}+\frac{1}{2!}\left(-\frac{i\theta}{2}\right)^2 I+\frac{1}{3!}\left(-\frac{i\theta}{2}\right)^3(\hat{\theta}\cdot\boldsymbol{\sigma})+\cdots \end{aligned}$$

Grouping together the coefficients of I and $\hat{\theta}\cdot\boldsymbol{\sigma}$, we get

$$U[R(\boldsymbol{\theta})]=\cos(\theta/2)I - i\sin(\theta/2)\hat{\theta}\cdot\boldsymbol{\sigma} \tag{14.3.44}$$

Let us put this operator to a test. Suppose we have a particle with spin up along the z direction, i.e., in the state $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$. If we want to get from this a particle in the state $|\hat{n},+\rangle$, it is clear that we must rotate $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ by an angle θ about an axis perpendicular to the z axis and the $\hat{n}$ axis. Thus the rotation angle is

$$\boldsymbol{\theta}=\theta\hat{\theta}=\theta\,\frac{\mathbf{k}\times\hat{n}}{|\mathbf{k}\times\hat{n}|} \tag{14.3.45}$$

‡ The inner product between two matrices M and M' acting on $\mathbb{V}_s$ is actually $\mathrm{Tr}(MM'^{\dagger})$. However, the dagger is irrelevant for the Hermitian σ's. It is an interesting exercise to check that this inner product obeys the three axioms.

where $\mathbf{k}$ is the unit vector along the z axis. Since $\hat{n}=(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, it follows that

$$\hat{\theta}=\frac{1}{\sin\theta}(-\sin\theta\sin\phi, \sin\theta\cos\phi, 0)=(-\sin\phi, \cos\phi, 0) \tag{14.3.46}$$

The rotation matrix is, from Eq. (14.3.44),

$$\exp\left(-\frac{i\theta}{2}\hat{\theta}\cdot\boldsymbol{\sigma}\right)=\begin{bmatrix}\cos(\theta/2) & -\sin(\theta/2)\,e^{-i\phi}\\ \sin(\theta/2)\,e^{i\phi} & \cos(\theta/2)\end{bmatrix} \tag{14.3.47}$$

According to our mnemonic, the first column gives the rotated version of $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$. We see that it agrees with $|\hat{n}, +\rangle$ given in Eq. (14.3.28) up to an overall phase.

Here is a summary of useful formulas that were derived or simply stated:

$$\mathbf{S}=\frac{\hbar}{2}\boldsymbol{\sigma}$$

$$[\sigma_i, \sigma_j]_+=2I\delta_{ij}$$

$$[\sigma_i, \sigma_j]=2i\sum_k \varepsilon_{ijk}\sigma_k$$

$$(\hat{n}\cdot\boldsymbol{\sigma})^2=I$$

$$\mathrm{Tr}\,\sigma_i=0$$

$$\mathrm{Tr}(\sigma_\alpha\sigma_\beta)=2\delta_{\alpha\beta} \qquad (\alpha, \beta=x, y, z, 0)$$

$$\exp\left(-i\frac{\theta}{2}\hat{\theta}\cdot\boldsymbol{\sigma}\right)=\cos\left(\frac{\theta}{2}\right)I-i\sin\left(\frac{\theta}{2}\right)\hat{\theta}\cdot\boldsymbol{\sigma}$$

$$(\mathbf{A}\cdot\boldsymbol{\sigma})(\mathbf{B}\cdot\boldsymbol{\sigma})=(\mathbf{A}\cdot\mathbf{B})I+i(\mathbf{A}\times\mathbf{B})\cdot\boldsymbol{\sigma}$$

*Exercise 14.3.2.** (1) Show that the eigenvectors of $\boldsymbol{\sigma}\cdot\hat{n}$ are given by Eq. (14.3.28).
(2) Verify Eq. (14.3.29).

*Exercise 14.3.3.** Using Eqs. (14.3.32) and (14.3.33) show that the Pauli matrices are traceless.

*Exercise 14.3.4.** Derive Eq. (14.3.39) in two different ways.
(1) Write $\sigma_i\sigma_j$ in terms of $[\sigma_i, \sigma_j]_+$ and $[\sigma_i, \sigma_j]$.
(2) Use Eqs. (14.3.42) and (14.3.43).

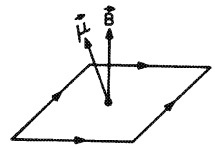

Figure 14.1. In the figure, **B** is the magnetic field and **μ** is the magnetic moment of the loop. The direction of the arrows in the loop is that of the current.

Exercise 14.3.5. Express the following matrix M in terms of the Pauli matrices:

$$M=\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

Exercise 14.3.6. (1) Argue that $|\hat{n}, +\rangle = U[R(\phi\mathbf{k})]U[R(\theta\mathbf{j})]|s_z=\hbar/2\rangle$. (2) Verify by explicit calculation.

Exercise 14.3.7. Express the following as linear combinations of the Pauli matrices and I:

(1) $(I+i\sigma_x)^{1/2}$. (Relate it to half a certain rotation.)

(2) $(2I+\sigma_x)^{-1}$.

(3) σ_x^{-1}.

*Exercise 14.3.8.** (1) Show that any matrix that commutes with $\boldsymbol{\sigma}$ is a multiple of the unit matrix.

(2) Show that we cannot find a matrix that anticommutes with all three Pauli matrices. (If such a matrix exists, it must equal zero.)

14.4. Spin Dynamics

Since the quest for the spin Hamiltonian is based on classical analogy, let us recall some basic ideas from classical magnetostatics. Consider a square loop (Fig. 14.1) carrying a current I, in a magnetic field **B**. From standard magnetostatics (force per unit length on a current-carrying conductor etc.) one can show that the torque on the loop is

$$\mathbf{T}=\boldsymbol{\mu}\times\mathbf{B} \tag{14.4.1}$$

where **μ**, the *magnetic moment*, is given by

$$\boldsymbol{\mu}=\frac{I\cdot A}{c}\mathbf{e}_\perp \tag{14.4.2}$$

where A is the area of the loop, c is the velocity of light, and $\mathbf{e}_\perp$ is a unit vector perpendicular to the plane of the loop.‡ The effect of **T** will be to rotate the loop until **μ** and **B** are parallel.

