6

Spin One-Half†

6–1 Transforming amplitudes

In the last chapter, using a system of spin one as an example, we outlined the general principles of quantum mechanics:

> Any state ψ can be described in terms of a set of base states by giving the amplitudes to be in each of the base states.
>
> The amplitude to go from any state to another can, in general, be written as a sum of products, each product being the amplitude to go into one of the base states times the amplitude to go from that base state to the final condition, with the sum including a term for each base state:
>
> $$\langle\chi\mid\psi\rangle=\sum_i\langle\chi\mid i\rangle\langle i\mid\psi\rangle. \tag{6.1}$$
>
> The base states are orthogonal—the amplitude to be in one if you are in the other is zero:
>
> $$\langle i\mid j\rangle=\delta_{ij}. \tag{6.2}$$
>
> The amplitude to get from one state to another directly is the complex conjugate of the reverse:
>
> $$\langle\chi\mid\psi\rangle^*=\langle\psi\mid\chi\rangle. \tag{6.3}$$

We also discussed a little bit about the fact that there can be more than one base for the states and that we can use Eq. (6.1) to convert from one base to another. Suppose, for example, that we have the amplitudes $\langle iS\mid\psi\rangle$ to find the state ψ in every one of the base states i of a base system S, but that we then decide that we would prefer to describe the state in terms of another set of base states, say the states j belonging to the base T. In the general formula, Eq. (6.1), we could substitute jT for χ and obtain this formula:

$$\langle jT\mid\psi\rangle=\sum_i\langle jT\mid iS\rangle\langle iS\mid\psi\rangle. \tag{6.4}$$

The amplitudes for the state (ψ) to be in the base states (iT) are related to the amplitudes to be in the base states (iS) by the set of coefficients $\langle jT\mid iS\rangle$. If there are N base states, there are N^2 such coefficients. Such a set of coefficients is often called the "*transformation matrix* to go from the *S-representation* to the *T-representation*." This looks rather formidable mathematically, but with a little renaming we can see that it is really not so bad. If we call C_i the amplitude that the state ψ is in the base state iS—that is, $C_i=\langle iS\mid\psi\rangle$—and call C_j' the corresponding amplitudes for the base system T—that is, $C_j'=\langle jT\mid\psi\rangle$, then Eq. (6.4) can be written as

$$C_j'=\sum_i R_{ji}C_i, \tag{6.5}$$

where R_{ji} means the same thing as $\langle jT\mid iS\rangle$. Each amplitude C_j' is equal to a sum

† This chapter is a rather long and abstract side tour, and it does not introduce any idea which we will not also come to by a different route in later chapters. You can, therefore, skip over it, and come back later if you are interested.

over all i of one of the coefficients R_{ji} times each amplitude C_i. It has the same form as the transformation of a vector from one coordinate system to another.

In order to avoid being too abstract for too long, we have given you some examples of these coefficients for the spin-one case, so you can see how to use them in practice. On the other hand, there is a very beautiful thing in quantum mechanics—that from the sheer fact that there are three states and from the symmetry properties of space under rotations, these coefficients can be found purely by abstract reasoning. Showing you such arguments at this early stage has a disadvantage in that you are immersed in another set of abstractions before we get "down to earth." However, the thing is so beautiful that we are going to do it anyway.

We will show you in this chapter how the transformation coefficients can be derived for spin one-half particles. We pick this case, rather than spin one, because it is somewhat easier. Our problem is to determine the coefficients R_{ji} for a particle—an atomic system—which is split into two beams in a Stern-Gerlach apparatus. We are going to derive all the coefficients for the transformation from one representation to another by pure reasoning—plus a few assumptions. *Some* assumptions are always necessary in order to use "pure" reasoning! Although the arguments will be abstract and somewhat involved, the result we get will be relatively simple to state and easy to understand—and the result is the most important thing. You may, if you wish, consider this as a sort of cultural excursion. We have, in fact, arranged that all the essential results derived here are also derived in some other way when they are needed in later chapters. So you need have no fear of losing the thread of our study of quantum mechanics if you omit this chapter entirely, or study it at some later time. The excursion is "cultural" in the sense that it is intended to show that the principles of quantum mechanics are not only interesting, but are so deep that by adding only a few extra hypotheses about the structure of space, we can deduce a great many properties of physical systems. Also, it is important that we know where the different consequences of quantum mechanics come from, because so long as our laws of physics are incomplete—as we know they are—it is interesting to find out whether the places where our theories fail to agree with experiment is where our logic is the best or where our logic is the worst. Until now, it appears that where our logic is the most abstract it always gives correct results—it agrees with experiment. Only when we try to make specific models of the internal machinery of the fundamental particles and their interactions are we unable to find a theory that agrees with experiment. The theory then that we are about to describe agrees with experiment wherever it has been tested—for the strange particles as well as for electrons, protons, and so on.

One remark on an annoying, but interesting, point before we proceed: It is not possible to determine the coefficients R_{ji} uniquely, because there is always some arbitrariness in the probability amplitudes. If you have a set of amplitudes of any kind, say the amplitudes to arrive at some place by a whole lot of different routes, and if you multiply every single amplitude by the same phase factor—say by $e^{i\delta}$—you have another set that is just as good. So, it is always possible to make an arbitrary change in phase of all the amplitudes in any given problem if you want to.

Suppose you calculate some probability by writing a sum of several amplitudes, say $(A + B + C + \cdots)$ and taking the absolute square. Then somebody else calculates the same thing by using the sum of the amplitudes $(A' + B' + C' + \cdots)$ and taking the absolute square. If all the A', B', C', etc., are equal to the A, B, C, etc., except for a factor $e^{i\delta}$, all probabilities obtained by taking the absolute squares will be exactly the same, since $(A' + B' + C' + \cdots)$ is then equal to $e^{i\delta}(A + B + C + \cdots)$. Or suppose, for instance, that we were computing something with Eq. (6.1), but then we suddenly change all of the phases of a certain base system. Every one of the amplitudes $\langle i \mid \psi\rangle$ would be multiplied by the same factor $e^{i\delta}$. Similarly, the amplitudes $\langle i \mid \chi\rangle$ would also be changed by $e^{i\delta}$, but the amplitudes $\langle \chi \mid i\rangle$ are the complex conjugates of the amplitudes $\langle i \mid \chi\rangle$; therefore, the former gets changed by the factor $e^{-i\delta}$. The plus and minus $i\delta$'s in the exponents cancel out, and we would have the same expression we had before. So it is a general rule that if we change all the amplitudes with respect to a given base system by the same phase—or even if we just change *all* the amplitudes in any problem by the same phase—it makes no difference. There is, therefore, some freedom to choose the phases in our transformation matrix. Every now and then we will make such an arbitrary choice—usually following the conventions that are in general use.

