5

Spin One

5–1 Filtering atoms with a Stern-Gerlach apparatus

Review: Chapter 35, Vol. II, *Paramagnetism and Magnetic Resonance*. For your convenience this chapter is reproduced in the Appendix of this volume.

In this chapter we really begin the quantum mechanics proper—in the sense that we are going to describe a quantum mechanical phenomenon in a completely quantum mechanical way. We will make no apologies and no attempt to find connections to classical mechanics. We want to talk about something new in a new language. The particular situation which we are going to describe is the behavior of the so-called quantization of the angular momentum, for a particle of *spin one*. But we won't use words like "angular momentum" or other concepts of classical mechanics until later. We have chosen this particular example because it is relatively simple, although not the simplest possible example. It is sufficiently complicated that it can stand as a prototype which can be generalized for the description of all quantum mechanical phenomena. Thus, although we are dealing with a particular example, all the laws which we mention are immediately generalizable, and we will give the generalizations so that you will see the general characteristics of a quantum mechanical description. We begin with the phenomenon of the splitting of a beam of atoms into three separate beams in a Stern-Gerlach experiment.

You remember that if we have an inhomogeneous magnetic field made by a magnet with a pointed pole tip and we send a beam through the apparatus, the beam of particles may be split into a number of beams—the number depending on the particular kind of atom and its state. We are going to take the case of an atom which gives three beams, and we are going to call that a particle of *spin one*. You can do for yourself the case of five beams, seven beams, two beams, etc.—you just copy everything down and where we have three terms, you will have five terms, seven terms, and so on.

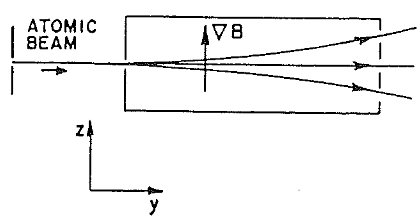


Fig. 5–1. In a Stern-Gerlach experiment, atoms of spin one are split into three beams.

Imagine the apparatus drawn schematically in Fig. 5–1. A beam of atoms (or particles of any kind) is collimated by some slits and passes through a nonuniform field. Let's say that the beam moves in the y-direction and that the magnetic field and its gradient are both in the z-direction. Then, looking from the side, we will see the beam split vertically into three beams, as shown in the figure. Now at the output end of the magnet we could put small counters which count the rate of arrival of particles in any one of the three beams. Or we can block off two of the beams and let the third one go on.

Suppose we block off the lower two beams and let the top-most beam go on and enter a second Stern-Gerlach apparatus of the same kind, as shown in Fig. 5–2. What happens? There are *not* three beams in the second apparatus; there is only the top beam.† This is what you would expect if you think of the second apparatus as simply an extension of the first. Those atoms which are being pushed upward continue to be pushed upward in the second magnet.

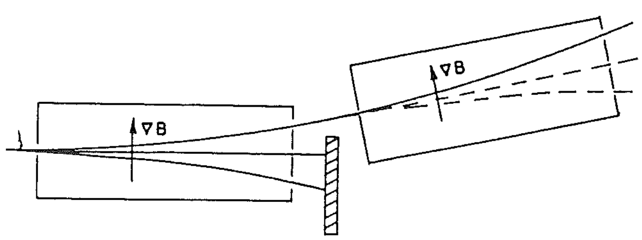


Fig. 5–2. The atoms from one of the beams are sent into a second identical apparatus.

† We are assuming that the deflection angles are very small.

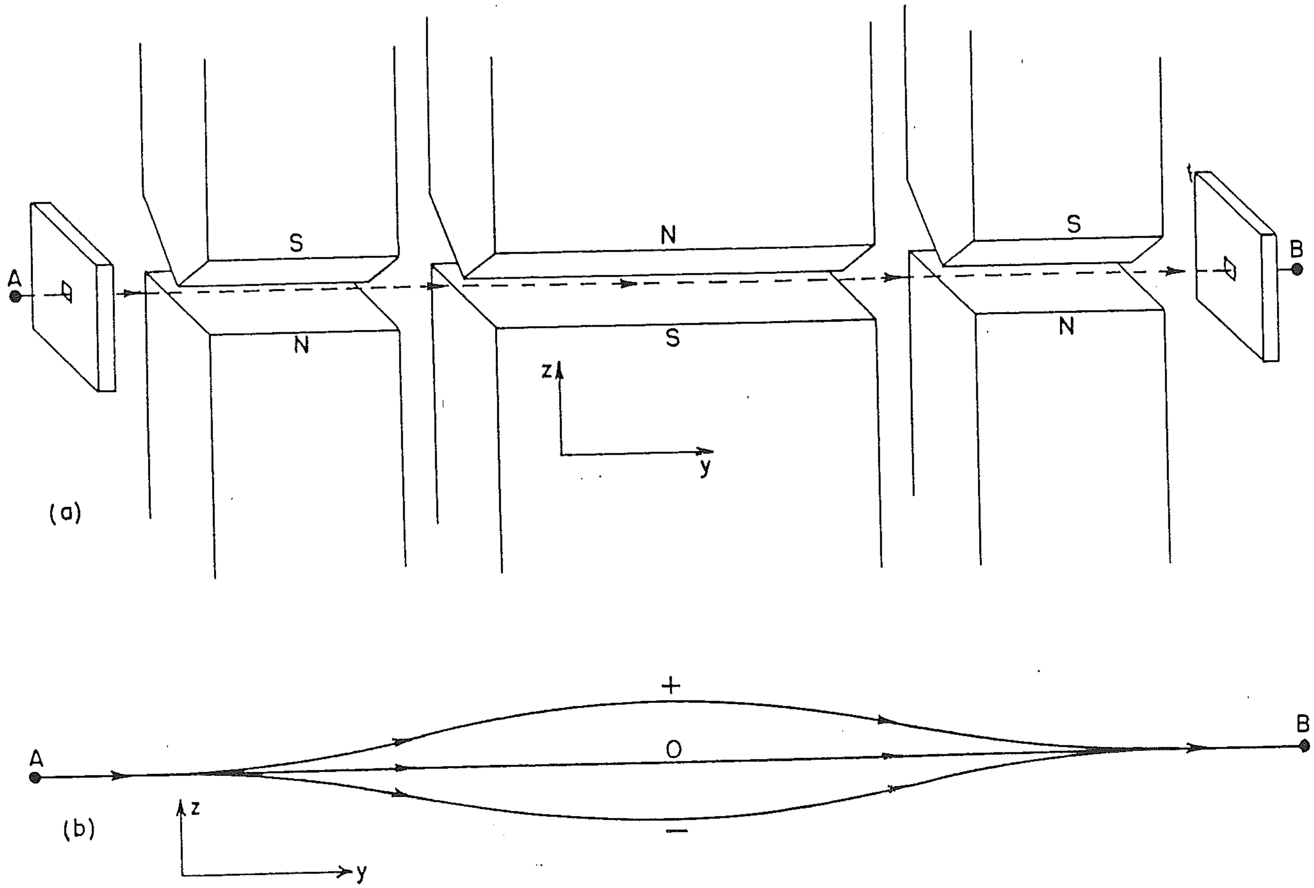


Fig. 5–3. (a) An imagined modification of a Stern-Gerlach apparatus. (b) The paths of spin-one atoms.

