The curious properties of spin

February 2, 2017

1 The Stern-Gerlach experiment

The Schrödinger equation predicts degenerate energy levels for atoms – electron states that differ only in the z-component of their angular momentum will have equal energies. Specifically, with the angular momentum quantized to integer multiples of the reduced Planck constant, $l\hbar$, quantization still allows $2l+1$ different allowed orientations of the direction of the angular momentum vector. In an attempt to verify this prediction experimentally, Stern and Gerlach performed a series of experiments through the 1920s. By passing atoms through an inhomogeneous magnetic field, they could separate the degenerate states of an atomic beam into an odd number, $(2l+1)$, of separate beams. Surprisingly, some of these experiments produced an even number of beams. Uhlenbeck and Goudsmit interpreted this as the presence of intrinsic spin, of magnitude $l\hbar = \frac{1}{2}\hbar$, of the electron. The half-integer value, $l = \frac{1}{2}$ gives $2l+1 = 2$-fold degeneracy.

1.1 The classical prediction

1.1.1 Magnetic moment of a spinning charged particle

The magnetic moment of a distribution of currents is given by integrating

$$d\boldsymbol{\mu} = \frac{1}{2}\mathbf{r}\times\mathbf{j}$$

where $\mathbf{j}$ is the current density. If $\mathbf{j}$ is given by a moving charge density, $\mathbf{j} = \rho_q\mathbf{v}$ Then we may write this as

$$d\boldsymbol{\mu} = \frac{1}{2}\rho_q\mathbf{r}\times\mathbf{v}$$

Consider a rotating sphere of radius R, of uniform charge Q, and uniform mass M. Then

$$\begin{aligned}\rho_q &= \frac{3Q}{4\pi R^3} \\ &= \frac{Q}{M}\rho_{mass}\end{aligned}$$

so that

$$\begin{aligned}\boldsymbol{\mu} &= \frac{1}{2}\int\mathbf{r}\times\left(\frac{Q}{M}\rho_{mass}\right)\mathbf{v}d^3x \\ &= \frac{Q}{2M}\mathbf{S}\end{aligned}$$

where $\mathbf{S} = \int\rho_{mass}\mathbf{r}\times\mathbf{v}d^3x$ is the total angular momentum vector of the spinning ball. For a nonuniform distribution of charge (and to include quantum field theory effects), we may include an additional factor g, expected to be of order one,

$$\boldsymbol{\mu} = \frac{gQ}{2M}\mathbf{S}$$

Exercise:

How large could g be? Suppose all the charge resided on a ring around the equator of the sphere instead of being distributed throughout the volume. Then the charge density would be

$$\rho_q = \frac{Q}{2\pi R^2}\delta(r-R)\,\delta\left(\theta - \frac{\pi}{2}\right)$$

Show that, in terms of the angular momentum of a sphere, $L = I\omega = \frac{2}{5}MR^2\omega$, that $\boldsymbol{\mu} = \frac{5}{2}\frac{Q}{2M}\mathbf{S}$

Exercise:

Suppose the intrinsic angular momentum of an electron, $\frac{\hbar}{2}$, is given by the classical form,

$$\frac{2}{5}mr^2\omega = \frac{\hbar}{2}$$

where m is the electron mass, r the Compton radius of the electron, $\frac{\hbar}{mc}$ where c is the speed of light. Let $\omega = \frac{v}{r}$ where v is the tangential velocity of the surface of the electron. What is v?

If a classical particle with magnetic moment $\boldsymbol{\mu}$ passes through a magnetic field $\mathbf{B}$, it experiences a torque, $\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$. In addition, if the magnetic field is nonuniform, (but with vanishing curl) there is a net translational force as well, arising from a potential, $V = -\boldsymbol{\mu} \cdot \mathbf{B}$. For the magnetic fields we discuss, the force may be written in the form

$$\mathbf{F} = \boldsymbol{\nabla}(\boldsymbol{\mu} \cdot \mathbf{B})$$

Let the magnetic field have its gradient in the z-direction. This means that if the magnetic moment of the spinning, charged particle is directed parallel to $\mathbf{B}$ it is deflected upward; antialignment leads to a downward deflection. For a beam of such particles, we expect the directions to be randomly directed, leading to a distribution of deflections.