6–2 Transforming to a rotated coordinate system

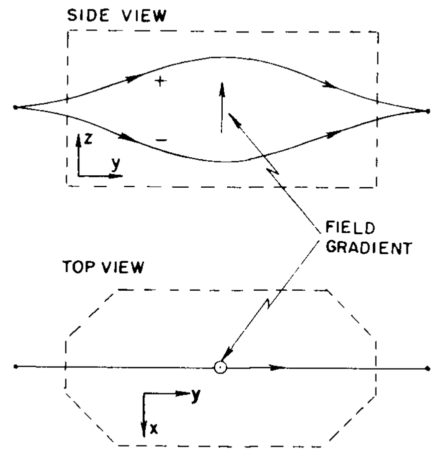

Fig. 6–1. Top and side views of an "improved" Stern-Gerlach apparatus with beams of a spin one-half particle.

We consider again the "improved" Stern-Gerlach apparatus described in the last chapter. A beam of spin one-half particles, entering at the left, would, in general, be split into *two* beams, as shown schematically in Fig. 6–1. (There were three beams for spin *one*.) As before, the beams are put back together again unless one or the other of them is blocked off by a "stop" which intercepts the beam at its half-way point. In the figure we show an arrow which points in the direction of the increase of the *magnitude* of the field—say toward the magnet pole with the sharp edges. This arrow we take to represent *the "up" axis* of any particular apparatus. It is fixed relative to the apparatus and will allow us to indicate the relative orientations when we use several apparatuses together. We also assume that the direction of the magnetic field in each magnet is always the same with respect to the arrow.

We will say that those atoms which go in the "upper" beam are in the $(+)$ state *with respect to that apparatus* and that those in the "lower" beam are in the $(-)$ state. (There is no "zero" state for spin one-half particles.)

Now suppose we put two of our modified Stern-Gerlach apparatuses in sequence, as shown in Fig. 6–2(a). The first one, which we call S, can be used to prepare a pure $(+S)$ or a pure $(-S)$ state by blocking one beam or the other. [As shown it prepares a pure $(+S)$ state.] For each condition, there is some amplitude for a particle that comes out of S to be in either the $(+T)$ or the $(-T)$ beam of the second apparatus. There are, in fact, just four amplitudes: the amplitude to go from $(+S)$ to $(+T)$, from $(+S)$ to $(-T)$, from $(-S)$ to $(+T)$, from $(-S)$ to $(-T)$. These amplitudes are just the four coefficients of the transformation matrix R_{ji} to go from the S-representation to the T-representation. We can consider that the first apparatus "prepares" a particular state in one representation and that the second apparatus "analyzes" that state in terms of the second representation. The kind of question we want to answer, then, is this: If an atom has been prepared in a given condition—say the $(+S)$ state—by blocking one of the beams in the apparatus S, what is the chance that it will get through the second apparatus T if this is set for, say, the $(-T)$ state. The result will depend, of course, on the angles between the two systems S and T.

We should explain why it is that we could have any hope of finding the coefficients R_{ji} by deduction. You know that it is almost impossible to believe that if a particle has its spin lined up in the $+z$-direction, that there is some chance of finding the same particle with its spin pointing in the $+x$-direction—or in any other direction at all. In fact, it *is* almost impossible, but not quite. It is so nearly impossible that there is *only one way* it can be done, and that is the reason we can find out what that unique way is.

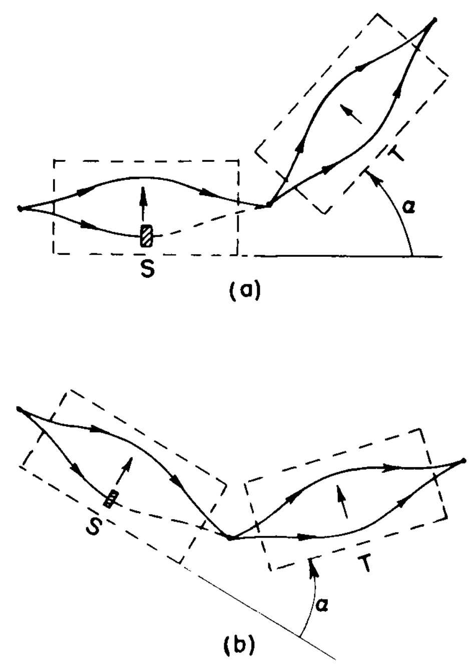

Fig. 6–2. Two equivalent experiments.

The first kind of argument we can make is this. Suppose we have a setup like the one in Fig. 6–2(a), in which we have the two apparatuses S and T, with T cocked at the angle α with respect to S, and we let only the $(+)$ beam through S and the $(-)$ beam through T. We would observe a certain number for the probability that the particles coming out of S get through T. Now suppose we make another measurement with the apparatus of Fig. 6–2(b). The *relative* orientation of S and T is the same, but the whole system sits at a different angle in space. We want to *assume* that both of these experiments give the same number for the chance that a particle in a pure state with respect to S will get into some particular state with respect to T. We are assuming, in other words, that the result of any experiment of this type is the same—that the *physics* is the same—no matter

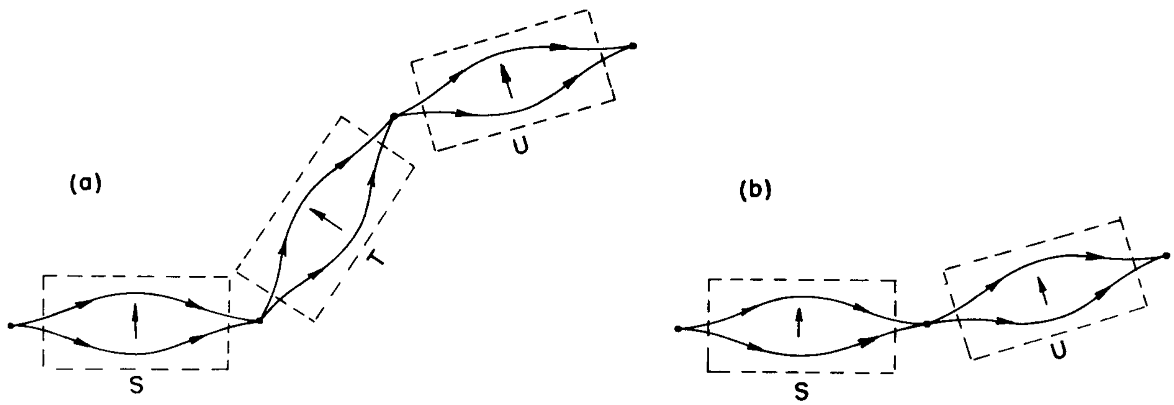

Fig. 6–3. If T is "wide open," (b) is equivalent to (a).

how the *whole* apparatus is oriented in space. (You say, "That's obvious." But it *is* an assumption, and it is "right" only if it is actually what happens.) That means that the coefficients R_{ji} depend only on the relation in space of S and T, and not on the absolute situation of S and T. To say this in another way, R_{ji} depends only on the *rotation* which carries S to T, for evidently what is the same in Fig. 6–2(a) and Fig. 6–2(b) is the three-dimensional rotation which would carry apparatus S into the orientation of apparatus T. When the transformation matrix R_{ji} depends only on a rotation, as it does here, it is called a *rotation matrix.*

For our next step we will need one more piece of information. Suppose we add a third apparatus which we can call U, which follows T at some arbitrary angle, as in Fig. 6–3(a). (It's beginning to look horrible, but that's the fun of abstract thinking—you can make the most weird experiments just by drawing lines!) Now what is the $S \to T \to U$ transformation? What we really want to ask for is the amplitude to go from some state with respect to S to some other state with respect to U, when we know the transformation from S to T and from T to U. We are then asking about an experiment in which both channels of T are open. We can get the answer by applying Eq. (6.5) twice in succession. For going from the S-representation to the T-representation, we have

$$C'_j = \sum_i R_{ji}^{TS} C_i, \tag{6.6}$$

where we put the superscripts TS on the R, so that we can distinguish it from the coefficients R^{UT} we will have for going from T to U.

