3

Electrons and Their Interactions

This is the third of four lectures on a rather difficult subject—the theory of quantum electrodynamics—and since there are obviously more people here tonight than there were before, some of you haven't heard the other two lectures and will find this lecture almost incomprehensible. Those of you who *have* heard the other two lectures will also find this lecture incomprehensible, but you know that that's all right: as I explained in the first lecture, the way we have to describe Nature is generally incomprehensible to us.

In these lectures I want to tell you about the part of physics that we know best, the interaction of light and electrons. Most of the phenomena you are familiar with involve the interaction of light and electrons—all of chemistry and biology, for example. The only phenomena that are not covered by this theory are phenomena of gravitation and nuclear phenomena; everything else is contained in this theory.

We found out in the first lecture that we have no satisfactory mechanism to describe even the simplest of phenomena, such as partial reflection of light by glass. We also have no way to predict whether a given photon will be reflected or transmitted by the glass. All we can do is calculate the *probability* that a particular event will happen—

whether the light will be reflected, in this case. (This is about 4%, when the light shines straight down on a single surface of glass; the probability of reflection increases as the light hits the glass at more of a slant.)

When we deal with probabilities under *ordinary* circumstances, there are the following "rules of composition": 1) if something can happen in *alternative ways*, we *add* the probabilities for each of the different ways; 2) if the event occurs as a *succession of steps*—or depends on a number of things happening "concomitantly" (independently)—then we *multiply* the probabilities of each of the steps (or things).

In the wild and wonderful world of quantum physics, probabilities are calculated as the *square of the length of an arrow*: where we would have expected to add the probabilities under ordinary circumstances, we find ourselves "adding" *arrows*; where we normally would have multiplied the probabilities, we "multiply" *arrows*. The peculiar answers that we get from calculating probabilities in this manner match perfectly the results of experiment. I'm rather delighted that we must resort to such peculiar rules and strange reasoning in order to understand Nature, and I enjoy telling people about it. There are no "wheels and gears" beneath this analysis of Nature; if you want to understand Her, this is what you have to take.

Before I go into the main part of this lecture, I'd like to show you another example of how light behaves. What I would like to talk about is very weak light of one color—one photon at a time—going from a source, at S, to a detector, at D (see Fig. 49). Let's put a screen in between the source and the detector and make two very tiny holes a few millimeters apart from each other, at A and B. (If the source and detector are 100 centimeters apart, the holes have to be smaller than a tenth of a millimeter.) Let's put A in line with S and D, and put B somewhere to the side of A, not in line with S and D.

When we close the hole at B, we get a certain number of clicks at D—which represents the photons that came through A (let's say the detector clicks an average of one time for every 100 photons that leave S, or 1%). When we close the hole at A and open the hole at B, we know from the second lecture that we get nearly the same number of clicks, on average, because the holes are so small. (When we "squeeze" light too much, the rules of the ordinary world—such as light goes in straight lines—fall apart.)

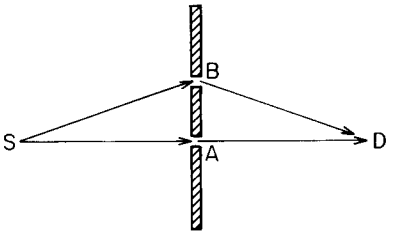


FIGURE 49. *Two tiny holes (at A and B) in a screen that is between a source S and a detector D let nearly the same amount of light through (in this case 1%) when one or the other hole is open. When both holes are open, "interference" occurs: the detector clicks from zero to 4% of the time, depending on the separation of A and B—shown in Figure 51 (a).*

When we open both holes we get a complicated answer, because interference is present: If the holes are a certain distance apart, we get more clicks than the expected 2% (the maximum is about 4%); if the two holes are a slightly different distance apart, we get no clicks at all.

One would normally think that opening a second hole would *always* increase the amount of light reaching the detector, but that's not what actually happens. And so saying that the light goes "either one way or the other" is false. I still catch myself saying, "Well, it goes either this way or that way," but when I say that, I have to keep in mind that

I mean in the sense of adding amplitudes: the photon has an amplitude to go one way, *and* an amplitude to go the other way. If the amplitudes oppose each other, the light won't get there—even though, in this case, both holes are open.

Now, here's an extra twist to the strangeness of Nature that I'd like to tell you about. Suppose we put in some special detectors—one at A and one at B (it is possible to design a detector that can tell whether a photon went through it)—so we can tell through which hole(s) the photon goes when both holes are open (see Fig. 50). Since the

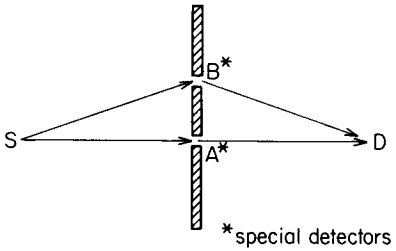


FIGURE 50. *When special detectors are put in at A and B to tell which way the light went when both holes are open, the experiment has been changed. Because a photon always goes through one hole or the other (when you are checking the holes), there are two distinguishable final conditions: 1) the detectors at A and D go off, and 2) the detectors at B and D go off. The probability of either event happening is about 1%. The probabilities of the two events are added in the normal way, which accounts for a 2% probability that the detector at D goes off—shown in Figure 51(b).*

probability that a single photon will get from S to D is affected only by the distance between the holes, there must be some sneaky way that the photon divides in two and then comes back together again, right? According to this hypothesis, the detectors at A and B should always go off together (at half strength, perhaps?), while the detector at D should go off with a probability of from zero to 4%, depending on the distance between A and B.

Here's what actually happens: the detectors at A and B *never* go off together—either A *or* B goes off. The photon does not divide in two; it goes one way or the other.

Furthermore, under such conditions the detector at D goes off 2% of the time—the simple sum of the probabilities for A and B (1% + 1%). The 2% is not affected by the spacing between A and B; the interference *disappears* when detectors are put in at A and B!

Nature has got it cooked up so we'll never be able to figure out how She does it: if we put instruments in to find out which way the light goes, we can find out, all right, but the wonderful interference effects disappear. But if we don't have instruments that can tell which way the light goes, the interference effects come back! Very strange, indeed!

To understand this paradox, let me remind you of a most important principle: in order to correctly calculate the probability of an event, one must be very careful to *define the complete event clearly*—in particular, what the initial conditions and the final conditions of the experiment are. You look at the equipment before and after the experiment, and look for changes. When we were calculating the probability that a photon gets from S to D with no detectors at A or B, the event was, simply, the detector at D makes a click. When a click at D was the only change in conditions, there was no way to tell which way the photon went, so there was interference.

When we put in detectors at A and B, we changed the problem. Now, it turns out, there are *two* complete events—two sets of final conditions—that are distinguishable: 1) the detectors at A and D go off, or 2) the detectors at B and D go off. When there are a number of possible final conditions in an experiment, we must calculate the probability of each as a separate, complete event.

To calculate the amplitude that the detectors at A and D go off, we multiply the arrows that represent the follow-

ing steps: a photon goes from S to A, the photon goes from A to D, and the detector at D goes off. The square of the final arrow is the probability of this event—1%—the same as when the hole at B was closed, because both cases have exactly the same steps. The other complete event is the detectors at B and D go off. The probability of this event is calculated in a similar way, and is also the same as before—about 1%.

If we want to know how often the detector at D goes off and we don't care whether it was A or B that went off in the process, the probability is the simple sum of the two events—2%. In principle, if there is something left in the system that we *could have* observed to tell which way the photon went, we have different "final states" (distinguishable final conditions), and we add the *probabilities*—not the amplitudes—for each final state.[1]

I have pointed out these things because the more you see how strangely Nature behaves, the harder it is to make a model that explains how even the simplest phenomena actually work. So theoretical physics has given up on that.

We saw in the first lecture how an event can be divided into alternative ways and how the arrow for each way can be "added." In the second lecture, we saw how each way can be divided into successive steps, how the arrow for each step can be regarded as the transformation of a unit arrow,

[1] The complete story on this situation is very interesting: if the detectors at A and B are not perfect, and detect photons only *some* of the time, then there are *three* distinguishable final conditions: 1) the detectors at A and D go off; 2) the detectors at B and D go off, and 3) the detector at D goes off alone, with A and B unchanged (they are left in their initial state). The probabilities for the first two events are calculated in the way explained above (except that there will be an extra step—a shrink for the probability that the detector at A [or B] goes off, since the detectors are not perfect). When D goes off alone, we can't separate the two cases, and Nature plays with us by bringing in interference—the same peculiar answer we would have had if there were no detectors (except that the final arrow is shrunk by the amplitude that the detectors do *not* go off). The final result is a mixture, the simple sum of all three cases (see Fig. 51). As the reliability of the detectors increases, we get less interference.

and how the arrows for each step can be "multiplied" by successive shrinks and turns. We are thus familiar with all the necessary rules for drawing and combining arrows (that represent bits and pieces of events) to obtain a final arrow, whose square is the probability of an observed physical event.

