CHAPTER

1 Fundamental Concepts

The revolutionary change in our understanding of microscopic phenomena that took place during the first 27 years of the twentieth century is unprecedented in the history of natural sciences. Not only did we witness severe limitations in the validity of classical physics, but we found the alternative theory that replaced the classical physical theories to be far broader in scope and far richer in its range of applicability.

The most traditional way to begin a study of quantum mechanics is to follow the historical developments—Planck's radiation law, the Einstein-Debye theory of specific heats, the Bohr atom, de Broglie's matter waves, and so forth—together with careful analyses of some key experiments such as the Compton effect, the Franck-Hertz experiment, and the Davisson-Germer-Thompson experiment. In that way we may come to appreciate how the physicists in the first quarter of the twentieth century were forced to abandon, little by little, the cherished concepts of classical physics and how, despite earlier false starts and wrong turns, the great masters—Heisenberg, Schrödinger, and Dirac, among others—finally succeeded in formulating quantum mechanics as we know it today.

However, we do not follow the historical approach in this book. Instead, we start with an example that illustrates, perhaps more than any other example, the inadequacy of classical concepts in a fundamental way. We hope that, exposing readers to a "shock treatment" at the onset will result in their becoming attuned to what we might call the "quantum-mechanical way of thinking" at a very early stage.

This different approach is not merely an academic exercise. Our knowledge of the physical world comes from making assumptions about nature, formulating these assumptions into postulates, deriving predictions from those postulates, and testing such predictions against experiment. If experiment does not agree with the prediction, then, presumably, the original assumptions were incorrect. Our approach emphasizes the fundamental assumptions we make about nature, upon which we have come to base all of our physical laws, and which aim to accommodate profoundly quantum-mechanical observations at the outset.

1.1 ■ THE STERN-GERLACH EXPERIMENT

The example we concentrate on in this section is the Stern-Gerlach experiment, originally conceived by O. Stern in 1921 and carried out in Frankfurt by him in

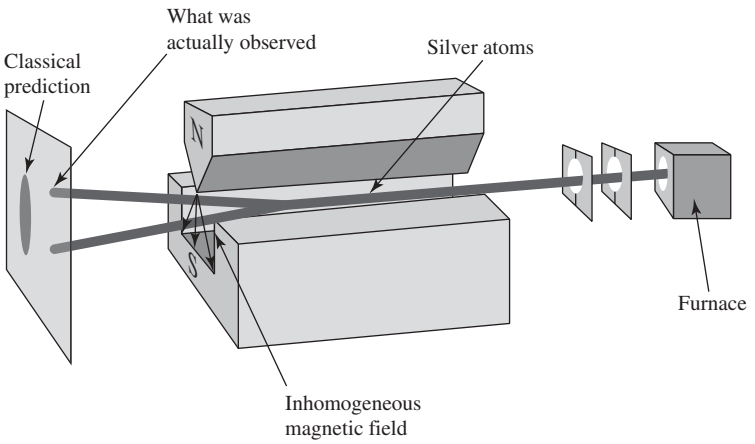

FIGURE 1.1 The Stern-Gerlach experiment.

collaboration with W. Gerlach in 1922.* This experiment illustrates in a dramatic manner the necessity for a radical departure from the concepts of classical mechanics. In the subsequent sections the basic formalism of quantum mechanics is presented in a somewhat axiomatic manner but always with the example of the Stern-Gerlach experiment in the back of our minds. In a certain sense, a two-state system of the Stern-Gerlach type is the least classical, most quantum-mechanical system. A solid understanding of problems involving two-state systems will turn out to be rewarding to any serious student of quantum mechanics. It is for this reason that we refer repeatedly to two-state problems throughout this book.