Assuming the amplitudes to be in the base states of the U-representation C''_k, we can relate them to the T-amplitudes by using Eq. (6.5) once more; we get

$$C''_k = \sum_j R_{kj}^{UT} C'_j. \tag{6.7}$$

Now we can combine Eqs. (6.6) and (6.7) to get the transformation to U directly from S. Substituting C'_j from Eq. (6.6) in Eq. (6.7), we have

$$C''_k = \sum_j R_{kj}^{UT} \sum_i R_{ji}^{TS} C_i. \tag{6.8}$$

Or, since i does not appear in R_{kj}^{UT}, we can put the i-summation also in front, and write

$$C''_k = \sum_i \sum_j R_{kj}^{UT} R_{ji}^{TS} C_i. \tag{6.9}$$

This is the formula for a double transformation.

Notice, however, that so long as all the beams in T are unblocked, the state coming out of T is the same as the one that went in. We could just as well have made a transformation from the S-representation directly to the U-representation. It should be the same as putting the U apparatus right after S, as in Fig.

6–3(b). In that case, we would have written

$$C_k'' = \sum_i R_{ki}^{US} C_i, \tag{6.10}$$

with the coefficients R_{ki}^{US} belonging to this transformation. Now, clearly, Eqs. (6.9) and (6.10) should give the same amplitudes C_k'', and this should be true no matter what the original state ϕ was which gave us the amplitudes C_i. So it must be that

$$R_{ki}^{US} = \sum_j R_{kj}^{UT} R_{ji}^{TS}. \tag{6.11}$$

In other words, for any rotation $S \rightarrow U$ of a reference base, which is viewed as a compounding of two successive rotations $S \rightarrow T$ and $T \rightarrow U$, the rotation matrix R_{ki}^{US} can be obtained from the matrices of the two partial rotations by Eq. (6.11). If you wish, you can find Eq. (6.11) directly from Eq. (6.1), for it is only a different notation for $\langle kU \mid iS\rangle = \sum_j \langle kU \mid jT\rangle\langle jT \mid iS\rangle$.

To be thorough, we should add the following parenthetical remarks. They are not terribly important, however, so you can skip to the next section if you want. What we have said is not quite right. We cannot really say that Eq. (6.9) and Eq. (6.10) must give *exactly* the same amplitudes. Only the *physics* should be the same; all the amplitudes could be different by some common phase factor like $e^{i\delta}$ without changing the result of any calculation about the real world. So, instead of Eq. (6.11), all we can say, really, is that

$$e^{i\delta} R_{ki}^{US} = \sum_j R_{kj}^{UT} R_{ji}^{TS}, \tag{6.12}$$

where δ is *some* real constant. What this extra factor of $e^{i\delta}$ means, of course, is that the amplitudes we get if we use the matrix R^{US} might all differ by the same phase $(e^{-i\delta})$ from the amplitude we would get using the two rotations R^{UT} and R^{TS}. We know that it doesn't matter if all amplitudes are changed by the same phase, so we could just ignore this phase factor if we wanted to. It turns out, however, that if we define all of our rotation matrices in a particular way, this extra phase factor will never appear—the δ in Eq. (6.12) will always be zero. Although it is not important for the rest of our arguments, we can give a quick proof by using a mathematical theorem about determinants. [If you don't yet know much about determinants, don't worry about the proof and just skip to the definition of Eq. (6.15).]

First, we should say that Eq. (6.11) is the mathematical definition of a "product" of two matrices. (It is just convenient to be able to say: "R^{US} is the product of R^{UT} and R^{TS}.") Second, there is a theorem of mathematics—which you can easily prove for the two-by-two matrices we have here—which says that the determinant of a "product" of two matrices is the product of their determinants. Applying this theorem to Eq. (6.12), we get

$$e^{i2\delta}\,(\text{Det } R^{US}) = (\text{Det } R^{UT}) \cdot (\text{Det } R^{TS}). \tag{6.13}$$

(We leave off the subscripts, because they don't tell us anything useful.) Yes, the 2δ is right. Remember that we are dealing with two-by-two matrices; every term in the matrix R_{ki}^{US} is multiplied by $e^{i\delta}$, so each product in the determinant—which has *two* factors—gets multiplied by $e^{i2\delta}$. Now let's take the square root of Eq. (6.13) and divide it into Eq. (6.12); we get

$$\frac{R_{ki}^{US}}{\sqrt{\text{Det } R^{US}}} = \sum_j \frac{R_{kj}^{UT}}{\sqrt{\text{Det } R^{UT}}} \frac{R_{ji}^{TS}}{\sqrt{\text{Det } R^{TS}}}. \tag{6.14}$$

The extra phase factor has disappeared.

Now it turns out that if we want all of our amplitudes in any given representation to be normalized (which means, you remember, that $\sum_i \langle\phi|i\rangle\langle i|\phi\rangle = 1$), the rotation matrices will all have determinants that are pure imaginary exponentials, like $e^{i\alpha}$. (We won't prove it; you will see that it always comes out that way.) So we can, if we wish, choose to make all our rotation matrices R have a unique phase by making Det $R = 1$. It is done like this. Suppose we find a rotation matrix R in some arbitrary way. We make it a rule to "convert" it to "standard form" by defining

$$R_{\text{standard}} = \frac{R}{\sqrt{\text{Det } R}}. \tag{6.15}$$

We can do this because we are just multiplying each term of R by the same phase factor, to get the phases we want. In what follows, we will always assume that our matrices have been put in the "standard form"; then we can use Eq. (6.11) without having any extra phase factors.

6–3 Rotations about the z-axis

We are now ready to find the transformation matrix R_{ji} between two different representations. With our rule for compounding rotations and our assumption that space has no preferred direction, we have the keys we need for finding the matrix of any arbitrary rotation. There is only *one* solution. We begin with the transformation which corresponds to a rotation about the z-axis. Suppose we have two apparatuses S and T placed in series along a straight line with their axes parallel and pointing out of the page, as shown in Fig. 6–4(a). We take our "z-axis" in this direction. Surely, if the beam goes "up" (toward $+z$) in the S apparatus, it will do the same in the T apparatus. Similarly, if it goes down in S, it will go down in T. Suppose, however, that the T apparatus were placed at some other angle, but still with its axis parallel to the axis of S, as in Fig. 6–4(b). Intuitively, you would say that a $(+)$ beam in S would still go with a $(+)$ beam in T, because the fields and field gradients are still in the same physical direction. And that would be quite right. Also, a $(-)$ beam in S would still go into a $(-)$ beam in T. The same result would apply for any orientation of T in the xy-plane of S. What does this tell us about the relation between $C'_+ = \langle +T \mid \psi\rangle$, $C'_- = \langle -T \mid \psi\rangle$ and $C_+ = \langle +S \mid \psi\rangle$, $C_- = \langle -S \mid \psi\rangle$? You might conclude that any rotation about the z-axis of the "frame of reference" for base states leaves the amplitudes C_+ to be "up" and "down," the same as before. We could write $C'_+ = C_+$ and $C'_- = C_-$ —but that is *wrong*. All we *can* conclude is that for such rotations the probabilities to be in the "up" beam are the same for the S and T apparatuses. That is,

$$|C'_+| = |C_+| \qquad \text{and} \qquad |C'_-| = |C_-|.$$

We cannot say that the *phases* of the amplitudes referred to the T apparatus may not be different for the two different orientations in (a) and (b) of Fig. 6–4.