1.1.2 The classical distribution of deflections

For a uniform distribution of magnetic moments, $\boldsymbol{\mu} = \mu\hat{\mathbf{n}}$ the unit vector $\hat{\mathbf{n}}$ is equally likely to point in any direction. The probability that $\hat{\mathbf{n}}$ points into any element of solid angle, $d\Omega = \sin\theta d\theta d\varphi$ is therefore

$$P(\theta,\varphi)\sin\theta d\theta d\varphi = \frac{\sin\theta d\theta d\varphi}{4\pi}$$

If we have a total of N particles, then the number dn in $d\Omega$ is

$$dn = \frac{N}{4\pi}\sin\theta d\theta d\varphi$$

Since $\boldsymbol{\mu} \cdot \mathbf{B} = \mu B\cos\theta$ is independent of θ, we integrate over all φ to get

$$dn = \frac{N}{2}\sin\theta d\theta$$

Let $\boldsymbol{\nabla}B_z(z)$ be constant. Then the impulse $p_z - 0 = \Delta p = F\Delta t$ delivered to a particle with z-component of magnetic moment μ_z will be approximately $\mu_z\frac{dB_z}{dz}\Delta t = \mu\cos\theta\frac{dB_z}{dz}\Delta t$ where Δt is the (short) time spent in the magnetic field. A particle with its magnetic moment oriented at an angle θ will be acquire an upward component of velocity of

$$\begin{aligned} v &= \frac{p_z}{m} \\ &= \frac{\mu\Delta t}{m}\frac{dB_z}{dz}\cos\theta \end{aligned}$$

so a change in the angle of the magnetic moment of $d\theta$ changes the z-component of the velocity by

$$\begin{aligned} dv &= -\frac{\mu\Delta t}{m}\frac{dB_z}{dz}\sin\theta d\theta \\ \frac{dv}{d\theta} &= -\frac{\mu\Delta t}{m}\frac{dB_z}{dz}\sin\theta \end{aligned}$$

Travelling for a time T, a particle with orientation θ strikes a distant screen a height vT above the centerline. The number of particles in an interval $d\theta$ is $dn = \frac{N}{2}\sin\theta d\theta$ so the number of particles per unit distance $dL = Tdv$ along the screen is

$$\begin{aligned} \frac{dn}{dL} &= \frac{dn}{d\theta}\left|\frac{d\theta}{dL}\right| \\ &= \frac{dn}{d\theta}\frac{1}{T}\left|\frac{d\theta}{dv}\right| \\ &= \frac{N}{2}\sin\theta\frac{1}{T}\frac{1}{\frac{\mu\Delta t}{m}\frac{dB_z}{dz}\sin\theta} \\ &= \frac{mN}{2T\mu\Delta t\frac{dB_z}{dz}} \end{aligned}$$

The distribution is therefore constant across the full interval, out to maximum displacements of $L = \pm\frac{\mu\Delta t}{m}\frac{dB_z}{dz}T$.

Classically, we expect the the beam to become a uniform vertical smear on the screen.

1.2 The experimental result

When the experiment is carried out on atoms such as silver or cesium, which have total angular momentum depending only on the spin of a single outer electron, the result is strikingly different from this. Just as the orbital angular momentum within atoms is quantized into discrete total angular momentum and discrete z-components given as integer multiples of $\hbar$, the electron spin is only ever measured to have one of two discrete values, $\pm\frac{1}{2}\hbar$. The Stern-Gerlach beam discribed above produces only two sharp spots, above and below the centerline, and corresponding to a magnetic moment of

$$\mu_{electron} = \frac{ge}{2m_e}S$$

where the *intrinsic spin angular momentum*, or just *spin*, of the electron is $S = \frac{\hbar}{2}$ and the g factor is measured experimentally[1] to be $g = 2.00231930419922 \pm \left(1.5\times10^{-12}\right)$.