Description of the Experiment

We now present a brief discussion of the Stern-Gerlach experiment, which is discussed in almost every book on modern physics.† First, silver (Ag) atoms are heated in an oven. The oven has a small hole through which some of the silver atoms escape. As shown in Figure 1.1, the beam goes through a collimator and is then subjected to an inhomogeneous magnetic field produced by a pair of pole pieces, one of which has a very sharp edge.

We must now work out the effect of the magnetic field on the silver atoms. For our purpose the following oversimplified model of the silver atom suffices. The silver atom is made up of a nucleus and 47 electrons, where 46 out of the 47 electrons can be visualized as forming a spherically symmetrical electron cloud with no net angular momentum. If we ignore the nuclear spin, which is irrelevant to our discussion, we see that the atom as a whole does have an angular momentum, which is due solely to the spin—intrinsic as opposed to orbital—angular

*For an excellent historical discussion of the Stern-Gerlach experiment, see "Stern and Gerlach: How a Bad Cigar Helped Reorient Atomic Physics," by Bretislav Friedrich and Dudley Herschbach, *Physics Today*, December (2003) 53.

†For an elementary but enlightening discussion of the Stern-Gerlach experiment, see French and Taylor (1978), pp. 432–38.

momentum of the single 47th ($5s$) electron. The 47 electrons are attached to the nucleus, which is $\sim 2 \times 10^5$ times heavier than the electron; as a result, the heavy atom as a whole possesses a magnetic moment equal to the spin magnetic moment of the 47th electron. In other words, the magnetic moment $\boldsymbol{\mu}$ of the atom is proportional to the electron spin $\mathbf{S}$,

$$\boldsymbol{\mu} \propto \mathbf{S}, \tag{1.1.1}$$

where the precise proportionality factor turns out to be $e/m_e c$ ($e < 0$ in this book) to an accuracy of about 0.2%.

Because the interaction energy of the magnetic moment with the magnetic field is just $-\boldsymbol{\mu}\cdot\mathbf{B}$, the z-component of the force experienced by the atom is given by

$$F_z = \frac{\partial}{\partial z}(\boldsymbol{\mu}\cdot\mathbf{B}) \simeq \mu_z \frac{\partial B_z}{\partial z}, \tag{1.1.2}$$

where we have ignored the components of $\mathbf{B}$ in directions other than the z-direction. Because the atom as a whole is very heavy, we expect that the classical concept of trajectory can be legitimately applied, a point that can be justified using the Heisenberg uncertainty principle to be derived later. With the arrangement of Figure 1.1, the $\mu_z > 0$ ($S_z < 0$) atom experiences a downward force, while the $\mu_z < 0$ ($S_z > 0$) atom experiences an upward force. The beam is then expected to get split according to the values of μ_z. In other words, the SG (Stern-Gerlach) apparatus "measures" the z-component of $\boldsymbol{\mu}$ or, equivalently, the z-component of $\mathbf{S}$ up to a proportionality factor.

The atoms in the oven are randomly oriented; there is no preferred direction for the orientation of $\boldsymbol{\mu}$. If the electron were like a classical spinning object, we would expect all values of μ_z to be realized between $|\boldsymbol{\mu}|$ and $-|\boldsymbol{\mu}|$. This would lead us to expect a continuous bundle of beams coming out of the SG apparatus, as indicated in Figure 1.1, spread more or less evenly over the expected range. Instead, what we experimentally observe is more like the situation also shown in Figure 1.1, where two "spots" are observed, corresponding to one "up" and one "down" orientation. In other words, the SG apparatus splits the original silver beam from the oven into *two distinct* components, a phenomenon referred to in the early days of quantum theory as "space quantization." To the extent that $\boldsymbol{\mu}$ can be identified within a proportionality factor with the electron spin $\mathbf{S}$, only two possible values of the z-component of $\mathbf{S}$ are observed to be possible: S_z up and S_z down, which we call S_z+ and S_z-. The two possible values of S_z are multiples of some fundamental unit of angular momentum; numerically it turns out that $S_z = \hbar/2$ and $-\hbar/2$, where

$$\begin{aligned} \hbar &= 1.0546 \times 10^{-27}\,\text{erg-s} \\ &= 6.5822 \times 10^{-16}\,\text{eV-s}. \end{aligned} \tag{1.1.3}$$

This "quantization" of the electron spin angular momentum* is the first important feature we deduce from the Stern-Gerlach experiment.