You can see then that the first apparatus has produced a beam of "purified" objects—atoms that get bent upward in the particular inhomogeneous field. The atoms, as they enter the original Stern-Gerlach apparatus, are of three "varieties," and the three kinds take different trajectories. By filtering out all but one of the varieties, we can make a beam whose future behavior in the same kind of apparatus is determined and predictable. We will call this a *filtered* beam, or a *polarized* beam, or a beam in which the atoms all are known to be in a *definite state.*

For the rest of our discussion, it will be more convenient if we consider a somewhat modified apparatus of the Stern-Gerlach type. The apparatus looks more complicated at first, but it will make all the arguments simpler. Anyway, since they are only "thought experiments," it doesn't cost anything to complicate the equipment. (Incidentally, no one has ever done all of the experiments we will describe in just this way, but we know what *would* happen from the laws of quantum mechanics, which are, of course, based on other similar experiments. These other experiments are harder to understand at the beginning, so we want to describe some idealized—but possible—experiments.)

Figure 5–3(a) shows a drawing of the "modified Stern-Gerlach apparatus" we would like to use. It consists of a sequence of three high-gradient magnets. The first one (on the left) is just the usual Stern-Gerlach magnet and splits the incoming beam of spin-one particles into three separate beams. The second magnet has the same cross section as the first, but is twice as long *and* the polarity of its magnetic field is opposite the field in magnet 1. The second magnet pushes in the opposite direction on the atomic magnets and bends their paths back toward the axis, as shown in the trajectories drawn in the lower part of the figure. The third magnet is just like the first, and brings the three beams back together again, so that leaves the exit hole along the axis. Finally, we would like to imagine that in front of the hole at *A* there is some mechanism which can get the atoms started from rest and that after the exit hole at *B* there is a decelerating mechanism that brings the atoms back to rest at *B*. That is not essential, but it will mean that in our analysis we won't have to worry about including any effects of the motion as the atoms come out, and can concentrate on those matters having only to do with the spin. The whole purpose of the "improved" apparatus is just to bring all the particles to the same place, and with zero speed.

Now if we want to do an experiment like the one in Fig. 5–2, we can first make a filtered beam by putting a plate in the middle of the apparatus that blocks two of the beams, as shown in Fig. 5–4. If we now put the polarized atoms through a second identical apparatus, all the atoms will take the upper path, as can be verified by putting similar plates in the way of the various beams of the second S filter and seeing whether particles get through.

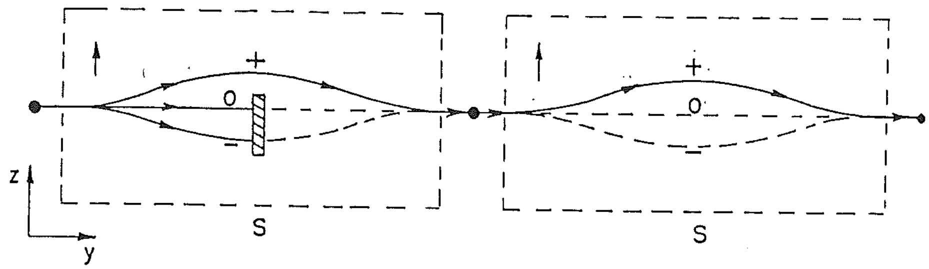


Fig. 5–4. The "improved" Stern-Gerlach apparatus as a filter.

Suppose we call the first apparatus by the name S. (We are going to consider all sorts of combinations, and we will need labels to keep things straight.) We will say that the atoms which take the top path in S are in the "plus state with respect to S"; the ones which take the middle path are in the "zero state with respect to S"; and the ones which take the lowest path are in the "minus state with respect to S." (In the more usual language we would say that the z-component of the angular momentum was $+1\hbar$, 0, and $-1\hbar$, but we are not using that language now.) Now in Fig. 5–4 the second apparatus is oriented just like the first, so the filtered atoms will all go on the upper path. Or if we had blocked off the upper and lower beams in the first apparatus and let only the zero state through, all the filtered atoms would go through the middle path of the second apparatus. And if we had blocked off all but the lowest beam in the first, there would be only a low beam in the second. We can say that in each case our first apparatus has produced a filtered beam in a *pure* state with respect to S ($+$, 0, or $-$), and we can test which state is present by putting the atoms through a second, identical apparatus.

We can make our second apparatus so that it transmits only atoms of a particular state—by putting masks inside it as we did for the first one—and then we can test the state of the incoming beam just by seeing whether anything comes out the far end. For instance, if we block off the two lower paths in the second apparatus, 100 percent of the atoms will still come through; but if we block off the upper path, nothing will get through.

To make this kind of discussion easier, we are going to invent a shorthand symbol to represent one of our improved Stern-Gerlach apparatuses. We will let the symbol

$$\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_{S} \tag{5.1}$$

stand for one complete apparatus. (This is *not* a symbol you will ever find used in quantum mechanics; we've just invented it for this chapter. It is simply meant to be a shorthand picture of the apparatus of Fig. 5–3.) Since we are going to want to use several apparatuses at once, and with various orientations, we will identify each with a letter underneath. So the symbol in (5.1) stands for the apparatus S. When we block off one or more of the beams inside, we will show that by some vertical bars indicating which beam is blocked, like this:

$$\begin{Bmatrix} + & \\ 0 & | \\ - & | \end{Bmatrix}_S . \tag{5.2}$$

The various possible combinations we will be using are shown in Fig. 5–5.

If we have two filters in succession (as in Fig. 5–4), we will put the two symbols next to each other, like this:

$$\begin{Bmatrix} + & \\ 0 & | \\ - & | \end{Bmatrix}_S \begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_S . \tag{5.3}$$

For this setup, everything that comes through the first also gets through the second. In fact, even if we block off the "zero" and "minus" channels of the second apparatus, so that we have

$$\begin{Bmatrix} + & \\ 0 & | \\ - & | \end{Bmatrix}_S \begin{Bmatrix} + & \\ 0 & | \\ - & | \end{Bmatrix}_S , \tag{5.4}$$

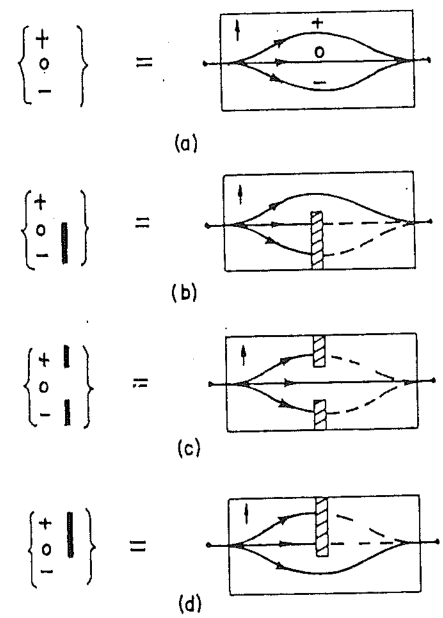


Fig. 5–5. Special shorthand symbols for Stern-Gerlach type filters.

we still get 100 percent transmission through the second apparatus. On the other hand, if we have

$$\begin{Bmatrix} + & \\ 0 & | \\ - & | \end{Bmatrix}_S \begin{Bmatrix} + & | \\ 0 & | \\ - & \end{Bmatrix}_S , \tag{5.5}$$

nothing at all comes out of the far end. Similarly,

$$\begin{Bmatrix} + & | \\ 0 & | \\ - & \end{Bmatrix}_S \begin{Bmatrix} + & \\ 0 & | \\ - & | \end{Bmatrix}_S \tag{5.6}$$

would give nothing out. On the other hand,

$$\begin{Bmatrix} + & | \\ 0 & | \\ - & \end{Bmatrix}_S \begin{Bmatrix} + & | \\ 0 & | \\ - & \end{Bmatrix}_S \tag{5.7}$$

would be just equivalent to

$$\begin{Bmatrix} + & | \\ 0 & | \\ - & \end{Bmatrix}_S$$

by itself.