It is natural to wonder how far we can push this process of splitting events into simpler and simpler subevents. What are the smallest possible bits and pieces of events? Is there

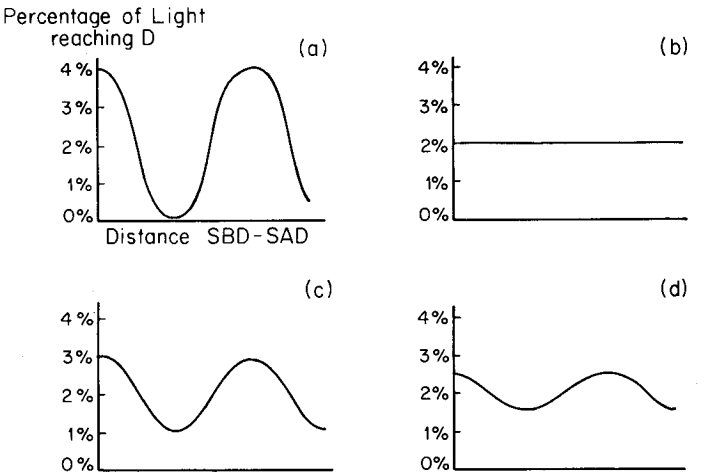


FIGURE 51. *When there are no detectors at A or B, there is interference—the amount of light varies from zero to 4% (a). When there are detectors at A and B that are 100% reliable, there is no interference—the amount of light reaching D is a constant 2% (b). When the detectors at A and B are not 100% reliable (i.e., when* sometimes *there is nothing left in A or in B that can be detected), there are now three possible final conditions—A and D go off, B and D go off, and D goes off alone. The final curve is thus a mixture, made up of contributions from each possible final condition. When the detectors at A and B are less reliable, there is more interference present. Thus the detectors in case (c) are less reliable than in case (d). The principle regarding interference is: The probability of each of the different possible final conditions must be independently calculated by adding arrows and squaring the length of the final arrow; after that, the several probabilities are added together in the normal fashion.*

a limited number of bits and pieces that can be compounded to form *all* the phenomena that involve light and electrons? Is there a limited number of "letters" in this language of quantum electrodynamics that can be combined to form "words" and "phrases" that describe nearly every phenomenon of Nature?

The answer is yes; the number is three. There are only three basic actions needed to produce all of the phenomena associated with light and electrons.

Before I tell you what these three basic actions are, I should properly introduce you to the actors. The actors are photons and electrons. The photons, particles of light, have been discussed at length in the first two lectures. Electrons were discovered in 1895 as particles: you could count them; you could put one of them on an oil drop and measure its electric charge. It gradually became apparent that the motion of these particles accounted for electricity in wires.

Shortly after electrons were discovered it was thought that atoms were like little solar systems, made up of a central, heavy part (called the nucleus) and electrons, which went around in "orbits," much like the planets do when they go around the sun. If you think that's the way atoms are, then you're back in 1910. In 1924 Louis De Broglie found that there was a wavelike character associated with electrons, and soon afterwards, C. J. Davisson and L. H. Germer of the Bell Laboratories bombarded a nickel crystal with electrons and showed that they, too, bounced off at crazy angles (just like X-rays do), and that these angles could be calculated from De Broglie's formula for the wavelength of an electron.

When we look at photons on a large scale—much larger than the distance required for one stopwatch turn—the phenomena that we see are very well approximated by rules such as "light travels in straight lines," because there are enough paths around the path of minimum time to reinforce each other, and enough other paths to cancel each other out. But when the space through which a photon moves becomes too small (such as the tiny holes in the screen), these rules fail—we discover that light doesn't have to go in straight lines, there are interferences created by two holes, and so on. The same situation exists with electrons: when seen on a large scale, they travel like particles, on definite paths. But on a small scale, such as inside an atom, the space is so small that there is no main path, no "orbit"; there are all sorts of ways the electron could go, each with an amplitude. The phenomenon of interference becomes very important, and we have to sum the arrows to predict where an electron is likely to be.

It's rather interesting to note that electrons looked like particles at first, and their wavish character was later discovered. On the other hand, apart from Newton making a mistake and thinking that light was "corpuscular," light looked like waves at first, and its characteristics as a particle were discovered later. In fact, both objects behave somewhat like waves, and somewhat like particles. In order to save ourselves from inventing new words such as "wavicles," we have chosen to call these objects "particles," but we all know that they obey these rules for drawing and combining arrows that I have been explaining. It appears that *all* the "particles" in Nature—quarks, gluons, neutrinos, and so forth (which will be discussed in the next lecture)—behave in this quantum mechanical way.

So now, I present to you the three basic actions, from which all the phenomena of light and electrons arise.

—ACTION #1: A photon goes from place to place.
—ACTION #2: An electron goes from place to place.
—ACTION #3: An electron emits or absorbs a photon.

Each of these actions has an amplitude—an arrow—that can be calculated according to certain rules. In a moment, I'll tell you those rules, or laws, out of which we can make

the whole world (aside from the nuclei, and gravitation, as always!).

Now, the stage on which these actions take place is not just space, it is space and time. Until now, I have disregarded problems concerning time, such as exactly when a photon leaves the source and exactly when it arrives at the detector. Although space is really three-dimensional, I'm going to reduce it to one dimension on the graphs that I'm going to draw: I will show a particular object's location in space on the horizontal axis, and the time on the vertical axis.

The first event I am going to draw in space and time—or space-time, as I might inadvertently call it—is a baseball standing still (See Fig. 52). On Thursday morning, which

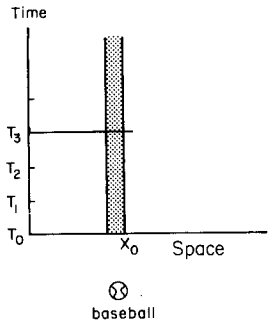


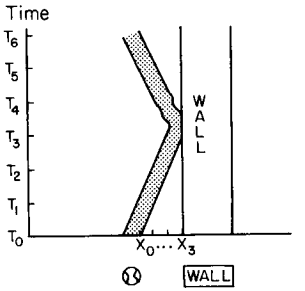
FIGURE 52. *The stage on which all actions in the universe take place is space-time. Usually consisting of four dimensions (three for space and one for time), space-time will be represented here in two dimensions—one for space, in the horizontal dimension, and one for time, in the vertical. Each time we look at the baseball (such as at time T_3), it is in the same place. This produces a "band of baseball" going straight up, as time goes on.*

I will label as T_0, the baseball occupies a certain space, which I will label as X_0. A few moments later, at T_1, it occupies the same space, because it's standing still. A few moments later, at T_2, the baseball is still at X_0. So the diagram of a baseball standing still is a vertical band, going straight up, with baseball all over it inside.

What happens if we have a baseball drifting in the weightlessness of outer space, going straight toward a wall? Well,

on Thursday morning (T_0) it starts at X_0 (see Fig. 53), but a little bit later, it's not in the same place—it has drifted over a little bit, to X_1. As the baseball continues to drift, it creates a slanted "band of baseball" on the diagram of space-time. When the baseball hits the wall (which is standing still and is therefore a vertical band), it goes back the other way, exactly where it came from in space (X_0), but to a different point in time (T_6).

FIGURE 53. *A baseball drifting directly toward a wall at right angles and then bouncing back to its original location (shown below the graph) is moving in one dimension and appears as a slanted "band of baseball." At times T_1 and T_2, the baseball is getting closer to the wall; at T_3 it hits the wall, and begins to go back.*



As for the time scale, it is most convenient to represent the time not in seconds, but in much smaller units. Since we will be dealing with photons and electrons, which move very rapidly, I am going to have a 45° angle represent something going the speed of light. For example, for a particle moving at the speed of light from X_1T_1 to X_2T_2, the horizontal distance between X_1 and X_2 is the same as the vertical distance between T_1 and T_2 (see Fig. 54). The factor by which time is stretched out (to make a 45° angle represent a particle going the speed of light) is called c, and you'll find c's flying around everywhere in Einstein's formulas—they are the result of the unfortunate choice of the second as the unit of time, rather than the time it takes light to go one meter.

Now, let's look at the first basic action in detail—a photon

goes from place to place. I will draw this action as a wiggly line from A to B for no good reason. I should be more careful: I should say, a photon that is known to be at a given place at a given time has a certain amplitude to get to another place at another time. On my space-time graph (see Fig. 55), the photon at point A—at X_1 and T_1—has an amplitude to appear at point B—X_2 and T_2. The size of this amplitude I will call P(A to B).

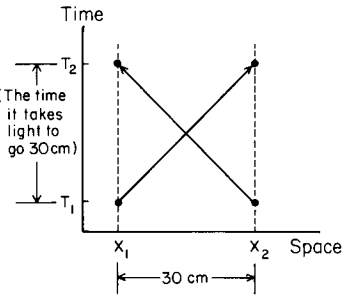


FIGURE 54. *The time scale I will use in these graphs will show particles going at the speed of light to be travelling at a 45-degree angle through space-time. The amount of time it takes light to go 30 centimeters—from X_1 to X_2 or from X_2 to X_1—is about one-billionth of a second.*

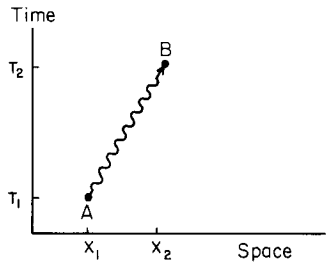


FIGURE 55. *A photon (represented by a wavy line) has an amplitude to go from a point A in space-time to another point, B. This amplitude, which I will call P(A to B), is calculated from a formula that depends only on the difference in location—$(X_2 - X_1)$—and the difference of the time—$(T_2 - T_1)$. In fact, it's a simple function that is the inverse of the difference of their squares—an "interval," I, that can be written as $(X_2 - X_1)^2 - (T_2 - T_1)^2$.*

There is a formula for the size of this arrow, P(A to B). It is one of the great laws of Nature, and it's very simple. It depends on the difference in *distance* and the difference in *time* between the two points. These differences can be expressed mathematically[2] as $(X_2 - X_1)$ and $(T_2 - T_1)$.

The major contribution to P(A to B) occurs at the conventional speed of light—when $(X_2 - X_1)$ is equal to $(T_2 - T_1)$—where one would expect it all to occur, but there is also an amplitude for light to go faster (or slower) than the conventional speed of light. You found out that in the last lecture that light doesn't go only in straight lines; now, you find out that it doesn't go only at the speed of light!