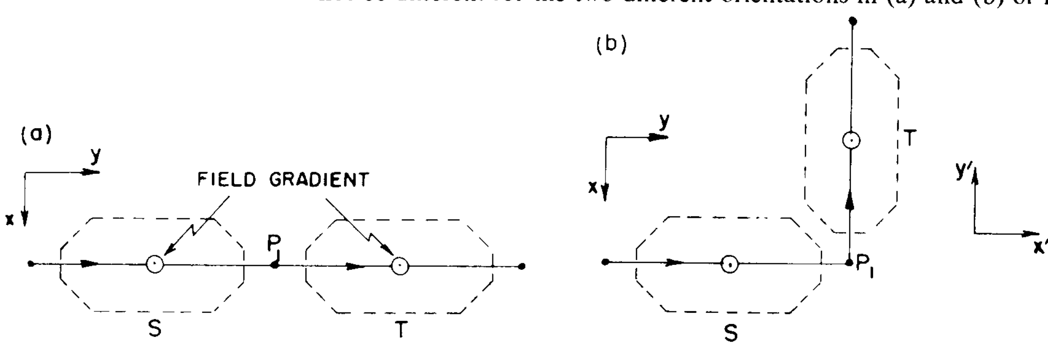

Fig. 6–4. Rotating 90° about the z-axis.

The two apparatuses in (a) and (b) of Fig. 6–4 are, in fact, different, as we can see in the following way. Suppose that we put an apparatus in front of S which produces a pure $(+x)$ state. (The x-axis points toward the bottom of the figure.) Such particles would be split into $(+z)$ and $(-z)$ beams in S, but the two beams would be recombined to give a $(+x)$ state again at P_1—the exit of S. The same thing happens again in T. If we follow T by a third apparatus U, whose axis is in the $(+x)$ direction and, as shown in Fig. 6–5(a), all the particles would go into the $(+)$ beam of U. Now imagine what happens if T and U are swung around *together* by 90° to the positions shown in Fig. 6–5(b). Again, the T apparatus puts out just what it takes in, so the particles that enter U are in a $(+x)$ state with respect to S. But U now analyzes for the $(+y)$ state with respect to S, which is different. (By symmetry, we would now expect only one-half of the particles to get through.)

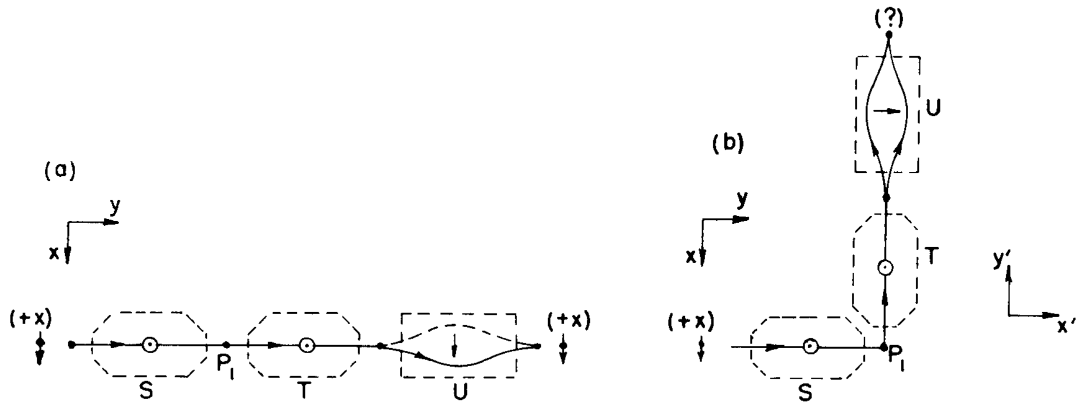

Fig. 6–5. Particle in a (+x) state behaves differently in (a) and (b).

What could have changed? The apparatuses T and U are still in the same *physical* relationship to each other. Can the *physics* be changed just because T and U are in a different orientation? Our original assumption is that it should not. It must be that the *amplitudes* with respect to T are different in the two cases shown in Fig. 6–5—and, therefore, also in Fig. 6–4. There must be some way for a particle to know that it has turned the corner at P_1. How could it tell? Well, all we have decided is that the *magnitudes* of C_1' and C_2' are the same in the two cases, but they could—in fact, *must*—have different *phases*. We conclude that C_+' and C_+ must be related by

$$C_+' = e^{i\lambda}C_+,$$

and that C_-' and C_- must be related by

$$C_-' = e^{i\mu}C_-,$$

where λ and μ are real numbers which must be related in some way to the angle between S and T.

The only thing we can say at the moment about λ and μ is that they must not be equal [except for the special case shown in Fig. 6–5(a), when T is in the same orientation as S]. We have seen that equal phase changes in all amplitudes have no physical consequence. For the same reason, we can always add the same arbitrary amount to both λ and μ without changing anything. So we are permitted to *choose* to make λ and μ equal to plus and minus the same number. That is, we can always take

$$\lambda' = \lambda - \frac{(\lambda + \mu)}{2}, \qquad \mu' = \mu - \frac{(\lambda + \mu)}{2}.$$

Then

$$\lambda' = \frac{\lambda}{2} - \frac{\mu}{2} = -\mu'.$$

So we adopt the convention† that $\mu = -\lambda$. We have then the general rule that for a rotation of the reference apparatus by some angle about the z-axis, the transformation is

$$C_+' = e^{+i\lambda}C_+, \qquad C_-' = e^{-i\lambda}C_-. \tag{6.16}$$

The absolute values are the same, only the phases are different. These phase factors are responsible for the different results in the two experiments of Fig. 6–5.

Now we would like to know the law that relates λ to the angle between S and T. We already know the answer for one case. If the angle is zero, λ is zero. Now we will *assume* that the phase shift λ is a continuous function of angle ϕ between S and T (see Fig. 6–4) as ϕ goes to zero—as only seems reasonable. In

† Looking at it another way, we are just putting the transformation in the "standard form" described in Section 6–2 by using Eq. (6.15).

ther words, if we rotate T from the straight line through S by the small angle ϵ, the λ is also a small quantity, say $m\epsilon$, where m is some number. We write it this way because we can show that λ must be proportional to ϵ. Suppose we were to put after T another apparatus T' which makes the angle ϵ with T, and, therefore, the angle 2ϵ with S. Then, with respect to T, we have

$$C'_+ = e^{i\lambda}C_+,$$

and with respect to T', we have

$$C''_+ = e^{i\lambda}C'_+ = e^{i2\lambda}C_+.$$

But we know that we should get the same result if we put T' right after S. Thus, when the angle is doubled, the phase is doubled. We can evidently extend the argument and build up any rotation at all by a sequence of infinitesimal rotations. We conclude that for *any* angle ϕ, λ is proportional to the angle. We can, therefore, write $\lambda = m\phi$.