Since we finally wish to address a quantum mechanical problem, it is preferable to summarize the interaction between the loop and the magnetic field in terms of

‡ The sense of $\mathbf{e}_\perp$ is related to the current flow by the right-hand rule.

the potential energy associated with the torque: If θ is the angle between $\boldsymbol{\mu}$ and $\mathbf{B}$, the interaction energy is‡

$$\mathscr{H}_{\text{int}} = \int T(\theta)\, d\theta = \int \mu B \sin\theta\, d\theta = -\mu B \cos\theta = -\boldsymbol{\mu}\cdot\mathbf{B} \tag{14.4.3}$$

As we would expect, this energy is minimized, i.e., a stable configuration obtains, when $\boldsymbol{\mu}$ and $\mathbf{B}$ are parallel.

Although we derived the above equations for a square loop, they are true for any tiny planar loop, over whose extent $\mathbf{B}$ is constant. So we may apply it to the following problem. Imagine a particle of mass m, charge q, moving in a circular orbit of radius r. The current associated with this charge is

$$\begin{aligned} I &= \text{charge flow past any point in the circle per second} \\ &= \frac{qv}{2\pi r} \end{aligned} \tag{14.4.4}$$

and the magnetic moment has a magnitude

$$\mu = \frac{qv}{2\pi r}\cdot\frac{\pi r^2}{c} = \frac{qvr}{2c} = \left(\frac{q}{2mc}\right) mvr = \frac{q}{2mc}\cdot l \tag{14.4.5}$$

where l is the magnitude of the angular momentum. Since $\boldsymbol{\mu}$ and $\mathbf{l}$ are parallel,

$$\boldsymbol{\mu} = \left(\frac{q}{2mc}\right)\mathbf{l} \tag{14.4.6}$$

The ratio of $\boldsymbol{\mu}$ to $\mathbf{l}$ is called the *gyromagnetic ratio* γ. For the particle considered above,

$$\gamma = \frac{q}{2mc} \tag{14.4.7}$$

In the case of the current loop, it was stated that the effect of the torque $\mathbf{T}$ is to cause $\boldsymbol{\mu}$ to align with $\mathbf{B}$. This picture changes when $\boldsymbol{\mu}$ has its origin in angular momentum, as is the case for the particle in question. In this case, $\mathbf{T}$ causes a

‡ This is not the full Hamiltonian (for it does not include the kinetic energy of the loop) but just the potential energy of interaction with the magnetic field.

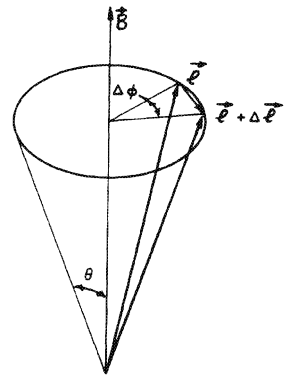

Figure 14.2. In a small time Δt, the tip of the **l** vector precesses by an angle $\Delta\phi$ around the magnetic field vector.

precession of $\boldsymbol{\mu}$ around **B**. We may see this as follows (see Fig. 14.2). The equation of motion is

$$\mathbf{T}=\frac{d\mathbf{l}}{dt}=\boldsymbol{\mu}\times\mathbf{B}=\gamma(\mathbf{l}\times\mathbf{B}) \tag{14.4.8}$$

So in a small time Δt,

$$\Delta\mathbf{l}=\gamma(\mathbf{l}\times\mathbf{B})\Delta t$$

or

$$\Delta l=\gamma lB\sin\theta\,\Delta t$$

Since $\Delta\mathbf{l}$ is perpendicular to **l**, the tip of the **l** vector moves by an angle

$$\Delta\phi=\left(\frac{-\Delta l}{l\sin\theta}\right)=(-\gamma B)\,\Delta t \tag{14.4.9}$$

i.e., precesses at a frequency

$$\boldsymbol{\omega}_0=-\gamma\mathbf{B} \tag{14.4.10}$$

Orbital Magnetic Moment in Quantum Theory

These ideas reemerge in the quantum theory. The Hamiltonian for a particle of mass m and charge q in a magnetic field is

$$H=\frac{(\mathbf{P}-q\mathbf{A}/c)^2}{2m}=\frac{|\mathbf{P}|^2}{2m}-\frac{q}{2mc}(\mathbf{P}\cdot\mathbf{A}+\mathbf{A}\cdot\mathbf{P})+\frac{q^2|\mathbf{A}|^2}{2mc^2} \tag{14.4.11}$$

Let

$$\mathbf{A}=\frac{B}{2}(-y\mathbf{i}+x\mathbf{j}) \tag{14.4.12}$$

so that

$$\nabla\times\mathbf{A}=\mathbf{B}=B\mathbf{k} \tag{14.4.13}$$

is constant and along the z axis. We will assume B is small and drop the last term in H, quadratic in B. When the middle term acts on any $|\psi\rangle$,

$$\begin{aligned}(\mathbf{P}\cdot\mathbf{A})|\psi\rangle &\to -i\hbar\nabla\cdot(\mathbf{A}\psi)\\ &=-i\hbar[(\nabla\cdot\mathbf{A})\psi+\mathbf{A}\cdot\nabla\psi]\\ &=(-i\hbar\mathbf{A}\cdot\nabla)\psi\to(\mathbf{A}\cdot\mathbf{P})|\psi\rangle\end{aligned}$$

since $\nabla\cdot\mathbf{A}=0$ here.‡ Thus the interaction Hamiltonian is

$$\begin{aligned}H_{\text{int}}&=-\frac{q}{2mc}(2\mathbf{A}\cdot\mathbf{P})\\ &=-\frac{q}{mc}\frac{B}{2}(-YP_x+XP_y)\\ &=-\frac{q}{2mc}\mathbf{L}\cdot\mathbf{B}\equiv-\boldsymbol{\mu}\cdot\mathbf{B}\end{aligned} \tag{14.4.14}$$

so that

$$\boldsymbol{\mu}=\frac{q}{2mc}\mathbf{L} \tag{14.4.15}$$

exactly as in the classical case. (We use the same symbol $\boldsymbol{\mu}$ to denote the classical variable and the quantum operator. We will occasionally violate our convention in this manner, so as to follow other widely used conventions.)