2 Implications of sequential Stern-Gerlach measurements

If the initial beam directed into a Stern-Gerlach device, we may easily check that the distribution is random by rotating the orientation of the inhomogeneous magnetic field. No matter what orientation we choose, the beam splits into two bright spots of equal intensity. There is no preferred orientation for a suitably prepared initial beam. We always get two resultant beams, one called “spin up” and the other, “spin down”, measured relative to the magnetic field gradient.

However, if we follow one such division of the beam by a second splitting applied to only the “spin up” part, using a Stern-Gerlach device with the same orientation as the first, only a single, spin up beam results. It is as if the use of the first device has fixed the direction of the subsequent beams. If we rotate the second device (say, orienting the field in the x-direction instead of the z-direction) we once again get two beams. The behavior is much like polarizing lenses.

[1] [http://physics.nist.gov/cgi-bin/cuu/Value?eqae%7Csearch_{f}or=electron+magnetic+moment]

Suppose we introduce two symbols,

$$\begin{aligned} |z;+\rangle &= |+\rangle \\ |z;-\rangle &= |-\rangle \end{aligned}$$

for spin up and spin down, respectively. The z tells us that that we have oriented the magnetic gradient in the z direction. In general, we may write $|\hat{\mathbf{n}};+\rangle, |\hat{\mathbf{n}};-\rangle$ when the field gradient is in the $\hat{\mathbf{n}}$ direction. If no direction is explicitly specified, we always mean the $z$direction.

We can correctly model the effect of a Stern-Gerlach device if we assume the initial beam is an equal mixture of these,

$$\frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$$

and that choosing the up beam after the beam separates acts like a projection,

$$\frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) \longrightarrow \frac{1}{\sqrt{2}}|+\rangle$$

Think of the pair $|+\rangle, |-\rangle$ as basis vectors in a 2-dimensional vector space, with the vector describing the original beam having components $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ in this basis. A second projection in the same orientation leaves the $\frac{1}{\sqrt{2}}|+\rangle$ state unchanged, with no down beam emerging.

Introducing a vector norm, the magnitude of $\frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$ is $\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = 1$. We will soon introduce additional notation for this, but for now we simply use what we know about vectors. Like the norm of the wave function, $\psi^*\psi$, it is the square of the amplitude that is measured. Since the state $\frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$ has amplitude $\frac{1}{\sqrt{2}}$ for both the up and down states, we find probability $\left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$ of measuring either of the two possible spins.

Now suppose we split the $\frac{1}{\sqrt{2}}|+\rangle$ state using a second Stern-Gerlach device oriented in the x-direction. Then this weaker beam, all in the $|z;+\rangle$ direction, breaks *equally* into two new beams, up and down with respect to the x-direction,

$$\frac{1}{\sqrt{2}}|+\rangle \quad \longrightarrow \quad \frac{1}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}(|x;+\rangle + |x;-\rangle)\right]$$

Notice what has happened to the intensities. Since our spin up beam had half the amplitude of the original random beam, each of the $|x;+\rangle$ and $|x;-\rangle$ beams has $\frac{1}{4}$ the overall amplitude.

We may think of the selection of the up beam as a measurement. "*Measuring*" the initial beam, we find it to be up half the time; measuring the resulting up beam, it is always up. Once we have measured the spin property, and determined it to be $|z;+\rangle = |+\rangle$, it remains so. The initial state $\frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$ is a mixture of up and down, essentially a list of what we might possibly measure. After *measuring* the beam to be all spin up, that list is reduced and we have only $\frac{1}{\sqrt{2}}|+\rangle$ left on the list of possible measurements.