*An understanding of the roots of this quantization lies in the application of relativity to quantum mechanics. See Section 8.2 of this book for a discussion.

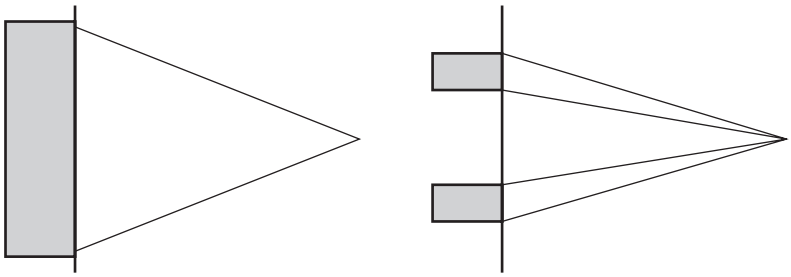

FIGURE 1.2 (a) Classical physics prediction for results from the Stern-Gerlach experiment. The beam should have been spread out vertically, over a distance corresponding to the range of values of the magnetic moment times the cosine of the orientation angle. Stern and Gerlach, however, observed the result in (b), namely that only two orientations of the magnetic moment manifested themselves. These two orientations did not span the entire expected range.

Figure 1.2a shows the result one would have expected from the experiment. According to classical physics, the beam should have spread itself over a vertical distance corresponding to the (continuous) range of orientation of the magnetic moment. Instead, one observes Figure 1b, which is completely at odds with classical physics. The beam mysteriously splits itself into two parts, one corresponding to spin "up" and the other to spin "down."

Of course, there is nothing sacred about the up-down direction or the z-axis. We could just as well have applied an inhomogeneous field in a horizontal direction, say in the x-direction, with the beam proceeding in the y-direction. In this manner we could have separated the beam from the oven into an S_x+ component and an S_x- component.

Sequential Stern-Gerlach Experiments

Let us now consider a sequential Stern-Gerlach experiment. By this we mean that the atomic beam goes through two or more SG apparatuses in sequence. The first arrangement we consider is relatively straightforward. We subject the beam coming out of the oven to the arrangement shown in Figure 1.3a, where SG$\hat{\mathbf{z}}$ stands for an apparatus with the inhomogeneous magnetic field in the z-direction, as usual. We then block the S_z- component coming out of the first SG$\hat{\mathbf{z}}$ apparatus and let the remaining S_z+ component be subjected to another SG$\hat{\mathbf{z}}$ apparatus. This time there is only one beam component coming out of the second apparatus—just the S_z+ component. This is perhaps not so surprising; after all, if the atom spins are up, they are expected to remain so, short of any external field that rotates the spins between the first and the second SG$\hat{\mathbf{z}}$ apparatuses.

A little more interesting is the arrangement shown in Figure 1.3b. Here the first SG apparatus is the same as before, but the second one (SG$\hat{\mathbf{x}}$) has an inhomogeneous magnetic field in the x-direction. The S_z+ beam that enters the second apparatus (SG$\hat{\mathbf{x}}$) is now split into two components, an S_x+ component and an