If we project this relation along the z axis, we get

$$\mu_z=\frac{q}{2mc}L_z=\frac{q\hbar}{2mc}(0,\pm1,\pm2,\ldots)$$

‡ It is shown in Section 18.4 that $\mathbf{A}$ corresponding to a given $\mathbf{B}$ can always be chosen divergenceless.

The quantity $q\hbar/2mc$ is called the *Bohr magneton* of the particle. The *electron Bohr magneton*, simply called the *Bohr magneton*, has a magnitude

$$\frac{e\hbar}{2mc} \simeq 0.6 \times 10^{-8} \text{ eV/G} \tag{14.4.16}$$

where m is the mass of the electron and G stands for gauss. The *nucleon Bohr magneton* is about 2000 times smaller:

$$\frac{e\hbar}{2Mc} \simeq 0.3 \times 10^{-11} \text{ eV/G} \tag{14.4.17}$$

where M is the nucleon (proton or neutron)‡ mass. (The nucleon Bohr magneton is also called the *nuclear Bohr magneton*.)

It may be verified, by the use of Ehrenfest's theorem, that $\langle \mathbf{L} \rangle$ precesses around the constant field $\mathbf{B}$ just as $\mathbf{l}$ would (Exercise 14.4.1).

Spin Magnetic Moment

Armed with all this knowledge, we now address the problem of how the electron interacts with an external magnetic field. We assume once again that there is a magnetic moment operator $\boldsymbol{\mu}$ associated with the spin angular momentum. Since any operator on $\mathcal{V}_s$ is a linear combination of the identity and the spin operators, and since $\boldsymbol{\mu}$ is a vector operator, we conclude that

$$\boldsymbol{\mu} = \gamma \mathbf{S} \tag{14.4.18a}$$

where γ is a constant. Since $\gamma = -e/2mc$ for the orbital case, let us write

$$\boldsymbol{\mu} = g(-e/2mc)\mathbf{S} \tag{14.4.18b}$$

where g is a constant. We also assume that

$$H_{\text{int}} = -\boldsymbol{\mu} \cdot \mathbf{B} = \frac{ge}{2mc} \mathbf{S} \cdot \mathbf{B}$$

$$= \frac{ge\hbar}{4mc} \boldsymbol{\sigma} \cdot \mathbf{B} \tag{14.4.19}$$

The intrinsic magnetic moment due to spin is $g/2$ magnetons. Our present formalism does not tell us what g is; to find it we must confront the above H with experiment and hope that for some value of g it gives the right physics. This happens to be the case, and the experimental value for g is very close to 2. We assume

‡ Recall that these two are nearly equal: $M_p c^2 = 938.28$ MeV, while $M_n c^2 = 939.57$ MeV.

hereafter that

$$g=2 \tag{14.4.20}$$

Thus the gyromagnetic ratio for spin is *twice as big* as for orbital angular momentum.

Why is $g \simeq 2$? And why isn't it exactly equal to 2, which would be much prettier? Our formalism doesn't tell us. But it is irresistible to digress and mention that the Dirac equation, which we will discuss in Chapter 20, predicts that $g=2$ exactly. Quantum electrodynamics, which we will not discuss in this book, predicts that the Dirac result will receive corrections that can be calculated in a power series in α, the fine-structure constant. The physics behind the corrections is the following. Recall that the interaction between the electron and other charged particles is mediated by the exchange of photons. Occasionally, an electron will recapture the photon it emitted. Between the emission and reabsorption, the system that originally contained just the electron will contain an electron and the photon. If the magnetic moment of the system is probed at this time, we can get a result that corresponds to $g \neq 2$, since the electron in the two-particle system has both spin and orbital angular momentum. In fact, quantum electrodynamics predicts that what we call the electron is really a superposition of states that contain one Dirac electron, a Dirac electron and a photon, a Dirac electron, several photons, and several electron–positron pairs, etc.‡ The reason the observed value of g is so close to the Dirac value of 2 is that configurations of increasing complexity are suppressed by increasing powers of the fine-structure constant in the superposition. Thus the simplest configuration, with just the Dirac electron, will dominate the picture and the complicated states will provide smaller and smaller corrections to the result $g=2$. The corrections may be calculated in a power series in α:

$$g=2\left[1+\frac{1}{2\pi}\cdot\alpha+O(\alpha^2)+\cdots\right]$$

which has been evaluated to order α^3. The result is§

$$g_{\text{theory}}=2[1.001159652140(\pm 28)]$$

where the error ± 28 in the last two digits is mostly due to uncertainties in the value of α itself and in the numerical evaluation of some of the integrals in the calculation.

In addition to higher-order corrections, this result also receives corrections due to other interactions of the electron, i.e., due to its ability to exchange other quanta such as the graviton. But these effects are negligible to the accuracy considered above. The experimental value of g is‖

$$g_{\text{exp}}=2[1.0011596521884(\pm 43)]$$

‡ The time-energy uncertainty relation allows the production of these particles for short times.

§ T. Kinoshita and W. B. Lindquist, *Phys. Rev.* **D42**, 636, 1990.

‖ R. S. Van Dyck, P. B. Schwinberg, and H. G. Dehmelt, *Phys. Rev. Lett.* **59**, 26, 1987.

in splendid agreement with theory. Feynman has pointed out that this is equivalent to predicting and measuring the distance between New York and Los Angeles to within the width of a human hair!