Now we want to describe these experiments quantum mechanically. We will say that an atom is in the $(+S)$ state if it has gone through the apparatus of Fig. 5–5(b), that it is in a $(0\,S)$ state if it has gone through (c), and in a $(-S)$ state if it has gone through (d).† Then we let $\langle b \mid a \rangle$ be the *amplitude* that an atom which is in state a will get through an apparatus into the b state. We can say: $\langle b \mid a \rangle$ is the amplitude for an atom *in* the state a to *get into* the state b. The experiment (5.4) gives us that

$$\langle +S \mid +S \rangle = 1,$$

† Read: $(+S)$ = "plus-S"; $(0\,S)$ = "zero-S"; $(-S)$ = "minus-S."

whereas (5.5) gives us

$$\langle -S \mid +S \rangle = 0.$$

Similarly, the result of (5.6) is

$$\langle +S \mid -S \rangle = 0,$$

and of (5.7) is

$$\langle -S \mid -S \rangle = 1.$$

As long as we deal only with "pure" states—that is, we have only one channel open—there are nine such amplitudes, and we can write them in a table:

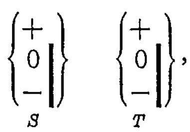
		from		
		$+S$	$0\,S$	$-S$
to	$+S$	1	0	0
	$0\,S$	0	1	0
	$-S$	0	0	1

(5.8)

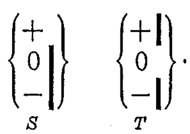
This array of nine numbers—called a *matrix*—summarizes the phenomena we've been describing.

5–2 Experiments with filtered atoms

Now comes the big question: What happens if the second apparatus is tipped to a different angle, so that its field axis is no longer parallel to the first? It could be not only tipped, but also pointed in a different direction—for instance, it could take the beam off at 90° with respect to the original direction. To take it easy at first, let's first think about an arrangement in which the second Stern-Gerlach experiment is tilted by some angle α about the y-axis, as shown in Fig. 5–6. We'll call the second apparatus T. Suppose that we now set up the following experiment:



or the experiment:



What comes out at the far end in these cases?

The answer is this: If the atoms are in a definite state with respect to S, they are *not* in the same state with respect to T—a $(+S)$ state is *not* also a $(+T)$ state. There *is*, however, a certain *amplitude* to find the atom in a $(+T)$ state—or a $(0\,T)$ state or a $(-T)$ state.

In other words, as careful as we have been to make sure that we have the atoms in a definite condition, the fact of the matter is that if it goes through an apparatus which is tilted at a different angle it has, so to speak, to "reorient"

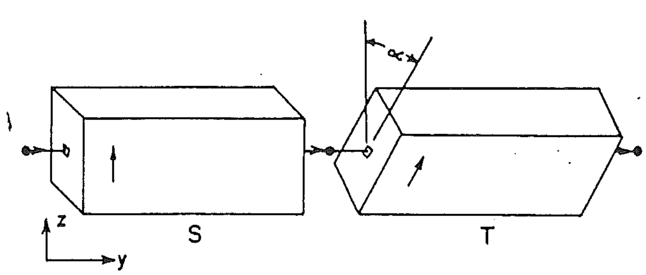


Fig. 5–6. Two Stern-Gerlach type filters in series; the second is tilted at the angle α with respect to the first.

itself—which it does, don't forget, by luck. We can put only one particle through at a time, and then we can only ask the question: What is the probability that it gets through? Some of the atoms that have gone through S will end in a $(+T)$ state, some of them will end in a $(0\,T)$, and some in a $(-T)$ state—all with different odds. These odds can be calculated by the absolute squares of complex amplitudes; what we want is some mathematical method, or quantum mechanical description, for these amplitudes. What we need to know are various quantities like

$$\langle -T \mid +S \rangle,$$

by which we mean the amplitude that an atom initially in the $(+S)$ state can get into the $(-T)$ condition (which is *not* zero unless T and S are lined up parallel to each other). There are other amplitudes like

$$\langle +T \mid 0\,S \rangle, \qquad \text{or} \qquad \langle 0\,T \mid -S \rangle, \qquad \text{etc.}$$

There are, in fact, nine such amplitudes—another matrix—that a theory of particles should tell us how to calculate. Just as $F = ma$ tells us how to calculate what happens to a classical particle in any circumstance, the laws of quantum mechanics permit us to determine the amplitude that a particle will get through a particular apparatus. The central problem, then, is to be able to calculate—for any given tilt angle α, or in fact for any orientation whatever—the nine amplitudes:

$$\begin{array}{lll} \langle +T \mid +S \rangle, & \langle +T \mid 0\,S \rangle, & \langle +T \mid -S \rangle, \\ \langle 0\,T \mid +S \rangle, & \langle 0\,T \mid 0\,S \rangle, & \langle 0\,T \mid -S \rangle, \\ \langle -T \mid +S \rangle, & \langle -T \mid 0\,S \rangle, & \langle -T \mid -S \rangle. \end{array} \tag{5.9}$$

We can already figure out some relations among these amplitudes. First, according to our definitions, the absolute square

$$|\langle +T \mid +S \rangle|^2$$

is the *probability* that an atom in a $(+S)$ state will enter a $(+T)$ state. We will often find it more convenient to write such squares in the equivalent form

$$\langle +T \mid +S \rangle \langle +T \mid +S \rangle^*.$$

In the same notation the number

$$\langle 0\,T \mid +S \rangle \langle 0\,T \mid +S \rangle^*$$

is the probability that a particle in the $(+S)$ state will enter the $(0\,T)$ state, and

$$\langle -T \mid +S \rangle \langle -T \mid +S \rangle^*$$

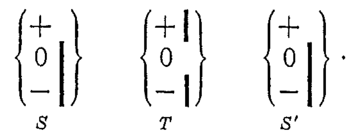
is the probability that it will enter the $(-T)$ state. But the way our apparatuses are made, every atom which enters the T apparatus must be found in *some* one of the three states of the T apparatus—there's nowhere else for a given kind of atom to go. So the sum of the three probabilities we've just written must be equal to 100 percent. We have the relation

$$\langle +T \mid +S \rangle \langle +T \mid +S \rangle^* + \langle 0\,T \mid +S \rangle \langle 0\,T \mid +S \rangle^* + \langle -T \mid +S \rangle \langle -T \mid +S \rangle^* = 1. \tag{5.10}$$

There are, of course, two other such equations that we get if we start with a $(0\,S)$ or a $(-S)$ state. But they are all we can easily get, so we'll go on to some other general questions.