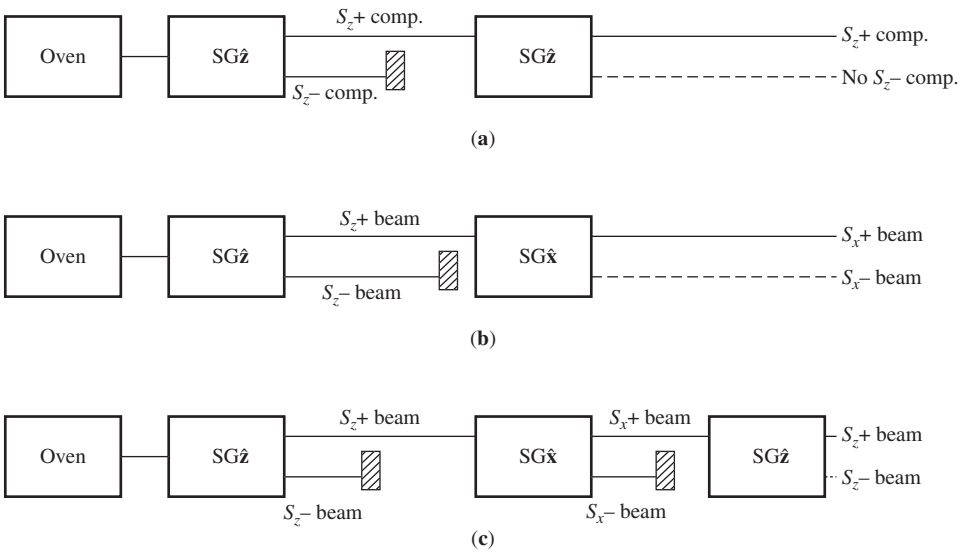

FIGURE 1.3 Sequential Stern-Gerlach experiments.

S_x- component, with equal intensities. How can we explain this? Does it mean that 50% of the atoms in the S_z+ beam coming out of the first apparatus (SG$\hat{\mathbf{z}}$) are made up of atoms characterized by both S_z+ and S_x+, while the remaining 50% have both S_z+ and S_x-? It turns out that such a picture runs into difficulty, as we will see below.

We now consider a third step, the arrangement shown in Figure 1.3c, which most dramatically illustrates the peculiarities of quantum-mechanical systems. This time we add to the arrangement of Figure 1.3b yet a third apparatus, of the SG$\hat{\mathbf{z}}$ type. It is observed experimentally that *two* components emerge from the third apparatus, not one; the emerging beams are seen to have *both* an S_z+ component and an S_z- component. This is a complete surprise because after the atoms emerged from the first apparatus, we made sure that the S_z- component was completely blocked. How is it possible that the S_z- component, which we thought, we eliminated earlier, reappears? The model in which the atoms entering the third apparatus are visualized to have both S_z+ and S_x+ is clearly unsatisfactory.

This example is often used to illustrate that in quantum mechanics we cannot determine both S_z and S_x simultaneously. More precisely, we can say that the selection of the S_x+ beam by the second apparatus (SG$\hat{\mathbf{x}}$) completely destroys any *previous* information about S_z.

It is amusing to compare this situation with that of a spinning top in classical mechanics, where the angular momentum

$$\mathbf{L} = I\boldsymbol{\omega} \tag{1.1.4}$$

can be measured by determining the components of the angular-velocity vector $\boldsymbol{\omega}$. By observing how fast the object is spinning in which direction, we can determine ω_x, ω_y, and ω_z simultaneously. The moment of inertia I is computable if we

know the mass density and the geometric shape of the spinning top, so there is no difficulty in specifying both L_z and L_x in this classical situation.

It is to be clearly understood that the limitation we have encountered in determining S_z and S_x is not due to the incompetence of the experimentalist. We cannot make the S_z- component out of the third apparatus in Figure 1.3c disappear by improving the experimental techniques. The peculiarities of quantum mechanics are imposed upon us by the experiment itself. The limitation is, in fact, inherent in microscopic phenomena.

Analogy with Polarization of Light

Because this situation looks so novel, some analogy with a familiar classical situation may be helpful here. To this end we now digress to consider the polarization of light waves. This analogy will help us develop a mathematical framework for formulating the postulates of quantum mechanics.

