Chapter 2

The Stern-Gerlach Experiment

Let us now talk about a particular property of an atom, called its *magnetic dipole moment.* It is simplest to first recall what an *electric* dipole moment is.

Consider an electrically neutral system consisting of two separated charges, one of charge $+e$ at $\mathbf{r}_+$ and one of charge $-e$ at position $\mathbf{r}_-$. Immerse this system in an external electric field

$$\mathbf{E} = -\boldsymbol{\nabla}\phi, \tag{2.1}$$

where ϕ the electrostatic potential. The energy of this system in the field is

$$\mathcal{E} = e\phi(\mathbf{r}_+) - e\phi(\mathbf{r}_-) \approx e(\mathbf{r}_+ - \mathbf{r}_-) \cdot \boldsymbol{\nabla}\phi(\mathbf{r}), \tag{2.2}$$

where $\mathbf{r}$ is some representative point in the atom, if the atom is very small compared to the distance over which $\mathbf{E}$ varies. Here

$$\boldsymbol{\nabla}\phi = \text{gradient}\,\phi = \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}\right). \tag{2.3}$$

We define $\mathbf{d} = e(\mathbf{r}_+ - \mathbf{r}_-)$ as the electric dipole moment, so

$$\mathcal{E} = -\mathbf{d} \cdot \mathbf{E}. \tag{2.4}$$

What is the force that $\mathbf{E}$ exerts on the atom?

$$\begin{aligned}
\mathbf{F} &= e\mathbf{E}(\mathbf{r}_+) - e\mathbf{E}(\mathbf{r}_-) \approx e(\mathbf{r}_+ - \mathbf{r}_-) \cdot \boldsymbol{\nabla}\mathbf{E}(\mathbf{r}) \\
&= (\mathbf{d} \cdot \boldsymbol{\nabla})\mathbf{E}(\mathbf{r}) = -(\mathbf{d} \cdot \boldsymbol{\nabla})\boldsymbol{\nabla}\phi(\mathbf{r}) \\
&= -\boldsymbol{\nabla}(\mathbf{d} \cdot \boldsymbol{\nabla})\phi(\mathbf{r}) = \boldsymbol{\nabla}(\mathbf{d} \cdot \mathbf{E}) = -\boldsymbol{\nabla}\mathcal{E}.
\end{aligned} \tag{2.5}$$

Note here that $\mathbf{d}$, the electric dipole moment, is a property of the atom, independent of where the atom is.

In addition, the electric field exerts a torque on the atom. About the center of mass of the atom, this is

$$\begin{aligned}\boldsymbol{\tau} &= \sum_i \mathbf{r}_i \times \mathbf{F}_i = \frac{1}{2}(\mathbf{r}_+ - \mathbf{r}_-) \times e\mathbf{E}(\mathbf{r}_+) - \frac{1}{2}(\mathbf{r}_+ - \mathbf{r}_-) \times (-e)\mathbf{E}(\mathbf{r}_-) \\ &= e(\mathbf{r}_+ - \mathbf{r}_-) \times \mathbf{E}(\mathbf{r}), \end{aligned} \tag{2.6}$$

in terms of the average electric field in the atom,

$$\mathbf{E}(\mathbf{r}) = \frac{1}{2}\left[\mathbf{E}(\mathbf{r}_+) + \mathbf{E}(\mathbf{r}_-)\right]. \tag{2.7}$$

Thus, the torque on the atom is

$$\boldsymbol{\tau} = \mathbf{d} \times \mathbf{E}(\mathbf{r}). \tag{2.8}$$

Torque is also derivable from the energy, by considering how the energy changes under a rotation of the dipole:

$$\mathcal{E} = -\mathbf{d} \cdot \mathbf{E} = -Ed\cos\theta. \tag{2.9}$$

where θ is the angle between $\mathbf{d}$ and $\mathbf{E}$. Then the torque on the dipole is

$$\tau = -\frac{\partial}{\partial\theta}\mathcal{E} = \frac{\partial}{\partial\theta}Ed\cos\theta = -Ed\sin\theta. \tag{2.10}$$

The negative sign means that the torque acts in such a direction as to decrease θ.