The theoretical situation is bad for the nucleons. The reason is that these participate in strong interactions as well, i.e., can emit and absorb pions etc., and the counterpart of α is large ($\simeq 15$). In other words, the state with just the Dirac particle no longer dominates, and the corrections are no longer tiny. We can of course measure g experimentally, and the result is (to two places)

$$\gamma_{\text{proton}} = 5.6\ (e/2Mc)$$
$$\gamma_{\text{neutron}} = -3.8\ (e/2Mc)$$

Dirac theory predicts $\gamma = e/Mc$ or $g = 2$ for the proton and $\gamma = 0$ for the neutral neutron. The nonzero γ of the neutron reflects the fact that the neutron can be in a state that has particles with compensating electrical charges but not necessarily compensating magnetic moments.

Because of their large masses, the magnetic moments of the nucleons are negligible compared to that of the electron.‡

Let us now return to the dynamics of spin in a magnetic field **B**. All we need from now on is the Hamiltonian

$$H = -\boldsymbol{\mu}\cdot\mathbf{B} = -\gamma\mathbf{S}\cdot\mathbf{B} \tag{14.4.21}$$

where

$$\gamma = \frac{-e\cdot 2}{2mc} = \frac{-e}{mc} \tag{14.4.22}$$

Let $|\psi(0)\rangle$ be the initial state of the electron. The state at a later time is

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle$$

where

$$U(t) = e^{-iHt/\hbar} = e^{+i\gamma t(\mathbf{S}\cdot\mathbf{B})/\hbar} \tag{14.4.23}$$

Since $\exp(-i\boldsymbol{\theta}\cdot\mathbf{S}/\hbar)$ is the operator that rotates by $\boldsymbol{\theta}$, *the effect of* $U(t)$ *is clearly to rotate the state by an angle*

$$\boldsymbol{\theta}(t) = -\gamma\mathbf{B}t \tag{14.4.24}$$

‡ End of digression.

It follows that $\langle \mathbf{S} \rangle$ will precess around $\mathbf{B}$ at a frequency $\boldsymbol{\omega}_0 = -\gamma \mathbf{B}$. If this seems too abstract, let us consider a concrete example. Let $\mathbf{B}$ be along the z axis: $\mathbf{B} = B\mathbf{k}$. In this case

$$\begin{aligned} U(t) &= \exp(i\gamma t S_z B/\hbar) \\ &= \exp(i\omega_0 t \sigma_z/2) \qquad (\omega_0 = \gamma B) \end{aligned}$$

Since σ_z is diagonal,

$$U(t) \rightarrow \begin{bmatrix} e^{i\omega_0 t/2} & 0 \\ 0 & e^{-i\omega_0 t/2} \end{bmatrix}$$

Consider an electron that starts out in the state $|\hat{n}, +\rangle$:

$$|\psi(0)\rangle = |\hat{n}, +\rangle \rightarrow \begin{bmatrix} \cos(\theta/2)\, e^{-i\phi/2} \\ \sin(\theta/2)\, e^{i\phi/2} \end{bmatrix}$$

in which case

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \rightarrow \begin{bmatrix} \cos(\theta/2)\, e^{-i(\phi - \omega_0 t)/2} \\ \sin(\theta/2)\, e^{i(\phi - \omega_0 t)/2} \end{bmatrix}$$

i.e., ϕ decreases at a rate ω_0.

Paramagnetic Resonance

Consider a classical magnetic moment $\boldsymbol{\mu}$ in a field $\mathbf{B}_0 = B_0\mathbf{k}$. It will precess around $\mathbf{B}_0$ at a frequency

$$\boldsymbol{\omega}_0 = -\gamma \mathbf{B}_0$$

Suppose we view this process in a frame that is rotating at a frequency $\boldsymbol{\omega}$ parallel to $\mathbf{B}_0$. In this rotating frame, the precession frequency will be

$$\boldsymbol{\omega}_r = \boldsymbol{\omega}_0 - \boldsymbol{\omega} = -\gamma \mathbf{B}_0 - \boldsymbol{\omega} = -\gamma(\mathbf{B}_0 + \boldsymbol{\omega}/\gamma) \tag{14.4.25}$$

Thus the effective field in this rotating frame will be

$$\mathbf{B}_r = \mathbf{B}_0 + \boldsymbol{\omega}/\gamma \tag{14.4.26}$$

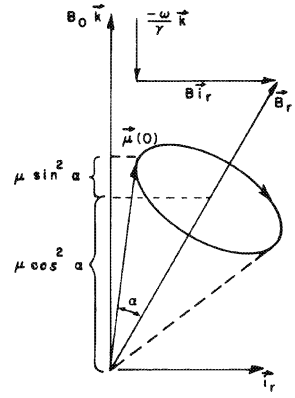

Figure 14.3. The situation in the rotating frame. The effective magnetic field is $\mathbf{B}_r$. The magnetic moment starts out along the z axis (but is slightly displaced in the figure for clarity) and precesses around $\mathbf{B}_r$. The z component of the moment oscillates with an amplitude $\mu \sin^2 \alpha$, where α is the opening angle of the cone. At resonance, $\mathbf{B}_r$ lies along the x axis and $\boldsymbol{\mu}$ precesses in the plane normal to it. The amplitude of the μ_z oscillation is then at its maximum value of μ.

This result is valid even if $\boldsymbol{\omega}$ and $\mathbf{B}_0$ are not parallel (Exercise 14.4.5). Consider now the problem of interest, where, in a nonrotating (lab) frame

$$\mathbf{B} = B\cos\omega t\mathbf{i} - B\sin\omega t\mathbf{j} + B_0\mathbf{k} \qquad (B \ll B_0) \tag{14.4.27}$$

and at $t=0$,

$$\boldsymbol{\mu}(0) = \mu\mathbf{k} \tag{14.4.28}$$

We would like to find out the evolution of $\boldsymbol{\mu}(t)$. Since $\mathbf{B}$ depends on time, it proves advantageous to first consider the problem in a frame that rotates at the same frequency $\boldsymbol{\omega} = -\omega\mathbf{k}$ as the tiny clockwise rotating field B. In this frame, the rotating component of $\mathbf{B}$ gets frozen (say along the x axis) and the constant component $B_0\mathbf{k}$ gets reduced as per Eq. (14.4.26) so that the effective, time-independent field is

$$\mathbf{B}_r = B\mathbf{i}_r + (B_0 - \omega/\gamma)\mathbf{k} \tag{14.4.29}$$

where $\mathbf{i}_r$ is the unit vector in the x direction in the rotating frame. ($\mathbf{k} = \mathbf{k}_r$ of course.) In this frame, $\boldsymbol{\mu}$ will precess around $\mathbf{B}_r$ at a frequency

$$\boldsymbol{\omega}_r = -\gamma\mathbf{B}_r \tag{14.4.30a}$$

where

$$|\boldsymbol{\omega}_r| = \omega_r = \gamma[B^2 + (B_0 - \omega/\gamma)^2]^{1/2} \tag{14.4.30b}$$

