1

ATOMS IN MOTION

Introduction

This two-year course in physics is presented from the point of view that you, the reader, are going to be a physicist. This is not necessarily the case of course, but that is what every professor in every subject assumes! If you are going to be a physicist, you will have a lot to study: two hundred years of the most rapidly developing field of knowledge that there is. So much knowledge, in fact, that you might think that you cannot learn all of it in four years, and truly you cannot; you will have to go to graduate school too!

Surprisingly enough, in spite of the tremendous amount of work that has been done for all this time it is possible to condense the enormous mass of results to a large extent—that is, to find *laws* which summarize all our knowledge. Even so, the laws are so hard to grasp that it is unfair to you to start exploring this tremendous subject without some kind of map or outline of the relationship of one part of the subject of science to another. Following these preliminary remarks, the first three chapters will therefore outline the relation of physics to the rest of the sciences, the relations of the sciences to each other, and the meaning of science, to help us develop a "feel" for the subject.

You might ask why we cannot teach physics by just giving the basic laws on page one and then showing how they work in all possible circumstances, as we do in Euclidean geometry, where we

state the axioms and then make all sorts of deductions. (So, not satisfied to learn physics in four years, you want to learn it in four minutes?) We cannot do it in this way for two reasons. First, we do not yet *know* all the basic laws: there is an expanding frontier of ignorance. Second, the correct statement of the laws of physics involves some very unfamiliar ideas which require advanced mathematics for their description. Therefore, one needs a considerable amount of preparatory training even to learn what the *words* mean. No, it is not possible to do it that way. We can only do it piece by piece.

Each piece, or part, of the whole of nature is always merely an *approximation* to the complete truth, or the complete truth so far as we know it. In fact, everything we know is only some kind of approximation, because *we know that we do not know all the laws* as yet. Therefore, things must be learned only to be unlearned again or, more likely, to be corrected.

The principle of science, the definition, almost, is the following: *The test of all knowledge is experiment.* Experiment is the *sole judge* of scientific "truth." But what is the source of knowledge? Where do the laws that are to be tested come from? Experiment, itself, helps to produce these laws, in the sense that it gives us hints. But also needed is *imagination* to create from these hints the great generalizations—to guess at the wonderful, simple, but very strange patterns beneath them all, and then to experiment to check again whether we have made the right guess. This imagining process is so difficult that there is a division of labor in physics: there are *theoretical* physicists who imagine, deduce, and guess at new laws, but do not experiment; and then there are *experimental* physicists who experiment, imagine, deduce, and guess.

We said that the laws of nature are approximate: that we first find the "wrong" ones, and then we find the "right" ones. Now, how can an experiment be "wrong"? First, in a trivial way: if something is wrong with the apparatus that you did not notice. But these things are easily fixed, and checked back and forth. So without

snatching at such minor things, how *can* the results of an experiment be wrong? Only by being inaccurate. For example, the mass of an object never seems to change: a spinning top has the same weight as a still one. So a "law" was invented: mass is constant, independent of speed. That "law" is now found to be incorrect. Mass is found to increase with velocity, but appreciable increases require velocities near that of light. A *true* law is: if an object moves with a speed of less than one hundred miles a second the mass is constant to within one part in a million. In some such approximate form this is a correct law. So in practice one might think that the new law makes no significant difference. Well, yes and no. For ordinary speeds we can certainly forget it and use the simple constant-mass law as a good approximation. But for high speeds we are wrong, and the higher the speed, the more wrong we are.

Finally, and most interesting, *philosophically we are completely wrong* with the approximate law. Our entire picture of the world has to be altered even though the mass changes only by a little bit. This is a very peculiar thing about the philosophy, or the ideas, behind the laws. Even a very small effect sometimes requires profound changes in our ideas.

Now, what should we teach first? Should we teach the *correct* but unfamiliar law with its strange and difficult conceptual ideas, for example the theory of relativity, four-dimensional space-time, and so on? Or should we first teach the simple "constant-mass" law, which is only approximate, but does not involve such difficult ideas? The first is more exciting, more wonderful, and more fun, but the second is easier to get at first, and is a first step to a real understanding of the first idea. This point arises again and again in teaching physics. At different times we shall have to resolve it in different ways, but at each stage it is worth learning what is now known, how accurate it is, how it fits into everything else, and how it may be changed when we learn more.

Let us now proceed with our outline, or general map, of our understanding of science today (in particular, physics, but also of other

sciences on the periphery), so that when we later concentrate on some particular point we will have some idea of the background, why that particular point is interesting, and how it fits into the big structure. So, what *is* our overall picture of the world?

Matter is made of atoms

If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, what statement would contain the most information in the fewest words? I believe it is the *atomic hypothesis* (or the atomic *fact,* or whatever you wish to call it) that *all things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another.* In that one sentence, you will see, there is an *enormous* amount of information about the world, if just a little imagination and thinking are applied.

To illustrate the power of the atomic idea, suppose that we have a drop of water a quarter of an inch on the side. If we look at it very closely we see nothing but water—smooth, continuous water. Even if we magnify it with the best optical microscope available—roughly two thousand times—then the water drop will be roughly forty feet across, about as big as a large room, and if we looked rather closely, we would *still* see relatively smooth water—but here and there small football-shaped things swimming back and forth. Very interesting. These are paramecia. You may stop at this point and get so curious about the paramecia with their wiggling cilia and twisting bodies that you go no further, except perhaps to magnify the paramecia still more and see inside. This, of course, is a subject for biology, but for the present we pass on and look still more closely at the water material itself, magnifying it two thousand times again. Now the drop of water extends about fifteen miles across, and if we look very closely at it we see a kind of teeming, something which no longer has a smooth appearance—it looks something like a crowd at a football game as seen from a very great

distance. In order to see what this teeming is about, we will magnify it another two hundred and fifty times and we will see something similar to what is shown in Fig. 1-1. This is a picture of water magnified a billion times, but idealized in several ways. In the first place, the particles are drawn in a simple manner with sharp edges, which is inaccurate. Secondly, for simplicity, they are sketched almost schematically in a two-dimensional arrangement, but of course they are moving around in three dimensions. Notice that there are two kinds of "blobs" or circles to represent the atoms of oxygen (black) and hydrogen (white), and that each oxygen has two hydrogens tied to it. (Each little group of an oxygen with its two hydrogens is called a molecule.) The picture is idealized further in that the real particles in nature are continually jiggling and bouncing, turning and twisting around one another. You will have to imagine this as a dynamic rather than a static picture. Another thing that cannot be illustrated in a drawing is the fact that the particles are "stuck together"—that they attract each other, this one pulled by that one, etc. The whole group is "glued together," so to speak. On the other hand, the particles do not squeeze through each other. If you try to squeeze two of them too close together, they repel.

The atoms are 1 or 2×10^{-8} cm in radius. Now 10^{-8} cm is called an *angstrom* (just as another name), so we say they are 1 or 2 angstroms (Å) in radius. Another way to remember their size is this: if an apple is magnified to the size of the earth, then the atoms in the apple are approximately the size of the original apple.

Now imagine this great drop of water with all of these jiggling particles stuck together and tagging along with each other. The water keeps its volume; it does not fall apart, because of the attraction of the molecules for each other. If the drop is on a slope, where it can move from one place to another, the water will flow, but it does not just disappear—things do not just fly apart—because of the molecular attraction. Now the jiggling motion is what we represent as *heat:* when we increase the temperature, we increase the motion. If we heat the water, the jiggling increases and the volume between the atoms increases, and if the heating continues there

WATER MAGNIFIED ONE BILLION TIMES

Figure 1-1

comes a time when the pull between the molecules is not enough to hold them together and they *do* fly apart and become separated from one another. Of course, this is how we manufacture steam out of water—by increasing the temperature; the particles fly apart because of the increased motion.

In Fig. 1-2 we have a picture of steam. This picture of steam fails in one respect: at ordinary atmospheric pressure there certainly would not be as many as three in this figure. Most squares this size would contain none—but we accidentally have two and a half or three in the picture (just so it would not be completely blank). Now in the case of steam we see the characteristic molecules more clearly than in the case of water. For simplicity, the molecules are drawn so that there is a 120° angle between the hydrogen atoms. In actual fact the angle is 105°3', and the distance between the center of a hydrogen and the center of the oxygen is 0.957 Å, so we know this molecule very well.

Let us see what some of the properties of steam vapor or any other gas are. The molecules, being separated from one another, will bounce against the walls. Imagine a room with a number of tennis balls (a hundred or so) bouncing around in perpetual motion. When they bombard the wall, this pushes the wall away. (Of course we would have to push the wall back.) This means that the gas exerts a jittery force which our coarse senses (not being ourselves magnified a billion times) feel only as an *average push.* In order to confine a gas we must apply a pressure. Figure 1-3 shows a standard

Figure 1-2

vessel for holding gases (used in all textbooks), a cylinder with a piston in it. Now, it makes no difference what the shapes of water molecules are, so for simplicity we shall draw them as tennis balls or little dots. These things are in perpetual motion in all directions. So many of them are hitting the top piston all the time that to keep it from being patiently knocked out of the tank by this continuous banging, we shall have to hold the piston down by a certain force, which we call the *pressure* (really, the pressure times the area is the force). Clearly, the force is proportional to the area, for if we increase the area but keep the number of molecules per cubic centimeter the same, we increase the number of collisions with the piston in the same proportion as the area was increased.

Now let us put twice as many molecules in this tank, so as to double the density, and let them have the same speed, i.e., the same

Figure 1-3

temperature. Then, to a close approximation, the number of collisions will be doubled, and since each will be just as "energetic" as before, the pressure is proportional to the density. If we consider the true nature of the forces between the atoms, we would expect a slight decrease in pressure because of the attraction between the atoms, and a slight increase because of the finite volume they occupy. Nevertheless, to an excellent approximation, if the density is low enough that there are not many atoms, *the pressure is proportional to the density.*

We can also see something else: If we increase the temperature without changing the density of the gas, i.e., if we increase the speed of the atoms, what is going to happen to the pressure? Well, the atoms hit harder because they are moving faster, and in addition they hit more often, so the pressure increases. You see how simple the ideas of atomic theory are.

Let us consider another situation. Suppose that the piston moves inward, so that the atoms are slowly compressed into a smaller space. What happens when an atom hits the moving piston? Evidently it picks up speed from the collision. You can try it by bouncing a ping-pong ball from a forward-moving paddle, for example, and you will find that it comes off with more speed than that with which it struck. (Special example: if an atom happens to be standing still and the piston hits it, it will certainly move.) So the atoms are "hotter" when they come away from the piston than they were before they struck it. Therefore all the atoms which are in the vessel will have picked up speed. This means that *when we compress a gas slowly, the temperature of the gas increases.* So, under slow *compression,* a gas will *increase* in temperature, and under slow *expansion* it will *decrease* in temperature.

We now return to our drop of water and look in another direction. Suppose that we decrease the temperature of our drop of water. Suppose that the jiggling of the molecules of the atoms in the water is steadily decreasing. We know that there are forces of attraction between the atoms, so that after a while they will not be able to jiggle so well. What will happen at very low temperatures is

indicated in Fig. 1-4: the molecules lock into a new pattern which is *ice.* This particular schematic diagram of ice is wrong because it is in two dimensions, but it is right qualitatively. The interesting point is that the material has a *definite place for every atom,* and you can easily appreciate that if somehow or other we were to hold all the atoms at one end of the drop in a certain arrangement, each atom in a certain place, then because of the structure of interconnections, which is rigid, the other end miles away (at our magnified scale) will have a definite location. So if we hold a needle of ice at one end, the other end resists our pushing it aside, unlike the case of water, in which the structure is broken down because of the increased jiggling so that the atoms all move around in different ways. The difference between solids and liquids is, then, that in a solid the atoms are arranged in some kind of an array, called a *crystalline array,* and they do not have a random position at long distances; the position of the atoms on one side of the crystal is determined by that of other atoms millions of atoms away on the other side of the crystal. Figure 1-4 is an invented arrangement for ice, and although it contains many of the correct features of ice, it is not the true arrangement. One of the correct features is that there is a part of the symmetry that is hexagonal. You can see that if we turn the picture around an axis by 60°, the picture returns to itself. So there is a *symmetry* in the ice which accounts for the six-sided appearance of snowflakes. Another thing we can see from Fig. 1-4 is why ice

ICE

Figure 1-4

shrinks when it melts. The particular crystal pattern of ice shown here has many "holes" in it, as does the true ice structure. When the organization breaks down, these holes can be occupied by molecules. Most simple substances, with the exception of water and type metal, *expand* upon melting, because the atoms are closely packed in the solid crystal and upon melting need more room to jiggle around, but an open structure collapses, as in the case of water.

Now although ice has a "rigid" crystalline form, its temperature can change—ice has heat. If we wish, we can change the amount of heat. What is the heat in the case of ice? The atoms are not standing still. They are jiggling and vibrating. So even though there is a definite order to the crystal—a definite structure—all of the atoms are vibrating "in place." As we increase the temperature, they vibrate with greater and greater amplitude, until they shake themselves out of place. We call this *melting.* As we decrease the temperature, the vibration decreases and decreases until, at absolute zero, there is a minimum amount of vibration that the atoms can have, but *not zero.* This minimum amount of motion that atoms can have is not enough to melt a substance, with one exception: helium. Helium merely decreases the atomic motions as much as it can, but even at absolute zero there is still enough motion to keep it from freezing. Helium, even at absolute zero, does not freeze, unless the pressure is made so great as to make the atoms squash together. If we increase the pressure, we *can* make it solidify.

Atomic processes

So much for the description of solids, liquids, and gases from the atomic point of view. However, the atomic hypothesis also describes *processes,* and so we shall now look at a number of processes from an atomic standpoint. The first process that we shall look at is associated with the surface of the water. What happens at the surface of the water? We shall now make the picture more complicated—and more realistic—by imagining that the surface is in air. Figure 1-5 shows the surface of water in air. We see the water molecules

Figure 1-5

as before, forming a body of liquid water, but now we also see the surface of the water. Above the surface we find a number of things: First of all there are water molecules, as in steam. This is *water vapor,* which is always found above liquid water. (There is an equilibrium between the steam vapor and the water which will be described later.) In addition we find some other molecules—here two oxygen atoms stuck together by themselves, forming an *oxygen molecule,* there two nitrogen atoms also stuck together to make a nitrogen molecule. Air consists almost entirely of nitrogen, oxygen, some water vapor, and lesser amounts of carbon dioxide, argon, and other things. So above the water surface is the air, a gas, containing some water vapor. Now what is happening in this picture? The molecules in the water are always jiggling around. From time to time, one on the surface happens to be hit a little harder than usual, and gets knocked away. It is hard to see that happening in the picture because it is a *still* picture. But we can imagine that one molecule near the surface has just been hit and is flying out, or perhaps another one has been hit and is flying out. Thus, molecule by molecule, the water disappears—it evaporates. But if we *close* the vessel above, after a while we shall find a large number of molecules of water amongst the air molecules. From time to time, one of these vapor molecules comes flying down to the water and gets stuck again. So we see that what looks like a dead, uninteresting thing—a glass of water with a cover, that has been sitting there for perhaps

twenty years—really contains a dynamic and interesting phenomenon which is going on all the time. To our eyes, our crude eyes, nothing is changing, but if we could see it a billion times magnified, we would see that from its own point of view it is always changing: molecules are leaving the surface, molecules are coming back.

Why do *we* see *no change*? Because just as many molecules are leaving as are coming back! In the long run "nothing happens." If we then take the top of the vessel off and blow the moist air away, replacing it with dry air, then the number of molecules leaving is just the same as it was before, because this depends on the jiggling of the water, but the number coming back is greatly reduced because there are so many fewer water molecules above the water. Therefore there are more going out than coming in, and the water evaporates. Hence, if you wish to evaporate water turn on the fan!

Here is something else: Which molecules leave? When a molecule leaves it is due to an accidental, extra accumulation of a little bit more than ordinary energy, which it needs if it is to break away from the attractions of its neighbors. Therefore, since those that leave have more energy than the average, the ones that are left have *less* average motion than they had before. So the liquid gradually *cools* if it evaporates. Of course, when a molecule of vapor comes from the air to the water below there is a sudden great attraction as the molecule approaches the surface. This speeds up the incoming molecule and results in generation of heat. So when they leave they take away heat; when they come back they generate heat. Of course when there is no net evaporation the result is nothing—the water is not changing temperature. If we blow on the water so as to maintain a continuous preponderance in the number evaporating, then the water is cooled. Hence, blow on soup to cool it!

Of course you should realize that the processes just described are more complicated than we have indicated. Not only does the water go into the air, but also, from time to time, one of the oxygen or nitrogen molecules will come in and "get lost" in the mass of water molecules, and work its way into the water. Thus the air dissolves in the water; oxygen and nitrogen molecules will work their way

into the water and the water will contain air. If we suddenly take the air away from the vessel, then the air molecules will leave more rapidly than they come in, and in doing so will make bubbles. This is very bad for divers, as you may know.

Now we go on to another process. In Fig. 1-6 we see, from an atomic point of view, a solid dissolving in water. If we put a crystal of salt in the water, what will happen? Salt is a solid, a crystal, an organized arrangement of "salt atoms." Figure 1-7 is an illustration of the three-dimensional structure of common salt, sodium chloride. Strictly speaking, the crystal is not made of atoms, but of what we call *ions.* An ion is an atom which either has a few extra electrons or has lost a few electrons. In a salt crystal we find chlorine ions (chlorine atoms with an extra electron) and sodium ions (sodium atoms with one electron missing). The ions all stick together by electrical attraction in the solid salt, but when we put them in the water we find, because of the attractions of the negative oxygen and positive hydrogen for the ions, that some of the ions jiggle loose. In Fig. 1-6 we see a chlorine ion getting loose, and other atoms floating in the water in the form of ions. This picture was made with some care. Notice, for example, that the hydrogen ends of the water molecules are more likely to be near the chlorine ion, while near the sodium ion we are more likely to find the oxygen end, because the

Figure 1-6

Crystal	●	○	a (Å)
Rocksalt	Na	Cl	5.64
Sylvine	K	Cl	6.28
	Ag	Cl	5.54
	Mg	O	4.20
Galena	Pb	S	5.97
	Pb	Se	6.14
	Pb	Te	6.34

Nearest neighbor distance $d = a/2$

Figure 1-7

sodium is positive and the oxygen end of the water is negative, and they attract electrically. Can we tell from this picture whether the salt is *dissolving in* water or *crystallizing out* of water? Of course we *cannot* tell, because while some of the atoms are leaving the crystal other atoms are rejoining it. The process is a *dynamic* one, just as in the case of evaporation, and it depends on whether there is more or less salt in the water than the amount needed for equilibrium. By equilibrium we mean that situation in which the rate at which atoms are leaving just matches the rate at which they are coming back. If there is almost no salt in the water, more atoms leave than return, and the salt dissolves. If, on the other hand, there are too many "salt atoms," more return than leave, and the salt is crystallizing.

In passing, we mention that the concept of a *molecule* of a substance is only approximate and exists only for a certain class of substances. It is clear in the case of water that the three atoms are actually stuck together. It is not so clear in the case of sodium chloride in the solid. There is just an arrangement of sodium and chlorine ions in a cubic pattern. There is no natural way to group them as "molecules of salt."

Returning to our discussion of solution and precipitation, if we increase the temperature of the salt solution, then the rate at which

atoms are taken away is increased, and so is the rate at which atoms are brought back. It turns out to be very difficult, in general, to predict which way it is going to go, whether more or less of the solid will dissolve. Most substances dissolve more, but some substances dissolve less, as the temperature increases.

Chemical reactions

In all of the processes which have been described so far, the atoms and the ions have not changed partners, but of course there are circumstances in which the atoms do change combinations, forming new molecules. This is illustrated in Fig. 1-8. A process in which the rearrangement of the atomic partners occurs is what we call a *chemical reaction.* The other processes so far described are called physical processes, but there is no sharp distinction between the two. (Nature does not care what we call it, she just keeps on doing it.) This figure is supposed to represent carbon burning in oxygen. In the case of oxygen, *two* oxygen atoms stick together very strongly. (Why do not *three* or even *four* stick together? That is one of the very peculiar characteristics of such atomic processes. Atoms are very special: they like certain particular partners, certain particular directions, and so on. It is the job of physics to analyze why each one wants what it wants. At any rate, two oxygen atoms form, saturated and happy, a molecule.)

CARBON BURNING IN OXYGEN

Figure 1-8

The carbon atoms are supposed to be in a solid crystal (which could be graphite or diamond*). Now, for example, one of the oxygen molecules can come over to the carbon, and each atom can pick up a carbon atom and go flying off in a new combination—"carbon-oxygen"—which is a molecule of the gas called carbon monoxide. It is given the chemical name CO. It is very simple: the letters "CO" are practically a picture of that molecule. But carbon attracts oxygen much more than oxygen attracts oxygen or carbon attracts carbon. Therefore in this process the oxygen may arrive with only a little energy, but the oxygen and carbon will snap together with a tremendous vengeance and commotion, and everything near them will pick up the energy. A large amount of motion energy, kinetic energy, is thus generated. This of course is *burning;* we are getting *heat* from the combination of oxygen and carbon. The heat is ordinarily in the form of the molecular motion of the hot gas, but in certain circumstances it can be so enormous that it generates *light.* That is how one gets *flames.*

In addition, the carbon monoxide is not quite satisfied. It is possible for it to attach another oxygen, so that we might have a much more complicated reaction in which the oxygen is combining with the carbon, while at the same time there happens to be a collision with a carbon monoxide molecule. One oxygen atom could attach itself to the CO and ultimately form a molecule, composed of one carbon and two oxygens, which is designated CO_2 and called carbon dioxide. If we burn the carbon with very little oxygen in a very rapid reaction (for example, in an automobile engine, where the explosion is so fast that there is not time for it to make carbon dioxide) a considerable amount of carbon monoxide is formed. In many such rearrangements, a very large amount of energy is released, forming explosions, flames, etc., depending on the reactions. Chemists have studied these arrangements of the atoms, and found that every substance is some type of *arrangement of atoms.*

---

*One *can* burn a diamond in air.

To illustrate this idea, let us consider another example. If we go into a field of small violets, we know what "that smell" is. It is some kind of *molecule,* or arrangement of atoms, that has worked its way into our noses. First of all, *how* did it work its way in? That is rather easy. If the smell is some kind of molecule in the air, jiggling around and being knocked every which way, it might have *accidentally* worked its way into the nose. Certainly it has no particular desire to get into our nose. It is merely one helpless part of a jostling crowd of molecules, and in its aimless wanderings this particular chunk of matter happens to find itself in the nose.

Now chemists can take special molecules like the odor of violets, and analyze them and tell us the *exact arrangement* of the atoms in space. We know that the carbon dioxide molecule is straight and symmetrical: O—C—O. (That can be determined easily, too, by physical methods.) However, even for the vastly more complicated arrangements of atoms that there are in chemistry, one can, by a long, remarkable process of detective work, find the arrangements of the atoms. Figure 1-9 is a picture of the air in the neighborhood of a violet; again we find nitrogen and oxygen in the air, and water vapor. (Why is there water vapor? Because the violet is *wet.* All plants transpire.) However, we also see a "monster" composed of carbon atoms, hydrogen atoms, and oxygen atoms, which have picked a certain particular pattern in which to be arranged. It is a much more complicated arrangement than that of carbon dioxide;

ODOR OF VIOLETS

Figure 1-9

in fact, it is an enormously complicated arrangement. Unfortunately, we cannot picture all that is really known about it chemically, because the precise arrangement of all the atoms is actually known in three dimensions, while our picture is in only two dimensions. The six carbons which form a ring do not form a flat ring, but a kind of "puckered" ring. All of the angles and distances are known. So a chemical *formula* is merely a picture of such a molecule. When the chemist writes such a thing on the blackboard, he is trying to "draw," roughly speaking, in two dimensions. For example, we see a "ring" of six carbons, and a "chain" of carbons hanging on the end, with an oxygen second from the end, three hydrogens tied to that carbon, two carbons and three hydrogens sticking up here, etc.

How does the chemist find what the arrangement is? He mixes bottles full of stuff together, and if it turns red, it tells him that it consists of one hydrogen and two carbons tied on here; if it turns blue, on the other hand, that is not the way it is at all. This is one of the most fantastic pieces of detective work that has ever been done—organic chemistry. To discover the arrangement of the atoms in these enormously complicated arrays the chemist looks at what happens when he mixes two different substances together. The physicist could never quite believe that the chemist knew what he was talking about when he described the arrangement of the atoms. For about twenty years it has been possible, in some cases, to look at such molecules (not quite as complicated as this one, but some which contain parts of it) by a physical method, and it has been possible to locate every atom, not by looking at colors, but by *measuring where they are.* And lo and behold!, the chemists are almost always correct.

It turns out, in fact, that in the odor of violets there are three slightly different molecules, which differ only in the arrangement of the hydrogen atoms.

One problem of chemistry is to name a substance, so that we will know what it is. Find a name for this shape! Not only must the name

Figure 1-10 The substance pictured is α-irone.

tell the shape, but it must also tell that here is an oxygen atom, there a hydrogen—exactly what and where each atom is. So we can appreciate that the chemical names must be complex in order to be complete. You see that the name of this thing in the more complete form that will tell you the structure of it is 4-(2, 2, 3, 6 tetramethyl-5-cyclohexenyl)-3-buten-2-one, and that tells you that this is the arrangement. We can appreciate the difficulties that the chemists have, and also appreciate the reason for such long names. It is not that they wish to be obscure, but they have an extremely difficult problem in trying to describe the molecules in words!

How do we *know* that there are atoms? By one of the tricks mentioned earlier: we make the *hypothesis* that there are atoms, and one after the other results come out the way we predict, as they ought to if things *are* made of atoms. There is also somewhat more direct evidence, a good example of which is the following: The atoms are so small that you cannot see them with a light microscope—in fact, not even with an *electron* microscope. (With a light microscope you can only see things which are much bigger.) Now if the atoms are always in motion, say in water, and we put a big ball of something in the water, a ball much bigger than the atoms, the ball will jiggle around—much as in a push ball game, where a great big ball is pushed around by a lot of people. The people are pushing in various directions, and the ball moves around the field in an irregular fashion. So, in the same way, the "large ball" will move because of

the inequalities of the collisions on one side to the other, from one moment to the next. Therefore, if we look at very tiny particles (colloids) in water through an excellent microscope, we see a perpetual jiggling of the particles, which is the result of the bombardment of the atoms. This is called the *Brownian motion.*

We can see further evidence for atoms in the structure of crystals. In many cases the structures deduced by x-ray analysis agree in their spatial "shapes" with the forms actually exhibited by crystals as they occur in nature. The angles between the various "faces" of a crystal agree, within seconds of arc, with angles deduced on the assumption that a crystal is made of many "layers" of atoms.

Everything is made of atoms. That is the key hypothesis. The most important hypothesis in all of biology, for example, is that *everything that animals do, atoms do.* In other words, *there is nothing that living things do that cannot be understood from the point of view that they are made of atoms acting according to the laws of physics.* This was not known from the beginning: it took some experimenting and theorizing to suggest this hypothesis, but now it is accepted, and it is the most useful theory for producing new ideas in the field of biology.

If a piece of steel or a piece of salt, consisting of atoms one next to the other, can have such interesting properties; if water—which is nothing but these little blobs, mile upon mile of the same thing over the earth—can form waves and foam, and make rushing noises and strange patterns as it runs over cement; if all of this, all the life of a stream of water, can be nothing but a pile of atoms, *how much more is possible*? If instead of arranging the atoms in some definite pattern, again and again repeated, on and on, or even forming little lumps of complexity like the odor of violets, we make an arrangement which is *always different* from place to place, with different kinds of atoms arranged in many ways, continually changing, not repeating, how much more marvelously is it possible that this thing might behave? Is it possible that that "thing" walking back and forth in front of you, talking to you, is a great glob of these atoms

in a very complex arrangement, such that the sheer complexity of it staggers the imagination as to what it can do? When we say we are a pile of atoms, we do not mean we are *merely* a pile of atoms, because a pile of atoms which is not repeated from one to the other might well have the possibilities which you see before you in the mirror.

2

BASIC PHYSICS

Introduction

In this chapter, we shall examine the most fundamental ideas that we have about physics—the nature of things as we see them at the present time. We shall not discuss the history of how we know that all these ideas are true; you will learn these details in due time.

The things with which we concern ourselves in science appear in myriad forms, and with a multitude of attributes. For example, if we stand on the shore and look at the sea, we see the water, the waves breaking, the foam, the sloshing motion of the water, the sound, the air, the winds and the clouds, the sun and the blue sky, and light; there is sand and there are rocks of various hardness and permanence, color and texture. There are animals and seaweed, hunger and disease, and the observer on the beach; there may be even happiness and thought. Any other spot in nature has a similar variety of things and influences. It is always as complicated as that, no matter where it is. Curiosity demands that we ask questions, that we try to put things together and try to understand this multitude of aspects as perhaps resulting from the action of a relatively small number of elemental things and forces acting in an infinite variety of combinations.

For example: Is the sand other than the rocks? That is, is the sand perhaps nothing but a great number of very tiny stones? Is the moon a great rock? If we understood rocks, would we also understand the

sand and the moon? Is the wind a sloshing of the air analogous to the sloshing motion of the water in the sea? What common features do different movements have? What is common to different kinds of sound? How many different colors are there? And so on. In this way we try gradually to analyze all things, to put together things which at first sight look different, with the hope that we may be able to *reduce* the number of *different* things and thereby understand them better.

A few hundred years ago, a method was devised to find partial answers to such questions. *Observation, reason,* and *experiment* make up what we call the *scientific method.* We shall have to limit ourselves to a bare description of our basic view of what is sometimes called *fundamental physics,* or fundamental ideas which have arisen from the application of the scientific method.

What do we mean by "understanding" something? We can imagine that this complicated array of moving things which constitutes "the world" is something like a great chess game being played by the gods, and we are observers of the game. We do not know what the rules of the game are; all we are allowed to do is to *watch* the playing. Of course, if we watch long enough, we may eventually catch on to a few of the rules. *The rules of the game* are what we mean by *fundamental physics.* Even if we knew every rule, however, we might not be able to understand why a particular move is made in the game, merely because it is too complicated and our minds are limited. If you play chess you must know that it is easy to learn all the rules, and yet it is often very hard to select the best move or to understand why a player moves as he does. So it is in nature, only much more so; but we may be able at least to find all the rules. Actually, we do not have all the rules now. (Every once in a while something like castling is going on that we still do not understand.) Aside from not knowing all of the rules, what we really can explain in terms of those rules is very limited, because almost all situations are so enormously complicated that we cannot follow the plays of the game using the rules, much less tell what is going to happen next. We must, therefore, limit ourselves to the more basic question

of the rules of the game. If we know the rules, we consider that we "understand" the world.

How can we tell whether the rules which we "guess" at are really right if we cannot analyze the game very well? There are, roughly speaking, three ways. First, there may be situations where nature has arranged, or we arrange nature, to be simple and to have so few parts that we can predict exactly what will happen, and thus we can check how our rules work. (In one corner of the board there may be only a few chess pieces at work, and that we can figure out exactly.)

A second good way to check rules is in terms of less specific rules derived from them. For example, the rule on the move of a bishop on a chessboard is that it moves only on the diagonal. One can deduce, no matter how many moves may be made, that a certain bishop will always be on a red square. So, without being able to follow the details, we can always check our idea about the bishop's motion by finding out whether it is always on a red square. Of course it will be, for a long time, until all of a sudden we find that it is on a *black* square (what happened, of course, is that in the meantime it was captured, another pawn crossed for queening, and it turned into a bishop on a black square). That is the way it is in physics. For a long time we will have a rule that works excellently in an overall way, even when we cannot follow the details, and then sometime we may discover a *new rule.* From the point of view of basic physics, the most interesting phenomena are of course in the *new* places, the places where the rules do not work—not the places where they *do* work! That is the way in which we discover new rules.

The third way to tell whether our ideas are right is relatively crude but probably the most powerful of them all. That is, by rough *approximation.* While we may not be able to tell why Alekhine moves *this particular piece,* perhaps we can *roughly* understand that he is gathering his pieces around the king to protect it, more or less, since that is the sensible thing to do in the circumstances. In the same way, we can often understand nature, more or less, without

being able to see what *every little piece* is doing, in terms of our understanding of the game.

At first the phenomena of nature were roughly divided into classes, like heat, electricity, mechanics, magnetism, properties of substances, chemical phenomena, light or optics, x-rays, nuclear physics, gravitation, meson phenomena, etc. However, the aim is to see *complete nature* as different aspects of *one set* of phenomena. That is the problem in basic theoretical physics today—to *find the laws behind experiment;* to *amalgamate these classes.* Historically, we have always been able to amalgamate them, but as time goes on new things are found. We were amalgamating very well, when all of a sudden x-rays were found. Then we amalgamated some more, and mesons were found. Therefore, at any stage of the game, it always looks rather messy. A great deal is amalgamated, but there are always many wires or threads hanging out in all directions. That is the situation today, which we shall try to describe.

Some historic examples of amalgamation are the following. First, take *heat* and *mechanics.* When atoms are in motion, the more motion, the more heat the system contains, and so *heat and all temperature effects can be represented by the laws of mechanics.* Another tremendous amalgamation was the discovery of the relation between electricity, magnetism, and light, which were found to be different aspects of the same thing, which we call today the *electromagnetic field.* Another amalgamation is the unification of chemical phenomena, the various properties of various substances, and the behavior of atomic particles, which is in the *quantum mechanics of chemistry.*

The question is, of course, is it going to be possible to amalgamate *everything,* and merely discover that this world represents different aspects of *one* thing? Nobody knows. All we know is that as we go along, we find that we can amalgamate pieces, and then we find some pieces that do not fit, and we keep trying to put the jigsaw puzzle together. Whether there are a finite number of pieces, and whether there is even a border to the puzzle, are of course un-

known. It will never be known until we finish the picture, if ever. What we wish to do here is to see to what extent this amalgamation process has gone on, and what the situation is at present, in understanding basic phenomena in terms of the smallest set of principles. To express it in a simple manner, *what are things made of and how few elements are there?*

Physics before 1920

It is a little difficult to begin at once with the present view, so we shall first see how things looked in about 1920 and then take a few things out of that picture. Before 1920, our world picture was something like this: The "stage" on which the universe goes is the three-dimensional *space* of geometry, as described by Euclid, and things change in a medium called *time.* The elements on the stage are *particles,* for example the atoms, which have some *properties.* First, the property of inertia: if a particle is moving it keeps on going in the same direction unless *forces* act upon it. The second element, then, is *forces,* which were then thought to be of two varieties: First, an enormously complicated, detailed kind of interaction force which held the various atoms in different combinations in a complicated way, which determined whether salt would dissolve faster or slower when we raise the temperature. The other force that was known was a long-range interaction—a smooth and quiet attraction—which varied inversely as the square of the distance, and was called *gravitation.* This law was known and was very simple. *Why* things remain in motion when they are moving, or *why* there is a law of gravitation, was, of course, not known.

A description of nature is what we are concerned with here. From this point of view, then, a gas, and indeed *all* matter, is a myriad of moving particles. Thus many of the things we saw while standing at the seashore can immediately be connected. First the pressure: this comes from the collisions of the atoms with the walls or whatever; the drift of the atoms, if they are all moving in one

direction on the average, is wind; the *random* internal motions are the *heat.* There are waves of excess density, where too many particles have collected, and so as they rush off they push up piles of particles farther out, and so on. This wave of excess density is *sound.* It is a tremendous achievement to be able to understand so much. Some of these things were described in the previous chapter.

What *kinds* of particles are there? There were considered to be 92 at that time: 92 different kinds of atoms were ultimately discovered. They had different names associated with their chemical properties.

The next part of the problem was, *what are the short-range forces?* Why does carbon attract one oxygen or perhaps two oxygens, but not three oxygens? What is the machinery of interaction between atoms? Is it gravitation? The answer is no. Gravity is entirely too weak. But imagine a force analogous to gravity, varying inversely with the square of the distance, but enormously more powerful and having one difference. In gravity everything attracts everything else, but now imagine that there are *two kinds* of "things," and that this new force (which is the electrical force, of course) has the property that likes *repel* but unlikes *attract.* The "thing" that carries this strong interaction is called *charge.*

Then what do we have? Suppose that we have two unlikes that attract each other, a plus and a minus, and that they stick very close together. Suppose we have another charge some distance away. Would it feel any attraction? It would feel *practically none,* because if the first two are equal in size, the attraction for the one and the repulsion for the other balance out. Therefore there is very little force at any appreciable distance. On the other hand, if we get *very close* with the extra charge, *attraction* arises, because the repulsion of likes and attraction of unlikes will tend to bring unlikes closer together and push likes farther apart. Then the repulsion will be *less* than the attraction. This is the reason why the atoms, which are constituted out of plus and minus electric charges, feel very little

force when they are separated by appreciable distance (aside from gravity). When they come close together, they can "see inside" each other and rearrange their charges, with the result that they have a very strong interaction. The ultimate basis of an interaction between the atoms is *electrical.* Since this force is so enormous, all the plusses and all minuses will normally come together in as intimate a combination as they can. All things, even ourselves, are made of fine-grained, enormously strongly interacting plus and minus parts, all neatly balanced out. Once in a while, by accident, we may rub off a few minuses or a few plusses (usually it is easier to rub off minuses), and in those circumstances we find the force of electricity *unbalanced,* and we can then see the effects of these electrical attractions.

To give an idea of how much stronger electricity is than gravitation, consider two grains of sand, a millimeter across, thirty meters apart. If the force between them were not balanced, if everything attracted everything else instead of likes repelling, so that there were no cancellations, how much force would there be? There would be a force of *three million tons* between the two! You see, there is very, *very* little excess or deficit of the number of negative or positive charges necessary to produce appreciable electrical effects. This is, of course, the reason why you cannot see the difference between an electrically charged and an uncharged thing—so few particles are involved that they hardly make a difference in the weight or size of an object.

With this picture the atoms were easier to understand. They were thought to have a "nucleus" at the center, which is positively electrically charged and very massive, and the nucleus is surrounded by a certain number of "electrons" which are very light and negatively charged. Now we go a little ahead in our story to remark that in the nucleus itself there were found two kinds of particles, protons and neutrons, almost of the same weight and very heavy. The protons are electrically charged and the neutrons are neutral. If we have an atom with six protons inside its nucleus, and this is surrounded

by six electrons (the negative particles in the ordinary world of matter are all electrons, and these are very light compared with the protons and neutrons which make nuclei), this would be atom number six in the chemical table, and it is called carbon. Atom number eight is called oxygen, etc., because the chemical properties depend upon the electrons on the *outside,* and in fact only upon *how many* electrons there are. So the *chemical* properties of a substance depend only on a number, the number of electrons. (The whole list of elements of the chemists really could have been called 1, 2, 3, 4, 5, etc. Instead of saying "carbon," we could say "element six," meaning six electrons, but of course, when the elements were first discovered, it was not known that they could be numbered that way, and secondly, it would make everything look rather complicated. It is better to have names and symbols for these things, rather than to call everything by number.)

More was discovered about the electrical force. The natural interpretation of electrical interaction is that two objects simply attract each other: plus against minus. However, this was discovered to be an inadequate idea to represent it. A more adequate representation of the situation is to say that the existence of the positive charge, in some sense, distorts, or creates a "condition" in space, so that when we put the negative charge in, it feels a force. This potentiality for producing a force is called an *electric field.* When we put an electron in an electric field, we say it is "pulled." We then have two rules: (a) charges make a field, and (b) charges in fields have forces on them and move. The reason for this will become clear when we discuss the following phenomena: If we were to charge a body, say a comb, electrically, and then place a charged piece of paper at a distance and move the comb back and forth, the paper will respond by always pointing to the comb. If we shake it faster, it will be discovered that the paper is a little behind, *there is a delay* in the action. (At the first stage, when we move the comb rather slowly, we find a complication which is *magnetism.* Magnetic influences have to do with *charges in relative motion,* so magnetic forces and electric forces can really be attributed to one field, as two

different aspects of exactly the same thing. A changing electric field cannot exist without magnetism.) If we move the charged paper farther out, the delay is greater. Then an interesting thing is observed. Although the forces between two charged objects should go inversely as the *square* of the distance, it is found, when we shake a charge, that the influence extends *very much farther out* than we would guess at first sight. That is, the effect falls off more slowly than the inverse square.

Here is an analogy: If we are in a pool of water and there is a floating cork very close by, we can move it "directly" by pushing the water with another cork. If you looked only at the two *corks,* all you would see would be that one moved immediately in response to the motion of the other—there is some kind of *"interaction"* between them. Of course, what we really do is to disturb the *water;* the *water* then disturbs the other cork. We could make up a "law" that if you pushed the water a little bit, an object close by in the water would move. If it were farther away, of course, the second cork would scarcely move, for we move the water *locally.* On the other hand, if we jiggle the cork a new phenomenon is involved, in which the motion of the water moves the water there, etc., and *waves* travel away, so that by jiggling, there is an influence *very much farther out,* an oscillatory influence, that cannot be understood from the direct interaction. Therefore the idea of direct interaction must be replaced with the existence of the water, or in the electrical case, with what we call the *electromagnetic field.*

The electromagnetic field can carry waves; some of these waves are *light,* others are used in *radio broadcasts,* but the general name is *electromagnetic waves.* These oscillatory waves can have various *frequencies.* The only thing that is really different from one wave to another is the *frequency of oscillation.* If we shake a charge back and forth more and more rapidly, and look at the effects, we get a whole series of different kinds of effects, which are all unified by specifying but one number, the number of oscillations per second. The usual "pickup" that we get from electric currents in the circuits in the walls of a building has a frequency of about one hundred cycles per

Frequency in oscillations/sec	Name	Rough behavior
10^2	Electrical disturbance	Field
$5 \times 10^5 - 10^6$	Radio broadcast	Waves
10^8	FM—TV	Waves
10^{10}	Radar	Waves
$5 \times 10^{14} - 10^{15}$	Light	Waves
10^{18}	X-rays	Particle
10^{21}	γ-rays, nuclear	Particle
10^{24}	γ-rays, "artificial"	Particle
10^{27}	γ-rays, in cosmic rays	Particle

Table 2-1 The Electromagnetic Spectrum.

second. If we increase the frequency to 500 or 1000 kilocycles (1 kilocycle = 1000 cycles) per second, we are "on the air," for this is the frequency range which is used for radio broadcasts. (Of course it has nothing to do with the *air!* We can have radio broadcasts without any air.) If we again increase the frequency, we come into the range that is used for FM and TV. Going still further, we use certain short waves, for example for *radar.* Still higher, and we do not need an instrument to "see" the stuff, we can see it with the human eye. In the range of frequency from 5×10^{14} to 10^{15} cycles per second our eyes would see the oscillation of the charged comb, if we could shake it that fast, as red, blue, or violet light, depending on the frequency. Frequencies below this range are called infrared, and above it, ultraviolet. The fact that we can see in a particular frequency range makes that part of the electromagnetic spectrum no more impressive than the other parts from a physicist's standpoint, but from a human standpoint, of course, it *is* more interesting. If we go up even higher in frequency, we get x-rays. X-rays are nothing but very high-frequency light. If we go still higher, we get gamma rays. These two terms, x-rays and gamma rays, are used al-

most synonymously. Usually electromagnetic rays coming from nuclei are called gamma rays, while those of high energy from atoms are called x-rays, but at the same frequency they are indistinguishable physically, no matter what their source. If we go to still higher frequencies, say to 10^{24} cycles per second, we find that we can make those waves artificially, for example with the synchrotron here at Caltech. We can find electromagnetic waves with stupendously high frequencies—with even a thousand times more rapid oscillation—in the waves found in *cosmic rays.* These waves cannot be controlled by us.

Quantum physics

Having described the idea of the electromagnetic field, and that this field can carry waves, we soon learn that these waves actually behave in a strange way which seems very unwavelike. At higher frequencies they behave much more like *particles!* It is *quantum mechanics,* discovered just after 1920, which explains this strange behavior. In the years before 1920, the picture of space as a three-dimensional space, and of time as a separate thing, was changed by Einstein, first into a combination which we call space-time, and then still further into a *curved* space-time to represent gravitation. So the "stage" is changed into space-time, and gravitation is presumably a modification of space-time. Then it was also found that the rules for the motions of particles were incorrect. The mechanical rules of "inertia" and "forces" are *wrong*—Newton's laws are *wrong*—in the world of atoms. Instead, it was discovered that things on a small scale behave *nothing like* things on a large scale. That is what makes physics difficult—and very interesting. It is hard because the way things behave on a small scale is so "unnatural"; we have no direct experience with it. Here things behave like nothing we know of, so that it is impossible to describe this behavior in any other than analytic ways. It is difficult, and takes a lot of imagination.

Quantum mechanics has many aspects. In the first place, the

idea that a particle has a definite location and a definite speed is no longer allowed; that is wrong. To give an example of how wrong classical physics is, there is a rule in quantum mechanics that says that one cannot know both where something is and how fast it is moving. The uncertainty of the momentum and the uncertainty of the position are complementary, and the product of the two is bounded by a small constant. We can write the law like this: $\Delta x\, \Delta p \geq \hbar/2$, but we shall explain it in more detail later. This rule is the explanation of a very mysterious paradox: If the atoms are made out of plus and minus charges, why don't the minus charges simply sit on top of the plus charges (they attract each other) and get so close as to completely cancel them out? *Why are atoms so big?* Why is the nucleus at the center with the electrons around it? It was first thought that this was because the nucleus was so big; but no, the nucleus is *very small.* An atom has a diameter of about 10^{-8} cm. The nucleus has a diameter of about 10^{-13} cm. If we had an atom and wished to see the nucleus, we would have to magnify it until the whole atom was the size of a large room, and then the nucleus would be a bare speck which you could just about make out with the eye, but very nearly *all the weight* of the atom is in that infinitesimal *nucleus.* What keeps the electrons from simply falling in? This principle: If they were in the nucleus, we would know their position precisely, and the uncertainty principle would then require that they have a very *large* (but uncertain) momentum, i.e., a very large *kinetic energy.* With this energy they would break away from the nucleus. They make a compromise: they leave themselves a little room for this uncertainty and then jiggle with a certain amount of minimum motion in accordance with this rule. (Remember that when a crystal is cooled to absolute zero, we said that the atoms do not stop moving, they still jiggle. Why? If they stopped moving, we would know where they were and that they had zero motion, and that is against the uncertainty principle. We cannot know where they are and how fast they are moving, so they must be continually wiggling in there!)

Another most interesting change in the ideas and philosophy

of science brought about by quantum mechanics is this: it is not possible to predict *exactly* what will happen in any circumstance. For example, it is possible to arrange an atom which is ready to emit light, and we can measure when it has emitted light by picking up a photon particle, which we shall describe shortly. We cannot, however, predict *when* it is going to emit the light or, with several atoms, *which one* is going to. You may say that this is because there are some internal "wheels" which we have not looked at closely enough. No, there *are* no internal wheels; nature, as we understand it today, behaves in such a way that it is *fundamentally impossible* to make a precise prediction of *exactly what will happen* in a given experiment. This is a horrible thing; in fact, philosophers have said before that one of the fundamental requisites of science is that whenever you set up the same conditions, the same thing must happen. This is simply *not true;* it is *not* a fundamental condition of science. The fact is that the same thing does not happen, that we can find only an average, statistically, as to what happens. Nevertheless, science has not completely collapsed. Philosophers, incidentally, say a great deal about what is *absolutely necessary* for science, and it is always, so far as one can see, rather naive, and probably wrong. For example, some philosopher or other said it is fundamental to the scientific effort that if an experiment is performed in, say, Stockholm, and then the same experiment is done in, say, Quito, the *same results* must occur. That is quite false. It is not necessary that *science* do that; it may be a *fact of experience,* but it is not necessary. For example, if one of the experiments is to look out at the sky and see the aurora borealis in Stockholm, you do not see it in Quito; that is a different phenomenon. "But," you say, "that is something that has to do with the outside; can you close yourself up in a box in Stockholm and pull down the shade and get any difference?" Surely. If we take a pendulum on a universal joint, and pull it out and let go, then the pendulum will swing almost in a plane, but not quite. Slowly the plane keeps changing in Stockholm, but not in Quito. The blinds are down,

too. The fact that this happened does not bring on the destruction of science. What *is* the fundamental hypothesis of science, the fundamental philosophy? We stated it in the first chapter: *the sole test of the validity of any idea is experiment.* If it turns out that most experiments work out the same in Quito as they do in Stockholm, then those "most experiments" will be used to formulate some general law, and those experiments which do not come out the same we will say were a result of the environment near Stockholm. We will invent some way to summarize the results of the experiment, and we do not have to be told ahead of time what this way will look like. If we are told that the same experiment will always produce the same result, that is all very well, but if when we try it, it does *not,* then it does *not.* We just have to take what we see, and then formulate all the rest of our ideas in terms of our actual experience.

Returning again to quantum mechanics and fundamental physics, we cannot go into details of the quantum-mechanical principles at this time, of course, because these are rather difficult to understand. We shall assume that they are there, and go on to describe what some of the consequences are. One of the consequences is that things which we used to consider as waves also behave like particles, and particles behave like waves; in fact everything behaves the same way. There is no distinction between a wave and a particle. So quantum mechanics *unifies* the idea of the field and its waves, and the particles, all into one. Now it is true that when the frequency is low, the field aspect of the phenomenon is more evident, or more useful as an approximate description in terms of everyday experiences. But as the frequency increases, the particle aspects of the phenomenon become more evident with the equipment with which we usually make the measurements. In fact, although we mentioned many frequencies, no phenomenon directly involving a frequency has yet been detected above approximately 10^{12} cycles per second. We only *deduce* the higher frequencies from the energy of the particles, by a rule

which assumes that the particle-wave idea of quantum mechanics is valid.

Thus we have a new view of electromagnetic interaction. We have a new kind of *particle* to add to the electron, the proton, and the neutron. That new particle is called a *photon.* The new view of the interaction of electrons and photons that is electromagnetic theory, but with everything quantum-mechanically correct, is called *quantum electrodynamics.* This fundamental theory of the interaction of light and matter, or electric field and charges, is our greatest success so far in physics. In this one theory we have the basic rules for all ordinary phenomena except for gravitation and nuclear processes. For example, out of quantum electrodynamics come all known electrical, mechanical, and chemical laws: the laws for the collision of billiard balls, the motions of wires in magnetic fields, the specific heat of carbon monoxide, the color of neon signs, the density of salt, and the reactions of hydrogen and oxygen to make water are all consequences of this one law. All these details can be worked out if the situation is simple enough for us to make an approximation, which is almost never, but often we can understand more or less what is happening. At the present time no exceptions are found to the quantum-electrodynamic laws outside the nucleus, and there we do not know whether there is an exception because we simply do not know what is going on in the nucleus.

In principle, then, quantum electrodynamics is the theory of all chemistry, and of life, if life is ultimately reduced to chemistry and therefore just to physics because chemistry is already reduced (the part of physics which is involved in chemistry being already known). Furthermore, the same quantum electrodynamics, this great thing, predicts a lot of new things. In the first place, it tells the properties of very high-energy photons, gamma rays, etc. It predicted another very remarkable thing: besides the electron, there should be another particle of the same mass, but of opposite charge, called a *positron,* and these two, coming together, could annihilate each other with the emission of light or gamma rays.

(After all, light and gamma rays are all the same; they are just different points on a frequency scale.) The generalization of this, that for each particle there is an antiparticle, turns out to be true. In the case of electrons, the antiparticle has another name—it is called a positron, but for most other particles, it is called anti-so-and-so, like antiproton or antineutron. In quantum electrodynamics, *two numbers* are put in and most of the other numbers in the world are supposed to come out. The two numbers that are put in are called the mass of the electron and the charge of the electron. Actually, that is not quite true, for we have a whole set of numbers for chemistry which tells how heavy the nuclei are. That leads us to the next part.

Nuclei and particles

What are the nuclei made of, and how are they held together? It is found that the nuclei are held together by enormous forces. When these are released, the energy released is tremendous compared with chemical energy, in the same ratio as the atomic bomb explosion is to a TNT explosion, because, of course, the atomic bomb has to do with changes inside the nucleus, while the explosion of TNT has to do with the changes of the electrons on the outside of the atoms. The question is, what are the forces which hold the protons and neutrons together in the nucleus? Just as the electrical interaction can be connected to a particle, a photon, Yukawa suggested that the forces between neutrons and protons also have a field of some kind, and that when this field jiggles it behaves like a particle. Thus there could be some other particles in the world besides protons and neutrons, and he was able to deduce the properties of these particles from the already known characteristics of nuclear forces. For example, he predicted they should have a mass of two or three hundred times that of an electron; and lo and behold, in cosmic rays there was discovered a particle of the right mass! But it later turned out to be the wrong particle. It was called a μ-meson, or muon.

However, a little while later, in 1947 or 1948, another particle was found, the π-meson, or pion, which satisfied Yukawa's criterion. Besides the proton and the neutron, then, in order to get nuclear forces we must add the pion. Now, you say, "Oh great!, with this theory we make quantum nucleodynamics using the pions just like Yukawa wanted to do, and see if it works, and everything will be explained." Bad luck. It turns out that the calculations that are involved in this theory are so difficult that no one has ever been able to figure out what the consequences of the theory are, or to check it against experiment, and this has been going on now for almost twenty years!

