The eigenket-eigenvalue relation

$$S_z|\pm\rangle = \pm(\hbar/2)|\pm\rangle \tag{1.3.37}$$

immediately follows from the orthonormality property of $|\pm\rangle$.

It is also instructive to look at two other operators,

$$S_+ \equiv \hbar|+\rangle\langle -|, \qquad S_- \equiv \hbar|-\rangle\langle +|, \tag{1.3.38}$$

which are both seen to be *non*-Hermitian. The operator S_+, acting on the spin-down ket $|-\rangle$, turns $|-\rangle$ into the spin-up ket $|+\rangle$ multiplied by $\hbar$. On the other hand, the spin-up ket $|+\rangle$, when acted upon by S_+, becomes a null ket. So the physical interpretation of S_+ is that it raises the spin component by one unit of $\hbar$; if the spin component cannot be raised any further, we automatically get a null state. Likewise, S_- can be interpreted as an operator that lowers the spin component by one unit of $\hbar$. Later we will show that $S_\pm$ can be written as $S_x \pm iS_y$.

In constructing the matrix representations of the angular momentum operators, it is customary to label the column (row) indices in *descending* order of angular momentum components, that is, the first entry corresponds to the maximum angular momentum component, the second, the next highest, and so forth. In our particular case of spin $\frac{1}{2}$ systems, we have

$$|+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{1.3.39a}$$

$$S_z \doteq \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad S_+ \doteq \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad S_- \doteq \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{1.3.39b}$$

We will come back to these explicit expressions when we discuss the Pauli two-component formalism in Chapter 3.

1.4. MEASUREMENTS, OBSERVABLES, AND THE UNCERTAINTY RELATIONS

Measurements

Having developed the mathematics of ket spaces, we are now in a position to discuss the quantum theory of measurement processes. This is not a particularly easy subject for beginners, so we first turn to the words of the great master, P. A. M. Dirac, for guidance (Dirac 1958, 36): "A measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured." What does all this mean? We interpret Dirac's words as follows: Before a measurement of observable A is made, the system is assumed to be represented by some linear combination

$$|\alpha\rangle = \sum_{a'} c_{a'}|a'\rangle = \sum_{a'}|a'\rangle\langle a'|\alpha\rangle. \tag{1.4.1}$$

When the measurement is performed, the system is "thrown into" one of the eigenstates, say $|a'\rangle$ of observable A. In other words,

$$|\alpha\rangle \xrightarrow{A \text{ measurement}} |a'\rangle. \tag{1.4.2}$$

For example, a silver atom with an arbitrary spin orientation will change into either $|S_z;+\rangle$ or $|S_z;-\rangle$ when subjected to a SG apparatus of type SG$\hat{z}$. Thus *a measurement usually changes the state*. The only exception is when the state is already in one of the eigenstates of the observable being measured, in which case

$$|a'\rangle \xrightarrow{A \text{ measurement}} |a'\rangle \tag{1.4.3}$$

with certainty, as will be discussed further. When the measurement causes $|\alpha\rangle$ to change into $|a'\rangle$, it is said that A is measured to be a'. It is in this sense that the result of a measurement yields one of the eigenvalues of the observable being measured.

Given (1.4.1), which is the state ket of a physical system before the measurement, we do not know in advance into which of the various $|a'\rangle$'s the system will be thrown as the result of the measurement. We do postulate, however, that the probability for jumping into some particular $|a'\rangle$ is given by

$$\text{Probability for } a' = |\langle a'|\alpha\rangle|^2, \tag{1.4.4}$$

provided that $|\alpha\rangle$ is normalized.

Although we have been talking about a single physical system, to determine probability (1.4.4) empirically, we must consider a great number of measurements performed on an ensemble—that is, a collection—of identically prepared physical systems, all characterized by the same ket $|\alpha\rangle$. Such an ensemble is known as a **pure ensemble**. (We will say more about ensembles in Chapter 3.) As an example, a beam of silver atoms which survive the first SG$\hat{z}$ apparatus of Figure 1.3 with the S_z- component blocked is an example of a pure ensemble because every member atom of the ensemble is characterized by $|S_z;+\rangle$.