If we choose the down beam using our Stern-Gerlach device, we find its intensity is also $\left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$ that of the original beam.

We found that the pure up beam $\frac{1}{\sqrt{2}}|+\rangle$ may be split into two x-orientations, so we may write

$$\frac{1}{\sqrt{2}}|+\rangle = \frac{1}{2}(|x;+\rangle + |x;-\rangle)$$

or simply

$$|+\rangle = \frac{1}{\sqrt{2}}(|x;+\rangle + |x;-\rangle)$$

In this way, we are treating the states as vectors, and the two possible different orientations as different pairs of basis vectors. Any state may be written as a linear combination of $|+\rangle, |-\rangle$, or equally well as a linear combination of $|x;+\rangle, |x;-\rangle$.

3 Spin vectors

Now, we study the relationships between the various pairs of basis vectors more carefully.

Contemplating the splitting of beams using Stern-Gerlach devices oriented in the x, y or z directions, we see that we may write a given spin state as a linear combination of any of three pairs,

$$
\begin{aligned}
&|+\rangle, |-\rangle \\
&|x;+\rangle, |x;-\rangle \\
&|y;+\rangle, |y;-\rangle
\end{aligned}
$$

Furthermore, the six basis states here are mutually orthogonal, in the sense that a pure state in any of the three directions is found to split into *equal measures* of the other two. For example

$$
\begin{aligned}
|x;+\rangle &= \frac{1}{\sqrt{2}}\left(\mu_+ |+\rangle + \mu_+ |-\rangle\right) \\
|x;+\rangle &= \frac{1}{\sqrt{2}}\left(\sigma_+ |y;+\rangle + \sigma_- |y;-\rangle\right)
\end{aligned}
$$

where we must have

$$
\begin{aligned}
|\mu_+| &= |\mu_-| = 1 \\
|\sigma_+| &= |\nu_-| = 1
\end{aligned}
$$

This cannot be accomplished with real coefficients since only ± 1 have norm 1. Since complex numbers have infinitely many numbers with norm 1, we take the spin space to be a *complex two dimensional space.*

Replacing the constants $\mu_\pm$ with phases, the x basis may be written as

$$
\begin{aligned}
|x;+\rangle &= \frac{1}{\sqrt{2}}\left(e^{i\varphi_1} |+\rangle + e^{i\varphi_2} |-\rangle\right) \\
|x;-\rangle &= \frac{1}{\sqrt{2}}\left(e^{i\varphi_1} |+\rangle - e^{i\varphi_2} |-\rangle\right)
\end{aligned}
$$

where the relationship between the spin up and spin down coefficients follows from othogonality, $\langle x;+ \mid x;-\rangle = 0$. Quantum states include an arbitrary overall factor. The real part of this factor is used to normalize the state, leaving an arbitrary phase with no physical consequence. We therefore may multiply these states by $e^{-i\varphi_1}$. Defining $\varphi = \varphi_2 - \varphi_1$,

$$
\begin{aligned}
|x;+\rangle &= \frac{1}{\sqrt{2}}\left(|+\rangle + e^{i\varphi} |-\rangle\right) \\
|x;-\rangle &= \frac{1}{\sqrt{2}}\left(|+\rangle - e^{i\varphi} |-\rangle\right)
\end{aligned}
$$

Inverting these equalities, we have

$$
\begin{aligned}
|+\rangle &= \frac{1}{\sqrt{2}}\left(|x;+\rangle + |x;-\rangle\right) \\
|-\rangle &= \frac{1}{\sqrt{2}} e^{-i\varphi}\left(|x;+\rangle - |x;-\rangle\right)
\end{aligned}
$$