So we are stuck with a theory, and we do not know whether it is right or wrong, but we do know that it is a *little* wrong, or at least incomplete. While we have been dawdling around theoretically, trying to calculate the consequences of this theory, the experimentalists have been discovering some things. For example, they had already discovered this μ-meson or muon, and we do not yet know where it fits. Also, in cosmic rays, a large number of other "extra" particles were found. It turns out that today we have approximately thirty particles, and it is very difficult to understand the relationships of all these particles, and what nature wants them for, or what the connections are from one to another. We do not today understand these various particles as different aspects of the same thing, and the fact that we have so many unconnected particles is a representation of the fact that we have so much unconnected information without a good theory. After the great successes of quantum electrodynamics, there is a certain amount of knowledge of nuclear physics which is rough knowledge, sort of half experience and half theory, assuming a type of force between protons and neutrons and seeing what will happen, but not really understanding where the force comes from. Aside from that, we have made very little progress. We have collected an enormous number of chemical elements. In the chemical case, there suddenly appeared a relationship among these elements which was unexpected, and which is embodied in the periodic table of Mendeléev. For example, sodium and

potassium are about the same in their chemical properties and are found in the same column in the Mendeléev chart. We have been seeking a Mendeléev-type chart for the new particles. One such chart of the new particles was made independently by Gell-Mann in the USA and Nishijima in Japan. The basis of their classification is a new number, like the electric charge, which can be assigned to each particle, called its "strangeness," *S.* This number is conserved, like the electric charge, in reactions which take place by nuclear forces.

In Table 2-2 are listed all the particles. We cannot discuss them much at this stage, but the table will at least show you how much we do not know. Underneath each particle its mass is given in a certain unit, called the Mev. One Mev is equal to 1.782×10^{-27} gram. The reason this unit was chosen is historical, and we shall not go into it now. More massive particles are put higher up on the chart; we see that a neutron and a proton have almost the same mass. In vertical columns we have put the particles with the same electrical charge, all neutral objects in one column, all positively charged ones to the right of this one, and all negatively charged objects to the left.

Particles are shown with a solid line and "resonances" with a dashed one. Several particles have been omitted from the table. These include the important zero-mass, zero-charge particles, the photon and the graviton, which do not fall into the baryon-meson-lepton classification scheme, and also some of the newer resonances (K^*, φ, η). The antiparticles of the mesons are listed in the table, but the antiparticles of the leptons and baryons would have to be listed in another table which would look exactly like this one reflected on the zero-charge column. Although all of the particles except the electron, neutrino, photon, graviton, and proton are unstable, decay products have been shown only for the resonances. Strangeness assignments are not applicable for leptons, since they do not interact strongly with nuclei.

All particles which are together with the neutrons and protons are called *baryons,* and the following ones exist: There is a "lambda," with a mass of 1115 Mev, and three others, called sigmas, minus,

Table 2-2 Elementary Particles.

neutral, and plus, with several masses almost the same. There are groups or multiplets with almost the same mass, within 1 or 2 percent. Each particle in a multiplet has the same strangeness. The first multiplet is the proton-neutron doublet, and then there is a singlet (the lambda), then the sigma triplet, and finally the xi doublet. Very recently, in 1961, even a few more particles were found. Or *are* they particles? They live so short a time, they disintegrate almost instantaneously, as soon as they are formed, that we do not know whether they should be considered as new particles, or some kind of "resonance" interaction of a certain definite energy between the Λ and π products into which they disintegrate.

In addition to the baryons the other particles which are involved in the nuclear interaction are called *mesons.* There are first the pions, which come in three varieties, positive, negative, and neutral; they form another multiplet. We have also found some new things called K-mesons, and they occur as a doublet, K^+ and K^0. Also, every particle has its antiparticle, unless a particle *is its own* antiparticle. For example, the $\pi-$ and the $\pi+$ are antiparticles, but the π^0 is its own antiparticle. The K^- and K^+ are antiparticles, and the K^0 and $\overline{K}^0$. In addition, in 1961 we also found some more mesons or *maybe* mesons which disintegrate almost immediately. A thing called ω which goes into three pions has a mass 780 on this scale, and somewhat less certain is an object which disintegrates into two pions. These particles, called mesons and baryons, and the antiparticles of the mesons are on the same chart, but the antiparticles of the baryons must be put on another chart, "reflected" through the charge-zero column.

Just as Mendeléev's chart was very good, except for the fact that there were a number of rare earth elements which were hanging out loose from it, so we have a number of things hanging out loose from this chart—particles which do not interact strongly in nuclei, have nothing to do with a nuclear interaction, and do not have a strong interaction (I mean the powerful kind of interaction of nuclear energy). These are called leptons, and they are the following:

there is the electron, which has a very small mass on this scale, only 0.510 Mev. Then there is that other, the μ-meson, the muon, which has a mass much higher, 206 times as heavy as an electron. So far as we can tell, by all experiments so far, the difference between the electron and the muon is nothing but the mass. Everything works exactly the same for the muon as for the electron, except that one is heavier than the other. Why is there another one heavier; what is the use for it? We do not know. In addition, there is a lepton which is neutral, called a neutrino, and this particle has zero mass. In fact, it is now known that there are *two* different kinds of neutrinos, one related to electrons and the other related to muons.

Finally, we have two other particles which do not interact strongly with the nuclear ones: one is a photon, and perhaps, if the field of gravity also has a quantum-mechanical analog (a quantum theory of gravitation has not yet been worked out), then there will be a particle, a graviton, which will have zero mass.

What is this "zero mass"? The masses given here are the masses of the particles *at rest.* The fact that a particle has zero mass means, in a way, that it cannot *be* at *rest.* A photon is never at rest; it is always moving at 186,000 miles a second. We will understand more what mass means when we understand the theory of relativity, which will come in due time.

Thus we are confronted with a large number of particles, which together seem to be the fundamental constituents of matter. Fortunately, these particles are not *all* different in their *interactions* with one another. In fact, there seem to be just *four kinds* of interaction between particles which, in the order of decreasing strength, are the nuclear force, electrical interactions, the beta-decay interaction, and gravity. The photon is coupled to all charged particles and the strength of the interaction is measured by some number, which is $^1/_{137}$. The detailed law of this coupling is known, that is quantum electrodynamics. Gravity is coupled to all *energy,* but its coupling is extremely weak, much weaker than that of electricity. This law is also known. Then there are the so-called weak decays—beta decay,

which causes the neutron to disintegrate into proton, electron, and neutrino, relatively slowly. This law is only partly known. The so-called strong interaction, the meson-baryon interaction, has a strength of 1 in this scale, and the law is completely unknown, although there are a number of known rules, such as that the number of baryons does not change in any reaction.

Coupling	Strength*	Law
Photon to charged particles	$\sim 10^{-2}$	Law known
Gravity to all energy	$\sim 10^{-40}$	Law known
Weak decays	$\sim 10^{-5}$	Law partly known
Mesons to baryons	~ 1	Law unknown (some rules known)

Table 2-3 Elementary Interactions.

This, then, is the horrible condition of our physics today. To summarize it, I would say this: outside the nucleus, we seem to know all; inside it, quantum mechanics is valid—the principles of quantum mechanics have not been found to fail. The stage on which we put all of our knowledge, we would say, is relativistic space-time; perhaps gravity is involved in space-time. We do not know how the universe got started, and we have never made experiments which check our ideas of space and time accurately, below some tiny distance, so we only *know* that our ideas work above that distance. We should also add that the rules of the game are the quantum-mechanical principles, and those principles apply, so far as we can tell, to the new particles as well as to the old. The origin of the forces in nuclei leads us to new particles, but unfortunately

* The "strength" is a dimensionless measure of the coupling constant involved in each interaction ($\sim$ means "of the order").

they appear in great profusion and we lack a complete understanding of their interrelationship, although we already know that there are some very surprising relationships among them. We seem gradually to be groping toward an understanding of the world of subatomic particles, but we really do not know how far we have yet to go in this task.

3

THE RELATION OF PHYSICS TO OTHER SCIENCES

Introduction

Physics is the most fundamental and all-inclusive of the sciences, and has had a profound effect on all scientific development. In fact, physics is the present-day equivalent of what used to be called *natural philosophy,* from which most of our modern sciences arose. Students of many fields find themselves studying physics because of the basic role it plays in all phenomena. In this chapter we shall try to explain what the fundamental problems in the other sciences are, but of course it is impossible in so small a space really to deal with the complex, subtle, beautiful matters in these other fields. Lack of space also prevents our discussing the relation of physics to engineering, industry, society, and war, or even the most remarkable relationship between mathematics and physics. (Mathematics is not a science from our point of view, in the sense that it is not a *natural* science. The test of its validity is not experiment.) We must, incidentally, make it clear from the beginning that if a thing is not a science, it is not necessarily bad. For example, love is not a science. So, if something is said not to be a science, it does not mean that there is something wrong with it; it just means that it is not a science.

Chemistry

The science which is perhaps the most deeply affected by physics is chemistry. Historically, the early days of chemistry dealt almost entirely with what we now call inorganic chemistry, the chemistry of substances which are not associated with living things. Considerable analysis was required to discover the existence of the many elements and their relationships—how they make the various relatively simple compounds found in rocks, earth, etc. This early chemistry was very important for physics. The interaction between the two sciences was very great because the theory of atoms was substantiated to a large extent by experiments in chemistry. The theory of chemistry, i.e., of the reactions themselves, was summarized to a large extent in the periodic chart of Mendeléev, which brings out many strange relationships among the various elements, and it was the collection of rules as to which substance is combined with which, and how, that constituted inorganic chemistry. All these rules were ultimately explained in principle by quantum mechanics, so that theoretical chemistry is in fact physics. On the other hand, it must be emphasized that this explanation is *in principle.* We have already discussed the difference between knowing the rules of the game of chess and being able to play. So it is that we may know the rules, but we cannot play very well. It turns out to be very difficult to predict precisely what will happen in a given chemical reaction; nevertheless, the deepest part of theoretical chemistry must end up in quantum mechanics.

There is also a branch of physics and chemistry which was developed by both sciences together, and which is extremely important. This is the method of statistics applied in a situation in which there are mechanical laws, which is aptly called *statistical mechanics.* In any chemical situation a large number of atoms are involved, and we have seen that the atoms are all jiggling around in a very random and complicated way. If we could analyze each collision, and be able to follow in detail the motion of each molecule, we

might hope to figure out what would happen, but the many numbers needed to keep track of all these molecules exceed so enormously the capacity of any computer, and certainly the capacity of the mind, that it was important to develop a method for dealing with such complicated situations. Statistical mechanics, then, is the science of the phenomena of heat, or thermodynamics. Inorganic chemistry is, as a science, now reduced essentially to what are called physical chemistry and quantum chemistry: physical chemistry to study the rates at which reactions occur and what is happening in detail (How do the molecules hit? Which pieces fly off first?, etc.), and quantum chemistry to help us understand what happens in terms of the physical laws.

The other branch of chemistry is *organic chemistry,* the chemistry of the substances which are associated with living things. For a time it was believed that the substances which are associated with living things were so marvelous that they could not be made by hand, from inorganic materials. This is not at all true—they are just the same as the substances made in inorganic chemistry, but more complicated arrangements of atoms are involved. Organic chemistry obviously has a very close relationship to the biology which supplies its substances, and to industry, and furthermore, much physical chemistry and quantum mechanics can be applied to organic as well as to inorganic compounds. However, the main problems of organic chemistry are not in these aspects, but rather in the analysis and synthesis of the substances which are formed in biological systems, in living things. This leads imperceptibly, in steps, toward biochemistry, and then into biology itself, or molecular biology.

Biology

Thus we come to the science of *biology,* which is the study of living things. In the early days of biology, the biologists had to deal with the purely descriptive problem of finding out *what* living things there were, and so they just had to count such things as the hairs

of the limbs of fleas. After these matters were worked out with a great deal of interest, the biologists went into the *machinery* inside the living bodies, first from a gross standpoint, naturally, because it takes some effort to get into the finer details.

There was an interesting early relationship between physics and biology in which biology helped physics in the discovery of the *conservation of energy,* which was first demonstrated by Mayer in connection with the amount of heat taken in and given out by a living creature.

If we look at the processes of biology of living animals more closely, we see *many* physical phenomena: the circulation of blood, pumps, pressure, etc. There are nerves: we know what is happening when we step on a sharp stone, and that somehow or other the information goes from the leg up. It is interesting how that happens. In their study of nerves, the biologists have come to the conclusion that nerves are very fine tubes with a complex wall which is very thin; through this wall the cell pumps ions, so that there are positive ions on the outside and negative ions on the inside, like a capacitor. Now this membrane has an interesting property; if it "discharges" in one place, i.e., if some of the ions were able to move through one place, so that the electric voltage is reduced there, that electrical influence makes itself felt on the ions in the neighborhood, and it affects the membrane in such a way that it lets the ions through at neighboring points also. This in turn affects it farther along, etc., and so there is a wave of "penetrability" of the membrane which runs down the fiber when it is "excited" at one end by stepping on the sharp stone. This wave is somewhat analogous to a long sequence of vertical dominoes; if the end one is pushed over, that one pushes the next, etc. Of course this will transmit only one message unless the dominoes are set up again; and similarly in the nerve cell, there are processes which pump the ions slowly out again, to get the nerve ready for the next impulse. So it is that we know what we are doing (or at least where we are). Of course the electrical effects associated with

this nerve impulse can be picked up with electrical instruments, and because there *are* electrical effects, obviously the physics of electrical effects has had a great deal of influence on understanding the phenomenon.

The opposite effect is that, from somewhere in the brain, a message is sent out along a nerve. What happens at the end of the nerve? There the nerve branches out into fine little things, connected to a structure near a muscle, called an endplate. For reasons which are not exactly understood, when the impulse reaches the end of the nerve, little packets of a chemical called acetylcholine are shot off (five or ten molecules at a time) and they affect the muscle fiber and make it contract—how simple! What makes a muscle contract? A muscle is a very large number of fibers close together, containing two different substances, myosin and actomyosin, but the machinery by which the chemical reaction induced by acetylcholine can modify the dimensions of the muscle is not yet known. Thus the fundamental processes in the muscle that make mechanical motions are not known.

Biology is such an enormously wide field that there are hosts of other problems that we cannot mention at all—problems on how vision works (what the light does in the eye), how hearing works, etc. (The way in which *thinking* works we shall discuss later under psychology.) Now, these things concerning biology which we have just discussed are, from a biological standpoint, really not fundamental, at the bottom of life, in the sense that even if we understood them we still would not understand life itself. To illustrate: the men who study nerves feel their work is very important, because after all you cannot have animals without nerves. But you *can* have *life* without nerves. Plants have neither nerves nor muscles, but they are working, they are alive, just the same. So for the fundamental problems of biology we must look deeper; when we do, we discover that all living things have a great many characteristics in common. The most common feature is that they are made of *cells,* within each of which is complex machinery for doing things chemically.

In plant cells, for example, there is machinery for picking up light and generating sucrose, which is consumed in the dark to keep the plant alive. When the plant is eaten the sucrose itself generates in the animal a series of chemical reactions very closely related to photosynthesis (and its opposite effect in the dark) in plants.

In the cells of living systems there are many elaborate chemical reactions, in which one compound is changed into another and another. To give some impression of the enormous efforts that have gone into the study of biochemistry, the chart in Fig. 3-1 summarizes our knowledge to date on just one small part of the many series of reactions which occur in cells, perhaps a percent or so of it.

Here we see a whole series of molecules which change from one to another in a sequence or cycle of rather small steps. It is called the Krebs cycle, the respiratory cycle. Each of the chemicals and each of the steps is fairly simple, in terms of what change is made in the molecule, but—and this is a centrally important discovery in biochemistry—these changes are *relatively difficult to accomplish in a laboratory.* If we have one substance and another very similar substance, the one does not just turn into the other, because the two forms are usually separated by an energy barrier or "hill." Consider this analogy: If we wanted to take an object from one place to another, at the same level but on the other side of a hill, we could push it over the top, but to do so requires the addition of some energy. Thus most chemical reactions do not occur, because there is what is called an *activation energy* in the way. In order to add an extra atom to our chemical requires that we get it *close* enough that some rearrangement can occur; then it will stick. But if we cannot give it enough energy to get it close enough, it will not go to completion it will just go partway up the "hill" and back down again. However, if we could literally take the molecules in our hands and push and pull the atoms around in such a way as to open a hole to let the new atom in, and then let it snap back, we would have found another way, *around* the hill, which would not require extra energy,

Figure 3-1 The Krebs cycle.

and the reaction would go easily. Now there actually *are*, in the cells, *very* large molecules, much larger than the ones whose changes we have been describing, which in some complicated way hold the smaller molecules just right, so that the reaction can occur easily. These very large and complicated things are called *enzymes*. (They were first called ferments, because they were originally discovered in the fermentation of sugar. In fact, some of the first reactions in the cycle were discovered there.) In the presence of an enzyme the reaction will go.

An enzyme is made of another substance called *protein*. Enzymes are very big and complicated, and each one is different, each being built to control a certain special reaction. The names of the enzymes are written in Fig. 3-1 at each reaction. (Sometimes the same enzyme may control two reactions.) We emphasize that the enzymes themselves are not involved in the reaction directly. They do not

change; they merely let an atom go from one place to another. Having done so, the enzyme is ready to do it to the next molecule, like a machine in a factory. Of course, there must be a supply of certain atoms and a way of disposing of other atoms. Take hydrogen, for example: there are enzymes which have special units on them which carry the hydrogen for all chemical reactions. For example, there are three or four hydrogen-reducing enzymes which are used all over our cycle in different places. It is interesting that the machinery which liberates some hydrogen at one place will take that hydrogen and use it somewhere else.

The most important feature of the cycle of Fig. 3-1 is the transformation from GDP to GTP (guanosine-di-phosphate to guanosine-tri-phosphate) because the one substance has much more energy in it than the other. Just as there is a "box" in certain enzymes for carrying hydrogen atoms around, there are special *energy*-carrying "boxes" which involve the triphosphate group. So, GTP has more energy than GDP and if the cycle is going one way, we are producing molecules which have extra energy and which can go drive some other cycle which *requires* energy, for example the contraction of muscle. The muscle will not contract unless there is GTP. We can take muscle fiber, put it in water, and add GTP, and the fibers contract, changing GTP to GDP if the right enzymes are present. So the real system is in the GDP-GTP transformation; in the dark the GTP which has been stored up during the day is used to run the whole cycle around the other way. An enzyme, you see, does not care in which direction the reaction goes, for if it did it would violate one of the laws of physics.

Physics is of great importance in biology and other sciences for still another reason, that has to do with *experimental techniques.* In fact, if it were not for the great development of experimental physics, these biochemistry charts would not be known today. The reason is that the most useful tool of all for analyzing this fantastically complex system is to *label* the atoms which are used in the reactions. Thus, if we could introduce into the cycle some

carbon dioxide which has a "green mark" on it, and then measure after three seconds where the green mark is, and again measure after ten seconds, etc., we could trace out the course of the reactions. What are the "green marks"? They are different *isotopes.* We recall that the chemical properties of atoms are determined by the number of *electrons,* not by the mass of the nucleus. But there can be, for example in carbon, six neutrons or seven neutrons, together with the six protons which all carbon nuclei have. Chemically, the two atoms C^{12} and C^{13} are the same, but they differ in weight and they have different nuclear properties, and so they are distinguishable. By using these isotopes of different weights, or even radioactive isotopes like C^{14}, which provide a more sensitive means for tracing very small quantities, it is possible to trace the reactions.

Now, we return to the description of enzymes and proteins. All proteins are not enzymes, but all enzymes are proteins. There are many proteins, such as the proteins in muscle, the structural proteins which are, for example, in cartilage and hair, skin, etc., that are not themselves enzymes. However, proteins are a very characteristic substance of life: first of all they make up all the enzymes, and second, they make up much of the rest of living material. Proteins have a very interesting and simple structure. They are a series, or chain, of different *amino acids.* There are twenty different amino acids, and they all can combine with each other to form chains in which the backbone is CO-NH, etc. Proteins are nothing but chains of various ones of these twenty amino acids. Each of the amino acids probably serves some special purpose. Some, for example, have a sulfur atom at a certain place; when two sulfur atoms are in the same protein, they form a bond, that is, they tie the chain together at two points and form a loop. Another has extra oxygen atoms which make it an acidic substance; another has a basic characteristic. Some of them have big groups hanging out to one side, so that they take up a lot of space. One of the amino acids, called proline, is not really an amino acid, but imino acid. There is a slight

difference, with the result that when proline is in the chain, there is a kink in the chain. If we wished to manufacture a particular protein, we would give these instructions: put one of those sulfur hooks here; next, add something to take up space; then attach something to put a kink in the chain. In this way, we will get a complicated-looking chain, hooked together and having some complex structure; this is presumably just the manner in which all the various enzymes are made. One of the great triumphs in recent times (since 1960) was at last to discover the exact spatial atomic arrangement of certain proteins, which involve some fifty-six or sixty amino acids in a row. Over a thousand atoms (more nearly two thousand, if we count the hydrogen atoms) have been located in a complex pattern in two proteins. The first was hemoglobin. One of the sad aspects of this discovery is that we cannot see anything from the pattern; we do not understand why it works the way it does. Of course, that is the next problem to be attacked.

Another problem is how do the enzymes know what to be? A red-eyed fly makes a red-eyed fly baby, and so the information for the whole pattern of enzymes to make red pigment must be passed from one fly to the next. This is done by a substance in the nucleus of the cell, not a protein, called DNA (short for deoxyribonucleic acid). This is the key substance which is passed from one cell to another (for instance, sperm cells consist mostly of DNA) and carries the information as to how to make the enzymes. DNA is the "blueprint." What does the blueprint look like and how does it work? First, the blueprint must be able to reproduce itself. Secondly, it must be able to instruct the protein. Concerning the reproduction, we might think that this proceeds like cell reproduction. Cells simply grow bigger and then divide in half. Must it be thus with DNA molecules, then, that they too grow bigger and divide in half? Every *atom* certainly does not grow bigger and divide in half! No, it is impossible to reproduce a molecule except by some more clever way.

The structure of the substance DNA was studied for a long time, first chemically to find the composition, and then with x-rays

to find the pattern in space. The result was the following remarkable discovery: The DNA molecule is a pair of chains, twisted upon each other. The backbone of each of these chains, which are analogous to the chains of proteins but chemically quite different, is a series of sugar and phosphate groups, as shown in Fig. 3-2. Now we see how the chain can contain instructions, for if we could split this chain down the middle, we would have a series *BAADC* . . . and every living thing could have a different series. Thus perhaps, in some way, the specific *instructions* for the manufacture of proteins are contained in the specific *series* of the DNA.

Attached to each sugar along the line, and linking the two chains together, are certain pairs of cross-links. However, they are not all of the same kind; there are four kinds, called adenine, thymine, cytosine, and guanine, but let us call them *A, B, C,* and *D*. The interesting thing is that only certain pairs can sit opposite each other, for example *A* with *B* and *C* with *D.* These pairs are put on the two chains in such a way that they "fit together," and have a strong energy of interaction. However, *C* will not fit with *A,* and *B* will not fit with *C*; they will only fit in pairs, *A* against *B* and *C* against *D*. Therefore if one is *C*, the other must be *D*, etc. Whatever the letters may be in one chain, each one must have its specific complementary letter on the other chain.

What then about reproduction? Suppose we split this chain in two. How can we make another one just like it? If, in the substances of the cells, there is a manufacturing department which brings up phosphate, sugar, and *A, B, C, D* units not connected in a chain, the only ones which will attach to our split chain will be the correct ones, the complements of *BAADC* . . . , namely, *ABBCD* . . . Thus what happens is that the chain splits down the middle during cell division, one half ultimately to go with one cell, the other half to end up in the other cell; when separated, a new complementary chain is made by each half-chain.

Next comes the question, precisely how does the order of the *A, B, C, D* units determine the arrangement of the amino acids in the

Figure 3-2 Schematic diagram of DNA.

protein? This is the central unsolved problem in biology today. The first clues, or pieces of information, however, are these: There are in the cell tiny particles called ribosomes, and it is now known that that is the place where proteins are made. But the ribosomes are not in the nucleus, where the DNA and its instructions are. Something seems to be the matter. However, it is also known that little molecule pieces come off the DNA—not as long as the big DNA molecule that carries all the information itself, but like a small section of it. This is called RNA, but that is not essential. It is a kind of copy of the DNA, a short copy. The RNA, which somehow car-

ries a message as to what kind of protein to make goes over to the ribosome; that is known. When it gets there, protein is synthesized at the ribosome. That is also known. However, the details of how the amino acids come in and are arranged in accordance with a code that is on the RNA are, as yet, still unknown. We do not know how to read it. If we knew, for example, the "lineup" *A, B, C, C, A,* we could not tell you what protein is to be made.

Certainly no subject or field is making more progress on so many fronts at the present moment than biology, and if we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that *all things are made of atoms,* and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms.

Astronomy

In this rapid-fire explanation of the whole world, we must now turn to astronomy. Astronomy is older than physics. In fact, it got physics started by showing the beautiful simplicity of the motion of the stars and planets, the understanding of which was the *beginning* of physics. But the most remarkable discovery in all of astronomy is that *the stars are made of atoms of the same kind as those on the earth*.*

---

* How I'm rushing through this! How much each sentence in this brief story contains. "The stars are made of the same atoms as the earth." I usually pick one small topic like this to give a lecture on. Poets say science takes away from the beauty of the stars—mere globs of gas atoms. Nothing is "mere." I too can see the stars on a desert night, and feel them. But do I see less or more? The vastness of the heavens stretches my imagination—stuck on this carousel my little eye can catch one-million-year-old light. A vast pattern—of which I am a part—perhaps my stuff was belched from some forgotten star, as one is belching there. Or see them with the greater eye of Palomar, rushing all apart from some common starting point when they were perhaps all together. What is the pattern, or the meaning, or the *why*? It does not do harm to the mystery to know a little about it. For far more

How was this done? Atoms liberate light which has definite frequencies, something like the timbre of a musical instrument, which has definite pitches or frequencies of sound. When we are listening to several different tones we can tell them apart, but when we look with our eyes at a mixture of colors we cannot tell the parts from which it was made, because the eye is nowhere near as discerning as the ear in this connection. However, with a spectroscope we *can* analyze the frequencies of the light waves and in this way we can see the very tunes of the atoms that are in the different stars. As a matter of fact, two of the chemical elements were discovered on a star before they were discovered on the earth. Helium was discovered on the sun, whence its name, and technetium was discovered in certain cool stars. This, of course, permits us to make headway in understanding the stars, because they are made of the same kinds of atoms which are on the earth. Now we know a great deal about the atoms, especially concerning their behavior under conditions of high temperature but not very great density, so that we can analyze by statistical mechanics the behavior of the stellar substance. Even though we cannot reproduce the conditions on the earth, using the basic physical laws we often can tell precisely, or very closely, what will happen. So it is that physics aids astronomy. Strange as it may seem, we understand the distribution of matter in the interior of the sun far better than we understand the interior of the earth. What goes on *inside* a star is better understood than one might guess from the difficulty of having to look at a little dot of light through a telescope, because we can *calculate* what the atoms in the stars should do in most circumstances.

One of the most impressive discoveries was the origin of the en-

---

marvelous is the truth than any artists of the past imagined! Why do the poets of the present not speak of it? What men are poets who can speak of Jupiter if he were like a man, but if he is an immense spinning sphere of methane and ammonia must be silent?

ergy of the stars, that makes them continue to burn. One of the men who discovered this was out with his girlfriend the night after he realized that *nuclear reactions* must be going on in the stars in order to make them shine. She said, "Look at how pretty the stars shine!" He said, "Yes, and right now I am the only man in the world who knows *why* they shine." She merely laughed at him. She was not impressed with being out with the only man who, at that moment, knew why stars shine. Well, it is sad to be alone, but that is the way it is in this world.

It is the nuclear "burning" of hydrogen which supplies the energy of the sun; the hydrogen is converted into helium. Furthermore, ultimately, the manufacture of various chemical elements proceeds in the centers of the stars, from hydrogen. The stuff of which *we* are made was "cooked" once, in a star, and spit out. How do we know? Because there is a clue. The proportion of the different isotopes—how much C^{12}, how much C^{13}, etc., is something which is never changed by *chemical* reactions, because the chemical reactions are so much the same for the two. The proportions are purely the result of *nuclear* reactions. By looking at the proportions of the isotopes in the cold, dead ember which we are, we can discover what the *furnace* was like in which the stuff of which we are made was formed. That furnace was like the stars, and so it is very likely that our elements were "made" in the stars and spit out in the explosions which we call novae and supernovae. Astronomy is so close to physics that we shall study many astronomical things as we go along.

Geology

We turn now to what are called *earth sciences,* or *geology.* First, meteorology and the weather. Of course the *instruments* of meteorology are physical instruments, and the development of experimental physics made these instruments possible, as was explained before. However, the theory of meteorology has never been satisfactorily

worked out by the physicist. "Well," you say, "there is nothing but air, and we know the equations of the motions of air." Yes we do. "So if we know the condition of air today, why can't we figure out the condition of the air tomorrow?" First, we do not *really* know what the condition is today, because the air is swirling and twisting everywhere. It turns out to be very sensitive, and even unstable. If you have ever seen water run smoothly over a dam, and then turn into a large number of blobs and drops as it falls, you will understand what I mean by unstable. You know the condition of the water before it goes over the spillway; it is perfectly smooth; but the moment it begins to fall, where do the drops begin? What determines how big the lumps are going to be and where they will be? That is not known, because the water is unstable. Even a smooth moving mass of air going over a mountain turns into complex whirlpools and eddies. In many fields we find this situation of *turbulent flow* that we cannot analyze today. Quickly we leave the subject of weather, and discuss geology!

The question basic to geology is, what makes the earth the way it is? The most obvious processes are in front of your very eyes, the erosion processes of the rivers, the winds, etc. It is easy enough to understand these, but for every bit of erosion there is an equal amount of something else going on. Mountains are no lower today, on the average, than they were in the past. There must be mountain-*forming* processes. You will find, if you study geology, that there *are* mountain-forming processes and volcanism, which nobody understands but which is half of geology. The phenomenon of volcanoes is really not understood. What makes an earthquake is, ultimately, not understood. It is understood that if something is pushing something else, it snaps and will slide—that is all right. But what pushes, and why? The theory is that there are currents inside the earth—circulating currents, due to the difference in temperature inside and outside—which, in their motion, push the surface slightly. Thus if there are two opposite circulations next to each other, the matter will collect in

the region where they meet and make belts of mountains which are in unhappy stressed conditions, and so produce volcanoes and earthquakes.

What about the inside of the earth? A great deal is known about the speed of earthquake waves through the earth and the density of distribution of the earth. However, physicists have been unable to get a good theory as to how dense a substance should be at the pressures that would be expected at the center of the earth. In other words, we cannot figure out the properties of matter very well in these circumstances. We do much less well with the earth than we do with the conditions of matter in the stars. The mathematics involved seems a little too difficult, so far, but perhaps it will not be too long before someone realizes that it is an important problem, and really works it out. The other aspect, of course, is that even if we did know the density, we cannot figure out the circulating currents. Nor can we really work out the properties of rocks at high pressure. We cannot tell how fast the rocks should "give"; that must all be worked out by experiment.

Psychology

Next, we consider the science of *psychology.* Incidentally, psychoanalysis is not a science: it is at best a medical process, and perhaps even more like witch-doctoring. It has a theory as to what causes disease—lots of different "spirits," etc. The witch doctor has a theory that a disease like malaria is caused by a spirit which comes into the air; it is not cured by shaking a snake over it, but quinine does help malaria. So, if you are sick, I would advise that you go to the witch doctor because he is the man in the tribe who knows the most about the disease; on the other hand, his knowledge is not science. Psychoanalysis has not been checked carefully by experiment, and there is no way to find a list of the number of cases in which it works, the number of cases in which it does not work, etc.

The other branches of psychology, which involve things like the physiology of sensation—what happens in the eye, and what happens in the brain—are, if you wish, less interesting. But some small but real progress has been made in studying them. One of the most interesting technical problems may or may not be called psychology. The central problem of the mind, if you will, or the nervous system, is this: when an animal learns something, it can do something different than it could before, and its brain cell must have changed too, if it is made out of atoms. *In what way is it different?* We do not know where to look, or what to look for, when something is memorized. We do not know what it means, or what change there is in the nervous system, when a fact is learned. This is a very important problem which has not been solved at all. Assuming, however, that there is some kind of memory thing, the brain is such an enormous mass of interconnecting wires and nerves that it probably cannot be analyzed in a straightforward manner. There is an analog of this to computing machines and computing elements, in that they also have a lot of lines, and they have some kind of element, analogous, perhaps, to the synapse, or connection of one nerve to another. This is a very interesting subject which we have not the time to discuss further—the relationship between thinking and computing machines. It must be appreciated, of course, that this subject will tell us very little about the real complexities of ordinary human behavior. All human beings are so different. It will be a long time before we get there. We must start much further back. If we could even figure out how a *dog* works, we would have gone pretty far. Dogs are easier to understand, but nobody yet knows how dogs work.

How did it get that way?

In order for physics to be useful to other sciences in a *theoretical* way, other than in the invention of instruments, the science in question must supply to the physicist a description of the object in

a physicist's language. They can say "why does a frog jump?," and the physicist cannot answer. If they tell him what a frog is, that there are so many molecules, there is a nerve here, etc., that is different. If they will tell us, more or less, what the earth or the stars are like, then we can figure it out. In order for physical theory to be of any use, we must know where the atoms are located. In order to understand the chemistry, we must know exactly what atoms are present, for otherwise we cannot analyze it. That is but one limitation, of course.

There is another *kind* of problem in the sister sciences which does not exist in physics; we might call it, for lack of a better term, the historical question. How did it get that way? If we understand all about biology, we will want to know how all the things which are on the earth got there. There is the theory of evolution, an important part of biology. In geology, we not only want to know how the mountains are forming, but how the entire earth was formed in the beginning, the origin of the solar system, etc. That, of course, leads us to want to know what kind of matter there was in the world. How did the stars evolve? What were the initial conditions? That is the problem of astronomical history. A great deal has been found out about the formation of stars, the formation of elements from which we were made, and even a little about the origin of the universe.

There is no historical question being studied in physics at the present time. We do not have a question, "Here are the laws of physics, how did they get that way?" We do not imagine, at the moment, that the laws of physics are somehow changing with time, that they were different in the past than they are at present. Of course they *may* be, and the moment we find they *are,* the historical question of physics will be wrapped up with the rest of the history of the universe, and then the physicist will be talking about the same problems as astronomers, geologists, and biologists.

Finally, there is a physical problem that is common to many fields, that is very old, and that has not been solved. It is not the

problem of finding new fundamental particles, but something left over from a long time ago—over a hundred years. Nobody in physics has really been able to analyze it mathematically satisfactorily in spite of its importance to the sister sciences. It is the analysis of *circulating or turbulent fluids.* If we watch the evolution of a star, there comes a point where we can deduce that it is going to start convection, and thereafter we can no longer deduce what should happen. A few million years later the star explodes, but we cannot figure out the reason. We cannot analyze the weather. We do not know the patterns of motions that there should be inside the earth. The simplest form of the problem is to take a pipe that is very long and push water through it at high speed. We ask: to push a given amount of water through that pipe, how much pressure is needed? No one can analyze it from first principles and the properties of water. If the water flows very slowly, or if we use a thick goo like honey, then we can do it nicely. You will find that in your textbook. What we really cannot do is deal with actual, wet water running through a pipe. That is the central problem which we ought to solve someday, and we have not.

A poet once said, "The whole universe is in a glass of wine." We will probably never know in what sense he meant that, for poets do not write to be understood. But it is true that if we look at a glass of wine closely enough we see the entire universe. There are the things of physics: the twisting liquid which evaporates depending on the wind and weather, the reflections in the glass, and our imagination adds the atoms. The glass is a distillation of the earth's rocks, and in its composition we see the secrets of the universe's age, and the evolution of stars. What strange array of chemicals are in the wine? How did they come to be? There are the ferments, the enzymes, the substrates, and the products. There in wine is found the great generalization: all life is fermentation. Nobody can discover the chemistry of wine without discovering, as did Louis Pasteur, the cause of much disease. How vivid is the claret, pressing its existence into the consciousness that watches it! If our small

minds, for some convenience, divide this glass of wine, this universe, into parts—physics, biology, geology, astronomy, psychology, and so on—remember that nature does not know it! So let us put it all back together, not forgetting ultimately what it is for. Let it give us one more final pleasure: drink it and forget it all!

4

CONSERVATION OF ENERGY

What is energy?

In this chapter, we begin our more detailed study of the different aspects of physics, having finished our description of things in general. To illustrate the ideas and the kind of reasoning that might be used in theoretical physics, we shall now examine one of the most basic laws of physics, the conservation of energy.

There is a fact, or if you wish, a *law,* governing all natural phenomena that are known to date. There is no known exception to this law—it is exact so far as we know. The law is called the *conservation of energy.* It states that there is a certain quantity, which we call energy, that does not change in the manifold changes which nature undergoes. That is a most abstract idea, because it is a mathematical principle; it says that there is a numerical quantity which does not change when something happens. It is not a description of a mechanism, or anything concrete; it is just a strange fact that we can calculate some number and when we finish watching nature go through her tricks and calculate the number again, it is the same. (Something like the bishop on a red square, and after a number of moves—details unknown—it is still on some red square. It is a law

of this nature.) Since it is an abstract idea, we shall illustrate the meaning of it by an analogy.

Imagine a child, perhaps "Dennis the Menace," who has blocks which are absolutely indestructible, and cannot be divided into pieces. Each is the same as the other. Let us suppose that he has 28 blocks. His mother puts him with his 28 blocks into a room at the beginning of the day. At the end of the day, being curious, she counts the blocks very carefully, and discovers a phenomenal law—no matter what he does with the blocks, there are always 28 remaining! This continues for a number of days, until one day there are only 27 blocks, but a little investigating shows that there is one under the rug—she must look everywhere to be sure that the number of blocks has not changed. One day, however, the number appears to change—there are only 26 blocks. Careful investigation indicates that the window was open, and upon looking outside, the other two blocks are found. Another day, careful count indicates that there are 30 blocks! This causes considerable consternation, until it is realized that Bruce came to visit, bringing his blocks with him, and he left a few at Dennis's house. After she has disposed of the extra blocks, she closes the window, does not let Bruce in, and then everything is going along all right, until one time she counts and finds only 25 blocks. However, there is a box in the room, a toy box, and the mother goes to open the toy box, but the boy says, "No, do not open my toy box," and screams. Mother is not allowed to open the toy box. Being extremely curious, and somewhat ingenious, she invents a scheme! She knows that a block weighs three ounces, so she weighs the box at a time when she sees 28 blocks, and it weighs 16 ounces. The next time she wishes to check, she weighs the box again, subtracts 16 ounces, and divides by 3. She discovers the following:

$$\begin{pmatrix}\text{number of}\\ \text{blocks seen}\end{pmatrix} + \frac{(\text{weight of box}) - 16\text{ ounces}}{3\text{ ounces}} = \text{constant.} \quad (4.1)$$

There then appear to be some new deviations, but careful study indicates that the dirty water in the bathtub is changing its level. The child is throwing blocks into the water, and she cannot see them because it is so dirty, but she can find out how many blocks are in the water by adding another term to her formula. Since the original height of the water was 6 inches and each block raises the water a quarter of an inch, this new formula would be:

$$\begin{pmatrix}\text{number of}\\ \text{blocks seen}\end{pmatrix} + \frac{(\text{weight of box}) - 16 \text{ ounces}}{3 \text{ ounces}} + \frac{(\text{height of water}) - 6 \text{ inches}}{1/4 \text{ inch}} = \text{constant.} \tag{4.2}$$

In the gradual increase in the complexity of her world, she finds a whole series of terms representing ways of calculating how many blocks are in places where she is not allowed to look. As a result, she finds a complex formula, a quantity which *has to be computed,* which always stays the same in her situation.

What is the analogy of this to the conservation of energy? The most remarkable aspect that must be abstracted from this picture is that *there are no blocks.* Take away the first terms in (4.1) and (4.2) and we find ourselves calculating more or less abstract things. The analogy has the following points: First, when we are calculating the energy, sometimes some of it leaves the system and goes away, or sometimes some comes in. In order to verify the conservation of energy, we must be careful that we have not put any in or taken any out. Second, the energy has a large number of *different forms,* and there is a formula for each one. These are gravitational energy, kinetic energy, heat energy, elastic energy, electrical energy, chemical energy, radiant energy, nuclear energy, mass energy. If we total up the formulas for each of these contributions, it will not change except for energy going in and out.

It is important to realize that in physics today, we have no knowledge of what energy *is.* We do not have a picture that energy comes

in little blobs of a definite amount. It is not that way. However, there are formulas for calculating some numerical quantity, and when we add it all together it gives "28"—always the same number. It is an abstract thing in that it does not tell us the mechanism or the *reasons* for the various formulas.

Gravitational potential energy

Conservation of energy can be understood only if we have the formula for all of its forms. I wish to discuss the formula for gravitational energy near the surface of the earth, and I wish to derive this formula in a way which has nothing to do with history but is simply a line of reasoning invented for this particular lecture to give you an illustration of the remarkable fact that a great deal about nature can be extracted from a few facts and close reasoning. It is an illustration of the kind of work theoretical physicists become involved in. It is patterned after a most excellent argument by Mr. Carnot on the efficiency of steam engines.*

Consider weight-lifting machines—machines which have the property that they lift one weight by lowering another. Let us also make a hypothesis: that *there is no such thing as perpetual motion* with these weight-lifting machines. (In fact, that there is no perpetual motion at all is a general statement of the law of conservation of energy.) We must be careful to define perpetual motion. First, let us do it for weight-lifting machines. If, when we have lifted and lowered a lot of weights and restored the machine to the original condition, we find that the net result is to have *lifted a weight,* then we have a perpetual motion machine because we can use that lifted weight to run something else. That is, *provided* the machine which lifted the weight is brought back to its exact *original condition,* and furthermore that it is completely *self-contained*—that it has not received the energy to lift that weight from some external source—like Bruce's blocks.

A very simple weight-lifting machine is shown in Fig. 4-1. This

* Our point here is not so much the result, (4.3), which in fact you may already know, as the possibility of arriving at it by theoretical reasoning.

Figure 4-1 Simple weight-lifting machine.

machine lifts weights three units "strong." We place three units on one balance pan, and one unit on the other. However, in order to get it actually to work, we must lift a little weight off the left pan. On the other hand, we could lift a one-unit weight by lowering the three-unit weight, if we cheat a little by lifting a little weight off the other pan. Of course, we realize that with any *actual* lifting machine, we must add a little extra to get it to run. This we disregard, *temporarily.* Ideal machines, although they do not exist, do not require anything extra. A machine that we actually use can be, in a sense, *almost* reversible: that is, if it will lift the weight of three by lowering a weight of one, then it will also lift nearly the weight of one the same amount by lowering the weight of three.

We imagine that there are two classes of machines, those that are *not* reversible, which includes all real machines, and those that *are* reversible, which of course are actually not attainable no matter how careful we may be in our design of bearings, levers, etc. We suppose, however, that there is such a thing—a reversible machine—which lowers one unit of weight (a pound or any other unit) by one unit of distance, and at the same time lifts a three-unit weight. Call this reversible machine Machine *A.* Suppose this particular reversible machine lifts the three-unit weight a distance *X.* Then suppose we have another machine, Machine *B,* which is not necessarily reversible, which also lowers a unit weight a unit distance, but which lifts three units a distance *Y.* We can now prove that *Y* is not higher than *X;* that is, it is impossible to build a machine that will lift a weight *any higher* than it will be lifted by a reversible machine. Let us see why. Let us suppose that *Y* was higher than *X.* We take a one-unit weight and lower it one unit height with Machine *B,* and that lifts the three-unit weight up a distance *Y.* Then we could lower

the weight from *Y* to *X*, *obtaining free power,* and use the reversible Machine *A,* running backwards, to lower the three-unit weight a distance *X* and lift the one-unit weight by one unit height. This will put the one-unit weight back where it was before, and leave both machines ready to be used again! We would therefore have perpetual motion if *Y* were higher than *X,* which we assumed was impossible. With those assumptions, we thus deduce that *Y is not higher than X,* so that of all machines that can be designed, the reversible machine is the best.

We can also see that all reversible machines must lift to *exactly the same height.* Suppose that *B* was really reversible also. The argument that *Y* is not higher than *X* is, of course, just as good as it was before, but we can also make our argument the other way around, using the machines in the opposite order, and prove that *X is not higher than Y.* This, then, is a very remarkable observation because it permits us to analyze the height to which different machines are going to lift something *without looking at the interior mechanism.* We know at once that if somebody makes an enormously elaborate series of levers that lift three units a certain distance by lowering one unit by one unit distance, and we compare it with a simple lever which does the same thing and is fundamentally reversible, his machine will lift it no higher, but perhaps less high. If his machine is reversible, we also know exactly *how* high it will lift. To summarize: every reversible machine, no matter how it operates, which drops one pound one foot and lifts a three-pound weight always lifts it the same distance, *X.* This is clearly a universal law of great utility. The next question is, of course, what is *X?*

Suppose we have a reversible machine which is going to lift this distance *X,* three for one. We set up three balls in a rack which does not move, as shown in Fig. 4-2. One ball is held on a stage at a distance one foot above the ground. The machine can lift three balls, lowering one by a distance 1. Now, we have arranged that the platform which holds three balls has a floor and two shelves, exactly spaced at distance *X,* and further, that the rack which holds the balls is spaced at distance *X,* (a). First we roll the balls horizontally

Figure 4-2 A reversible machine.

from the rack to the shelves, (b), and we suppose that this takes no energy because we do not change the height. The reversible machine then operates: it lowers the single ball to the floor, and it lifts the rack a distance *X,* (c). Now we have ingeniously arranged the rack so that these balls are again even with the platforms. Thus we unload the balls onto the rack, (d); having unloaded the balls, we can restore the machine to its original condition. Now we have three balls on the upper three shelves and one at the bottom. But the strange thing is that, in a certain way of speaking, we have not lifted *two* of them at all because, after all, there were balls on shelves 2 and 3 before. The resulting effect has been to lift *one ball* a distance 3*X*. Now, if 3*X* exceeds one foot, then we can *lower* the ball to return the machine to the initial condition, (f), and we can run the apparatus again. Therefore 3*X* cannot exceed one foot, for if 3*X* exceeds one foot we can make perpetual motion. Likewise, we can prove that *one foot cannot exceed* 3*X,* by making the whole machine run the opposite

way, since it is a reversible machine. Therefore $3X$ is neither *greater nor less than a foot,* and we discover then, by argument alone, the law that $X = {}^1/_3$ foot. The generalization is clear: one pound falls a certain distance in operating a reversible machine; then the machine can lift p pounds this distance divided by p. Another way of putting the result is that three pounds times the height lifted, which in our problem was X, is equal to one pound times the distance lowered, which is one foot in this case. If we take all the weights and multiply them by the heights at which they are now, above the floor, let the machine operate, and then multiply all the weights by all the heights again, *there will be no change.* (We have to generalize the example where we moved only one weight to the case where when we lower one, we lift several different ones—but that is easy.)

We call the sum of the weights times the heights *gravitational potential energy*—the energy which an object has because of its relationship in space, relative to the earth. The formula for gravitational energy, then, so long as we are not too far from the earth (the force weakens as we go higher) is

$$\begin{pmatrix}\text{gravitational}\\ \text{potential energy}\\ \text{for one object}\end{pmatrix} = (\text{weight}) \times (\text{height}). \tag{4.3}$$

It is a very beautiful line of reasoning. The only problem is that perhaps it is not true. (After all, nature does not *have* to go along with our reasoning.) For example, perhaps perpetual motion is, in fact, possible. Some of the assumptions may be wrong, or we may have made a mistake in reasoning, so it is always necessary to check. *It turns out experimentally,* in fact, to be true.

The general name of energy which has to do with location relative to something else is called *potential* energy. In this particular case, of course, we call it *gravitational potential energy.* If it is a question of electrical forces against which we are working, instead of gravitational forces, if we are "lifting" charges away from other charges with a lot of levers, then the energy content is called *elec-*

trical potential energy. The general principle is that the change in the energy is the force times the distance that the force is pushed, and that this is a change in energy in general:

$$\begin{pmatrix}\text{change in}\\ \text{energy}\end{pmatrix} = (\text{force}) \times \begin{pmatrix}\text{distance force}\\ \text{acts through}\end{pmatrix}. \tag{4.4}$$

We will return to many of these other kinds of energy as we continue the course.

The principle of the conservation of energy is very useful for deducing what will happen in a number of circumstances. In high school we learned a lot of laws about pulleys and levers used in different ways. We can now see that these "laws" are *all the same thing,* and that we did not have to memorize 75 rules to figure it out. A simple example is a smooth inclined plane which is, happily, a three-four-five triangle (Fig. 4-3). We hang a one-pound weight on the inclined plane with a pulley, and on the other side of the pulley, a weight *W.* We want to know how heavy *W* must be to balance the one pound on the plane. How can we figure that out? If we say it is just balanced, it is reversible and so can move up and down, and we can consider the following situation. In the initial circumstance, (a), the one-pound weight is at the bottom and weight *W* is at the top. When *W* has slipped down in a reversible way, we have a one-pound weight at the top and the weight *W* the slant distance, (b), or five feet, from the plane in which it was before. We *lifted* the one-pound weight only *three* feet and we lowered *W* pounds by *five* feet.

Figure 4-3 Inclined plane.

Therefore $W = {}^3/_5$ of a pound. Note that we deduced this from the *conservation of energy,* and not from force components. Cleverness, however, is relative. It can be deduced in a way which is even more brilliant, discovered by Stevinus and inscribed on his tombstone. Figure 4-4 explains that it has to be $^3/_5$ of a pound, because the chain does not go around. It is evident that the lower part of the chain is balanced by itself, so that the pull of the five weights on one side must balance the pull of three weights on the other, or whatever the ratio of the legs. You see, by looking at this diagram, that W must be $^3/_5$ of a pound. (If you get an epitaph like that on your gravestone, you are doing fine.)

Let us now illustrate the energy principle with a more complicated problem, the screw jack shown in Fig. 4-5. A handle 20 inches long is used to turn the screw, which has 10 threads to the inch. We would like to know how much force would be needed at the handle to lift one ton (2000 pounds). If we want to lift the ton 1 inch, say, then we must turn the handle around ten times. When it goes around once it goes approximately 126 inches. The handle must thus travel 1260 inches, and if we used various pulleys, etc., we would be lifting our one ton with an unknown smaller weight W applied to the end of the handle. So we find out that W is about 1.6 pounds. This is a result of the conservation of energy.

Figure 4-4 The epitaph of Stevinus.

Figure 4-5 A screw jack.

Take now the somewhat more complicated example shown in Fig. 4-6. A rod or bar, 8 feet long, is supported at one end. In the middle of the bar is a weight of 60 pounds, and at a distance of two feet from the support there is a weight of 100 pounds. How hard do we have to lift the end of the bar in order to keep it balanced, disregarding the weight of the bar? Suppose we put a pulley at one end and hang a weight on the pulley. How big would the weight W have to be in order for it to balance? We imagine that the weight falls any arbitrary distance—to make it easy for ourselves suppose it goes down 4 inches—how high would the two load weights rise? The center rises 2 inches, and the point a quarter of the way from the fixed end lifts 1 inch. Therefore, the principle that the sum of the heights times the weights does not change tells us that the weight W times 4 inches down, plus 60 pounds times 2 inches up, plus 100 pounds times 1 inch, has to add up to nothing:

$$-4W + (2)(60) + (1)(100) = 0, \qquad W = 55\,\text{lb}. \tag{4.5}$$

Thus we must have a 55-pound weight to balance the bar. In this way we can work out the laws of "balance"—the statics of complicated bridge arrangements, and so on. This approach is called the *principle of virtual work,* because in order to apply this argument we had to *imagine* that the structure moves a little—even though

Figure 4-6 Weighted rod supported on one end.

it is not *really* moving or even *movable.* We use the very small imagined motion to apply the principle of conservation of energy.

Kinetic energy

To illustrate another type of energy we consider a pendulum (Fig. 4-7). If we pull the mass aside and release it, it swings back and forth. In its motion, it loses height in going from either end to the center. Where does the potential energy go? Gravitational energy disappears when it is down at the bottom; nevertheless, it will climb up again. The gravitational energy must have gone into another form. Evidently it is by virtue of its *motion* that it is able to climb up again, so we have the conversion of gravitational energy into some other form when it reaches the bottom.

We must get a formula for the energy of motion. Now, recalling our arguments about reversible machines, we can easily see that in the motion at the bottom must be a quantity of energy which permits it to rise a certain height, and which has nothing to do with the *machinery* by which it comes up or the *path* by which it comes up. So we have an equivalence formula something like the one we wrote for the child's blocks. We have another form to represent the energy. It is easy to say what it is. The kinetic energy at the bottom equals the weight times the height that it could go, corresponding to its velocity: K.E. = *WH.* What we need is the formula which tells us the height by some rule that has to do with the motion of objects. If we start something out with a certain velocity, say, straight

Figure 4-7 Pendulum.

up, it will reach a certain height; we do not know what it is yet, but it depends on the velocity—there is a formula for that. Then to find the formula for kinetic energy for an object moving with velocity V, we must calculate the height that it could reach, and multiply by the weight. We shall soon find that we can write it this way:

$$\text{K.E.} = WV^2/2g. \tag{4.6}$$

Of course, the fact that motion has energy has nothing to do with the fact that we are in a gravitational field. It makes no difference *where* the motion came from. This is a general formula for various velocities. Both (4.3) and (4.6) are approximate formulas, the first because it is incorrect when the heights are great, i.e., when the heights are so high that gravity is weakening; the second, because of the relativistic correction at high speeds. However, when we do finally get the exact formula for the energy, then the law of conservation of energy is correct.