Consider a monochromatic light wave propagating in the z-direction. A linearly polarized (or plane polarized) light with a polarization vector in the x-direction, which we call for short an *x-polarized light*, has a space-time–dependent electric field oscillating in the x-direction

$$\mathbf{E} = E_0\hat{\mathbf{x}}\cos(kz - \omega t). \tag{1.1.5}$$

Likewise, we may consider a y-polarized light, also propagating in the z-direction,

$$\mathbf{E} = E_0\hat{\mathbf{y}}\cos(kz - \omega t). \tag{1.1.6}$$

Polarized light beams of type (1.1.5) or (1.1.6) can be obtained by letting an unpolarized light beam go through a Polaroid filter. We call a filter that selects only beams polarized in the x-direction an *x-filter*. An x-filter, of course, becomes a y-filter when rotated by 90° about the propagation (z) direction. It is well known that when we let a light beam go through an x-filter and subsequently let it impinge on a y-filter, no light beam comes out (provided, of course, that we are dealing with 100% efficient Polaroids); see Figure 1.4a.

The situation is even more interesting if we insert between the x-filter and the y-filter yet another Polaroid that selects only a beam polarized in the direction—which we call the x'-direction—that makes an angle of 45° with the x-direction in the xy-plane; see Figure 1.4b. This time, there is a light beam coming out of the y-filter despite the fact that right after the beam went through the x-filter it did not have any polarization component in the y-direction. In other words, once the x'-filter intervenes and selects the x'-polarized beam, it is immaterial whether the beam was previously x-polarized. The selection of the x'-polarized beam by the second Polaroid destroys any previous information on light polarization. Notice that this situation is quite analogous to the situation that we encountered earlier with the SG arrangement of Figure 1.3b, provided that the following correspondence is made:

$$\begin{aligned} S_z \pm \text{atoms} &\leftrightarrow x\text{-},\ y\text{-polarized light} \\ S_x \pm \text{atoms} &\leftrightarrow x'\text{-},\ y'\text{-polarized light,} \end{aligned} \tag{1.1.7}$$

where the x'- and y'-axes are defined as in Figure 1.5.

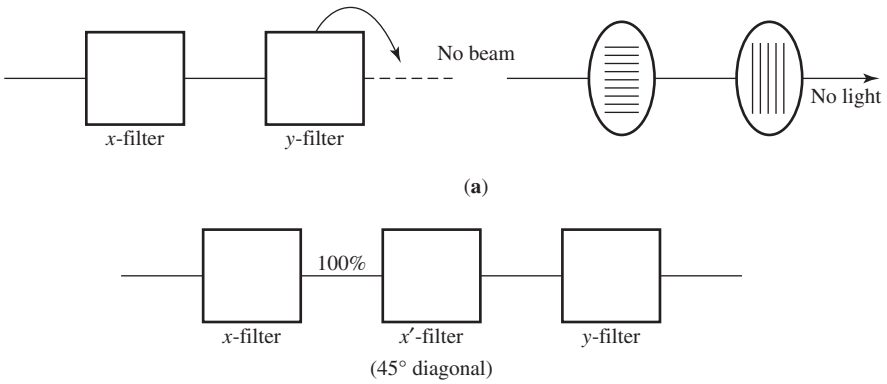

FIGURE 1.4 Light beams subjected to Polaroid filters.

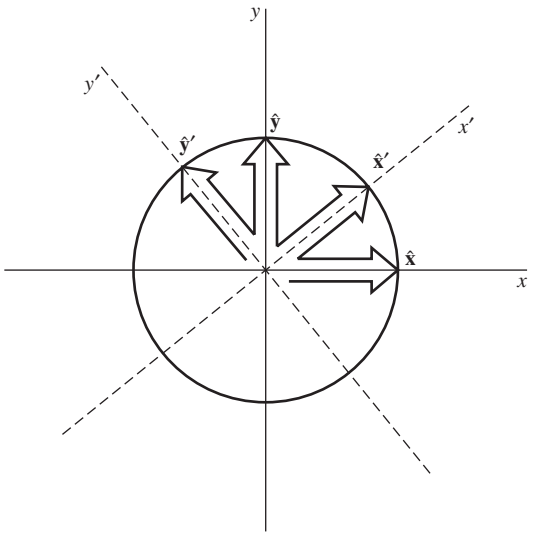

FIGURE 1.5 Orientations of the x'- and y'-axes.