The probabilistic interpretation (1.4.4) for the squared inner product $|\langle a'|\alpha\rangle|^2$ is one of the fundamental postulates of quantum mechanics, so it cannot be proven. Let us note, however, that it makes good sense in extreme cases. Suppose the state ket is $|a'\rangle$ itself even before a measurement is made; then according to (1.4.4), the probability for getting a'—or, more precisely, for being thrown into $|a'\rangle$—as the result of the measurement is predicted to be 1, which is just what we expect. By measuring A once again, we, of course, get $|a'\rangle$ only; quite generally, repeated measurements of the same observable in succession yield the same result.* If, on the other hand, we are interested in the probability for the system initially characterized by $|a'\rangle$ to be thrown into some other eigenket $|a''\rangle$ with $a'' \neq a'$, then (1.4.4) gives zero because of the orthogonality between $|a'\rangle$ and $|a''\rangle$. From the point of view of measurement theory, orthogonal kets correspond to mutually exclusive alternatives; for example, if a spin $\frac{1}{2}$ system is in $|S_z; +\rangle$, it is not in $|S_z; -\rangle$ with certainty.

Quite generally, the probability for anything must be nonnegative. Furthermore, the probabilities for the various alternative possibilities must add up to unity. Both of these expectations are met by our probability postulate (1.4.4).

We define the **expectation value** of A taken with respect to state $|\alpha\rangle$ as

$$\langle A \rangle \equiv \langle \alpha | A | \alpha \rangle. \tag{1.4.5}$$

To make sure that we are referring to state $|\alpha\rangle$, the notation $\langle A \rangle_\alpha$ is sometimes used. Equation (1.4.5) is a definition; however, it agrees with our intuitive notion of *average measured value* because it can be written as

$$\begin{aligned} \langle A \rangle &= \sum_{a'} \sum_{a''} \langle \alpha | a'' \rangle \langle a'' | A | a' \rangle \langle a' | \alpha \rangle \\ &= \sum_{a'} \underset{\substack{\uparrow \\ \text{measured value } a'}}{a'} \qquad \underbrace{|\langle a' | \alpha \rangle|^2}_{\text{probability for obtaining } a'} \end{aligned} \tag{1.4.6}$$

It is very important not to confuse eigenvalues with expectation values. For example, the expectation value of S_z for spin $\frac{1}{2}$ systems can assume *any* real value between $-\hbar/2$ and $+\hbar/2$, say $0.273\hbar$; in contrast, the eigenvalue of S_z assumes only two values, $\hbar/2$ and $-\hbar/2$.

To clarify further the meaning of measurements in quantum mechanics we introduce the notion of a **selective measurement**, or *filtration*. In Section 1.1 we considered a Stern-Gerlach arrangement where we let only one of the spin components pass out of the apparatus while we completely blocked the other component. More generally, we imagine a measurement process with a device that selects only one of the eigenkets of A, say $|a'\rangle$, and rejects all others; see Figure 1.6. This is what we mean by a selective measurement; it is also called filtration because only one of the A eigenkets filters through the ordeal. Mathematically we can say that such a selective

*Here successive measurements must be carried out immediately afterward. This point will become clear when we discuss the time evolution of a state ket in Chapter 2.

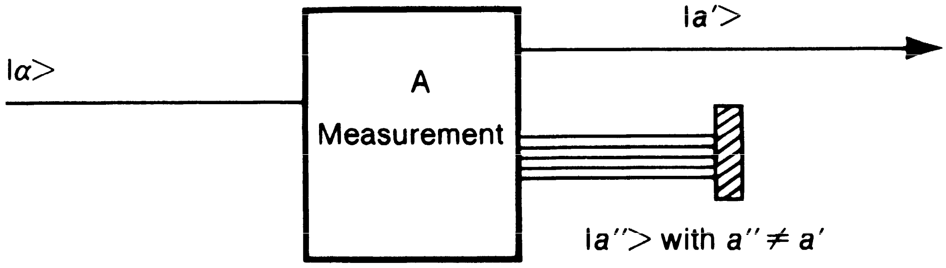

FIGURE 1.6. Selective measurement.

measurement amounts to applying the projection operator $\Lambda_{a'}$ to $|\alpha\rangle$:

$$\Lambda_{a'}|\alpha\rangle = |a'\rangle\langle a'|\alpha\rangle. \tag{1.4.7}$$

J. Schwinger has developed a formalism of quantum mechanics based on a thorough examination of selective measurements. He introduces a measurement symbol $M(a')$ in the beginning, which is identical to $\Lambda_{a'}$ or $|a'\rangle\langle a'|$ in our notation, and deduces a number of properties of $M(a')$ (and also of $M(b', a')$ which amount to $|b'\rangle\langle a'|$) by studying the outcome of various Stern-Gerlach–type experiments. In this way he motivates the entire mathematics of kets, bras, and operators. In this book we do not follow Schwinger's path; the interested reader may consult Gottfried's book. (Gottfried 1966, 192–9).