The general result we get, then, is that for T rotated about the z-axis by the angle ϕ with respect to S

$$C'_+ = e^{im\phi}C_+, \qquad C'_- = e^{-im\phi}C_-. \tag{6.17}$$

For the angle ϕ, and for all rotations we speak of in the future, we adopt the standard convention that a *positive* rotation is a *right-handed* rotation about the positive direction of the reference axis. A positive ϕ has the sense of rotation of a right-handed screw advancing in the positive z-direction.

Now we have to find what m must be. First, we might try this argument: Suppose T is rotated by 360°; then, clearly, it is right back at zero degrees, and we should have $C'_+ = C_+$ and $C'_- = C_-$, or, what is the same thing, $e^{im2\pi} = 1$. We get $m = 1$. *This argument is wrong!* To see that it is, consider that T is rotated by 180°. If m were equal to 1, we would have $C'_+ = e^{i\pi}C_+ = -C_+$ and $C'_- = e^{-i\pi}C_- = -C_-$. However, this is just the *original* state all over again. *Both* amplitudes are just multiplied by -1 which gives back the original physical system. (It is again a case of a common phase change.) This means that if the angle between T and S in Fig. 6–5(b) is increased to 180°, the system (with respect to T) would be indistinguishable from the zero-degree situation, and the particles would again go through the $(+)$ state of the U apparatus. At 180°, though, the $(+)$ state of the U apparatus is the $(-x)$ state of the original S apparatus. So a $(+x)$ state would become a $(-x)$ state. But we have done nothing to *change* the original state; the answer is wrong. We cannot have $m = 1$.

We must have the situation that a rotation by 360° and *no smaller angle* reproduces the same physical state. This will happen if $m = \frac{1}{2}$. Then, and only then, will the first angle that reproduces the same *physical* state be $\phi = 360°$.† It gives

$$\left.\begin{aligned} C'_+ &= -C_+ \\ C'_- &= -C_- \end{aligned}\right\} 360^\circ \text{ about } z\text{-axis}. \tag{6.18}$$

It is very curious to say that if you turn the apparatus 360° you get new amplitudes. They aren't really new, though, because the common change of sign doesn't give any different physics. If someone else had decided to change all the signs of the amplitudes because he thought he had turned 360°, that's all right; he gets the same physics.‡ So our final answer is that if we know the amplitudes C_+ and C_- for spin one-half particles with respect to a reference frame S, and we then use a base

† It appears that $m = -\frac{1}{2}$ would also work. However, we see in (6.17) that the change in sign merely redefines the notation for a spin-up particle.

‡ Also, if something has been rotated by a sequence of small rotations whose net result is to return it to the original orientation, it is possible to define the idea that it has been rotated 360°—as distinct from zero net rotation—if you have kept track of the whole history. (Interestingly enough, this is *not* true for a net rotation of 720°.)

system referred to T which is obtained from S by a rotation of ϕ around the z-axis, the new amplitudes are given in terms of the old by

$$\left.\begin{aligned} C'_+ &= e^{i\phi/2}C_+ \\ C'_- &= e^{-i\phi/2}C_- \end{aligned}\right\} \phi \text{ about } z. \tag{6.19}$$

6–4 Rotations of 180° and 90° about y

Next, we will try to guess the transformation for a rotation of T with respect to S of 180° around an axis *perpendicular* to the z-axis—say, about the y-axis. (We have defined the coordinate axes in Fig. 6–1.) In other words, we start with two identical Stern-Gerlach equipments, with the second one, T, turned "upside down" with respect to the first one, S, as in Fig. 6–6. Now if we think of our particles as little magnetic dipoles, a particle that is the $(+S)$ state—so that it goes on the "upper" path in the first apparatus—will also take the "upper" path in the second, so that it will be in the *minus* state with respect to T. (In the inverted T apparatus, both the gradients *and* the field direction are reversed; for a particle with its magnetic moment in a given direction, the force is unchanged.) Anyway, what is "up" with respect to S will be "down" with respect to T. For these relative positions of S and T, then, we know that the transformation must give

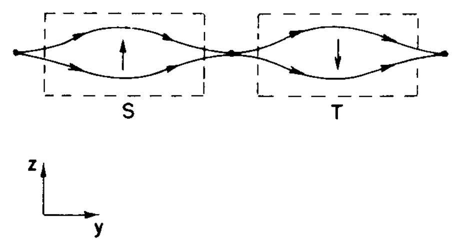

Fig. 6–6. A rotation of 180° about the y-axis.

$$|C'_+| = |C_-|, \qquad |C'_-| = |C_+|.$$

As before, we cannot rule out some additional phase factors; we could have (for 180° about the y-axis)

$$C'_+ = e^{i\beta}C_- \quad \text{and} \quad C'_- = e^{i\gamma}C_+, \tag{6.20}$$

where β and γ are still to be determined.

What about a rotation of 360° about the y-axis? Well, we already know the answer for a rotation of 360° about the z-axis—the amplitude to be in any state changes sign. A rotation of 360° around any axis always brings us back to the original position. It must be that for *any* 360° rotation, the result is the same as a 360° rotation about the z-axis—all amplitudes simply change sign. Now suppose we imagine two successive rotations of 180° about y—using Eq. (6.20)—we should get the result of Eq. (6.18). In other words,

$$C''_+ = e^{i\beta}C'_- = e^{i\beta}e^{i\gamma}C_+ = -C_+$$

and

$$C''_- = e^{i\gamma}C'_+ = e^{i\gamma}e^{i\beta}C_- = -C_-. \tag{6.21}$$

This means that

$$e^{i\beta}e^{i\gamma} = -1 \quad \text{or} \quad e^{i\gamma} = -e^{-i\beta}.$$

So the transformation for a rotation of 180° about the y-axis can be written

$$C'_+ = e^{i\beta}C_-, \qquad C'_- = -e^{-i\beta}C_+. \tag{6.22}$$

The arguments we have just used would apply equally well to a rotation of 180° about *any* axis in the xy-plane, although different axes can, of course, give different numbers for β. However, that is the only way they can differ. Now there is a certain amount of arbitrariness in the number β, but once it is specified for one axis of rotation in the xy-plane it is determined for any other axis. It is *conventional* to choose to set $\beta = 0$ for a 180° rotation about the y-axis.