It may surprise you that there is an amplitude for a photon to go at speeds faster or slower than the conventional speed, c. The amplitudes for these possibilities are very small compared to the contribution from speed c; in fact, they cancel out when light travels over long distances. However, when the distances are short—as in many of the

[2] In these lectures, I am plotting a point's location in space in one dimension, along the x-axis. To locate a point in three-dimensional space, a "room" has to be set up, and the distance of the point from the floor and from each of two adjacent walls (all at right angles to each other) has to be measured. These three measurements can be labeled X_1, Y_1, and Z_1. The actual distance from this point to a second point with measurements X_2, Y_2, Z_2 can be calculated using a "three-dimensional Pythagorean Theorem": the square of this actual distance is

$$(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2.$$

The excess of *this* over the time difference, squared—

$$(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2 - (T_2 - T_1)^2$$

—is sometimes called "the Interval," or I, and is the combination that Einstein's theory of relativity says that P(A to B) must depend on. Most of the contribution to the final arrow for P(A to B) is just where you would expect it—where the difference in distance is equal to the difference in time (that is, when I is zero). But in addition; there is a contribution when I is not zero, that is inversely proportional to I: it points in the direction of 3 o'clock when I is more than zero (when light is going faster than c), and points toward 9 o'clock when I is less than zero. These later contributions cancel out in many circumstances (see Fig. 56).

diagrams I will be drawing—these other possibilities become vitally important and must be considered.

So that's the first basic action, the first basic law of physics—a photon goes from point to point. That explains all about optics; that's the entire theory of light! Well, not quite: I left out polarization (as always), and the interaction of light with matter, which brings me to the second law.

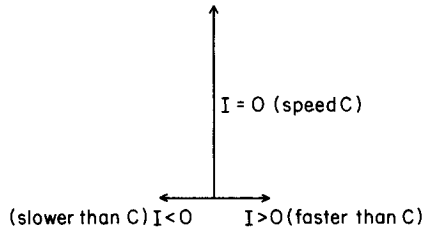


FIGURE 56. *When light goes at the speed* C, *the "interval,"* I, *equals zero, and there is a large contribution in the 12 o'clock direction. When I is greater than zero, there is a small contribution in the three o'clock direction inversely proportional to I; when I is less than zero, there is a similar contribution in the nine o'clock direction. Thus light has an amplitude to go faster or slower than speed* C, *but these amplitudes cancel out over long distances.*

The second action fundamental to quantum electrodynamics is: An electron goes from point A to point B in space-time. (For the moment we will imagine this electron as a simplified, fake electron, with no polarization—what the physicists call a "spin-zero" electron. In reality, electrons have a type of polarization, which doesn't add anything to the main ideas; it only complicates the formulas a little bit.) The formula for the amplitude for this action, which I will call E(A to B) also depends on $(X_2 - X_1)$ and $(T_2 - T_1)$ (in the same combination as described in note 2) as well as on a number I will call "n," a number that, once determined, enables all our calculations to agree with experiment. (We will see later how we determine n's value.) It is a rather complicated formula, and I'm sorry that I don't know how

to explain it in simple terms. However, you might be interested to know that the formula for P(A to B)—a photon going from place to place in space-time—is the same as that for E(A to B)—an electron going from place to place—if n is set to zero.[3]

The third basic action is: an electron emits or absorbs a photon—it doesn't make any difference which. I will call this action a "junction," or "coupling." To distinguish electrons from photons in my diagrams, I will draw each electron going through space-time as a straight line. Every coupling, therefore, is a junction between two straight lines and a wavy line (see Fig. 58). There is no complicated formula for the amplitude of an electron to emit or absorb a photon; it doesn't depend on anything—it's just a number! This junction number I will call j—its value is about -0.1: a shrink to about one-tenth, and half a turn.[4]

Well, that's all there is to these basic actions—except for some slight complications due to this polarization that we're

[3] The formula for E(A to B) is complicated, but there is an interesting way to explain what it amounts to. E(A to B) can be represented as a giant sum of a lot of different ways an electron could go from point A to point B in space-time (see Fig. 57): the electron could take a "one-hop flight," going directly from A to B; it could take a "two-hop flight," stopping at an intermediate point C; it could take a "three-hop flight," stopping at points D and E, and so on. In such an analysis, the amplitude for each "hop"—from one point F to another point G—is P(F to G), the same as the amplitude for a photon to go from a point F to a point G. The amplitude for each "stop" is represented by n^2, n being the same number I mentioned before which we used to make our calculations come out right.

The formula for E(A to B) is thus a series of terms: P(A to B) [the "one-hop" flight] + P(A to C)$*n^2*$P(C to B) ["two-hop" flights, stopping at C] + P(A to D)$*n^2*$P(D to E) $* n^2*$P(E to B) ["three-hop" flights, stopping at D and E] + . . . for *all possible intermediate points* C, D, E, and so on.

Note that when n increases, the nondirect paths make a greater contribution to the final arrow. When n is zero (as for the photon), all terms with an n drop out (because they are also equal to zero), leaving only the first term, which is P(A to B). Thus E(A to B) and P(A to B) are closely related.

[4] This number, the amplitude to emit or absorb a photon, is sometimes called the "charge" of a particle.

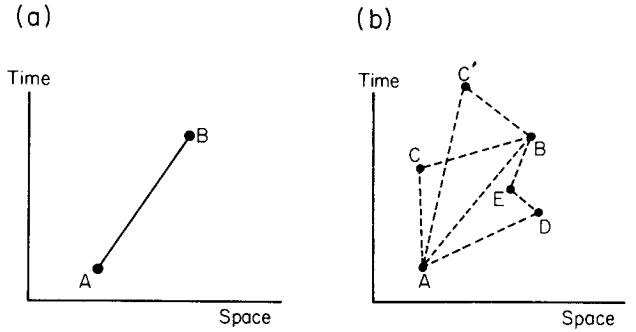


FIGURE 57. *An electron has an amplitude to go from point to point in space-time, which I will call "E(A to B)." Although I will represent E(A to B) as a straight line between two points (a), we can think of it as the sum of many amplitudes (b)—among them, the amplitude for the electron to change direction at points C or C' on a "two-hop" path, and the amplitude to change direction at D and E on a "three-hop" path—in addition to the direct path from A to B. The number of times an electron can change direction is anywhere from zero to infinity, and the points at which the electron can change direction on its way from A to B in space-time are infinite. All are included in E(A to B).*

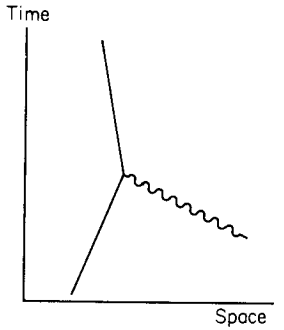


FIGURE 58. *An electron, depicted by a straight line, has a certain amplitude to emit or absorb a photon, shown by a wavy line. Since the amplitude to emit or absorb is the same, I will call either case a "coupling." The amplitude for a coupling is a number that I will call* j*; it is about −0.1 for the electron (this number is sometimes called the "charge").*

always leaving out. Our next job is to put these three actions together to represent circumstances that are somewhat more complicated.

For our first example, let's calculate the probability that two electrons, at points 1 and 2 in space-time, end up at

points 3 and 4 (see Fig. 59). This event can happen in several ways. The first way is that the electron at 1 goes to 3—computed by putting 1 and 3 into the formula E(A to B), which I will write as E(1 to 3)—and the electron at 2 goes to 4—computed by E(2 to 4). These are two "subevents" happening concomitantly, so the two arrows are multiplied to produce an arrow for this first way the event could happen. Therefore we write the formula for the "first-way arrow" as E(1 to 3) * E(2 to 4).

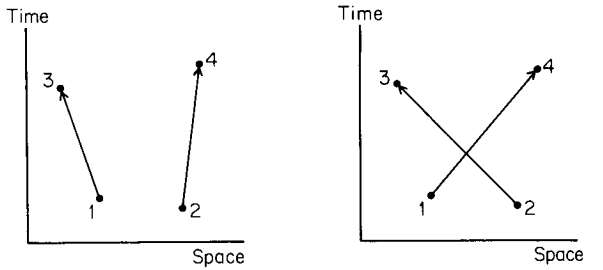


FIGURE 59. *To calculate the probability that electrons at points 1 and 2 in space-time end up at points 3 and 4, we calculate the "first way" arrow for 1 going to 3 and 2 going to 4 with the formula for E(A to B); then we calculate the "second way" arrow for 1 going to 4 and 2 going to 3 (a "crossover"). Finally, we add the "first way" and "second way" arrows to arrive at a good approximation of the final arrow. (This is true for the fake, simplified "spin zero" electron. Had we included the polarization of the electron, we would have subtracted—rather than added—the two arrows.)*

Another way this event could happen is that the electron at 1 goes to 4 and the electron at 2 goes to 3—again, two concomitant subevents. The "second-way arrow" is E(1 to 4) * E(2 to 3), and we add it to the "first-way" arrow.[5]

This is a good approximation for the amplitude of this event. To make a more exact calculation that will agree more closely with the results of experiment, we must con-

[5] Had I included the effects of the polarization of the electron, the "second-way" arrow would have been "subtracted"—turned 180° and added. (More on this comes later in this lecture.)

sider other ways this event could happen. For instance, for each of the two main ways the event can happen, one electron could go charging off to some new and wonderful place and emit a photon (see Fig. 60). Meanwhile, the other electron could go to some other place and absorb the pho-

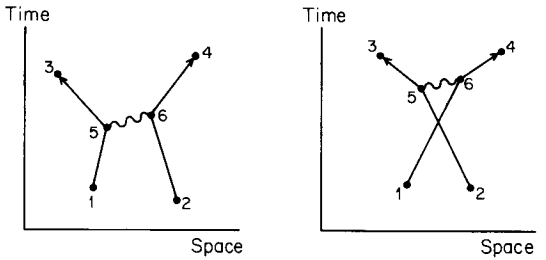


FIGURE 60. *Two "other ways" the event in Fig. 59 could happen are: a photon is emitted at 5 and absorbed at 6 for each case. The final conditions of these alternatives are the same as for the other cases—two electrons went in, and two electrons came out—and these results are indistinguishable from the other alternatives. Therefore the arrows for these "other ways" must be added to the arrows in Fig. 59 to arrive at a better approximation of the final arrow for the event.*

ton. Calculating the amplitude for the first of these new ways involves multiplying the amplitudes for: an electron goes from 1 to the new and wonderful place, 5 (where it emits a photon), and then goes from 5 to 3; the other electron goes from 2 to the other place, 6 (where it absorbs the photon), and then goes from 6 to 4. We must remember to include the amplitude that the photon goes from 5 to 6. I'm going to write the amplitude for this way the event could happen in a high-class mathematical fashion, and you can follow along: E(1 to 5)**j**E(5 to 3) * E(2 to 6)**j**E(6 to 4) * P(5 to 6)—a lot of shrinking and turning. (I'll let you figure out the notation for the other case, where the electron at 1 ends up at 4, and the electron at 2 ends up at 3.)[6]

[6] The final conditions of the experiment for these more complicated ways are the same as for the simpler ways—electrons start at points 1 and 2 and end up at points 3 and 4—so we cannot distinguish between these alternatives and the first two. Therefore we must add the arrows for these two ways to the two ways just previously considered.