5-3 Stern-Gerlach filters in series

Here is an interesting question: Suppose we had atoms filtered into the $(+S)$ state, then we put them through a second filter, say into a $(0\,T)$ state, and *then* through *another* $+S$ filter. (We'll call the last filter S' just so we can distinguish it from the first S-fiter.) Do the atoms remember that they were once in a $(+S)$ state? In other words, we have the following experiment:

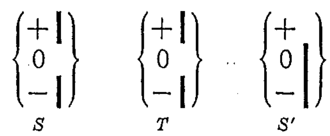


(5.11)

We want to know whether all those that get through T also get through S'. *They do not.* Once they have been filtered by T, they *do not remember* in any way that they were in a $(+S)$ state when they entered T. Note that the second S apparatus in (5.11) is oriented exactly the same as the first, so it is still an S-type filter. The states filtered by S' are, of course, still $(+S)$, $(0\,S)$, and $(-S)$.

The important point is this: *If the T filter passes only one beam*, the *fraction* that gets through the second S filter depends only on the setup of the T filter, and is completely independent of what precedes it. The fact that the same atoms were once sorted by an S filter has no influence whatever on what they will do once they have been sorted again into a pure beam by a T apparatus. From then on, the probability for getting into different states is the same no matter what happened before they got into the T apparatus.

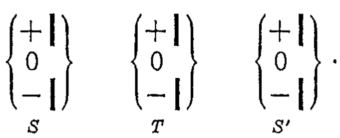
As an example, let's compare the experiment of (5.11) with the following experiment:



(5.12)

in which only the first S is changed. Let's say that the angle α (between S and T) is such that in experiment (5.11) one-third of the atoms that get through T also get through S'. In experiment (5.12), although there will, in general, be a different number of atoms coming through T, the *same fraction of these*—one-third—will also get through S'.

We can, in fact, show from what you have learned earlier that the fraction of the atoms that come out of T and get through any particular S' depends only on T and S', not on anything that happened earlier. Let's compare experiment (5.12) with



(5.13)

The amplitude that an atom that comes out of S will also get through both T and S' is, for the experiments of (5.12),

$$\langle +S \mid 0\,T\rangle\langle 0\,T \mid 0\,S\rangle.$$

The corresponding probability is

$$|\langle +S \mid 0\,T\rangle\langle 0\,T \mid 0\,S\rangle|^2 = |\langle +S \mid 0\,T\rangle|^2\,|\langle 0\,T \mid 0\,S\rangle|^2.$$

The probability for experiment (5.13) is

$$|\langle 0\,S \mid 0\,T\rangle\langle 0\,T \mid 0\,S\rangle|^2 = |\langle 0\,S \mid 0\,T\rangle|^2\,|\langle 0\,T \mid 0\,S\rangle|^2.$$

The ratio is

$$\frac{|\langle 0\,S \mid 0\,T\rangle|^2}{|\langle +S \mid 0\,T\rangle|^2}$$

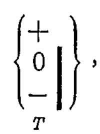
and depends only on T and S', and not at all on which beam $(+S)$, $(0\,S)$, or $(-S)$ is selected by S. (The absolute numbers may go up and down together depending on how much gets through T.) We would, of course, find the same result if we compared the probabilities that the atoms would go into the plus or the minus states with respect to S', or the ratio of the probabilities to go into the zero or minus states.

In fact, since these ratios depend only on which beam is allowed to pass through T, and not on the selection made by the first S filter, it is clear that we would get the same result even if the last apparatus were not an S filter. If we use for the third apparatus—which we will now call R—one rotated by some arbitrary angle with respect to T, we would find that a ratio such as $|\langle 0\,R \mid 0\,T\rangle|^2/|\langle +R \mid 0\,T\rangle|^2$ was independent of which beam was passed by the first filter S.

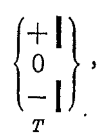
5–4 Base states

These results illustrate one of the basic principles of quantum mechanics: Any atomic system can be separated by a filtering process into a certain set of what we will call *base states*, and the future behavior of the atoms in any single given base state depends only on the nature of the base state—it is independent of any previous history.† The base states depend, of course, on the filter used; for instance, the three states $(+T)$, $(0\,T)$, and $(-T)$ are one set of base states; the three states $(+S)$, $(0\,S)$, and $(-S)$ are another. There are any number of possibilities each as good as any other.

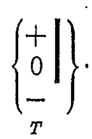
We should be careful to say that we are considering *good* filters which do indeed produce "pure" beams. If, for instance, our Stern-Gerlach apparatus didn't produce a good separation of the three beams so that we could not separate them cleanly by our masks, then we could not make a complete separation into base states. We can tell if we have pure base states by seeing whether or not the beams can be split again in another filter of the same kind. If we have a pure $(+T)$ state, for instance, all the atoms will go through



and none will go through



or through



Our statement about base states means that it is *possible* to filter to some pure state, so that no further filtering by an identical apparatus is possible.

We must also point out that what we are saying is exactly true only in rather idealized situations. In any real Stern-Gerlach apparatus, we would have to worry about diffraction by the slits that could cause some atoms to go into states corresponding to different angles, or about whether the beams might contain atoms with different excitations of their internal states, and so on. We have idealized the situation so that we are talking only about the states that are split in a magnetic field; we are ignoring things having to do with position, momentum, internal excitations, and the like. In general, one would need to consider also base states which are sorted out with respect to such things also. But to keep the concepts simple, we are considering only our set of three states, which is sufficient for the exact treatment of the idealized situation in which the atoms don't get torn up in

† We do not intend the word "base state" to imply anything more than what is said here. They are not to be thought of as "basic" in any sense. We are using the word base with the thought of a *basis* for a description, somewhat in the sense that one speaks of "numbers to the *base* ten."

going through the apparatus, or otherwise badly treated, and come to rest when they leave the apparatus.

You will note that we always begin our thought experiments by taking a filter with only one channel open, so that we start with some definite base state. We do this because atoms come out of a furnace in various states determined at random by the accidental happenings inside the furnace. (It gives what is called an "unpolarized" beam.) This randomness involves probabilities of the "classical" kind—as in coin tossing—which are different from the quantum mechanical probabilities we are worrying about now. Dealing with an unpolarized beam would get us into additional complications that are better to avoid until after we understand the behavior of polarized beams. So don't try to consider at this point what happens if the *first* apparatus lets more than one beam through. (We will tell you how you can handle such cases at the end of the chapter.)