It is a simple matter to deduce from Fig. 14.3 that μ_z oscillates as follows:

$$\begin{aligned}\mu_z(t) &= \mu\cos^2\alpha + \mu\sin^2\alpha\cos\omega_r t \\ &= \mu_z(0)\left[\frac{(\omega_0-\omega)^2}{(\omega_0-\omega)^2+\gamma^2B^2} + \frac{\gamma^2B^2\cos\omega_r t}{(\omega_0-\omega)^2+\gamma^2B^2}\right]\end{aligned} \tag{14.4.31}$$

This formula for $\mu_z(t)$ applies in the lab frame as well, since μ_z is invariant under z rotations. As ω increases from 0, the z component of $\mathbf{B}_r$ steadily decreases; α, the opening angle of the cone, increases, and the amplitude of oscillation, $\mu \sin^2 \alpha$, grows. At *paramagnetic resonance*, $\omega = \omega_0$, $\mathbf{B}_r = B\mathbf{i}_r$, $\alpha = \pi/2$, the cone becomes a circle in the y–z plane, and μ_z oscillates with the largest amplitude μ at a frequency γB. The behavior for $\omega > \omega_0$ is obvious.

What if we apply the rotating field at the resonance frequency, but for a time τ such that

$$\gamma B \tau = \pi/2?$$

Such a pulse, called a *90° pulse*, will swing the magnetic moment into the x–y plane (in either frame). Thereafter $\boldsymbol{\mu}$ will precess around $B_0\mathbf{k}$ at the frequency ω_0 in the lab frame. If we apply a 180° pulse, i.e., choose τ such that

$$\gamma B \tau = \pi$$

the pulse will reverse the sign of $\boldsymbol{\mu}$ and leave it pointing down the z axis, where it will stay (in either frame).

These results for the classical moment $\boldsymbol{\mu}$ apply to the expectation value $\langle \boldsymbol{\mu} \rangle$ in the quantum problem, as you may verify by doing Exercise 14.4.1, where it is proved in general, and Exercise 14.4.3, where the explicit verification in this case is discussed.

Negative Absolute Temperature (Optional Digression)

The absolute zero of temperature, 0 K, ($\simeq -273$°C) is defined so that nothing can be colder, yet here we speak of negative absolute temperatures! There is no conflict, however, since we will see that negative temperatures are hotter than positive temperatures! Before you give up all faith, let us quickly sort this thing out.

The absolute temperature T is defined as follows:

$$\beta = \frac{1}{kT} = \frac{1}{k}\frac{\partial S}{\partial E} = \frac{\partial \ln \Omega(E)}{\partial E} \tag{14.4.32}$$

where β is the *thermodynamic temperature*, k is Boltzmann's constant, $S = k \ln \Omega$ is the entropy and $\Omega(E)$ is the number of states available to the system as a function of its energy. (Ω depends on other variables, assumed to be fixed.) In most systems, β is positive because adding energy only opens up more states and increases Ω. For instance, if we have a box of gas molecules, they all stay in the ground state at $T = 0$. So, $S = k \ln \Omega = k \ln 1 = 0$. As we pump in energy, they can occupy higher states, and S and Ω can increase without limit.

Consider now a collection of N spin-half particles sitting on some crystal lattice which is immersed in a field $\mathbf{B} = B_0\mathbf{k}$. Each magnetic moment (or spin) has two states only, with energies $E = \pm\mu B_0$, where μ is the magnitude of the magnetic moment. At $T = 0$ K, all are in the ground state ($\boldsymbol{\mu}$ parallel to $\mathbf{B}$); $\Omega = 1$, and $S = 0$. The system has a magnetic moment $\mathbf{M} = n\mu\mathbf{k}$. If we pump in energy $2\mu B_0$, one of the moments can move to the upper energy state; there are N ways to pick the one that moves

up, so that $\Omega = N$ and $S = k \ln N$. Clearly β and T are positive. As we pump in more and more energy, S keeps growing until half are up and half are down. At this point, S reaches a maximum, $\beta = \partial S/\partial E = 0$, and $T = +\infty$. The system has no mean magnetic moment along the z axis. Pumping in more energy only reduces S, with more and more particles piling up in the *upper* state. So β and T become negative. Finally, when $E = N\mu B_0$, all moments are in the upper energy state (antiparallel to **B**), $\mathbf{M} = -N\mu\mathbf{k}$, there is only one such state; $\Omega = 1$ and $S = 0$. This corresponds to $\beta = -\infty$, $T = 0^-$. Thus the sequence of temperatures is $T = 0^+, \ldots, 300, \ldots, \infty, -\infty, \ldots, -300, \ldots, 0^-$. In terms of β, there is more continuity: $\beta = \infty, \ldots, 0^+, 0^-, \ldots, -\infty$. (We should have chosen $-\beta$ as the temperature, for it rises monotonically from $-\infty$ to $+\infty$ as we heat the system.) It should be clear that negative temperatures are hotter than positive temperatures since we go from the latter to the former by pumping in energy. We can also see this by imagining a system at $T = -300$ K brought in contact with identical system at $T = +300$ K. Since the populations of the two systems are identical, except for the change, parallel $\leftrightarrow$ antiparallel, they can increase their entropies by moving toward the state with equal numbers up and down. In this process energy clearly flows from the negative temperature system to the positive temperature system, i.e., the former is hotter. Also note that the final equilibrium temperature is not 0 K but ∞ K.