But wait: positions 5 and 6 could be anywhere in space and time—yes, anywhere—and the arrows for *all* of those positions have to be calculated and added together. You see it's getting to be a lot of work. Not that the rules are so difficult—it's like playing checkers: the rules are simple, but you use them over and over. So our difficulty in calculating comes from having to pile so many arrows together. That's why it takes four years of graduate work for the students to learn how to do this efficiently—and we're looking at an *easy* problem! (When the problems get too difficult, we just put them on the computer!)

I would like to point out something about photons being emitted and absorbed: if point 6 is later than point 5, we might say that the photon was emitted at 5 and absorbed at 6 (see Fig. 61). If 6 is earlier than 5, we might prefer to say the photon was emitted at 6 and absorbed at 5, but we could just as well say that the photon is going backwards in time! However, we don't have to worry about which way in space-time the photon went; it's all included in the formula for P(5 to 6), and we say a photon was "exchanged." Isn't it beautiful how simple Nature is![7]

Now, in addition to the photon that is exchanged between 5 and 6, another photon could be exchanged—between two points, 7 and 8 (see Fig. 62). I'm too tired to write down all the basic actions whose arrows have to be multiplied, but—as you may have noticed—every straight line gets an E(A to B), every wavy line gets a P(A to B), and every coupling gets a *j*. Thus, there are six E(A to B)'s, two P(A to B)'s, and four *j*'s—for *every possible 5, 6, 7, and 8*! That makes billions of tiny arrows that have to be multiplied and then added together!

[7] Such an exchanged photon that never really appears in the initial or final conditions of the experiment is sometimes called a "virtual photon."

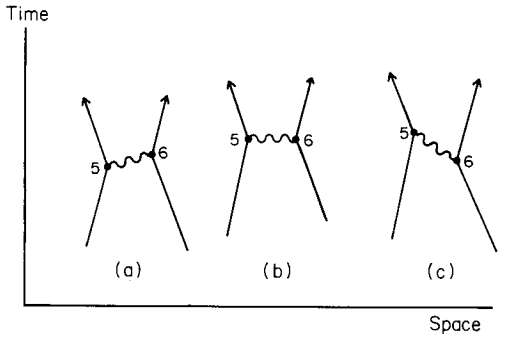


FIGURE 61. *Since light has an amplitude to go faster or slower than the conventional speed of light, the photons in all three examples above can be thought of as being emitted from point 5 and absorbed at point 6, even though the photon in example (b) is emitted at the* same time *that it is absorbed, and the photon in (c) is emitted* later *than it is absorbed—a situation in which you might have preferred to say that it was emitted by 6 and absorbed by 5; otherwise, the photon would have to go* backwards in time*! As far as calculating (and Nature) is concerned, it's all the same (and it's all possible), so we simply say a photon is "exchanged" and plug the locations in space-time into the formula for P(A to B).*

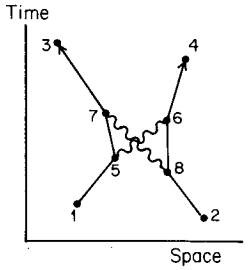


FIGURE 62. *Yet another way the event in Fig. 59 could happen is that* two *photons could be exchanged. Many diagrams of this way are possible (as we will see in more detail later); one of them is shown here. The arrow for this way involves all possible intermediate points 5, 6, 7, and 8, and is calculated with great difficulty. Because* j *is less than 0.1, the length of this arrow is generally less than 1 part in 10,000 (because there are four couplings involved) compared to the "first way" and "second way" arrows in Fig. 59 that contained no* j*'s.*

It appears that calculating the amplitude for this simple event is a hopeless business, but when you're a graduate student you've got to get your degree, so you keep on going.

But there *is* hope for success. It is found in that magic number, *j*. The first two ways the event could happen had

no *j*'s in the calculation; the next way had *j*j*, and the last way we looked at had *j*j*j*j*. Since *j*j* is less than 0.01, it means the length of the arrow for this way is generally less than 1% of the arrow for the first two ways; an arrow with *j*j*j*j* in it is less than 1% of 1%—one part in 10,000—compared to the arrows that have no *j*. If you've got enough time on the computer, you can work out the possibilities that involve j^6—one part in a million—and match the accuracy of the experiments. That's how the calculations of simple events are made. That's the way it works; that's all there is to it!

Let's look at another event now. We begin with a photon and an electron, and we end with a photon and an electron. One way this event can happen is: a photon is absorbed by an electron, the electron continues on a bit, and a new photon comes out. This process is called the scattering of light. When we make the diagrams and calculations for scattering, we must include some peculiar possibilities (see Fig. 63). For example, the electron could emit a photon *before* absorbing one (b). Even more strange is the possibility

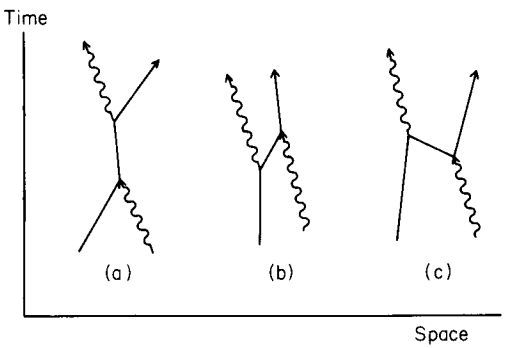


FIGURE 63. *The scattering of light involves a photon going into an electron and a photon coming out—not necessarily in that order, as seen in example (b). The example in (c) shows a strange but real possibility: the electron emits a photon, rushes backwards in time to absorb a photon, and then continues forwards in time.*

(c) that the electron emits a photon, then *travels backwards in time* to absorb a photon, and then proceeds forwards in time again. The path of such a "backwards-moving" electron can be so long as to appear real in an actual physical experiment in the laboratory. Its behavior is included in these diagrams and the equation for E(A to B).

The backwards-moving electron when viewed with time moving forwards appears the same as an ordinary electron, except it's attracted to normal electrons—we say it has a "positive charge." (Had I included the effects of polarization, it would be apparent why the sign of j for the backwards-moving electron appears reversed, making the charge appear positive.) For this reason it's called a "positron." The positron is a sister particle to the electron, and is an example of an "anti-particle."[8]

This phenomenon is general. Every particle in Nature has an amplitude to move backwards in time, and therefore has an anti-particle. When a particle and its anti-particle collide, they annihilate each other and form other particles. (For positrons and electrons annihilating, it is usually a photon or two.) And what about photons? Photons look exactly the same in all respects when they travel backwards in time—as we saw earlier—so they are their own anti-particles. You see how clever we are at making an exception part of the rule!

I'd like to show you what this backwards-moving electron looks like to us, as we move forwards in time. With a sequence of parallel lines to aid the eye, I'm going to divide the diagram into blocks of time, T_0 to T_{10} (see Fig. 64). We start at T_0 with an electron moving toward a photon, which is moving in the opposite direction. All of a sudden—at T_3—the photon turns into two particles, a positron and an

[8] Dirac proposed the reality of "anti-electrons" in 1931; in the following year, Carl Anderson found them experimentally and called them "positrons." Today, positrons can be easily made (for example, by making two photons collide with each other) and kept for weeks in a magnetic field.

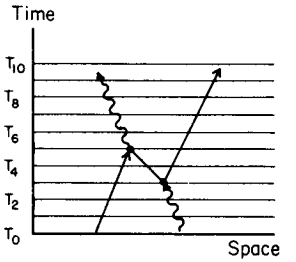


FIGURE 64. *Looking at example (c) from Fig. 63 going only forwards in time (as we are forced to do in the laboratory), from T_0 to T_3 we see the electron and photon moving toward each other. All of a sudden, at T_3 the photon "disintegrates" and two particles appear—an electron and a new kind of particle (called a "positron") which is an electron going backwards in time and which appears to move toward the original electron (itself!). At T_5 the positron annihilates with the original electron to produce a new photon. Meanwhile, the electron created by the earlier photon continues forwards in space-time. This sequence of events has been observed in the laboratory, and is included automatically in the formula for E(A to B) without any modification.*

electron. The positron doesn't last very long: it soon runs into the electron—at T_5, where they annihilate and produce a new photon. Meanwhile, the electron created earlier by the original photon continues on through space-time.

The next thing I would like to talk about is an electron in an atom. In order to understand the behavior of electrons in atoms, we have to add one other feature, the nucleus—the heavy part at the center of an atom that contains at least one proton (a proton is a "Pandora's Box" that we will open in the next lecture). I will not give you the correct laws for the behavior of the nucleus in this lecture; they are very complicated. But in this case, where the nucleus is quiet, we can approximate its behavior as that of a particle with an amplitude to go from one place to another in space-time according to the formula for E(A to B), but with a much higher number for n. Since the nucleus is so heavy compared to an electron, we can deal with it approximately here by saying that it stays in essentially one place as it moves through time.