Let's now go back and see what happens when we go from a base state for one filter to a base state for a different filter. Suppose we start again with

$$\underset{S}{\begin{Bmatrix} + & \\ 0 & \vert \\ - & \vert \end{Bmatrix}} \quad \underset{T}{\begin{Bmatrix} + & \vert \\ 0 & \\ - & \vert \end{Bmatrix}}.$$

The atoms which come out of T are in the base state $(0\,T)$ and have no memory that they were once in the state $(+S)$. Some people would say that in the filtering by T we have "lost the information" about the previous state $(+S)$ because we have "disturbed" the atoms when we separated them into three beams in the apparatus T. But that is not true. The past information is not lost by the *separation* into three beams, but by the *blocking masks* that are put in—as we can see by the following set of experiments.

We start with a $+S$ filter and will call N the number of atoms that come through it. If we follow this by a $0\,T$ filter, the number of atoms that come out is some fraction of the original number, say αN. If we then put another $+\ S$ filter, only some fraction β of these atoms will get to the far end. We can indicate this in the following way:

$$\underset{S}{\begin{Bmatrix} + & \\ 0 & \vert \\ - & \vert \end{Bmatrix}} \xrightarrow{N} \underset{T}{\begin{Bmatrix} + & \vert \\ 0 & \\ - & \vert \end{Bmatrix}} \xrightarrow{\alpha N} \underset{S'}{\begin{Bmatrix} + & \\ 0 & \vert \\ - & \vert \end{Bmatrix}} \xrightarrow{\beta\alpha N}. \tag{5.14}$$

If our third apparatus S' selected a different state, say the $(0\,S)$ state, a different fraction, say γ, would get through.† We would have

$$\underset{S}{\begin{Bmatrix} + & \\ 0 & \vert \\ - & \vert \end{Bmatrix}} \xrightarrow{N} \underset{T}{\begin{Bmatrix} + & \vert \\ 0 & \\ - & \vert \end{Bmatrix}} \xrightarrow{\alpha N} \underset{S'}{\begin{Bmatrix} + & \vert \\ 0 & \\ - & \vert \end{Bmatrix}} \xrightarrow{\gamma\alpha N}. \tag{5.15}$$

Now suppose we repeat these two experiments but remove all the masks from T. We would then find the remarkable results as follows:

$$\underset{S}{\begin{Bmatrix} + & \\ 0 & \vert \\ - & \vert \end{Bmatrix}} \xrightarrow{N} \underset{T}{\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}} \xrightarrow{N} \underset{S'}{\begin{Bmatrix} + & \\ 0 & \vert \\ - & \vert \end{Bmatrix}} \xrightarrow{N}, \tag{5.16}$$

$$\underset{S}{\begin{Bmatrix} + & \\ 0 & \vert \\ - & \vert \end{Bmatrix}} \xrightarrow{N} \underset{T}{\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}} \xrightarrow{N} \underset{S'}{\begin{Bmatrix} + & \vert \\ 0 & \\ - & \vert \end{Bmatrix}} \xrightarrow{0}. \tag{5.17}$$

† In terms of our earlier notation $\alpha = |\langle 0\,T|+S\rangle|^2$, $\beta = |\langle +S|0\,T\rangle|^2$, and $\gamma = |\langle 0\,S|0\,T\rangle|^2$.

All the atoms get through S' in the first case, but *none* in the second case! This is one of the great laws of quantum mechanics. That nature works this way is not self-evident, but the results we have given correspond for our idealized situation to the quantum mechanical behavior observed in innumerable experiments.

5-5 Interfering amplitudes

How can it be that in going from (5.15) to (5.17)—by *opening more channels*—we let *fewer* atoms through? This is the old, deep mystery of quantum mechanics—the interference of amplitudes. It's the same kind of thing we first saw in the two-slit interference experiment with electrons. We saw that we could get fewer electrons at some places with both slits open than we got with one slit open. It works quantitatively this way. We can write the amplitude that an atom will get through T and S' in the apparatus of (5.17) as the sum of three amplitudes, one for each of the three beams in T; the sum is equal to zero:

$$\langle 0\,S \mid +T\rangle\langle +T \mid +S\rangle + \langle 0\,S \mid 0\,T\rangle\langle 0\,T \mid +S\rangle + \langle 0\,S \mid -T\rangle\langle -T \mid +S\rangle = 0. \tag{5.18}$$

None of the three individual amplitudes is zero—for example, the absolute square of the second amplitude is $\gamma\alpha$, see (5.15)—but the *sum is zero.* We would have also the same answer if S' were set to select the $(-S)$ state. However, in the setup of (5.16), the answer is different. If we call a the amplitude to get through T and S', in this case we have†

$$a = \langle +S \mid +T\rangle\langle +T \mid +S\rangle + \langle +S \mid 0\,T\rangle\langle 0\,T \mid +S\rangle + \langle +S \mid -T\rangle\langle -T \mid +S\rangle = 1. \tag{5.19}$$

In the experiment (5.16) the beam has been split and recombined. Humpty Dumpty has been put back together again. The information about the original $(+S)$ state is retained—it is just as though the T apparatus were not there at all. This is true whatever is put after the "wide-open" T apparatus. We could follow it with an R filter—a filter at some odd angle—or anything we want. The answer will always be the same as if the atoms were taken directly from the first S filter.

So this is the important principle: A T filter—or any filter—with wide-open masks produces no change at all. We should make one additional condition. The wide-open filter must not only transmit all three beams, but it must also *not* produce unequal disturbances on the three beams. For instance, it should not have a strong electric field near one beam and not the others. The reason is that even if this extra disturbance would still let all the atoms through the filter, it could change the *phases* of some of the amplitudes. Then the interference would be changed, and the amplitudes in Eqs. (5.18) and (5.19) would be different. We will always assume that there are no such extra disturbances.

Let's rewrite Eqs. (5.18) and (5.19) in an improved notation. We will let i stand for any one of the three states $(+T)$, $(0\,T)$, or $(-T)$; then the equations can be written:

$$\sum_{\text{all } i} \langle 0\,S \mid i\rangle\langle i \mid +S\rangle = 0 \tag{5.20}$$

and

$$\sum_{\text{all } i} \langle +S \mid i\rangle\langle i \mid +S\rangle = 1. \tag{5.21}$$

Similarly, for an experiment where S' is replaced by a completely arbitrary filter R, we have

$$\underset{S}{\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}\!\Big|}\quad \underset{T}{\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}}\quad \underset{R}{\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}\!\Big|}\,. \tag{5.22}$$

† We really cannot conclude from the experiment that $a = 1$, but only that $|a|^2 = 1$, so a might be $e^{i\delta}$, but it can be shown that the choice $\delta = 0$ represents no real loss of generality.

The results will always be the same as if the T apparatus were left out and we had only

$$\underset{S}{\left\{\begin{array}{c|}+\\0\\-\end{array}\right\}}\quad\underset{R}{\left\{\begin{array}{c|}+\\0\\-\end{array}\right\}}.$$

Or, expressed mathematically,

$$\sum_{\text{all } i}\langle +R \mid i\rangle\langle i \mid +S\rangle = \langle +R \mid +S\rangle. \tag{5.23}$$

This is our fundamental law, and it is generally true so long as i stands for the three base states of any filter.