This was all by way of a prologue. What we really want to discuss are magnetic dipoles. But we can proceed by analogy:

- Electric dipoles and electric fields are described by $\mathbf{d}$ and $\mathbf{E}$.
- Magnetic dipoles and magnetic fields are described by $\boldsymbol{\mu}$ and $\mathbf{H}$.
- So to transcribe results for electric fields and dipoles to magnetic fields and dipoles we simply make the replacement:

$$\mathbf{d} \to \boldsymbol{\mu}, \quad \mathbf{E} \to \mathbf{H}. \tag{2.11}$$

So the corresponding results for a magnetic dipole in a magnetic field is

$$\mathcal{E} = -\boldsymbol{\mu} \cdot \mathbf{H}, \tag{2.12a}$$

$$\mathbf{F} = (\boldsymbol{\mu} \cdot \boldsymbol{\nabla})\mathbf{H} = \boldsymbol{\nabla}(\boldsymbol{\mu} \cdot \mathbf{H}), \tag{2.12b}$$

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{H}. \tag{2.12c}$$

How does one measure the magnetic moment of an individual atom? By using a beam of atoms. In an atomic beam, one can do repeated measurements all at once, but the atoms are widely enough separated that we can ignore the interactions between them. The experimental apparatus is schematically as

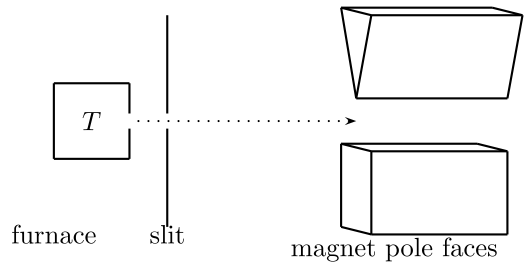

Figure 2.1: Sketch of a Stern-Gerlach experiment. Atoms are heated in the furnace, exit through a small hole, and collimated by a slit in an otherwise opaque screen. The atomic beam then enters an inhomogeneous magnetic field, where the atoms are deflected because of the interaction between the dipole moment of the atoms and the magnetic field.

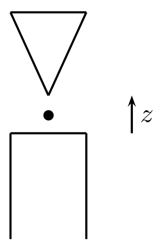

Figure 2.2: Sketch of the end-on view of the magnet. The dot represents the position of the beam traveling down the axis of the magnet.

shown in Fig. 2.1. Because of the knife edge of the upper pole piece, the magnetic field gets stronger as that knife edge is approached. Because the atoms have magnetic dipole moments, and the field is inhomogeneous, a force is exerted on the atoms when they enter the region between the pole pieces:

$$\mathbf{F} = \boldsymbol{\nabla}(\boldsymbol{\mu} \cdot \mathbf{H}). \tag{2.13}$$

The beam is arranged so that it goes down the central axis of the apparatus, and it therefore only sees a z component of the field, which depends only on z, $H_z(z)$, where z is the symmetry direction of the magnet: See Fig. 2.2. H_z increases as z increases. Because of the symmetry, the force on the atom has only a z component,

$$F_z = \frac{\partial}{\partial z}\mu_z H_z = \mu_z \frac{\partial H_z}{\partial z}. \tag{2.14}$$

We know $\partial H_z/\partial z$ by knowing how the magnet is constructed, what the current through its coils are, etc., and we measure F_z, as now described. This enables us to perform a determination of μ_z.

F_z is measured from the deflection of the atoms in the beam. The change in the momentum is given by Newton's law,

$$\frac{dp_z}{dt} = \mu_z \frac{\partial H_z}{\partial z}. \tag{2.15}$$

Initially, the atom has no momentum in the z direction. Suppose $\partial H_z/\partial z > 0$; then if $\mu_z > 0$ the atom is deflected up, and if $\mu_z < 0$ it is deflected down. Since the deflection is small, $\partial H_z/\partial z$ is nearly constant along the trajectory of the atom. If the length of the magnet is ℓ, the time during which the atom is inside the magnet is

$$\Delta t = \frac{\ell}{v}, \tag{2.16}$$

if the speed of the atoms in the beam is v, which hardly changes. So the net momentum acquired in the z direction is

$$p_z = \mu_z \frac{\partial H_z}{\partial z} \Delta t = \mu_z \frac{\partial H_z}{\partial z} \frac{\ell}{v}. \tag{2.17}$$

The beam is deflected upward or downward by an angle

$$\theta = \frac{p_z}{p} = \frac{p_z}{mv}, \tag{2.18}$$

where m is the mass of the atom, or

$$\theta = \mu_z \frac{\partial H_z}{\partial z} \frac{\ell}{mv^2}. \tag{2.19}$$

Take some typical numbers: $\frac{1}{2}mv^2 \sim kT$, where T is the temperature of the furnace, say 10^3 K. The magnetic moment is something like $\mu \sim 10^{-20}$ erg/G, the field gradient might be $\partial H_z/\partial z \sim 10^4$ G/cm, and the length $\ell \sim 10$ cm. This gives an estimate

$$\theta \sim \frac{10^{-20} \times 10^4 \times 10}{10^{-16} \times 10^3} = 10^{-2} \tag{2.20}$$

radians, or about $1/2°$. That means about a 1 cm deflection over a distance of 1 m, which is easily observable.