How does one prepare negative temperatures in the lab? One takes a sample at room temperature, say at $T = 300$ K. It will have more moments parallel than antiparallel:

$$\frac{N(\text{parallel})}{N(\text{antiparallel})} = \frac{e^{-(-B\mu B_0)}}{e^{-\beta\mu B_0}} = e^{2\beta\mu B_0} > 1 \tag{14.4.33}$$

and a net magnetic moment **M** along the z axis. If one applies a 180° pulse, there will be *population inversion* (parallel$\leftrightarrow$antiparallel), which amounts to a change in the sign of β and T [see Eq. (14.4.33)]. The spin system cannot stay in this hot state ($T = -300$ K) forever, because it is in contact with the lattice, which will eventually cool it down to room temperature.

The return to thermal equilibrium is easier to observe if one applies a 90° pulse which swings **M** into the x–y plane. The temperature now is $T = \infty$ K, since $M_z = 0 \rightarrow N(\text{parallel}) = N(\text{antiparallel}) \rightarrow T = \infty$. Thus **M**, which will initially begin to precess around $\mathbf{B} = B_0\mathbf{k}$, will eventually realign itself with **B**. The decay of its rotating components in the x–y plane may be observed as follows. Suppose the specimen is a long cylinder whose axis lies in the x–y plane. If one winds a coil around it, the transverse (x–y) components of **M**, which simulate a bar magnet rotating in the x–y plane, will induce an oscillating voltage in the coil. The frequency of the (damped) oscillation will be ω_0 and the half-life will be a time τ, called the *transverse relaxation time*.‡

*Exercise 14.4.1.** Show that if $H = -\gamma\mathbf{L}\cdot\mathbf{B}$, and **B** is position independent,

$$\frac{d\langle\mathbf{L}\rangle}{dt} = \langle\boldsymbol{\mu}\times\mathbf{B}\rangle = \langle\boldsymbol{\mu}\rangle\times\mathbf{B}$$

‡ The transverse components of **M** decay for other reasons, besides restoration of thermal equilibrium. See R. Schumacher, *Magnetic Resonance*, W. A. Benjamin, New York (1970).

Comparing this to Eq. (14.4.8), we see that $\langle \boldsymbol{\mu} \rangle$ evolves exactly like $\boldsymbol{\mu}$. Notice that this conclusion is valid even if $\mathbf{B}$ depends on time and also if we are talking about spin instead of orbital angular momentum. A more explicit verification follows in Exercise 14.4.3.

Exercise 14.4.2. Derive (14.4.31) by studying Fig. 14.3.

*Exercise 14.4.3.** We would like to study here the evolution of a state that starts out as $\binom{1}{0}$ and is subject to the $\mathbf{B}$ field given in Eq. (14.4.27). This state obeys

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H|\psi\rangle \tag{14.4.34}$$

where $H = -\gamma \mathbf{S} \cdot \mathbf{B}$, and $\mathbf{B}$ is time dependent. Since classical reasoning suggests that in a frame rotating at frequency $(-\omega \mathbf{k})$ the Hamiltonian should be time independent and governed by $\mathbf{B}_r$ [Eq. (14.4.29)], consider the ket in the rotating frame, $|\psi_r(t)\rangle$, related to $|\psi(t)\rangle$ by a rotation angle ωt:

$$|\psi_r(t)\rangle = e^{-i\omega t S_z/\hbar} |\psi(t)\rangle \tag{14.4.35}$$

Combine Eqs. (14.4.34) and (14.4.35) to derive Schrödinger's equation for $|\psi_r(t)\rangle$ in the S_z basis and verify that the classical expectation is borne out. Solve for $|\psi_r(t)\rangle = U_r(t)|\psi_r(0)\rangle$ by computing $U_r(t)$, the propagator in the rotating frame. Rotate back to the lab and show that

$$|\psi(t)\rangle \xrightarrow[S_z \text{ basis}]{} \begin{bmatrix} \left[\cos\left(\frac{\omega_r t}{2}\right) + i\frac{\omega_0 - \omega}{\omega_r}\sin\left(\frac{\omega_r t}{2}\right)\right] e^{+i\omega t/2} \\ \frac{i\gamma B}{\omega_r}\sin\left(\frac{\omega_r t}{2}\right) e^{-i\omega t/2} \end{bmatrix} \tag{14.4.36}$$

Compare this to the state $|\hat{\mathbf{n}}, +\rangle$ and see what is happening to the spin for the case $\omega_0 = \omega$.

Calculate $\langle \mu_z(t) \rangle$ and verify that it agrees with Eq. (14.4.31).

Exercise 14.4.4. At $t = 0$, an electron is in the state with $s_z = \hbar/2$. A steady field $\mathbf{B} = B\mathbf{i}$, $B = 100$ G, is turned on. How many seconds will it take for the spin to flip?

Exercise 14.4.5. We would like to establish the validity of Eq. (14.4.26) when $\boldsymbol{\omega}$ and $\mathbf{B}_0$ are not parallel.

(1) Consider a vector $\mathbf{V}$ in the inertial (nonrotating) frame which changes by $\Delta\mathbf{V}$ in a time Δt. Argue, using the results from Exercise 12.4.3, that the change as seen in a frame rotating at an angular velocity $\boldsymbol{\omega}$, is $\Delta\mathbf{V} - \boldsymbol{\omega} \times \mathbf{V}\Delta t$. Obtain a relation between the time derivatives of $\mathbf{V}$ in the two frames.

(2) Apply this result to the case of $\mathbf{I}$ [Eq. (14.4.8)], and deduce the formula for the effective field in the rotating frame.

Exercise 14.4.6 (A Density Matrix Problem). (1) Show that the density matrix for an ensemble of spin-1/2 particles may be written as

$$\rho = \tfrac{1}{2}(I + \mathbf{a}\cdot\boldsymbol{\sigma})$$

where $\mathbf{a}$ is a c-number vector.

(2) Show that $\mathbf{a}$ is the mean polarization, $\langle\vec{\boldsymbol{\sigma}}\rangle$.

(3) An ensemble of electrons in a magnetic field $\mathbf{B} = B\mathbf{k}$, is in thermal equilibrium at temperature T. Construct the density matrix for this ensemble. Calculate $\langle\vec{\boldsymbol{\mu}}\rangle$.