Other forms of energy

We can continue in this way to illustrate the existence of energy in other forms. First, consider elastic energy. If we pull down on a spring, we must do some work, for when we have it down, we can lift weights with it. Therefore in its stretched condition it has a possibility of doing some work. If we were to evaluate the sums of

weights times heights, it would not check out—we must add something else to account for the fact that the spring is under tension. Elastic energy is the formula for a spring when it is stretched. How much energy is it? If we let go, the elastic energy, as the spring passes through the equilibrium point, is converted to kinetic energy and it goes back and forth between compressing or stretching the spring and kinetic energy of motion. (There is also some gravitational energy going in and out, but we can do this experiment "sideways" if we like.) It keeps going until the losses—Aha! We have cheated all the way through by putting on little weights to move things or saying that the machines are reversible, or that they go on forever, but we can see that things do stop, eventually. Where is the energy when the spring has finished moving up and down? This brings in *another* form of energy: *heat energy.*

Inside a spring or a lever there are crystals which are made up of lots of atoms, and with great care and delicacy in the arrangement of the parts one can try to adjust things so that as something rolls on something else, none of the atoms do any jiggling at all. But one must be very careful. Ordinarily when things roll, there is bumping and jiggling because of the irregularities of the material, and the atoms start to wiggle inside. So we lose track of that energy; we find the atoms are wiggling inside in a random and confused manner after the motion slows down. There is still kinetic energy, all right, but it is not associated with visible motion. What a dream! How do we *know* there is still kinetic energy? It turns out that with thermometers you can find out that, in fact, the spring or the lever is *warmer*, and that there is really an increase of kinetic energy by a definite amount. We call this form of energy *heat energy*, but we know that it is not really a new form, it is just kinetic energy—internal motion. (One of the difficulties with all these experiments with matter that we do on a large scale is that we cannot really demonstrate the conservation of energy and we cannot really make our reversible machines, because every time we move a large clump of stuff, the atoms do not remain absolutely undisturbed, and so a

certain amount of random motion goes into the atomic system. We cannot see it, but we can measure it with thermometers, etc.)

There are many other forms of energy, and of course we cannot describe them in any more detail just now. There is electrical energy, which has to do with pushing and pulling by electric charges. There is radiant energy, the energy of light, which we know is a form of electrical energy because light can be represented as wigglings in the electromagnetic field. There is chemical energy, the energy which is released in chemical reactions. Actually, elastic energy is, to a certain extent, like chemical energy, because chemical energy is the energy of the attraction of the atoms, one for the other, and so is elastic energy. Our modern understanding is the following: chemical energy has two parts, kinetic energy of the electrons inside the atoms, so part of it is kinetic, and electrical energy of interaction of the electrons and the protons—the rest of it, therefore, is electrical. Next we come to nuclear energy, the energy which is involved with the arrangement of particles inside the nucleus, and we have formulas for that, but we do not have the fundamental laws. We know that it is not electrical, not gravitational, and not purely chemical, but we do not know what it is. It seems to be an additional form of energy. Finally, associated with the relativity theory, there is a modification of the laws of kinetic energy, or whatever you wish to call it, so that kinetic energy is combined with another thing called *mass energy.* An object has energy from its sheer *existence.* If I have a positron and an electron, standing still doing nothing—never mind gravity, never mind anything—and they come together and disappear, radiant energy will be liberated, in a definite amount, and the amount can be calculated. All we need know is the mass of the object. It does not depend on what it is—we make two things disappear, and we get a certain amount of energy. The formula was first found by Einstein; it is $E = mc^2$.

It is obvious from our discussion that the law of conservation of energy is enormously useful in making analyses, as we have illustrated in a few examples without knowing all the formulas. If we

had all the formulas for all kinds of energy, we could analyze how many processes should work without having to go into the details. Therefore conservation laws are very interesting. The question naturally arises as to what other conservation laws there are in physics. There are two other conservation laws which are analogous to the conservation of energy. One is called the conservation of linear momentum. The other is called the conservation of angular momentum. We will find out more about these later. In the last analysis, we do not understand the conservation laws deeply. We do not understand the conservation of energy. We do not understand energy as a certain number of little blobs. You may have heard that photons come out in blobs and that the energy of a photon is Planck's constant times the frequency. That is true, but since the frequency of light can be anything, there is no law that says that energy has to be a certain definite amount. Unlike Dennis's blocks, there can be any amount of energy, at least as presently understood. So we do not understand this energy as counting something at the moment, but just as a mathematical quantity, which is an abstract and rather peculiar circumstance. In quantum mechanics it turns out that the conservation of energy is very closely related to another important property of the world, *things do not depend on the absolute time.* We can set up an experiment at a given moment and try it out, and then do the same experiment at a later moment, and it will behave in exactly the same way. Whether this is strictly true or not, we do not know. If we assume that it *is* true, and add the principles of quantum mechanics, then we can deduce the principle of the conservation of energy. It is a rather subtle and interesting thing, and it is not easy to explain. The other conservation laws are also linked together. The conservation of momentum is associated in quantum mechanics with the proposition that it makes no difference *where* you do the experiment, the results will always be the same. As independence in space has to do with the conservation of momentum, independence of time has to do with the conservation of energy, and finally, if we *turn* our apparatus, this too makes no difference, and so the invariance of the world to angular orientation

is related to the conservation of *angular momentum.* Besides these, there are three other conservation laws, that are exact so far as we can tell today, which *are* much simpler to understand because they are in the nature of counting blocks.

The first of the three is the *conservation of charge,* and that merely means that you count how many positive, minus how many negative electrical charges you have, and the number is never changed. You may get rid of a positive with a negative, but you do not create any net excess of positives over negatives. Two other laws are analogous to this one—one is called the *conservation of baryons*. There are a number of strange particles, a neutron and a proton are examples, which are called baryons. In any reaction whatever in nature, if we count how many baryons are coming into a process, the number of baryons* which come out will be exactly the same. There is another law, the *conservation of leptons*. We can say that the group of particles called leptons are electron, mu meson, and neutrino. There is an antielectron which is a positron, that is, a –1 lepton. Counting the total number of leptons in a reaction reveals that the number in and out never changes, at least so far as we know at present.

These are the six conservation laws, three of them subtle, involving space and time, and three of them simple, in the sense of counting something.

With regard to the conservation of energy, we should note that *available* energy is another matter—there is a lot of jiggling around in the atoms of the water of the sea, because the sea has a certain temperature, but it is impossible to get them herded into a definite motion without taking energy from somewhere else. That is, although we know for a fact that energy is conserved, the energy available for human utility is not conserved so easily. The laws which govern how much energy is available are called the *laws of thermodynamics* and involve a concept called entropy for irreversible thermodynamic processes.

---

* Counting antibaryons as –1 baryon.

Finally, we remark on the question of where we can get our supplies of energy today. Our supplies of energy are from the sun, rain, coal, uranium, and hydrogen. The sun makes the rain, and the coal also, so that all these are from the sun. Although energy is conserved, nature does not seem to be interested in it; she liberates a lot of energy from the sun, but only one part in two billion falls on the earth. Nature has conservation of energy, but does not really care; she spends a lot of it in all directions. We have already obtained energy from uranium; we can also get energy from hydrogen, but at present only in an explosive and dangerous condition. If it can be controlled in thermonuclear reactions, it turns out that the energy that can be obtained from 10 quarts of water per second is equal to all of the electrical power generated in the United States. With 150 gallons of running water a minute, you have enough fuel to supply all the energy which is used in the United States today! Therefore it is up to the physicist to figure out how to liberate us from the need for having energy. It can be done.

5

THE THEORY OF GRAVITATION

Planetary motions

In this chapter we shall discuss one of the most far-reaching generalizations of the human mind. While we are admiring the human mind, we should take some time off to stand in awe of a *nature* that could follow with such completeness and generality such an elegantly simple principle as the law of gravitation. What is this law of gravitation? It is that every object in the universe attracts every other object with a force which for any two bodies is proportional to the mass of each and varies inversely as the square of the distance between them. This statement can be expressed mathematically by the equation

$$F = G\frac{mm'}{r^2}.$$

If to this we add the fact that an object responds to a force by accelerating in the direction of the force by an amount that is inversely proportional to the mass of the object, we shall have said everything required, for a sufficiently talented mathematician could then deduce all the consequences of these two principles. However,

since you are not assumed to be sufficiently talented yet, we shall discuss the consequences in more detail, and not just leave you with only these two bare principles. We shall briefly relate the story of the discovery of the law of gravitation and discuss some of its consequences, its effects on history, the mysteries that such a law entails, and some refinements of the law made by Einstein; we shall also discuss the relationships of the law to the other laws of physics. All this cannot be done in one chapter, but these subjects will be treated in due time in subsequent chapters.

The story begins with the ancients observing the motions of planets among the stars, and finally deducing that they went around the sun, a fact that was rediscovered later by Copernicus. Exactly *how* the planets went around the sun, with exactly *what motion*, took a little more work to discover. In the beginning of the fifteenth century there were great debates as to whether they really went around the sun or not. Tycho Brahe had an idea that was different from anything proposed by the ancients: his idea was that these debates about the nature of the motions of the planets would best be resolved if the actual positions of the planets in the sky were measured sufficiently accurately. If measurement showed exactly how the planets moved, then perhaps it would be possible to establish one or another viewpoint. This was a tremendous idea—that to find something out, it is better to perform some careful experiments than to carry on deep philosophical arguments. Pursuing this idea, Tycho Brahe studied the positions of the planets for many years in his observatory on the island of Hven, near Copenhagen. He made voluminous tables, which were then studied by the mathematician Kepler, after Tycho's death. Kepler discovered from the data some very beautiful and remarkable, but simple, laws regarding planetary motion.

Kepler's laws

First of all, Kepler found that each planet goes around the sun in a curve called an *ellipse*, with the sun at a focus of the ellipse. An ellipse is not just an oval, but is a very specific and precise curve that

can be obtained by using two tacks, one at each focus, a loop of string, and a pencil; more mathematically, it is the locus of all points the sum of whose distances from two fixed points (the foci) is a constant. Or, if you will, it is a foreshortened circle (Fig. 5-1).

Kepler's second observation was that the planets do not go around the sun at a uniform speed, but move faster when they are nearer the sun and more slowly when they are farther from the sun, in precisely this way: Suppose a planet is observed at any two successive times, let us say a week apart, and that the radius vector* is drawn to the planet for each observed position. The orbital arc traversed by the planet during the week, and the two radius vectors, bound a certain plane area, the shaded area shown in Fig. 5-2. If two similar observations are made a week apart, at a part of the orbit farther from the sun (where the planet moves more slowly), the similarly bounded area is exactly the same as in the first case. So, in accordance with the second law, the orbital speed of each planet is such that the radius "sweeps out" equal areas in equal times.

Finally, a third law was discovered by Kepler much later; this law is of a different category from the other two, because it deals not with only a single planet, but relates one planet to another. This law says that when the orbital period and orbit size of any two planets are compared, the periods are proportional to the $^3/_2$ power of

Figure 5-1 An ellipse.

* A radius vector is a line drawn from the sun to any point in a planet's orbit.

Figure 5-2 Kepler's law of areas.

the orbit size. In this statement the period is the time interval it takes a planet to go completely around its orbit, and the size is measured by the length of the greatest diameter of the elliptical orbit, technically known as the major axis. More simply, if the planets went in circles, as they nearly do, the time required to go around the circle would be proportional to the $^3/_2$ power of the diameter (or radius). Thus Kepler's three laws are:

I. Each planet moves around the sun in an ellipse, with the sun at one focus.
II. The radius vector from the sun to the planet sweeps out equal areas in equal intervals of time.
III. The squares of the periods of any two planets are proportional to the cubes of the semimajor axes of their respective orbits: $T \propto a^{3/2}$.

Development of dynamics

While Kepler was discovering these laws, Galileo was studying the laws of motion. The problem was, what makes the planets go around? (In those days, one of the theories proposed was that the planets went around because behind them were invisible angels, beating their wings and driving the planets forward. You will see that this theory is now modified! It turns out that in order to keep the planets going around, the invisible angels must fly in a different

direction and they have no wings. Otherwise, it is a somewhat similar theory!) Galileo discovered a very remarkable fact about motion, which was essential for understanding these laws. That is the principle of *inertia*—if something is moving, with nothing touching it and completely undisturbed, it will go on forever, coasting at a uniform speed in a straight line. (*Why* does it keep on coasting? We do not know, but that is the way it is.)

Newton modified this idea, saying that the only way to change the motion of a body is to use *force.* If the body speeds up, a force has been applied *in the direction of motion.* On the other hand, if its motion is changed to a new *direction,* a force has been applied *sideways.* Newton thus added the idea that a force is needed to change the speed *or the direction* of motion of a body. For example, if a stone is attached to a string and is whirling around in a circle, it takes a force to keep it in the circle. We have to *pull* on the string. In fact, the law is that the acceleration produced by the force is inversely proportional to the mass, or the force is proportional to the mass times the acceleration. The more massive a thing is, the stronger the force required to produce a given acceleration. (The mass can be measured by putting other stones on the end of the same string and making them go around the same circle at the same speed. In this way it is found that more or less force is required, the more massive object requiring more force.) The brilliant idea resulting from these considerations is that no *tangential* force is needed to keep a planet in its orbit (the angels do not have to fly tangentially) because the planet would coast in that direction anyway. If there were nothing at all to disturb it, the planet would go off in a *straight line.* But the actual motion deviates from the line on which the body would have gone if there were no force, the deviation being essentially *at right angles* to the motion, not in the direction of the motion. In other words, because of the principle of inertia, the force needed to control the motion of a planet around the sun is not a force *around* the sun but *toward* the sun. (If there is a force toward the sun, the sun might be the angel, of course!)

Newton's law of gravitation

From his better understanding of the theory of motion, Newton appreciated that *the sun* could be the seat or organization of forces that govern the motion of the planets. Newton proved to himself (and perhaps we shall be able to prove it soon) that the very fact that equal areas are swept out in equal times is a precise signpost of the proposition that all deviations are precisely *radial*—that the law of areas is a direct consequence of the idea that all of the forces are directed exactly *toward the sun.*

Next, by analyzing Kepler's third law it is possible to show that the farther away the planet, the weaker the forces. If two planets at different distances from the sun are compared, the analysis shows that the forces are inversely proportional to the squares of the respective distances. With the combination of the two laws, Newton concluded that there must be a force, inversely as the square of the distance, directed in a line between the two objects.

Being a man of considerable feeling for generalities, Newton supposed, of course, that this relationship applied more generally than just to the sun holding the planets. It was already known, for example, that the planet Jupiter had moons going around it as the moon of the earth goes around the earth, and Newton felt certain that each planet held its moons with a force. He already knew of the force holding *us* on the earth, so he proposed that this was a *universal* force—*that everything pulls everything else.*

The next problem was whether the pull of the earth on its people was the "same" as its pull on the moon, i.e., inversely as the square of the distance. If an object on the surface of the earth falls 16 feet in the first second after it is released from rest, how far does the moon fall in the same time? We might say that the moon does not fall at all. But if there were no force on the moon, it would go off in a straight line, whereas it goes in a circle instead, so it really *falls in* from where it would have been if there were no force at all. We can calculate from the radius of the moon's orbit (which is about 240,000 miles) and how long it takes to go around the earth (ap-

proximately 29 days) how far the moon moves in its orbit in one second, and can then calculate how far it falls in one second.* This distance turns out to be roughly $^1/_{20}$ of an inch in a second. That fits very well with the inverse square law, because the earth's radius is 4000 miles, and if something which is 4000 miles from the center of the earth falls 16 feet in a second, something 240,000 miles, or 60 times as far away, should fall only $^1/_{3600}$ of 16 feet, which also is roughly $^1/_{20}$ of an inch. Wishing to put this theory of gravitation to a test by similar calculations, Newton made his calculations very carefully and found a discrepancy so large that he regarded the theory as contradicted by facts, and did not publish his results. Six years later a new measurement of the size of the earth showed that the astronomers had been using an incorrect distance to the moon. When Newton heard of this, he made the calculation again, with the corrected figures, and obtained beautiful agreement.

This idea that the moon "falls" is somewhat confusing, because, as you see, it does not come any *closer.* The idea is sufficiently interesting to merit further explanation: the moon falls in the sense that *it falls away from the straight line that it would pursue if there were no forces.* Let us take an example on the surface of the earth. An object released near the earth's surface will fall 16 feet in the first second. An object shot out *horizontally* will also fall 16 feet; even though it is moving horizontally, it still falls the same 16 feet in the same time. Figure 5-3 shows an apparatus which demonstrates this. On the horizontal track is a ball which is going to be driven forward a little distance away. At the same height is a ball which is going to fall vertically, and there is an electrical switch arranged so that at the moment the first ball leaves the track, the second ball is released. That they come to the same depth at the same time is witnessed by the fact that they collide in midair. An object like a bullet, shot horizontally, might go a long way in one second—perhaps 2000 feet—but it will still fall 16 feet if it is aimed

* That is, how far the circle of the moon's orbit falls below the straight line tangent to it at the point where the moon was one second before.

Figure 5-3 Apparatus for showing the independence of vertical and horizontal motions.

horizontally. What happens if we shoot a bullet faster and faster? Do not forget that the earth's surface is curved. If we shoot it fast enough, then when it falls 16 feet it may be at just the same height above the ground as it was before. How can that be? It still falls, but the earth curves away, so it falls "around" the earth. The question is, how far does it have to go in one second so that the earth is 16 feet below the horizon? In Fig. 5-4 we see the earth with its 4000-mile radius, and the tangential, straight-line path that the bullet would take if there were no force. Now, if we use one of those wonderful theorems in geometry, which says that our tangent is the mean proportional between the two parts of the diameter cut by an equal chord, we see that the horizontal distance travelled is the mean proportional between the 16 feet fallen and the 8000-mile diameter of the earth. The square root of $(16/5280) \times 8000$ comes out very close to 5 miles. Thus we see that if the bullet moves at 5 miles a second, it then will continue to fall toward the earth at the same rate of 16 feet each second, but will never get any closer because the earth keeps curving away from it. Thus it was that Mr. Gagarin maintained himself in space while going 25,000 miles around the earth at approximately 5 miles per second. (He took a little longer because he was a little higher.)

Figure 5-4 Acceleration toward the center of a circular path. From plane geometry, $x/S = (2R - S)/x \approx 2R/x$, where R is the radius of the earth, 4000 miles; x is the distance "travelled horizontally" in one second; and S is the distance "fallen" in one second (16 feet).

Any great discovery of a new law is useful only if we can take more out than we put in. Now, Newton *used* the second and third of Kepler's laws to deduce his law of gravitation. What did he *predict?* First, his analysis of the moon's motion was a prediction because it connected the falling of objects on the earth's surface with that of the moon. Second, the question is, *is the orbit an ellipse?* We shall see in a later chapter how it is possible to calculate the motion exactly, and indeed one can prove that it should be an ellipse,* so no extra fact is needed to explain Kepler's *first* law. Thus Newton made his first powerful prediction.

The law of gravitation explains many phenomena not previously understood. For example, the pull of the moon on the earth causes the tides, hitherto mysterious. The moon pulls the water up under it and makes the tides—people had thought of that before, but they were not as clever as Newton, and so they thought there ought to be only one tide during the day. The reasoning was that the moon pulls the water up under it, making a high tide and a low tide, and

* The proof is not given in this course.

since the earth spins underneath, that makes the tide at one station go up and down every 24 hours. Actually the tide goes up and down in 12 hours. Another school of thought claimed that the high tide should be on the other side of the earth because, so they argued, the moon pulls the earth away from the water! Both of these theories are wrong. It actually works like this: The pull of the moon for the earth and for the water is "balanced" at the center. But the water which is closer to the moon is pulled *more* than the average and the water which is farther away from it is pulled *less* than the average. Furthermore, the water can flow while the more rigid earth cannot. The true picture is a combination of these two things.

What do we mean by "balanced"? What balances? If the moon pulls the whole earth toward it, why doesn't the earth fall right "up" to the moon? Because the earth does the same trick as the moon: it goes in a circle around a point which is inside the earth but not at its center. The moon does not just go around the earth, the earth and the moon both go around a central position, each falling toward this common position, as shown in Fig. 5-5. This motion around the common center is what balances the fall of each. So the earth is not going in a straight line either; it travels in a circle. The water on the far side is "unbalanced" because the moon's attraction there is weaker than it is at the center of the earth, where it just balances the "centrifugal force." The result of this imbalance is that the water rises up, away from the center of the earth. On the near side, the attraction from the moon is stronger, and the imbalance is in the opposite direction in space, but again *away* from the center of the earth. The net result is that we get *two* tidal bulges.

Universal gravitation

What else can we understand when we understand gravity? Everyone knows the earth is round. Why is the earth round? That is easy; it is due to gravitation. The earth can be understood to be round merely because everything attracts everything else and so it has attracted it-

Figure 5-5 The earth-moon system, with tides.

self together as far as it can! If we go even further, the earth is not *exactly* a sphere because it is rotating, and this brings in centrifugal effects which tend to oppose gravity near the equator. It turns out that the earth should be elliptical, and we even get the right shape for the ellipse. We can thus deduce that the sun, the moon, and the earth should be (nearly) spheres, just from the law of gravitation.

What else can you do with the law of gravitation? If we look at the moons of Jupiter we can understand everything about the way they move around that planet. Incidentally, there was once a certain difficulty with the moons of Jupiter that is worth remarking on. These satellites were studied very carefully by Roemer, who noticed that the moons sometimes seemed to be ahead of schedule, and sometimes behind. (One can find their schedules by waiting a very long time and finding out how long it takes on the average for the moons to go around.) Now they were *ahead* when Jupiter was particularly *close* to the earth and they were *behind* when Jupiter was *farther* from the earth. This would have been a very difficult thing to explain according to the law of gravitation—it would have been, in fact, the death of this wonderful theory if there were no other explanation. If a law does not work even in *one place* where it ought to, it is just wrong. But the reason for this discrepancy was very simple and beautiful: it takes a little while to *see* the moons of Jupiter because of the time it takes light to travel from Jupiter to the earth.

When Jupiter is closer to the earth the time is a little less, and when it is farther from the earth, the time is more. This is why moons appear to be, on the average, a little ahead or a little behind, depending on whether they are closer to or farther from the earth. This phenomenon showed that light does not travel instantaneously, and furnished the first estimate of the speed of light. This was done in 1656.

If all of the planets push and pull on each other, the force which controls, let us say, Jupiter in going around the sun is not just the force from the sun; there is also a pull from, say, Saturn. This force is not really strong, since the sun is much more massive than Saturn, but there is *some* pull, so the orbit of Jupiter should not be a perfect ellipse, and it is not; it is slightly off, and "wobbles" around the correct elliptical orbit. Such a motion is a little more complicated. Attempts were made to analyze the motions of Jupiter, Saturn, and Uranus on the basis of the law of gravitation. The effects of each of these planets on each other were calculated to see whether or not the tiny deviations and irregularities in these motions could be completely understood from this one law. Lo and behold, for Jupiter and Saturn, all was well, but Uranus was "weird." It behaved in a very peculiar manner. It was not travelling in an exact ellipse, but that was understandable, because of the attractions of Jupiter and Saturn. But even if allowance were made for these attractions, Uranus *still* was not going right, so the laws of gravitation were in danger of being overturned, a possibility that could not be ruled out. Two men, Adams and Le Verrier, in England and France, independently, arrived at another possibility: perhaps there is *another* planet, dark and invisible, which men had not seen. This planet, *N,* could pull on Uranus. They calculated where such a planet would have to be in order to cause the observed perturbations. They sent messages to the respective observatories, saying, "Gentlemen, point your telescope to such and such a place, and you will see a new planet." It often depends on with whom you are working as to whether they pay any attention to you or not. They did pay attention to Le Verrier; they looked, and there planet *N* was! The other observatory then also looked very quickly in the next few days and saw it too.

This discovery shows that Newton's laws are absolutely right in the solar system, but do they extend beyond the relatively small distances of the nearest planets? The first test lies in the question, do *stars* attract *each other* as well as planets? We have definite evidence that they do in the *double stars.* Figure 5-6 shows a double star—two stars very close together (there is also a third star in the picture so that we will know that the photograph was not turned). The stars are also shown as they appeared several years later. We see that, relative to the "fixed" star, the axis of the pair has rotated, i.e., the two stars are going around each other. Do they rotate according to Newton's laws? Careful measurements of the relative positions of one such double star system are shown in Fig. 5-7. There we see a beautiful ellipse, the measures starting in 1862 and going all the way around to 1904 (by now it must have gone around once more). Everything coincides with Newton's laws, except that the star Sirius A is *not at the focus.* Why should that be? Because the plane of the ellipse is not in the "plane of the sky." We are not looking at right angles to the orbit plane, and when an ellipse is viewed at a tilt, it remains an ellipse but the focus is no longer at the same place. Thus we can analyze double stars, moving about each other, according to the requirements of the gravitational law.

Figure 5-6 A double-star system.

Figure 5-7 Orbit of Sirius B with respect to Sirius A.

That the law of gravitation is true at even bigger distances is indicated in Fig. 5-8. If one cannot see gravitation acting here, he has no soul. This figure shows one of the most beautiful things in the sky—a globular star cluster. All of the dots are stars. Although they look as if they are packed solid toward the center, that is due to the fallibility of our instruments. Actually, the distances between even the centermost stars are very great and they very rarely collide. There are more stars in the interior than farther out, and as we move outward there are fewer and fewer. It is obvious that there is an attraction among these stars. It is clear that gravitation exists at these enormous dimensions, perhaps 100,000 times the size of the solar system. Let us now go further, and look at an *entire galaxy,* shown in Fig. 5-9. The shape of this galaxy indicates an obvious tendency for its matter to agglomerate. Of course we cannot prove that the law here is precisely inverse square, only that there is still an attraction, at this enormous dimension, that holds the whole thing together. One may say, "Well, that is all very clever but why is it not just a ball?" Because it is *spinning* and has *angular momentum* which it cannot give up as it contracts;

Figure 5-8 A globular star cluster.

it must contract mostly in a plane. (Incidentally, if you are looking for a good problem, the exact details of how the arms are formed and what determines the shapes of these galaxies have not been worked out.) It is, however, clear that the shape of the galaxy is due to gravitation even though the complexities of its structure have not yet allowed us to analyze it completely. In a galaxy we have a scale of perhaps 50,000 to 100,000 light-years. The earth's distance from the sun is $8^1/_3$ light-*minutes,* so you can see how large these dimensions are.

Gravity appears to exist at even bigger dimensions, as indicated by Fig. 5-10, which shows many "little" things clustered together. This is a *cluster of galaxies,* just like a star cluster. Thus galaxies attract each other at such distances that they too are agglomerated into clusters. Perhaps gravitation exists even over distances of *tens of millions* of light-years; so far as we now know, gravity seems to go out forever inversely as the square of the distance.

Not only can we understand the nebulae, but from the law of gravitation we can even get some ideas about the origin of the stars. If we have a big cloud of dust and gas, as indicated in Fig. 5-11, the gravitational attractions of the pieces of dust for one another

Figure 5-9 A galaxy.

might make them form little lumps. Barely visible in the figure are "little" black spots which may be the beginning of the accumulations of dust and gases which, due to their gravitation, begin to form stars. Whether we have ever seen a star form or not is still debatable. Figure 5-12 shows the one piece of evidence which suggests that we have. At the left is a picture of a region of gas with some stars in it taken in 1947, and at the right is another picture, taken only seven years later, which shows two new bright spots. Has gas accumulated, has gravity acted hard enough and collected it into a ball big enough that the stellar nuclear reaction starts in the interior and turns it into a star? Perhaps, and perhaps not. It is unreasonable that in only seven years we should be so lucky as to see a star change itself into visible form; it is much less probable that we should see *two!*

Cavendish's experiment

Gravitation, therefore, extends over enormous distances. But if there is a force between *any* pair of objects, we ought to be able to measure the force between our own objects. Instead of having to

Figure 5-10 A cluster of galaxies.

watch the stars go around each other, why can we not take a ball of lead and a marble and watch the marble go toward the ball of lead? The difficulty of this experiment when done in such a simple manner is the very weakness or delicacy of the force. It must be done with extreme care, which means covering the apparatus to keep the air out, making sure it is not electrically charged, and so on; then the force can be measured. It was first measured by Cavendish with an apparatus which is schematically indicated in Fig. 5-13. This first demonstrated the direct force between two large, fixed balls of lead and two smaller balls of lead on the ends of an arm supported by a very fine fiber, called a torsion fiber. By measuring how much the fiber gets twisted, one can measure the strength of the force, verify that it is inversely proportional to the square of the distance, and determine how strong it is. Thus, one may accurately determine the coefficient G in the formula

$$F = G\frac{mm'}{r^2}.$$

Figure 5-11 An interstellar dust cloud.

All the masses and distances are known. You say, "We knew it already for the earth." Yes, but we did not know the *mass* of the earth. By knowing G from this experiment and by knowing how strongly the earth attracts, we can indirectly learn how great is the mass of the earth! This experiment has been called "weighing the earth" by some people, and it can be used to determine the coefficient G of the gravity law. This is the only way in which the mass of the earth can be determined. G turns out to be

$$6.670 \times 10^{-11} \text{ newton} \cdot \text{m}^2/\text{kg}^2.$$

It is hard to exaggerate the importance of the effect on the history of science produced by this great success of the theory of gravitation. Compare the confusion, the lack of confidence, the incomplete knowledge that prevailed in the earlier ages, when there were endless debates and paradoxes, with the clarity and simplicity of this law—this fact that all the moons and planets and stars have such a *simple*

Figure 5-12 The formation of new stars?

rule to govern them, and further that man could *understand* it and deduce how the planets should move! This is the reason for the success of the sciences in following years, for it gave hope that the other phenomena of the world might also have such beautifully simple laws.

What is gravity?

But is this such a simple law? What about the machinery of it? All we have done is to describe *how* the earth moves around the sun, but we have not said *what makes it go.* Newton made no hypotheses about this; he was satisfied to find *what* it did without getting into the machinery of it. *No one has since given any machinery.* It is characteristic of the physical laws that they have this abstract character. The law of conservation of energy is a theorem concerning quantities that have to be calculated and added together, with no mention of the machinery, and likewise the great laws of mechanics are quantitative mathematical laws for which no machinery is available.

Figure 5-13 A simplified diagram of the apparatus used by Cavendish to verify the law of universal gravitation for small objects and to measure the gravitational constant *G*.

Why can we use mathematics to describe nature without a mechanism behind it? No one knows. We have to keep going because we find out more that way.

Many mechanisms for gravitation have been suggested. It is interesting to consider one of these, which many people have thought of from time to time. At first, one is quite excited and happy when he "discovers" it, but he soon finds that it is not correct. It was first discovered about 1750. Suppose there were many particles moving in space at a very high speed in all directions and being only slightly absorbed in going through matter. When they *are* absorbed, they give an impulse to the earth. However, since there are as many going one way as another, the impulses all balance. But when the sun is nearby, the particles coming toward the earth through the sun are partially absorbed, so fewer of them are coming from the sun than are coming from the other side. Therefore, the earth feels a net impulse toward the sun and it does not take one long to see that it is inversely proportional to the square of the distance—because of the variation of the solid angle that the sun subtends as we vary the distance. What is wrong with that machinery? It involves some new consequences which are *not true.* This particular idea has the fol-

lowing trouble: the earth, in moving around the sun, would impinge on more particles which are coming from its forward side than from its hind side (when you run in the rain, the rain in your face is stronger than that on the back of your head!). Therefore there would be more impulse given the earth from the front, and the earth would feel a *resistance to motion* and would be slowing up in its orbit. One can calculate how long it would take for the earth to stop as a result of this resistance, and it would not take long enough for the earth to still be in its orbit, so this mechanism does not work. No machinery has ever been invented that "explains" gravity without also predicting some other phenomenon that does *not* exist.

Next we shall discuss the possible relation of gravitation to other forces. There is no explanation of gravitation in terms of other forces at the present time. It is not an aspect of electricity or anything like that, so we have no explanation. However, gravitation and other forces are very similar, and it is interesting to note analogies. For example, the force of electricity between two charged objects looks just like the law of gravitation: the force of electricity is a constant, with a minus sign, times the product of the charges, and varies inversely as the square of the distance. It is in the opposite direction—likes repel. But is it still not very remarkable that the two laws involve the same function of distance? Perhaps gravitation and electricity are much more closely related than we think. Many attempts have been made to unify them; the so-called unified field theory is only a very elegant attempt to combine electricity and gravitation; but, in comparing gravitation and electricity, the most interesting thing is the *relative strengths* of the forces. Any theory that contains them both must also deduce how strong the gravity is.

If we take, in some natural units, the repulsion of two electrons (nature's universal charge) due to electricity, and the attraction of two electrons due to their masses, we can measure the ratio of electrical repulsion to the gravitational attraction. The ratio is independent of the distance and is a fundamental constant of nature. The

ratio is shown in Fig. 5-14. The gravitational attraction relative to the electrical repulsion between two electrons is 1 divided by 4.17×10^{42}! The question is, where does such a large number come from? It is not accidental, like the ratio of the volume of the earth to the volume of a flea. We have considered two natural aspects of the same thing, an electron. This fantastic number is a natural constant, so it involves something deep in nature. Where could such a tremendous number come from? Some say that we shall one day find the "universal equation," and in it, one of the roots will be this number. It is very difficult to find an equation for which such a fantastic number is a natural root. Other possibilities have been thought of; one is to relate it to the age of the universe. Clearly, we have to find *another* large number somewhere. But do we mean the age of the universe in *years?* No, because years are not "natural"; they were devised by men. As an example of something natural, let us consider the time it takes light to go across a proton, 10^{-24} second. If we compare this time with the *age of the universe,* 2×10^{10} years, the answer is 10^{-42}. It has about the same number of zeros going off it, so it has been proposed that the gravitational constant is related to the age of the universe. If that were the case, the gravitational constant would change with time, because as the universe

$$\frac{\text{Gravitation Attraction}}{\text{Electrical Repulsion}} = 1/4.17 \times 10^{42}$$

$$= 1/4{,}170{,}000{,}000{,}000

got older the ratio of the age of the universe to the time which it takes for light to go across a proton would be gradually increasing. Is it possible that the gravitational constant *is* changing with time? Of course the changes would be so small that it is quite difficult to be sure.

One test which we can think of is to determine what would have been the effect of the change during the past 10^9 years, which is approximately the age from the earliest life on the earth to now, and one-tenth of the age of the universe. In this time, the gravity constant would have increased by about 10 percent. It turns out that if we consider the structure of the sun—the balance between the weight of its material and the rate at which radiant energy is generated inside it—we can deduce that if the gravity were 10 percent stronger, the sun would be much more than 10 percent brighter—by the *sixth power* of the gravity constant! If we calculate what happens to the orbit of the earth when the gravity is changing, we find that the earth was then *closer in.* Altogether, the earth would be about 100 degrees centigrade hotter, and all of the water would not have been in the sea, but vapor in the air, so life would not have started in the sea. So we do *not* now believe that the gravity constant is changing with the age of the universe. But such arguments as the one we have just given are not very convincing, and the subject is not completely closed.

It is a fact that the force of gravitation is proportional to the *mass,* the quantity which is fundamentally a measure of *inertia*—of how hard it is to hold something which is going around in a circle. Therefore two objects, one heavy and one light, going around a larger object in the same circle at the same speed because of gravity, will stay together because to go in a circle *requires* a force which is stronger for a bigger mass. That is, the gravity is stronger for a given mass in *just the right proportion* so that the two objects will go around together. If one object were inside the other it would *stay* inside; it is a perfect balance. Therefore, Gagarin or Titov would find things "weightless" inside a spaceship; if they happened

to let go of a piece of chalk, for example, it would go around the earth in exactly the same way as the whole spaceship, and so it would appear to remain suspended before them in space. It is very interesting that this force is *exactly* proportional to the mass with great precision, because if it were not exactly proportional there would be some effect by which inertia and weight would differ. The absence of such an effect has been checked with great accuracy by an experiment done first by Eötvös in 1909 and more recently by Dicke. For all substances tried, the masses and weights are exactly proportional within 1 part in 1,000,000,000, or less. This is a remarkable experiment.

Gravity and relativity

Another topic deserving discussion is Einstein's modification of Newton's law of gravitation. In spite of all the excitement it created, Newton's law of gravitation is not correct! It was modified by Einstein to take into account the theory of relativity. According to Newton, the gravitational effect is instantaneous, that is, if we were to move a mass, we would at once feel a new force because of the new position of that mass; by such means we could send signals at infinite speed. Einstein advanced arguments which suggest that we *cannot send signals faster than the speed of light,* so the law of gravitation must be wrong. By correcting it to take the delays into account, we have a new law, called Einstein's law of gravitation. One feature of this new law which is quite easy to understand is this: In the Einstein relativity theory, anything which has *energy* has mass—mass in the sense that it is attracted gravitationally. Even light, which has an energy, has a "mass." When a light beam, which has energy in it, comes past the sun there is an attraction on it by the sun. Thus the light does not go straight, but is deflected. During the eclipse of the sun, for example, the stars which are around the sun should appear displaced from where they would be if the sun were not there, and this has been observed.

Finally, let us compare gravitation with other theories. In recent

years we have discovered that all mass is made of tiny particles and that there are several kinds of interactions, such as nuclear forces, etc. None of these nuclear or electrical forces has yet been found to explain gravitation. The quantum-mechanical aspects of nature have not yet been carried over to gravitation. When the scale is so small that we need the quantum effects, the gravitational effects are so weak that the need for a quantum theory of gravitation has not yet developed. On the other hand, for consistency in our physical theories it would be important to see whether Newton's law modified to Einstein's law can be further modified to be consistent with the uncertainty principle. This last modification has not yet been completed.

6

QUANTUM BEHAVIOR

Atomic mechanics

In the last few chapters we have treated the essential ideas necessary for an understanding of most of the important phenomena of light—or electromagnetic radiation in general. (We have left a few special topics for next year. Specifically, the theory of the index of dense materials and total internal reflection.) What we have dealt with is called the "classical theory" of electric waves, which turns out to be a completely adequate description of nature for a large number of effects. We have not had to worry yet about the fact that light energy comes in lumps or "photons."

We would like to take up as our next subject the problem of the behavior of relatively large pieces of matter—their mechanical and thermal properties, for instance. In discussing these, we will find that the "classical" (or older) theory fails almost immediately, because matter is really made up of atomic-sized particles. Still, we will deal only with the classical part, because that is the only part that we can understand using the classical mechanics we have been learning. But we shall not be very successful. We shall find that in the case of matter, unlike the case of light, we shall be in difficulty relatively soon. We could, of course, continuously skirt away from the atomic effects, but we shall instead interpose here a short excursion in which we will describe the basic ideas of the quantum properties of matter, i.e., the quantum ideas of atomic physics, so

that you will have some feeling for what it is we are leaving out. For we will have to leave out some important subjects that we cannot avoid coming close to.

So we will give now the *introduction* to the subject of quantum mechanics, but will not be able actually to get into the subject until much later.

"Quantum mechanics" is the description of the behavior of matter in all its details and, in particular, of the happenings on an atomic scale. Things on a very small scale behave like nothing that you have any direct experience about. They do not behave like waves, they do not behave like particles, they do not behave like clouds, or billiard balls, or weights on springs, or like anything that you have ever seen.

Newton thought that light was made up of particles, but then it was discovered, as we have seen here, that it behaves like a wave. Later, however (in the beginning of the twentieth century), it was found that light did indeed sometimes behave like a particle. Historically, the electron, for example, was thought to behave like a particle, and then it was found that in many respects it behaved like a wave. So it really behaves like neither. Now we have given up. We say: "It is like *neither*."

There is one lucky break, however—electrons behave just like light. The quantum behavior of atomic objects (electrons, protons, neutrons, photons, and so on) is the same for all; they are all "particle waves," or whatever you want to call them. So what we learn about the properties of electrons (which we shall use for our examples) will apply also to all "particles," including photons of light.

The gradual accumulation of information about atomic and small-scale behavior during the first quarter of this century, which gave some indications about how small things do behave, produced an increasing confusion which was finally resolved in 1926 and 1927 by Schrödinger, Heisenberg, and Born. They finally obtained a consistent description of the behavior of matter on a small scale. We take up the main features of that description in this chapter.

Because atomic behavior is so unlike ordinary experience, it is very difficult to get used to and it appears peculiar and mysterious to everyone, both to the novice and to the experienced physicist. Even the experts do not understand it the way they would like to, and it is perfectly reasonable that they should not, because all of direct, human experience and human intuition applies to large objects. We know how large objects will act, but things on a small scale just do not act that way. So we have to learn about them in a sort of abstract or imaginative fashion and not by connection with our direct experience.

In this chapter we shall tackle immediately the basic element of the mysterious behavior in its most strange form. We choose to examine a phenomenon which is impossible, *absolutely* impossible, to explain in any classical way, and which has in it the heart of quantum mechanics. In reality, it contains the *only* mystery. We cannot explain the mystery in the sense of "explaining" how it works. We will *tell* you how it works. In telling you how it works we will have told you about the basic peculiarities of all quantum mechanics.

An experiment with bullets

To try to understand the quantum behavior of electrons, we shall compare and contrast their behavior, in a particular experimental setup, with the more familiar behavior of particles like bullets, and with the behavior of waves like water waves. We consider first the behavior of bullets in the experimental setup shown diagrammatically in Fig. 6-1. We have a machine gun that shoots a stream of bullets. It is not a very good gun, in that it sprays the bullets (randomly) over a fairly large angular spread, as indicated in the figure. In front of the gun we have a wall (made of armor plate) that has in it two holes just about big enough to let a bullet through. Beyond the wall is a backstop (say a thick wall of wood) which will "absorb" the bullets when they hit it. In front of the wall we have an object which we shall call a "detector" of bullets. It might be a

Figure 6-1 Interference experiment with bullets.

box containing sand. Any bullet that enters the detector will be stopped and accumulated. When we wish, we can empty the box and count the number of bullets that have been caught. The detector can be moved back and forth (in what we will call the x-direction). With this apparatus, we can find out experimentally the answer to the question: "What is the probability that a bullet which passes through the holes in the wall will arrive at the backstop at the distance x from the center?" First, you should realize that we should talk about probability, because we cannot say definitely where any particular bullet will go. A bullet which happens to hit one of the holes may bounce off the edges of the hole, and may end up anywhere at all. By "probability" we mean the chance that the bullet will arrive at the detector, which we can measure by counting the number which arrive at the detector in a certain time and then taking the ratio of this number to the *total* number that hit the backstop during that time. Or, if we assume that the gun always shoots at the same rate during the measurements, the probability we want is just proportional to the number that reach the detector in some standard time interval.

For our present purposes we would like to imagine a somewhat idealized experiment in which the bullets are not real bullets, but

are *indestructible* bullets—they cannot break in half. In our experiment we find that bullets always arrive in lumps, and when we find something in the detector, it is always one whole bullet. If the rate at which the machine gun fires is made very low, we find that at any given moment either nothing arrives or one and only one—exactly one—bullet arrives at the backstop. Also, the size of the lump certainly does not depend on the rate of firing of the gun. We shall say: "Bullets *always* arrive in identical lumps." What we measure with our detector is the probability of arrival of a lump. And we measure the probability as a function of x. The result of such measurements with this apparatus (we have not yet done the experiment, so we are really imagining the result) is plotted in the graph drawn in part (c) of Fig. 6-1. In the graph we plot the probability to the right and x vertically, so that the x-scale fits the diagram of the apparatus. We call the probability P_{12} because the bullets may have come either through hole 1 or through hole 2. You will not be surprised that P_{12} is large near the middle of the graph but gets small if x is very large. You may wonder, however, why P_{12} has its maximum value at $x = 0$. We can understand this fact if we do our experiment again after covering up hole 2, and once more while covering up hole 1. When hole 2 is covered, bullets can pass only through hole 1, and we get the curve marked P_1 in part (b) of the figure. As you would expect, the maximum of P_1 occurs at the value of x which is on a straight line with the gun and hole 1. When hole 1 is closed, we get the symmetric curve P_2 drawn in the figure. P_2 is the probability distribution for bullets that pass through hole 2. Comparing parts (b) and (c) of Fig. 6-1, we find the important result that

$$P_{12} = P_1 + P_2. \tag{6.1}$$

The probabilities just add together. The effect with both holes open is the sum of the effects with each hole open alone. We shall call this result an observation of "*no interference*," for a reason that you will see later. So much for bullets. They come in lumps, and their probability of arrival shows no interference.

An experiment with waves

Now we wish to consider an experiment with water waves. The apparatus is shown diagrammatically in Fig. 6-2. We have a shallow trough of water. A small object labeled the "wave source" is jiggled up and down by a motor and makes circular waves. To the right of the source we have again a wall with two holes, and beyond that is a second wall, which, to keep things simple, is an "absorber," so that there is no reflection of the waves that arrive there. This can be done by building a gradual sand "beach." In front of the beach we place a detector which can be moved back and forth in the x-direction, as before. The detector is now a device which measures the "intensity" of the wave motion. You can imagine a gadget which measures the height of the wave motion, but whose scale is calibrated in proportion to the *square* of the actual height, so that the reading is proportional to the intensity of the wave. Our detector reads, then, in proportion to the *energy* being carried by the wave—or rather, the rate at which energy is carried to the detector.

With our wave apparatus, the first thing to notice is that the intensity can have *any* size. If the source just moves a very small

Figure 6-2 Interference experiment with water waves.

amount, then there is just a little bit of wave motion at the detector. When there is more motion at the source, there is more intensity at the detector. The intensity of the wave can have any value at all. We would *not* say that there was any "lumpiness" in the wave intensity.

Now let us measure the wave intensity for various values of x (keeping the wave source operating always in the same way). We get the interesting-looking curve marked I_{12} in part (c) of the figure.

We have already worked out how such patterns can come about when we studied the interference of electric waves. In this case we would observe that the original wave is diffracted at the holes, and new circular waves spread out from each hole. If we cover one hole at a time and measure the intensity distribution at the absorber we find the rather simple intensity curves shown in part (b) of the figure. I_1 is the intensity of the wave from hole 1 (which we find by measuring when hole 2 is blocked off) and I_2 is the intensity of the wave from hole 2 (seen when hole 1 is blocked).

The intensity I_{12} observed when both holes are open is certainly *not* the sum of I_1 and I_2. We say that there is "interference" of the two waves. At some places (where the curve I_{12} has its maxima) the waves are "in phase" and the wave peaks add together to give a large amplitude and, therefore, a large intensity. We say that the two waves are "interfering constructively" at such places. There will be such constructive interference wherever the distance from the detector to one hole is a whole number of wavelengths larger (or shorter) than the distance from the detector to the other hole.

At those places where the two waves arrive at the detector with a phase difference of π (where they are "out of phase") the resulting wave motion at the detector will be the difference of the two amplitudes. The waves "interfere destructively," and we get a low value for the wave intensity. We expect such low values wherever the distance between hole 1 and the detector is different from the distance between hole 2 and the detector by an odd number of half-wavelengths. The low values of I_{12} in Fig. 6-2 correspond to the places where the two waves interfere destructively.

You will remember that the quantitative relationship between

I_1, I_2, and I_{12} can be expressed in the following way: The instantaneous height of the water wave at the detector for the wave from hole 1 can be written as (the real part of) $\hat{h}_1 e^{i\omega t}$, where the "amplitude" $\hat{h}_1$ is, in general, a complex number. The intensity is proportional to the mean squared height or, when we use the complex numbers, to $|\hat{h}_1|^2$. Similarly, for hole 2 the height is $\hat{h}_2 e^{i\omega t}$ and the intensity is proportional to $|\hat{h}_2|^2$. When both holes are open, the wave heights add to give the height $(\hat{h}_1 + \hat{h}_2)e^{i\omega t}$ and the intensity $|\hat{h}_1 + \hat{h}_2|^2$. Omitting the constant of proportionality for our present purposes, the proper relations for *interfering waves* are

$$I_1 = |\hat{h}_1|^2, \quad I_2 = |\hat{h}_2|^2, \quad I_{12} = |\hat{h}_1 + \hat{h}_2|^2. \tag{6.2}$$

You will notice that the result is quite different from that obtained with bullets (6.1). If we expand $|\hat{h}_1 + \hat{h}_2|^2$ we see that

$$|\hat{h}_1 + \hat{h}_2|^2 = |\hat{h}_1|^2 + |\hat{h}_2|^2 + 2\,|\hat{h}_1||\hat{h}_2| \cos\delta, \tag{6.3}$$

where δ is the phase difference between $\hat{h}_1$ and $\hat{h}_2$. In terms of the intensities, we could write

$$I_{12} = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta. \tag{6.4}$$

The last term in (6.4) is the "interference term." So much for water waves. The intensity can have any value, and it shows interference.

An experiment with electrons

Now we imagine a similar experiment with electrons. It is shown diagrammatically in Fig. 6-3. We make an electron gun which consists of a tungsten wire heated by an electric current and surrounded by a metal box with a hole in it. If the wire is at a negative voltage with respect to the box, electrons emitted by the wire will be accelerated toward the walls and some will pass through the hole. All the electrons which come out of the gun will have (nearly) the same

energy. In front of the gun is again a wall (just a thin metal plate) with two holes in it. Beyond the wall is another plate which will serve as a "backstop." In front of the backstop we place a movable detector. The detector might be a geiger counter or, perhaps better, an electron multiplier, which is connected to a loudspeaker.

We should say right away that you should not try to set up this experiment (as you could have done with the two we have already described). This experiment has never been done in just this way. The trouble is that the apparatus would have to be made on an impossibly small scale to show the effects we are interested in. We are doing a "thought experiment," which we have chosen because it is easy to think about. We know the results that *would* be obtained because there *are* many experiments that have been done, in which the scale and the proportions have been chosen to show the effects we shall describe.

The first thing we notice with our electron experiment is that we hear sharp "clicks" from the detector (that is, from the loudspeaker). And all "clicks" are the same. There are *no* "half-clicks."

We would also notice that the "clicks" come very erratically. Something like: click . . . . . click-click . . . click . . . . . click . . . .

Figure 6-3 Interference experiment with electrons.

click-click . . . . . click . . . , etc., just as you have, no doubt, heard a geiger counter operating. If we count the clicks which arrive in a sufficiently long time—say, for many minutes—and then count again for another equal period, we find that the two numbers are very nearly the same. So we can speak of the *average rate* at which the clicks are heard (so-and-so-many clicks per minute on the average).

As we move the detector around, the *rate* at which the clicks appear is faster or slower, but the size (loudness) of each click is always the same. If we lower the temperature of the wire in the gun the rate of clicking slows down, but still each click sounds the same. We would notice also that if we put two separate detectors at the backstop, one *or* the other would click, but never both at once. (Except that once in a while, if there were two clicks very close together in time, our ear might not sense the separation.) We conclude, therefore, that whatever arrives at the backstop arrives in "lumps." All the "lumps" are the same size: only whole "lumps" arrive, and they arrive one at a time at the backstop. We shall say: "Electrons always arrive in identical lumps."

Just as for our experiment with bullets, we can now proceed to find experimentally the answer to the question: "What is the relative probability that an electron 'lump' will arrive at the backstop at various distances x from the center?" As before, we obtain the relative probability by observing the rate of clicks, holding the operation of the gun constant. The probability that lumps will arrive at a particular x is proportional to the average rate of clicks at that x.

The result of our experiment is the interesting curve marked P_{12} in part (c) of Fig. 6-3. Yes! That is the way electrons go.

The interference of electron waves

Now let us try to analyze the curve of Fig. 6-3 to see whether we can understand the behavior of the electrons. The first thing we would say is that since they come in lumps, each lump, which we may as

well call an electron, has come either through hole 1 or through hole 2. Let us write this in the form of a "Proposition":

> *Proposition A*: Each electron *either* goes through hole 1 *or* it goes through hole 2.

Assuming Proposition A, all electrons that arrive at the backstop can be divided into two classes: (1) those that come through hole 1, and (2) those that come through hole 2. So our observed curve must be the sum of the effects of the electrons which come through hole 1 and the electrons which come through hole 2. Let us check this idea by experiment. First, we will make a measurement for those electrons that come through hole 1. We block off hole 2 and make our counts of the clicks from the detector. From the clicking rate, we get P_1. The result of the measurement is shown by the curve marked P_1 in part (b) of Fig. 6-3. The result seems quite reasonable. In a similar way, we measure P_2, the probability distribution for the electrons that come through hole 2. The result of this measurement is also drawn in the figure.

The result P_{12} obtained with *both* holes open is clearly not the sum of P_1 and P_2, the probabilities for each hole alone. In analogy with our water-wave experiment, we say: "There is interference."

$$\textit{For electrons:} \qquad P_{12} \neq P_1 + P_2. \tag{6.5}$$

How can such an interference come about? Perhaps we should say: "Well, that means, presumably, that it is *not true* that the lumps go either through hole 1 or hole 2, because if they did, the probabilities should add. Perhaps they go in a more complicated way. They split in half and . . ." But no! They cannot, they always arrive in lumps . . . "Well, perhaps some of them go through 1, and then they go around through 2, and then around a few more times, or by some other complicated path . . . then by closing hole 2, we changed the chance that an electron that *started out* through hole 1 would finally get to the backstop . . ." But notice! There are some

points at which very few electrons arrive when *both* holes are open, but which receive many electrons if we close one hole, so *closing* one hole *increased* the number from the other. Notice, however, that at the center of the pattern, P_{12} is more than twice as large as $P_1 + P_2$. It is as though closing one hole *decreased* the number of electrons which come through the other hole. It seems hard to explain *both* effects by proposing that the electrons travel in complicated paths.

It is all quite mysterious. And the more you look at it the more mysterious it seems. Many ideas have been concocted to try to explain the curve for P_{12} in terms of individual electrons going around in complicated ways through the holes. None of them has succeeded. None of them can get the right curve for P_{12} in terms of P_1 and P_2.

Yet, surprisingly enough, the *mathematics* for relating P_1 and P_2 to P_{12} is extremely simple. For P_{12} is just like the curve I_{12} of Fig. 6-2, and *that* was simple. What is going on at the backstop can be described by two complex numbers that we can call $\hat{\phi}_1$ and $\hat{\phi}_2$ (they are functions of x, of course). The absolute square of $\hat{\phi}_1$ gives the effect with only hole 1 open. That is, $P_1 = |\hat{\phi}_1|^2$. The effect with only hole 2 open is given by $\hat{\phi}_2$ in the same way. That is, $P_2 = |\hat{\phi}_1|^2$. And the combined effect of the two holes is just $P_{12} = |\hat{\phi}_1 + \hat{\phi}_2|$. The *mathematics* is the same as what we had for the water waves! (It is hard to see how one could get such a simple result from a complicated game of electrons going back and forth through the plate on some strange trajectory.)