To show that we have this choice, suppose we imagine that β was not equal to zero for a rotation about the y-axis; then we can show that there is *some other* axis in the xy-plane, for which the corresponding phase factor *will* be zero. Let's find the phase factor β_A for an axis A that makes the angle α with the y-axis, as shown in Fig. 6–7(a). (For clarity, the figure is drawn with α equal to a negative number, but that doesn't matter.) Now if we take a T apparatus which is initially lined up with the S apparatus and is then rotated 180° about the axis A, its axes—which we will call x'', y'', and z''—will be as shown in Fig 6–7(a). The amplitudes with respect to T will then be

$$C''_+ = e^{i\beta_A}C_-, \qquad C''_- = -e^{-i\beta_A}C_+. \tag{6.23}$$

We can now think of getting to the same orientation by the two successive rotations shown in (b) and (c) of the figure. First, we imagine an apparatus U which is rotated with respect to S by 180° about the y-axis. The axes x', y', and z' of U will be as shown in Fig. 6–7(b), and the amplitudes *with respect to* U are given by (6.22).

Now notice that we can go from U to T by a rotation about the "z-axis" of U, namely about z', as shown in Fig. 6–7(c). From the figure you can see that the angle required is two times the angle α but in the opposite direction (with respect to z'). Using the transformation of (6.19) with $\phi = -2\alpha$, we get

$$C''_+ = e^{-i\alpha}C'_+, \qquad C''_- = e^{+i\alpha}C'_-. \tag{6.24}$$

Combining Eqs. (6.24) and (6.22), we get that

$$C''_+ = e^{i(\beta-\alpha)}C_-, \qquad C''_- = -\,e^{-i(\beta-\alpha)}C_+. \tag{6.25}$$

These amplitudes must, of course, be the same as we got in 6.23). So β_A must be related to α and β by

$$\beta_A = \beta - \alpha. \tag{6.26}$$

This means that if the angle α between the A-axis and the y-axis (of S) is equal to β, the transformation for a rotation of 180° about A will have $\beta_A = 0$.

Now so long as *some* axis perpendicular to the z-axis is going to have $\beta = 0$, we may as well take it to be the y-axis. It is purely a matter of *convention*, and we adopt the one in general use. Our result: For a rotation of 180° about the y-axis, we have

$$\left.\begin{aligned} C'_+ &= C_- \\ C'_- &= -C_+ \end{aligned}\right\} 180^\circ \text{ about } y. \tag{6.27}$$

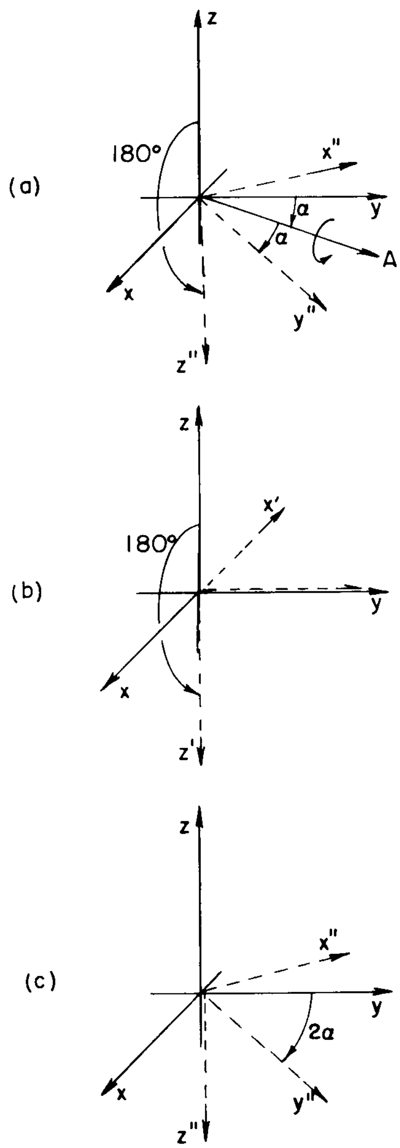

Fig. 6–7. A 180° rotation about the axis A is equivalent to a rotation of 180° about y, followed by a rotation about z'.

While we are thinking about the y-axis, let's next ask for the transformation matrix for a rotation of 90° about y. We can find it because we know that two successive 90° rotations about the same axis must equal one 180° rotation. We start by writing the transformation for 90° in the most general form:

$$C'_+ = aC_+ + bC_-, \qquad C'_- = cC_+ + dC_-. \tag{6.28}$$

A second rotation of 90° about the same axis would have the same coefficients:

$$C''_+ = aC'_+ + bC'_-, \qquad C''_- = cC'_+ + dC'_-. \tag{6.29}$$

Combining Eqs. (6.28) and (6.29), we have

$$\begin{aligned} C''_+ &= a(aC_+ + bC_-) + b(cC_+ + dC_-), \\ C''_- &= c(aC_+ + bC_-) + d(cC_+ + dC_-). \end{aligned} \tag{6.30}$$

However, from (6.27) we know that

$$C''_+ = C_-, \qquad C''_- = -C_+,$$

so that we must have that

$$\begin{aligned} ab + bd &= 1, \\ a^2 + bc &= 0, \\ ac + cd &= -1, \\ bc + d^2 &= 0. \end{aligned} \tag{6.31}$$

These four equations are enough to determine all our unknowns: a, b, c, and d.

It is not hard to do. Look at the second and fourth equations. Deduce that $a^2 = d^2$, which means that $a = d$ or else that $a = -d$. But $a = -d$ is out, because then the first equation wouldn't be right. So $d = a$. Using this, we have immediately that $b = 1/2a$ and that $c = -1/2a$. Now we have everything in terms of a. Putting, say, the second equation all in terms of a, we have

$$a^2 - \frac{1}{4a^2} = 0 \qquad \text{or} \qquad a^4 = \frac{1}{4}.$$

This equation has four different solutions, but only two of them give the standard value for the determinant. We might as well take $a = 1/\sqrt{2}$; then†

$$a = 1/\sqrt{2}, \qquad b = 1/\sqrt{2},$$
$$c = -1/\sqrt{2}, \qquad d = 1/\sqrt{2}.$$

In other words, for two apparatuses S and T, with T rotated with respect to S by 90° about the y-axis, the transformation is

$$\left.\begin{aligned} C'_+ &= \frac{1}{\sqrt{2}}(C_+ + C_-) \\ C'_- &= \frac{1}{\sqrt{2}}(-C_+ + C_-) \end{aligned}\right\} 90^\circ \text{ about } y. \qquad (6.32)$$

We can, of course, solve these equations for C_+ and C_-, which will give us the transformation for a rotation of *minus* 90° about y. Changing the primes around, we would conclude that

$$\left.\begin{aligned} C'_+ &= \frac{1}{\sqrt{2}}(C_+ - C_-) \\ C'_- &= \frac{1}{\sqrt{2}}(C_+ + C_-) \end{aligned}\right\} -90^\circ \text{ about } y. \qquad (6.33)$$

6-5 Rotations about x

You may be thinking: "This is getting ridiculous. What are they going to do next, 47° around y, then 33° about x, and so on, forever?" No, we are almost finished. With just two of the transformations we have—90° about y, and an arbitrary angle about z (which we did first if you remember)—we can generate any rotation at all.

As an illustration, suppose that we want the angle α around x. We know how to deal with the angle α around z, but now we want it around x. How do we get it? First, we turn the axis z down onto x—which is a rotation of $+90°$ about y, as shown in Fig. 6-8. Then we turn through the angle α around z'. Then we rotate $-90°$ about y''. The net result of the three rotations is the same as turning around x by the angle α. It is a property of space.