Spin $\frac{1}{2}$ Systems, Once Again

Before proceeding with a general discussion of observables, we once again consider spin $\frac{1}{2}$ systems. This time we show that the results of sequential Stern-Gerlach experiments, when combined with the postulates of quantum mechanics discussed so far, are sufficient to determine not only the $S_{x,y}$ eigenkets, $|S_x; \pm\rangle$ and $|S_y; \pm\rangle$, but also the operators S_x and S_y themselves.

First, we recall that when the S_x+ beam is subjected to an apparatus of type SG$\hat{z}$, the beam splits into two components with equal intensities. This means that the probability for the S_x+ state to be thrown into $|S_z; \pm\rangle$, simply denoted as $|\pm\rangle$, is $\frac{1}{2}$ each; hence,

$$|\langle + |S_x; +\rangle| = |\langle - |S_x; +\rangle| = \frac{1}{\sqrt{2}}. \tag{1.4.8}$$

We can therefore construct the S_x+ ket as follows:

$$|S_x; +\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}e^{i\delta_1}|-\rangle, \tag{1.4.9}$$

with δ_1 real. In writing (1.4.9) we have used the fact that the *overall* phase (common to both $|+\rangle$ and $|-\rangle$) of a state ket is immaterial; the coefficient just $|\langle c'|a'\rangle|^2$, which can also be written as follows:

$$|\langle c'|a'\rangle|^2 = |\sum_{b'} \langle c'|b'\rangle\langle b'|a'\rangle|^2 = \sum_{b'}\sum_{b''} \langle c'|b'\rangle\langle b'|a'\rangle\langle a'|b''\rangle\langle b''|c'\rangle. \tag{1.4.47}$$

Notice that expressions (1.4.46) and (1.4.47) are different! This is remarkable because in both cases the pure $|a'\rangle$ beam coming out of the first (A) filter can be regarded as being made up of the B eigenkets

$$|a'\rangle = \sum_{b'} |b'\rangle\langle b'|a'\rangle, \tag{1.4.48}$$

where the sum is over all possible values of b'. The crucial point to be noted is that the result coming out of the C filter depends on whether or not B measurements have actually been carried out. In the first case we experimentally ascertain which of the B eigenvalues are actually realized; in the second case, we merely imagine $|a'\rangle$ to be built up of the various $|b'\rangle$'s in the sense of (1.4.48). Put in another way, actually recording the probabilities of going through the various b' routes makes all the difference even though we sum over b' afterwards. Here lies the heart of quantum mechanics.

Under what conditions do the two expressions become equal? It is left as an exercise for the reader to show that for this to happen, in the absence of degeneracy, it is sufficient that

$$[A, B] = 0 \quad \text{or} \quad [B, C] = 0. \tag{1.4.49}$$

In other words, the peculiarity we have illustrated is characteristic of incompatible observables.

The Uncertainty Relation

The last topic to be discussed in this section is the uncertainty relation. Given an observable A, we define an **operator**

$$\Delta A \equiv A - \langle A\rangle, \tag{1.4.50}$$

where the expectation value is to be taken for a certain physical state under consideration. The expectation value of $(\Delta A)^2$ is known as the **dispersion** of A. Because we have

$$\langle(\Delta A)^2\rangle = \langle(A^2 - 2A\langle A\rangle + \langle A\rangle^2)\rangle = \langle A^2\rangle - \langle A\rangle^2, \tag{1.4.51}$$

the last line of (1.4.51) may be taken as an alternative definition of dispersion. Sometimes the terms **variance** and **mean square deviation** are used for the same quantity. Clearly, the dispersion vanishes when the state in question is an eigenstate of A. Roughly speaking, the dispersion of an observable characterizes "fuzziness." For example, for the S_z+ state of a spin $\frac{1}{2}$ system, the dispersion of S_x can be computed to be

$$\langle S_x^2\rangle - \langle S_x\rangle^2 = \hbar^2/4. \tag{1.4.52}$$

In contrast the dispersion $\langle \Delta S_z)^2\rangle$ obviously vanishes for the S_z+ state. So, for the S_z+ state, S_z is "sharp"—a vanishing dispersion for S_z—while S_x is fuzzy.