14.5. Return of Orbital Degrees of Freedom

Let us now put back the orbital degrees of freedom. The simplest case is when H is separable:

$$H = H_0 + H_s \tag{14.5.1}$$

so that the energy eigenstates factorize

$$|\psi\rangle = |\psi_0\rangle \otimes |\chi_s\rangle$$

An example is provided by the hydrogen atom, where the Coulomb interaction is independent of spin:

$$H = H_0 \tag{14.5.2}$$

Here the spin is a constant in time, and all that happens is that we attach a constant spinor χ to the wave functions we found in Chapter 13. If we choose χ to be an eigenstate of S_z, we have‡

$$\begin{aligned} |nlmm_s = 1/2\rangle &\to \psi_{nlm}(r, \theta, \phi)\chi_+ \qquad \left[\chi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right] \\ |nlmm_s = -1/2\rangle &\to \psi_{nlm}(r, \theta, \phi)\chi_- \qquad \left[\chi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right] \end{aligned} \tag{14.5.3}$$

The energy levels are of course unaffected. All we have is a doubling of states, with the electron spin being up or down (the z axis) in each of the orbital states (nlm).

Consider next the problem of the hydrogen atom in a weak magnetic field $\mathbf{B} = B\mathbf{k}$. Although both the proton and the electron couple to $\mathbf{B}$, the smallness of the ratio m/M allows us to ignore, in the first approximation, the coupling of the proton's intrinsic and orbital magnetic moments [these are of order m/M and $(m/M)^2$ relative

‡ We use the subscript s on m_s to remind us that it measures the *spin* projection: $s_z = m_s\hbar$. It will be dropped whenever it is obvious that we are dealing with spin.

to that of the electron; see Exercise 14.5.1]. Thus we have, from Eqs. (14.4.14) and (14.4.19),

$$H = H_{\text{Coulomb}} - \left(\frac{-eB}{2mc}\right)L_z - \left(\frac{-eB}{mc}\right)S_z \tag{14.5.4}$$

Since the additional terms in H commute with H_{Coulomb}, L^2, L_z, and S_z, this H is diagonalized by the same states as before, namely, $|nlmm_s\rangle$. The eigenvalues are, however, different:

$$H|nlmm_s\rangle = \left[\frac{-\text{Ry}}{n^2} + \frac{eB\hbar}{2mc}(m + 2m_s)\right]|nlmm_s\rangle \tag{14.5.5}$$

The degeneracy is greatly reduced by the **B** field. The ground state, which was twofold degenerate, splits into two levels:

$$E_{n=1} = -\text{Ry} \pm \frac{e\hbar B}{2mc} \tag{14.5.6}$$

The second, which was eightfold degenerate, splits into five levels:

$$E_{n=2} = -\frac{\text{Ry}}{4} + \frac{eB\hbar}{2mc} \times \begin{bmatrix} 2(m=1,\, m_s=1/2) \\ 1(m=0,\, m_s=1/2)(l=0 \text{ or } 1) \\ 0(m=1,\, m_s=-1/2, \text{ or } m=-1,\, m_s=1/2) \\ -1(m=0,\, m_s=-1/2)\ (l=0 \text{ or } 1) \\ -2(m=-1,\, m_s=-1/2) \end{bmatrix} \tag{14.5.7}$$

and so on. In a multielectron atom, one simply adds the contributions from all the electrons. The splitting of levels leads to an increase in the number of spectral lines; where there was one, there will now be several, and the spacing between them may be varied by varying B. This phenomenon is called the *Zeeman effect*.

Consider lastly the Hamiltonian

$$H = H_{\text{Coulomb}} + a\mathbf{L}\cdot\mathbf{S} \tag{14.5.8}$$

whose origin will be explained in a later chapter. For the present, we note that it is not separable, and consequently the spin and orbital degrees of freedom are coupled in their time evolution. The eigenstates of H will not be simply products of orbital and spin parts, but instead superpositions of such states that diagonalize $\mathbf{L}\cdot\mathbf{S}$. The details will be explained in the next chapter.

*Exercise 14.5.1.** (1) Why is the coupling of the proton's intrinsic moment to **B** an order m/M correction to Eq. (14.5.4)?

(2) Why is the coupling of its orbital motion an order $(m/M)^2$ correction? (You may reason classically in both parts.)

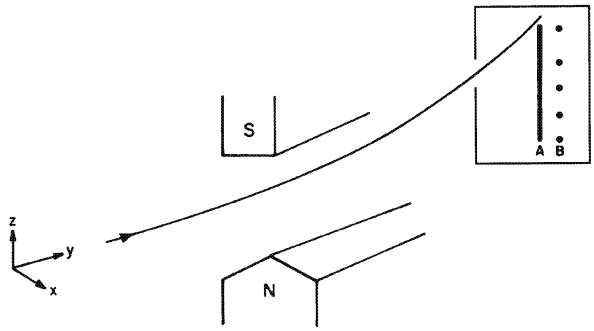

Figure 14.4. The Stern–Gerlach experiment. A beam of particles endowed with magnetic moments enters the inhomogeneous field. Classically the beam is expected to fan out and produce a continuous trace (A) on the screen. What one observes is a set of discrete dots (B). This implies the quantization of magnetic moment and angular momentum.

*Exercise 14.5.2.** (1) Estimate the relative size of the level splitting in the $n=1$ state to the unperturbed energy of the $n=1$ state, when a field $\mathbf{B}=1000\text{kG}$ is applied.

(2) Recall that we have been neglecting the order B^2 term in H. Estimate its contribution in the $n=1$ state relative to the linear $(-\boldsymbol{\mu}\cdot\mathbf{B})$ term we have kept, by assuming the electron moves on a classical orbit of radius a_0. Above what $|\mathbf{B}|$ does it begin to be a poor approximation?