(These facts of the combinations of rotations, and what they produce, are hard to grasp intuitively. It is rather strange, because we live in three dimensions, but it is hard for us to appreciate what happens if we turn this way and then that way. Perhaps, if we were fish or birds and had a real appreciation of what happens when we turn somersaults in space, we could more easily appreciate such things.)

Anyway, let's work out the transformation for a rotation by α around the x-axis by using what we know. From the first rotation by $+90°$ around y the amplitudes go according to Eq. (6.32). Calling the rotated axes x', y', and z', the

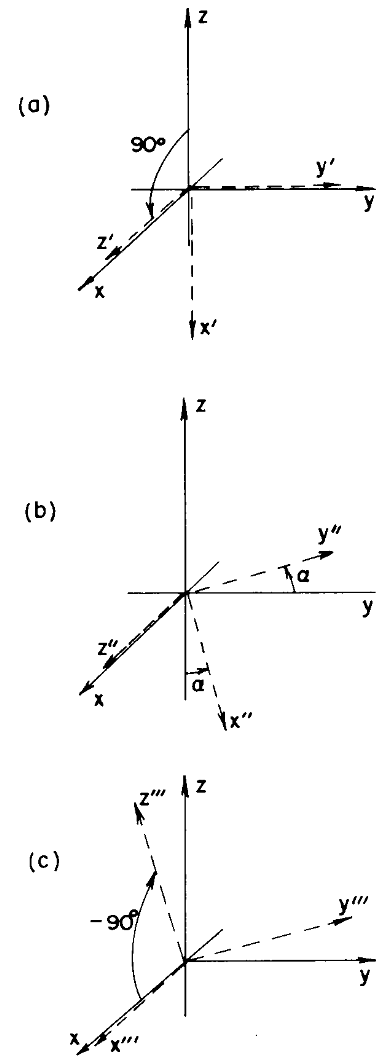

Fig. 6-8. A rotation by α about the x-axis is equivalent to: (a) a rotation by $+90°$ about y, followed by (b) a rotation by α about z', followed by (c) a rotation of $-90°$ about y''.

† The other solution changes all signs of a, b, c, and d and corresponds to a $-270°$ rotation.

next rotation by the angle α around z' takes us to a frame x'', y'', z'', for which

$$C''_+ = e^{i\alpha/2}C'_+, \qquad C''_- = e^{-i\alpha/2}C'_- .$$

The last rotation of $-90°$ about y'' takes us to x''', y''', z'''; by (6.33),

$$C'''_+ = \frac{1}{\sqrt{2}}(C''_+ - C''_-), \qquad C'''_- = \frac{1}{\sqrt{2}}(C''_+ + C''_-).$$

Combining these last two transformations, we get

$$C'''_+ = \frac{1}{\sqrt{2}}(e^{+i\alpha/2}C'_+ - e^{-i\alpha/2}C'_-),$$

$$C'''_- = \frac{1}{\sqrt{2}}(e^{+i\alpha/2}C'_+ + e^{-i\alpha/2}C'_-).$$

Using Eqs. (6.32) for C'_+ and C'_-, we get the complete transformation:

$$C'''_+ = \tfrac{1}{2}\{e^{+i\alpha/2}(C_+ + C_-) - e^{-i\alpha/2}(-C_+ + C_-)\},$$
$$C'''_- = \tfrac{1}{2}\{e^{+i\alpha/2}(C_+ + C_-) + e^{-i\alpha/2}(-C_+ + C_-)\}.$$

We can put these formulas in a simpler form by remembering that

$$e^{i\theta} + e^{-i\theta} = 2\cos\theta, \qquad \text{and} \qquad e^{i\theta} - e^{i\theta} = 2i\sin\theta.$$

We get

$$\left.\begin{aligned} C'''_+ &= \left(\cos\frac{\alpha}{2}\right)C_+ + i\left(\sin\frac{\alpha}{2}\right)C_- \\ C'''_- &= i\left(\sin\frac{\alpha}{2}\right)C_+ + \left(\cos\frac{\alpha}{2}\right)C_- \end{aligned}\right\}\ \alpha \text{ about } x. \tag{6.34}$$

Here is our transformation for a rotation about the x-axis by *any* angle α. It is only a little more complicated than the others.

6–6 Arbitrary rotations

Now we can see how to do *any* angle at all. First, notice that any relative orientation of two coordinate frames can be described in terms of three angles, as shown in Fig. 6–9. If we have a set of axes x', y', and z' oriented in any way at all with respect to x, y, and z, we can describe the relationship between the two frames by means of the three Euler angles α, β, and γ, which define three successive rotations that will bring the x, y, z frame into the x', y', z' frame. Starting at x, y, z, we rotate our frame through the angle β about the z-axis, bringing the x-axis to the line x_1. Then, we rotate by α about this temporary x-axis, to bring z down to z'. Finally, a rotation about the new z-axis (that is, z') by the angle γ will bring the x-axis into x' and the y-axis into y'.† We know the transformations for each of the three rotations—they are given in (6.19) and (6.34). Combining them in the proper order, we get

$$\begin{aligned} C'_+ &= \cos\frac{\alpha}{2}\,e^{i(\beta+\gamma)/2}C_+ + i\sin\frac{\alpha}{2}\,e^{-i(\beta-\gamma)/2}C_-, \\ C'_- &= i\sin\frac{\alpha}{2}\,e^{i(\beta-\gamma)/2}C_+ + \cos\frac{\alpha}{2}\,e^{-i(\beta+\gamma)/2}C_-. \end{aligned} \tag{6.35}$$

So just starting from some assumptions about the properties of space, we have derived the amplitude transformation for any rotation at all. That means that if

† With a little work you can show that the frame x, y, z can also be brought into the frame x', y', z' by the following three rotations about the *original* axes: (1) rotate by the angle γ around the original z-axis; (2) rotate by the angle α around the original x-axis; (3) rotate by the angle β around the original z-axis.

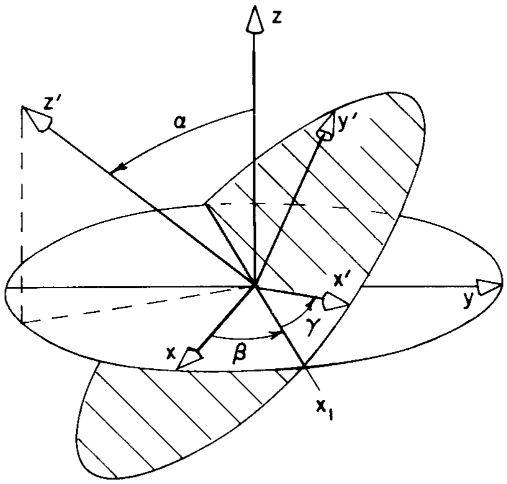

Fig. 6–9. The orientation of any coordinate frame x', y', z' relative to another frame x, y, z can be defined in terms of Euler's angles α, β, γ.