The simplest atom, called hydrogen, is a proton and an electron. By exchanging photons, the proton keeps the electron nearby, dancing around it (see Fig. 65).[9] Atoms that contain more than one proton and the corresponding number of electrons also scatter light (atoms in the air scatter light from the sun and make the sky blue), but the diagrams for these atoms would involve so many straight and wiggly lines that they'd be a complete mess!

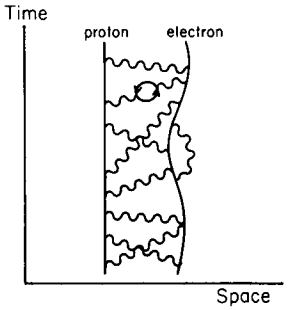


FIGURE 65. *An electron is kept within a certain range of distance to the nucleus of an atom by photon exchanges with a proton (a "Pandora's Box" that we will look into in Chapter 4). For now, the proton can be approximated as a stationary particle. Shown here is a hydrogen atom, consisting of a proton and an electron exchanging photons.*

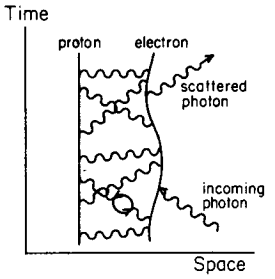


FIGURE 66. *The scattering of light by an electron in an atom is the phenomenon that accounts for partial reflection in a layer of glass. The diagram shows one way this event can happen in a hydrogen atom.*

Now, I'd like to show you a diagram of an electron in a hydrogen atom scattering light (see Fig. 66). As the electron and the nucleus are exchanging photons, a photon comes

[9] The amplitude for the photon exchange is $(-j) * P(A-B) * j$—two couplings and the amplitude for a photon to go from place to place. The amplitude for a proton to have a coupling with a photon is $-j$.

from outside the atom, hits the electron and is absorbed; then a new photon is emitted. (As usual, there are other possibilities to be considered, such as the new photon is emitted before the old photon is absorbed.) The total amplitude for all the ways an electron can scatter a photon can be summed up as a single arrow, a certain amount of shrink and turn. (Later, we will call this arrow "S.") This amount depends on the nucleus and the arrangement of the electrons in the atoms, and is different for different materials.

Now, let's look again at the partial reflection of light by a layer of glass. How does it work? I talked about light being reflected from the front surface and the back surface. This idea of surfaces was a simplification I made in order to keep things easy at the beginning. Light is really not affected by surfaces. An incoming photon is scattered by the electrons in the atoms inside the glass, and a *new* photon comes back up to the detector. It's interesting that instead of adding up all the billions of tiny arrows that represent the amplitude for all the electrons inside the glass to scatter an incoming photon, we can add just two arrows—for the "front surface" and "back surface" reflections—and come out with the same answer. Let's see why.

To discuss reflection by a layer from our new point of view we must take into account the dimension of time. Previously, when we talked about light from a monochromatic source, we used an imaginary stopwatch that times a photon as it moves—the hand of this stopwatch determined the angle of the amplitude for a given path. In the formula for P(A to B) (the amplitude for a photon to go from point to point) there is no mention of any turning. What happened to the stopwatch? What happened to the turning?

In the first lecture I simply said that the light source was monochromatic. To correctly analyze partial reflection by a layer, we need to know more about a monochromatic

light source. The amplitude for a photon to be emitted by a source varies, in general, with the *time*: as time goes on, the angle of the amplitude for a photon to be emitted by a source changes. A source of white light—many colors mixed together—emits photons in a chaotic manner: the angle of the amplitude changes abruptly and irregularly in fits and starts. But when we construct a *monochromatic* source, we are making a device that has been carefully arranged so that the amplitude for a photon to be emitted at a certain time is easily calculated: it changes its angle at a *constant* speed, like a stopwatch hand. (Actually, this arrow turns at the same speed as the imaginary stopwatch we used before, but in the opposite direction—see Fig. 67.)

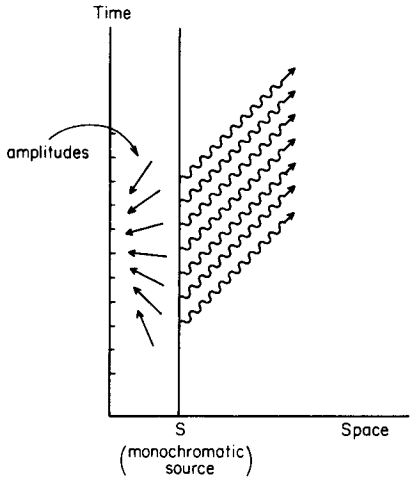


FIGURE 67. *A monochromatic source is a beautifully constructed apparatus that emits a photon in a very predictable way: the amplitude for a photon to be emitted at a certain* time *rotates counterclockwise as time moves forwards. Thus the amplitude for the source to emit a photon at a later time has a lesser angle. It will be assumed that all the light emitted from the source goes at speed* c *(since the distances are large).*

The rate of turning depends on the color of the light: the amplitude for a blue source turns nearly twice as fast as that for a red source, just as before. So the timer we used for the "imaginary stopwatch" was the monochromatic source:—in reality, the angle of the amplitude for a given

path depends on what *time* the photon is emitted from the source.

Once a photon has been emitted, there is no further turning of the arrow as a photon goes from one point to another in space-time. Although the formula P(A to B) says that there is an amplitude for light to go from one place to another at speeds *other* than c, the distance from the source to the detector in our experiment is relatively large (compared to an atom), so the only surviving contribution to P(A to B)'s length that counts comes from speed c.

To begin our new calculation of partial reflection, let's start by defining the event completely: the detector at A makes a click at a certain *time*, T. Then, let's divide the layer of glass into a number of very thin sections—let's say, six (see Fig. 68a). From the analysis we did in the second lecture in which we found that nearly all the light is reflected from the middle of a mirror, we know that although each electron is scattering light in all directions, when all the arrows for each section are added, the only place where they *don't* cancel out is where light goes straight down to the middle of the section and scatters in one of two directions—straight back up to the detector or straight down through the glass. The final arrow for the event will thus be determined by adding the six arrows representing the scattering of light from the six middle points—X_1 to X_6—arranged vertically throughout the glass.

All right, let's calculate the arrow for each of these ways the light could go—via the six points, X_1 to X_6. There are four steps involved in each way (which means four arrows will be multiplied):

—STEP #1: A photon is emitted by the source at a certain time.
—STEP #2: The photon goes from the source to one of the points in the glass.

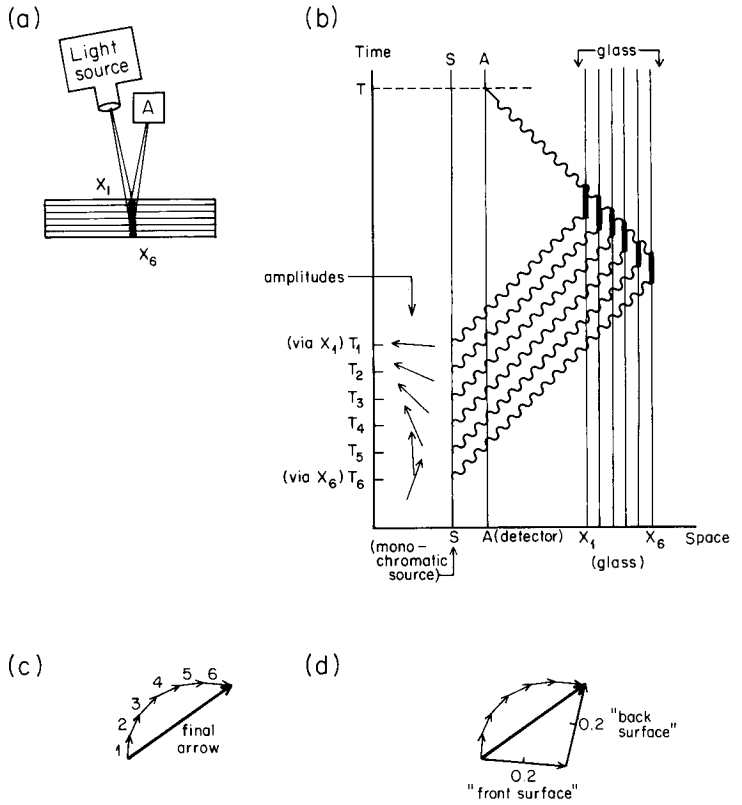


FIGURE 68. *We begin our new analysis of partial reflection by dividing a layer of glass into a number of sections (here, six), and looking at the various ways the light could go from the source to the glass and back up to the detector at A. The only important points in the glass (where the amplitudes for scattering light don't cancel out) are located at the middle of each section; X_1 to X_6 are shown in (a) at their physical location inside the glass, and in (b) as vertical lines on the space-time graph. The event whose probability we are calculating is: the detector at A makes a click at a certain time, T. Thus the event appears as a point (where A and T intersect) on the space-time graph.*

For each of the ways the event can happen, four steps must occur in succession, so four arrows have to be multiplied. The steps are shown in (b): 1) a photon leaves the source at a certain time (the arrows at T_1 to T_6 represent the amplitude to do that for six different times); 2) the photon goes from the source to one of the points in the glass (the six alternatives are depicted as wavy lines going up to the right); 3) an electron at one of the points scatters a photon (shown as short, wide vertical lines); and 4) a new photon goes to the detector and arrives at the appointed time, T (shown as a wavy line going up to the left). The amplitudes for steps 2, 3, and 4 are the same for the six alternatives, while the amplitudes for step 1 are different: compared to a photon scattered by an electron at the top of the glass (at X_1), a photon scattered deeper in the glass—at X_2, for example—must leave the source earlier, *at T_2.*

When we are finished multiplying the four arrows for each alternative, the resulting arrows, shown in (c), are shorter than those in (b); each has been turned 90° (in accordance with the scattering characteristics of electrons in glass). When these six arrows are added together in order, they form an arc; the final arrow is its chord. The same final arrow can be obtained by drawing two radius arrows, shown in (d), and "subtracting" them (turning the "front surface" arrow around in the opposite direction and adding it to the "back surface" arrow). This shortcut was used as a simplification in the first lecture.