Let us examine how we can quantitatively describe the behavior of 45°-polarized beams (x'- and y'-polarized beams) within the framework of classical electrodynamics. Using Figure 1.5 we obtain

$$\begin{aligned} E_0\hat{\mathbf{x}}'\cos(kz-\omega t) &= E_0\left[\frac{1}{\sqrt{2}}\hat{\mathbf{x}}\cos(kz-\omega t)+\frac{1}{\sqrt{2}}\hat{\mathbf{y}}\cos(kz-\omega t)\right], \\ E_0\hat{\mathbf{y}}'\cos(kz-\omega t) &= E_0\left[-\frac{1}{\sqrt{2}}\hat{\mathbf{x}}\cos(kz-\omega t)+\frac{1}{\sqrt{2}}\hat{\mathbf{y}}\cos(kz-\omega t)\right]. \end{aligned} \tag{1.1.8}$$

In the triple-filter arrangement of Figure 1.4b, the beam coming out of the first Polaroid is an $\hat{\mathbf{x}}$-polarized beam, which can be regarded as a linear combination of an x'-polarized beam and a y'-polarized beam. The second Polaroid selects the x'-polarized beam, which can in turn be regarded as a linear combination of an x-polarized and a y-polarized beam. And finally, the third Polaroid selects the y-polarized component.

Applying correspondence (1.1.7) from the sequential Stern-Gerlach experiment of Figure 1.3c to the triple-filter experiment of Figure 1.4b suggests that we might be able to represent the spin state of a silver atom by some kind of vector in a new kind of two-dimensional vector space, an abstract vector space not to be confused with the usual two-dimensional (xy) space. Just as $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ in (1.1.8) are the base vectors used to decompose the polarization vector $\hat{\mathbf{x}}'$ of the $\hat{\mathbf{x}}'$-polarized light, it is reasonable to represent the S_x+ state by a vector, which we call a ket in the Dirac notation to be developed fully in the next section. We denote this vector by $|S_x;+\rangle$ and write it as a linear combination of two base vectors, $|S_z;+\rangle$ and $|S_z;-\rangle$, which correspond to the S_z+ and the S_z- states, respectively. So we may conjecture

$$|S_x;+\rangle \stackrel{?}{=} \frac{1}{\sqrt{2}}|S_z;+\rangle + \frac{1}{\sqrt{2}}|S_z;-\rangle \tag{1.1.9a}$$

$$|S_x;-\rangle \stackrel{?}{=} -\frac{1}{\sqrt{2}}|S_z;+\rangle + \frac{1}{\sqrt{2}}|S_z;-\rangle \tag{1.1.9b}$$

in analogy with (1.1.8). Later we will show how to obtain these expressions using the general formalism of quantum mechanics.

Thus the unblocked component coming out of the second (SG$\hat{\mathbf{x}}$) apparatus of Figure 1.3c is to be regarded as a superposition of S_z+ and S_z- in the sense of (1.1.9a). It is for this reason that two components emerge from the third (SG$\hat{\mathbf{z}}$) apparatus.