We conclude the following: The electrons arrive in lumps, like particles, and the probability of arrival of these lumps is distributed like the distribution of intensity of a wave. It is in this sense that an electron behaves "sometimes like a particle and sometimes like a wave."

Incidentally, when we were dealing with classical waves we defined the intensity as the mean over time of the square of the wave amplitude, and we used complex numbers as a mathematical trick to simplify the analysis. But in quantum mechanics it turns out that the amplitudes *must* be represented by complex numbers. The

real parts alone will not do. That is a technical point, for the moment, because the formulas look just the same.

Since the probability of arrival through both holes is given so simply, although it is not equal to $(P_1 + P_2)$, that is really all there is to say. But there are a large number of subtleties involved in the fact that nature does work this way. We would like to illustrate some of these subtleties for you now. First, since the number that arrives at a particular point is *not* equal to the number that arrives through 1 plus the number that arrives through 2, as we would have concluded from Proposition A, undoubtedly we should conclude that *Proposition A is false.* It is *not* true that the electrons go *either* through hole 1 or hole 2. But that conclusion can be tested by another experiment.

Watching the electrons

We shall now try the following experiment. To our electron apparatus we add a very strong light source, placed behind the wall and between the two holes, as shown in Fig. 6-4. We know that electric charges scatter light. So when an electron passes, however it does pass, on its way to the detector, it will scatter some light to our eye, and we can *see* where the electron goes. If, for instance, an electron were to take the path via hole 2 that is sketched in Fig. 6-4, we should see a flash of light coming from the vicinity of the place marked *A* in the figure. If an electron passes through hole 1 we would expect to see a flash from the vicinity of the upper hole. If it should happen that we get light from both places at the same time, because the electron divides in half . . . Let us just do the experiment!

Here is what we see: *every* time that we hear a "click" from our electron detector (at the backstop), we *also see* a flash of light *either* near hole 1 *or* near hole 2, but *never* both at once! And we observe the same result no matter where we put the detector. From this observation we conclude that when we look at the electrons we find that the electrons go either through one hole or the other. Experimentally, Proposition A is necessarily true.

Figure 6-4 A different electron experiment.

What, then, is wrong with our argument *against* Proposition A? Why *isn't* P_{12} just equal to $P_1 + P_2$? Back to experiment! Let us keep track of the electrons and find out what they are doing. For each position (x-location) of the detector we will count the electrons that arrive and *also* keep track of which hole they went through, by watching for the flashes. We can keep track of things this way: whenever we hear a "click" we will put a count in Column 1 if we see the flash near hole 1, and if we see the flash near hole 2, we will record a count in Column 2. Every electron which arrives is recorded in one of two classes: those which come through 1 and those which come through 2. From the number recorded in Column 1 we get the probability P'_1 that an electron will arrive at the detector via hole 1; and from the number recorded in Column 2 we get P'_2, the probability that an electron will arrive at the detector via hole 2. If we now repeat such a measurement for many values of x, we get the curves for P'_1 and P'_2 shown in part (b) of Fig. 6-4.

Well, that is not too surprising! We get for P'_1 something quite similar to what we got before for P, by blocking off hole 2; and P'_2 is similar to what we got by blocking hole 1. So there is *not* any complicated business like going through both holes. When we

watch them, the electrons come through just as we would expect them to come through. Whether the holes are closed or open, those which we see come through hole 1 are distributed in the same way whether hole 2 is open or closed.

But wait! What do we have *now* for the *total* probability, the probability that an electron will arrive at the detector by any route? We already have that information. We just pretend that we never looked at the light flashes, and we lump together the detector clicks which we have separated into the two columns. We *must* just *add* the numbers. For the probability that an electron will arrive at the backstop by passing through *either* hole, we do find $P'_{12} = P'_1 + P'_2$. That is, although we succeeded in watching which hole our electrons come through, we no longer get the old interference curve P_{12}, but a new one, P'_{12}, showing no interference! If we turn out the light P_{12} is restored.

We must conclude that *when we look at the electrons* the distribution of them on the screen is different than when we do not look. Perhaps it is turning on our light source that disturbs things? It must be that the electrons are very delicate, and the light, when it scatters off the electrons, gives them a jolt that changes their motion. We know that the electric field of the light acting on a charge will exert a force on it. So perhaps we *should* expect the motion to be changed. Anyway, the light exerts a big influence on the electrons. By trying to "watch" the electrons we have changed their motions. That is, the jolt given to the electron when the photon is scattered by it is such as to change the electron's motion enough so that if it *might* have gone to where P_{12} was at a maximum, it will instead land where P_{12} was a minimum; that is why we no longer see the wavy interference effects.

You may be thinking: "Don't use such a bright source! Turn the brightness down! The light waves will then be weaker and will not disturb the electrons so much. Surely, by making the light dimmer and dimmer, eventually the wave will be weak enough that it will have a negligible effect." OK. Let's try it. The first thing we observe is that the flash of light scattered from the electrons as they pass by

does *not* get weaker. *It is always the same-sized flash.* The only thing that happens as the light is made dimmer is that sometimes we hear a "click" from the detector but see *no flash at all.* The electron has gone by without being "seen." What we are observing is that light *also* acts like electrons; we *knew* that it was "wavy," but now we find that it is also "lumpy." It always arrives—or is scattered—in lumps that we call "photons." As we turn down the *intensity* of the light source we do not change the *size* of the photons, only the *rate* at which they are emitted. *That* explains why, when our source is dim, some electrons get by without being seen. There did not happen to be a photon around at the time the electron went through.

This is all a little discouraging. If it is true that whenever we "see" the electron we see the same-sized flash, then those electrons we see are *always* the disturbed ones. Let us try the experiment with a dim light anyway. Now whenever we hear a click in the detector we will keep a count in three columns: in Column 1 those electrons seen by hole 1, in Column 2 those electrons seen by hole 2, and in Column 3 those electrons not seen at all. When we work up our data (computing the probabilities) we find these results: Those "seen by hole 1" have a distribution like P'_1; those "seen by hole 2" have a distribution like P'_2 (so that those "seen by either hole 1 or 2" have a distribution like P'_{12}); and those "not seen at all" have a "wavy" distribution just like P_{12} of Fig. 6-3! *If the electrons are not seen, we have interference!*

That is understandable. When we do not see the electron, no photon disturbs it, and when we do see it, a photon has disturbed it. There is always the same amount of disturbance because the light photons all produce the same-sized effects and the effect of the photons being scattered is enough to smear out any interference effect.

Is there not *some* way we can see the electrons without disturbing them? We learned in an earlier chapter that the momentum carried by a "photon" is inversely proportional to its wavelength ($p = h/\lambda$). Certainly the jolt given to the electron when the photon is scattered

toward our eye depends on the momentum that photon carries. Aha! If we want to disturb the electrons only slightly we should not have lowered the *intensity* of the light; we should have lowered its *frequency* (the same as increasing its wavelength). Let us use light of a redder color. We could even use infrared light, or radiowaves (like radar), and "see" where the electron went with the help of some equipment that can "see" light of these longer wavelengths. If we use "gentler" light perhaps we can avoid disturbing the electrons so much.

Let us try the experiment with longer waves. We shall keep repeating our experiment, each time with light of a longer wavelength. At first, nothing seems to change. The results are the same. Then a terrible thing happens. You remember that when we discussed the microscope we pointed out that, due to the *wave nature* of the light, there is a limitation on how close two spots can be and still be seen as two separate spots. This distance is of the order of the wavelength of light. So now, when we make the wavelength longer than the distance between our holes, we see a *big* fuzzy flash when the light is scattered by the electrons. We can no longer tell which hole the electron went through! We just know it went somewhere! And it is just with light of this color that we find that the jolts given to the electron are small enough so that P'_{12} begins to look like P_{12}—that we begin to get some interference effect. And it is only for wavelengths much longer than the separation of the two holes (when we have no chance at all of telling where the electron went) that the disturbance due to the light gets sufficiently small that we again get the curve P_{12} shown in Fig. 6-3.

In our experiment we find that it is impossible to arrange the light in such a way that one can tell which hole the electron went through, and at the same time not disturb the pattern. It was suggested by Heisenberg that the then-new laws of nature could only be consistent if there were some basic limitation on our experimental capabilities not previously recognized. He proposed, as a general principle, his *uncertainty principle,* which we can state in terms of our experiment as follows: "It is impossible to design an apparatus

to determine which hole the electron passes through, that will not at the same time disturb the electrons enough to destroy the interference pattern." If an apparatus is capable of determining which hole the electron goes through, it *cannot* be so delicate that it does not disturb the pattern in an essential way. No one has ever found (or even thought of) a way around the uncertainty principle. So we must assume that it describes a basic characteristic of nature.

The complete theory of quantum mechanics which we now use to describe atoms and, in fact, all matter depends on the correctness of the uncertainty principle. Since quantum mechanics is such a successful theory, our belief in the uncertainty principle is reinforced. But if a way to "beat" the uncertainty principle were ever discovered, quantum mechanics would give inconsistent results and would have to be discarded as a valid theory of nature.

"Well," you say, "what about Proposition A? It is true, or is it *not* true, that the electron either goes through hole 1 or it goes through hole 2?" The only answer that can be given is that we have found from experiment that there is a certain special way that we have to think in order that we do not get into inconsistencies. What we must say (to avoid making wrong predictions) is the following: If one looks at the holes or, more accurately, if one has a piece of apparatus which is capable of determining whether the electrons go through hole 1 or hole 2, then one *can* say that it goes either through hole 1 or hole 2. *But,* when one does *not* try to tell which way the electron goes, when there is nothing in the experiment to disturb the electrons, then one may *not* say that an electron goes either through hole 1 or hole 2. If one does say that, and starts to make any deductions from the statement, he will make errors in the analysis. This is the logical tightrope on which we must walk if we wish to describe nature successfully.

◐ ◐ ◐

If the motion of all matter—as well as electrons—must be described in terms of waves, what about the bullets in our first experiment? Why didn't we see an interference pattern there? It turns

out that for the bullets the wavelengths were so tiny that the interference patterns became very fine. So fine, in fact, that with any detector of finite size one could not distinguish the separate maxima and minima. What we saw was only a kind of average, which is the classical curve. In Fig. 6-5 we have tried to indicate schematically what happens with large-scale objects. Part (a) of the figure shows the probability distribution one might predict for bullets, using quantum mechanics. The rapid wiggles are supposed to represent the interference pattern one gets for waves of very short wavelength. Any physical detector, however, straddles several wiggles of the

Figure 6-5 Interference pattern with bullets: (a) actual (schematic), (b) observed.

probability curve, so that the measurements show the smooth curve drawn in part (b) of the figure.

First principles of quantum mechanics

We will now write a summary of the main conclusions of our experiments. We will, however, put the results in a form which makes them true for a general class of such experiments. We can write our

summary more simply if we first define an "ideal experiment" as one in which there are no uncertain external influences, i.e., no jiggling or other things going on that we cannot take into account. We would be quite precise if we said: "An ideal experiment is one in which all of the initial and final conditions of the experiment are completely specified." What we will call "an event" is, in general, just a specific set of initial and final conditions. (For example: "An electron leaves the gun, arrives at the detector, and nothing else happens.") Now for our summary.

Summary

(1) The probability of an event in an ideal experiment is given by the square of the absolute value of a complex number ϕ which is called the probability amplitude.

$$\begin{aligned} P &= \text{probability}, \\ \phi &= \text{probability amplitude}, \\ P &= |\phi|^2. \end{aligned} \tag{6.6}$$

(2) When an event can occur in several alternative ways, the probability amplitude for the event is the sum of the probability amplitudes for each way considered separately. There is interference.

$$\begin{aligned} \phi &= \phi_1 + \phi_2, \\ P &= |\phi_1 + \phi_2|^2. \end{aligned} \tag{6.7}$$

(3) If an experiment is performed which is capable of determining whether one or another alternative is actually taken, the probability of the event is the sum of the probabilities for each alternative. The interference is lost.

$$P = P_1 + P_2. \tag{6.8}$$

One might still like to ask: "How does it work? What is the machinery behind the law?" No one has found any machinery behind the law. No one can "explain" any more than we have just "explained." No one will give you any deeper representation of the situation. We have no ideas about a more basic mechanism from which these results can be deduced.

We would like to emphasize a very important difference between classical and quantum mechanics. We have been talking about the probability that an electron will arrive in a given circumstance. We have implied that in our experimental arrangement (or even in the best possible one) it would be impossible to predict exactly what would happen. We can only predict the odds! This would mean, if it were true, that physics has given up on the problem of trying to predict exactly what will happen in a definite circumstance. Yes! Physics *has* given up. *We do not know how to predict what would happen in a given circumstance,* and we believe now that it is impossible, that the only thing that can be predicted is the probability of different events. It must be recognized that this is a retrenchment in our earlier ideal of understanding nature. It may be a backward step, but no one has seen a way to avoid it.

We make now a few remarks on a suggestion that has sometimes been made to try to avoid the description we have given: "Perhaps the electron has some kind of internal works—some inner variables—that we do not yet know about. Perhaps that is why we cannot predict what will happen. If we could look more closely at the electron we would be able to tell where it would end up." So far as we know, that is impossible. We would still be in difficulty. Suppose we were to assume that inside the electron there is some kind of machinery that determines where it is going to end up. That machine must *also* determine which hole it is going to go through on its way. But we must not forget that what is inside the electron should not be dependent on what *we* do, and in particular upon whether we open or close one of the holes. So if an electron, before it starts, has already made up its mind (a) which hole it is going to use, and (b) where it is going

to land, we should find P_1 for those electrons that have chosen hole 1, P_2 for those that have chosen hole 2, *and necessarily* the sum $P_1 + P_2$ for those that arrive through the two holes. There seems to be no way around this. But we have verified experimentally that that is not the case. And no one has figured a way out of this puzzle. So at the present time we must limit ourselves to computing probabilities. We say "at the present time," but we suspect very strongly that it is something that will be with us forever—that it is impossible to beat that puzzle—that this is the way nature really *is*.

The uncertainty principle

This is the way Heisenberg stated the uncertainty principle originally: If you make the measurement on any object, and you can determine the x-component of its momentum with an uncertainty Δp, you cannot, at the same time, know its x-position more accurately than $\Delta x \geq \hbar/2\Delta p$. The uncertainties in the position and momentum at any instant must have their product greater than half the reduced Planck constant. This is a special case of the uncertainty principle that was stated above more generally. The more general statement was that one cannot design equipment in any way to determine which of two alternatives is taken, without, at the same time, destroying the pattern of interference.

Let us show for one particular case that the kind of relation given by Heisenberg must be true in order to keep from getting into trouble. We imagine a modification of the experiment of Fig. 6-3, in which the wall with the holes consists of a plate mounted on rollers so that it can move freely up and down (in the x-direction), as shown in Fig. 6-6. By watching the motion of the plate carefully we can try to tell which hole an electron goes through. Imagine what happens when the detector is placed at $x = 0$. We would expect that an electron which passes through hole 1 must be deflected downward by the plate to reach the detector. Since the vertical component of the electron momentum is changed, the plate must recoil with an equal momentum in the opposite direction. The plate will

Figure 6-6 An experiment in which the recoil of the wall is measured.

get an upward kick. If the electron goes through the lower hole, the plate should feel a downward kick. It is clear that for every position of the detector, the momentum received by the plate will have a different value for a traversal via hole 1 than for a traversal via hole 2. So! Without disturbing the electrons *at all,* but just by watching the *plate,* we can tell which path the electron used.

Now in order to do this it is necessary to know what the momentum of the screen is, before the electron goes through. So when we measure the momentum after the electron goes by, we can figure out how much the plate's momentum has changed. But remember, according to the uncertainty principle we cannot at the same time know the position of the plate with an arbitrary accuracy. But if we do not know exactly *where* the plate is we cannot say precisely where the two holes are. They will be in a different place for every electron that goes through. This means that the center of our interference pattern will have a different location for each electron. The wiggles of the interference pattern will be smeared out. We shall show quantitatively in the next chapter that if we determine the momentum of the plate sufficiently accurately to determine from the recoil measurement which hole was used, then the uncertainty in the x-position of the plate will, according to the uncertainty principle, be enough to shift the pattern observed at the

detector up and down in the x-direction about the distance from a maximum to its nearest minimum. Such a random shift is just enough to smear out the pattern so that no interference is observed.

The uncertainty principle "protects" quantum mechanics. Heisenberg recognized that if it were possible to measure the momentum and the position simultaneously with a greater accuracy, the quantum mechanics would collapse. So he proposed that it must be impossible. Then people sat down and tried to figure out ways of doing it, and nobody could figure out a way to measure the position and the momentum of anything—a screen, an electron, a billiard ball, anything—with any greater accuracy. Quantum mechanics maintains its perilous but accurate existence.

1

VECTORS

1-1 Symmetry in physics

In this chapter we introduce a subject that is technically known in physics as *symmetry in physical law*. The word "symmetry" is used here with a special meaning, and therefore needs to be defined. When is a thing symmetrical—how can we define it? When we have a picture that is symmetrical, one side is somehow the same as the other side. Professor Hermann Weyl has given this definition of symmetry: a thing is symmetrical if one can subject it to a certain operation and it appears exactly the same after the operation. For instance, if we look at a silhouette of a vase that is left-and-right symmetrical, then turn it 180° around the vertical axis, it looks the same. We shall adopt the definition of symmetry in Weyl's more general form, and in that form we shall discuss symmetry of physical laws.

Suppose we build a complex machine in a certain place, with a lot of complicated interactions, and balls bouncing around with forces between them, and so on. Now suppose we build exactly the same kind of equipment at some other place, matching part by part, with the same dimensions and the same orientation, everything the same only displaced laterally by some distance. Then, if we start the two machines in the same initial circumstances, in exact correspondence, we ask: Will one machine behave exactly the same as the other? Will it follow all the motions in exact parallelism? Of course the answer may well be *no*, because if we choose the wrong

place for our machine it might be inside a wall and interferences from the wall would make the machine not work.

All of our ideas in physics require a certain amount of common sense in their application; they are not purely mathematical or abstract ideas. We have to understand what we mean when we say that the phenomena are the same when we move the apparatus to a new position. We mean that we move everything that we believe is relevant; if the phenomenon is not the same, we suggest that something relevant has not been moved, and we proceed to look for it. If we never find it, then we claim that the laws of physics do not have this symmetry. On the other hand, we may find it—we expect to find it—if the laws of physics do have this symmetry; looking around, we may discover, for instance, that the wall is pushing on the apparatus. The basic question is, if we define things well enough, if all the essential forces are included inside the apparatus, if all the relevant parts are moved from one place to another, will the laws be the same? Will the machinery work the same way?

It is clear that what we want to do is to move all the equipment and *essential* influences, but not *everything* in the world—planets, stars, and all—for if we do that, we have the same phenomenon again for the trivial reason that we are right back where we started. No, we cannot move *everything*. But it turns out in practice that with a certain amount of intelligence about what to move, the machinery will work. In other words, if we do not go inside a wall, if we know the origin of the outside forces, and arrange that those are moved too, then the machinery *will* work the same in one location as in another.

1-2 Translations

We shall limit our analysis to just mechanics, for which we now have sufficient knowledge. In previous chapters we have seen that the laws of mechanics can be summarized by a set of three equations for each particle:

$$m(d^2x/dt^2) = F_x, \quad m(d^2y/dt^2) = F_y, \quad m(d^2z/dt^2) = F_z. \quad (1.1)$$

Now this means that there exists a way to *measure x, y,* and z on three perpendicular axes, and the forces along those directions, such that these laws are true. These must be measured from some origin, but *where do we put the origin*? All that Newton would tell us at first is that there *is* some place that we can measure from, perhaps the center of the universe, such that these laws are correct. But we can show immediately that we can never find the center, because if we use some other origin it would make no difference. In other words, suppose that there are two people—Joe, who has an origin in one place, and Moe, who has a parallel system whose origin is somewhere else (Figure 1-1). Now when Joe measures the location of the point in space, he finds it at x, y, and z (we shall usually leave z out because it is too confusing to draw in a picture). Moe, on the other hand, when measuring the same point, will obtain a different x (in order to distinguish it, we will call it x'), and in principle a different y, although in our example they are numerically equal. So we have

$$x' = x - a, \quad y' = y, \quad z' = z. \tag{1.2}$$

Now in order to complete our analysis we must know what Moe would obtain for the forces. The force is supposed to act along some line, and by the force in the x-direction we mean the part of the total which is in the x-direction, which is the magnitude of the force times this cosine of its angle with the x-axis. Now we see that

Figure 1-1 Two parallel coordinate systems.

Moe would use exactly the same projection as Joe would use, so we have a set of equations

$$F_{x'} = F_x, \quad F_{y'} = F_y, \quad F_{z'} = F_z. \tag{1.3}$$

These would be the relationships between quantities as seen by Joe and Moe.

The question is, if Joe knows Newton's laws, and if Moe tries to write down Newton's laws, will they also be correct for him? Does it make any difference from which origin we measure the points? In other words, assuming that equations (1.1) are true, and the Eqs. (1.2) and (1.3) give the relationship of the measurements, is it or is it not true that

$$\begin{aligned} &\text{(a)} \quad m(d^2x'/dt^2) = F_{x'}, \\ &\text{(b)} \quad m(d^2y'/dt^2) = F_{y'}, \\ &\text{(c)} \quad m(d^2z'/dt^2) = F_{z'}? \end{aligned} \tag{1.4}$$

In order to test these equations we shall differentiate the formula for x' twice. First of all

$$\frac{dx'}{dt} = \frac{d}{dt}(x - a) = \frac{dx}{dt} - \frac{da}{dt}.$$

Now we shall assume that Moe's origin is fixed (not moving) relative to Joe's; therefore a is a constant and $da/dt = 0$, so we find that

$$dx'/dt = dx/dt$$

and therefore

$$d^2x'/dt^2 = d^2x/dt^2;$$

therefore we know that Eq. (1.4a) becomes

$$m(d^2x/dt^2) = F_{x'}.$$

(We also suppose that the masses measured by Joe and Moe are equal.) Thus the acceleration times the mass is the same as the other fellow's. We have also found the formula for $F_{x'}$, for, substituting from Eq. (1.1), we find that

$$F_{x'} = F_x.$$

Therefore the laws as seen by Moe appear the same; he can write Newton's laws too, with different coordinates, and they will still be right. That means that there is no unique way to define the origin of the world, because the laws will appear the same, from whatever position they are observed.

This is also true: if there is a piece of equipment in one place with a certain kind of machinery in it, the same equipment in another place will behave in the same way. Why? Because one machine, when analyzed by Moe, has exactly the same equations as the other one, analyzed by Joe. Since the *equations* are the same, the *phenomena* appear the same. So the proof that an apparatus in a new position behaves the same as it did in the old position is the same as the proof that the equations when displaced in space reproduce themselves. Therefore we say that *the laws of physics are symmetrical for translational displacements*, symmetrical in the sense that the laws do not change when we make a translation of our coordinates. Of course it is quite obvious intuitively that this is true, but it is interesting and entertaining to discuss the mathematics of it.

1-3 Rotations

The above is the first of a series of ever more complicated propositions concerning the symmetry of a physical law. The next proposition is that it should make no difference in which *direction* we choose the axes. In other words, if we build a piece of equipment in some place and watch it operate, and nearby we build the same kind of apparatus but put it up on an angle, will it operate in the same way? Obviously it will not if it is a grandfather clock, for example! If a pendulum clock stands upright, it works fine, but if it is tilted the pendulum falls against the side of the case and nothing happens. The theorem is then false in the case of the pendulum clock, unless we include the earth, which is pulling on the pendulum. Therefore we can make a prediction about pendulum clocks if we believe in the symmetry of physical law for rotation: something else is involved in the operation of a pendulum clock besides the machinery of the clock, something outside it that we should

look for. We may also predict that pendulum clocks will not work the same way when located in different places relative to this mysterious source of asymmetry, perhaps the earth. Indeed, we know that a pendulum clock up in an artificial satellite, for example, would not tick either, because there is no effective force, and on Mars it would go at a different rate. Pendulum clocks *do* involve something more than just the machinery inside, they involve something on the outside. Once we recognize this factor, we see that we must turn the earth along with the apparatus. Of course we do not have to worry about that, it is easy to do; one simply waits a moment or two and the earth turns; then the pendulum clock ticks again in the new position the same as it did before. While we are rotating in space our angles are always changing, absolutely; this change does not seem to bother us very much, for in the new position we seem to be in the same condition as in the old. This has a certain tendency to confuse one, because it is true that in the new turned position the laws are the same as in the unturned position, but it is *not* true that *as we turn* a thing it follows the same laws as it does when we are not turning it. If we perform sufficiently delicate experiments, we can tell that the earth *is rotating*, but not that it *had rotated*. In other words, we cannot locate its angular position, but we can tell that it is changing.

Now we may discuss the effects of angular orientation upon physical laws. Let us find out whether the same game with Joe and Moe works again. This time, to avoid needless complication, we shall suppose that Joe and Moe use the same origin (we have already shown that the axes can be moved by translation to another place). Assume that Moe's axes have rotated relative to Joe's by an angle θ. The two coordinate systems are shown in Figure 1-2, which is restricted to two dimensions. Consider any point P having coordinates (x, y) in Joe's system and (x', y') in Moe's system. We shall begin, as in the previous case, by expressing the coordinates x' and y' in terms of x, y, and θ. To do so, we first drop perpendiculars from P to all four axes and draw AB perpendicular to PQ. Inspection of the figure shows that x' can be written as the sum of two

Figure 1-2 Two coordinate systems having different angular orientations.

lengths along the x'-axis, and y' as the difference of two lengths along AB. All these lengths are expressed in terms of x, y, and θ in equations (1.5), to which we have added an equation for the third dimension.

$$\begin{aligned} x' &= x \cos\theta + y \sin\theta, \\ y' &= y \cos\theta - x \sin\theta, \\ z' &= z. \end{aligned} \tag{1.5}$$

The next step is to analyze the relationship of forces as seen by the two observers, following the same general method as before. Let us assume that a force F, which has already been analyzed as having components F_x and F_y (as seen by Joe), is acting on a particle of mass m, located at point P in Figure 1-2. For simplicity, let us move both sets of axes so that the origin is at P, as shown in Figure 1-3. Moe sees the components of F along his axes as $F_{x'}$, and $F_{y'}$. F_x has components along both the x'- and y'-axes, and F_y likewise has components along both these axes. To express $F_{x'}$, in terms of F_x and F_y, we sum these components along the x'-axis, and in a like manner we can express $F_{y'}$ in terms of F_x and F_y. The results are

$$\begin{aligned} F_{x'} &= F_x \cos\theta + F_y \sin\theta, \\ F_{y'} &= F_y \cos\theta - F_x \sin\theta, \\ F_{z'} &= F_z. \end{aligned} \tag{1.6}$$

Figure 1-3 Components of a force in the two systems.

It is interesting to note an accident of sorts, which is of extreme importance: the formulas (1.5) and (1.6), for coordinates of P and components of F, respectively, *are of identical form*.

As before, Newton's laws are assumed to be true in Joe's system, and are expressed by equations (1.1). The question, again, is whether Moe can apply Newton's laws—will the results be correct for his system of rotated axes? In other words, if we assume that Eqs. (1.5) and (1.6) give the relationship of the measurements, is it true or not true that

$$\begin{aligned} m(d^2x'/dt^2) &= F_{x'}, \\ m(d^2y'/dt^2) &= F_{y'}, \\ m(d^2z'/dt^2) &= F_{z'}? \end{aligned} \tag{1.7}$$

To test these equations, we calculate the left and right sides independently, and compare the results. To calculate the left sides, we multiply equations (1.5) by m, and differentiate twice with respect to time, assuming the angle θ to be constant. This gives

$$\begin{aligned} m(d^2x'/dt^2) &= m(d^2x/dt^2)\cos\theta + m(d^2y/dt^2)\sin\theta, \\ m(d^2y'/dt^2) &= m(d^2y/dt^2)\cos\theta - m(d^2x/dt^2)\sin\theta, \\ m(d^2z'/dt^2) &= m(d^2z/dt^2). \end{aligned} \tag{1.8}$$

We calculate the right sides of equations (1.7) by substituting equations (1.1) into equations (1.6). This gives

$$
\begin{aligned}
F_{x'} &= m(d^2x/dt^2)\cos\theta + m(d^2y/dt^2)\sin\theta,\\
F_{y'} &= m(d^2y/dt^2)\cos\theta - m(d^2x/dt^2)\sin\theta,\\
F_{z'} &= m(d^2z/dt^2).
\end{aligned}
\tag{1.9}
$$

Behold! The right sides of Eqs. (1.8) and (1.9) are identical, so we conclude that if Newton's laws are correct on one set of axes, they are also valid on any other set of axes. This result, which has now been established for both translation and rotation of axes, has certain consequences: First, no one can claim his particular axes are unique, but of course they can be more *convenient* for certain particular problems. For example, it is handy to have gravity along one axis, but this is not physically necessary. Second, it means that any piece of equipment which is completely self-contained, with all the force-generating equipment completely inside the apparatus, would work the same when turned at an angle.

1-4 Vectors

Not only Newton's laws, but also the other laws of physics, so far as we know today, have the two properties which we call invariance (or symmetry) under translation of axes and rotation of axes. These properties are so important that a mathematical technique has been developed to take advantage of them in writing and using physical laws.

The foregoing analysis involved considerable tedious mathematical work. To reduce the details to a minimum in the analysis of such questions, a very powerful mathematical machinery has been devised. This system, called *vector analysis*, supplies the title of this chapter; strictly speaking, however, this is a chapter on the symmetry of physical laws. By the methods of the preceding analysis, we were able to do everything required for obtaining the results that we sought, but in practice we should like to do things more easily and rapidly, so we employ the vector technique.

We begin by noting some characteristics of two kinds of quantities that are important in physics. (Actually there are more than

two, but let us start out with two.) One of them, like the number of potatoes in a sack, we call an ordinary quantity, or an undirected quantity, or a *scalar*. Temperature is an example of such a quantity. Other quantities that are important in physics do have direction, for instance velocity: we have to keep track of which way a body is going, not just its speed. Momentum and force also have direction, as does displacement: when someone steps from one place to another in space, we can keep track of how far he went, but if we wish also to know *where* he went, we have to specify a direction.

All quantities that have a direction, like a step in space, are called *vectors*.

A vector is three numbers. In order to represent a step in space, say from the origin to some particular point P whose location is (x, y, z), we really need three numbers, but we are going to invent a single mathematical symbol, $\mathbf{r}$, which is unlike any other mathematical symbols we have so far used.* It is *not* a single number, it represents *three* numbers: x, y, and z. It means three numbers, but not really only *those* three numbers, because if we were to use a different coordinate system, the three numbers would be changed to x', y', and z'. However, we want to keep our mathematics simple and so we are going to use the *same mark* to represent the three numbers (x, y, z) and the three numbers (x', y', z'). That is, we use the same mark to represent the first set of three numbers for one coordinate system, but the second set of three numbers if we are using the other coordinate system. This has the advantage that when we change the coordinate system, we do not have to change the letters of our equations. If we write an equation in terms of x, y, z, and then use another system, we have to change to x', y', z', but we shall just write $\mathbf{r}$, with the convention that it represents (x, y, z) if we use one set of axes, or (x', y', z') if we use another set of axes, and so on. The three numbers which describe the quantity in a given coordinate system are called the *components* of the vector

* In type, vectors are represented by boldface; in handwritten form an arrow is used: $\vec{r}$.

in the direction of the coordinate axes of that system. That is, we use the same symbol for the three letters that correspond to the *same object*, *as seen from different axes*. The very fact that we can say "the same object" implies a physical intuition about the reality of a step in space, that is independent of the components in terms of which we measure it. So the symbol **r** will represent the same thing no matter how we turn the axes.

Now suppose there is another directed physical quantity, any other quantity, which also has three numbers associated with it, like force, and these three numbers change to three other numbers by a certain mathematical rule, if we change the axes. It must be the same rule that changes (x, y, z) into (x', y', z'). In other words, any physical quantity associated with three numbers which transform as do the components of a step in space is a vector. An equation like

$$\mathbf{F} = \mathbf{r}$$

would thus be true in *any* coordinate system if it were true in one. This equation, of course, stands for the three equations

$$F_x = x, \quad F_y = y, \quad F_z = z,$$

or, alternatively, for

$$F_{x'} = x', \quad F_{y'} = y', \quad F_{z'} = z'.$$

The fact that a physical relationship can be expressed as a vector equation assures us the relationship is unchanged by a mere rotation of the coordinate system. That is the reason why vectors are so useful in physics.

Now let us examine some of the properties of vectors. As examples of vectors we may mention velocity, momentum, force, and acceleration. For many purposes it is convenient to represent a vector quantity by an arrow that indicates the direction in which it is acting. Why can we represent force, say, by an arrow? Because it has the same mathematical transformation properties as a "step in space." We thus represent it in a diagram as if it were a step, using a scale such that one unit of force, or one newton, corresponds to

a certain convenient length. Once we have done this, all forces can be represented as lengths, because an equation like

$$\mathbf{F} = k\mathbf{r},$$

where k is some constant, is a perfectly legitimate equation. Thus we can always represent forces by lines, which is very convenient, because once we have drawn the line we no longer need the axes. Of course, we can quickly calculate the three components as they change upon turning the axes, because that is just a geometric problem.

1-5 Vector algebra

Now we must describe the laws, or rules, for combining vectors in various ways. The first such combination is the *addition* of two vectors: suppose that **a** is a vector which in some particular coordinate system has the three components (a_x, a_y, a_z), and that **b** is another vector which has the three components (b_x, b_y, b_z). Now let us invent three new numbers $(a_x + b_x, a_y + b_y, a_z + b_z)$. Do these form a vector? "Well," we might say, "they are three numbers, and every three numbers form a vector." No, *not* every three numbers form a vector! In order for it to be a vector, not only must there be three numbers, but these must be associated with a coordinate system in such a way that if we turn the coordinate system, the three numbers "revolve" on each other, get "mixed up" in each other, by the precise laws we have already described. So the question is, if we now rotate the coordinate system so that (a_x, a_y, a_z) become $(a_{x'}, a_{y'}, a_{z'})$ and (b_x, b_y, b_z) become $(b_{x'}, b_{y'}, b_{z'})$, what do $(a_x + b_x, a_y + b_y, a_z + b_z)$ become? Do they become $(a_{x'} + b_{x'}, a_{y'} + b_{y'}, a_{z'} + b_{z'})$ or not? The answer is, of course, yes, because the prototype transformations of Eq. (1.5) constitute what we call a *linear* transformation. If we apply those transformations to a_x and b_x to get $a_{x'}$, $+ b_{x'}$, we find that the transformed $a_x + b_x$ is indeed the same as $a_{x'} + b_{x'}$. When **a** and **b** are "added together" in this sense, they will form a vector which we may call **c**. We would write this as

$$\mathbf{c} = \mathbf{a} + \mathbf{b}.$$

Now **c** has the interesting property

$$\mathbf{c} = \mathbf{b} + \mathbf{a},$$

as we can immediately see from its components. Thus also,

$$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}.$$

We can add vectors in any order.

What is the geometric significance of **a** + **b**? Suppose that **a** and **b** were represented by lines on a piece of paper, what would **c** look like? This is shown in Figure 1-4. We see that we can add the components of **b** to those of **a** most conveniently if we place the rectangle representing the components of **b** next to that representing the components of **a** in the manner indicated. Since **b** just "fits" into its rectangle, as does **a** into its rectangle, this is the same as putting the "tail" of **b** on the "head" of **a**, the arrow from the "tail" of **a** to the "head" of **b** being the vector **c**. Of course, if we added **a** to **b** the other way around, we would put the "tail" of **a** on the "head" of **b**, and by the geometrical properties of parallelograms we would get the same result for **c**. Note that vectors can be added in this way without reference to any coordinate axes.

Suppose we multiply a vector by a number α, what does this mean? We *define* it to mean a new vector whose components are

Figure 1-4 The addition of vectors.

αa_x, αa_y, and αa_z. We leave it as a problem for the student to prove that it *is* a vector.

Now let us consider vector subtraction. We may define subtraction in the same way as addition, but instead of adding, we subtract the components. Or we might define subtraction by defining a negative vector, $-\mathbf{b} = -1\mathbf{b}$, and then we would add the components. It comes to the same thing. The result is shown in Figure 1-5. This figure shows $\mathbf{d} = \mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$; we also note that the difference $\mathbf{a} - \mathbf{b}$ can be found very easily from $\mathbf{a}$ and $\mathbf{b}$ by using the equivalent relation $\mathbf{a} = \mathbf{b} + \mathbf{d}$. Thus the difference is even easier to find than the sum: we just draw the vector from $\mathbf{b}$ to $\mathbf{a}$, to get $\mathbf{a} - \mathbf{b}$!

Next we discuss velocity. Why is velocity a vector? If position is given by the three coordinates (x, y, z), what is the velocity? The velocity is given by dx/dt, dy/dt, and dz/dt. Is that a vector, or not? We can find out by differentiating the expressions in Eq. (1.5) to find out whether dx'/dt *transforms* in the right way. We see that the components dx/dt and dy/dt *do* transform according to the same law as x and y, and therefore the time derivative is a vector. So the velocity *is* a vector. We can write the velocity in an interesting way as

$$\mathbf{v} = d\mathbf{r}/dt.$$

What the velocity is, and why it is a vector, can also be understood more pictorially: How far does a particle move in a short time Δt?

Figure 1-5 The subtraction of vectors.

Answer: $\Delta\mathbf{r}$, so if a particle is "here" at one instant and "there" at another instant, then the vector difference of the positions $\Delta\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$, which is in the direction of motion shown in Figure 1-6, divided by the time interval $\Delta t = t_2 - t_1$, is the "average velocity" vector.

In other words, by vector velocity we mean the limit, as Δt goes to 0, of the difference between the radius vectors at the time $t + \Delta t$ and the time t, divided by Δt:

$$\mathbf{v} = \lim_{\Delta t \to 0}(\Delta\mathbf{r}/\Delta t) = d\mathbf{r}/dt. \tag{1.10}$$

Thus velocity is a vector because it is the difference of two vectors. It is also the right definition of velocity because its components are dx/dt, dy/dt, and dz/dt. In fact, we see from this argument that if we differentiate *any* vector with respect to time we produce a new vector. So we have several ways of producing new vectors: (1) multiply by a constant, (2) differentiate with respect to time, (3) add or subtract two vectors.

1-6 Newton's laws in vector notation

In order to write Newton's laws in vector form, we have to go just one step further, and define the acceleration vector. This is the time derivative of the velocity vector, and it is easy to demonstrate that

Figure 1-6 The displacement of a particle in a short time interval $\Delta t = t_2 - t_1$.

its components are the second derivatives of x, y, and z with respect to t:

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \left(\frac{d}{dt}\right)\left(\frac{d\mathbf{r}}{dt}\right) = \frac{d^2\mathbf{r}}{dt^2}, \tag{1.11}$$

$$a_x = \frac{dv_x}{dt} = \frac{d^2x}{dt^2}, \quad a_y = \frac{dv_y}{dt} = \frac{d^2y}{dt^2}, \quad a_z = \frac{dv_z}{dt} = \frac{d^2z}{dt^2}. \tag{1.12}$$

With this definition, then, Newton's laws can be written in this way:

$$m\mathbf{a} = \mathbf{F} \tag{1.13}$$

or

$$m(d^2\mathbf{r}/dt^2) = \mathbf{F}. \tag{1.14}$$

Now the problem of proving the invariance of Newton's laws under rotation of coordinates is this: prove that $\mathbf{a}$ is a vector; this we have just done. Prove that $\mathbf{F}$ is a vector; we *suppose* it is. So if force is a vector, then, since we know acceleration is a vector, Eq. (1.13) will look the same in any coordinate system. Writing it in a form which does not explicitly contain x's, y's, and z's has the advantage that from now on we need not write *three* laws every time we write Newton's equations or other laws of physics. We write what looks like *one* law, but really, of course, it is the three laws for any particular set of axes, because any vector equation involves the statement that *each of the components is equal*.

The fact that the acceleration is the rate of change of the vector velocity helps us to calculate the acceleration in some rather complicated circumstances. Suppose, for instance, that a particle is moving on some complicated curve (Figure 1-7) and that, at a given instant t, it had a certain velocity $\mathbf{v}_1$, but that when we go to another instant t_2 a little later, it has a different velocity $\mathbf{v}_2$. What is the acceleration? answer: Acceleration is the difference in the velocity divided by the small time interval, so we need the difference of the two velocities. How do we get the difference of the velocities? To subtract two vectors, we put the vector across the ends of $\mathbf{v}_2$ and

Figure 1-7 A curved trajectory.

$\mathbf{v}_1$; that is, we draw $\Delta\mathbf{v}$ as the difference of the two vectors, right? *No!* That only works when the *tails* of the vectors are in the same place! It has no meaning if we move the vector somewhere else and then draw a line across, so watch out! We have to draw a new diagram to subtract the vectors. In Figure 1-8, $\mathbf{v}_1$ and $\mathbf{v}_2$ are both drawn parallel and equal to their counterparts in Figure 1-7, and now we can discuss the acceleration. Of course the acceleration is simply $\Delta\mathbf{v}/\Delta t$. It is interesting to note that we can compose the velocity difference out of two parts; we can think of acceleration as having *two components*, $\Delta\mathbf{v}_{||}$ in the direction tangent to the path and $\Delta\mathbf{v}_{\perp}$ at right angles to the path, as indicated in Figure 1-8. The acceleration tangent to the path is, of course, just the change in the *length* of the vector, i.e., the change in the *speed v*:

$$a_{||} = dv/dt. \tag{1.15}$$

The other component of acceleration, at right angles to the curve, is easy to calculate, using Figures 1-7 and 1-8. In the short time Δt let the change in angle between $\mathbf{v}_1$ and $\mathbf{v}_2$ be the small angle $\Delta\theta$. If

Figure 1-8 Diagram for calculating the acceleration.

the magnitude of the velocity is called v, then of course

$$\Delta v_\perp = v\Delta\theta$$

and the acceleration a will be

$$a_\perp = v(\Delta\theta/\Delta t).$$

Now we need to know $\Delta\theta/\Delta t$, which can be found this way: If, at the given moment, the curve is approximated as a circle of a certain radius R, then in a time Δt the distance s is, of course, $v\,\Delta t$, where v is the speed.

$$\Delta\theta = (v\Delta t)/R, \quad \text{or} \quad \Delta\theta/\Delta t = v/R.$$

Therefore, we find

$$a_\perp = v^2/R, \tag{1.16}$$

as we have seen before.

1-7 Scalar product of vectors

Now let us examine a little further the properties of vectors. It is easy to see that the *length* of a step in space would be the same in any coordinate system. That is, if a particular step $\mathbf{r}$ is represented by x, y, z, in one coordinate system, and by x', y', z' in another coordinate system, surely the distance $r = |\mathbf{r}|$ would be the same in both. Now

$$r = \sqrt{x^2 + y^2 + z^2}$$

and also

$$r' = \sqrt{x'^2 + y'^2 + z'^2}.$$

So what we wish to verify is that these two quantities are equal. It is much more convenient not to bother to take the square root, so let us talk about the square of the distance; that is, let us find out whether

$$x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2. \tag{1.17}$$

It had better be—and if we substitute Eq. (1.5) we do indeed find that it is. So we see that there are other kinds of equations which are true for any two coordinate systems.

Something new is involved. We can produce a new quantity, a function of x, y, and z, called a *scalar function*, a quantity which has no direction but which is the same in both systems. Out of a vector we can make a scalar. We have to find a general rule for that. It is clear what the rule is for the case just considered: add the squares of the components. Let us now define a new thing, which we call $\mathbf{a} \cdot \mathbf{a}$. This is not a vector, but a scalar; it is a number that is the same in all coordinate systems, and it is defined to be the sum of the squares of the three components of the vector:

$$\mathbf{a} \cdot \mathbf{a} = a_x^2 + a_y^2 + a_z^2. \tag{1.18}$$

Now you say, "But with what axes?" It does not depend on the axes, the answer is the same in *every* set of axes. So we have a new *kind* of quantity, a new *invariant* or *scalar* produced by one vector "squared." If we now define the following quantity for any two vectors $\mathbf{a}$ and $\mathbf{b}$:

$$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z, \tag{1.19}$$

we find that this quantity, calculated in the primed and unprimed systems, also stays the same. To prove it we note that it is true of $\mathbf{a} \cdot \mathbf{a}$, $\mathbf{b} \cdot \mathbf{b}$, and $\mathbf{c} \cdot \mathbf{c}$, where $\mathbf{c} = \mathbf{a} + \mathbf{b}$. Therefore the sum of the squares $(a_x + b_x)^2 + (a_y + b_y)^2 + (a_z + b_z)^2$ will be invariant:

$$(a_x + b_x)^2 + (a_y + b_y)^2 + (a_z + b_z)^2 = (a_{x'} + b_{x'})^2 + (a_{y'} + b_{y'})^2 + (a_{z'} + b_{z'})^2. \tag{1.20}$$

If both sides of this equation are expanded, there will be cross products of just the type appearing in Eq. (1.19), as well as the sums of squares of the components of $\mathbf{a}$ and $\mathbf{b}$. The invariance of terms of the form of Eq. (1.18) then leaves the cross product terms (1.19) invariant also.

The quantity $\mathbf{a} \cdot \mathbf{b}$ is called the *scalar product* of two vectors, $\mathbf{a}$ and $\mathbf{b}$, and it has many interesting and useful properties. For instance, it is easily proved that

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}. \tag{1.21}$$

Also, there is a simple geometrical way to calculate $\mathbf{a} \cdot \mathbf{b}$, without having to calculate the components of **a** and **b**: $\mathbf{a} \cdot \mathbf{b}$ is the product of the length of **a** and the length of **b** times the cosine of the angle between them. Why? Suppose that we choose a special coordinate system in which the x-axis lies along **a**; in those circumstances, the only component of **a** that will be there is a_x, which is of course the whole length of **a**. Thus Eq. (1.19) reduces to $\mathbf{a} \cdot \mathbf{b} = a_x b_x$ for this case, and this is the length of **a** times the component of **b** in the direction of **a**, that is, $b \cos \theta$:

$$\mathbf{a} \cdot \mathbf{b} = ab \cos \theta.$$

Therefore, in that special coordinate system, we have proved that $\mathbf{a} \cdot \mathbf{b}$ is the length of **a** times the length of **b** times $\cos \theta$. But *if it is true in one coordinate system, it is true in all*, because $\mathbf{a} \cdot \mathbf{b}$ is independent of the coordinate system; that is our argument.

What good is the dot product? Are there any cases in physics where we need it? Yes, we need it all the time. For instance, in Chapter 4* the kinetic energy was called $\frac{1}{2}mv^2$, but if the object is moving in space it should be the velocity squared in the x-direction, the y-direction, and the z-direction, and so the formula for kinetic energy according to vector analysis is

$$\text{K.E.} = \frac{1}{2}m(\mathbf{v} \cdot \mathbf{v}) = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2). \quad (1.22)$$

Energy does not have direction. Momentum has direction; it is a vector, and it is the mass times the velocity vector.

Another example of a dot product is the work done by a force when something is pushed from one place to the other. We have not yet defined work, but it is equivalent to the energy change, the weights lifted, when a force **F** acts through a distance **s**:

$$\text{Work} = \mathbf{F} \cdot \mathbf{s}. \quad (1.23)$$

It is sometimes very convenient to talk about the component of a vector in a certain direction (say the vertical direction because that is the direction of gravity). For such purposes, it is useful to

* of *The Feynman Lectures on Physics*, vol. I.

invent what we call a *unit vector* in the direction that we want to study. By a unit vector we mean one whose dot product with itself is equal to unity. Let us call this unit vector **i**; then $\mathbf{i} \cdot \mathbf{i} = 1$. Then, if we want the component of some vector in the direction of **i**, we see that the dot product $\mathbf{a} \cdot \mathbf{i}$ will be $a \cos \theta$, i.e., the component of **a** in the direction of **i**. This is a nice way to get the component; in fact, it permits us to get *all* the components and to write a rather amusing formula. Suppose that in a given system of coordinates, x, y, and z, we invent three vectors: **i**, a unit vector in the direction x; **j**, a unit vector in the direction y; and **k**, a unit vector in the direction z. Note first that $\mathbf{i} \cdot \mathbf{i} = 1$. What is $\mathbf{i} \cdot \mathbf{j}$? When two vectors are at right angles, their dot product is zero. Thus

$$\begin{aligned} &\mathbf{i} \cdot \mathbf{i} = 1 \\ &\mathbf{i} \cdot \mathbf{j} = 0 \quad \mathbf{j} \cdot \mathbf{j} = 1 \\ &\mathbf{i} \cdot \mathbf{k} = 0 \quad \mathbf{j} \cdot \mathbf{k} = 0 \quad \mathbf{k} \cdot \mathbf{k} = 1 \end{aligned} \tag{1.24}$$

Now with these definitions, any vector whatsoever can be written this way:

$$\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} = a_z\mathbf{k}. \tag{1.25}$$

By this means we can go from the components of a vector to the vector itself.

This discussion of vectors is by no means complete. However, rather than try to go more deeply into the subject now, we shall first learn to use in physical situations some of the ideas so far discussed. Then, when we have properly mastered this basic material, we shall find it easier to penetrate more deeply into the subject without getting too confused. We shall later find that it is useful to define another kind of product of two vectors, called the vector product, and written as $\mathbf{a} \times \mathbf{b}$. However, we shall undertake a discussion of such matters in a later chapter.

2

SYMMETRY IN PHYSICAL LAWS

2-1 Symmetry operations

The subject of this chapter is what we may call *symmetry in physical laws*. We have already discussed certain features of symmetry in physical laws in connection with vector analysis (Chapter 1), the theory of relativity (which follows in Chapter 4), and rotation (Chapter 20*).

Why should we be concerned with symmetry? In the first place, symmetry is fascinating to the human mind, and everyone likes objects or patterns that are in some way symmetrical. It is an interesting fact that nature often exhibits certain kinds of symmetry in the objects we find in the world around us. Perhaps the most symmetrical object imaginable is a sphere, and nature is full of spheres—stars, planets, water droplets in clouds. The crystals found in rocks exhibit many different kinds of symmetry, the study of which tells us some important things about the structure of solids. Even the animal and vegetable worlds show some degree of symmetry, although the symmetry of a flower or of a bee is not as perfect or as fundamental as is that of a crystal.

But our main concern here is not with the fact that the *objects*

* of *The Feynman Lectures on Physics*, vol. I.

of nature are often symmetrical. Rather, we wish to examine some of the even more remarkable symmetries of the universe—the symmetries that exist in the *basic laws themselves* which govern the operation of the physical world.

First, what *is* symmetry? How can a physical *law* be "symmetrical"? The problem of defining symmetry is an interesting one and we have already noted that Weyl gave a good definition, the substance of which is that a thing is symmetrical if there is something we can do to it so that after we have done it, it looks the same as it did before. For example, a symmetrical vase is of such a kind that if we reflect or turn it, it will look the same as it did before. The question we wish to consider here is what we can do to physical phenomena, or to a physical situation in an experiment, and yet leave the result the same. A list of the known operations under which various physical phenomena remain invariant is shown in Table 2-1.

2-2 Symmetry in space and time

The first thing we might try to do, for example, is to *translate* the phenomenon in space. If we do an experiment in a certain region, and then build another apparatus at another place in space (or move the original one over), then, whatever went on in one apparatus, in a certain order in time, will occur in the same way if we

Table 2-1 Symmetry Operations

Translation in space
Translation in time
Rotation through a fixed angle
Uniform velocity in a straight line (Lorentz transformation)
Reversal of time
Reflection of space
Interchange of identical atoms or identical particles
Quantum-mechanical phase
Matter-antimatter (charge conjugation)

have arranged the same condition with all due attention to the restrictions that we mentioned before: that all of those features of the environment which make it not behave the same way have also been moved over—we talked about how to define how much we should include in those circumstances, and we shall not go into those details again.

In the same way, we also believe today that *displacement in time* will have no effect on physical laws. (That is, *as far as we know today*—all of these things are as far as we know today!) That means that if we build a certain apparatus and start it at a certain time, say on Thursday at 10:00 A.M., and then build the same apparatus and start it, say, three days later in the same condition, the two apparatuses will go through the same motions in exactly the same way as a function of time no matter what the starting time, provided again, of course, that the relevant features of the environment are also modified appropriately in *time*. That symmetry means, of course, that if one bought General Motors stock three months ago, the same thing would happen to it if he bought it now!

We have to watch out for geographical differences too, for there are, of course, variations in the characteristics of the earth's surface. So, for example, if we measure the magnetic field in a certain region and move the apparatus to some other region, it may not work in precisely the same way because the magnetic field is different, but we say that is because the magnetic field is associated with the earth. We can imagine that if we move the whole earth and the equipment, it would make no difference in the operation of the apparatus.

Another thing that we discussed in considerable detail was rotation in space: if we turn an apparatus at an angle it works just as well, provided we turn everything else that is relevant along with it. In fact, we discussed the problem of symmetry under rotation in space in some detail in Chapter 1, and we invented a mathematical system called *vector analysis* to handle it as neatly as possible.

On a more advanced level we had another symmetry—the symmetry under uniform velocity in a straight line. That is to say—a rather remarkable effect—that if we have a piece of apparatus

working a certain way and then take the same apparatus and put it in a car, and move the whole car, plus all the relevant surroundings, at a uniform velocity in a straight line, then so far as the phenomena inside the car are concerned there is no difference: all the laws of physics appear the same. We even know how to express this more technically, and that is that the mathematical equations of the physical laws must be unchanged under a *Lorentz transformation*. As a matter of fact, it was a study of the relativity problem that concentrated physicists' attention most sharply on symmetry in physical laws.