—STEP #3: The photon is scattered by an electron at that point.

—STEP #4: A new photon makes its way up to the detector.

We will say the amplitudes for steps 2 and 4 (a photon goes to or from a point in the glass) involve no shrinking or turning, because we can assume that none of the light gets lost or spread out between the source and the glass or between the glass and the detector. For step 3 (an electron scatters a photon) the amplitude for scattering is a constant—a shrink and a turn by a certain amount, S—and is the same everywhere in the glass. (This amount is, as I mentioned before, different for different materials. For glass, the turn of S is 90°.) Therefore, of the four arrows to be multiplied, only the arrow for step 1—the amplitude for a photon to be emitted from the source at a certain time—is different from one alternative to the next.

The time at which a photon would have to have been emitted to reach the detector A at time T (see Fig. 68b) is

not the same for the six different paths. A photon scatterd by X_2 would have to have been emitted slightly *earlier* than a photon scattered by X_1, because that path is longer. Thus the arrow at T_2 is turned slightly more than the arrow at T_1 because the amplitude for a monochromatic source to emit a photon at a certain time rotates counterclockwise as time goes on. The same goes for each arrow down to T_6: all six arrows have the same length, but they are turned at different angles—that is, they are pointing in different directions—because they represent a photon emitted by the source at different times.

After shrinking the arrow at T_1 by the amounts prescribed in steps 2, 3 and 4—and turning it the 90° prescribed in step 3—we end up with arrow 1 (see Fig. 68c). The same goes for the arrows 2 through 6. Thus arrows 1 through 6 are all the same (shortened) length, and are turned relative to each other in exactly the same amount as the arrows at T_1 through T_6.

Next, we add arrows 1 to 6. Connecting the arrows in order from 1 to 6, we get something like an arc, or part of a circle. The final arrow forms the chord of this arc. The length of the final arrow increases with the thickness of the glass—thicker glass means more sections, more arrows, and therefore more of a circle—until half a circle is reached (and the final arrow is its diameter). Then the length of the final arrow *decreases* as the thickness of the glass continues to increase, and the circle becomes complete to begin a new cycle. The square of this length is the probability of the event, and it varies in the cycle of zero to 16%.

There is a mathematical trick we can use to get the same answer (see Fig. 68d): If we draw arrows from the center of the "circle" to the tail of arrow 1 and to the head of arrow 6, we get two radii. If the radius arrow from the center to arrow 1 is turned 180° ("subtracted"), then it can be combined with the other radius arrow to give us the

same final arrow! That's what I was doing in the first lecture: these two radii are the two arrows I said represented the "front surface" and "back surface" reflections. They each have the famous length of 0.2.[10]

Thus we can get the correct answer for the probability of partial reflection by imagining (falsely) that all reflection comes from only the front and back surfaces. In this intuitively easy analysis, the "front surface" and "back surface" arrows are mathematical constructions that give us the right answer, whereas the analysis we just did—with the space-time drawing and the arrows forming part of a circle—is a more accurate representation of what is really going on: partial reflection is the scattering of light by electrons *inside* the glass.

Now, what about the light that goes *through* the layer of glass? First, there is an amplitude that the photon goes straight through the glass without hitting any electrons (see Fig. 69a). This is the most important arrow in terms of length. But there are six other ways a photon could reach the detector below the glass: a photon could hit X_1 and scatter the new photon down to B; a photon could hit X_2 and scatter the new photon down to B, and so on. These six arrows all have the same length as the arrows that formed the "circle" in the previous example: their length

[10] The radius of the arc evidently depends on the *length* of the arrow for each section, which is ultimately determined by the amplitude S that an electron in an atom of glass scatters a photon. This radius can be calculated using the formulas for the three basic actions for the multitude of photon exchanges involved and summing up the amplitudes. It is a very difficult problem, but the radius has been calculated for relatively simple substances with considerable success, and the variation of the radius from substance to substance is fairly well understood using these ideas of quantum electrodynamics. It must be said, however, that no direct calculation from first principles for a substance as complex as glass has ever actually been done. In such cases, the radius is determined by experiment. For glass, it has been determined from experiment that the radius is approximately 0.2 (when the light shines directly onto the glass at right angles).

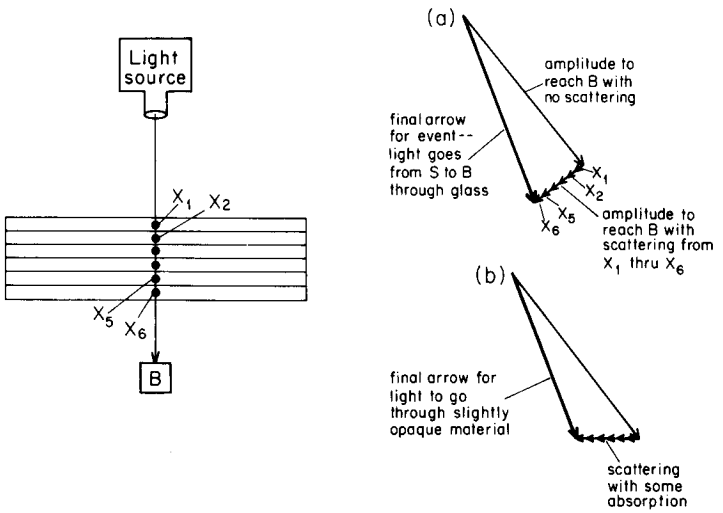


FIGURE 69. *The largest amplitude for light that is transmitted through the layer of glass to the detector at B comes from the part that represents no scattering by the electrons inside the glass, shown in (a). To this arrow we add six small arrows that represent the scattering of light from each of the sections, represented by points X_1 to X_6. These six arrows have the same length (because the amplitude for scattering is the same anywhere in the glass) and point in the same direction (because the length of each path from the source through any point X to B is the same). After adding the small arrows to the large one, we find the final arrow for the transmission of light through a layer of glass is turned more than what we would have expected if the light came only directly. For this reason it appears to us that light takes longer to go through glass than it takes to go through a vacuum or through air. The amount of turning by the final arrow caused by the electrons in a material is called the "index of refraction."*

For transparent materials, the little arrows are at right angles to the main arrow (they actually curve around when we include double and triple scatterings, keeping the final arrow from being longer than the main arrow: Nature always has it worked out so we never get more light out than we put in). For materials that are partially opaque—that absorb light to an extent—the little arrows point toward the main arrow, resulting in a final arrow that is significantly shorter than expected, shown in (b). This shorter final arrow represents a reduced probability of a photon being transmitted through partially opaque material.

is based on that same amplitude of an electron in the glass to scatter a photon, S. But this time, all six arrows point in the same direction, because the length of all six paths that involve one scattering is the same. The direction of these minor arrows is at right angles to the main arrow for transparent substances such as glass. When the minor arrows are added to the main arrow, they result in a final arrow that has the same length as the main arrow, but is turned in a slightly different direction. The thicker the glass, the more minor arrows there are, and the more the final arrow is turned. That's how a focusing lens really works: the final arrows for all the paths can be made to point in the same direction by inserting extra thicknesses of glass into the shorter paths.

The same effect would appear if photons went slower through glass than through air: there would be extra turning of the final arrow. That's why I said earlier that light appears to go slower through glass (or water) than through air. In reality, the "slowing" of the light is extra turning caused by the atoms in the glass (or water) scattering the light. The degree to which there is extra turning of the final arrow as light goes through a given material is called its "index of refraction."[11]

For substances that absorb light, the minor arrows are at

[11] Each of the arrows for reflection by a section (that form a "circle") has the same length as each of the arrows that make the final arrow from transmission appear to turn more. Thus there is a relationship between the partial reflection of a material and its index of refraction.

It appears that the final arrow has become longer than 1, which means that more light comes out through the glass than went into it! It looks that way because I disregarded the amplitudes for a photon to go down to one section, a new photon to scatter up to another section, and then a third photon to scatter back down through the glass—and other, more complicated possibilities—which result in the little arrows curving around and keeping the length of the final arrow between 0.92 and 1 (so the total probability of light being reflected or transmitted by the layer of glass is always 100%).

less than right angles to the main arrow (see Fig. 69b). This causes the final arrow to be shorter than the main arrow, indicating that the probability of a photon going through partially opaque glass is smaller than through transparent glass.

Thus it is that all the phenomena and the arbitrary numbers mentioned in the first two lectures—such as partial reflection with an amplitude of 0.2, the "slowing" of light in water and glass, and so on—are explained in more detail by just the three basic actions—three actions that do, in fact, explain nearly everything else, too.

It is hard to believe that nearly all the vast apparent variety in Nature results from the monotony of repeatedly combining just these three basic actions. But it does. I'll outline a bit of how some of this variety arises.

We may start with photons (see Fig. 70). What is the probability that two photons, at points 1 and 2 in space-time, go to two detectors, at points 3 and 4? There are two main ways this event could happen and each depends on two things happening concomitantly: the photons could go directly—P(1 to 3)*P(2 to 4)—or they could "cross over"—P(1 to 4)*P(2 to 3). The resulting amplitudes for these two possibilities are added, and there is interference (as we saw in the second lecture), making the final arrow vary in length, depending on the relative location of the points in space-time.