The next question of immediate concern is, How are we going to represent the $S_y\pm$ states? Symmetry arguments suggest that if we observe an $S_z\pm$ beam going in the x-direction and subject it to an SG$\hat{\mathbf{y}}$ apparatus, the resulting situation will be very similar to the case where an $S_z\pm$ beam going in the y-direction is subjected to an SG$\hat{\mathbf{x}}$ apparatus. The kets for $S_y\pm$ should then be regarded as a linear combination of $|S_z;\pm\rangle$, but it appears from (1.1.9) that we have already used up the available possibilities in writing $|S_x;\pm\rangle$. How can our vector space formalism distinguish $S_y\pm$ states from $S_x\pm$ states?

An analogy with polarized light again rescues us here. This time we consider a circularly polarized beam of light, which can be obtained by letting a linearly polarized light pass through a quarter-wave plate. When we pass such a circularly polarized light through an x-filter or a y-filter, we again obtain either an x-polarized beam or a y-polarized beam of equal intensity. Yet everybody knows that the circularly polarized light is totally different from the 45°-linearly polarized (x'-polarized or y'-polarized) light.

Mathematically, how do we represent a circularly polarized light? A right circularly polarized light is nothing more than a linear combination of an x-polarized

light and a y-polarized light, where the oscillation of the electric field for the y-polarized component is 90° out of phase with that of the x-polarized component:*

$$\mathbf{E} = E_0\left[\frac{1}{\sqrt{2}}\hat{\mathbf{x}}\cos(kz-\omega t)+\frac{1}{\sqrt{2}}\hat{\mathbf{y}}\cos\left(kz-\omega t+\frac{\pi}{2}\right)\right]. \tag{1.1.10}$$

It is more elegant to use complex notation by introducing $\boldsymbol{\epsilon}$ as follows:

$$\mathrm{Re}(\boldsymbol{\epsilon}) = \mathbf{E}/E_0. \tag{1.1.11}$$

For a right circularly polarized light, we can then write

$$\boldsymbol{\epsilon} = \left[\frac{1}{\sqrt{2}}\hat{\mathbf{x}}e^{i(kz-\omega t)}+\frac{i}{\sqrt{2}}\hat{\mathbf{y}}e^{i(kz-\omega t)}\right], \tag{1.1.12}$$

where we have used $i = e^{i\pi/2}$.

We can make the following analogy with the spin states of silver atoms:

$$\begin{aligned} &S_y + \text{atom} \leftrightarrow \text{right circularly polarized beam},\\ &S_y - \text{atom} \leftrightarrow \text{left circularly polarized beam}. \end{aligned} \tag{1.1.13}$$

Applying this analogy to (1.1.12), we see that if we are allowed to make the coefficients preceding base kets complex, there is no difficulty in accommodating the $S_y\pm$ atoms in our vector space formalism:

$$|S_y;\pm\rangle \stackrel{?}{=} \frac{1}{\sqrt{2}}|S_z;+\rangle \pm \frac{i}{\sqrt{2}}|S_z;-\rangle, \tag{1.1.14}$$

which are obviously different from (1.1.9). We thus see that the two-dimensional vector space needed to describe the spin states of silver atoms must be a *complex* vector space; an arbitrary vector in the vector space is written as a linear combination of the base vectors $|S_z;\pm\rangle$ with, in general, complex coefficients. The fact that the necessity of complex numbers is already apparent in such an elementary example is rather remarkable.

The reader must have noted by this time that we have deliberately avoided talking about photons. In other words, we have completely ignored the quantum aspect of light; nowhere did we mention the polarization states of individual photons. The analogy we worked out is between kets in an abstract vector space that describes the spin states of individual atoms with the polarization vectors of the *classical electromagnetic field*. Actually, we could have made the analogy even more vivid by introducing the photon concept and talking about the probability of finding a circularly polarized photon in a linearly polarized state, and so forth; however, that is not needed here. Without doing so, we have already accomplished the main goal of this section: to introduce the idea that quantum-mechanical states are to be represented by vectors in an abstract complex vector space.†

*Unfortunately, there is no unanimity in the definition of right versus left circularly polarized light in the literature.