Now the above-mentioned symmetries have all been of a geometrical nature, time and space being more or less the same, but there are other symmetries of a different kind. For example, there is a symmetry which describes the fact that we can replace one atom by another of the same kind; to put it differently, there *are* atoms of the same kind. It is possible to find groups of atoms such that if we change a pair around, it makes no difference—the atoms are identical. Whatever one atom of oxygen of a certain type will do, another atom of oxygen of that type will do. One may say, "That is ridiculous, that is the *definition* of equal types!" That may be merely the definition, but then we still do not know whether there *are* any "atoms of the same type"; the *fact* is that there are many, many atoms of the same type. Thus it does mean something to say that it makes no difference if we replace one atom by another of the same type. The so-called elementary particles of which the atoms are made are also identical particles in the above sense—all electrons are the same; all protons are the same; all positive pions are the same; and so on.

After such a long list of things that can be done without changing the phenomena, one might think we could do practically anything; so let us give some examples to the contrary, just to see the difference. Suppose that we ask: "Are the physical laws symmetrical under a change of scale?" Suppose we build a certain piece of apparatus, and then build another apparatus five times bigger in every part, will it work exactly the same way? The answer is, in this case,

no! The wavelength of light emitted, for example, by the atoms inside one box of sodium atoms and the wavelength of light emitted by a gas of sodium atoms five times in volume is not five times longer, but is in fact exactly the same as the other. So the ratio of the wavelength to the size of the emitter will change.

Another example: we see in the newspaper, every once in a while, pictures of a great cathedral made with little matchsticks—a tremendous work of art by some retired fellow who keeps gluing matchsticks together. It is much more elaborate and wonderful than any real cathedral. If we imagine that this wooden cathedral were actually built on the scale of a real cathedral, we see where the trouble is; it would not last—the whole thing would collapse because of the fact that scaled-up matchsticks are just not strong enough. "Yes," one might say, "but we also know that when there is an influence from the outside, it also must be changed in proportion!" We are talking about the ability of the object to withstand gravitation. So what we should do is first to take the model cathedral of real matchsticks and the real earth, and then we know it is stable. Then we should take the larger cathedral and take a bigger earth. But then it is even worse, because the gravitation is increased still more!

Today, of course, we understand the fact that phenomena depend on the scale on the grounds that matter is atomic in nature, and certainly if we built an apparatus that was so small there were only five atoms in it, it would clearly be something we could not scale up and down arbitrarily. The scale of an individual atom is not at all arbitrary—it is quite definite.

The fact that the laws of physics are not unchanged under a change of scale was discovered by Galileo. He realized that the strengths of materials were not in exactly the right proportion to their sizes, and he illustrated this property that we were just discussing, about the cathedral of matchsticks, by drawing two bones, the bone of one dog, in the right proportion for holding up his weight, and the imaginary bone of a "super dog" that would be, say, ten or a hundred times bigger—that bone was a big, solid thing

with quite different proportions. We do not know whether he ever carried the argument quite to the conclusion that the laws of nature must have a definite scale, but he was so impressed with this discovery that he considered it to be as important as the discovery of the laws of motion, because he published them both in the same volume, called "On Two New Sciences."

Another example in which the laws are not symmetrical, that we know quite well, is this: a system in rotation at a uniform angular velocity does not give the same apparent laws as one that is not rotating. If we make an experiment and then put everything in a spaceship and have the spaceship spinning in empty space, all alone at a constant angular velocity, the apparatus will not work the same way because, as we know, things inside the equipment will be thrown to the outside, and so on, by the centrifugal or Coriolis forces, etc. In fact, we can tell that the earth is rotating by using a so-called Foucault pendulum, without looking outside.

Next we mention a very interesting symmetry which is obviously false, i.e., *reversibility in time*. The physical laws apparently cannot be reversible in time, because, as we know, all obvious phenomena are irreversible on a large scale: "The moving finger writes, and having writ, moves on." So far as we can tell, this irreversibility is due to the very large number of particles involved, and if we could see the individual molecules, we would not be able to discern whether the machinery was working forwards or backwards. To make it more precise: we build a small apparatus in which we know what all the atoms are doing, in which we can watch them jiggling. Now we build another apparatus like it, but which starts its motion in the final condition of the other one, with all the velocities precisely reversed. *It will then go through the same motions, but exactly in reverse*. Putting it another way: if we take a motion picture, with sufficient detail, of all the inner works of a piece of material and shine it on a screen and run it backwards, no physicist will be able to say, "That is against the laws of physics, that is doing something wrong!" If we do not see all the details, of course, the situation will be perfectly clear. If we see the egg splattering on the sidewalk and

the shell cracking open, and so on, then we will surely say, "That is irreversible, because if we run the moving picture backwards the egg will all collect together and the shell will go back together, and that is obviously ridiculous!" But if we look at the individual atoms themselves, the laws look completely reversible. This is, of course, a much harder discovery to have made, but apparently it is true that the fundamental physical laws, on a microscopic and fundamental level, are completely reversible in time!

2-3 Symmetry and conservation laws

The symmetries of the physical laws are very interesting at this level, but they turn out, in the end, to be even more interesting and exciting when we come to quantum mechanics. For a reason which we cannot make clear at the level of the present discussion—a fact that most physicists still find somewhat staggering, a most profound and beautiful thing, is that, in quantum mechanics, *for each of the rules of symmetry there is a corresponding conservation law*; there is a definite connection between the laws of conservation and the symmetries of physical laws. We can only state this at present, without any attempt at explanation.

The fact, for example, that the laws are symmetrical for translation in space when we add the principles of quantum mechanics turns out to mean that *momentum is conserved*.

That the laws are symmetrical under translation in time means, in quantum mechanics, that *energy is conserved*.

Invariance under rotation through a fixed angle in space corresponds to the *conservation of angular momentum*. These connections are very interesting and beautiful things, among the most beautiful and profound things in physics.

Incidentally, there are a number of symmetries which appear in quantum mechanics which have no classical analog, which have no method of description in classical physics. One of these is as follows: If ψ is the amplitude for some process or other, we know that the absolute square of ψ is the probability that the process will

occur. Now if someone else were to make his calculations, not with this ψ, but with a ψ' which differs merely by a change in phase (let Δ be some constant, and multiply $e^{i\Delta}$ times the old $\boldsymbol{\psi}$), the absolute square of ψ', which is the probability of the event, is then equal to the absolute square of ψ:

$$\psi' = \psi e^{i\Delta}; \quad |\psi'|^2 = |\psi|^2. \tag{2.1}$$

Therefore the physical laws are unchanged if the phase of the wave function is shifted by an arbitrary constant. That is another symmetry. Physical laws must be of such a nature that a shift in the quantum-mechanical phase makes no difference. As we have just mentioned, in quantum mechanics there is a conservation law for every symmetry. The conservation law which is connected with the quantum-mechanical phase seems to be the *conservation of electrical charge*. This is altogether a very interesting business!

2-4 Mirror reflections

Now the next question, which is going to concern us for most of the rest of this chapter, is the question of symmetry under *reflection in space*. The problem is this: Are the physical laws symmetrical under reflection? We may put it this way: Suppose we build a piece of equipment, let us say a clock, with lots of wheels and hands and numbers; it ticks, it works, and it has things wound up inside. We look at the clock in the mirror. How it *looks* in the mirror is not the question. But let us actually *build* another clock which is exactly the same as the way the first clock looks in the mirror—every time there is a screw with a right-hand thread in one, we use a screw with a left-hand thread in the corresponding place of the other; where one is marked "2" on the face, we mark a "ᘔ" on the face of the other; each coiled spring is twisted one way in one clock and the other way in the mirror-image clock; when we are all finished, we have two clocks, both physical, which bear to each other the relation of an object and its mirror image, although they are both actual, material objects, we emphasize. Now the question is: If the

two clocks are started in the same condition, the springs wound to corresponding tightnesses, will the two clocks tick and go around, forever after, as exact mirror images? (This is a physical question, not a philosophical question.) Our intuition about the laws of physics would suggest that they *would*.

We would suspect that, at least in the case of these clocks, reflection in space is one of the symmetries of physical laws, that if we change everything from "right" to "left" and leave it otherwise the same, we cannot tell the difference. Let us, then, suppose for a moment that this is true. If it is true, then it would be impossible to distinguish "right" and "left" by any physical phenomenon, just as it is, for example, impossible to define a particular absolute velocity by a physical phenomenon. So it should be impossible, by any physical phenomenon, to define absolutely what we mean by "right" as opposed to "left," because the physical laws should be symmetrical.

Of course, the world does not *have* to be symmetrical. For example, using what we may call "geography," surely "right" can be defined. For instance, we stand in New Orleans and look at Chicago, and Florida is to our right (when our feet are on the ground!). So we can define "right" and "left" by geography. Of course, the actual situation in any system does not have to have the symmetry that we are talking about; it is a question of whether the *laws* are symmetrical—in other words, whether it is *against the physical laws* to have a sphere like the earth with "left-handed dirt" on it and a person like ourselves standing looking at a city like Chicago from a place like New Orleans, but with everything the other way around, so Florida is on the other side. It clearly seems not impossible, not against the physical laws, to have everything changed left for right.

Another point is that our definition of "right" should not depend on history. An easy way to distinguish right from left is to go to a machine shop and pick up a screw at random. The odds are it has a right-hand thread—not necessarily, but it is much more likely to have a right-hand thread than a left-hand one. This is a question of history or convention, or the way things happen to be, and is

again not a question of fundamental laws. As we can well appreciate, everyone could have started out making left-handed screws!

So we must try to find some phenomenon in which "right hand" is involved fundamentally. The next possibility we discuss is the fact that polarized light rotates its plane of polarization as it goes through, say, sugar water. As we saw in Chapter 33,* it rotates, let us say, to the right in a certain sugar solution. That is a way of defining "right-hand," because we may dissolve some sugar in the water and then the polarization goes to the right. But sugar has come from living things, and if we try to make the sugar artificially, then we discover that it *does not* rotate the plane of polarization! But if we then take that same sugar which is made artificially and which does not rotate the plane of polarization, and put bacteria in it (they eat some of the sugar) and then filter out the bacteria, we find that we still have sugar left (almost half as much as we had before), and this time it does rotate the plane of polarization, but *the other way*! It seems very confusing, but is easily explained.

Take another example: One of the substances which is common to all living creatures and that is fundamental to life is protein. Proteins consist of chains of amino acids. Figure 2-1 shows a model of an amino acid that comes out of a protein. This amino acid is called alanine, and the molecular arrangement would look like that in Figure 2-1(a) if it came out of a protein of a real living thing. On the other hand, if we try to make alanine from carbon dioxide, ethane, and ammonia (and we *can* make it, it is not a complicated molecule), we discover that we are making equal amounts of this molecule and the one shown in Figure 2-1(b)! The first molecule, the one that comes from the living thing, is called *L-alanine*. The other one, which is the same chemically, in that it has the same kinds of atoms and the same connections of the atoms, is a "right-hand" molecule, compared with the "left-hand" L-alanine, and it is called *D-alanine*. The interesting thing is that when we make alanine at

* of *The Feynman Lectures on Physics*, vol. I.

Figure 2-1 (a) L-alanine (left), and (b) D-alanine (right).

home in a laboratory from simple gases, we get an equal mixture of both kinds. However, the only thing that life uses is L-alanine. (This is not exactly true. Here and there in living creatures there is a special use for D-alanine, but it is very rare. All proteins use L-alanine exclusively.) Now if we make both kinds, and we feed the mixture to some animal which likes to "eat," or use up, alanine, it cannot use D-alanine, so it only uses the L-alanine; that is what happened to our sugar—after the bacteria eat the sugar that works well for them, only the "wrong" kind is left! (Left-handed sugar tastes sweet, but not the same as right-handed sugar.)

So it looks as though the phenomena of life permit a distinction between "right" and "left," or chemistry permits a distinction, because the two molecules are chemically different. But no, it does not! So far as physical measurements can be made, such as of energy, the rates of chemical reactions, and so on, the two kinds work exactly the same way if we make everything else in a mirror image too. One molecule will rotate light to the right, and the other will rotate it to the left in precisely the same amount, through the same amount of fluid. Thus, so far as physics is concerned, these two amino acids are equally satisfactory. So far as we understand things today, the fundamentals of the Schrödinger equation have it that the two molecules should behave in exactly corresponding ways, so

that one is to the right as the other is to the left. Nevertheless, in life it is all one way!

It is presumed that the reason for this is the following. Let us suppose, for example, that life is somehow at one moment in a certain condition in which all the proteins in some creatures have left-handed amino acids, and all the enzymes are lopsided—every substance in the living creature is lopsided—it is not symmetrical. So when the digestive enzymes try to change the chemicals in the food from one kind to another, one kind of chemical "fits" into the enzyme, but the other kind does not (like Cinderella and the slipper, except that it is a "left foot" that we are testing). So far as we know, in principle, we could build a frog, for example, in which every molecule is reversed, everything is like the "left-hand" mirror image of a real frog; we have a left-hand frog. This left-hand frog would go on all right for a while, but he would find nothing to eat, because if he swallows a fly, his enzymes are not built to digest it. The fly has the wrong "kind" of amino acids (unless we give him a left-hand fly). So as far as we know, the chemical and life processes would continue in the same manner if everything were reversed.

If life is entirely a physical and chemical phenomenon, then we can understand that the proteins are all made in the same corkscrew only from the idea that at the very beginning some living molecules, by accident, got started and a few won. Somewhere, once, one organic molecule was lopsided in a certain way, and from this particular thing the "right" happened to evolve in our particular geography; a particular historical accident was one-sided, and ever since then the lopsidedness has propagated itself. Once having arrived at the state that it is in now, of course, it will always continue—all the enzymes digest the right things, manufacture the right things: when the carbon dioxide and the water vapor, and so on, go in the plant leaves, the enzymes that make the sugars make them lopsided because the enzymes are lopsided. If any new kind of virus or living thing were to originate at a later time, it would survive only if it could "eat" the kind of living matter already present. Thus it, too, must be of the same kind.

There is no conservation of the number of right-handed molecules. Once started, we could keep increasing the number of right-handed molecules. So the presumption is, then, that the phenomena in the case of life do not show a lack of symmetry in physical laws, but do show, on the contrary, the universal nature and the commonness of ultimate origin of all creatures on earth, in the sense described above.

2-5 Polar and axial vectors

Now we go further. We observe that in physics there are a lot of other places where we have "right-hand" and "left-hand" rules. As a matter of fact, when we learned about vector analysis we learned about the right-hand rules we have to use in order to get the angular momentum, torque, magnetic field, and so on, to come out right. The force on a charge moving in a magnetic field, for example, is $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$. In a given situation, in which we know $\mathbf{F}$, $\mathbf{v}$, and $\mathbf{B}$, isn't that equation enough to define right-handedness? As a matter of fact, if we go back and look at where the vectors came from, we know that the "right-hand rule" was merely a convention; it was a trick. The original quantities, like the angular momenta and the angular velocities, and things of this kind, were not really vectors at all! They are all somehow associated with a certain plane, and it is just because there are three dimensions in space that we can associate the quantity with a direction perpendicular to that plane. Of the two possible directions, we chose the "right-hand" direction.

So if the laws of physics are symmetrical, we should find that if some demon were to sneak into all the physics laboratories and replace the word "right" for "left" in every book in which "right-hand rules" are given, and instead we were to use all "left-hand rules" uniformly, then it should make no difference whatever in the physical laws.

Let us give an illustration. There are two kinds of vectors. There are "honest" vectors, for example a step $\Delta\mathbf{r}$ in space. If in our apparatus there is a piece here and something else there, then in a

mirror apparatus there will be the image piece and the image something else, and if we draw a vector from the "piece" to the "something else," one vector is the mirror image of the other (Figure 2-2). The vector arrow changes its head, just as the whole space turns inside out; such a vector we call a *polar vector*.

But the other kind of vector, which has to do with rotations, is of a different nature. For example, suppose that in three dimensions something is rotating as shown in Figure 2-3. Then if we look at it in a mirror, it will be rotating as indicated, namely, as the mirror image of the original rotation. Now we have agreed to represent the mirror rotation by the same rule, it is a "vector" which, on reflection, does *not* change about as the polar vector does, but is reversed relative to the polar vectors and to the geometry of the space; such a vector is called an *axial vector*.

Now if the law of reflection symmetry is right in physics, then it must be true that the equations must be so designed that if we change the sign of each axial vector and each cross-product of vectors, which would be what corresponds to reflection, nothing will happen. For instance, when we write a formula which says that the angular momentum is $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, that equation is all right, because if we change to a left-hand coordinate system, we change the sign of $\mathbf{L}$, but $\mathbf{p}$ and $\mathbf{r}$ do not change; the cross-product sign is changed, since we must change from a right-hand rule to a left-hand rule. As another example, we know that the force on a charge moving in a magnetic field is $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, but if we change from a right- to a left-handed system, since $\mathbf{F}$ and $\mathbf{v}$ are known to be polar vectors the sign change required by the cross-product must be cancelled by a sign change in $\mathbf{B}$, which means that $\mathbf{B}$ must be an axial vector.

Figure 2-2 A step in space and its mirror image.

Figure 2-3 A rotating wheel and its mirror image. Note that the angular velocity "vector" is not reversed in direction.

In other words, if we make such a reflection, **B** must go to –**B**. So if we change our coordinates from right to left, we must also change the poles of magnets from north to south.

Let us see how that works in an example. Suppose that we have two magnets, as in Figure 2-4. One is a magnet with the coils going around a certain way, and with current in a given direction. The other magnet looks like the reflection of the first magnet in a mirror—the coil will wind the other way, everything that happens inside the coil is exactly reversed, and the current goes as shown. Now, from the laws for the production of magnetic fields, which we do not know yet officially, but which we most likely learned in high school, it turns out that the magnetic field is as shown in the figure. In one case the pole is a south magnetic pole, while in the other magnet the current is going the other way and the magnetic field is reversed—it is a north magnetic pole. So we see that when we go from right to left we must indeed change from north to south!

Figure 2-4 A magnet and its mirror image.

Never mind changing north to south; these too are mere conventions. Let us talk about *phenomena*. Suppose, now, that we have an electron moving through one field, going into the page. Then, if we use the formula for the force, $\mathbf{v} \times \mathbf{B}$ (remember the charge is minus), we find that the electron will deviate in the indicated direction according to the physical law. So the phenomenon is that we have a coil with a current going in a specified sense and an electron curves in a certain way—that is the physics—never mind how we label everything.

Now let us do the same experiment with a mirror: we send an electron through in a corresponding direction and now the force is reversed, if we calculate it from the same rule, and that is very good because the corresponding *motions* are then mirror images!

2-6 Which hand is right?

So the fact of the matter is that in studying any phenomenon there are always two right-hand rules, or an even number of them, and the net result is that the phenomena always look symmetrical. In short, therefore, we cannot tell right from left if we also are not able to tell north from south. However, it may seem that we *can* tell the north pole of a magnet. The north pole of a compass needle, for example, is one that points to the north. But of course that is again a local property that has to do with geography of the earth; that is just like talking about in which direction is Chicago, so it does not count. If we have seen compass needles, we may have noticed that the north-seeking pole is a sort of bluish color. But that is just due to the man who painted the magnet. These are all local, conventional criteria.

However, if a magnet were to have the property that if we looked at it closely enough we would see small hairs growing on its north pole but not on its south pole, if that were the general rule, or if there were *any* unique way to distinguish the north from the south pole of a magnet, then we could tell which of the two cases we actually had, and *that would be the end of the law of reflection symmetry*.

To illustrate the whole problem still more clearly, imagine that we were talking to a Martian, or someone very far away, by telephone. We are not allowed to send him any actual samples to inspect; for instance, if we could send light, we could send him right-hand circularly polarized light and say, "That is right-hand light—just watch the way it is going." But we cannot *give* him anything, we can only talk to him. He is far away, or in some strange location, and he cannot see anything we can see. For instance, we cannot say, "Look at Ursa major; now see how those stars are arranged. What we mean by 'right' is . . ." We are only allowed to telephone him.

Now we want to tell him all about us. Of course, first we start defining numbers, and say, "Tick, tick, *two*, tick, tick, tick, *three* . . . ," so that gradually he can understand a couple of words, and so on. After a while we may become very familiar with this fellow, and he says, "What do you guys look like?" We start to describe ourselves, and say, "Well, we are six feet tall." He says, "Wait a minute, what is six feet?" Is it possible to tell him what six feet is? Certainly! We say, "You know about the diameter of hydrogen atoms—we are 17,000,000,000 hydrogen atoms high!" That is possible because physical laws are not invariant under change of scale, and therefore we *can* define an absolute length. And so we define the size of the body, and tell him what the general shape is—it has prongs with five bumps sticking out on the ends, and so on, and he follows us along, and we finish describing how we look on the outside, presumably without encountering any particular difficulties. He is even making a model of us as we go along. He says, "My, you are certainly very handsome fellows; now what is on the inside?" So we start to describe the various organs on the inside, and we come to the heart, and we carefully describe the shape of it, and say, "Now put the heart on the left side." He says, "Duhhh—the left side?" Now our problem is to describe to him which side the heart goes on without his ever seeing anything that we see, and without our ever sending any sample to him of what we mean by "right"—no standard right-handed object. Can we do it?

2-7 Parity is not conserved!

It turns out that the laws of gravitation, the laws of electricity and magnetism, nuclear forces, all satisfy the principle of reflection symmetry, so these laws, or anything derived from them, cannot be used. But associated with the many particles that are found in nature there is a phenomenon called *beta decay*, or *weak decay*. One of the examples of weak decay, in connection with a particle discovered in about 1954, posed a strange puzzle. There was a certain charged particle which disintegrated into three π-mesons, as shown schematically in Figure 2-5. This particle was called, for a while, a τ-meson. Now in Figure 2-5 we also see another particle which disintegrates into *two* mesons; one must be neutral, from the conservation of charge. This particle was called a θ-meson. So on the one hand we have a particle called a τ, which disintegrates into three π-mesons, and a θ, which disintegrates into two π-mesons. Now it was soon discovered that the τ and the θ are almost equal in mass; in fact, within the experimental error, they are equal. Next, the length of time it took for them to disintegrate into three π's and two π's was found to be almost exactly the same; they live the same length of time. Next, whenever they were made, they were made in the same proportions, say, 14 percent τ's to 86 percent θ's.

Anyone in his right mind realizes immediately that they must be the same particle, that we merely produce an object which has two different ways of disintegrating—not two different particles. This object that can disintegrate in two different ways has, there-

Figure 2-5 A schematic diagram of the disintegration of a τ^+ and a θ^+ particle.

fore, the same lifetime and the same production ratio (because this is simply the ratio of the odds with which it disintegrates into these two kinds).

However, it was possible to prove (and we cannot here explain at all *how*), from the principle of reflection symmetry in quantum mechanics, that it was *impossible* to have these both come from the same particle—the same particle *could not* disintegrate in both of these ways. The conservation law corresponding to the principle of reflection symmetry is something which has no classical analog, and so this kind of quantum-mechanical conservation was called the *conservation of parity*. So, it was a result of the conservation of parity or, more precisely, from the symmetry of the quantum-mechanical equations of the weak decays under reflection, that the same particle could not go into both, so it must be some kind of coincidence of masses, lifetimes, and so on. But the more it was studied, the more remarkable the coincidence, and the suspicion gradually grew that possibly the deep law of the reflection symmetry of nature may be false.

As a result of this apparent failure, the physicists Lee and Yang suggested that other experiments be done in related decays to try to test whether the law was correct in other cases. The first such experiment was carried out by Miss Wu from Columbia, and was done as follows. Using a very strong magnet at a very low temperature, it turns out that a certain isotope of cobalt, which disintegrates by emitting an electron, is magnetic, and if the temperature is low enough that the thermal oscillations do not jiggle the atomic magnets about too much, they line up in the magnetic field. So the cobalt atoms will all line up in this strong field. They then disintegrate, emitting an electron, and it was discovered that when the atoms were lined up in a field whose **B** vector points upward, most of the electrons were emitted in a downward direction.

If one is not really "hep" to the world, such a remark does not sound like anything of significance, but if one appreciates the problems and interesting things in the world, then he sees that it is a

most dramatic discovery: When we put cobalt atoms in an extremely strong magnetic field, more disintegration electrons go down than up. Therefore if we were to put it in a corresponding experiment in a "mirror," in which the cobalt atoms would be lined up in the opposite direction, they would spit their electrons *up*, not *down*; the action is *unsymmetrical. The magnet has grown hairs!* The south pole of a magnet is of such a kind that the electrons in a β-disintegration tend to go away from it; that distinguishes, in a physical way, the north pole from the south pole.

After this, a lot of other experiments were done: the disintegration of the π into μ and v; μ into an electron and two neutrinos; nowadays, the Λ into proton and π; disintegration of Σ's; and many other disintegrations. In fact, in almost all cases where it could be expected, all have been found *not* to obey reflection symmetry! Fundamentally, the law of reflection symmetry, at this level in physics, is incorrect.

In short, we can tell a Martian where to put the heart: we say, "Listen, build yourself a magnet, and put the coils in, and put the current on, and then take some cobalt and lower the temperature. Arrange the experiment so the electrons go from the foot to the head, then the direction in which the current goes through the coils is the direction that goes in on what we call the right and comes out on the left." So it is possible to define right and left, now, by doing an experiment of this kind.

There are a lot of other features that were predicted. For example, it turns out that the spin, the angular momentum, of the cobalt nucleus before disintegration is 5 units of $\hbar$, and after disintegration it is 4 units. The electron carries spin angular momentum, and there is also a neutrino involved. It is easy to see from this that the electron must carry its spin angular momentum aligned along its direction of motion, the neutrino likewise. So it looks as though the electron is spinning to the left, and that was also checked. In fact, it was checked right here at Caltech by Boehm and Wapstra, that the electrons spin mostly to the left.

(There were some other experiments that gave the opposite answer, but they were wrong!)

The next problem, of course, was to find the law of the failure of parity conservation. What is the rule that tells us how strong the failure is going to be? The rule is this: it occurs only in these very slow reactions, called weak decays, and when it occurs, the rule is that the particles which carry spin, like the electron, neutrino, and so on, come out with a spin tending to the left. That is a lopsided rule; it connects a polar vector velocity and an axial vector angular momentum, and says that the angular momentum is more likely to be opposite to the velocity than along it.

Now that is the rule, but today we do not really understand the whys and wherefores of it. *Why* is this the right rule, what is the fundamental reason for it, and how is it connected to anything else? At the moment we have been so shocked by the fact that this thing is unsymmetrical that we have not been able to recover enough to understand what it means with regard to all the other rules. However, the subject is interesting, modern, and still unsolved, so it seems appropriate that we discuss some of the questions associated with it.

2-8 Antimatter

The first thing to do when one of the symmetries is lost is to immediately go back over the list of known or assumed symmetries and ask whether any of the others are lost. Now we did not mention one operation on our list, which must necessarily be questioned, and that is the relation between matter and antimatter. Dirac predicted that in addition to electrons there must be another particle, called the positron (discovered at Caltech by Anderson), that is necessarily related to the electron. All the properties of these two particles obey certain rules of correspondence: the energies are equal; the masses are equal; the charges are reversed; but, more important than anything, the two of them, when they come together,

can annihilate each other and liberate their entire mass in the form of energy, say γ-rays. The positron is called an *antiparticle* to the electron, and these are the characteristics of a particle and its antiparticle. It was clear from Dirac's argument that all the rest of the particles in the world should also have corresponding antiparticles. For instance, for the proton there should be an antiproton, which is now symbolized by a $\bar{p}$. The $\bar{p}$ would have a negative electrical charge and the same mass as a proton, and so on. The most important feature, however, is that a proton and an antiproton coming together can annihilate each other. The reason we emphasize this is that people do not understand it when we say there is a neutron and also an antineutron, because they say, "A neutron is neutral, so how *can* it have the opposite charge?" The rule of the "anti" is not just that it has the opposite charge, it has a certain set of properties, the whole lot of which are opposite. The antineutron is distinguished from the neutron in this way: if we bring two neutrons together, they just stay as two neutrons, but if we bring a neutron and an antineutron together, they annihilate each other with a great explosion of energy being liberated, with various π-mesons, γ-rays, and whatnot.

Now if we have antineutrons, antiprotons, and antielectrons, we can make antiatoms, in principle. They have not been made yet, but it is possible in principle. For instance, a hydrogen atom has a proton in the center with an electron going around outside. Now imagine that somewhere we can make an antiproton with a positron going around, would it go around? Well, first of all, the antiproton is electrically negative and the antielectron is electrically positive, so they attract each other in a corresponding manner—the masses are all the same; everything is the same. It is one of the principles of the symmetry of physics, the equations seem to show, that if a clock, say, were made of matter on one hand, and then we made the same clock of antimatter, it would run in this way. (Of course, if we put the clocks together, they would annihilate each other, but that is different.)

An immediate question then arises. We can build, out of matter,

two clocks, one which is "left-hand" and one which is "right-hand." For example, we could build a clock which is not built in a simple way, but has cobalt and magnets and electron detectors which detect the presence of β-decay electrons and count them. Each time one is counted, the second hand moves over. Then the mirror clock, receiving fewer electrons, will not run at the same rate. So evidently we can make two clocks such that the left-hand clock does not agree with the right-hand one. Let us make, out of matter, a clock which we call the standard or right-hand clock. Now let us make, also out of matter, a clock which we call the left-hand clock. We have just discovered that, in general, these two will *not* run the same way; prior to that famous physical discovery, it was thought that they would. Now it was also supposed that matter and antimatter were equivalent. That is, if we made an antimatter clock, right-hand, the same shape, then it would run the same as the right-hand matter clock, and if we made the same clock to the left it would run the same. In other words, in the beginning it was believed that *all four* of these clocks were the same; now of course we know that the right-hand and left-hand matter are not the same. Presumably, therefore, the right-handed antimatter and the left-handed antimatter are not the same.

So the obvious question is, which goes with which, if either? In other words, does the right-handed matter behave the same way as the right-handed antimatter? Or does the right-handed matter behave the same way as the left-handed antimatter? β-decay experiments, using positron decay instead of electron decay, indicate that this is the interconnection: matter to the "right" works the same way as antimatter to the "left."

Therefore, at long last, it is really true that right and left symmetry is still maintained! If we made a left-hand clock, but made it out of the other kind of matter, antimatter instead of matter, it would run in the same way. So what has happened is that instead of having two independent rules in our list of symmetries, two of these rules go together to make a new rule, which says that matter to the right is symmetrical with antimatter to the left.

So if our Martian is made of antimatter and we give him instructions to make this "right"-handed model like us, it will, of course, come out the other way around. What would happen when, after much conversation back and forth, we each have taught the other to make spaceships and we meet halfway in empty space? We have instructed each other on our traditions, and so forth, and the two of us come rushing out to shake hands. Well, if he puts out his left hand, watch out!

2-9 Broken symmetries

The next question is, what can we make out of laws which are *nearly* symmetrical? The marvelous thing about it all is that for such a wide range of important, strong phenomena—nuclear forces, electrical phenomena, and even weak ones like gravitation—over a tremendous range of physics, all the laws for these seem to be symmetrical. On the other hand, this little extra piece says, "No, the laws are not symmetrical!" How is it that nature can be almost symmetrical, but not perfectly symmetrical? What shall we make of this? First, do we have any other examples? The answer is, we do, in fact, have a few other examples. For instance, the nuclear part of the force between proton and proton, between neutron and neutron, and between neutron and proton, is all exactly the same—there is a symmetry for nuclear forces, a new one, that we can interchange neutron and proton—but it evidently is not a general symmetry, for the electrical repulsion between two protons at a distance does not exist for neutrons. So it is not generally true that we can *always* replace a proton with a neutron, but only to a good approximation. Why *good*? Because the nuclear forces are much stronger than the electrical forces. So this is an "almost" symmetry also. So we do have examples in other things.

We have, in our minds, a tendency to accept symmetry as some kind of perfection. In fact it is like the old idea of the Greeks that circles were perfect, and it was rather horrible to believe that the planetary orbits were not circles, but only nearly circles. The dif-

ference between being a circle and being nearly a circle is not a small difference, it is a fundamental change so far as the mind is concerned. There is a sign of perfection and symmetry in a circle that is not there the moment the circle is slightly off—that is the end of it—it is no longer symmetrical. Then the question is why it is only *nearly* a circle—that is a much more difficult question. The actual motion of the planets, in general, should be ellipses, but during the ages, because of tidal forces, and so on, they have been made almost symmetrical. Now the question is whether we have a similar problem here. The problem from the point of view of the circles is if they were perfect circles there would be nothing to explain, that is clearly simple. But since they are only nearly circles, there is a lot to explain, and the result turned out to be a big dynamical problem, and now our problem is to explain why they are nearly symmetrical by looking at tidal forces and so on.

So our problem is to explain where symmetry comes from. Why is nature so nearly symmetrical? No one has any idea why. The only thing we might suggest is something like this: There is a gate in Japan, a gate in Neiko, which is sometimes called by the Japanese the most beautiful gate in all Japan; it was built in a time when there was great influence from Chinese art. This gate is very elaborate, with lots of gables and beautiful carving and lots of columns and dragon heads and princes carved into the pillars, and so on. But when one looks closely he sees that in the elaborate and complex design along one of the pillars, one of the small design elements is carved upside down; otherwise the thing is completely symmetrical. If one asks why this is, the story is that it was carved upside down so that the gods will not be jealous of the perfection of man. So they purposely put an error in there, so that the gods would not be jealous and get angry with human beings.

We might like to turn the idea around and think that the true explanation of the near symmetry of nature is this: that God made the laws only nearly symmetrical so that we should not be jealous of His perfection!

CALTECH

3

THE SPECIAL THEORY OF RELATIVITY

3-1 The principle of relativity

For over 200 years the equations of motion enunciated by Newton were believed to describe nature correctly, and the first time that an error in these laws was discovered, the way to correct it was also discovered. Both the error and its correction were discovered by Einstein in 1905.

Newton's Second Law, which we have expressed by the equation

$$F = d(mv)/dt,$$

was stated with the tacit assumption that m is a constant, but we now know that this is not true, and that the mass of a body increases with velocity. In Einstein's corrected formula m has the value

$$m = \frac{m_0}{\sqrt{1 - v^2/c^2}}, \tag{3.1}$$

where the "rest mass" m_0 represents the mass of a body that is not moving and c is the speed of light, which is about 3×10^5 km · sec^{-1} or about 186,000 mi · sec^{-1}.

For those who want to learn just enough about it so they can solve problems, that is all there is to the theory of relativity—it just changes Newton's laws by introducing a correction factor to the

mass. From the formula itself it is easy to see that this mass increase is very small in ordinary circumstances. If the velocity is even as great as that of a satellite, which goes around the earth at 5 mi/sec, then $v/c = 5/186{,}000$: putting this value into the formula shows that the correction to the mass is only one part in two to three billion, which is nearly impossible to observe. Actually, the correctness of the formula has been amply confirmed by the observation of many kinds of particles, moving at speeds ranging up to practically the speed of light. However, because the effect is ordinarily so small, it seems remarkable that it was discovered theoretically before it was discovered experimentally. Empirically, at a sufficiently high velocity, the effect is very large, but it was not discovered that way. Therefore it is interesting to see how a law that involved so delicate a modification (at the time when it was first discovered) was brought to light by a combination of experiments and physical reasoning. Contributions to the discovery were made by a number of people, the final result of whose work was Einstein's discovery.

There are really two Einstein theories of relativity. This chapter is concerned with the Special Theory of Relativity, which dates from 1905. In 1915 Einstein published an additional theory, called the General Theory of Relativity. This latter theory deals with the extension of the Special Theory to the case of the law of gravitation; we shall not discuss the General Theory here.

The principle of relativity was first stated by Newton, in one of his corollaries to the laws of motion: "The motions of bodies included in a given space are the same among themselves, whether that space is at rest or moves uniformly forward in a straight line." This means, for example, that if a spaceship is drifting along at a uniform speed, all experiments performed in the spaceship and all the phenomena in the spaceship will appear the same as if the ship were not moving, provided, of course, that one does not look outside. That is the meaning of the principle of relativity. This is a simple enough idea, and the only question is whether it is *true* that in all experiments performed inside a moving system the laws of physics will appear the same as they would if the system were stand-

ing still. Let us first investigate whether Newton's laws appear the same in the moving system.

Suppose that Moe is moving in the x-direction with a uniform velocity u, and he measures the position of a certain point, shown in Figure 3-1. He designates the "x-distance" of the point in his coordinate system as x'. Joe is at rest, and measures the position of the same point, designating its x-coordinate in his system as x. The relationship of the coordinates in the two systems is clear from the diagram. After time t Moe's origin has moved a distance ut, and if the two systems originally coincided,

$$\begin{aligned} x' &= x - ut, \\ y' &= y, \\ z' &= z, \\ t' &= t. \end{aligned} \tag{3.2}$$

If we substitute this transformation of coordinates into Newton's laws we find that these laws transform to the same laws in the primed system; that is, the laws of Newton are of the same form in a moving system as in a stationary system, and therefore it is impossible to tell, by making mechanical experiments, whether the system is moving or not.

The principle of relativity has been used in mechanics for a long time. It was employed by various people, in particular Huygens, to obtain the rules for the collision of billiard balls, in much the same

Figure 3-1 Two coordinate systems in uniform relative motion along their x-axes.

way as we used it in Chapter 10* to discuss the conservation of momentum. In the past century interest in it was heightened as the result of investigations into the phenomena of electricity, magnetism, and light. A long series of careful studies of these phenomena by many people culminated in Maxwell's equations of the electromagnetic field, which describe electricity, magnetism, and light in one uniform system. However, the Maxwell equations did *not* seem to obey the principle of relativity. That is, if we transform Maxwell's equations by the substitution of equations 3.2, *their form does not remain the same*; therefore, in a moving spaceship the electrical and optical phenomena should be different from those in a stationary ship. Thus one could use these optical phenomena to determine the speed of the ship; in particular, one could determine the absolute speed of the ship by making suitable optical or electrical measurements. One of the consequences of Maxwell's equations is that if there is a disturbance in the field such that light is generated, these electromagnetic waves go out in all directions equally and at the same speed c, or 186,000 mi/sec. Another consequence of the equations is that if the source of the disturbance is moving, the light emitted goes through space at the same speed c. This is analogous to the case of sound, the speed of sound waves being likewise independent of the motion of the source.

This independence of the motion of the source, in the case of light, brings up an interesting problem:

Suppose we are riding in a car that is going at a speed u, and light from the rear is going past the car with speed c. Differentiating the first equation in (3.2) gives

$$dx'/dt = dx/dt - u,$$

which means that according to the Galilean transformation the apparent speed of the passing light, as we measure it in the car, should not be c but should be $c - u$. For instance, if the car is going

* of *The Feynman Lectures on Physics,* vol. I.

100,000 mi/sec, and the light is going 186,000 mi/sec, then apparently the light going past the car should go 86,000 mi/sec. In any case, by measuring the speed of the light going past the car (if the Galilean transformation is correct for light), one could determine the speed of the car. A number of experiments based on this general idea were performed to determine the velocity of the earth, but they all failed—they gave *no velocity at all*. We shall discuss one of these experiments in detail, to show exactly what was done and what was the matter; something *was* the matter, of course, something was wrong with the equations of physics. What could it be?

3-2 The Lorentz transformation

When the failure of the equations of physics in the above case came to light, the first thought that occurred was that the trouble must lie in the new Maxwell equations of electrodynamics, which were only 20 years old at the time. It seemed almost obvious that these equations must be wrong, so the thing to do was to change them in such a way that under the Galilean transformation the principle of relativity would be satisfied. When this was tried, the new terms that had to be put into the equations led to predictions of new electrical phenomena that did not exist at all when tested experimentally, so this attempt had to be abandoned. Then it gradually became apparent that Maxwell's laws of electrodynamics were correct, and the trouble must be sought elsewhere.

In the meantime, H. A. Lorentz noticed a remarkable and curious thing when he made the following substitutions in the Maxwell equations:

$$\begin{aligned} x' &= \frac{x - ut}{\sqrt{1 - u^2/c^2}}, \\ y' &= y, \\ z' &= z, \\ t' &= \frac{t - ux/c^2}{\sqrt{1 - u^2/c^2}}, \end{aligned} \tag{3.3}$$

namely, Maxwell's equations remain in the same form when this transformation is applied to them! Equations (3.3) are known as a *Lorentz transformation*. Einstein, following a suggestion originally made by Poincaré, then proposed that *all the physical laws* should be of such a kind that they *remain unchanged under a Lorentz transformation*. In other words, we should change, not the laws of electrodynamics, but the laws of mechanics. How shall we change Newton's laws so that *they* will remain unchanged by the Lorentz transformation? If this goal is set, we then have to rewrite Newton's equations in such a way that the conditions we have imposed are satisfied. As it turned out, the only requirement is that the mass m in Newton's equations must be replaced by the form shown in Eq. (3.1). When this change is made, Newton's laws and the laws of electrodynamics will harmonize. Then if we use the Lorentz transformation in comparing Moe's measurements with Joe's, we shall never be able to detect whether either is moving, because the form of all the equations will be the same in both coordinate systems!

It is interesting to discuss what it means that we replace the old transformation between the coordinates and time with a new one, because the old one (Galilean) seems to be self-evident, and the new one (Lorentz) looks peculiar. We wish to know whether it is logically and experimentally possible that the new, and not the old, transformation can be correct. To find that out, it is not enough to study the laws of mechanics but, as Einstein did, we too must analyze our ideas of *space* and *time* in order to understand this transformation. We shall have to discuss these ideas and their implications for mechanics at some length, so we say in advance that the effort will be justified, since the results agree with experiment.

3-3 The Michelson-Morley experiment

As mentioned above, attempts were made to determine the absolute velocity of the earth through the hypothetical "ether" that was supposed to pervade all space. The most famous of these experiments

is one performed by Michelson and Morley in 1887. It was 18 years later before the negative results of the experiment were finally explained, by Einstein.

The Michelson-Morley experiment was performed with an apparatus like that shown schematically in Figure 3-2. This apparatus is essentially comprised of a light source *A*, a partially silvered glass plate *B*, and two mirrors *C* and *E*, all mounted on a rigid base. The mirrors are placed at equal distances *L* from *B*. The plate *B* splits an oncoming beam of light, and the two resulting beams continue in mutually perpendicular directions to the mirrors, where they are reflected back to *B*. On arriving back at *B*, the two beams are recombined as two superposed beams, *D* and *F*. If the time taken for the light to go from *B* to *E* and back is the same as the time from *B* to *C* and back, the emerging beams *D* and *F* will be in phase and will reinforce each other, but if the two times differ slightly, the beams will be slightly out of phase and interference will result. If the apparatus is "at rest" in the ether, the times should be precisely equal, but if it is moving toward the right with a velocity *u*, there should be a difference in the times. Let us see why.

Figure 3-2 Schematic diagram of the Michelson-Morley experiment.

First, let us calculate the time required for the light to go from B to E and back. Let us say that the time for light to go from plate B to mirror E is t_1, and the time for the return is t_2. Now, while the light is on its way from B to the mirror, the apparatus moves a distance ut_1, so the light must traverse a distance $L + ut_1$, at the speed c. We can also express this distance as ct_1, so we have

$$ct_1 = L + ut_1, \quad \text{or} \quad t_1 = L/(c - u).$$

(This result is also obvious from the point of view that the velocity of light relative to the apparatus is $c - u$, so the time is the length L divided by $c - u$.) In a like manner, the time t_2 can be calculated. During this time the plate B advances a distance ut_2, so the return distance of the light is $L - ut_2$. Then we have

$$ct_2 = L - ut_2, \quad \text{or} \quad t_2 = L/(c + u).$$

Then the total time is

$$t_1 + t_2 = 2Lc/(c^2 - u^2).$$

For convenience in later comparison of times we write this as

$$t_1 + t_2 = \frac{2L/c}{1 - u^2/c^2}. \tag{3.4}$$

Our second calculation will be of the time t_3 for the light to go from B to the mirror C. As before, during time t_3 the mirror C moves to the right a distance ut_3 to the position C'; in the same time, the light travels a distance ct_3 along the hypotenuse of a triangle, which is BC'. For this right triangle we have

$$(ct_3)^2 = L^2 + (ut_3)^2$$

or

$$L^2 = c^2 t_3^2 - u^2 t_3^2 = (c^2 - u^2)\, t_3^2,$$

from which we get

$$t_3 = L/\sqrt{c^2 - u^2}\,.$$

For the return trip from C' the distance is the same, as can be seen from the symmetry of the figure; therefore the return time is also the same, and the total time is $2t_3$. With a little rearrangement of the form we can write

$$2t_3 = \frac{2L}{\sqrt{c^2 - u^2}} = \frac{2L/c}{\sqrt{1 - u^2/c^2}}. \tag{3.5}$$

We are now able to compare the times taken by the two beams of light. In expressions (3.4) and (3.5) the numerators are identical, and represent the time that would be taken if the apparatus were at rest. In the denominators, the term u^2/c^2 will be small, unless u is comparable in size to c. The denominators represent the modifications in the times caused by the motion of the apparatus. And behold, these modifications are *not the same*—the time to go to C and back is a little less than the time to E and back, even though the mirrors are equidistant from B, and all we have to do is to measure that difference with precision.

Here a minor technical point arises—suppose the two lengths L are not exactly equal? In fact, we surely cannot make them exactly equal. In that case we simply turn the apparatus 90 degrees, so that BC is in the line of motion and BE is perpendicular to the motion. Any small difference in length then becomes unimportant, and what we look for is a *shift* in the interference fringes when we rotate the apparatus.

In carrying out the experiment, Michelson and Morley oriented the apparatus so that the line BE was nearly parallel to the earth's motion in its orbit (at certain times of the day and night). This orbital speed is about 18 miles per second, and any "ether drift" should be at least that much at some time of the day or night and at some time during the year. The apparatus was amply sensitive to observe such an effect, but no time difference was found—the velocity of the earth through the ether could not be detected. The result of the experiment was null.

The result of the Michelson-Morley experiment was very puzzling

and most disturbing. The first fruitful idea for finding a way out of the impasse came from Lorentz. He suggested that material bodies contract when they are moving, and that this foreshortening is only in the direction of the motion, and also, that if the length is L_0 when a body is at rest, then when it moves with speed u parallel to its length, the new length, which we call $L_{\|}$ (L-parallel), is given by

$$L_{\|} = L_0 \sqrt{1 - u^2/c^2}\,. \tag{3.6}$$

When this modification is applied to the Michelson-Morley interferometer apparatus the distance from B to C does not change, but the distance from B to E is shortened to $L\sqrt{1 - u^2/c^2}$. Therefore Eq. (3.5) is not changed, but the L of Eq. (3.4) must be changed in accordance with Eq. (3.6). When this is done we obtain

$$t_1 + t_2 = \frac{(2L/c)\sqrt{1 - u^2/c^2}}{1 - u^2/c^2} = \frac{2L/c}{\sqrt{1 - u^2/c^2}}. \tag{3.7}$$

Comparing this result with Eq. (3.5), we see that $t_1 + t_2 = 2t_3$. So if the apparatus shrinks in the manner just described, we have a way of understanding why the Michelson-Morley experiment gives no effect at all. Although the contraction hypothesis successfully accounted for the negative result of the experiment, it was open to the objection that it was invented for the express purpose of explaining away the difficulty, and was too artificial. However, in many other experiments to discover an ether wind, similar difficulties arose, until it appeared that nature was in a "conspiracy" to thwart man by introducing some new phenomenon to undo every phenomenon that he thought would permit a measurement of u.

It was ultimately recognized, as Poincaré pointed out, that *a complete conspiracy is itself a law of nature*! Poincaré then proposed that there *is* such a law of nature, that it is not possible to discover an ether wind by *any* experiment; that is, there is no way to determine an absolute velocity.

3-4 Transformation of time

In checking out whether the contraction idea is in harmony with the facts in other experiments, it turns out that everything is correct provided that the *times* are also modified, in the manner expressed in the fourth equation of the set (3.3). That is because the time t_3, calculated for the trip from B to C and back, is not the same when calculated by a man performing the experiment in a moving spaceship as when calculated by a stationary observer who is watching the spaceship. To the man in the ship the time is simply $2L/c$, but to the other observer it is $(2L/c)/\sqrt{1-u^2/c^2}$ (Eq. 3.5). In other words, when the outsider sees the man in the spaceship lighting a cigar, all the actions appear to be slower than normal, while to the man inside, everything moves at a normal rate. So not only must the lengths shorten, but also the time-measuring instruments ("clocks") must apparently slow down. That is, when the clock in the spaceship records 1 second elapsed, as seen by the man in the ship, it shows $1/\sqrt{1-u^2/c^2}$ second to the man outside.

This slowing of the clocks in a moving system is a very peculiar phenomenon, and is worth an explanation. In order to understand this, we have to watch the machinery of the clock and see what happens when it is moving. Since that is rather difficult, we shall take a very simple kind of clock. The one we choose is rather a silly kind of clock, but it will work in principle: it is a rod (meter stick) with a mirror at each end, and when we start a light signal between the mirrors, the light keeps going up and down, making a click every time it comes down, like a standard ticking clock. We build two such clocks, with exactly the same lengths, and synchronize them by starting them together; then they agree always thereafter, because they are the same in length, and light always travels with speed c. We give one of these clocks to the man to take along in his spaceship, and he mounts the rod perpendicular to the direction of motion of the ship; then the length of the rod will not change. How do we know that perpendicular lengths do not change? The

men can agree to make marks on each other's y-meter stick as they pass each other. By symmetry, the two marks must come at the same y- and y'-coordinates, since otherwise, when they get together to compare results, one mark will be above or below the other, and so we could tell who was really moving.

Now let us see what happens to the moving clock. Before the man took it aboard, he agreed that it was a nice, standard clock, and when he goes along in the spaceship he will not see anything peculiar. If he did, he would know he was moving—if anything at all changed because of the motion, he could tell he was moving. But the principle of relativity says this is impossible in a uniformly moving system, so nothing has changed. On the other hand, when the external observer looks at the clock going by, he sees that the light, in going from mirror to mirror, is "really" taking a zigzag path, since the rod is moving sideways all the while. We have already analyzed such a zigzag motion in connection with the Michelson-Morley experiment. If in a given time the rod moves forward a distance proportional to u in Figure 3-3, the distance the

Figure 3-3 (a) A "light clock" at rest in the S' system. (b) The same clock, moving through the S system. (c) Illustration of the diagonal path taken by the light beam in a moving "light clock."

light travels in the same time is proportional to c, and the vertical distance is therefore proportional to $\sqrt{c^2 - u^2}$.

That is, it takes a *longer time* for light to go from end to end in the moving clock than in the stationary clock. Therefore the apparent time between clicks is longer for the moving clock, in the same proportion as shown in the hypotenuse of the triangle (that is the source of the square root expressions in our equations). From the figure it is also apparent that the greater u is, the more slowly the moving clock appears to run. Not only does this particular kind of clock run more slowly, but if the theory of relativity is correct, any other clock, operating on any principle whatsoever, would also appear to run slower, and in the same proportion—we can say this without further analysis. Why is this so?

To answer the above question, suppose we had two other clocks made exactly alike with wheels and gears, or perhaps based on radioactive decay, or something else. Then we adjust these clocks so they both run in precise synchronism with our first clocks. When light goes up and back in the first clocks and announces its arrival with a click, the new models also complete some sort of cycle, which they simultaneously announce by some doubly coincident flash, or bong, or other signal. One of these clocks is taken into the spaceship, along with the first kind. Perhaps *this* clock will not run slower, but will continue to keep the same time as its stationary counterpart, and thus disagree with the other moving clock. Ah no, if that should happen, the man in the ship could use this mismatch between his two clocks to determine the speed of his ship, which we have been supposing is impossible. *We need not know anything about the machinery* of the new clock that might cause the effect—we simply know that whatever the reason, it will appear to run slow, just like the first one.

Now if *all* moving clocks run slower, if no way of measuring time gives anything but a slower rate, we shall just have to say, in a certain sense, that *time itself* appears to be slower in a spaceship. All the phenomena there—the man's pulse rate, his thought processes, the time

he takes to light a cigar, how long it takes to grow up and get old—all these things must be slowed down in the same proportion, because he cannot tell he is moving. The biologists and medical men sometimes say it is not quite certain that the time it takes for a cancer to develop will be longer in a spaceship, but from the viewpoint of a modern physicist it is nearly certain; otherwise one could use the rate of cancer development to determine the speed of the ship!

A very interesting example of the slowing of time with motion is furnished by mu-mesons (muons), which are particles that disintegrate spontaneously after an average lifetime of 2.2×10^{-6} sec. They come to the earth in cosmic rays, and can also be produced artificially in the laboratory. Some of them disintegrate in midair, but the remainder disintegrate only after they encounter a piece of material and stop. It is clear that in its short lifetime a muon cannot travel, even at the speed of light, much more than 600 meters. But although the muons are created at the top of the atmosphere, some 10 kilometers up, yet they are actually found in a laboratory down here, in cosmic rays. How can that be? The answer is that different muons move at various speeds, some of which are very close to the speed of light. While from their own point of view they live only about 2 μsec, from our point of view they live considerably longer—enough longer that they may reach the earth. The factor bv which the time is increased has already been given as $1/\sqrt{1 - u^2/c^2}$. The average life has been measured quite accurately for muons of different velocities, and the values agree closely with the formula.

We do not know why the meson disintegrates or what its machinery is, but we do know its behavior satisfies the principle of relativity. That is the utility of the principle of relativity—it permits us to make predictions, even about things that otherwise we do not know much about. For example, before we have any idea at all about what makes the meson disintegrate, we can still predict that when it is moving at nine-tenths of the speed of light, the apparent length of time that it lasts is $(2.2 \times 10^{-6})/\sqrt{1 - 9^2/10^2}$ sec; and our prediction works—that is the good thing about it.