We now state the uncertainty relation, which is the generalization of the well-known x-p uncertainty relation to be discussed in Section 1.6. Let A and B be observables. Then for any state we must have the following inequality:

$$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle \geq \frac{1}{4}|\langle[A, B]\rangle|^2. \tag{1.4.53}$$

To prove this we first state three lemmas.

Lemma 1. *The Schwarz inequality*

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\alpha|\beta\rangle|^2, \tag{1.4.54}$$

which is analogous to

$$|\mathbf{a}|^2|\mathbf{b}|^2 \geq |\mathbf{a}\cdot\mathbf{b}|^2 \tag{1.4.55}$$

in real Euclidian space.

Proof. First note

$$(\langle\alpha| + \lambda^*\langle\beta|)\cdot(|\alpha\rangle + \lambda|\beta\rangle) \geq 0, \tag{1.4.56}$$

where λ can be any complex number. This inequality must hold when λ is set equal to $-\langle\beta|\alpha\rangle/\langle\beta|\beta\rangle$:

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle - |\langle\alpha|\beta\rangle|^2 \geq 0, \tag{1.4.57}$$

which is the same as (1.4.54). □

Lemma 2. *The expectation value of a Hermitian operator is purely real.*

Proof. The proof is trivial—just use (1.3.21). □

Lemma 3. *The expectation value of an anti-Hermitian operator, defined by $C = -C^\dagger$, is purely imaginary.*

Proof. The proof is also trivial. □

Armed with these lemmas, we are in a position to prove the uncertainty relation (1.4.53). Using Lemma 1 with

$$\begin{aligned} |\alpha\rangle &= \Delta A|\ \rangle, \\ |\beta\rangle &= \Delta B|\ \rangle, \end{aligned} \tag{1.4.58}$$

where the blank ket $|\ \rangle$ emphasizes the fact that our consideration may be applied to *any* ket, we obtain

$$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle \geq |\langle\Delta A \Delta B\rangle|^2, \tag{1.4.59}$$

where the Hermiticity of ΔA and ΔB has been used. To evaluate the right-hand side of (1.4.59), we note

$$\Delta A \Delta B = \frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\}, \tag{1.4.60}$$

where the commutator $[\Delta A, \Delta B]$, which is equal to $[A, B]$, is clearly anti-Hermitian

$$([A, B])^\dagger = (AB - BA)^\dagger = BA - AB = -[A, B]. \tag{1.4.61}$$

In contrast, the anticommutator $\{\Delta A, \Delta B\}$ is obviously Hermitian, so

$$\langle\Delta A \Delta B\rangle = \frac{1}{2}\underbrace{\langle[A, B]\rangle}_{\substack{\text{purely}\\ \text{imaginary}}} + \frac{1}{2}\underbrace{\langle\{\Delta A, \Delta B\}\rangle}_{\substack{\text{purely}\\ \text{real}}}, \tag{1.4.62}$$

where Lemmas 2 and 3 have been used. The right-hand side of (1.4.59) now becomes

$$|\langle\Delta A \Delta B\rangle|^2 = \frac{1}{4}|\langle[A, B]\rangle|^2 + \frac{1}{4}|\langle\{\Delta A, \Delta B\}\rangle|^2. \tag{1.4.63}$$

The proof of (1.4.53) is now complete because the omission of the second (the anticommutator) term of (1.4.63) can only make the inequality relation stronger.*

Applications of the uncertainty relation to spin $\frac{1}{2}$ systems will be left as exercises. We come back to this topic when we discuss the fundamental x-p commutation relation in Section 1.6.

1.5. CHANGE OF BASIS

Transformation Operator

Suppose we have two incompatible observables A and B. The ket space in question can be viewed as being spanned either by the set $\{|a'\rangle\}$ or by the set $\{|b'\rangle\}$. For example, for spin $\frac{1}{2}$ systems $|S_z \pm\rangle$ may be used as our base kets; alternatively, $|S_x \pm\rangle$ may be used as our base kets. The two different sets of base kets, of course, span the same ket space. We are interested in finding out how the two descriptions are related. Changing the

*In the literature most authors use ΔA for our $\sqrt{\langle(\Delta A)^2\rangle}$ so the uncertainty relation is written as $\Delta A \Delta B \geq \frac{1}{2}|\langle[A, B]\rangle|$. In this book, however, ΔA and ΔB are to be understood as operators [see (1.4.50)], not numbers.