You will notice that in the experiment (5.22) there is no special relation of S and R to T. Furthermore, the arguments would be the same no matter what states they selected. To write the equation in a general way, without having to refer to the specific states selected by S and R, let's call ϕ ("phi") the state prepared by the first filter (in our special example, $+S$) and χ ("khi") the state tested by the final filter (in our example, $+R$). Then we can state our fundamental law of Eq. (5.23) in the form

$$\langle \chi \mid \phi\rangle = \sum_{\text{all } i}\langle \chi \mid i\rangle\langle i \mid \phi\rangle, \tag{5.24}$$

where i is to range over the three base states of some particular filter.

We want to emphasize again what we mean by base states. They are like the three states which can be selected by one of our Stern-Gerlach apparatuses. One condition is that if you have a base state, then the future is independent of the past. Another condition is that if you have a complete set of base states, Eq. (5.24) is true for any set of beginning and ending states ϕ and χ. There is, however, *no unique* set of base states. We began by considering base states *with respect to* a particular apparatus T. We could equally well consider a *different set* of base states with respect to an apparatus S, or with respect to R, etc.† We usually speak of the base states "in a certain representation."

Another condition on a set of base states in any particular representation is that they are all completely different. By that we mean that if we have a $(+T)$ state, there is no amplitude for it to go into a $(0\,T)$ or a $(-T)$ state. If we let i and j stand for any two base states of a particular set, the general rules discussed in connection with (5.8) are that

$$\langle j \mid i\rangle = 0$$

for all i and j that are not equal. Of course, we know that

$$\langle i \mid i\rangle = 1.$$

These two equations are usually written as

$$\langle j \mid i\rangle = \delta_{ji}, \tag{5.25}$$

where δ_{ji} (the "Kronecker delta") is a symbol that is defined to be zero for $i \neq j$, and to be one for $i = j$.

Equation (5.25) is not independent of the other laws we have mentioned. It happens that we are not particularly interested in the mathematical problem of finding the minimum set of independent axioms that will give all the laws as consequences.‡ We are satisfied if we have a set that is complete and not apparently inconsistent. We can, however, show that Eqs. (5.25) and (5.24) are not independent. Suppose we let ϕ in Eq. (5.24) represent one of the base states of the

† In fact, for atomic systems with three or more base states, there exist other kinds of filters—quite different from a Stern-Gerlach apparatus—which can be used to get more choices for the set of base states (each set with the same *number* of states).

‡ Redundant *truth* doesn't bother us!

same set as i, say the jth state; then we have

$$\langle \chi \mid j \rangle = \sum_i \langle \chi \mid i \rangle \langle i \mid j \rangle.$$

But Eq. (5.25) says that $\langle i \mid j \rangle$ is zero unless $i = j$, so the sum becomes just $\langle \chi \mid j \rangle$ and we have an identity, which shows that the two laws are not independent.

We can see that there must be another relation among the amplitudes if both Eqs. (5.10) and (5.24) are true. Equation (5.10) is

$$\langle +T \mid +S \rangle \langle +T \mid +S \rangle^* + \langle 0\,T \mid +S \rangle \langle 0\,T \mid +S \rangle^* + \langle -T \mid +S \rangle \langle -T \mid +S \rangle^* = 1.$$

If we write Eq. (5.24), letting both ϕ and χ be the state $(+S)$, the left-hand side is $\langle +S \mid +S \rangle$, which is clearly $= 1$; so we get once more Eq. (5.19),

$$\langle +S \mid +T \rangle \langle +T \mid +S \rangle + \langle +S \mid 0\,T \rangle \langle 0\,T \mid +S \rangle + \langle +S \mid -T \rangle \langle -T \mid +S \rangle = 1.$$

These two equations are consistent (for all relative orientations of the T and S apparatuses) only if

$$\begin{aligned} \langle +S \mid +T \rangle &= \langle +T \mid +S \rangle^*, \\ \langle +S \mid 0\,T \rangle &= \langle 0\,T \mid +S \rangle^*, \\ \langle +S \mid -T \rangle &= \langle -T \mid +S \rangle^*. \end{aligned}$$

And it follows that for any states ϕ and χ,

$$\langle \phi \mid \chi \rangle = \langle \chi \mid \phi \rangle^*. \tag{5.26}$$

If this were not true, probability wouldn't be "conserved," and particles would get "lost."

Before going on, we want to summarize the three important general laws about amplitudes. They are Eqs. (5.24), (5.25), and (5.26):

$$\begin{aligned} &\text{I} \quad \langle j \mid i \rangle = \delta_{ji}, \\ &\text{II} \quad \langle \chi \mid \phi \rangle = \sum_{\text{all } i} \langle \chi \mid i \rangle \langle i \mid \phi \rangle, \\ &\text{III} \quad \langle \phi \mid \chi \rangle = \langle \chi \mid \phi \rangle^*. \end{aligned} \tag{5.27}$$

In these equations the i and j refer to *all* the base states of some *one* representation, while ϕ and χ represent any possible states of the atom. It is important to note that II is valid only if the sum is carried out over *all* the base states of the system (in our case, three: $+T$, $0\,T$, $-T$). These laws say nothing about what we should choose for a base for our set of base states. We began by using a T apparatus, which is a Stern-Gerlach experiment with some arbitrary orientation; but any other orientation, say W, would be just as good. We would have a different set of states to use for i and j, but all the laws would still be good—there is no unique set. One of the great games of quantum mechanics is to make use of the fact that things can be calculated in more than one way.

5–6 The machinery of quantum mechanics

We want to show you why these laws are useful. Suppose we have an atom in a given condition (by which we mean that it was prepared in a certain way), and we want to know what will happen to it in some experiment. In other words, we start with our atom in the state ϕ and want to know what are the *odds* that it will go through some apparatus which accepts atoms only in the condition χ. The laws say that we can describe the apparatus completely in terms of three complex numbers $\langle \chi \mid i \rangle$, the amplitudes for each base state to be in the condition χ; and that we can tell what will happen if an atom is put into the apparatus if we describe the state of the atom by giving three numbers $\langle i \mid \phi \rangle$, the amplitudes for the atom in its original condition to be found in each of the three base states. This is an important idea.

Let's consider another illustration. Think of the following problem: We start with an S apparatus; then we have a complicated mess of junk, which we can call A, and then an R apparatus—like this:

$$\underset{S}{\begin{Bmatrix} + & \\ 0 & | \\ - & | \end{Bmatrix}} \quad \begin{Bmatrix} \\ A \\ \\ \end{Bmatrix} \quad \underset{R}{\begin{Bmatrix} + & | \\ 0 & \\ - & | \end{Bmatrix}}. \tag{5.28}$$