†The reader who is interested in grasping the basic concepts of quantum mechanics through a careful study of photon polarization may find Chapter 1 of Baym (1969) extremely illuminating.

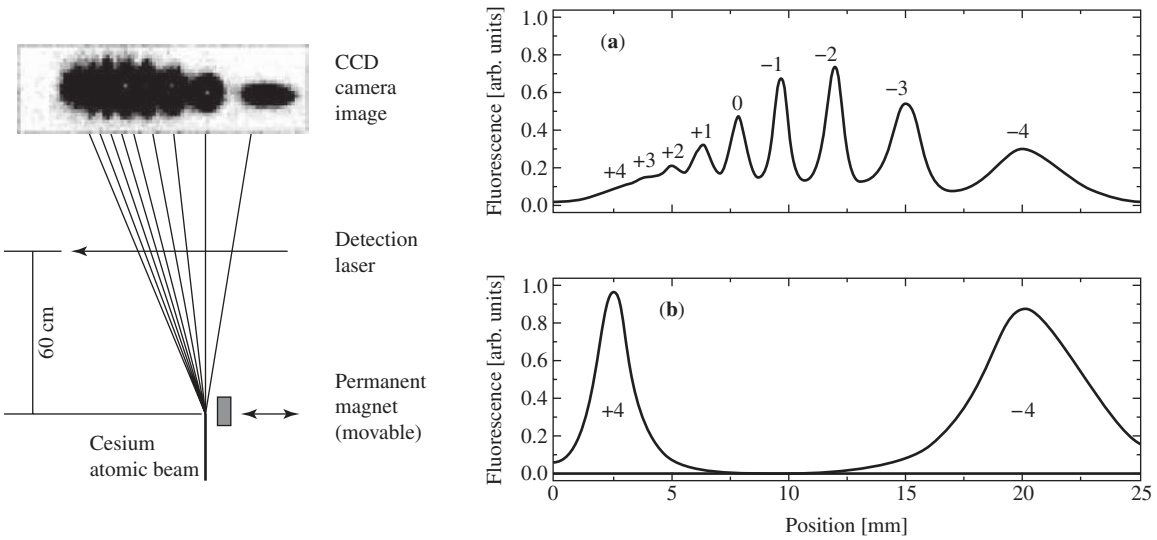

FIGURE 1.6 A modern Stern-Gerlach apparatus, used to separate spin states of atomic cesium, taken from F. Lison et al., *Phys. Rev.* A **61** (1999) 013405. The apparatus is shown on the left, while the data show the nine different projections for the spin-four atom, (a) before and (b) after optical pumping is used to populate only extreme spin projections. The spin quantum number $F = 4$ is a coupling between the outermost electron in the atom and the nuclear spin $I = 7/2$.

Finally, before outlining the mathematical formalism of quantum mechanics, we remark that the physics of a Stern-Gerlach apparatus is of far more than simply academic interest. The ability to separate spin states of atoms has tremendous practical interest as well. Figure 1.6 shows the use of the Stern-Gerlach technique to analyze the result of spin manipulation in an atomic beam of cesium atoms. The only stable isotope, ^{133}Cs, of this alkali atom has a nuclear spin $I = 7/2$, and the experiment sorts out the $F = 4$ hyperfine magnetic substate, giving nine spin orientations. This is only one of many examples where this once mysterious effect is used for practical devices. Of course, all of these uses only go to firmly establish this effect, as well as the quantum-mechanical principles that we will now present and further develop.

1.2 ■ KETS, BRAS, AND OPERATORS

In the preceding section we showed how analyses of the Stern-Gerlach experiment lead us to consider a complex vector space. In this and the following section we formulate the basic mathematics of vector spaces as used in quantum mechanics. Our notation throughout this book is the bra and ket notation developed by P. A. M. Dirac. The theory of linear vector spaces had, of course, been known to mathematicians prior to the birth of quantum mechanics, but Dirac's way of intro-