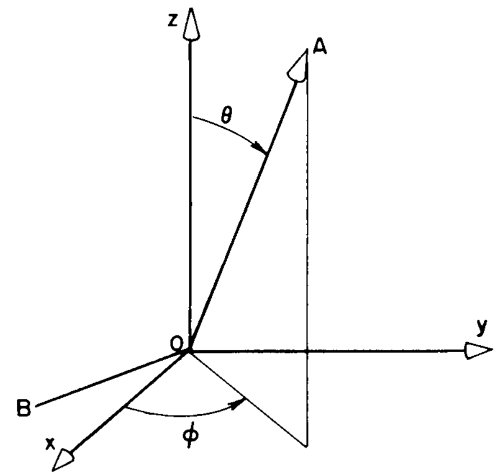

Fig. 6–10. An axis A defined by the polar angles θ and ϕ.

we know the amplitudes for any state of a spin one-half particle to go into the two beams of a Stern-Gerlach apparatus S, whose axes are x, y, and z, we can calculate what fraction would go into either beam of an apparatus T with the axes x', y', and z'. In other words, if we have a state ψ of a spin one-half particle, whose amplitudes are $C_+ = \langle + \mid \psi \rangle$ and $C_- = \langle - \mid \psi \rangle$ to be "up" and "down" with respect to the z-axis of the x, y, z frame, we also know the amplitudes C'_+ and C'_- to be "up" and "down" with respect to the z'-axis of any other frame x', y', z'. The four coefficients in Eqs. (6.35) are the terms of the "transformation matrix" with which we can project the amplitudes of a spin one-half particle into any other coordinate system.

We will now work out a few examples to show you how it all works. Let's take the following simple question. We put a spin one-half atom through a Stern-Gerlach apparatus that transmits only the $(+z)$ state. What is the amplitude that it will be in the $(+x)$ state? The $+x$ axis is the same as the $+z'$ axis of a system rotated 90° about the y-axis. For this problem, then, it is simplest to use Eqs. (6.32)—although you could, of course, use the complete equations of (6.35). Since $C_+ = 1$ and $C_- = 0$, we get $C'_+ = 1/\sqrt{2}$. The probabilities are the absolute square of these amplitudes; there is a 50 percent chance that the particle will go through an apparatus that selects the $(+x)$ state. If we had asked about the $(-x)$ state the amplitude would have been $-1/\sqrt{2}$, which also gives a probability 1/2—as you would expect from the symmetry of space. So if a particle is in the $(+z)$ state, it is equally likely to be in $(+x)$ or $(-x)$, but with opposite phase.

There's no prejudice in y either. A particle in the $(+z)$ state has a 50–50 chance of being in $(+y)$ or in $(-y)$. However, for these (using the formula for rotating $-90°$ about x), the amplitudes are $1/\sqrt{2}$ and $-i/\sqrt{2}$. In this case, the two amplitudes have a phase difference of 90° instead of 180°, as they did for the $(+x)$ and $(-x)$. In fact, that's how the distinction between x and y shows up.

As our final example, suppose that we know that a spin one-half particle is in a state ψ such that it is polarized "up" along some axis A, defined by the angles θ and ϕ in Fig. 6–10. We want to know the amplitude $\langle C_+ \mid \psi \rangle$ that the particle is "up" along z and the amplitude $\langle C_- \mid \psi \rangle$ that it is "down" along z. We can find these amplitudes by imagining that A is the z-axis of a system whose x-axis lies in some arbitrary direction—say in the plane formed by A and z. We can then bring the frame of A into x, y, z by three rotations. First, we make a rotation by $-\pi/2$ about the axis A, which brings the x-axis into the line B in the figure. Then we rotate by θ about line B (the new x-axis of frame A) to bring A to the z-axis. Finally, we rotate by the angle $(\pi/2 - \phi)$ about x. Remembering that we have only a $(+)$

state with respect to A, we get

$$C_+ = \cos\frac{\theta}{2}\,e^{-i\phi/2}, \qquad C_- = \sin\frac{\theta}{2}\,e^{+i\phi/2}. \tag{6.36}$$

We would like, finally, to summarize the results of this chapter in a form that will be useful for our later work. First, we remind you that our primary result in Eqs. (6.35) can be written in another notation. Note that Eqs. (6.35) mean just the same thing as Eq. (6.4). That is, in Eqs. (6.35) the coefficients of $C_+ = \langle +S \mid \psi\rangle$ and $C_- = \langle -S \mid \psi\rangle$ are just the amplitudes $\langle jT \mid iS\rangle$ of Eq. (6.4)—the amplitudes that a particle in the i-state with respect to S will be in the j-state with respect to T (when the orientation of T with respect to S is given in terms of the angles α, β, and γ). We also called them R_{ji}^{TS} in Eq. (6.6). (We have a plethora of notations!) For example, $R_{-+}^{TS} = \langle -T \mid +S\rangle$ is the coefficient of C_+ in the formula for C'_-, namely, $i\sin(\alpha/2)\,e^{i(\beta-\gamma)/2}$. We can, therefore, make a summary of our results in the form of a table, as we have done in Table 6–1.

It will occasionally be handy to have these amplitudes already worked out for some simple special cases. Let's let $R_z(\phi)$ stand for a rotation by the angle ϕ about the z-axis. We can also let it stand for the corresponding rotation matrix (omitting the subscripts i and j, which are to be implicitly understood). In the same spirit $R_x(\phi)$ and $R_y(\phi)$ will stand for rotations by the angle ϕ about the x-axis or the y-axis. We give in Table 6–2 the matrices—the tables of amplitudes $\langle jT \mid iS\rangle$—which project the amplitudes from the S-frame into the T-frame, where T is obtained from S by the rotation specified.

Table 6–1

The amplitudes $\langle jT \mid iS\rangle$ for a rotation defined by the Euler angles α, β, γ of Fig. 6–9

$R_{ji}(\alpha,\beta,\gamma)$

$\langle jT \mid iS\rangle$	$+S$	$-S$
$+T$	$\cos\frac{\alpha}{2}\,e^{i(\beta+\gamma)/2}$	$i\sin\frac{\alpha}{2}\,e^{-i(\beta-\gamma)/2}$
$-T$	$i\sin\frac{\alpha}{2}\,e^{i(\beta-\gamma)/2}$	$\cos\frac{\alpha}{2}\,e^{-i(\beta+\gamma)/2}$

Table 6–2

The amplitudes $\langle jT \mid iS\rangle$ for a rotation $R(\phi)$ by the angle ϕ about the z-axis, x-axis, or y-axis

$R_z(\phi)$

$\langle jT \mid iS\rangle$	$+S$	$-S$
$+T$	$e^{i\phi/2}$	0
$-T$	0	$e^{-i\phi/2}$

$R_x(\phi)$

$\langle jT \mid iS\rangle$	$+S$	$-S$
$+T$	$\cos\phi/2$	$i\sin\phi/2$
$-T$	$i\sin\phi/2$	$\cos\phi/2$

$R_y(\phi)$

$\langle jT \mid iS\rangle$	$+S$	$-S$
$+T$	$\cos\phi/2$	$\sin\phi/2$
$-T$	$-\sin\phi/2$	$\cos\phi/2$