Now consider two ways or writing the spin up state along the y axis. Since it will split equally if measured along either the x- or the z-axis, and we may apply the same phase arguments as for the x-basis, we may write

$$
\begin{aligned}
|y;+\rangle &= \frac{1}{\sqrt{2}}\left(|x;+\rangle + e^{i\xi} |x;-\rangle\right) \\
|y;+\rangle &= \frac{1}{\sqrt{2}}\left(|+\rangle + e^{i\theta} |-\rangle\right)
\end{aligned}
$$

for phases ξ and θ. Equating these up to a phase, then substituting for the x-basis terms,

$$\begin{aligned}
\frac{1}{\sqrt{2}} e^{i\zeta} \left(|+\rangle + e^{i\theta} |-\rangle \right) &= \frac{1}{\sqrt{2}} \left(|x;+\rangle + e^{i\xi} |x;-\rangle \right) \\
&= \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} \left(|+\rangle + e^{i\varphi} |-\rangle \right) + e^{i\xi} \frac{1}{\sqrt{2}} \left(|+\rangle - e^{i\varphi} |-\rangle \right) \right) \\
&= \frac{1}{\sqrt{2}} \left(\frac{1+e^{i\xi}}{\sqrt{2}} |+\rangle + \frac{1-e^{i\xi}}{\sqrt{2}} e^{i\varphi} |-\rangle \right)
\end{aligned}$$

Equating the phases,

$$\begin{aligned}
e^{i\zeta} &= \frac{1+e^{i\xi}}{\sqrt{2}} \\
e^{i\zeta} e^{i\theta} &= \frac{1-e^{i\xi}}{\sqrt{2}} e^{i\varphi}
\end{aligned}$$

The first requires $\frac{1+e^{i\xi}}{\sqrt{2}}$ to have norm 1,

$$\begin{aligned}
\frac{1+e^{i\xi}}{\sqrt{2}} \frac{1+e^{-i\xi}}{\sqrt{2}} &= 1 \\
1 + 2\cos\xi + 1 &= 2 \\
\cos\xi &= 0 \\
\xi &= (2n+1)\frac{\pi}{2}
\end{aligned}$$

With $e^{i\xi} = e^{i(2n+1)\frac{\pi}{2}} = (-1)^n i$, the second becomes

$$\begin{aligned}
\frac{1+(-1)^n i}{\sqrt{2}} e^{i\theta} &= \frac{1-(-1)^n i}{\sqrt{2}} e^{i\varphi} \\
e^{\pm\frac{i\pi}{4}} e^{i\theta} &= e^{\mp\frac{i\pi}{4}} e^{i\varphi} \\
\theta &= \varphi \mp \frac{\pi}{2}
\end{aligned}$$

with the upper sign holding for $n$even and the lower for $n$odd.

Substituting into the expressions for the x- and y-basis,

$$\begin{aligned}
|x;+\rangle &= \frac{1}{\sqrt{2}} \left(|+\rangle + e^{i\varphi} |-\rangle \right) \\
|x;-\rangle &= \frac{1}{\sqrt{2}} \left(|+\rangle - e^{i\varphi} |-\rangle \right) \\
|y;+\rangle &= \frac{1}{\sqrt{2}} \left(|+\rangle \mp i e^{i\varphi} |-\rangle \right) \\
|y;-\rangle &= \frac{1}{\sqrt{2}} \left(|+\rangle \pm i e^{i\varphi} |-\rangle \right)
\end{aligned}$$

The choices of $n$and φ correspond to a rotation about the z-axis and therefore does not affect the essential orthogonality relationships. Choosing the lower sign, and $\varphi = 0$ gives our choice of basis,

$$\begin{aligned}
|x;+\rangle &= \frac{1}{\sqrt{2}} \left(|+\rangle + |-\rangle \right) \\
|x;-\rangle &= \frac{1}{\sqrt{2}} \left(|+\rangle - |-\rangle \right)
\end{aligned}$$

$$
\begin{aligned}
|y;+\rangle &= \frac{1}{\sqrt{2}}\left(|+\rangle + i\,|-\rangle\right) \\
|y;-\rangle &= \frac{1}{\sqrt{2}}\left(|+\rangle - i\,|-\rangle\right)
\end{aligned}
$$