We would expect that $\mu_z = \mu \cos\theta$, where now $\cos\theta$ is the direction between the direction of the magnetic field, the z axis, and the direction of the dipole, ranges continuously in the beam from $-\mu$ to $+\mu$, since atoms emitted from the furnace are randomly polarized. This would imply a continuous distribution of deflections, as sketched in Fig. 2.3. But when the experiment was performed by Stern and Gerlach in 1921, using silver (Ag) atoms, they found discrete rather than continuous deflections. As shown in Fig. 2.4, either the atoms were deflected up by a certain amount, or deflected down by the same amount, with nothing in between. It is as though $\mu_z = \pm\mu$ only. This is an abrupt break with classical physics.

Of course, there is nothing special about *up* and *down*; if the apparatus lay in a horizontal plane we would get deflection left and right.

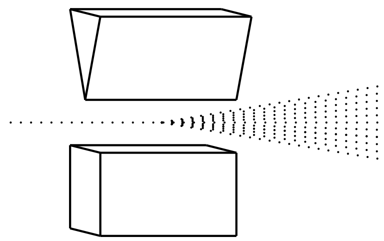

Figure 2.3: The expected distribution of atoms deflected by the Stern-Gerlach measurement

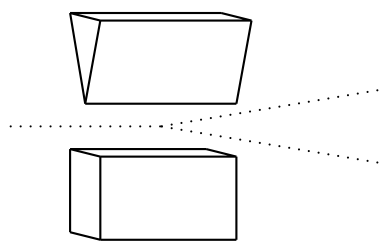

Figure 2.4: The actual distribution of atoms deflected by the Stern-Gerlach measurement

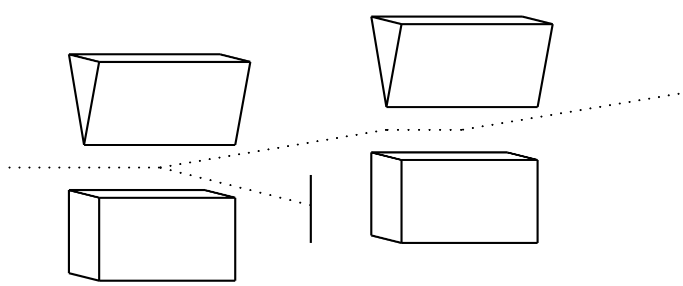

Figure 2.5: A first Stern-Gerlach experiment selects atoms with $\mu_z = +\mu$, and blocks those atoms deflected down, with $\mu = -\mu$. A second apparatus again measures μ_z for that selected beam. All the atoms emerging from the second apparatus are deflected up.

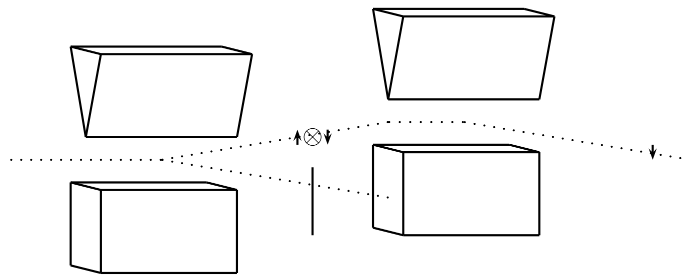

Figure 2.6: Now the atoms emerging from the first Stern-Gerlach apparatus are rotated 180° about the beam axis, so that they enter the second apparatus with $\mu_z = -\mu$. The symbol $\otimes$ represents the magnetic field apparatus that rotates the dipole moments of the atoms emerging from the first Stern-Gerlach apparatus by 180° before they enter the second Stern-Gerlach apparatus. All the atoms emerging from the second apparatus are deflected down.

Suppose we did a repeated measurement, where the beam of those atoms which were deflected up were sent through an second Stern-Gerlach apparatus, which also measure μ_z, as shown in Fig. 2.5. The second apparatus deflects all the atoms with $\mu_z = +\mu$ upward, by the same amount.

If the second apparatus (the "analyzing magnet") were rotated through an angle $\pi = 180°$ about the beam axis, all the atoms entering the analyzing magnet would be deflected down by it. In practice, one doesn't have to rotate the magnet, but one can rotate the orientation of the dipoles by applying a torque on them with an additional magnetic field. The second deflection above is equivalent to the deflection of downward polarized dipoles deflected downward by an upwardly oriented apparatus, as sketched in Fig. 2.6.

What if the analyzing magnet is rotated about the beam axis by an angle θ, so that the atoms entering that apparatus have an orientation misaligned the the magnet axis, as seen in Fig. 2.7. In this case, atoms will be deflected both "up" and "down", where the directions refer to the symmetry axis of the analyzing magnet. We know that if $\theta = 0$, all atoms will be deflected up, and if $\theta = \pi$ all atoms will be deflected down. $\theta = \pi/2 = 90°$ is a symmetrical situation; equal numbers of atoms will be deflected up and down. What does an individual atom do? It is either deflected up or down, and certainly it does not split in two. We *cannot predict* what it will do.