3-5 The Lorentz contraction

Now let us return to the Lorentz transformation (3.3) and try to get a better understanding of the relationship between the (x, y, z, t) and the (x', y', z', t') coordinate systems, which we shall call the S and S' systems, or Joe and Moe systems, respectively. We have already noted that the first equation is based on the Lorentz suggestion of contraction along the x-direction; how can we prove that a contraction takes place? In the Michelson-Morley experiment, we now appreciate that the *transverse* arm BC cannot change length, by the principle of relativity; yet the null result of the experiment demands that the *times* must be equal. So, in order for the experiment to give a null result, the longitudinal arm BE must appear shorter, by the square root $\sqrt{1 - u^2/c^2}$. What does this contraction mean, in terms of measurements made by Joe and Moe? Suppose that Moe, moving with the S' system in the x-direction, is measuring the x'-coordinate of some point with a meter stick. He lays the stick down x' times, so he thinks the distance is x' meters. From the viewpoint of Joe in the S system, however, Moe is using a foreshortened ruler, so the "real" distance measured is $x'\sqrt{1 - u^2/c^2}$ meters. Then if the S' system has travelled a distance ut away from the S system, the S observer would say that the same point, measured in his coordinates, is at a distance $x = x'\sqrt{1 - u^2/c^2} + ut$, or

$$x' = \frac{x - ut}{\sqrt{1 - u^2/c^2}},$$

which is the first equation of the Lorentz transformation.

3-6 Simultaneity

In an analogous way, because of the difference in time scales, the denominator expression is introduced into the fourth equation of the Lorentz transformation. The most interesting term in that equation is the ux/c^2 in the numerator, because that is quite new and unexpected. Now what does that mean? If we look at the situation

carefully we see that events that occur at two separate places at the same time, as seen by Moe in S', do *not* happen at the same time as viewed by Joe in S. If one event occurs at point x_1 at time t_0 and the other event at x_2 and t_0 (the same time), we find that the two corresponding times t'_1 and t'_2 differ by an amount

$$t'_2 - t'_1 = \frac{u(x_1 - x_2)/c^2}{\sqrt{1 - u^2/c^2}}.$$

This circumstance is called "failure of simultaneity at a distance," and to make the idea a little clearer let us consider the following experiment.

Suppose that a man moving in a spaceship (system S') has placed a clock at each end of the ship and is interested in making sure that the two clocks are in synchronism. How can the clocks be synchronized? There are many ways. One way, involving very little calculation, would be first to locate exactly the midpoint between the clocks. Then from this station we send out a light signal which will go both ways at the same speed and will arrive at both clocks, clearly, at the same time. This simultaneous arrival of the signals can be used to synchronize the clocks. Let us then suppose that the man in S' synchronizes his clocks by this particular method. Let us see whether an observer in system S would agree that the two clocks are synchronous. The man in S' has a right to believe they are, because he does not know that he is moving. But the man in S reasons that since the ship is moving forward, the clock in the front end was running away from the light signal, hence the light had to go more than halfway in order to catch up; the rear clock, however, was advancing to meet the light signal, so this distance was shorter. Therefore the signal reached the rear clock first, although the man in S' thought that the signals arrived simultaneously. We thus see that when a man in a spaceship thinks the times at two locations are simultaneous, equal values of t' in his coordinate system must correspond to *different* values of t in the other coordinate system!

3-7 Four-vectors

Let us see what else we can discover in the Lorentz transformation. It is interesting to note that the transformation between the x's and t's is analogous in form to the transformation of the x's and y's that we studied in Chapter 1 for a rotation of coordinates. We then had

$$\begin{aligned} x' &= x \cos\theta + y \sin\theta, \\ y' &= y \cos\theta - x \sin\theta, \end{aligned} \tag{3.8}$$

in which the new x' mixes the old x and y, and the new y' also mixes the old x and y; similarly, in the Lorentz transformation we find a new x' which is a mixture of x and t, and a new t' which is a mixture of t and x. So the Lorentz transformation is analogous to a rotation, only it is a "rotation" in *space and time*, which appears to be a strange concept. A check of the analogy to rotation can be made by calculating the quantity

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = x^2 + y^2 + z^2 - c^2t^2. \tag{3.9}$$

In this equation the first three terms on each side represent, in three-dimensional geometry, the square of the distance between a point and the origin (surface of a sphere) which remains unchanged (invariant) regardless of rotation of the coordinate axes. Similarly, Eq. (3.9) shows that there is a certain combination which includes time, that is invariant to a Lorentz transformation. Thus, the analogy to a rotation is complete, and is of such a kind that vectors, i.e., quantities involving "components" which transform the same way as the coordinates and time, are also useful in connection with relativity.

Thus we contemplate an extension of the idea of vectors, which we have so far considered to have only space components, to include a time component. That is, we expect that there will be vectors with four components, three of which are like the components

of an ordinary vector, and with these will be associated a fourth component, which is the analog of the time part.

This concept will be analyzed further in the next chapters, where we shall find that if the ideas of the preceding paragraph are applied to momentum, the transformation gives three space parts that are like ordinary momentum components, and a fourth component, the time part, which is the *energy*.

3-8 Relativistic dynamics

We are now ready to investigate, more generally, what form the laws of mechanics take under the Lorentz transformation. [We have thus far explained how length and time change, but not how we get the modified formula for m (Eq. 3.1). We shall do this in the next chapter.] To see the consequences of Einstein's modification of m for Newtonian mechanics, we start with the Newtonian law that force is the rate of change of momentum, or

$$\mathbf{F} = d(m\mathbf{v})/dt.$$

Momentum is still given by $m\mathbf{v}$, but when we use the new m this becomes

$$\mathbf{p} = m\mathbf{v} = \frac{m_0\mathbf{v}}{\sqrt{1 - v^2/c^2}} \tag{3.10}$$

This is Einstein's modification of Newton's laws. Under this modification, if action and reaction are still equal (which they may not be in detail, but are in the long run), there will be conservation of momentum in the same way as before, but the quantity that is being conserved is not the old $m\mathbf{v}$ with its constant mass, but instead is the quantity shown in (3.10), which has the modified mass. When this change is made in the formula for momentum, conservation of momentum still works.

Now let us see how momentum varies with speed. In Newtonian mechanics it is proportional to the speed and, according to (3.10),

over a considerable range of speed, but small compared with c, it is nearly the same in relativistic mechanics, because the square-root expression differs only slightly from 1. But when v is almost equal to c, the square-root expression approaches zero, and the momentum therefore goes toward infinity.

What happens if a constant force acts on a body for a long time? In Newtonian mechanics the body keeps picking up speed until it goes faster than light. But this is impossible in relativistic mechanics. In relativity, the body keeps picking up, not speed, but momentum, which can continually increase because the mass is increasing. After a while there is practically no acceleration in the sense of a change of velocity, but the momentum continues to increase. Of course, whenever a force produces very little change in the velocity of a body, we say that the body has a great deal of inertia, and that is exactly what our formula for relativistic mass says (see Eq. 3.10)—it says that the inertia is very great when v is nearly as great as c. As an example of this effect, to deflect the high-speed electrons in the synchrotron that is used here at Caltech, we need a magnetic field that is 2000 times stronger than would be expected on the basis of Newton's laws. In other words, the mass of the electrons in the synchrotron is 2000 times as great as their normal mass, and is as great as that of a proton! That m should be 2000 times m_0 means that $1 - v^2/c^2$ must be 1/4,000,000, and that means that v^2/c^2 differs from 1 by one part in 4,000,000, or that v differs from c by one part in 8,000,000, so the electrons are getting pretty close to the speed of light. If the electrons and light were both to start from the synchrotron (estimated as 700 feet away) and rush out to Bridge Lab, which would arrive first? The light, of course, because light always travels faster.* How much earlier? That is too hard to tell—instead, we tell by what distance the light is ahead: it is about 1/1000 of an inch, or ¼ the thickness of a piece of paper! When

* The electrons would actually win the race versus *visible* light because of the index of refraction of air. A gamma ray would make out better.

the electrons are going that fast their masses are enormous, but their speed cannot exceed the speed of light.

Now let us look at some further consequences of relativistic change of mass. Consider the motion of the molecules in a small tank of gas. When the gas is heated, the speed of the molecules is increased, and therefore the mass is also increased and the gas is heavier. An approximate formula to express the increase of mass, for the case when the velocity is small, can be found by expanding $m_0/\sqrt{1-v^2/c^2} = m_0(1-v^2/c^2)^{-\frac{1}{2}}$ in a power series, using the binomial theorem. We get

$$m_0(1-v^2/c^2)^{-\frac{1}{2}} = m_0(1+\tfrac{1}{2}v^2/c^2+\tfrac{3}{8}v^4/c^4+\ldots).$$

We see clearly from the formula that the series converges rapidly when v is small, and the terms after the first two or three are negligible. So we can write

$$m \cong m_0 + \tfrac{1}{2}m_0 v^2\left(\frac{1}{c^2}\right) \tag{3.11}$$

in which the second term on the right expresses the increase of mass due to molecular velocity. When the temperature increases the v^2 increases proportionately, so we can say that the increase in mass is proportional to the increase in temperature. But since $\frac{1}{2}m_0v^2$ is the kinetic energy in the old-fashioned Newtonian sense, we can also say that the increase in mass of all this body of gas is equal to the increase in kinetic energy divided by c^2, or $\Delta m = \Delta(\text{K.E.})/c^2$.

3-9 Equivalence of mass and energy

The above observation led Einstein to the suggestion that the mass of a body can be expressed more simply than by the formula (3.1), if we say that the mass is equal to the total energy content divided by c^2. If Eq. (3.11) is multiplied by c^2 the result is

$$mc^2 = m_0c^2 + \tfrac{1}{2}m_0v^2 + \ldots. \tag{3.12}$$

Here, the term on the left expresses the total energy of a body, and we recognize the last term as the ordinary kinetic energy. Einstein

interpreted the large constant term, m_0c^2, to be part of the total energy of the body, an intrinsic energy known as the "rest energy."

Let us follow out the consequences of assuming, with Einstein, that *the energy of a body always equals* mc^2. As an interesting result, we shall find the formula (3.1) for the variation of mass with speed, which we have merely assumed up to now. We start with the body at rest, when its energy is m_0c^2. Then we apply a force to the body, which starts it moving and gives it kinetic energy; therefore, since the energy has increased, the mass has increased—this is implicit in the original assumption. So long as the force continues, the energy and the mass both continue to increase. We have already seen (Chapter 13*) that the rate of change of energy with time equals the force times the velocity, or

$$\frac{dE}{dt} = \mathbf{F} \cdot \mathbf{v}. \tag{3.13}$$

We also have (Chapter 9*, Eq. 9.1) that $F = d(mv)/dt$. When these relations are put together with the definition of E, Eq. (3.13) becomes

$$\frac{d(mc^2)}{dt} = \mathbf{v} \cdot \frac{d(m\mathbf{v})}{dt}. \tag{3.14}$$

We wish to solve this equation for m. To do this we first use the mathematical trick of multiplying both sides by $2m$, which changes the equation to

$$c^2(2m)\frac{dm}{dt} = 2m\mathbf{v} \cdot \frac{d(m\mathbf{v})}{dt}. \tag{3.15}$$

We need to get rid of the derivatives, which can be accomplished by integrating both sides. The quantity $(2m)dm/dt$ can be recognized

* of *The Feynman Lectures on Physics*, vol. I.

as the time derivative of m^2, and $(2m\mathbf{v}) \cdot d(m\mathbf{v})/dt$ is the time derivative of $(mv)^2$. So, Eq. (3.15) is the same as

$$c^2 \frac{d(m^2)}{dt} = \frac{d(m^2 v^2)}{dt}. \tag{3.16}$$

If the derivatives of two quantities are equal, the quantities themselves differ at most by a constant, say C. This permits us to write

$$m^2c^2 = m^2v^2 + C. \tag{3.17}$$

We need to define the constant C more explicitly. Since Eq. (3.17) must be true for all velocities, we can choose a special case where $v = 0$, and say that in this case the mass is m_0. Substituting these values into Eq. (3.17) gives

$$m_0^2 c^2 = 0 + C.$$

We can now use this value of C in Eq. (3.17), which becomes

$$m^2 c^2 = m^2 v^2 + m_0^2 c^2. \tag{3.18}$$

Dividing by c^2 and rearranging terms gives

$$m^2(1 - v^2/c^2) = m_0^2,$$

from which we get

$$m = m_0/\sqrt{1 - v^2/c^2}. \tag{3.19}$$

This is the formula (3.1), and is exactly what is necessary for the agreement between mass and energy in Eq. (3.12).

Ordinarily these energy changes represent extremely slight changes in mass, because most of the time we cannot generate much energy from a given amount of material; but in an atomic bomb of explosive energy equivalent to 20 kilotons of TNT, for example, it can be shown that the dirt after the explosion is lighter by 1 gram than the initial mass of the reacting material, because of the energy that was released, i.e., the released energy had a mass of 1 gram, according to the relationship $\Delta E = \Delta(mc^2)$. This theory of equivalence of mass and energy has been beautifully verified by ex-

periments in which matter is annihilated—converted totally to energy: An electron and a positron come together at rest, each with a rest mass m_0. When they come together they disintegrate and two gamma rays emerge, each with the measured energy of m_0c^2. This experiment furnishes a direct determination of the energy associated with the existence of the rest mass of a particle.

4

RELATIVISTIC ENERGY AND MOMENTUM

4-1 Relativity and the philosophers

In this chapter we shall continue to discuss the principle of relativity of Einstein and Poincaré, as it affects our ideas of physics and other branches of human thought.

Poincaré made the following statement of the principle of relativity: "According to the principle of relativity, the laws of physical phenomena must be the same for a fixed observer as for an observer who has a uniform motion of translation relative to him, so that we have not, nor can we possibly have, any means of discerning whether or not we are carried along in such a motion."

When this idea descended upon the world, it caused a great stir among philosophers, particularly the "cocktail-party philosophers," who say, "Oh, it is very simple: Einstein's theory says all is relative!" In fact, a surprisingly large number of philosophers, not only those found at cocktail parties (but rather than embarrass them, we shall just call them "cocktail-party philosophers"), will say, "That all is relative is a consequence of Einstein, and it has profound influences on our ideas." In addition, they say "It has been demonstrated in physics that phenomena depend upon your frame of reference." We hear that a great deal, but it is difficult to find out what it means.

Probably the frames of reference that were originally referred to were the coordinate systems which we use in the analysis of the theory of relativity. So the fact that "things depend upon your frame of reference" is supposed to have had a profound effect on modern thought. One might well wonder why, because, after all, that things depend upon one's point of view is so simple an idea that it certainly cannot have been necessary to go to all the trouble of the physical relativity theory in order to discover it. That what one sees depends upon his frame of reference is certainly known to anybody who walks around, because he sees an approaching pedestrian first from the front and then from the back; there is nothing deeper in most of the philosophy which is said to have come from the theory of relativity than the remark that "a person looks different from the front than from the back." The old story about the elephant that several blind men describe in different ways is another example, perhaps, of the theory of relativity from the philosopher's point of view.

But certainly there must be deeper things in the theory of relativity than just this simple remark that "a person looks different from the front than from the back." Of course relativity is deeper than this, because *we can make definite predictions with it*. It certainly would be rather remarkable if we could predict the behavior of nature from such a simple observation alone.

There is another school of philosophers who feel very uncomfortable about the theory of relativity, which asserts that we cannot determine our absolute velocity without looking at something outside, and who would say, "It is obvious that one cannot measure his velocity without looking outside. It is self-evident that it is *meaningless* to talk about the velocity of a thing without looking outside; the physicists are rather stupid for having thought otherwise, but it has just dawned on them that this is the case. If only we philosophers had realized what the problems were that the physicists had, we could have decided immediately by brainwork that it is impossible to tell how fast one is moving without looking outside, and we could have made an enormous contribution to physics." These philosophers are always with us, struggling in the

periphery to try to tell us something, but they never really understand the subtleties and depths of the problem.

Our inability to detect absolute motion is a result of *experiment* and not a result of plain thought, as we can easily illustrate. In the first place, Newton believed that it was true that one could not tell how fast he is going if he is moving with uniform velocity in a straight line. In fact, Newton first stated the principle of relativity, and one quotation made in the last chapter was a statement of Newton's. Why then did the philosophers not make all this fuss about "all is relative," or whatever, in Newton's time? Because it was not until Maxwell's theory of electrodynamics was developed that there were physical laws that suggested that one *could* measure his velocity without looking outside; soon it was found *experimentally* that one could *not*.

Now, *is* it absolutely, definitely, philosophically *necessary* that one should not be able to tell how fast he is moving without looking outside? One of the consequences of relativity was the development of a philosophy which said, "You can only define what you can measure! Since it is self-evident that one cannot measure a velocity without seeing what he is measuring it relative to, therefore it is clear that there is no *meaning* to absolute velocity. The physicists should have realized that they can talk only about what they can measure." But *that is the whole problem*: whether or not one *can define* absolute velocity is the same as the problem of whether or not one *can detect in an experiment*, without looking outside, whether he is moving. In other words, whether or not a thing is measurable is not something to be decided *a priori* by thought alone, but something that can be decided only by experiment. Given the fact that the velocity of light is 186,000 mi/sec, one will find few philosophers who will calmly state that it is self-evident that if light goes 186,000 mi/sec inside a car, and the car is going 100,000 mi/sec, that the light also goes 186,000 mi/sec past an observer on the ground. That is a shocking fact to them; the very ones who claim it is obvious find, when you give them a specific fact, that it is not obvious.

Finally, there is even a philosophy which says that one cannot

detect *any* motion except by looking outside. It is simply not true in physics. True, one cannot perceive a *uniform* motion in a *straight line*, but if the whole room were *rotating* we would certainly know it, for everybody would be thrown to the wall—there would be all kinds of "centrifugal" effects. That the earth is turning on its axis can be determined without looking at the stars, by means of the so-called Foucault pendulum, for example. Therefore it is not true that "all is relative"; it is only *uniform velocity* that cannot be detected without looking outside. Uniform *rotation* about a fixed axis *can* be. When this is told to a philosopher, he is very upset that he did not really understand it, because to him it seems impossible that one should be able to determine rotation about an axis without looking outside. If the philosopher is good enough, after some time he may come back and say, "I understand. We really do not have such a thing as absolute rotation; we are really rotating *relative to the stars*, you see. And so some influence exerted by the stars on the object must cause the centrifugal force."

Now, for all we know, that is true; we have no way, at the present time, of telling whether there would have been centrifugal force if there were no stars and nebulae around. We have not been able to do the experiment of removing all the nebulae and then measuring our rotation, so we simply do not know. We must admit that the philosopher may be right. He comes back, therefore, in delight and says, "It is absolutely necessary that the world ultimately turn out to be this way: *absolute* rotation means nothing; it is only *relative* to the nebulae." Then we say to him, "*Now*, my friend, is it or is it not obvious that uniform velocity in a straight line, *relative to the nebulae*, should produce no effects inside a car?" Now that the motion is no longer absolute, but is a motion *relative to the nebulae*, it becomes a mysterious question, and a question that can be answered only by experiment.

What, then, *are* the philosophic influences of the theory of relativity? If we limit ourselves to influences in the sense of *what kind of new ideas and suggestions* are made to the physicist by the principle of relativity, we could describe some of them as follows. The

first discovery is, essentially, that even those ideas which have been held for a very long time and which have been very accurately verified might be wrong. It was a shocking discovery, of course, that Newton's laws are wrong, after all the years in which they seemed to be accurate. Of course it is clear, not that the experiments were wrong, but that they were done over only a limited range of velocities, so small that the relativistic effects would not have been evident. But nevertheless, we now have a much more humble point of view of our physical laws—everything *can* be wrong!

Secondly, if we have a set of "strange" ideas, such as that time goes slower when one moves, and so forth, whether we *like* them or do *not* like them is an irrelevant question. The only relevant question is whether the ideas are consistent with what is found experimentally. In other words, the "strange ideas" need only agree with *experiment*, and the only reason that we have to discuss the behavior of clocks and so forth is to demonstrate that although the notion of the time dilation is strange, it is *consistent* with the way we measure time.

Finally, there is a third suggestion which is a little more technical but which has turned out to be of enormous utility in our study of other physical laws, and that is to *look at the symmetry of the laws* or, more specifically, to look for the ways in which the laws can be transformed and leave their form the same. When we discussed the theory of vectors, we noted that the fundamental laws of motion are not changed when we rotate the coordinate system, and now we learn that they are not changed when we change the space and time variables in a particular way, given by the Lorentz transformation. So this idea of studying the patterns or operations under which the fundamental laws are not changed has proved to be a very useful one.

4-2 The twin paradox

To continue our discussion of the Lorentz transformation and relativistic effects, we consider a famous so-called paradox of Peter and Paul, who are supposed to be twins, born at the same time.

When they are old enough to drive a spaceship, Paul flies away at very high speed. Because Peter, who is left on the ground, sees Paul going so fast, all of Paul's clocks appear to go slower, his heartbeats go slower, his thoughts go slower, everything goes slower, from Peter's point of view. Of course, Paul notices nothing unusual, but if he travels around and about for a while and then comes back, he will be younger than Peter, the man on the ground! That is actually right; it is one of the consequences of the theory of relativity which has been clearly demonstrated. Just as the mu-mesons last longer when they are moving, so also will Paul last longer when he is moving. This is called a "paradox" only by the people who believe that the principle of relativity means that *all motion* is relative; they say, "Heh, heh, heh, from the point of view of Paul, can't we say that *Peter* was moving and should therefore appear to age more slowly? By symmetry, the only possible result is that both should be the same age when they meet." But in order for them to come back together and make the comparison, Paul must either stop at the end of the trip and make a comparison of clocks or, more simply, he has to come back, and the one who comes back must be the man who was moving, and he knows this, because he had to turn around. When he turned around, all kinds of unusual things happened in his spaceship—the rockets went off, things jammed up against one wall, and so on—while Peter felt nothing.

So the way to state the rule is to say that *the man who has felt the accelerations*, who has seen things fall against the walls, and so on, is the one who would be the younger; that is the difference between them in an "absolute" sense, and it is certainly correct. When we discussed the fact that moving mu-mesons live longer, we used as an example their straight-line motion in the atmosphere. But we can also make mu-mesons in a laboratory and cause them to go in a curve with a magnet, and even under this accelerated motion, they last exactly as much longer as they do when they are moving in a straight line. Although no one has arranged an experiment explicitly so that we can get rid of the paradox, one could compare a mu-meson which is left standing with one that had gone around a

complete circle, and it would surely be found that the one that went around the circle lasted longer. Although we have not actually carried out an experiment using a complete circle, it is really not necessary, of course, because everything fits together all right. This may not satisfy those who insist that every single fact be demonstrated directly, but we confidently predict the result of the experiment in which Paul goes in a complete circle.

4-3 Transformation of velocities

The main difference between the relativity of Einstein and the relativity of Newton is that the laws of transformation connecting the coordinates and times between relatively moving systems are different. The correct transformation law, that of Lorentz, is

$$\begin{aligned} x' &= \frac{x - ut}{\sqrt{1 - u^2/c^2}}, \\ y' &= y, \\ z' &= z, \\ t' &= \frac{t - ux/c^2}{\sqrt{1 - u^2/c^2}}. \end{aligned} \tag{4.1}$$

These equations correspond to the relatively simple case in which the relative motion of the two observers is along their common x-axes. Of course other directions of motion are possible, but the most general Lorentz transformation is rather complicated, with all four quantities mixed up together. We shall continue to use this simpler form, since it contains all the essential features of relativity.

Let us now discuss more of the consequences of this transformation. First, it is interesting to solve these equations in reverse. That is, here is a set of linear equations, four equations with four unknowns, and they can be solved in reverse, for x, y, z, t in terms of x', y', z', t'. The result is very interesting, since it tells us how a system of coordinates "at rest" looks from the point of view of one that is "moving." Of course, since the motions are relative and of uniform velocity, the man who is "moving" can say, if he wishes,

that it is really the other fellow who is moving and he himself who is at rest. And since he is moving in the opposite direction, he should get the same transformation, but with the opposite sign of velocity. That is precisely what we find by manipulation, so that is consistent. If it did not come out that way, we would have real cause to worry!

$$\begin{aligned} x &= \frac{x' + ut'}{\sqrt{1 - u^2/c^2}}, \\ y &= y', \\ z &= z', \\ t &= \frac{t' + ux'/c^2}{\sqrt{1 - u^2/c^2}}. \end{aligned} \tag{4.2}$$

Next we discuss the interesting problem of the addition of velocities in relativity. We recall that one of the original puzzles was that light travels at 186,000 mi/sec in all systems, even when they are in relative motion. This is a special case of the more general problem exemplified by the following. Suppose that an object inside a spaceship is going at 100,000 mi/sec and the spaceship itself is going at 100,000 mi/sec; how fast is the object inside the spaceship moving from the point of view of an observer outside? We might want to say 200,000 mi/sec, which is faster than the speed of light. This is very unnerving, because it is not supposed to be going faster than the speed of light! The general problem is as follows.

Let us suppose that the object inside the ship, from the point of view of the man inside, is moving with velocity v, and that the spaceship itself has a velocity u with respect to the ground. We want to know with what velocity v_x this object is moving from the point of view of the man on the ground. This is, of course, still but a special case in which the motion is in the x-direction. There will also be a transformation for velocities in the y-direction, or for any angle; these can be worked out as needed. Inside the spaceship the velocity is $v_{x'}$, which means that the displacement x' is equal to the velocity times the time:

$$x' = v_{x'}t'. \tag{4.3}$$

Now we have only to calculate what the position and time are from the point of view of the outside observer for an object which has the relation (4.2) between x' and t'. So we simply substitute (4.3) into (4.2), and obtain

$$x = \frac{v_{x'}t' + ut'}{\sqrt{1 - u^2/c^2}}. \tag{4.4}$$

But here we find x expressed in terms of t'. In order to get the velocity as seen by the man on the outside, we must divide *his distance* by *his time*, not by the *other man's time*! So we must also calculate the *time* as seen from the outside, which is

$$t = \frac{t' + u(v_{x'}t')/c^2}{\sqrt{1 - u^2/c^2}}. \tag{4.5}$$

Now we must find the ratio of x to t, which is

$$v_x = \frac{x}{t} = \frac{u + v_{x'}}{1 + uv_{x'}/c^2}, \tag{4.6}$$

the square roots having cancelled. This is the law that we seek: the resultant velocity, the "summing" of two velocities, is not just the algebraic sum of two velocities (we know that it cannot be or we get in trouble), but is "corrected" by $1 + uv/c^2$.

Now let us see what happens. Suppose that you are moving inside the spaceship at half the speed of light, and that the spaceship itself is going at half the speed of light. Thus u is $\frac{1}{2}c$ and v is $\frac{1}{2}c$, but in the denominator uv is one-fourth, so that

$$v = \frac{\frac{1}{2}c + \frac{1}{2}c}{1 + \frac{1}{4}} = \frac{4c}{5}.$$

So, in relativity, "half" and "half" does not make "one," it makes only "4/5." Of course low velocities can be added quite easily in the familiar way, because so long as the velocities are small compared with the speed of light we can forget about the $(1 + uv/c^2)$ factor; but things are quite different and quite interesting at high velocity.

Let us take a limiting case. Just for fun, suppose that inside the spaceship the man was observing *light itself*. In other words, $v = c$,

and yet the spaceship is moving. How will it look to the man on the ground? The answer will be

$$v = \frac{u + c}{1 + uc/c^2} = c\,\frac{u + c}{u + c} = c.$$

Therefore, if something is moving at the speed of light inside the ship, it will appear to be moving at the speed of light from the point of view of the man on the ground too! This is good, for it is, in fact, what the Einstein theory of relativity was designed to do in the first place—so it had *better* work!

Of course, there are cases in which the motion is not in the direction of the uniform translation. For example, there may be an object inside the ship which is just moving "upward" with the velocity $v_{y'}$ with respect to the ship, and the ship is moving "horizontally." Now, we simply go through the same thing, only using y's instead of x's, with the result

$$y = y' = v_{y'}t',$$

so that if $v_{x'} = 0$,

$$v_y = \frac{y}{t} = v_{y'}\sqrt{1 - u^2/c^2}. \tag{4.7}$$

Thus a sidewise velocity is no longer $v_{y'}$, but $v_{y'}\sqrt{1 - u^2/c^2}$. We found this result by substituting and combining the transformation equations, but we can also see the result directly from the principle of relativity for the following reason (it is always good to look again to see whether we can see the reason). We have already (Figure 3-3) seen how a possible clock might work when it is moving; the light appears to travel at an angle at the speed c in the fixed system, while it simply goes vertically with the same speed in the moving system. We found that the *vertical component* of the velocity in the fixed system is less than that of light by the factor $\sqrt{1 - u^2/c^2}$ (see Eq. 3-3). But now suppose that we let a material particle go back and forth in this same "clock," but at some integral fraction $1/n$ of the speed of light (Figure 4-1). Then when the particle has gone back and forth once, the light will have gone exactly n times. That is, each "click" of the "particle" clock will coincide

Figure 4-1 Trajectories described by a light ray and particle inside a moving clock.

with each *n*th "click" of the light clock. *This fact must still be true when the whole system is moving*, because the physical phenomenon of coincidence will be a coincidence in any frame. Therefore, since the speed c_y is less than the speed of light, the speed v_y of the particle must be slower than the corresponding speed by the same square-root ratio! That is why the square root appears in any vertical velocity.

4-4 Relativistic mass

We learned in the last chapter that the mass of an object increases with velocity, but no demonstration of this was given, in the sense that we made no arguments analogous to those about the way clocks have to behave. However, we *can* show that, as a consequence of relativity plus a few other reasonable assumptions, the mass must vary in this way. (We have to say "a few other assumptions" because we cannot prove anything unless we have some laws which we assume to be true, if we expect to make meaningful deductions.) To avoid the need to study the transformation laws of force, we shall analyze a *collision*, where we need know nothing about the laws of force, except that we shall assume the conservation of momentum and energy. Also, we shall assume that the momentum of a particle which is moving is a vector and is always directed in the direction of the

velocity. However, we shall not assume that the momentum is a *constant* times the velocity, as Newton did, but only that it is some *function* of velocity. We thus write the momentum vector as a certain coefficient times the vector velocity:

$$\mathbf{p} = m_v \mathbf{v}. \tag{4.8}$$

We put a subscript v on the coefficient to remind us that it is a function of velocity, and we shall agree to call this coefficient m_v the "mass." Of course, when the velocity is small, it is the same mass that we would measure in the slow-moving experiments that we are used to. Now we shall try to demonstrate that the formula for m_v must be $m_0/\sqrt{1 - v^2/c^2}$, by arguing from the principle of relativity that the laws of physics must be the same in every coordinate system.

Suppose that we have two particles, like two protons, that are absolutely equal, and they are moving toward each other with exactly equal velocities. Their total momentum is zero. Now what can happen? After the collision, their directions of motion must be exactly opposite to each other, because if they are not exactly opposite, there will be a nonzero total vector momentum, and momentum would not have been conserved. Also they must have the same speeds, since they are exactly similar objects; in fact, they must have the same speed they started with, since we suppose that the energy is conserved in these collisions. So the diagram of an elastic collision, a reversible collision, will look like Figure 4-2(a): all the arrows are the same length, all the speeds are equal. We shall suppose that such collisions can always be arranged, that any angle θ can occur, and that any speed could be used in such a collision. Next, we notice that this same collision can be viewed differently by turning the axes, and just for convenience we *shall* turn the axes, so that the horizontal splits it evenly, as in Figure 4-2(b). It is the same collision redrawn, only with the axes turned.

Now here is the real trick: let us look at this collision from the point of view of someone riding along in a car that is moving with a speed equal to the horizontal component of the velocity of one particle.

Figure 4-2 Two views of an elastic collision between equal objects moving at the same speed in opposite directions.

Then how does the collision look? It looks as though particle 1 is just going straight up, because it has lost its horizontal component, and it comes straight down again, also because it does not have that component. That is, the collision appears as shown in Figure 4-3(a). Particle 2, however, was going the other way, and as we ride past it appears to fly by at some terrific speed and at a smaller angle, but we can appreciate that the angles before and after the collision are the *same*. Let us denote by u the horizontal component of the velocity of particle 2, and by w the vertical velocity of particle 1.

Now the question is, what is the vertical velocity $u \tan \alpha$? If we knew that, we could get the correct expression for the momentum, using the law of conservation of momentum in the vertical direction. Clearly, the horizontal component of the momentum is conserved: it is the same before and after the collision for both particles, and is zero for particle 1. So we need use the conservation law only for the upward velocity $u \tan \alpha$. *But we can* get the upward velocity, simply by looking at the same collision going the other way! If we look at the collision of Figure 4-3(a) from a car moving to the left with speed u, we see the same collision, except "turned over," as shown in Figure 4-3(b). Now particle 2 is the one that goes up and down with speed w, and particle 1 has picked up the horizontal speed u. Of course, now we *know* what the velocity $u \tan \alpha$ is: it is $w\sqrt{1 - u^2/c^2}$ (see Eq. 4.7). We know that the change in the vertical momentum of the vertically moving particle is

$$\Delta p = 2m_w w$$

Figure 4-3 Two more views of the collision, from moving cars.

(2, because it moves up and back down). The obliquely moving particle has a certain velocity v whose components we have found to be u and $w\sqrt{1-u^2/c^2}$, and whose mass is m_v. The change in *vertical* momentum of this particle is therefore $\Delta p' = 2m_v w\sqrt{1-u^2/c^2}$ because, in accordance with our assumed law (4.8), the momentum component is always the mass corresponding to the magnitude of the velocity times the component of the velocity in the direction of interest. Thus in order for the total momentum to be zero the vertical momenta must cancel and the ratio of the mass moving with speed v and the mass moving with speed w must therefore be

$$\frac{m_w}{m_v} = \sqrt{1-u^2/c^2}\,. \tag{4.9}$$

Let us take the limiting case that w is infinitesimal. If w is very tiny indeed, it is clear that v and u are practically equal. In this case, $m_w \to m_0$ and $m_v \to m_u$. The grand result is

$$m_u = \frac{m_0}{\sqrt{1-u^2/c^2}}\,. \tag{4.10}$$

It is an interesting exercise now to check whether or not Eq. (4.9) is indeed true for arbitrary values of w, assuming that Eq. (4.10) is the right formula for the mass. Note that the velocity v needed in Eq. (4.9) can be calculated from the right-angle triangle:

$$v^2 = u^2 + w^2\,(1 - u^2/c^2).$$

Figure 4-4 Two views of an inelastic collision between equally massive objects.

It will be found to check out automatically, although we used it only in the limit of small w.

Now, let us accept that momentum is conserved and that the mass depends upon the velocity according to (4.10) and go on to find what else we can conclude. Let us consider what is commonly called an *inelastic collision*. For simplicity, we shall suppose that two objects of the same kind, moving oppositely with equal speeds w, hit each other and stick together, to become some new, stationary object, as shown in Figure 4-4(a). The mass m of each corresponds to w, which, as we know, is $m_0/\sqrt{1 - w^2/c^2}$. If we assume the conservation of momentum and the principle of relativity, we can demonstrate an interesting fact about the mass of the new object which has been formed. We imagine an infinitesimal velocity u at right angles to w (we can do the same with finite values of u, but it is easier to understand with an infinitesimal velocity), then look at this same collision as we ride by in an elevator at the velocity $-u$. What we see is shown in Figure 4-4(b). The composite object has an unknown mass M. Now object 1 moves with an upward component of velocity u and a horizontal component which is practically equal to w, and so also does object 2. After the collision we have the mass M moving upward with velocity u, considered very small compared with the speed of light, and also small compared with w. Momentum must be conserved, so let us estimate the momentum in the upward direction before and after the collision. Before the collision we have $p \approx 2m_w u$, and after the collision the momentum is evidently $p' = M_u u$, but M_u is essentially the same as M_0 because u is so small. These momenta must be equal because of the conservation of momentum, and therefore

$$M_0 = 2m_w. \tag{4.11}$$

The mass of the object which is formed when two equal objects collide must be twice the mass of the objects which come together. You might say, "Yes, of course, that is the conservation of mass." But not "Yes, of course," so easily, because *these masses have been enhanced* over the masses that they would be if they were standing still, yet they still contribute, to the total M, not the mass they have when standing still, but *more*. Astonishing as that may seem, in order for the conservation of momentum to work when two objects come together, the mass that they form must be greater than the rest masses of the objects, even though the objects are at rest after the collision!

4-5 Relativistic energy

In the last chapter we demonstrated that as a result of the dependence of the mass on velocity and Newton's laws, the changes in the kinetic energy of an object resulting from the total work done by the forces on it always comes out to be

$$\Delta T = (m_u - m_0)c^2 = \frac{m_0 c^2}{\sqrt{1 - u^2/c^2}} - m_0 c^2. \qquad (4.12)$$

We even went further, and guessed that the total energy is the total mass times c^2. Now we continue this discussion.

Suppose that our two equally massive objects that collide can still be "seen" inside M. For instance, a proton and a neutron are "stuck together," but are still moving about inside of M. Then, although we might at first expect the mass M to be $2m_0$, we have found that it is not $2m_0$, but $2m_w$. Since $2m_w$ is what is put in, but $2m_0$ are the rest masses of the things inside, the *excess* mass of the composite object is equal to the kinetic energy brought in. This means, of course, that *energy has inertia*. In the last chapter we discussed the heating of a gas, and showed that because the gas molecules are moving and moving things are heavier, when we put energy into the gas its molecules move faster and so the gas gets

heavier. But in fact the argument is completely general, and our discussion of the inelastic collision shows that the mass is there whether or not it is *kinetic* energy. In other words, if two particles come together and produce potential or any other form of energy; if the pieces are slowed down by climbing hills, doing work against internal forces, or whatever; then it is still true that the mass is the total energy that has been put in. So we see that the conservation of mass which we have deduced above is equivalent to the conservation of energy, and therefore there is no place in the theory of relativity for strictly inelastic collisions, as there was in Newtonian mechanics. According to Newtonian mechanics it is all right for two things to collide and so form an object of mass $2m_0$ which is in no way distinct from the one that would result from putting them together slowly. Of course we know from the law of conservation of energy that there is more kinetic energy inside, but that does not affect the mass, according to Newton's laws. But now we see that this is impossible; because of the kinetic energy involved in the collision, the resulting object will be *heavier*; therefore, it will be a *different* object. When we put the objects together gently they make something whose mass is $2m_0$; when we put them together forcefully, they make something whose mass is greater. When the mass is different, we can *tell* that it is different. So, necessarily, the conservation of energy must go along with the conservation of momentum in the theory of relativity.

This has interesting consequences. For example, suppose that we have an object whose mass M is measured, and suppose something happens so that it flies into two equal pieces moving with speed w, so that they each have a mass m_w. Now suppose that these pieces encounter enough material to slow them up until they stop; then they will have mass m_0. How much energy will they have given to the material when they have stopped? Each will give an amount $(m_w - m_0)c^2$, by the theorem that we proved before. This much energy is left in the material in some form, as heat, potential energy, or whatever. Now $2m_w = M$, so the liberated energy is $E = (M - 2m_0)c^2$. This equation was used to estimate how much energy would be

liberated under fission in the atomic bomb, for example. (Although the fragments are not exactly equal, they are nearly equal.) The mass of the uranium atom was known—it had been measured ahead of time—and the atoms into which it split, iodine, xenon, and so on, all were of known mass. By masses, we do not mean the masses while the atoms are moving, we mean the masses when the atoms are *at rest*. In other words, both M and m_0 are known. So by subtracting the two numbers one can calculate how much energy will be released if M can be made to split in "half." For this reason poor old Einstein was called the "father" of the atomic bomb in all the newspapers. Of course, all that meant was that he could tell us ahead of time how much energy would be released if we told him what process would occur. The energy that should be liberated when an atom of uranium undergoes fission was estimated about six months before the first direct test, and as soon as the energy was in fact liberated, someone measured it directly (and if Einstein's formula had not worked, they would have measured it anyway), and the moment they measured it they no longer needed the formula. Of course, we should not belittle Einstein, but rather should criticize the newspapers and many popular descriptions of what causes what in the history of physics and technology. The problem of how to get the thing to occur in an effective and rapid manner is a completely different matter.

The result is just as significant in chemistry. For instance, if we were to weigh the carbon dioxide molecule and compare its mass with that of the carbon and the oxygen, we could find out how much energy would be liberated when carbon and oxygen form carbon dioxide. The only trouble here is that the differences in masses are so small that it is technically very difficult to do.

Now let us turn to the question of whether we should add m_0c^2 to the kinetic energy and say from now on that the total energy of an object is mc^2. First, if we can still *see* the component pieces of rest mass m_0 inside M, then we could say that some of the mass M of the compound object is the mechanical rest mass of the parts, part of it is kinetic energy of the parts, and part of it is potential

energy of the parts. But we have discovered, in nature, particles of various kinds which undergo reactions just like the one we have treated above, in which with all the study in the world, *we cannot see the parts inside*. For instance, when a K-meson disintegrates into two pions it does so according to the law (4.11), but the idea that a K is made out of 2 π's is a useless idea, because it also disintegrates into 3 π's!

Therefore we have a *new idea*: we do not have to know what things are made of inside; we cannot and need not identify, inside a particle, which of the energy is rest energy of the parts into which it is going to disintegrate. It is not convenient and often not possible to separate the total mc^2 energy of an object into rest energy of the inside pieces, kinetic energy of the pieces, and potential energy of the pieces; instead, we simply speak of the *total energy* of the particle. We "shift the origin" of energy by adding a constant m_0c^2 to everything, and say that the total energy of a particle is the mass in motion times c^2, and when the object is standing still, the energy is the mass at rest times c^2.

Finally, we find that the velocity v, momentum P, and total energy E are related in a rather simple way. That the mass in motion at speed v is the mass m_0 at rest divided by $\sqrt{1 - v^2/c^2}$, surprisingly enough, is rarely used. Instead, the following relations are easily proved, and turn out to be very useful:

$$E^2 - P^2c^2 = m_0^2c^4 \tag{4.13}$$

and

$$Pc = Ev/c. \tag{4.14}$$

5

SPACE-TIME

5-1 The geometry of space-time

The theory of relativity shows us that the relationships of positions and times as measured in one coordinate system and another are not what we would have expected on the basis of our intuitive ideas. It is very important that we thoroughly understand the relations of space and time implied by the Lorentz transformation, and therefore we shall consider this matter more deeply in this chapter.

The Lorentz transformation between the positions and times (x, y, z, t) as measured by an observer "standing still," and the corresponding coordinates and time (x', y', z', t') measured inside a "moving" spaceship, moving with velocity u are

$$\begin{aligned} x' &= \frac{x - ut}{\sqrt{1 - u^2/c^2}}, \\ y' &= y, \\ z' &= z, \\ t' &= \frac{t - ux/c^2}{\sqrt{1 - u^2/c^2}}. \end{aligned} \tag{5.1}$$

Let us compare these equations with Eq. (1.5), which also relates measurements in two systems, one of which in this instance is *rotated* relative to the other:

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta, \\ y' &= y\cos\theta - x\sin\theta, \\ z' &= z. \end{aligned} \tag{5.2}$$

In this particular case, Moe and Joe are measuring with axes having an angle θ between the x'- and x-axes. In each case, we note that the "primed" quantities are "mixtures" of the "unprimed" ones: the new x' is a mixture of x and y, and the new y' is also a mixture of x and y.

An analogy is useful: When we look at an object, there is an obvious thing we might call the "apparent width," and another we might call the "depth." But the two ideas, width and depth, are not *fundamental* properties of the object, because if we step aside and look at the same thing from a different angle, we get a different width and a different depth, and we may develop some formulas for computing the new ones from the old ones and the angles involved. Equations (5.2) are these formulas. One might say that a given depth is a kind of "mixture" of all depth and all width. If it were impossible ever to move, and we always saw a given object from the same position, then this whole business would be irrelevant—we would always see the "true" width and the "true" depth, and they would appear to have quite different qualities, because one appears as a subtended optical angle and the other involves some focusing of the eyes or even intuition; they would seem to be very different things and would never get mixed up. It is because we *can* walk around that we realize that depth and width are, somehow or other, just two different aspects of the same thing.

Can we not look at the Lorentz transformations in the same way? Here also we have a mixture—of positions and the time. A difference between a space measurement and a time measurement produces a new space measurement. In other words, in the space measurements of one man there is mixed in a little bit of the time, as seen by the other. Our analogy permits us to generate this idea: the "reality" of an object that we are looking at is somehow greater (speaking crudely and intuitively) than its "width" and its "depth"

because *they* depend upon *how* we look at it; when we move to a new position, our brain immediately recalculates the width and the depth. But our brain does not immediately recalculate coordinates and time when we move at high speed, because we have had no effective experience of going nearly as fast as light to appreciate the fact that time and space are also of the same nature. It is as though we were always stuck in the position of having to look at just the width of something, not being able to move our heads appreciably one way or the other; if we could, we understand now, we would see some of the other man's time—we would see "behind," so to speak, a little bit.

Thus we shall try to think of objects in a new kind of world, of space and time mixed together, in the same sense that the objects in our ordinary space-world are real, and can be looked at from different directions. We shall then consider that objects occupying space and lasting for a certain length of time occupy a kind of "blob" in a new kind of world, and that we look at this "blob" from different points of view when we are moving at different velocities. This new world, this geometrical entity in which the "blobs" exist by occupying position and taking up a certain amount of time, is called *space-time*. A given point (x, y, z, t) in space-time is called an *event*. Imagine, for example, that we plot the x-positions horizontally, y and z in two other directions, both mutually at "right angles" and at "right angles" to the paper (!), and time, vertically. Now, how does a moving particle, say, look on such a diagram? If the particle is standing still, then it has a certain x, and as time goes on, it has the same x, the same x, the same x; so its "path" is a line that runs parallel to the t-axis (Figure 5-1[a]). On the other hand, if it drifts outward, then as the time goes on x increases (Figure 5-1[b]). So a particle, for example, which starts to drift out and then slows up should have a motion something like that shown in Figure 5-1(c). A particle, in other words, which is permanent and does not disintegrate is represented by a line in space-time. A particle which disintegrates would be represented by a forked line, because it would turn into two other things which would start from that point.

Figure 5-1 Three particle paths in space-time: (a) a particle at rest at $x = x_0$; (b) a particle which starts as $x = x_0$ and moves with constant speed; (c) a particle which starts at high speed but slows down.

What about light? Light travels at the speed c, and that would be represented by a line having a certain fixed slope (Figure 5-1[d]).

Now according to our new idea, if a given event occurs to a particle, say if it suddenly disintegrates at a certain space-time point into two new ones which follow some new tracks, and this interesting event occurred at a certain value of x and a certain value of t, then we would expect that, if this makes any sense, we just have to take a new pair of axes and turn them, and that will give us the new t and the new x in our new system, as shown in Figure 5-2(a). But this is wrong, because Eq. (5.1) is not *exactly* the same mathematical transformation as Eq. (5.2). Note, for example, the difference in sign between the two, and the fact that one is written in terms of cos θ and sin θ, while the other is written with algebraic quantities. (Of course, it is not impossible that the algebraic quantities could be written as cosine and sine, but actually they cannot.) But still, the two expressions *are* very similar.

Figure 5-2 Two views of a disintegrating particle.

As we shall see, it is not really possible to think of space-time as a real, ordinary geometry because of that difference in sign. In fact, although we shall not emphasize this point, it turns out that a man who is moving has to use a set of axes which are inclined equally to the light ray, using a special kind of projection parallel to the x'- and t'-axes, for his x' and t', as shown in Figure 5-2(b). We shall not deal with the geometry, since it does not help much; it is easier to work with the equations.

5-2 Space-time intervals

Although the geometry of space-time is not Euclidean in the ordinary sense, there *is* a geometry which is very similar, but peculiar in certain respects. If this idea of geometry is right, there ought to be some functions of coordinates and time which are independent of the coordinate system. For example, under ordinary rotations, if we take two points, one at the origin, for simplicity, and the other one somewhere else, both systems would have the same origin, and the distance from here to the other point is the same in both. That is one property that is independent of the particular way of measuring it. The square of the distance is $x^2 + y^2 + z^2$. Now what about space-time? It is not hard to demonstrate that we have here, also, something which stays the same, namely, the combination $c^2t^2 - x^2 - y^2 - z^2$ is the same before and after the transformation:

$$c^2t'^2 - x'^2 - y'^2 - z'^2 = c^2t^2 - x^2 - y^2 - z^2. \tag{5.3}$$

This quantity is therefore something which, like the distance, is "real" in some sense; it is called the *interval* between the two space-time points, one of which is, in this case, at the origin. (Actually, of course, it is the interval squared, just as $x^2 + y^2 + z^2$ is the distance squared.) We give it a different name because it is in a different geometry, but the interesting thing is only that some signs are reversed and there is a c in it.

Let us get rid of the c; that is an absurdity if we are going to have a wonderful space with x's and y's that can be interchanged. One of

the confusions that could be caused by someone with no experience would be to measure widths, say, by the angle subtended at the eye, and measure depth in a different way, like the strain on the muscles needed to focus them, so that the depths would be measured in feet and the widths in meters. Then one would get an enormously complicated mess of equations in making transformations such as (5.2), and would not be able to see the clarity and simplicity of the thing for a very simple technical reason, that the same thing is being measured in two different units. Now in Eqs. (5.1) and (5.3) nature is telling us that time and space are equivalent; time becomes space; *they should be measured in the same units*. What distance is a "second"? It is easy to figure out from (5.3) what it is. It is 3×10^8 meters, *the distance that light would go in one second*. In other words, if we were to measure all distances and times in the same units, seconds, then our unit of distance would be 3×10^8 meters, and the equations would be simpler. Or another way that we could make the units equal is to measure time in meters. What is a meter of time? A meter of time is the time it takes for light to go one meter, and is therefore $\frac{1}{3} \times 10^{-8}$ sec, or 3.3 billionths of a second! We would like, in other words, to put all our equations in a system of units in which $c = 1$. If time and space are measured in the same units, as suggested, then the equations are obviously much simplified. They are

$$\begin{aligned} x' &= \frac{x - ut}{\sqrt{1 - u^2}}, \\ y' &= y, \\ z' &= z, \\ t' &= \frac{t - ux}{\sqrt{1 - u^2}}, \end{aligned} \tag{5.4}$$

$$t'^2 - x'^2 - y'^2 - z'^2 = t^2 - x^2 - y^2 - z^2. \tag{5.5}$$

If we are ever unsure or "frightened" that after we have this system with $c = 1$ we shall never be able to get our equations right again, the answer is quite the opposite. It is much easier to remember

them without the c's in them, and it is always easy to put the c's back, by looking after the dimensions. For instance, in $\sqrt{1 - u^2}$, we know that we cannot subtract a velocity squared, which has units, from the pure number 1, so we know that we must divide u^2 by c^2 in order to make that unitless, and that is the way it goes.

The difference between space-time and ordinary space, and the character of an interval as related to the distance, is very interesting. According to formula (5.5), if we consider a point which in a given coordinate system had zero time, and only space, then the interval squared would be negative and we would have an imaginary interval, the square root of a negative number. Intervals can be either real or imaginary in the theory. The square of an interval may be either positive or negative, unlike distance, which has a positive square. When an interval is imaginary, we say that the two points have a *space-like interval* between them (instead of imaginary), because the interval is more like space than like time. On the other hand, if two objects are at the same place in a given coordinate system, but differ only in time, then the square of the time is positive and the distances are zero and the interval squared is positive; this is called a *time-like interval*. In our diagram of space-time, therefore, we would have a representation something like this: at 45° there are two lines (actually, in four dimensions these will be "cones," called light cones) and points on these lines are all at zero interval from the origin. Where light goes from a given point is always separated from it by a zero interval, as we see from Eq. (5.5). Incidentally, we have just proved that if light travels with speed c in one system, it travels with speed c in another, for if the interval is the same in both systems, i.e., zero in one and zero in the other, then to state that the propagation speed of light is invariant is the same as saying that the interval is zero.

5-3 Past, present, and future

The space-time region surrounding a given space-time point can be separated into three regions, as shown in Figure 5-3. In one region

Figure 5-3 The space-time region surrounding a point at the origin.

we have space-like intervals, and in two regions, time-like intervals. Physically, these three regions into which space-time around a given point is divided have an interesting physical relationship to that point: a physical object or a signal can get from a point in region 2 to the event O by moving along at a speed less than the speed of light. Therefore events in this region can affect the point O, can have an influence on it from the past. In fact, of course, an object at P on the negative t-axis is precisely in the "past" with respect to O; it is the same space-point as O, only earlier. What happened there then, affects O now. (Unfortunately, that is the way life is.) Another object at Q can get to O by moving with a certain speed less than c, so if this object were in a spaceship and moving, it would be, again, the past of the same space-point. That is, in another coordinate system, the axis of time might go through both O and Q. So all points of region 2 are in the "past" of O, and anything that happens in this region *can* affect O. Therefore region 2 is sometimes called the *affective past*, or affecting past; it is the locus of all events which can affect point O in any way.

Region 3, on the other hand, is a region which we can affect *from* O, we can "hit" things by shooting "bullets" out at speeds less than c. So this is the world whose future can be affected by us, and we may call that the *affective future*. Now the interesting thing about all the rest of space-time, i.e., region 1, is that we can neither affect it now *from* O, nor can it affect us now *at* O, because nothing can go faster than the speed of light. Of course, what happens at R

can affect us *later*; that is, if the sun is exploding "right now," it takes eight minutes before we know about it, and it cannot possibly affect us before then.