By A we mean any complicated arrangement of Stern-Gerlach apparatuses with masks or half-masks, oriented at peculiar angles, with odd electric and magnetic fields . . . almost anything you want to put. (It's nice to do thought experiments—you don't have to go to all the trouble of actually *building* the apparatus!) The problem then is: With what amplitude does a particle that enters the section A in a $(+S)$ state come out of A in the $(0\,R)$ state, so that it will get through the last R filter? There is a regular notation for such an amplitude; it is

$$\langle 0\,R \mid A \mid +S\rangle.$$

As usual, it is to be read from right to left (like Hebrew):

$$\langle \text{finish} \mid \text{through} \mid \text{start}\rangle.$$

If by chance A doesn't do anything—but is just an open channel—then we write

$$\langle 0\,R \mid 1 \mid +S\rangle = \langle 0\,R \mid +S\rangle; \tag{5.29}$$

the two symbols are equivalent. For a more general problem, we might replace $(+S)$ by a general starting state ϕ and $(0\,R)$ by a general finishing state χ, and we would want to know the amplitude

$$\langle \chi \mid A \mid \phi\rangle.$$

A complete analysis of the apparatus A would have to give the amplitude $\langle \chi \mid A \mid \phi\rangle$ for every possible pair of states ϕ and χ—an infinite number of combinations! How then can we give a concise description of the behavior of the apparatus A? We can do it in the following way. Imagine that the apparatus of (5.28) is modified to be

$$\underset{S}{\begin{Bmatrix} + & \\ 0 & | \\ - & | \end{Bmatrix}} \quad \underset{T}{\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}} \quad \begin{Bmatrix} \\ A \\ \\ \end{Bmatrix} \quad \underset{T}{\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}} \quad \underset{R}{\begin{Bmatrix} + & | \\ 0 & \\ - & | \end{Bmatrix}}. \tag{5.30}$$

This is really no modification at all since the wide-open T apparatuses don't do anything. But they do suggest how we can analyze the problem. There is a certain set of amplitudes $\langle i \mid +S\rangle$ that the atoms from S will get into the i state of T. Then there is another set of amplitudes that an i state (with respect to T) entering A will come out as a j state (with respect to T). And finally there is an amplitude that each j state will get through the last filter as a $(0\,R)$ state. For each possible alternative path, there is an amplitude of the form

$$\langle 0\,R \mid j\rangle\langle j \mid A \mid i\rangle\langle i \mid +S\rangle,$$

and the total amplitude is the sum of the terms we can get with all possible combinations of i and j. The amplitude we want is

$$\sum_{ij} \langle 0\,R \mid j\rangle\langle j \mid A \mid i\rangle\langle i \mid +S\rangle. \tag{5.31}$$

If $(0\,R)$ and $(+S)$ are replaced by general states χ and ϕ, we would have the same kind of expression; so we have the general result

$$\langle \chi \mid A \mid \phi\rangle = \sum_{ij} \langle \chi \mid j\rangle\langle j \mid A \mid i\rangle\langle i \mid \phi\rangle. \tag{5.32}$$

Now notice that the right-hand side of Eq. (5.32) is really "simpler" than the left-hand side. The apparatus A is completely described by the *nine* numbers $\langle j \mid A \mid i \rangle$ which tell the response of A with respect to the three base states of the apparatus T. Once we know these nine numbers, we can handle any two incoming and outgoing states ϕ and χ if we define each in terms of the three amplitudes for going into, or from, each of the three base states. The result of an experiment is predicted using Eq. (5.32).

This then is the machinery of quantum mechanics for a spin-one particle. Every *state* is described by three numbers which are the amplitudes to be in each of some selected set of base states. Every apparatus is described by nine numbers which are the amplitudes to go from one base state to another in the apparatus. From these numbers anything can be calculated.

The nine amplitudes which describe the apparatus are often written as a square matrix—called the matrix $\langle j \mid A \mid i \rangle$:

$$\begin{array}{cc|ccc|} & & & \text{from} & \\ & & + & 0 & - \\ \text{to} & + & \langle + \mid A \mid + \rangle & \langle + \mid A \mid 0 \rangle & \langle + \mid A \mid - \rangle \\ & 0 & \langle 0 \mid A \mid + \rangle & \langle 0 \mid A \mid 0 \rangle & \langle 0 \mid A \mid - \rangle \\ & - & \langle - \mid A \mid + \rangle & \langle - \mid A \mid 0 \rangle & \langle - \mid A \mid - \rangle \end{array} \qquad (5.33)$$

The mathematics of quantum mechanics is just an extension of this idea. We will give you a simple illustration. Suppose we have an apparatus C that we wish to analyze—that is, we want to calculate the various $\langle j \mid C \mid i \rangle$. For instance, we might want to know what happens in an experiment like

$$\underset{S}{\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}} \quad \begin{Bmatrix} C \end{Bmatrix} \quad \underset{R}{\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}}. \qquad (5.34)$$

But then we notice that C is just built of two pieces of apparatus A and B in series—the particles go through A and then through B—so we can write symbolically

$$\begin{Bmatrix} C \end{Bmatrix} = \begin{Bmatrix} A \end{Bmatrix} \cdot \begin{Bmatrix} B \end{Bmatrix}. \qquad (5.35)$$

We can call the C apparatus the "product" of A and B. Suppose also that we already know how to analyze the two parts; so we can get the matrices (with respect to T) of A and B. Our problem is then solved. We can easily find

$$\langle \chi \mid C \mid \phi \rangle$$

for any input and output states. First we write that

$$\langle \chi \mid C \mid \phi \rangle = \sum_k \langle \chi \mid B \mid k \rangle \langle k \mid A \mid \phi \rangle.$$

Do you see why? (*Hint:* Imagine putting a T apparatus between A and B.) Then if we consider the special case in which ϕ and χ are also base states (of T), say i and j, we have

$$\langle j \mid C \mid i \rangle = \sum_k \langle j \mid B \mid k \rangle \langle k \mid A \mid i \rangle. \qquad (5.36)$$

This equation gives the matrix for the "product" apparatus C in terms of the two matrices of the apparatuses A and B. Mathematicians call the new matrix $\langle j \mid C \mid i \rangle$ —formed from two matrices $\langle j \mid B \mid i \rangle$ and $\langle j \mid A \mid i \rangle$ according to the sum specified in Eq. (5.36)—the "product" matrix BA of the two matrices B and A. (Note that the *order* is important, $AB \neq BA$.) Thus, we can say that the matrix for a succession of two pieces of apparatus is the matrix product of the matrices for the two apparatuses (putting the *first* apparatus on the *right* in the product). Anyone who knows matrix algebra then understands that we mean just Eq. (5.36).

5–7 Transforming to a different base

We want to make one final point about the base states used in the calculations. Suppose we have chosen to work with some particular base—say the S base—and another fellow decides to do the same calculations with a different base—say the T base. To keep things straight let's call our base states the (iS) states, where $i = +, 0, -$. Similarly, we can call his base states (jT). How can we compare our work with his? The final answers for the result of any measurement should come out the same, but in the calculations the various amplitudes and matrices used will be different. How are they related? For instance, if we both start with the same ϕ, we will describe it in terms of the three amplitudes $\langle iS \mid \phi\rangle$ that ϕ goes into our base states in the S representation, whereas he will describe it by the amplitudes $\langle jT \mid \phi\rangle$ that the state ϕ goes into the base states is his T representation. How can we check that we are really both describing the same state ϕ? We can do it with the general rule II in (5.27). Replacing χ by any one of his states jT, we have

$$\langle jT \mid \phi\rangle = \sum_i \langle jT \mid iS\rangle\langle iS \mid \phi\rangle. \tag{5.37}$$

To relate the two representations, we need only give the nine complex numbers of the matrix $\langle jT \mid iS\rangle$. This matrix can then be used to convert all of his equations to our form. It tells us how to *transform* from one set of base states to another. (For this reason $\langle jT \mid iS\rangle$ is sometimes called "the transformation matrix from representation S to representation T." Big words!)