What if we make 3 and 4 the same point in space-time (see Fig. 71)? Let's say both photons end up at point 3, and see how this affects the probability of the event. Now we have P(1 to 3)*P(2 to 3) and P(2 to 3)*P(1 to 3), which result in two identical arrows. When added, their sum is twice the length of either one, and produces a final arrow whose square is four times the square of either arrow alone. Because the two arrows are identical, they are always "lined up." In other words, the interference doesn't fluctuate according to the relative separation between points 1 and 2; it is always positive. If we didn't think about the always positive interference of the two photons, we should have thought that we would get twice the probability, on average. Instead, we get four times the probability all the time. When

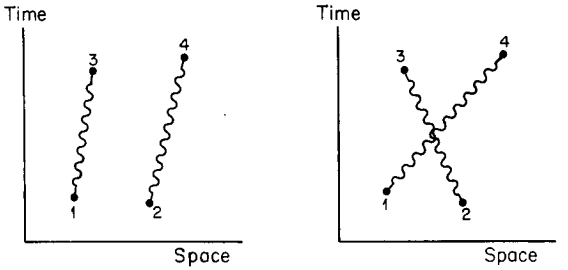


FIGURE 70. *Photons at points 1 and 2 in space-time have an amplitude to arrive at points 3 and 4 in space-time that is approximated by considering two main ways the event could happen: P(1 to 3) * P(2 to 4) and P(1 to 4) * P(2 to 3), shown above. Depending on the relative locations of points 1, 2, 3, and 4, there are varying degrees of interference.*

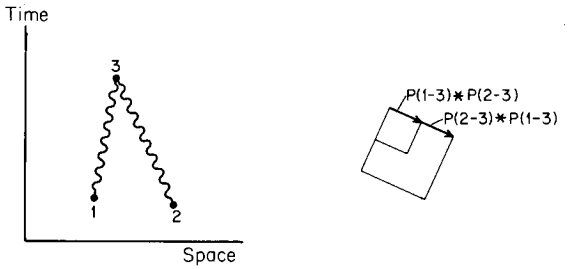


FIGURE 71. *When points 4 and 3 are made to converge, the two arrows—P(1 to 3) * P(2 to 3) and P(2 to 3) * P(1 to 3)—are identical in length and direction. When they are added they always "line up" and form an arrow with twice the length of either arrow alone, with a square four times as large. Thus photons tend to go to the same point in space-time. This effect is magnified even more by more photons. This is the basis of a laser's operation.*

many photons are involved, this more-than-expected probability increases even further.

This results in a number of practical effects. We can say that photons tend to get into the same condition, or "state" (the way the amplitude to find one varies in space). The chance that an atom emits a photon is enhanced if some photons (in a state that the atom can emit into) are already present. This phenomenon of "stimulated emission" was discovered by Einstein when he launched the quantum theory proposing the photon model of light. Lasers work on the basis of this phenomenon.

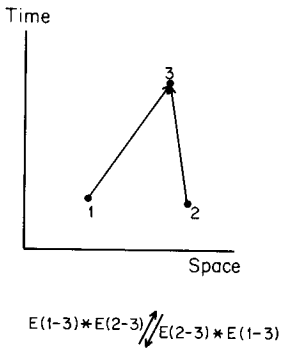


FIGURE 72. *If two electrons (with the same polarization) try to go to the same point in space-time, the interference is always negative because of the effects of polarization: the two identical arrows—E(1 to 3) * E(2 to 3) and E(2 to 3) * E(1 to 3)—are subtracted to make a final arrow of no length. The aversion of two electrons to occupy the same place in space-time is called the "Exclusion Principle," and accounts for the great variety of atoms in the universe.*

If we made the same comparison with our fake, spin-zero electrons, the same thing would happen. But in the real world, where electrons are polarized, something very different happens: the two arrows, E(1 to 3) * E(2 to 4) and E(1 to 4) * E(2 to 3), are subtracted—one of them is turned 180° before they are added. When points 3 and 4 are the same, the two arrows have the same length and direction and thus cancel out when they are subtracted (see Fig. 72). That means electrons, unlike photons, do not like to go to the same place; they avoid each other like the plague—no

two electrons with the same polarization can be at the same point in space-time—it's called the "exclusion principle."

This exclusion principle turns out to be the origin of the great variety of chemical properties of the atoms. One proton exchanging photons with one electron dancing around it is called a hydrogen atom. Two protons in the same nucleus exchanging photons with two electrons (polarized in opposite directions) is called a helium atom. You see, the chemists have a complicated way of counting: instead of saying "one, two, three, four, five protons," they say, "hydrogen, helium, lithium, beryllium, boron."

There are only two states of polarization available to electrons, so in an atom with three protons in the nucleus exchanging photons with three electrons—a condition called a lithium atom—the third electron is farther away from the nucleus than the other two (which have used up the nearest available space), and exchanges fewer photons. This causes the electron to easily break away from its own nucleus under the influence of photons from other atoms. A large number of such atoms close together easily lose their individual third electrons to form a sea of electrons swimming around from atom to atom. This sea of electrons reacts to any small electrical force (photons), generating a current of electrons—I am describing lithium metal conducting electricity. Hydrogen and helium atoms do not lose their electrons to other atoms. They are "insulators."

All the atoms—more than one hundred different kinds—are made up of a certain number of protons exchanging photons with the same number of electrons. The patterns in which they gather are complicated and offer an enormous variety of properties: some are metals, some are insulators, some are gases, others are crystals; there are soft things, hard things, colored things, and transparent things—a terrific cornucopia of variety and excitement that comes from the exclusion principle and the repetition again

and again and again of the three very simple actions P(A to B), E(A to B), and *j*. (If the electrons in the world were unpolarized, all the atoms would have very similar properties: the electrons would all cluster together, close to the nucleus of their own atom, and would not be easily attracted to other atoms to make chemical reactions.)

You might wonder how such simple actions could produce such a complex world. It's because phenomena we see in the world are the result of an enormous intertwining of tremendous numbers of photon exchanges and interferences. Knowing the three fundamental actions is only a very small beginning toward analyzing any *real* situation, where there is such a multitude of photon exchanges going on that it is impossible to calculate—experience has to be gained as to which possibilities are more important. Thus we invent such ideas as "index of refraction" or "compressibility" or "valence" to help us calculate in an approximate way when there's an enormous amount of detail going on underneath. It's analogous to knowing the rules of chess—which are fundamental and simple—compared to being able to play chess well, which involves understanding the character of each position and the nature of various situations—which is much more advanced and difficult.

The branches of physics that deal with questions such as why iron (with 26 protons) is magnetic, while copper (with 29) is not, or why one gas is transparent and another one is not, are called "solid-state physics," or "liquid-state physics," or "honest physics." The branch of physics that found these three simple little actions (the easiest part) is called "fundamental physics"—we stole that name in order to make the other physicists feel uncomfortable! The most interesting problems today—and certainly the most practical problems—are obviously in solid-state physics. But someone said there is nothing so practical as a good theory, and the theory of quantum electrodynamics is definitely a good theory!

Finally, I would like to return to that number 1.00115965221, the number that I told you about in the first lecture that has been measured and calculated so carefully. The number represents the response of an electron to an external magnetic field—something called the "magnetic moment." When Dirac first worked out the rules to calculate this number, he used the formula for E(A to B) and got a very simple answer, which we will consider in our units as 1. The diagram for this first approximation of the magnetic moment of an electron is very simple—an electron goes from place to place in space-time and couples with a photon from a magnet (see Fig. 73).

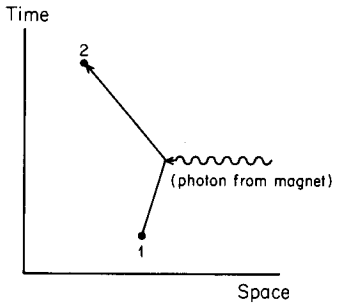


FIGURE 73. *The diagram for Dirac's calculation of the magnetic moment of an electron is very simple. The value represented by this diagram will be called 1.*

After some years it was discovered that this value was not exactly 1, but slightly more—something like 1.00116. This correction was worked out for the first time in 1948 by Schwinger as *j*j* divided by 2 pi, and was due to an alternative way the electron can go from place to place: instead of going directly from one point to another, the electron goes along for a while and suddenly emits a photon; then (horrors!) it absorbs its own photon (see Fig. 74). Perhaps there's something "immoral" about that, but the

electron does it! To calculate the arrow for this alternative, we have to make an arrow for every place in space-time that the photon can be emitted and every place it can be absorbed. Thus there will be two extra E(A to B)'s, a P(A to B) and two extra *j*'s, all multiplied together. Students learn how to do this simple calculation in their elementary quantum electrodynamics course, in their second year of graduate school.

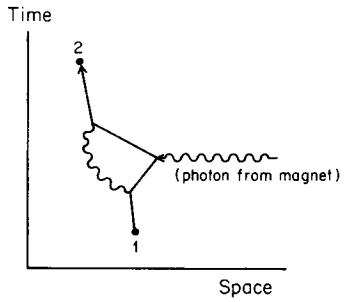


FIGURE 74. *Laboratory experiments show that the actual value of the magnetic moment of an electron is not 1, but a little bit more. This is because there are alternatives: the electron can emit a photon and then absorb it—requiring two extra E(A to B)'s, a P(A to B), and two extra j's. Schwinger calculated the adjustment that takes this alternative into account to be j*j divided by 2 pi. Since this alternative is indistinguishable experimentally from the original way the electron can go—an electron starts at point 1 and ends up at point 2—the arrows for the two alternatives are added, and there is interference.*

But wait: experiments have measured the behavior of an electron so accurately that we have to consider still other possibilities in our calculations—all the ways the electron can go from place to place with *four* extra couplings (see Fig. 75). There are three ways the electron can emit and absorb two photons. There's also a new, interesting possibility (shown at the right of Fig. 75): one photon is emitted; it makes a positron-electron pair, and—again, if you'll hold

your "moral" objections—the electron and positron annihilate, creating a new photon that is ultimately absorbed by the electron. That possibility also has to be figured in!