What we mean by "right now" is a mysterious thing which we cannot define and we cannot affect, but it can affect us later, or we could have affected it if we had done something far enough in the past. When we look at the star Alpha Centauri, we see it as it was four years ago; we might wonder what it is like "now." "Now" means at the same time from our special coordinate system. We can only see Alpha Centauri by the light that has come from our past, up to four years ago, but we do not know what it is doing "now"; it will take four years before what it is doing "now" can affect us. Alpha Centauri "now" is an idea or concept of our mind; it is not something that is really definable physically at the moment, because we have to wait to observe it; we cannot even define it right "now." Furthermore, the "now" depends on the coordinate system. If, for example, Alpha Centauri were moving, an observer there would not agree with us because he would put his axes at an angle, and his "now" would be a *different* time. We have already talked about the fact that simultaneity is not a unique thing.

There are fortune-tellers, or people who tell us they can know the future, and there are many wonderful stories about the man who suddenly discovers that he has knowledge about the affective future. Well, there are lots of paradoxes produced by that because if we know something is going to happen, then we can make sure we will avoid it by doing the right thing at the right time, and so on. But actually there is no fortune-teller who can even tell us the *present*! There is no one who can tell us what is really happening right now, at any reasonable distance, because that is unobservable. We might ask ourselves this question, which we leave to the student to try to answer: Would any paradox be produced if it were suddenly to become possible to know things that are in the space-like intervals of region 1?

5-4 More about four-vectors

Let us now return to our consideration of the analogy of the Lorentz transformation and rotations of the space axes. We have learned the utility of collecting together other quantities which have the same transformation properties as the coordinates, to form what we call *vectors*, directed lines. In the case of ordinary rotations, there are many quantities that transform the same way as x, y, and z under rotation: for example, the velocity has three components, an x, y, and z component; when seen in a different coordinate system, none of the components is the same, instead they are all transformed to new values. But, somehow or other, the velocity "itself" has a greater reality than do any of its particular components, and we represent it by a directed line.

We therefore ask: Is it or is it not true that there are quantities which transform, or which are related, in a moving system and in a nonmoving system, in the same way as x, y, z, and t? From our experience with vectors, we know that three of the quantities, like x, y, z, would constitute the three components of an ordinary space-vector, but the fourth quantity would look like an ordinary scalar under space rotation, because it does not change so long as we do not go into a moving coordinate system. Is it possible, then, to associate with some of our known "three-vectors" a fourth object, that we could call the "time component," in such a manner that the four objects together would "rotate" the same way as position and time in space-time? We shall now show that there is, indeed, at least one such thing (there are many of them, in fact): *the three components of momentum, and the energy as the time component, transform together* to make what we call a "four-vector." In demonstrating this, since it is quite inconvenient to have to write c's everywhere, we shall use the same trick concerning units of the energy, the mass, and the momentum that we used in Eq. (5.4). Energy and mass, for example, differ only by a factor c^2 which is merely a question of units, so we can say energy *is* the mass. Instead of having to write the c^2, we put $E = m$, and then, of course, if there were

any trouble we would put in the right amounts of c so that the units would straighten out in the last equation, but not in the intermediate ones.

Thus our equations for energy and momentum are

$$\begin{aligned} E &= m = m_0/\sqrt{1-v^2}, \\ \mathbf{p} &= m\mathbf{v} = m_0\mathbf{v}/\sqrt{1-v^2}. \end{aligned} \tag{5.6}$$

Also in these units, we have

$$E^2 - p^2 = m_0^2. \tag{5.7}$$

For example, if we measure energy in electron volts, what does a mass of 1 electron volt mean? It means the mass whose rest energy is 1 electron volt, that is, m_0c^2 is one electron volt. For example, the rest mass of an electron is 0.511×10^6 ev.

Now what would the momentum and energy look like in a new coordinate system? To find out, we shall have to transform Eq. (5.6), which we can do because we know how the velocity transforms. Suppose that, as we measure it, an object has a velocity v, but we look upon the same object from the point of view of a spaceship which itself is moving with a velocity u, and in that system we use a prime to designate the corresponding thing. In order to simplify things at first, we shall take the case that the velocity v is in the direction of u. (Later, we can do the more general case.) What is v', the velocity as seen from the spaceship? It is the composite velocity, the "difference" between v and u. By the law which we worked out before,

$$v' = \frac{v-u}{1-uv}. \tag{5.8}$$

Now let us calculate the new energy E', the energy as the fellow in the spaceship would see it. He would use the same rest mass, of course, but he would use v' for the velocity. What we have to do is square v', subtract it from one, take the square root, and take the reciprocal:

$$
\begin{aligned}
v'^2 &= \frac{v^2 - 2uv + u^2}{1 - 2uv + u^2v^2}, \\
1 - v'^2 &= \frac{1 - 2uv + u^2v^2 - v^2 + 2uv - u^2}{1 - 2uv + u^2v^2} \\
&= \frac{1 - v^2 - u^2 + u^2v^2}{1 - 2uv + u^2v^2} \\
&= \frac{(1 - v^2)(1 - u^2)}{(1 - uv)^2}.
\end{aligned}
$$

Therefore

$$
\frac{1}{\sqrt{1 - v'^2}} = \frac{1 - uv}{\sqrt{1 - v^2}\sqrt{1 - u^2}}. \tag{5.9}
$$

The energy E' is then simply m_0 times the above expression. But we want to express the energy in terms of the unprimed energy and momentum, and we note that

$$
E' = \frac{m_0 - m_0uv}{\sqrt{1 - v^2}\sqrt{1 - u^2}} = \frac{(m_0/\sqrt{1 - v^2}) - (m_0v/\sqrt{1 - v^2})u}{\sqrt{1 - u^2}},
$$

or

$$
E' = \frac{E - up_x}{\sqrt{1 - u^2}}, \tag{5.10}
$$

which we recognize as being exactly of the same form as

$$
t' = \frac{t - ux}{\sqrt{1 - u^2}}.
$$

Next we must find the new momentum p'_x. This is just the energy E' times v', and is also simply expressed in terms of E and p:

$$
p'_x = E'v' = \frac{m_0(1 - uv)}{\sqrt{1 - v^2}\sqrt{1 - u^2}} \cdot \frac{v - u}{(1 - uv)} = \frac{m_0v - m_0u}{\sqrt{1 - v^2}\sqrt{1 - u^2}}
$$

Thus

$$
p'_x = \frac{p_x - uE}{\sqrt{1 - u^2}}, \tag{5.11}
$$

which we recognize as being of precisely the same form as

$$x' = \frac{x - ut}{\sqrt{1 - u^2}}.$$

Thus the transformations for the new energy and momentum in terms of the old energy and momentum are exactly the same as the transformations for t' in terms of t and x, and x' in terms of x and t: all we have to do is, every time we see t in (5.4) substitute E, and every time we see x substitute p_x, and then the equations (5.4) will become the same as Eqs. (5.10) and (5.11). This would imply, if everything works right, an additional rule that $p'_y = p_y$ and that $p'_z = p_z$. To prove this would require our going back and studying the case of motion up and down. Actually, we did study the case of motion up and down in the last chapter. We analyzed a complicated collision and we noticed that, in fact, the transverse momentum is *not* changed when viewed from a moving system; so we have already verified that $p'_y = p_y$ and $p'_z = p_z$. The complete transformation, then, is

$$\begin{aligned} p'_x &= \frac{p_x - uE}{\sqrt{1 - u^2}}, \\ p'_y &= p_y, \\ p'_z &= p_z, \\ E' &= \frac{E - up_x}{\sqrt{1 - u^2}}. \end{aligned} \tag{5.12}$$

In these transformations, therefore, we have discovered four quantities which transform like x, y, z, and t, and which we call the *four-vector momentum*. Since the momentum is a four-vector, it can be represented on a space-time diagram of a moving particle as an "arrow" tangent to the path, as shown in Figure 5-4. This arrow has a time component equal to the energy, and its space components represent its three-vector momentum; this arrow is more "real" than either the energy or the momentum, because those just depend on how we look at the diagram.

Figure 5-4 The four-vector momentum of a particle.

5-5 Four-vector algebra

The notation for four-vectors is different than it is for three-vectors. In the case of three-vectors, if we were to talk about the ordinary three-vector momentum we would write it **p**. If we wanted to be more specific, we could say it has three components which are, for the axes in question, p_x, p_y, and p_z, or we could simply refer to a general component as p_i, and say that i could either be x, y, or z, and that these are the three components; that is, imagine that i is any one of three directions, x, y, or z. The notation that we use for four-vectors is analogous to this: we write p_μ for the four-vector, and μ stands for the *four* possible directions t, x, y, or z.

We could, of course, use any notation we want; do not laugh at notations; invent them, they are powerful. In fact, mathematics is, to a large extent, invention of better notations. The whole idea of a four-vector, in fact, is an improvement in notation so that the transformations can be remembered easily. A_μ, then, is a general four-vector, but for the special case of momentum, the p_t is identified as the energy, p_x is the momentum in the x-direction, p_y is that in the y-direction, and p_z is that in the z-direction. To add four-vectors, we add the corresponding components.

If there is an equation among four-vectors, then the equation is true for *each component*. For instance, if the law of conservation of three-vector momentum is to be true in particle collisions, i.e., if the sum of the momenta for a large number of interacting or colliding particles is to be a constant, that must mean that the sums of all momenta in the x-direction, in the y-direction, and in the z-

direction, for all the particles, must each be constant. This law alone would be impossible in relativity because it is *incomplete*; it is like talking about only two of the components of a three-vector. It is incomplete because if we rotate the axes, we mix the various components, so we must include all three components in our law. Thus, in relativity, we must complete the law of conservation of momentum by extending it to include the *time* component. This is *absolutely necessary* to go with the other three, or there cannot be relativistic invariance. The *conservation of energy* is the fourth equation which goes with the conservation of momentum to make a valid four-vector relationship in the geometry of space and time. Thus the law of conservation of energy and momentum in four-dimensional notation is

$$\sum_{\substack{\text{particles}\\ \text{in}}} p_\mu = \sum_{\substack{\text{particles}\\ \text{out}}} p_\mu, \tag{5.13}$$

or, in a slightly different notation

$$\sum_i p_{i\mu} = \sum_j p_{j\mu}, \tag{5.14}$$

where $i = 1, 2, \ldots$ refers to the particles going into the collision; $j = 1, 2, \ldots$ refers to the particles coming out of the collision; and $\mu = x, y, z$, or t. You say, "In which axes?" It makes no difference. The law is true for each component, using *any* axes.

In vector analysis we discussed one other thing, the dot product of two vectors. Let us now consider the corresponding thing in space-time. In ordinary rotation we discovered there was an unchanged quantity $x^2 + y^2 + z^2$. In four dimensions, we find that the corresponding quantity is $t^2 - x^2 - y^2 - z^2$ (Eq. 5.3). How can we write that? One way would be to write some kind of four-dimensional thing with a square dot between, like $A_\mu \diamondsuit B_\mu$; one of the notations which is actually used is

$$\sum_\mu{}' A_\mu A_\mu = A_t^2 - A_x^2 - A_y^2 - A_z^2. \tag{5.15}$$

The prime on $\sum$ means that the first term, the "time" term, is positive, but the other three terms have minus signs. This quantity, then, will be the same in any coordinate system, and we may call it the square of the length of the four-vector. For instance, what is the square of the length of the four-vector momentum of a single particle? This will be equal to $p_t^2 - p_x^2 - p_y^2 - p_z^2$ or, in other words, $E^2 - p^2$, because we know that p_t is E. What is $E^2 - p^2$? It must be something which is the same in every coordinate system. In particular, it must be the same for a coordinate system which is moving right along with the particle, in which the particle is standing still. If the particle is standing still, it would have no momentum. So in that coordinate system, it is purely its energy, which is the same as its rest mass. Thus $E^2 - p^2 = m_0^2$. So we see that the square of the length of this vector, the four-vector momentum, is equal to m_0^2.

From the square of a vector, we can go on to invent the "dot product," or the product which is a scalar: if a_μ is one four-vector and b_μ is another four-vector, then the scalar product is

$$\sum{}' a_\mu b_\mu = a_t b_t - a_x b_x - a_y b_y - a_z b_z. \qquad (5.16)$$

It is the same in all coordinate systems.

Finally, we shall mention certain things whose rest mass m_0 is zero. A photon of light, for example. A photon is like a particle, in that it carries an energy and a momentum. The energy of a photon is a certain constant, called Planck's constant, times the frequency of the photon: $E = h\nu$. Such a photon also carries a momentum, and the momentum of a photon (or of any other particle, in fact) is h divided by the wavelength: $p = h/\lambda$. But, for a photon, there is a definite relationship between the frequency and the wavelength: $\nu = c/\lambda$. (The number of waves per second, times the wavelength of each, is the distance that the light goes in one second, which, of course, is c.) Thus we see immediately that the energy of a photon must be the momentum times c, or if $c = 1$, *the energy and momentum are equal*. That is to say, the rest mass is zero. Let us look at

that again; that is quite curious. If it is a particle of zero rest mass, what happens when it stops? *It never stops!* It always goes at the speed c. The usual formula for energy is $m_0/\sqrt{1-v^2}$. Now can we say that $m_0 = 0$ and $v = 1$, so the energy is 0? We *cannot* say that it is zero; the photon really can (and does) have energy even though it has no rest mass, but this it possesses by perpetually going at the speed of light!

We also know that the momentum of any particle is equal to its total energy times its velocity: if $c = 1$, $p = vE$ or, in ordinary units, $p = vE/c^2$. For any particle moving at the speed of light, $p = E$ if $c = 1$. The formulas for the energy of a photon as seen from a moving system are, of course, given by Eq. (5.12), but for the momentum we must substitute the energy times c (or times 1 in this case). The different energies after transformation mean that there are different frequencies. This is called the Doppler effect, and one can calculate it easily from Eq. (5.12), using also $E = p$ and $E = h\nu$.

As Minkowski said, "Space of itself, and time of itself will sink into mere shadows, and only a kind of union between them shall survive."

6

CURVED SPACE

6-1 Curved spaces with two dimensions

According to Newton, everything attracts everything else with a force inversely proportional to the square of the distance from it, and objects respond to forces with accelerations proportional to the forces. They are Newton's laws of universal gravitation and of motion. As you know, they account for the motions of balls, planets, satellites, galaxies, and so forth.

Einstein had a different interpretation of the law of gravitation. According to him, space and time—which must be put together as space-time—are *curved* near heavy masses. And it is the attempt of things to go along "straight lines" in this curved space-time which makes them move the way they do. Now that is a complex idea—very complex. It is the idea we want to explain in this chapter.

Our subject has three parts. One involves the effects of gravitation. Another involves the ideas of space-time which we already studied. The third involves the idea of curved space-time. We will simplify our subject in the beginning by not worrying about gravity and by leaving out the time—discussing just curved space. We will talk later about the other parts, but we will concentrate now on the idea of curved space—what is meant by curved space, and, more specifically, what is meant by curved space in this application of Einstein. Now even that much turns out to be somewhat difficult

in three dimensions. So we will first reduce the problem still further and talk about what is meant by the words "curved space" in two dimensions.

In order to understand this idea of curved space in two dimensions you really have to appreciate the limited point of view of the character who lives in such a space. Suppose we imagine a bug with no eyes who lives on a plane, as shown in Figure 6-1. He can move only on the plane, and he has no way of knowing that there is any way to discover any "outside world." (He hasn't got your imagination.) We are, of course, going to argue by analogy. *We* live in a three-dimensional world, and we don't have any imagination about going off our three-dimensional world in a new direction; so we have to think the thing out by analogy. It is as though we were bugs living on a plane, and there was a space in another direction. That's why we will first work with the bug, remembering that he must live on his surface and can't get out.

As another example of a bug living in two dimensions, let's imagine one who lives on a sphere. We imagine that he can walk around on the surface of the sphere, as in Figure 6-2, but that he can't look "up," or "down," or "out."

Now we want to consider still a *third* kind of creature. He is also a bug like the others, and also lives on a plane, as our first bug did, but this time the plane is peculiar. The temperature is different at different places. Also, the bug and any rulers he uses are all made

Figure 6-1 A bug on a plane surface.

Figure 6-2 A bug on a sphere.

of the same material which expands when it is heated. Whenever he puts a ruler somewhere to measure something the ruler expands immediately to the proper length for the temperature at that place. Wherever he puts any object—himself, a ruler, a triangle, or anything—the thing stretches itself because of the thermal expansion. Everything is longer in the hot places than it is in the cold places, and everything has the same coefficient of expansion. We will call the home of our third bug a "hot plate," although we will particularly want to think of a special kind of hot plate that is cold in the center and gets hotter as we go out toward the edges (Figure 6-3).

Now we are going to imagine that our bugs begin to study geometry. Although we imagine that they are blind so that they can't see any "outside" world, they can do a lot with their legs and feelers.

Figure 6-3 A bug on a hot plate.

They can draw lines, and they can make rulers, and measure off lengths. First, let's suppose that they start with the simplest idea in geometry. They learn how to make a straight line—defined as the shortest line between two points. Our first bug—see Figure 6-4—learns to make very good lines. But what happens to the bug on the sphere? He draws his straight line as the shortest distance—*for him*—between two points, as in Figure 6-5. It may look like a curve to us, but he has no way of getting off the sphere and finding out that there is "really" a shorter line. He just knows that if he tries any other path *in his world* it is always longer than his straight line. So we will let him have his straight line as the shortest arc between two points. (It is, of course, an arc of a great circle.)

Finally, our third bug—the one in Figure 6-3—will also draw "straight lines" that look like curves to us. For instance, the shortest distance between *A* and *B* in Figure 6-6 would be on a curve like the one shown. Why? Because when his line curves out toward the warmer parts of his hot plate, the rulers get longer (from our omniscient point of view) and it takes fewer "yardsticks" laid end-to-end to get from *A* to *B*. So *for him* the line is straight—he has no way of knowing that there could be someone out in a strange three-dimensional world who would call a different line "straight."

We think you get the idea now that all the rest of the analysis will always be from the point of view of the creatures on the particular surfaces and not from *our* point of view. With that in mind let's see what the rest of their geometries look like. Let's assume that

Figure 6-4 Making a "straight line" on a plane.

Figure 6-5 Making a "straight line" on a sphere.

the bugs have all learned how to make two lines intersect at right angles. (You can figure out how they could do it.) Then our first bug (the one on the normal plane) finds an interesting fact. If he starts at the point *A* and makes a line 100 inches long, then makes a right angle and marks off another 100 inches, then makes another right angle and goes another 100 inches, then makes a third right angle and a fourth line 100 inches long, he ends up right at the starting point as shown in Figure 6-7(a). It is a property of his world—one of the facts of his "geometry."

Then he discovers another interesting thing. If he makes a triangle—a figure with three straight lines—the sum of the angles is equal to 180°, that is, to the sum of two right angles. See Figure 6-7(b).

Then he invents the circle. What's a circle? A circle is made this way: you rush off on straight lines in many many directions from

Figure 6-6 Making a "straight line" on the hot plate.

Figure 6-7 A square, triangle, and circle in a flat space.

a single point, and lay out a lot of dots that are all the same distance from that point. See Figure 6-7(c). (We have to be careful how we define these things because we've got to be able to make the analogs for the other fellows.) Of course, it's equivalent to the curve you can make by swinging a ruler around a point. Anyway, our bug learns how to make circles. Then one day he thinks of measuring the distance around a circle. He measures several circles and finds a neat relationship: The distance around is always the same number times the radius r (which is, of course, the distance from the center out to the curve). The circumference and the radius always have the same ratio—approximately 6.283—independent of the size of the circle.

Now let's see what our other bugs have been finding out about *their* geometries. First, what happens to the bug on the sphere when he tries to make a "square"? If he follows the prescription we gave above, he would probably think that the result was hardly worth the trouble. He gets a figure like the one shown in Figure 6-8. His end-point B isn't on top of the starting point A. It doesn't work out to a closed figure at all. Get a sphere and try it. A similar thing would happen to our friend on the hot plate. If he lays out four straight lines of equal length—as measured with his expanding rulers—joined by right angles he gets a picture like the one in Figure 6-9.

Figure 6-8 Trying to make a "square" on a sphere.

Now suppose that our bugs had each had their own Euclid who had told them what geometry "should" be like, and that they had checked him out roughly by making crude measurements on a *small* scale. Then as they tried to make accurate squares on a larger scale they would discover that something was wrong. The point is, that just by *geometrical measurements* they would discover that something was the matter with their space. We define a *curved space* to be a space in which the geometry is not what we expect for a plane. The geometry of the bugs on the sphere or on the hot plate is the geometry of a curved space. The rules of Euclidean geometry fail. And it isn't necessary to be able to lift yourself out of the plane in order to find out that the world that you live in is curved. It isn't necessary to circumnavigate the globe in order to find out that it is a ball. You can find out that you live on a ball by laying out a

Figure 6-9 Trying to make a "square" on the hot plate.

square. If the square is very small you will need a lot of accuracy, but if the square is large the measurement can be done more crudely.

Let's take the case of a triangle on a plane. The sum of the angles is 180 degrees. Our friend on the sphere can find triangles that are very peculiar. He can, for example, find triangles which have *three right angles*. Yes indeed! One is shown in Figure 6-10. Suppose our bug starts at the north pole and makes a straight line all the way down to the equator. Then he makes a right angle and another perfect straight line the same length. Then he does it again. For the very special length he has chosen he gets right back to his starting point, and also meets the first line with a right angle. So there is no doubt that for him this triangle has three right angles, or 270 degrees in the sum. It turns out that for him the sum of the angles of the triangle is *always* greater than 180 degrees. In fact, the excess (for the special case shown, the extra 90 degrees) is proportional to how much area the triangle has. If a triangle on a sphere is very small, its angles add up to very nearly 180 degrees, only a little bit over. As the triangle gets bigger the discrepancy goes up. The bug on the hot plate would discover similar difficulties with his triangles.

Let's look next at what our other bugs find out about circles. They make circles and measure their circumferences. For example,

Figure 6-10 On a sphere a "triangle" can have three 90° angles.

the bug on the sphere might make a circle like the one shown in Figure 6-11. And he would discover that the circumference is *less* than 2π times the radius. (You can see that because from the wisdom of our three-dimensional view it is obvious that what he calls the "radius" is a curve which is *longer* than the true radius of the circle.) Suppose that the bug on the sphere had read Euclid, and decided to predict a radius by dividing the circumference C by 2π, taking

$$r_{\text{pred}} = \frac{C}{2\pi}. \tag{6.1}$$

Then he would find that the measured radius was larger than the predicted radius. Pursuing the subject, he might define the difference to be the "excess radius," and write

$$r_{\text{meas}} - r_{\text{pred}} = r_{\text{excess}}, \tag{6.2}$$

and study how the excess radius effect depended on the size of the circle.

Our bug on the hot plate would discover a similar phenomenon. Suppose he was to draw a circle centered at the cold spot on the plate as in Figure 6-12. If we were to watch him as he makes the circle we would notice that his rulers are short near the center and get longer as they are moved outward—although the bug doesn't know it, of course. When he measures the circumference the ruler

Figure 6-11 Making a circle on a sphere.

Figure 6-12 Making a circle on the hot plate.

is long all the time, so he, too, finds out that the measured radius is longer than the predicted radius, $C/2\pi$. The effect depends on the radius of the circle.

We will *define* a "curved space" as one in which these types of geometrical errors occur: the sum of the angles of a triangle is different from 180 degrees; the circumference of a circle divided by 2π is not equal to the radius; the rule for making a square doesn't give a closed figure. You can think of others.

We have given two different examples of curved space: the sphere and the hot plate. But it is interesting that if we choose the right temperature variation as a function of distance on the hot plate, the two *geometries* will be exactly the same. It is rather amusing. We can make the bug on the hot plate get exactly the same answers as the bug on the ball. For those who like geometry and geometrical problems we'll tell you how it can be done. If you assume that the length of the rulers (as determined by the temperature) goes in proportion to one plus some constant times the square of the distance away from the origin, then you will find that the geometry of that hot plate is exactly the same in all details* as the geometry of the sphere.

There are, of course, other kinds of geometry. We could ask about the geometry of a bug who lived on a pear, namely something which has a sharper curvature in one place and a weaker curvature

* Except for the one point at infinity.

in the other place, so that the excess in angles in triangles is more severe when he makes little triangles in one part of his world than when he makes them in another part. In other words, the curvature of a space can vary from place to place. That's just a generalization of the idea. It can also be imitated by a suitable distribution of temperature on a hot plate.

We may also point out that the results could come out with the opposite kind of discrepancies. You could find out, for example, that all triangles when they are made too large have the sum of their angles *less* than 180 degrees. That may sound impossible, but it isn't at all. First of all, we could have a hot plate with the temperature decreasing with the distance from the center. Then all the effects would be reversed. But we can also do it purely geometrically by looking at the two-dimensional geometry of the surface of a saddle. Imagine a saddle-shaped surface like the one sketched in Figure 6-13. Now draw a "circle" on the surface, defined as the locus of all points the same distance from a center. This circle is a curve that oscillates up and down with a scallop effect. So its circumference is larger than you would expect from calculating $2\pi r$. So $C/2\pi$ is now less than r. The "excess radius" would be negative.

Spheres and pears and such are all surfaces of *positive* curvatures; and the others are called surfaces of *negative* curvature. In general, a two-dimensional world will have a curvature which varies from place to place and may be positive in some places and negative in

Figure 6-13 A "circle" on a saddle-shaped surface.

other places. In general, we mean by a curved space simply one in which the rules of Euclidean geometry break down with one sign of discrepancy or the other. The amount of curvature—defined, say, by the excess radius—may vary from place to place.

We might point out that, from our definition of curvature, a cylinder is, surprisingly enough, not curved. If a bug lived on a cylinder, as shown in Figure 6-14, he would find out that triangles, squares, and circles would all have the same behavior they have on a plane. This is easy to see, by just thinking about how all the figures will look if the cylinder is unrolled onto a plane. Then all the geometrical figures can be made to correspond exactly to the way they are in a plane. So there is no way for a bug living on a cylinder (assuming that he doesn't go all the way around, but just makes local measurements) to discover that his space is curved. In our technical sense, then, we consider that his space is *not* curved. What we want to talk about is more precisely called *intrinsic* curvature; that is, a curvature which can be found by measurements only in a local region. (A cylinder has no intrinsic curvature.) This

Figure 6-14 A two-dimensional space with zero intrinsic curvature.

was the sense intended by Einstein when he said that our space is curved. But we as yet only have defined a curved space in two dimensions; we must go onward to see what the idea might mean in three dimensions.

6-2 Curvature in three-dimensional space

We live in three-dimensional space and we are going to consider the idea that three-dimensional space is curved. You say, "But how can you imagine it being bent in any direction?" Well, we can't imagine space being bent in any direction because our imagination isn't good enough. (Perhaps it's just as well that we can't imagine too much, so that we don't get too free of the real world.) But we can still *define* a curvature without getting out of our three-dimensional world. All we have been talking about in two dimensions was simply an exercise to show how we could get a definition of curvature which didn't require that we be able to "look in" from the outside.

We can determine whether our world is curved or not in a way quite analogous to the one used by the gentlemen who live on the sphere and on the hot plate. We may not be able to distinguish between two such cases but we certainly can distinguish those cases from the flat space, the ordinary plane. How? Easy enough: We lay out a triangle and measure the angles. Or we make a great big circle and measure the circumference and the radius. Or we try to lay out some accurate squares, or try to make a cube. In each case we test whether the laws of geometry work. If they don't work, we say that our space is curved. If we lay out a big triangle and the sum of its angles exceeds 180 degrees, we can say our space is curved. Or if the measured radius of a circle is not equal to its circumference over 2π, we can say our space is curved.

You will notice that in three dimensions the situation can be much more complicated than in two. At any one place in two dimensions there is a certain amount of curvature. But in three dimensions there can be *several components* to the curvature. If we lay

out a triangle in some plane, we may get a different answer than if we orient the plane of the triangle in a different way. Or take the example of a circle. Suppose we draw a circle and measure the radius and it doesn't check with $C/2\pi$ so that there is some excess radius. Now we draw another circle at right angles—as in Figure 6-15. There's no need for the excess to be exactly the same for both circles. In fact, there might be a positive excess for a circle in one plane, and a defect (negative excess) for a circle in the other plane.

Perhaps you are thinking of a better idea: Can't we get around all of these components by using a *sphere* in three dimensions? We can specify a sphere by taking all the points that are the same distance from a given point in space. Then we can measure the surface area by laying out a fine-scale rectangular grid on the surface of the sphere and adding up all the bits of area. According to Euclid the total area A is supposed to be 4π times the square of the radius; so we can define a "predicted radius" as $\sqrt{A/4\pi}$. But we can also measure the radius directly by digging a hole to the center and measuring the distance. Again, we can take the measured radius minus the predicted radius and call the difference the radius excess.

$$r_{\text{excess}} = r_{\text{meas}} - \left(\frac{\text{measured area}}{4\pi}\right)^{1/2},$$

which would be a perfectly satisfactory measure of the curvature.

Figure 6-15 The excess radius may be different for circles with different orientations.

It has the great advantage that it doesn't depend upon how we orient a triangle or a circle.

But the excess radius of a sphere also has a disadvantage: it doesn't completely characterize the space. It gives what is called the *mean curvature* of the three-dimensional world, since there is an averaging effect over the various curvatures. Since it is an average, however, it does not solve completely the problem of defining the geometry. If you know only this number you can't predict all properties of the geometry of the space, because you can't tell what would happen with circles of different orientation. The complete definition requires the specification of six "curvature numbers" at each point. Of course the mathematicians know how to write all those numbers. You can read someday in a mathematics book how to write them all in a high-class and elegant form, but it is first a good idea to know in a rough way what it is that you are trying to write about. For most of our purposes the average curvature will be enough.*

6-3 Our space is curved

Now comes the main question. Is it true? That is, is the actual physical three-dimensional space we live in curved? Once we have enough imagination to realize the possibility that space might be curved, the human mind naturally gets curious about whether the real world is curved or not. People have made direct geometrical measurements to try to find out, and haven't found any deviations.

---

* We should mention one additional point for completeness. If you want to carry the hot-plate model of curved space over into three dimensions you must imagine that the length of the ruler depends not only on where you put it, but also on which orientation the ruler has when it is laid down. It is a generalization of the simple case in which the length of the ruler depends on where it is, but is the same if set north-south, or east-west, or up-down. This generalization is needed if you want to represent a three-dimensional space with any arbitrary geometry with such a model, although it happens not to be necessary for two dimensions.

On the other hand, by arguments about gravitation, Einstein discovered that space *is* curved, and we'd like to tell you what Einstein's law is for the amount of curvature, and also tell you a little bit about how he found out about it.

Einstein said that space is curved and that matter is the source of the curvature. (Matter is also the source of gravitation, so gravity is related to the curvature—but that will come later in the chapter.) Let us suppose, to make things a little easier, that the matter is distributed continuously with some density, which may vary, however, as much as you want from place to place.* The rule that Einstein gave for the curvature is the following: If there is a region of space with matter in it and we take a sphere small enough that the density ρ of matter inside it is effectively constant, then the *radius excess* for the sphere is proportional to the mass inside the sphere. Using the definition of excess radius, we have

$$\text{Radius excess} = r_{\text{meas}} - \sqrt{\frac{A}{4\pi}} = \frac{G}{3c^2} \cdot M\,. \qquad (6.3)$$

Here, G is the gravitational constant (of Newton's theory), c is the velocity of light, and $M = 4\pi\rho r^3/3$ is the mass of the matter inside the sphere. This is Einstein's law for the mean curvature of space.

Suppose we take the earth as an example and forget that the density varies from point to point—so we won't have to do any integrals. Suppose we were to measure the surface of the earth very carefully, and then dig a hole to the center and measure the radius. From the surface area we could calculate the predicted radius we would get from setting the area equal to $4\pi r^2$. When we compared the predicted radius with the actual radius, we would find that the actual radius exceeded the predicted radius by the amount given in Eq. (6.3). The constant $G/3c^2$ is about 2.5×10^{-29} cm per gram, so for each gram of material the measured radius is off by 2.5×10^{-29}cm. Putting in the mass of the earth, which is about 6×10^{27} grams, it turns out

* Nobody—not even Einstein—knows how to do it if mass comes concentrated at points.

that the earth has 1.5 millimeters more radius than it should have for its surface area.* Doing the same calculation for the sun, you find that the sun's radius is one-half a kilometer too long.

You should note that the law says that the *average* curvature *above* the surface area of the earth is zero. But that does *not* mean that all the components of the curvature are zero. There may still be—and, in fact, there is—some curvature above the earth. For a circle in a plane there will be an excess radius of one sign for some orientations and of the opposite sign for other orientations. It just turns out that the average over a sphere is zero when there is no mass *inside* it. Incidentally, it turns out that there is a relation between the various components of the curvature and the *variation* of the average curvature from place to place. So if you know the average curvature everywhere, you can figure out the details of the curvature components at each place. The average curvature inside the earth varies with altitude, and this means that some curvature components are nonzero both inside the earth and outside. It is that curvature that we see as a gravitational force.

Suppose we have a bug on a plane, and suppose that the "plane" has little pimples in the surface. Wherever there is a pimple the bug would conclude that his space has little local regions of curvature. We have the same thing in three dimensions. Wherever there is a lump of matter, our three-dimensional space has a local curvature—a kind of three-dimensional pimple.

If we make a lot of bumps on a plane there might be an overall curvature besides all the pimples—the surface might become like a ball. It would be interesting to know whether our space has a net average curvature as well as the local pimples due to the lumps of matter like the earth and the sun. The astrophysicists have been trying to answer that question by making measurements of galaxies at very large distances. For example, if the number of galaxies we see in a spherical shell at a large distance is different from what

* Approximately, because the density is not independent of radius as we are assuming.

we would expect from our knowledge of the radius of the shell, we would have a measure of the excess radius of a tremendously large sphere. From such measurements it is hoped to find out whether our whole universe is flat on the average, or round—whether it is "closed," like a sphere, or "open," like a plane. You may have heard about the debates that are going on about this subject. There are debates because the astronomical measurements are still completely inconclusive; the experimental data are not precise enough to give a definite answer. Unfortunately, we don't have the slightest idea about the overall curvature of our universe on a large scale.

6-4 Geometry in space-time

Now we have to talk about time. As you know from the special theory of relativity, measurements of space and measurements of time are interrelated. And it would be kind of crazy to have something happening to the space, without the time being involved in the same thing. You will remember that the measurement of time depends on the speed at which you move. For instance, if we watch a guy going by in a spaceship we see that things happen more slowly for him than for us. Let's say he takes off on a trip and returns in 100 seconds flat *by our watches*; his watch might say that he had been gone for only 95 seconds. In comparison with ours, his watch—and all other processes, like his heartbeat—have been running slow.

Now let's consider an interesting problem. Suppose you are the one in the spaceship. We ask you to start off at a given signal and return to your starting place just in time to catch a later signal—at, say, exactly 100 seconds later according to *our* clock. And you are also asked to make the trip in such a way that *your* watch will show the *longest possible* elapsed time. How should you move? You should stand still. If you move at all your watch will read less than 100 seconds when you get back.

Suppose, however, we change the problem a little. Suppose we ask you to start at point *A* on a given signal and go to point *B* (both fixed relative to us), and to do it in such a way that you arrive back just at the time of a second signal (say 100 seconds later according to our fixed clock). Again you are asked to make the trip in the way that lets you arrive with the latest possible reading on your watch. How would you do it? For which path and schedule will *your* watch show the greatest elapsed time when you arrive? The answer is that you will spend the longest time from *your* point of view if you make the trip by going at a uniform speed along a straight line. Reason: Any extra motions and any extra-high speeds will make your clock go slower. (Since the time deviations depend on the *square* of the velocity, what you lose by going extra fast at one place you can never make up by going extra slowly in another place.)

The point of all this is that we can use the idea to define "a straight line" in space-time. The analog of a straight line in space is for space-time a *motion* at uniform velocity in a constant direction.

The curve of shortest distance in space corresponds in space-time not to the path of shortest time, but to the one of *longest* time, because of the funny things that happen to signs of the *t*-terms in relativity. "Straight-line" motion—the analog of "uniform velocity along a straight line"—is then that motion which takes a watch from one place at one time to another place at another time in the way that gives the longest time reading for the watch. This will be our definition for the analog of a straight line in space-time.

6-5 Gravity and the principle of equivalence

Now we are ready to discuss the laws of gravitation. Einstein was trying to generate a theory of gravitation that would fit with the relativity theory that he had developed earlier. He was struggling along until he latched on to one important principle which guided him into getting the correct laws. That principle is based on the idea that when a thing is falling freely everything inside it seems

weightless. For example, a satellite in orbit is falling freely in the earth's gravity, and an astronaut in it feels weightless. This idea, when stated with greater precision, is called *Einstein's principle of equivalence*. It depends on the fact that all objects fall with exactly the same acceleration no matter what their mass, or what they are made of. If we have a spaceship that is "coasting"—so it's in a free fall—and there is a man inside, then the laws governing the fall of the man and the ship are the same. So if he puts himself in the middle of the ship he will stay there. He doesn't fall *with respect to the ship*. That's what we mean when we say he is "weightless."

Now suppose you are in a rocket ship which is accelerating. Accelerating with respect to what? Let's just say that its engines are on and generating a thrust so that it is not coasting in a free fall. Also imagine that you are way out in empty space so that there are practically no gravitational forces on the ship. If the ship is accelerating with "1g" you will be able to stand on the "floor" and will feel your normal weight. Also if you let go of a ball, it will "fall" toward the floor. Why? Because the ship is accelerating "upward," but the ball has no forces on it, so it will not accelerate; it will get left behind. Inside the ship the ball will appear to have a downward acceleration of "1g."

Now let's compare that with the situation in a spaceship sitting at rest on the surface of the earth. *Everything is the same!* You would be pressed toward the floor, a ball would fall with an acceleration of 1g, and so on. In fact, how could you tell inside a spaceship whether you are sitting on the earth or are accelerating in free space? According to Einstein's equivalence principle there is no way to tell if you only make measurements of what happens to things inside!

To be strictly correct, that is true only for one point inside the ship. The gravitational field of the earth is not precisely uniform, so a freely falling ball has a slightly different acceleration at different places—the direction changes and the magnitude changes. But if we imagine a strictly uniform gravitational field, it is completely imitated in every respect by a system with a constant acceleration. That is the basis of the principle of equivalence.

6-6 The speed of clocks in a gravitational field

Now we want to use the principle of equivalence for figuring out a strange thing that happens in a gravitational field. We'll show you something that happens in a rocket ship which you probably wouldn't have expected to happen in a gravitational field. Suppose we put a clock at the "head" of the rocket ship—that is, at the "front" end—and we put another identical clock at the "tail," as in Figure 6-16. Let's call the two clocks *A* and *B*. If we compare these two clocks when the ship is accelerating, the clock at the head seems

Figure 6-16 An accelerating rocket ship with two clocks.

to run fast relative to the one at the tail. To see that, imagine that the front clock emits a flash of light each second, and that you are sitting at the tail comparing the arrival of the light flashes with the ticks of clock *B*. Let's say that the rocket is in the position *a* of Figure 6-17 when clock *A* emits a flash, and at the position *b* when the flash arrives at clock *B*. Later on, the ship will be at position *c* when the clock *A* emits its next flash, and at position *d* when you see it arrive at clock *B*.

The first flash travels the distance L_1 and the second flash travels the shorter distance L_2. It is a shorter distance because the ship is accelerating and has a higher speed at the time of the second flash. You can see, then, that if the two flashes were emitted from clock *A* one second apart, they would arrive at clock *B* with a separation somewhat less than one second, since the second flash doesn't spend as much time on the way. The same thing will also happen for all the later flashes. So if you were sitting in the tail you would conclude that clock *A* was running faster than clock *B*. If you were to do the same thing in reverse—letting clock *B* emit light and observing it at clock *A*—you would conclude that *B* was running

Figure 6-17 A clock at the head of an accelerating rocket ship appears to run faster than a clock at the tail.

slower than *A*. Everything fits together and there is nothing mysterious about it all.

But now let's think of the rocket ship at rest in the earth's gravity. *The same thing happens*. If you sit on the floor with one clock and watch another one which is sitting on a high shelf, it will appear to run faster than the one on the floor! You say, "But that is wrong. The times should be the same. With no acceleration there's no reason for the clocks to appear to be out of step." But they must if the principle of equivalence is right. And Einstein insisted that the principle *was* right, and went courageously and correctly ahead. He proposed that clocks at different places in a gravitational field must appear to run at different speeds. But if one always *appears* to be running at a different speed with respect to the other, then so far as the first is concerned the other *is* running at a different rate.

But now you see we have the analog for clocks of the hot ruler we were talking about earlier, when we had the bug on a hot plate. We imagined that rulers and bugs and everything changed lengths in the same way at various temperatures so they could never tell that their measuring sticks were changing as they moved around on the hot plate. It's the same with clocks in a gravitational field. Every clock we put at a higher level is seen to go faster. Heartbeats go faster, all processes run faster.

If they didn't you would be able to tell the difference between a gravitational field and an accelerating reference system. The idea that time can vary from place to place is a difficult one, but it is the idea Einstein used, and it is correct—believe it or not.

Using the principle of equivalence we can figure out how much the speed of a clock changes with height in a gravitational field. We just work out the apparent discrepancy between the two clocks in the accelerating rocket ship. The easiest way to do this is to use the result we found in Chapter 34 of Vol. I* for the Doppler effect. There, we found—see Eq. (34.14)*—that if v is the *relative* velocity

*of *The Feynman Lectures on Physics*.

of a source and a receiver, the *received* frequency ω is related to *the emitted* frequency ω_0 by

$$\omega = \omega_0 \frac{1 + v/c}{\sqrt{1 - v^2/c^2}}. \tag{6.4}$$

Now if we think of the accelerating rocket ship in Figure 6-17 the emitter and receiver are moving with equal velocities at any one instant. But in the time that it takes the light signals to go from clock A to clock B the ship has accelerated. It has, in fact, picked up the additional velocity gt, where g is the acceleration and t is time it takes light to travel the distance H from A to B. This time is very nearly H/c. So when the signals arrive at B, the ship has increased its velocity by gH/c. The receiver always has this velocity *with respect to the emitter* at the instant the signal left it. So this is the velocity we should use in the Doppler shift formula, Eq. (6.4). Assuming that the acceleration and the length of the ship are small enough that this velocity is much smaller than c, we can neglect the term in v^2/c^2. We have that

$$\omega = \omega_0 \left(1 + \frac{gH}{c^2}\right). \tag{6.5}$$

So for the two clocks in the spaceship we have the relation

$$\text{(Rate at the receiver)} = \text{(Rate of emission)} \left(1 + \frac{gH}{c^2}\right), \tag{6.6}$$

where H is the height of the emitter *above* the receiver.

From the equivalence principle the same result must hold for two clocks separated by the height H in a gravitational field with the free fall acceleration g.

This is such an important idea we would like to demonstrate that it also follows from another law of physics—from the conservation of energy. We know that the gravitational force on an object is proportional to its mass M, which is related to its total internal energy E by $M = E/c^2$. For instance, the masses of nuclei determined from the *energies* of nuclear reactions which transmute one nucleus into another agree with the masses obtained from atomic *weights*.

Now think of an atom which has a lowest energy state of total energy E_0 and a higher energy state E_1, and which can go from the state E_1 to the state E_0 by emitting light. The frequency ω of the light will be given by

$$\hbar\omega = E_1 - E_0. \tag{6.7}$$

Now suppose we have such an atom in the state E_1 sitting on the floor, and we carry it from the floor to the height H. To do that we must do some work in carrying the mass $m_1 = E_1/c^2$ up against the gravitational force. The amount of work done is

$$\frac{E_1}{c^2} gH. \tag{6.8}$$

Then we let the atom emit a photon and go into the lower energy state E_0. Afterward we carry the atom back to the floor. On the return trip the mass is E_0/c^2; we get back the energy

$$\frac{E_0}{c^2} gH, \tag{6.9}$$

so we have done a net amount of work equal to

$$\Delta U = \frac{E_1 - E_0}{c^2} gH. \tag{6.10}$$

When the atom emitted the photon it gave up the energy $E_1 - E_0$. Now suppose that the photon happened to go down to the floor and be absorbed. How much energy would it deliver there? You might at first think that it would deliver just the energy $E_1 - E_0$. But that can't be right if energy is conserved, as you can see from the following argument. We started with the energy E_1 at the floor. When we finish, the energy at the floor level is the energy E_0 of the atom in its lower state plus the energy E_{ph} received from the photon. In the meantime we have had to supply the additional energy ΔU of Eq. (6.10). If energy is conserved, the energy we end up with at the floor must be greater than we started with by just the work we have done. Namely, we must have that

$$E_{\text{ph}} + E_0 = E_1 + \Delta U,$$

or (6.11)

$$E_{\text{ph}} = (E_1 - E_0) + \Delta U.$$

It must be that the photon does *not* arrive at the floor with just the energy $E_1 - E_0$ it started with, but with a *little more energy*. Otherwise some energy would have been lost. If we substitute in Eq. (6.11) the ΔU we got in Eq. (6.10), we get that the photon arrives at the floor with the energy

$$E_{\text{ph}} = (E_1 - E_0)\left(1 + \frac{gH}{c^2}\right). \tag{6.12}$$

But a photon of energy E_{ph} has the frequency $\omega = E_{\text{ph}}/\hbar$. Calling the frequency of the *emitted* photon ω_0—which is by Eq. (6.7) equal to $(E_1 - E_0)/\hbar$—our result in Eq. (6.12) gives again the relation of (6.5) between the frequency of the photon when it is absorbed on the floor and the frequency with which it was emitted.

The same result can be obtained in still another way. A photon of frequency ω_0 has the energy $E_0 = \hbar\omega_0$. Since the energy E_0 has the gravitational mass E_0/c^2 the photon has a mass (*not* rest mass) $\hbar\omega_0/c^2$, and is "attracted" by the earth. In falling the distance H it will gain an additional energy $(\hbar\omega_0/c^2)gH$, so it arrives with the energy

$$E = \hbar\omega_0\left(1 + \frac{gH}{c^2}\right).$$

But its frequency after the fall is $E/\hbar$, giving again the result in Eq. (6.5). Our ideas about relativity, quantum physics, and energy conservation all fit together only if Einstein's predictions about clocks in a gravitational field are right. The frequency changes we are talking about are normally very small. For instance, for an altitude difference of 20 meters at the earth's surface the frequency difference is only about two parts in 10^{15}. However, just such a change has recently been found experimentally using the Mössbauer effect.* Einstein was perfectly correct.

* R. V. Pound and G. A. Rebka, Jr., *Physical Review Letters,* vol. 4, p. 337 (1960).

6-7 The curvature of space-time

Now we want to relate what we have just been talking about to the idea of curved space-time. We have already pointed out that if the time goes at different rates in different places, it is analogous to the curved space of the hot plate. But it is more than an analogy; it means that space-time *is* curved. Let's try to do some geometry in space-time. That may at first sound peculiar, but we have often made diagrams of space-time with distance plotted along one axis and time along the other. Suppose we try to make a rectangle in space-time. We begin by plotting a graph of height H versus t as in Figure 6-18(a). To make the base of our rectangle we take an object which is *at rest* at the height H_1 and follow its world line for 100 seconds. We get the line BD in part (b) of the figure which is parallel to the t-axis. Now let's take another object which is 100 feet above the first one at $t = 0$. It starts at the point A in Figure 6-18(c). Now we follow its world line for 100 seconds as measured by a clock at A. The object goes from A to C, as shown in part (d) of the figure. But notice that since time goes at a different rate at the two heights—we are assuming that there is a gravitational field—the two points C and D are not simultaneous. If we try to complete the square by drawing a line to the point C' which is 100 feet above D at the same time, as in Figure 6-18(e), the pieces don't fit. And that's what we mean when we say that space-time is curved.

6-8 Motion in curved space-time

Let's consider an interesting little puzzle. We have two identical clocks, A and B, sitting together on the surface of the earth as in Figure 6-19. Now we lift clock A to some height H, hold it there awhile, and return it to the ground so that it arrives at just the instant when clock B has advanced by 100 seconds. Then clock A will read something like 107 seconds, because it was running faster when it was up in the air. Now here is the puzzle. How should we

Figure 6-18 Trying to make a rectangle in space-time.

move clock A so that it reads the latest possible time—always assuming that it returns when B reads 100 seconds? You say, "That's easy. Just take A as high as you can. Then it will run as fast as possible, and be the latest when you return." Wrong. You forgot something—we've only got 100 seconds to go up and back. If we go very high, we have to go very fast to get there and back in 100 seconds. And you mustn't forget the effect of special relativity which causes moving clocks to *slow down* by the factor $\sqrt{1 - v^2/c^2}$. This relativity effect works in the direction of making clock A read *less*

Figure 6-19 In a uniform gravitational field the trajectory with the maximum proper time for a fixed elapsed time is a parabola.

time than clock *B*. You see that we have a kind of game. If we stand still with clock *A* we get 100 seconds. If we go up slowly to a small height and come down slowly we can get a little more than 100 seconds. If we go a little higher, maybe we can gain a little more. But if we go too high we have to move fast to get there, and we may slow down the clock enough that we end up with less than 100 seconds. What program of height versus time—how high to go and with what speed to get there, carefully adjusted to bring us back to clock *B* when it has increased by 100 seconds—will give us the largest possible time reading on clock *A*?

Answer: Find out how fast you have to throw a ball up into the air so that it will fall back to earth in exactly 100 seconds. The ball's motion—rising fast, slowing down, stopping, and coming back down—is exactly the right motion to make the time the maximum on a wristwatch strapped to the ball.

Now consider a slightly different game. We have two points *A* and *B* both on the earth's surface at some distance from one another. We play the same game that we did earlier to find what we call the straight line. We ask how we should go from *A* to *B* so that the time on our moving watch will be the longest—assuming we start at *A* on a given signal and arrive at *B* on another signal at *B* which we will say is 100 seconds later by a fixed clock. Now you say, "Well, we found out before that the thing to do is to coast along

a straight line at a uniform speed chosen so that we arrive at B exactly 100 seconds later. If we don't go along a straight line it takes more speed, and our watch is slowed down." But wait! That was before we took gravity into account. Isn't it better to curve upward a little bit and then come down? Then during part of the time we are higher up and our watch will run a little faster? It is, indeed. If you solve the mathematical problem of adjusting the curve of the motion so that the elapsed time of the moving watch is the most it can possibly be, you will find that the motion is a parabola—the same curve followed by something that moves on a free ballistic path in the gravitational field, as in Figure 6-19. Therefore the law of motion in a gravitational field can also be stated: *An object always moves from one place to another so that a clock carried on it gives a longer time than it would on any other possible trajectory*—with, of course, the same starting and finishing conditions. The time measured by a moving clock is often called its "proper time." In free fall, the trajectory makes the proper time of an object a maximum.

Let's see how this all works out. We begin with Eq. (6.5) which says that the *excess* rate of the moving watch is

$$\frac{\omega_0 gH}{c^2}. \tag{6.13}$$

Besides this, we have to remember that there is a correction of the opposite sign for the speed. For this effect we know that

$$\omega = \omega_0 \sqrt{1 - v^2/c^2}\,.$$

Although the principle is valid for any speed, we take an example in which the speeds are always much less than c. Then we can write this equation as

$$\omega = \omega_0(1 - v^2/2c^2),$$

and the defect in the rate of our clock is

$$-\omega_0 \frac{v^2}{2c^2}. \tag{6.14}$$

Combining the two terms in (6.13) and (6.14) we have that

$$\Delta\omega = \frac{\omega_0}{c^2}\left(gH - \frac{v^2}{2}\right). \tag{6.15}$$

Such a frequency shift of our moving clock means that if we measure a time dt on a fixed clock, the moving clock will register the time

$$dt\left[1 + \left(\frac{gH}{c^2} - \frac{v^2}{2c^2}\right)\right]. \tag{6.16}$$

The total time excess over the trajectory is the integral of the extra term with respect to time, namely

$$\frac{1}{c^2}\int\left(gH - \frac{v^2}{2}\right)dt, \tag{6.17}$$

which is supposed to be a maximum.

The term gH is just the gravitational potential ϕ. Suppose we multiply the whole thing by a constant factor $-mc^2$, where m is the mass of the object. The constant won't change the condition for the maximum, but the minus sign will just change the maximum to a minimum. Equation (6.16) then says that the object will move so that

$$\int\left(\frac{mv^2}{2} - m\phi\right)dt = \text{a minimum}. \tag{6.18}$$

But now the integrand is just the difference of the kinetic and potential energies. And if you look in Chapter 19 of Volume II* you will see that when we discussed the principle of least action we showed that Newton's laws for an object in any potential could be written exactly in the form of Eq. (6.18).

6-9 Einstein's theory of gravitation

Einstein's form of the equations of motion—that the proper time should be a maximum in curved space-time—gives the same results

* of *The Feynman Lectures on Physics*.

as Newton's laws for low velocities. As he was circling around the earth, Gordon Cooper's watch was reading later than it would have in any other path you could have imagined for his satellite.*

So the law of gravitation can be stated in terms of the ideas of the geometry of space-time in this remarkable way. The particles always take the longest proper time—in space-time a quantity analogous to the "shortest distance." That's the law of motion in a gravitational field. The great advantage of putting it this way is that the law doesn't depend on any coordinates, or any other way of defining the situation.

Now let's summarize what we have done. We have given you two laws for gravity:

(1) How the geometry of space-time changes when matter is present—namely, that the curvature expressed in terms of the excess radius is proportional to the mass inside a sphere, Eq. (6.3).
(2) How objects move if there are only gravitational forces—namely, that objects move so that their proper time between two end conditions is a maximum.