There is an analogy here with the throwing of a die. We cannot predict an individual outcome, but we can say that the *probability* of a particular face appearing is 1/6. But with a die, in principle, we could predict the outcome, if we knew the initial conditions of throwing the die well enough. This is not the case with atoms. We cannot get beyond probabilities. We cannot get further information to tell us what a particular atom will do. The deterministic aspect of mechanics is lost. Individual events are not predictable; we can only predict

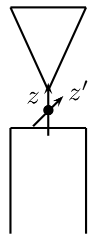

Figure 2.7: Now the atoms emerging from the first Stern-Gerlach apparatus are rotated by an angle of θ about the beam axis, relative to the axis of the second analyzing magnet. This is an end-on view of the second apparatus, showing the atoms entering, polarized in the direction z', where the symmetry axis of the second apparatus, the z direction, makes an angle θ with respect to the z' direction. Now some of the atoms emerging from the second apparatus will be deflected up, and some down.

outcomes of large numbers of experiments.

We'll see that we already know enough to calculate the probabilities for deflection up or down for an arbitrary dipole orientation θ. On the average, for a large number of atoms, the projection of $\boldsymbol{\mu}$ on the z axis is

$$\langle \mu_z \rangle_{+z'} = \mu \cos\theta, \tag{2.21}$$

where the subscript means that the first apparatus, oriented in the z' direction, selected atoms with $\mu_{z'} = +\mu$. The relation between the axes of the two apparatuses is shown in Fig. 2.7. The first is oriented in the z' direction, and selects either $\mu_{z'} = \pm\mu$; the second is oriented in the z direction, and selects either $\mu_z = \pm\mu$, where θ is the angle between these two directions. Now $\langle \mu_z \rangle_{+z'}$ means the average value of μ_z measured by the analyzing magnet, given that the atoms entered that magnet with $\mu_{z'} = +\mu$. Thus, by the meaning of probability,

$$\mu \cos\theta = +\mu\, p(+,+) - \mu\, p(-,+), \tag{2.22}$$

where $p(\pm,+)$ is the probability of finding the dipole with $\mu_z = \pm\mu$ when the incoming atom has $\mu_{z'} = +\mu$. Since the atoms come out either deflected up or deflected down,

$$1 = p(+,+) + p(-,+). \tag{2.23}$$

We can solve these two linear equations for the individual probabilities:

$$p(+,+) = \frac{1+\cos\theta}{2} = \cos^2\frac{\theta}{2}, \tag{2.24a}$$

$$p(-,+) = \frac{1-\cos\theta}{2} = \sin^2\frac{\theta}{2}. \tag{2.24b}$$

Evidently, the previous cases are reproduced:

$$\theta = 0: \quad p(+,+) = 1, \quad p(-,+) = 0, \tag{2.25a}$$

$$\theta = \pi : \quad p(+,+) = 0, \quad p(-,+) = 1, \tag{2.25b}$$

$$\theta = \frac{\pi}{2} : \quad p(+,+) = \frac{1}{2}, \quad p(-,+) = \frac{1}{2}. \tag{2.25c}$$

If the incoming atoms had orientation along the $-z'$ axis, $\mu_{z'} = -\mu$, results are obtained from the above by replacing $\cos\theta \to \cos(\pi - \theta) = -\cos\theta$:

$$p(+,-) = \frac{1-\cos\theta}{2} = \sin^2\frac{\theta}{2} = p(-,+), \tag{2.26a}$$

$$p(-,-) = \frac{1+\cos\theta}{2} = \cos^2\frac{\theta}{2} = p(-,-), \tag{2.26b}$$

where for example, $p(-,-)$ refers an initial measurement that selects $\mu_{z'} = -\mu$, followed by a second measurement which gives $\mu_{z'} = -\mu$. Evidently, the successive measurements $+$, $+$ are geometrically the same as $-$, $-$. There is really only one independent function here. A significant test of consistency here is that all the probabilities, as they must, lie between zero and one; the solution to our two simultaneous equations might have turned out otherwise.

The results of two successive Stern Gerlach measurements, the first of which selects magnetic moments along the z' direction, the second along the direction z, the angles between these two directions being θ, are summarized by the following probability array:

$$p(\pm,\pm) = \begin{pmatrix} \cos^2\frac{\theta}{2} & \sin^2\frac{\theta}{2} \\ \sin^2\frac{\theta}{2} & \cos^2\frac{\theta}{2} \end{pmatrix}. \tag{2.27}$$