The Stern–Gerlach (SG) Experiment

We now consider (in simplified form) the SG experiment, which clearly displays the quantization of angular momentum (along any direction). The apparatus (Fig. 14.4) consists of north and south pole pieces, between which is an *inhomogeneous* magnetic field. A beam of (particles with) magnetic moments, traveling along the y axis, enters the apparatus in a region where $\mathbf{B}$ is predominantly along the z axis and $\partial B_z/\partial z<0$. What do we expect will happen classically? If we pretend that the magnetic moment is due to a pair of equal and opposite (fictitious) magnetic charges, it is clear that any inhomogeneity in $\mathbf{B}$ can lead to a net force on the dipole. This is confirmed if we calculate the force associated with the gradient of the interaction energy

$$\mathbf{F}=-\nabla\mathscr{H}=\nabla(\boldsymbol{\mu}\cdot\mathbf{B})=(\boldsymbol{\mu}\cdot\nabla)\mathbf{B}=\mu_z\frac{\partial B_z}{\partial z}\mathbf{k} \tag{14.5.9}$$

[We have used the identity $\nabla(\boldsymbol{\mu}\cdot\mathbf{B})=(\boldsymbol{\mu}\cdot\nabla)\mathbf{B}+(\mathbf{B}\cdot\nabla)\boldsymbol{\mu}+\boldsymbol{\mu}\times(\nabla\times\mathbf{B})+\mathbf{B}\times(\nabla\times\boldsymbol{\mu})$. In the present case, $\boldsymbol{\mu}$ is not a function of $\mathbf{r}$, and by Maxwell's equations, $\nabla\times\mathbf{B}=0$. Both F_x and F_y vanish on average due to the precession of spin in the $x-y$ plane.] Classically, since μ_z is continuous, the beam is expected to fan out and produce a continuous trace (A in figure) on a screen placed behind the magnet. The actual experiment performed with atoms reveals a series of discrete dots (B in figure). We understand this in semiclassical terms, by saying that μ_z in Eq. (14.5.9) is discrete and therefore so is the angular momentum along the z axis.

This experiment can also be used to reveal the existence of electron spin. For example, if we send in a beam of hydrogen atoms in their ground state, the beam will split into two parts.

Let us describe the above-mentioned hydrogen atom experiment in quantum mechanical terms. Suppose the initial state of a hydrogen atom is

$$\psi_{\text{initial}} = \psi_y(\mathbf{r}_{\text{CM}})\psi_{100}(\mathbf{r})\begin{bmatrix}1\\0\end{bmatrix} \tag{14.5.10}$$

where ψ_y is a wave packet drifting along the y axis that describes the CM motion, ψ_{100} is the ground state wave function, and $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ is the electron spinor. (The proton spin is ignored, for the associated magnetic moment is too small to affect the dynamics.) Since the electron spin is up, its μ_z is down. Since $\partial B_z/\partial z < 0$, the classical force on the atom is up. So by Ehrenfest's theorem‡ we expect the atom to emerge from the apparatus in a state (up to a phase factor)

$$\psi_{\text{out}} = \psi_{y,+z}(\mathbf{r}_{\text{CM}})\psi_{100}(\mathbf{r})\begin{bmatrix}1\\0\end{bmatrix} \tag{14.5.11}$$

where $\psi_{y,+z}$ describes a wave packet that is displaced (relative to the incoming one) along the positive z axis and has also a small velocity in the same direction. Likewise, if the electron spinor had initially been $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$, the CM would have emerged in the state $\psi_{y,-z}$ (in the same notation). More generally, if

$$\psi_{\text{initial}} = \psi_y\psi_{100}\begin{bmatrix}\alpha\\\beta\end{bmatrix} = \psi_y\psi_{100}\begin{bmatrix}\alpha\\0\end{bmatrix} + \psi_y\psi_{100}\begin{bmatrix}0\\\beta\end{bmatrix} \tag{14.5.12}$$

then, by the linearity of Schrödinger's equation

$$\psi_{\text{out}} = \psi_{y,+z}\psi_{100}\begin{bmatrix}\alpha\\0\end{bmatrix} + \psi_{y,-z}\psi_{100}\begin{bmatrix}0\\\beta\end{bmatrix} \tag{14.5.13}$$

Assuming $\psi_{y,\pm z}$ are narrow packets with no overlap, we see that the SG apparatus has introduced a correlation between the spin and orbital coordinates: if we catch (by placing a screen) the outgoing atom above the original line of flight (i.e., in a region where $\psi_{y,+z}$ is peaked) it will have spin up, while if we catch it below, the spin is down.

The SG apparatus can be used to prepare a state of definite spin orientation: to get a pure spin up/down beam we simply block the lower/upper beam. But note that the filtering process changes the average z component of linear momentum. This can be undone and the particle restored its original momentum (but filtered with respect to spin) if we place some more magnets (with **B** along the z axis) behind this apparatus. With this modification (which is assumed in the following exercises) the

‡ Recall the warning at the end of Chapter 6. In the present case, the system follows the classical trajectory (approximately) thanks to the massive proton. If we send in just the electron, quantum fluctuations would wipe out the effect. See, for example, pages 324–330 of G. Baym, *Lectures on Quantum Mechanics*, Benjamin, New York (1969).

only effect of the SG apparatus with one or the other beams blocked is to filter the spin without affecting the orbital motion.

*Exercise 14.5.3.** A beam of spin-1/2 particles moving along the y axis goes through two collinear SG apparatuses, both with lower beams blocked. The first has its **B** field along the z axis and the second has its **B** field along the x axis (i.e., is obtained by rotating the first by an angle $\pi/2$ about the y axis). What fraction of particles leaving the first will exit the second? If a third filter that transmits only spin up along the z axis is introduced, what fraction of particles leaving the first will exit the third? If the middle filter transmits both spins up and down (no blocking) the x axis, but the last one transmits only spin down the z axis, what fraction of particles leaving the first will leave the last?

Exercise 14.5.4. A beam of spin-1 particles, moving along the y axis, is incident on two collinear SG apparatuses, the first with **B** along the z axis and the second with **B** along the z' axis, which lies in the x–z plane at an angle θ relative to the z axis. Both apparatuses transmit only the uppermost beams. What fraction leaving the first will pass the second?