It took two "independent" groups of physicists two years to calculate this next term, and then another year to find

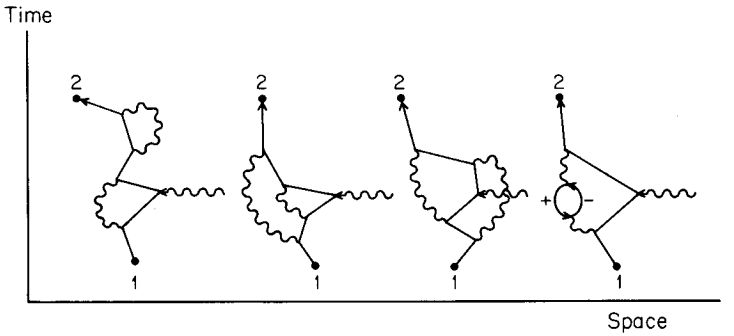


FIGURE 75. *Laboratory experiments became so accurate that further alternatives, involving four extra couplings (over all possible intermediate points in space-time), had to be calculated, some of which are shown here. The alternative on the right involves a photon disintegrating into a positron-electron pair (as described in Fig. 64), which annihilates to form a new photon, which is ultimately absorbed by the electron.*

out there was a mistake—experimenters had measured the value to be slightly different, and it looked for awhile that the theory didn't agree with experiment for the first time, but no: it was a mistake in arithmetic. How could two groups make the same mistake? It turns out that near the end of the calculation the two groups compared notes and ironed out the differences between their calculations, so they were not really independent.

The term with *six* extra *j*'s involves even more possible ways the event can happen, and I'll draw a few of them for you now (see Fig. 76). It took twenty years to get this extra accuracy figured into the theoretical value of the magnetic

moment of an electron. Meanwhile the experimenters made even more detailed experiments and added a few more digits onto their number—and the theory still agreed with it.

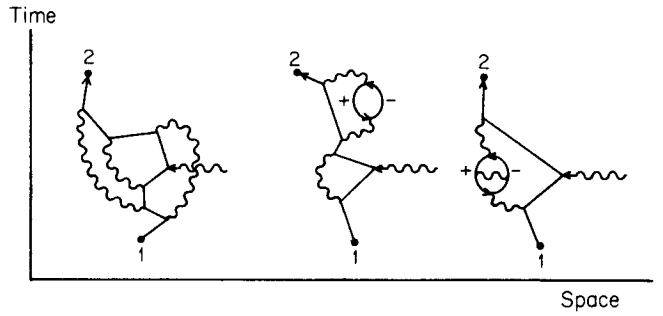


FIGURE 76. *Calculations are presently going on to make the theoretical value even more accurate. The next contribution to the amplitude, which represents all possibilities with six extra couplings, involves something like 70 diagrams, three of which are shown here. As of 1983, the theoretical number was 1.00115965246, with an uncertainty of about 20 in the last two digits; the experimental number was 1.00115965221, with an uncertainty of about 4 in the last digit. This accuracy is equivalent to measuring the distance from Los Angeles to New York, a distance of over 3,000 miles, to within the width of a human hair.*

So, to make our calculations we make these diagrams, write down what they correspond to mathematically, and add the amplitudes—a straightforward, "cookbook" process. Therefore, it can be done by machines. Now that we have super-duper computers, we have begun to compute the term with eight extra *j*'s. At the present time the theoretical number is 1.00115965246; experimentally, it's 1.00115965221, plus or minus 4 in the last decimal place. Some of the uncertainty in the theoretical value (about 4 in the last decimal place) is due to the computer's rounding off numbers; most of it (about 20) is due to the fact that

the value for *j* is not exactly known. The term for eight extra *j*'s involves something like nine hundred diagrams, with a hundred thousand terms each—a fantastic calculation—and it's being done right now.

I am sure that in a few more years, the theoretical and experimental numbers for the magnetic moment of an electron will be worked out to still more places. Of course, I am not sure whether the two values will still agree. That, one can never tell until one makes the calculation and does the experiments.

And so we have come full circle to the number I chose to "intimidate" you with at the beginning of these lectures. I hope you understand the significance of this number much better now: it represents the extraordinary degree to which we've been constantly checking that the strange theory of quantum electrodynamics is indeed correct.

Throughout these lectures I have delighted in showing you that the price of gaining such an accurate theory has been the erosion of our common sense. We must accept some very bizarre behavior: the amplification and suppression of probabilities, light reflecting from all parts of a mirror, light travelling in paths other than a straight line, photons going faster or slower than the conventional speed of light, electrons going backwards in time, photons suddenly disintegrating into a positron-electron pair, and so on. That we must do, in order to appreciate what Nature is really doing underneath nearly all the phenomena we see in the world.

With the exception of technical details of polarization, I have described to you the framework by which we understand all these phenomena. We draw *amplitudes* for every way an event can happen and add them when we would have expected to add probabilities under ordinary circumstances; we multiply amplitudes when we would have expected to multiply probabilities. Thinking of everything in

terms of amplitudes may cause difficulties at first because of their abstraction, but after a while, one gets used to this strange language. Underneath so many of the phenomena we see every day are only three basic actions: one is described by the simple coupling number, j; the other two by functions—P(A to B) and E(A to B)—both of which are closely related. That's all there is to it, and from it all the rest of the laws of physics come.

However, before I finish this lecture, I would like to make a few additional remarks. One can understand the spirit and character of quantum electrodynamics without including this technical detail of polarization. But I'm sure you'll all feel uncomfortable unless I say something about what I've been leaving out. Photons, it turns out, come in four different varieties, called polarizations, that are related geometrically to the directions of space and time. Thus there are photons polarized in the X, Y, Z, and T directions. (Perhaps you have heard somewhere that light comes in only two states of polarization—for example, a photon going in the Z direction can be polarized at right angles, either in the X or Y direction. Well, you guessed it: in situations where the photon goes a long distance and appears to go at the speed of light, the amplitudes for the Z and T terms exactly cancel out. But for virtual photons going between a proton and an electron in an atom, it is the T component that is the most important.)

In a similar manner, an electron can be in one of four conditions that are also related to geometry, but in a somewhat more subtle manner. We can call these conditions 1, 2, 3, and 4. Calculating the amplitude for an electron going from point A to point B in space-time becomes somewhat more complicated, because we can now ask questions such as, "What is the amplitude that an electron liberated in condition 2 at the point A arrives in condition 3 at the point B?" The sixteen possible combinations—coming from the

four different conditions an electron can start in at A and the four different conditions it can end up in at B—are related in a simple mathematical way to the formula for that E(A to B) I told you about.

For a photon, no such modification is necessary. Thus a photon polarized in the X direction at A will still be polarized in the X direction at B, arriving with the amplitude P(A to B).

Polarization produces a large number of different possible couplings. We could ask, for example, "What is the amplitude that an electron in condition 2 absorbs a photon polarized in the X direction and thereby turns into an electron in condition 3?" All the possible combinations of polarized electrons and photons do not couple, but those that do, do so with the same amplitude j, but sometimes with an additional turn of the arrow by some multiple of 90°.

These possibilities for the different kinds of polarization and the nature of the couplings can all be deduced in a very elegant and beautiful manner from the principles of quantum electrodynamics and two further assumptions: 1) the results of an experiment are not affected if the apparatus with which you are making experiments is turned in some other direction, and 2) it also doesn't make any difference if the apparatus is in a spaceship moving at some arbitrary speed. (This is the principle of relativity.)

This elegant and general analysis shows that every particle must be in one or another class of possible polarizations, which we call spin 0, spin 1/2, spin 1, spin 3/2, spin 2, and so on. The different classes behave in different ways. A spin 0 particle is the simplest—it has just one component, and is not effectively polarized at all. (The fake electrons and photons that we have been considering in this lecture are spin 0 particles. So far, no fundamental spin 0 particles have been found.) A real electron is an example of a spin 1/2 particle, and a real photon is an example of a spin 1

particle. Both spin 1/2 and spin 1 particles have four components. The other types would have more components, such as spin 2 particles, with ten components.

I said that the connection between relativity and polarization is simple and elegant, but I'm not sure I can explain it simply and elegantly! (It would take me at least one additional lecture to do it.) Although the details of polarization are not essential to understanding the spirit and character of quantum electrodynamics, they are, of course, essential to the correct calculation of any real process, and often have profound effects.

In these lectures we have been concentrating on relatively simple interactions between electrons and photons at very small distances, in which only a few particles are involved. But I would like to make one or two remarks about how these interactions appear in the larger world, where very, very large numbers of photons are being exchanged. On such a large scale, the calculation of arrows gets very complicated.

There are, however, some situations that are not so difficult to analyze. There are circumstances, for example, where the amplitude to emit a photon by a source is independent of whether another photon has been emitted. This can happen when the source is very heavy (the nucleus of an atom), or when a very large number of electrons are all moving the same way, such as up and down in the antenna of a broadcasting station or going around in the coils of an electromagnet. Under such circumstances a large number of photons are emitted, all of exactly the same kind. The amplitude of an electron to absorb a photon in such an environment is independent of whether it or any other electron has absorbed other photons before. Therefore its entire behavior can be given by just this amplitude for an electron to absorb a photon, which depends only on the electron's position in space and time. Physicists use ordinary words to describe this circumstance. They say the electron is moving in an external field.

Physicists use the word "field" to describe a quantity that depends on position in space and time. Temperatures in the air provide a good example: they vary according to where and when you make your measurements. When we take polarization into account, there are more components to the field. (There are four components—corresponding to the amplitude to absorb each of the different kinds of polarization (X, Y, Z, T) the photon might be in—technically called the vector and scalar electromagnetic potentials. From combinations of these, classical physics derives more convenient components called the electric and magnetic fields.)

In a situation where the electric and magnetic fields are varying slowly enough, the amplitude for an electron to travel over a very long distance depends on the path it takes. As we saw earlier in the case of light, the most important paths are the ones where the angles of the amplitudes from nearby paths are nearly the same. The result is that the particle doesn't necessarily go in a straight line.

This brings us all the way back to classical physics, which supposes that there are fields and that electrons move through them in such a way as to make a certain quantity least. (Physicists call this quantity "action" and formulate this rule as the "principle of least action.") This is one example of how the rules of quantum electrodynamics produce phenomena on a large scale. We could expand in many directions from here, but we have to limit the scope of these lectures somewhere. I just wanted to remind you that the effects that we see on a large scale and the strange phenomena we see on a small scale are both produced by the interaction of electrons and photons, and are all described, ultimately, by the theory of quantum electrodynamics.