For the case of spin-one particles for which we have only three base states (for higher spins, there are more) the mathematical situation is analogous to what we have seen in vector algebra. Every vector can be represented by giving three numbers—the components along the axes x, y, and z. That is, every vector can be resolved into three "base" vectors which are vectors along the three axes. But suppose someone else chooses to use a different set of axes—x', y', and z'. He will be using different numbers to represent any particular vector. His calculations will look different, but the final results will be the same. We have considered this before and know the rules for transforming vectors from one set of axes to another.

You may want to see how the quantum mechanical transformations work by trying some out; so we will give here, without proof, the transformation matrices for converting the spin-one amplitudes in one representation S to another representation T, for various special relative orientations of the S and T filters. (We will show you in a later chapter how to derive these same results.)

First case: The T apparatus has the same y-axis (along which the particles move) as the S apparatus, but is rotated about the common y-axis by the angle α (as in Fig. 5–6). (To be specific, a set of coordinates x', y', z' is fixed in the T apparatus, related to the x, y, z coordinates of the S apparatus by: $z' = z\cos\alpha + x\sin\alpha$, $x' = x\cos\alpha - z\sin\alpha$, $y' = y$.) Then the transformation amplitudes are:

$$\begin{aligned}
\langle +T \mid +S\rangle &= \tfrac{1}{2}(1 + \cos\alpha),\\
\langle 0T \mid +S\rangle &= -\frac{1}{\sqrt{2}}\sin\alpha,\\
\langle -T \mid +S\rangle &= \tfrac{1}{2}(1 - \cos\alpha),\\
\langle +T \mid 0S\rangle &= +\frac{1}{\sqrt{2}}\sin\alpha,\\
\langle 0T \mid 0S\rangle &= \cos\alpha,\\
\langle -T \mid 0S\rangle &= -\frac{1}{\sqrt{2}}\sin\alpha,\\
\langle +T \mid -S\rangle &= \tfrac{1}{2}(1 - \cos\alpha),\\
\langle 0T \mid -S\rangle &= +\frac{1}{\sqrt{2}}\sin\alpha,\\
\langle -T \mid -S\rangle &= \tfrac{1}{2}(1 + \cos\alpha).
\end{aligned} \tag{5.38}$$

Second Case: The T apparatus has the same z-axis as S, but is rotated around the z-axis by the angle β. (The coordinate transformation is $z' = z$, $x' = x\cos\beta + y\sin\beta$, $y' = y\cos\beta - x\sin\beta$.) Then the transformation amplitudes are:

$$\begin{aligned}\langle +T \mid +S \rangle &= e^{+i\beta},\\ \langle 0\,T \mid 0\,S \rangle &= 1,\\ \langle -T \mid -S \rangle &= e^{-i\beta},\\ \text{all others} &= 0.\end{aligned} \tag{5.39}$$

Note that any rotations of T whatever can be made up of the two rotations described.

If a state ϕ is defined by the three numbers

$$C_+ = \langle +S \mid \phi \rangle, \qquad C_0 = \langle 0\,S \mid \phi \rangle, \qquad C_- = \langle -S \mid \phi \rangle, \tag{5.40}$$

and the same state is described from the point of view of T by the three numbers

$$C'_+ = \langle +T \mid \phi \rangle, \qquad C'_0 = \langle 0\,T \mid \phi \rangle, \qquad C'_- = \langle -T \mid \phi \rangle, \tag{5.41}$$

then the coefficients $\langle jT \mid iS \rangle$ of (5.38) or (5.39) give the transformation connecting C'_i and C'_i. In other words, the C_i are very much like the components of a vector that appear different from the point of view of S and T.

For a spin-one particle *only*—because it requires *three* amplitudes—the correspondence with a vector is very close. In each case, there are three numbers that must transform with coordinate changes in a certain definite way. In fact, there is a set of base states *which transform just like the three components of a vector.* The three combinations

$$C_x = -\frac{1}{\sqrt{2}}(C_+ - C_-), \qquad C_y = -\frac{i}{\sqrt{2}}(C_+ + C_-), \qquad C_z = C_0 \tag{5.42}$$

transform to C'_x, C'_y, and C'_z just the way that x, y, z transform to x', y', z'. [You can check that this is so by using the transformation laws (5.38) and (5.39).] Now you see why a spin-one particle is often called a "vector particle."

5-8 Other situations

We began by pointing out that our discussion of spin-one particles would be a prototype for any quantum mechanical problem. The generalization has only to do with the numbers of states. Instead of only three base states, any particular situation may involve n base states.† Our basic laws in Eq. (5.27) have exactly the same form—with the understanding that i and j must range over all n base states. Any phenomenon can be analyzed by giving the amplitudes that it starts in each one of the base states and ends in any other one of the base states, and then summing over the complete set of base states. Any proper set of base states can be used, and if someone wishes to use a different set, it is just as good; the two can be connected by using an n by n transformation matrix. We will have more to say later about such transformations.

Finally, we promised to remark on what to do if atoms come directly from a furnace, go through some apparatus, say A, and are then analyzed by a filter which selects the state χ. You do not know what the state ϕ is that they start out in. It is perhaps best if you don't worry about this problem just yet, but instead concentrate on problems that always start out with pure states. But if you insist, here is how the problem can be handled.

First, you have to be able to make some reasonable guess about the way the states are distributed in the atoms that come from the furnace. For example, if

† The number of base states n may be, and generally is, infinite.

there were nothing "special" about the furnace, you might reasonably guess that atoms would leave the furnace with random "orientations." Quantum mechanically, that corresponds to saying that you don't know anything about the states, but that one-third are in the $(+S)$ state, one-third are in the $(0\,S)$ state, and one-third are in the $(-S)$ state. For those that are in the $(+S)$ state the amplitude to get through is $\langle x \mid A \mid +S\rangle$ and the probability is $|\langle x \mid A \mid +S\rangle|^2$, and similarly for the others. The overall probability is then

$$\tfrac{1}{3}|\langle x \mid A \mid +S\rangle|^2 + \tfrac{1}{3}|\langle x \mid A \mid 0\,S\rangle|^2 + \tfrac{1}{3}|\langle x \mid A \mid -S\rangle|^2.$$

Why did we use S rather than, say, T? The answer is, surprisingly, the same no matter what we choose for our initial resolution—so long as we are dealing with completely random orientations. It comes about in the same way that

$$\sum_i |\langle x \mid iS\rangle|^2 = \sum_j |\langle x \mid jT\rangle|^2$$

for any x. (We leave it for you to prove.)

Note that it is *not* correct to say that the input state has the amplitudes $\sqrt{1/3}$ to be in $(+S)$, $\sqrt{1/3}$ to be in $(0\,S)$, and $\sqrt{1/3}$ to be in $(-S)$; that would imply that certain interferences might be possible. It is simply that you do not *know* what the initial state is; you have to think in terms of the probability that the system starts out in the various possible initial states, and then you have to take a weighted average over the various possibilities.