Those two laws correspond to similar pairs of laws we have seen earlier. We originally described motion in a gravitational field in terms of Newton's inverse square law of gravitation and his laws of motion. Now laws (1) and (2) take their places. Our new pair of laws also correspond to what we have seen in electrodynamics. There we had our law—the set of Maxwell's equations—which determines the fields produced by charges. It tells how the character of "space" is changed by the presence of charged matter, which is what law (1) does for gravity. In addition, we had a law about how

* Strictly speaking it is only a *local* maximum. We should have said that the proper time is larger than for any *nearby* path. For example, the proper time on an elliptical orbit around the earth need not be longer than on a ballistic path of an object which is shot to a great height and falls back down.

particles move in the given fields—$d(m\mathbf{v})/dt = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$. This, for gravity, is done by law (2).

In the laws (1) and (2) you have a precise statement of Einstein's theory of gravitation—although you will usually find it stated in a more complicated mathematical form. We should, however, make one further addition. Just as time scales change from place to place in a gravitational field, so do the length scales. Rulers change lengths as you move around. It is impossible with space and time so intimately mixed to have something happen with time that isn't in some way reflected in space. Take even the simplest example: You are riding past the earth. What is *"time"* from *your* point of view is partly space from *our* point of view. So there must also be changes in space. It is the entire *space-time* which is distorted by the presence of matter, and this is more complicated than a change only in time scale. However, the rule that we gave in Eq. (6-3) is enough to determine completely all the laws of gravitation, provided that it is understood that this rule about the curvature of space applies not only from one man's point of view but is true for everybody. Somebody riding by a mass of material sees a different mass content because of the kinetic energy he calculates for its motion past him, and he must include the mass corresponding to that energy. The theory must be arranged so that everybody—no matter how he moves—will, when he draws a sphere, find that the excess radius is $G/3c^2$ times the total mass (or, better, $G/3c^4$ times the total energy content) inside the sphere. That this law—law (1)—should be true in any moving system is one of the great laws of gravitation, called *Einstein's field equation*. The other great law is (2)—that things must move so that the proper time is a maximum—and is called *Einstein's equation of motion*.

To write these laws in a complete algebraic form, to compare them with Newton's laws, or to relate them to electrodynamics is difficult mathematically. But it is the way our most complete laws of the physics of gravity look today.

Although they gave a result in agreement with Newton's mechanics for the simple example we considered, they do not always

do so. The three discrepancies first derived by Einstein have been experimentally confirmed: the orbit of Mercury is not a fixed ellipse; starlight passing near the sun is deflected twice as much as you would think; and the rates of clocks depend on their location in a gravitational field. Whenever the predictions of Einstein have been found to differ from the ideas of Newtonian mechanics, Nature has chosen Einstein's.

Let's summarize everything that we have said in the following way. First, time and distance rates depend on the place in space you measure them and on the time. This is equivalent to the statement that space-time is curved. From the measured area of a sphere we can define a predicted radius, $\sqrt{A/4\pi}$, but the actual measured radius will have an excess over this which is proportional (the constant is G/c^2) to the total mass contained inside the sphere. This fixes the exact degree of the curvature of space-time. And the curvature must be the same no matter who is looking at the matter or how it is moving. Second, particles move on "straight lines" (trajectories of maximum proper time) in this curved space-time. This is the content of Einstein's formulation of the laws of gravitation.

1

The Pleasure of Finding Things Out

This is the edited transcript of an interview with Feynman made for the BBC television program Horizon *in 1981, shown in the United States as an episode of* Nova. *Feynman had most of his life behind him by this time (he died in 1988), so he could reflect on his experiences and accomplishments with the perspective not often attainable by a younger person. The result is a candid, relaxed, and very personal discussion on many topics close to Feynman's heart: why knowing merely the name of something is the same as not knowing anything at all about it; how he and his fellow atomic scientists of the Manhattan Project could drink and revel in the success of the terrible weapon they had created while on the other side of the world in Hiroshima thousands of their fellow human beings were dead or dying from it; and why Feynman could just as well have gotten along without a Nobel Prize.*

2

◆

The Pleasure of Finding Things Out

The Beauty of a Flower

I have a friend who's an artist and he's sometimes taken a view which I don't agree with very well. He'll hold up a flower and say, "Look how beautiful it is," and I'll agree, I think. And he says—"you see, I as an artist can see how beautiful this is, but you as a scientist, oh, take this all apart and it becomes a dull thing." And I think that he's kind of nutty. First of all, the beauty that he sees is available to other people and to me, too, I believe, although I might not be quite as refined aesthetically as he is; but I can appreciate the beauty of a flower. At the same time I see much more about the flower than he sees. I can imagine the cells in there, the complicated actions inside which also have a beauty. I mean it's not just beauty at this dimension of one centimeter, there is also beauty at a smaller dimension, the inner structure. Also the processes, the fact that the colors in the flower evolved in order to attract insects to pollinate it is interesting—it means that insects can see the color. It adds a question: Does this aesthetic sense also exist in the lower forms? Why is it aesthetic? All kinds of interesting questions which shows that a science knowledge only adds to the excitement and mystery and the awe of a flower. It only adds; I don't understand how it subtracts.

Avoiding Humanities

I've always been very one-sided about science and when I was younger I concentrated almost all my effort on it. I didn't have time to learn and I didn't have much patience with what's called the humanities, even though in the university there were humanities that you had to take. I tried my best to avoid somehow learning anything and working at it. It was

only afterwards, when I got older, that I got more relaxed, that I've spread out a little bit. I've learned to draw and I read a little bit, but I'm really still a very one-sided person and I don't know a great deal. I have a limited intelligence and I use it in a particular direction.

Tyrannosaurus in the Window

We had the *Encyclopaedia Britannica* at home and even when I was a small boy [my father] used to sit me on his lap and read to me from the *Encyclopaedia Britannica*, and we would read, say, about dinosaurs and maybe it would be talking about the brontosaurus or something, or the tyrannosaurus rex, and it would say something like, "This thing is twenty-five feet high and the head is six feet across," you see, and so he'd stop all this and say, "Let's see what that means. That would mean that if he stood in our front yard he would be high enough to put his head through the window but not quite because the head is a little bit too wide and it would break the window as it came by."

Everything we'd read would be translated as best we could into some reality and so I learned to do that—everything that I read I try to figure out what it really means, what it's really saying by translating and so (LAUGHS) I used to read the *Encyclopaedia* when I was a boy but with translation, you see, so it was very exciting and interesting to think there were animals of such magnitude—I wasn't frightened that there would be one coming in my window as a consequence of this, I don't think, but I thought that it was very, very interesting, that they all died out and at that time nobody knew why.

We used to go to the Catskill Mountains. We lived in New York and the Catskill Mountains was the place where people went in the summer; and the fathers—there was a big group of

people there but the fathers would all go back to New York to work during the week and only come back on the weekends. When my father came he would take me for walks in the woods and tell me various interesting things that were going on in the woods–which I'll explain in a minute–but the other mothers seeing this, of course, thought this was wonderful and that the other fathers should take their sons for walks, and they tried to work on them but they didn't get anywhere at first and they wanted my father to take all the kids, but he didn't want to because he had a special relationship with me–we had a personal thing together–so it ended up that the other fathers had to take their children for walks the next weekend, and the next Monday when they were all back to work, all the kids were playing in the field and one kid said to me, "See that bird, what kind of a bird is that?" And I said, "I haven't the slightest idea what kind of a bird it is." He says, "It's a brown throated thrush," or something, "Your father doesn't tell you anything." But it was the opposite: my father *had* taught me. Looking at a bird he says, "Do you know what that bird is? It's a brown throated thrush; but in Portuguese it's a . . . in Italian a . . . ," he says "in Chinese it's a . . . , in Japanese a . . . ," etcetera. "Now," he says, "you know in all the languages you want to know what the name of that bird is and when you've finished with all that," he says, "you'll know absolutely nothing whatever about the bird. You only know about humans in different places and what they call the bird. Now," he says, "let's look at the bird."

He had taught me to notice things and one day when I was playing with what we call an express wagon, which is a little wagon which has a railing around it for children to play with that they can pull around. It had a ball in it–I remember this–it had a ball in it, and I pulled the wagon and I noticed something about the way the ball moved, so I went to my fa-

ther and I said, "Say, Pop, I noticed something: When I pull the wagon the ball rolls to the back of the wagon, and when I'm pulling it along and I suddenly stop, the ball rolls to the front of the wagon," and I says, "why is that?" And he said, "That nobody knows," he said. "The general principle is that things that are moving try to keep on moving and things that are standing still tend to stand still unless you push on them hard." And he says, "This tendency is called inertia but nobody knows why it's true." Now that's a deep understanding–he doesn't give me a name, he knew the difference between knowing the name of something and knowing something, which I learnt very early. He went on to say, "If you look close you'll find the ball does not rush to the back of the wagon, but it's the back of the wagon that you're pulling against the ball; that the ball stands still or as a matter of fact from the friction starts to move forward really and doesn't move back." So I ran back to the little wagon and set the ball up again and pulled the wagon from under it and looking sideways and seeing indeed he was right–the ball never moved backwards in the wagon when I pulled the wagon forward. It moved backward relative to the wagon, but relative to the sidewalk it was moved forward a little bit, it's just [that] the wagon caught up with it. So that's the way I was educated by my father, with those kinds of examples and discussions, no pressure, just lovely interesting discussions.

Algebra for the Practical Man

My cousin, at that time, who was three years older, was in high school and was having considerable difficulty with his algebra and had a tutor come, and I was allowed to sit in a corner while (LAUGHS) the tutor would try to teach my cousin algebra, problems like $2x$ plus something. I said to my

cousin then, "What're you trying to do?" You know, I hear him talking about x. He says, "What do *you* know–$2x + 7$ is equal to 15," he says "and you're trying to find out what x is." I says, "You mean 4." He says, "Yeah, but you did it with arithmetic, you have to do it by algebra," and that's why my cousin was never able to do algebra, because he didn't understand how he was supposed to do it. There was no way. I learnt algebra fortunately by not going to school and knowing the whole idea was to find out what x was and it didn't make any difference how you did it–there's no such thing as, you know, you do it by arithmetic, you do it by algebra–that was a false thing that they had invented in school so that the children who have to study algebra can all pass it. They had invented a set of rules which if you followed them without thinking could produce the answer: subtract 7 from both sides, if you have a multiplier divide both sides by the multiplier and so on, and a series of steps by which you could get the answer if you didn't understand what you were trying to do.

There was a series of math books, which started *Arithmetic for the Practical Man,* and then *Algebra for the Practical Man,* and then *Trigonometry for the Practical Man*, and I learned trigonometry for the practical man from that. I soon forgot it again because I didn't understand it very well but the series was coming out, and the library was going to get *Calculus for the Practical Man* and I knew by this time by reading the *Encyclopaedia* that calculus was an important subject and it was an interesting one and I ought to learn it. I was older now, I was perhaps thirteen; and then the calculus book finally came out and I was so excited and I went to the library to take it out and she looks at me and she says, "Oh, you're just a child, what are you taking this book out for, this book is a [book for adults]." So this was one of the few times in my life I was un-

comfortable and I lied and I said it was for my father, he selected it. So I took it home and I learnt calculus from it and I tried to explain it to my father and he'd start to read the beginning of it and he found it confusing and it really bothered me a little bit. I didn't know that he was so limited, you know, that he didn't understand, and I thought it was relatively simple and straightforward and he didn't understand it. So that was the first time I knew I had learnt more in some sense than he.

Epaulettes and the Pope

One of the things that my father taught me besides physics (LAUGHS), whether it's correct or not, was a disrespect for respectable . . . for certain kinds of things. For example, when I was a little boy, and a rotogravure–that's printed pictures in newspapers–first came out in the *New York Times,* he used to sit me again on his knee and he'd open a picture, and there was a picture of the Pope and everybody bowing in front of him. And he'd say, "Now look at these humans. Here is one human standing here, and all these others are bowing. Now what is the difference? This one is the Pope"–he hated the Pope anyway–and he'd say, "the difference is epaulettes"–of course not in the case of the Pope, but if he was a general–it was always the uniform, the position, "but this man has the same human problems, he eats dinner like anybody else, he goes to the bathroom, he has the same kind of problems as everybody, he's a human being. Why are they all bowing to him? Only because of his name and his position, because of his uniform, not because of something special he did, or his honor, or something like that." He, by the way, was in the uniform business, so he knew what the difference was be-

tween the man with the uniform off and the uniform on; it's the same man for him.

He was happy with me, I believe. Once, though, when I came back from MIT–I'd been there a few years–he said to me, "Now," he said, "you've become educated about these things and there's one question I've always had that I've never understood very well and I'd like to ask you, now that you've studied this, to explain it to me," and I asked him what it was. And he said that he understood that when an atom made a transition from one state to another it emits a particle of light called a photon. I said, "That's right." And he says, "Well, now, is the photon in the atom ahead of time that it comes out, or is there no photon in it to start with?" I says, "There's no photon in, it's just that when the electron makes a transition it comes" and he says "Well, where does it come from then, how does it come out?" So I couldn't just say, "The view is that photon numbers aren't conserved, they're just created by the motion of the electron." I couldn't try to explain to him something like: the sound that I'm making now wasn't in me. It's not like my little boy who when he started to talk, suddenly said that he could no longer say a certain word–the word was "cat"–because his word bag has run out of the word cat (LAUGHS). So there's no word bag that you have inside so that you use up the words as they come out, you just make them as they go along, and in the same sense there was no photon bag in an atom and when the photons come out they didn't come from somewhere, but I couldn't do much better. He was not satisfied with me in the respect that I never was able to explain any of the things that he didn't understand (LAUGHS). So he was unsuccessful, he sent me through all these universities in order to find out these things and he never did find out (LAUGHS).

Invitation to the Bomb

[While working on his PhD thesis, Feynman was asked to join the project to develop the atomic bomb.] It was a completely different kind of a thing. It would mean that I would have to stop the research in what I was doing, which is my life's desire, to take time off to do this, which I felt I should do in order to protect civilization. Okay? So that was what I had to debate with myself. My first reaction was, well, I didn't want to get interrupted in my normal work to do this odd job. There was also the problem, of course, of any moral thing involving war. I wouldn't have much to do with that, but it kinda scared me when I realized what the weapon would be, and that since it might be possible, it must be possible. There was nothing that I knew that indicated that if we could do it they couldn't do it, and therefore it was very important to try to cooperate.

[In early 1943 Feynman joined Oppenheimer's team at Los Alamos.] With regard to moral questions, I do have something I would like to say about it. The original reason to start the project, which was that the Germans were a danger, started me off on a process of action which was to try to develop this first system at Princeton and then at Los Alamos, to try to make the bomb work. All kinds of attempts were made to redesign it to make it a worse bomb and so on. It was a project on which we all worked very, very hard, all co-operating together. And with any project like that you continue to work trying to get success, having decided to do it. But what I did—immorally I would say—was to not remember the reason that I said I was doing it, so that when the reason changed, because Germany was defeated, not the singlest thought came to my mind at all about that, that that meant

now that I have to reconsider why I am continuing to do this. I simply didn't think, okay?

Success and Suffering

[On 6 August 1945 the atomic bomb was exploded over Hiroshima.] The only reaction that I remember–perhaps I was blinded by my own reaction–was a very considerable elation and excitement, and there were parties and people got drunk and it would make a tremendously interesting contrast, what was going on in Los Alamos at the same time as what was going on in Hiroshima. I was involved with this happy thing and also drinking and drunk and playing drums sitting on the hood of–the bonnet of–a Jeep and playing drums with excitement running all over Los Alamos at the same time as people were dying and struggling in Hiroshima.

I had a very strong reaction after the war of a peculiar nature–it may be from just the bomb itself and it may be for some other psychological reasons, I'd just lost my wife or something, but I remember being in New York with my mother in a restaurant, immediately after [Hiroshima], and thinking about New York, and I knew how big the bomb in Hiroshima was, how big an area it covered and so on, and I realized from where we were–I don't know, 59th Street–that to drop one on 34th Street, it would spread all the way out here and all these people would be killed and all the things would be killed and there wasn't only one bomb available, but it was easy to continue to make them, and therefore that things were sort of doomed because already it appeared to me–very early, earlier than to others who were more optimistic–that international relations and the way people were behaving were no different than they had ever been before and that it was just going to go on the same way as any other thing and I was sure that

it was going, therefore, to be used very soon. So I felt very uncomfortable and thought, really believed, that it was silly: I would see people building a bridge and I would say "they don't understand." I really believed that it was senseless to make anything because it would all be destroyed very soon anyway, but they didn't understand that and I had this very strange view of any construction that I would see, I would always think how foolish they are to try to make something. So I was really in a kind of depressive condition.

"I Don't Have to Be Good Because They Think I'm Going to Be Good."

[After the war Feynman joined Hans Bethe at Cornell University. He turned down the offer of a job at Princeton's Institute for Advanced Study.]* They [must have] expected me to be wonderful to offer me a job like this and I wasn't wonderful, and therefore I realized a new principle, which was that I'm not responsible for what other people think I am able to do; I don't have to be good because they think I'm going to be good. And somehow or other I could relax about this, and I thought to myself, I haven't done anything important and I'm never going to do anything important. But I used to enjoy physics and mathematical things and because I used to play with them it was in very short order [that I] worked the things out for which I later won the Nobel Prize.†

*(1906–) Winner of the 1967 Nobel Prize in Physics for contributions to the theory of nuclear reactions, especially for his discoveries concerning the energy production in stars. *Ed.*

†In 1965, the Nobel Prize for Physics was shared by Richard Feynman, Julian Schwinger, and Sin-Itiro Tomonaga for their fundamental work in quantum electrodynamics, and its deep consequences for the physics of elementary particles. *Ed.*

The Nobel Prize—Was It Worth It?

[Feynman was awarded a Nobel Prize for his work on quantum electrodynamics.] What I essentially did–and also it was done independently by two other people, [Sinitiro] Tomanaga in Japan and [Julian] Schwinger–was to figure out how to control, how to analyze and discuss the original quantum theory of electricity and magnetism that had been written in 1928; how to interpret it so as to avoid the infinities, to make calculations for which there were sensible results which have since turned out to be in exact agreement with every experiment which has been done so far, so that quantum electrodynamics fits experiment in every detail where it's applicable–not involving the nuclear forces, for instance–and it was the work that I did in 1947 to figure out how to do that, for which I won the Nobel Prize.

[BBC: *Was it worth the Nobel Prize?]* As a (LAUGHS) . . . I don't know anything about the Nobel Prize, I don't understand what it's all about or what's worth what, but if the people in the Swedish Academy decide that x, y, or z wins the Nobel Prize then so be it. I won't have anything to do with the Nobel Prize . . . it's a pain in the . . . (LAUGHS). I don't like honors. I appreciate it for the work that I did, and for people who appreciate it, and I know there's a lot of physicists who use my work, I don't need anything else, I don't think there's any sense to anything else. I don't see that it makes any point that someone in the Swedish Academy decides that this work is noble enough to receive a prize–I've already got the prize. The prize is the pleasure of finding the thing out, the kick in the discovery, the observation that other people use it [my work]–those are the real things, the honors are unreal to me. I don't believe in honors, it bothers me, honors bother, honors is epaulettes, honors is uni-

forms. My papa brought me up this way. I can't stand it, it hurts me.

When I was in high school, one of the first honors I got was to be a member of the Arista, which is a group of kids who got good grades–eh?–and everybody wanted to be a member of the Arista, and when I got into the Arista I discovered that what they did in their meetings was to sit around to discuss who else was worthy to join this wonderful group that we are–okay? So we sat around trying to decide who it was who would get to be allowed into this Arista. This kind of thing bothers me psychologically for one or another reason I don't understand myself–honors–and from that day to this [it] always bothered me. When I became a member of the National Academy of Sciences, I had ultimately to resign because that was another organization most of whose time was spent in choosing who was illustrious enough to join, to be allowed to join us in our organization, including such questions as [should] we physicists stick together because they've a very good chemist that they're trying to get in and we haven't got enough room for so-and-so. What's the matter with chemists? The whole thing was rotten because its purpose was mostly to decide who could have this honor–okay? I don't like honors.

The Rules of the Game

[From 1950 to 1988 Feynman was Professor of Theoretical Physics at the California Institute of Technology.] One way, that's kind of a fun analogy in trying to get some idea of what we're doing in trying to understand nature, is to imagine that the gods are playing some great game like chess, let's say, and you don't know the rules of the game, but you're allowed to look

at the board, at least from time to time, in a little corner, perhaps, and from these observations you try to figure out what the rules of the game are, what the rules of the pieces moving are. You might discover after a bit, for example, that when there's only one bishop around on the board that the bishop maintains its color. Later on you might discover the law for the bishop as it moves on the diagonal which would explain the law that you understood before–that it maintained its color– and that would be analagous to discovering one law and then later finding a deeper understanding of it. Then things can happen, everything's going good, you've got all the laws, it looks very good, and then all of a sudden some strange phenomenon occurs in some corner, so you begin to investigate that–it's castling, something you didn't expect. We're always, by the way, in fundamental physics, always trying to investigate those things in which we don't understand the conclusions. After we've checked them enough, we're okay.

The thing that doesn't fit is the thing that's the most interesting, the part that doesn't go according to what you expected. Also, we could have revolutions in physics: after you've noticed that the bishops maintain their color and they go along the diagonal and so on for such a long time and everybody knows that that's true, then you suddenly discover one day in some chess game that the bishop doesn't maintain its color, it changes its color. Only later do you discover a new possibility, that a bishop is captured and that a pawn went all the way down to the queen's end to produce a new bishop–that can happen but you didn't know it, and so it's very analagous to the way our laws are: They sometimes look positive, they keep on working and all of a sudden some little gimmick shows that they're wrong and then we have to investigate the conditions under which this bishop change of

color happened and so forth, and gradually learn the new rule that explains it more deeply. Unlike the chess game, though, in [which] the rules become more complicated as you go along, in physics, when you discover new things, it looks more simple. It appears on the whole to be more complicated because we learn about a greater experience–that is, we learn about more particles and new things–and so the laws look complicated again. But if you realize all the time what's kind of wonderful–that is, if we expand our experience into wilder and wilder regions of experience–every once in a while we have these integrations when everything's pulled together into a unification, in which it turns out to be simpler than it looked before.

If you are interested in the ultimate character of the physical world, or the complete world, and at the present time our only way to understand that is through a mathematical type of reasoning, then I don't think a person can fully appreciate, or in fact can appreciate much of, these particular aspects of the world, the great depth of character of the universality of the laws, the relationships of things, without an understanding of mathematics. I don't know any other way to do it, we don't know any other way to describe it accurately . . . or to see the interrelationships without it. So I don't think a person who hasn't developed some mathematical sense is capable of fully appreciating this aspect of the world–don't misunderstand me, there are many, many aspects of the world that mathematics is unnecessary for, such as love, which are very delightful and wonderful to appreciate and to feel awed and mysterious about; and I don't mean to say that the only thing in the world is physics, but you were talking about physics and if that's what you're talking about, then to not know mathematics is a severe limitation in understanding the world.

Smashing Atoms

Well, what I'm working on in physics right now is a special problem which we've come up against and I'll describe what it is. You know that everything's made out of atoms, we've got that far already and most people know that already, and that the atom has a nucleus with electrons going around. The behavior of the electrons on the outside is now completely [known], the laws for it are well understood as far as we can tell in this quantum electrodynamics that I told you about. And after that was evolved, then the problem was how does the nucleus work, how do the particles interact, how do they hold together? One of the by-products was to discover fission and to make the bomb. But investigating the forces that hold the nuclear particles together was a long task. At first it was thought that it was an exchange of some sort of particles inside, which were invented by Yukawa, called pions, and it was predicted that if you hit protons—the proton is one of the particles of the nucleus—against a nucleus, they would knock out such pions, and sure enough, such particles came out.

Not only pions came out but other particles, and we began to run out of names—kaons and sigmas and lamdas and so on; they're all called hadrons now—and as we increased the energy of the reaction and got more and more different kinds, until there were hundreds of different kinds of particles; then the problem, of course—this period is 1940 up to 1950, towards the present—was to find the pattern behind it. There seemed to be many many interesting relations and patterns among the particles, until a theory was evolved to explain these patterns, that all of these particles were really made of something else, that they were made of things called quarks—three quarks, for example, would form a proton—and that the proton is one of the particles of the nucleus; another one is a

neutron. The quarks came in a number of varieties–in fact, at first only three were needed to explain all the hundreds of particles and the different kinds of quarks–they are called u-type, d-type, s-type. Two us and a d made a proton, two ds and a u made a neutron. If they were moving in a different way inside they were some other particle. Then the problem came: What exactly is the behavior of the quarks and what holds them together? And a theory was thought of which is very simple, a very close analogy to quantum electrodynamics–not exactly the same but very close–in which the quarks are like the electron and the particles called gluons–which go between the electrons, which makes them attract each other electrically–are like the photons. The mathematics was very similar but there are a few terms slightly different. The difference in the form of the equations that were guessed at were guessed by principles of such beauty and simplicity that it isn't arbitrary, it's very, very determined. What is arbitrary is how many different kinds of quark there are, but not the character of the force between them.

Now unlike electrodynamics, in which two electrons can be pulled apart as far as you want, in fact when they are very far away the force is weakened; if this were true for quarks you would have expected that when you hit things together hard enough the quarks would have come out. But instead of that, when you're doing an experiment with enough energy that quarks could come out, instead of that you find a big jet–that is, all particles going about in the same direction as the old hadrons, no quarks–and from the theory, it was clear that what was required was that when the quark comes out, it kind of makes these new pairs of quarks and they come in little groups and make hadrons.

The question is, why is it so different in electrodynamics, how do these small-term differences, these little terms that are

different in the equation, produce such different effects, entirely different effects? In fact, it was very surprising to most people that this would really come out, that first you would think that the theory was wrong, but the more it's studied the clearer it became that it's very possible that these extra terms would produce these effects. Now we were in a position that's different in history than any other time in physics, that's always different. We have a theory, a complete and definite theory of all of these hadrons, and we have an enormous number of experiments and lots and lots of details, so why can't we test the theory right away to find out whether it's right or wrong? Because what we have to do is calculate the consequences of the theory. If this theory is right, what should happen, and has that happened? Well, this time the difficulty is in the first step. If the theory is right, what should happen is very hard to figure out. The mathematics needed to figure out what the consequences of this theory are have turned out to be, at the present time, insuperably difficult. At the present time–all right? And therefore it's obvious what my problem is–my problem is to try to develop a way of getting numbers out of this theory, to test it really carefully, not just qualitatively, to see if it might give the right result.

I spent a few years trying to invent mathematical things that would permit me to solve the equations, but I didn't get anywhere, and then I decided that in order to do that I must first understand more or less how the answer probably looks. It's hard to explain this very well, but I had to get a qualitative idea of how the phenomenon works before I could get a good quantitative idea. In other words, people didn't even understand roughly how it worked, and so I have been working most recently in the last year or two on understanding roughly how it works, not quantitatively yet, with the hope that in the future that rough understanding can be refined

into a precise mathematical tool, way, or algorithm to get from the theory to the particles. You see, we're in a funny position: It's not that we're looking for the theory, we've got the theory–a good, good candidate–but we're in the step in the science that we need to compare the theory to experiment by seeing what the consequences are and checking it. We're stuck in seeing what the consequences are, and it's my aim, it's my desire to see if I can work out a way to work out what the consequences of this theory are (LAUGHS). It's a kind of a crazy position to be in, to have a theory that you can't work out the consequences of . . . I can't stand it, I have to figure it out. Someday, maybe.

"Let George Do It."

To do high, real good physics work you do need absolutely solid lengths of time, so that when you're putting ideas together which are vague and hard to remember, it's very much like building a house of cards and each of the cards is shaky, and if you forget one of them the whole thing collapses again. You don't know how you got there and you have to build them up again, and if you're interrupted and kind of forget half the idea of how the cards went together–your cards being different-type parts of the ideas, ideas of different kinds that have to go together to build up the idea–the main point is, you put the stuff together, it's quite a tower and it's easy [for it] to slip, it needs a lot of concentration–that is, solid time to think–and if you've got a job in administrating anything like that, then you don't have the solid time. So I have invented another myth for myself–that I'm irresponsible. I tell everybody, I don't do anything. If anybody asks me to be on a committee to take care of admissions, no, I'm irresponsible, I don't give a

damn about the students–of course I give a damn about the students but I know that somebody else'll do it–and I take the view, "Let George do it," a view which you're not supposed to take, okay, because that's not right to do, but I do that because I like to do physics and I want to see if I can still do it, and so I'm selfish, okay? I want to do my physics.

Bored by the History

All those students are in the class: Now you ask me how should I best teach them? Should I teach them from the point of view of the history of science, from the applications? My theory is that the best way to teach is to have no philosophy, [it] is to be chaotic and [to] confuse it in the sense that you use every possible way of doing it. That's the only way I can see to answer it, so as to catch this guy or that guy on different hooks as you go along, [so] that during the time when the fellow who's interested in history's being bored by the abstract mathematics, on the other hand the fellow who likes the abstractions is being bored another time by the history–if you can do it so you don't bore them all, all the time, perhaps you're better off. I really don't know how to do it. I don't know how to answer this question of different kinds of minds with different kinds of interests–what hooks them on, what makes them interested, how you direct them to become interested. One way is by a kind of force, you have to pass this course, you have to take this examination. It's a very effective way. Many people go through schools that way and it may be a more effective way. I'm

sorry, after many, many years of trying to teach and trying all different kinds of methods, I really don't know how to do it.

Like Father, Like Son

I got a kick, when I was a boy, [out] of my father telling me things, so I tried to tell my son things that were interesting about the world. When he was very small we used to rock him to bed, you know, and tell him stories, and I'd make up a story about little people that were about so high [who] would walk along and they would go on picnics and so on and they lived in the ventilator; and they'd go through these woods which had great big long tall blue things like trees, but without leaves and only one stalk, and they had to walk between them and so on; and he'd gradually catch on [that] that was the rug, the nap of the rug, the blue rug, and he loved this game because I would describe all these things from an odd point of view and he liked to hear the stories and we got all kinds of wonderful things–he even went to a moist cave where the wind kept going in and out–it was coming in cool and went out warm and so on. It was inside the dog's nose that they went, and then of course I could tell him all about physiology by this way and so on. He loved that and so I told him lots of stuff, and I enjoyed it because I was telling him stuff that I liked, and we had fun when he would guess what it was and so on. And then I have a daughter and I tried the same thing–well, my daughter's personality was different, she didn't want to hear this story, she wanted the story that was in the book repeated again, and reread to her. She wanted me to read to her, not to make up stories, and it's a different personality. And so if I were to say a very good method for teaching children about science is to make up these stories of the

little people, it doesn't work at all on my daughter–it happened to work on my son–okay?

"Science Which Is Not a Science . . . "

Because of the success of science, there is, I think, a kind of pseudoscience. Social science is an example of a science which is not a science; they don't do [things] scientifically, they follow the forms–or you gather data, you do so-and-so and so forth but they don't get any laws, they haven't found out anything. They haven't got anywhere yet–maybe someday they will, but it's not very well developed, but what happens is on an even more mundane level. We get experts on everything that sound like they're sort of scientific experts. They're not scientific, they sit at a typewriter and they make up something like, oh, food grown with, er, fertilizer that's organic is better for you than food grown with fertilizer that's inorganic–may be true, may not be true, but it hasn't been demonstrated one way or the other. But they'll sit there on the typewriter and make up all this stuff as if it's science and then become an expert on foods, organic foods and so on. There's all kinds of myths and pseudoscience all over the place.

I may be quite wrong, maybe they do know all these things, but I don't think I'm wrong. You see, I have the advantage of having found out how hard it is to get to really know something, how careful you have to be about checking the experiments, how easy it is to make mistakes and fool yourself. I know what it means to know something, and therefore I see how they get their information and I can't believe that they know it, they haven't done the work necessary, haven't done the checks necessary, haven't done the care necessary. I have a great suspicion that they don't know, that this stuff is

[wrong] and they're intimidating people. I think so. I don't know the world very well but that's what I think.

Doubt and Uncertainty

If you expected science to give all the answers to the wonderful questions about what we are, where we're going, what the meaning of the universe is and so on, then I think you could easily become disillusioned and then look for some mystic answer to these problems. How a scientist can take a mystic answer I don't know because the whole spirit is to understand–well, never mind that. Anyhow, I don't understand that, but anyhow if you think of it, the way I think of what we're doing is we're exploring, we're trying to find out as much as we can about the world. People say to me, "Are you looking for the ultimate laws of physics?" No, I'm not, I'm just looking to find out more about the world and if it turns out there is a simple ultimate law which explains everything, so be it, that would be very nice to discover.

If it turns out it's like an onion with millions of layers and we're just sick and tired of looking at the layers, then that's the way it is, but whatever way it comes out its nature is there and she's going to come out the way she is, and therefore when we go to investigate it we shouldn't predecide what it is we're trying to do except to try to find out more about it. If you say your problem is, why do you find out more about it, if you thought you were trying to find out more about it because you're going to get an answer to some deep philosophical question, you may be wrong. It may be that you can't get an answer to that particular question by finding out more about the character of nature, but I don't look at it [like that]. My interest in science is to simply find out about the world, and the more I find out the better it is, like, to find out.

The Pleasure of Finding Things Out

There are very remarkable mysteries about the fact that we're able to do so many more things than apparently animals can do, and other questions like that, but those are mysteries I want to investigate without knowing the answer to them, and so altogether I can't believe these special stories that have been made up about our relationship to the universe at large because they seem to be too simple, too connected, too local, too provincial. The earth, He came to the earth, one of the aspects of God came to the earth, mind you, and look at what's out there. It isn't in proportion. Anyway, it's no use arguing, I can't argue it, I'm just trying to tell you why the scientific views that I have do have some effect on my belief. And also another thing has to do with the question of how you find out if something's true, and if all the different religions have all different theories about the thing, then you begin to wonder. Once you start doubting, just like you're supposed to doubt, you ask me if the science is true. You say no, we don't know what's true, we're trying to find out and everything is possibly wrong.

Start out understanding religion by saying everything is possibly wrong. Let us see. As soon as you do that, you start sliding down an edge which is hard to recover from and so on. With the scientific view, or my father's view, that we should look to see what's true and what may be or may not be true, once you start doubting, which I think to me is a very fundamental part of my soul, to doubt and to ask, and when you doubt and ask it gets a little harder to believe.

You see, one thing is, I can live with doubt and uncertainty and not knowing. I think it's much more interesting to live not knowing than to have answers which might be wrong. I have approximate answers and possible beliefs and different degrees of certainty about different things, but I'm not absolutely sure of anything and there are many things I don't

know anything about, such as whether it means anything to ask why we're here, and what the question might mean. I might think about it a little bit and if I can't figure it out, then I go on to something else, but I don't have to know an answer, I don't feel frightened by not knowing things, by being lost in a mysterious universe without having any purpose, which is the way it really is so far as I can tell. It doesn't frighten me.

2

Computing Machines in the Future

Forty years to the day after the atomic bombing of Nagasaki, Manhattan Project veteran Feynman delivers a talk in Japan, but the topic is a peaceful one, one that still occupies our sharpest minds: the future of the computing machine, including the topic that made Feynman seem a Nostradamus of computer science—the ultimate lower limit to the size of a computer. This chapter may be challenging for some readers; however, it is such an important part of Feynman's contribution to science that I hope they will take the time to read it, even if they have to skip over some of the more technical spots. It ends with a brief discussion of one of Feynman's favorite pet ideas, which launched the current revolution in nanotechnology.

Introduction

It's a great pleasure and an honor to be here as a speaker in memorial for a scientist that I have respected and admired as much as Professor Nishina. To come to Japan and talk about computers is like giving a sermon to Buddha. But I have been thinking about computers and this is the only subject I could think of when invited to talk.

◆

The Pleasure of Finding Things Out

The first thing I would like to say is what I am not going to talk about. I want to talk about the future of computing machines. But the most important possible developments in the future are things that I will not speak about. For example, there is a great deal of work to try to develop smarter machines, machines which have a better relationship with humans so that input and output can be made with less effort than the complex programming that's necessary today. This often goes under the name of artificial intelligence, but I don't like that name. Perhaps the unintelligent machines can do even better than the intelligent ones.

Another problem is the standardization of programming languages. There are too many languages today, and it would be a good idea to choose just one. (I hesitate to mention that in Japan, for what will happen will be that there will simply be more standard languages—you already have four ways of writing now, and attempts to standardize anything here result apparently in more standards and not fewer!)

Another interesting future problem that is worth working on but I will not talk about is automatic debugging programs. Debugging means fixing errors in a program or in a machine, and it is surprisingly difficult to debug programs as they get more complicated.

Another direction of improvement is to make physical machines three dimensional instead of all on a surface of a chip. That can be done in stages instead of all at once—you can have several layers and then add many more layers as time goes on. Another important device would be one that could automatically detect defective elements on a chip; then the chip would automatically rewire itself so as to avoid the defective elements. At the present time, when we try to make big chips there are often flaws or bad spots in the chips, and we throw the whole chip away. If we could make it so that we

could use the part of the chip that was effective, it would be much more efficient. I mention these things to try to tell you that I am aware of what the real problems are for future machines. But what I want to talk about is simple, just some small technical, physically good things that can be done in principle according to the physical laws. In other words, I would like to discuss the machinery and not the way we use the machines.

I will talk about some technical possibilities for making machines. There will be three topics. One is parallel processing machines, which is something of the very near future, almost present, that is being developed now. Further in the future is the question of the energy consumption of machines, which seems at the moment to be a limitation, but really isn't. Finally I will talk about the size. It is always better to make the machines smaller, and the question is, how much smaller is it still possible, in principle, to make machines according to the laws of Nature? I will not discuss which and what of these things will actually appear in the future. That depends on economic problems and social problems and I am not going to try to guess at those.

Parallel Computers

The first topic concerns parallel computers. Almost all the present computers, conventional computers, work on a layout or an architecture invented by von Neumann,* in which there is a very large memory that stores all the information, and one central location that does simple calculations. We take a number from this place in the memory and a number

*John von Neumann (1903–1957), a Hungarian-American mathematician who is credited as being one of the fathers of the computer. *Ed.*

from that place in the memory, send the two to the central arithmetical unit to add them, and then send the answer to some other place in the memory. There is, therefore, effectively one central processor which is working very, very fast and very hard, while the whole memory sits out there like a fast filing cabinet of cards which are very rarely used. It is obvious that if there were more processors working at the same time we ought to be able to do calculations faster. But the problem is that someone who might be using one processor may be using some information from the memory that another one needs, and it gets very confusing. For such reasons it has been said that it is very difficult to get many processors to work in parallel.

Some steps in that direction have been taken in the larger conventional machines called "vector processors." When sometimes you want to do exactly the same step on many different items, you can perhaps do that at the same time. The hope is that regular programs can be written in the ordinary way, and then an interpreter program will discover automatically when it is useful to use this vector possibility. That idea is used in the Cray and in "supercomputers" in Japan. Another plan is to take what is effectively a large number of relatively simple (but not very simple) computers, and connect them all together in some pattern. Then they can all work on a part of the problem. Each one is really an independent computer, and they will transfer information to each other as one or another needs it. This kind of a scheme is realized in the Caltech Cosmic Cube, for example, and represents only one of many possibilities. Many people are now making such machines. Another plan is to distribute very large numbers of very simple central processors all over the memory. Each one deals with just a small part of the memory and there is an elaborate system of interconnections between them. An ex-

ample of such a machine is the Connection Machine made at MIT. It has 64,000 processors and a system of routing in which every 16 can talk to any other 16 and thus has 4,000 routing connection possibilities.

It would appear that scientific problems such as the propagation of waves in some material might be very easily handled by parallel processing. This is because what happens in any given part of space at any moment can be worked out locally and only the pressures and the stresses from the neighboring volumes need to be known. These can be worked out at the same time for each volume and these boundary conditions communicated across the different volumes. That's why this type of design works for such problems. It has turned out that a very large number of problems of all kinds can be dealt with in parallel. As long as the problem is big enough so that a lot of calculating has to be done, it turns out that a parallel computation can speed up time to solution enormously, and this principle applies not just to scientific problems.

What happened to the prejudice of two years ago, which was that the parallel programming is difficult? It turns out that what was difficult, and almost impossible, is to take an ordinary program and automatically figure out how to use the parallel computation effectively on that program. Instead, one must start all over again with the problem, appreciating that we have the possibility of parallel calculation, and rewrite the program completely with a new [understanding of] what is inside the machine. It is not possible to effectively use the old programs. They must be rewritten. That is a great disadvantage to most industrial applications and has met with considerable resistance. But the big programs usually belong to scientists or other, unofficial, intelligent programmers who love computer science and are willing to start all over

again and rewrite the program if they can make it more efficient. So what's going to happen is that the hard programs, vast big ones, will be the first to be re-programmed by experts in the new way, and then gradually everybody will have to come around, and more and more programs will be programmed that way, and programmers will just have to learn how to do it.

Reducing the Energy Loss

The second topic I want to talk about is energy loss in computers. The fact that they must be cooled is an apparent limitation for the largest computers—a good deal of effort is spent in cooling the machine. I would like to explain that this is simply a result of very poor engineering and is nothing fundamental at all. Inside the computer a bit of information is controlled by a wire which has a voltage of either one value or another value. It is called "one bit," and we have to change the voltage of the wire from one value to the other and put charge on or take charge off. I make an analogy with water: We have to fill a vessel with water to get one level or empty it to get to the other level. This is just an analogy—if you like electricity better you can think more accurately electrically. What we do now is analogous, in the water case, to filling the vessel by pouring water in from a top level (Fig. 1), and lowering the level by opening the valve at the bottom and letting it all run out. In both cases there is a loss of energy because of the sudden drop in level of the water, through a height from the top level where it comes in, to the low bottom level, and also when you start pouring water in to fill it up again. In the cases of voltage and charge, the same thing occurs.

It's like, as Mr. Bennett has explained, operating an automobile which has to start by turning on the engine and stop

FIGURE 1

by putting on the brakes. By turning on the engine and then putting on the brakes, each time you lose power. Another way to arrange things for a car would be to connect the wheels to flywheels. Now when the car stops, the flywheel speeds up, thus saving the energy—it can then be reconnected to start the car again. The water analog of this would be to have a U-shaped tube with a valve in the center at the bottom, connecting the two arms of the U (Fig. 2). We start with it full on the right but empty on the left with the valve closed. If we now open the valve, the water will slip over to the other side, and we can close the valve again, just in time to catch the water in the left arm. Now when we want to go the other way, we open the valve again and the water slips back to the other side and we catch it again. There is some loss and the water doesn't climb as high as it did before, but all we have to do is to put a little water in to correct the loss—a much smaller energy loss than the direct fill method. This trick uses the inertia of the water and the analog for electricity is inductance. However, it is very difficult with the silicon transistors that we use today to make up inductance on the

FIGURE 2

chips. So this technique is not particularly practical with present technology.

Another way would be to fill the tank by a supply which stays only a little bit above the level of the water, lifting the water supply in time as we fill up the tank (Fig. 3), so that the dropping of water is always small during the entire effort. In the same way, we could use an outlet to lower the level in the tank, but just take water off near the top and lower the tube so that the heat loss would not appear at the position of the transistor, or would be small. The actual amount of loss will depend on how high the distance is between the supply and the surface as we fill it up. This method corresponds to changing the voltage supply with time. So if we could use a time varying voltage supply, we could use this method. Of course, there is energy loss in the voltage supply, but that is all located in one place and there it is simple to make one big inductance. This scheme is called "hot clocking," because the voltage supply operates at the same time as the clock which times everything. In addition, we don't need an extra clock signal to time the circuits as we do in conventional designs.

VARIABLE VOLTAGE SUPPLY
("HOT CLOCKING")
ENERGY LOSS · TIME = CONSTANT.

FIGURE 3

Both of these last two devices use less energy if they go slower. If I try to move the water supply level too fast, the water in the tube doesn't keep up with it and there ends being a big drop in water level. So to make the device work I must go slowly. Similarly, the U-tube scheme will not work unless that central valve can open and close faster than the time it takes for the water in the U-tube to slip back and forth. So my devices must be slower–I've saved an energy loss but I've made the devices slower. In fact the energy loss multiplied by the time it takes for the circuit to operate is constant. But nevertheless, this turns out to be very practical because the clock time is usually much larger than the circuit time for the transistors, and we can use that to decrease the energy. Also if we went, let us say, three times slower with our calculations, we could use one-third the energy over three times the time, which is nine times less power that has to be dissipated. Maybe this is worth it. Maybe by redesigning using parallel computations or other devices, we can spend a little longer than we could do at maximum circuit speed, in order to make

ENERGY · TIME FOR TRANSISTOR

$$= kT \cdot \frac{\text{LENGTH}}{\text{THERMAL VELOCITY}} \cdot \frac{\text{LENGTH}}{\text{MEAN FREE PATH}} \cdot \text{NUMBER OF ELECTRONS}$$

ENERGY $\sim 10^{9-11}$ kT

$\therefore$ DECREASE SIZE : FASTER
LESS ENERGY

FIGURE 4

a larger machine that is practical and from which we could still reduce the energy loss.

For a transistor, the energy loss multiplied by the time it takes to operate is a product of several factors (Fig. 4):

1. the thermal energy proportional to temperature, kT;
2. the length of the transistor between source and drain, divided by the velocity of the electrons inside (the thermal velocity $\sqrt{3kT/m}$);
3. the length of the transistor in units of the mean free path for collisions of electrons in the transistor;
4. the total number of the electrons that are inside the transistor when it operates.

Putting in appropriate values for all of these numbers tells us that the energy used in transistors today is somewhere between a billion to ten billion or more times the thermal energy kT. When the transistor switches, we use that much energy. This is very large amount of energy. It is obviously a

good idea to decrease the size of the transistor. We decrease the length between source and drain and we can decrease the number of the electrons, and so use much less energy. It also turns out that a smaller transistor is much faster, because the electrons can cross it faster and make their decisions to switch faster. For every reason, it is a good idea to make the transistor smaller, and everybody is always trying to do that.

But suppose we come to a circumstance in which the mean free path is longer than the size of the transistor; then we discover that the transistor doesn't work properly anymore. It does not behave the way we expected. This reminds me, years ago there was something called the sound barrier. Airplanes were supposed not to be able to go faster than the speed of sound because, if you designed them normally and then tried to put the speed of sound in the equations, the propeller wouldn't work and the wings don't lift and nothing works correctly. Nevertheless, airplanes can go faster than the speed of sound. You just have to know what the right laws are under the right circumstances, and design the device with the correct laws. You cannot expect old designs to work in new circumstances. But *new* designs can work in *new* circumstances, and I assert that it is perfectly possible to make transistor systems, or, more correctly, switching systems and computing devices, in which the dimensions are smaller than the mean free path. I speak, of course, "in principle," and I am not speaking about the actual manufacture of such devices. Let us therefore discuss what happens if we try to make the devices as small as possible.

Reducing the Size

So my third topic is the size of computing elements and now I speak entirely theoretically. The first thing that you would

BROWNIAN MOTION

2 VOLT = 80 kT

PROB. ERROR $e^{-80} = 10^{-43}$

10^9 TRANSISTORS

10^{10} CHANGES / SEC. EACH

10^9 SECONDS (30 YEARS)

10^{28}

FIGURE 5

worry about when things get very small is Brownian motion*—everything is shaking about and nothing stays in place. How can you control the circuits then? Furthermore, if a circuit does work, doesn't it now have a chance of accidentally jumping back? If we use two volts for the energy of this electric system, which is what we ordinarily use (Fig. 5), that is eighty times the thermal energy at room temperature ($kT =$ 1/40 volt) and the chance that something jumps backward against 80 times thermal energy is e, the base of the natural logarithm, to the power minus eighty, or 10^{-43}. What does that mean? If we had a billion transistors in a computer (which we don't yet have), all of them switching 10^{10} times a second (a switching time of a tenth of a nanosecond), switching perpetually, operating for 10^9 seconds, which is 30 years, the total number of switching operations in such a machine is 10^{28}. The chance of one of the transistors going backward

*The jerky movements of particles caused by the constant random collisions of molecules, first noted in print in 1928 by botanist Robert Brown, and explained by Albert Einstein in a 1905 paper in *Annalen der Physik*. Ed.

is only 10^{-43}, so there will be no error produced by thermal oscillations whatsoever in 30 years. If you don't like that, use 2.5 volts and then the probability gets even smaller. Long before that, real failures will come when a cosmic ray accidentally goes through the transistor, and we don't have to be more perfect than that.

However, much more is in fact possible and I would like to refer you to an article in a most recent *Scientific American* by C. H. Bennett and R. Landauer, "The Fundamental Physical Limits of Computation."* It is possible to make a computer in which each element, each transistor, can go forward and accidentally reverse and still the computer will operate. All the operations in the computer can go forward or backward. The computation proceeds for a while one way and then it undoes itself, "uncalculates," and then goes forward again and so on. If we just pull it along a little, we can make this computer go through and finish the calculation by making it just a little bit more likely that it goes forward than backward.

It is known that all possible computations can be made by putting together some simple elements like transistors; or, if we want to be more logically abstract, something called a NAND gate, for example (NAND means NOT-AND). A NAND gate has two "wires" in and one out (Fig. 6). Forget the NOT for the moment. What is an AND gate? An AND gate is a device whose output is 1 only if both input wires are 1, otherwise its output is 0. NOT-AND means the opposite, thus the output wire reads 1 (i.e., has the voltage level corresponding to 1) unless both input wires read 1; if both input wires read 1, then the output wire reads 0 (i.e., has the voltage level corresponding to 0). Figure 6 shows a little table of inputs and outputs for such a NAND gate. *A* and *B* are in-

**Sci. Am.* July 1985; Japanese Transl.—SAIENSU, Sept. 1985. *Ed.*

A	B	C'
0	0	1
0	1	1
1	0	1
1	1	0

NOT REVERSIBLE

INFORMATION LOST

FIGURE 6

puts and C is the output. If A and B are both 1, the output is 0, otherwise 1. But such a device is irreversible: Information is lost. If I only know the output, I cannot recover the input. The device can't be expected to flip forward and then come back and compute correctly anymore. For instance, if we know that the output is now 1, we don't know whether it came from A=0, B=1 or A=1, B=0 or A=0, B=0 and it cannot go back. Such a device is an irreversible gate. The great discovery of Bennett and, independently, of Fredkin is that it is possible to do computation with a different kind of fundamental gate unit, namely, a reversible gate unit. I have illustrated their idea–with a unit which I could call a reversible NAND gate. It has three inputs and three outputs (Fig. 7). Of the outputs, two, A' and B', are the same as two of the inputs, A and B, but the third input works this way. C' is the same as C unless A and B are both 1, in which case it changes whatever C is. For instance, if C is 1 it is changed to 0, if C is 0 it

FIGURE 7

is changed to 1—but these changes only happen if both A and B are 1. If you put two of these gates in succession, you see that A and B will go through, and if C is not changed in both it stays the same. If C is changed, it is changed twice so that it stays the same. So this gate can reverse itself and no information has been lost. It is possible to discover what went in if you know what came out.

A device made entirely with such gates will make calculations if everything moves forward. But if things go back and forth for a while, but then eventually go forward enough, it still operates correctly. If the things flip back and then go forward later it is still all right. It's very much the same as a particle in a gas which is bombarded by the atoms around it. Such a particle usually goes nowhere, but with just a little pull, a little prejudice that makes a chance to move one way a little higher than the other way, the thing will slowly drift forward and travel from one end to the other, in spite of the

Brownian motion that it has made. So our computer will compute provided we apply a drift force to pull the thing across the calculation. Although it is not doing the calculation in a smooth way, nevertheless, calculating like this, forward and backward, it eventually finishes the job. As with the particle in the gas, if we pull it very slightly, we lose very little energy, but it takes a long time to get to one side from the other. If we are in a hurry, and we pull hard, then we lose a lot of energy. It is the same with this computer. If we are patient and go slowly, we can make the computer operate with practically no energy loss, even less than kT per step, any amount as small as you like if you have enough time. But if you are in a hurry, you must dissipate energy, and again it's true that the energy lost to pull the calculation forward to complete it multiplied by the time you are allowed to make the calculation is a constant.

With these possibilities in mind, let's see how small we can make a computer. How big must a number be? We all know we can write numbers in base 2 as strings of "bits," each a one or a zero. And the next atom could be a one or a zero, so a little string of atoms are enough to hold a number, one atom for each bit. (Actually, since an atom can have more than just two states, we could use even fewer atoms, but one per bit is little enough!) So, for intellectual entertainment, we consider whether we could make a computer in which the writing of bits is of atomic size, in which a bit is, for example, whether the spin in the atom is up for 1 or down for 0. And then our "transistor," which changes the bits in different places, would correspond to some interaction between atoms which will change their states. The simplest example would be if a kind of 3-atom interaction were to be the fundamental element or gate in such a computer. But again, the device won't work right if we design it with the laws appropriate for large objects.

MUST NOW USE NEW LAWS OF PHYSICS

REVERSIBLE GATES
QUANTUM MECHANICS

NO FURTHER LIMITATIONS BESIDE
- CANNOT BE SMALLER THAN ATOM
- THERMO LOSS (BENNETT)
- SPEED OF LIGHT

FIGURE 8

We must use the new laws of physics, quantum mechanical laws, the laws that are appropriate to atomic motion (Fig. 8).

We therefore have to ask whether the principles of quantum mechanics permit an arrangement of atoms so small in number as a few times the number of gates in a computer that could operate as a computer. This has been studied in principle, and such an arrangement has been found. Since the laws of quantum mechanics are reversible, we must use the invention by Bennett and Fredkin of reversible logic gates. When this quantum mechanical situation is studied, it is found that quantum mechanics adds no further limitations to anything that Mr. Bennett has said from thermodynamic considerations. Of course there is a limitation, the practical limitation anyway, that the bits must be of the size of an atom and a transistor 3 or 4 atoms. The quantum mechanical gate I used has 3 atoms. (I would not try to write my bits onto nuclei, I'll wait till the technological development reaches atoms before I need to go any further!) That leaves us just with: (a) the limitations in size to the size of atoms; (b) the energy requirements depending on the time as worked out by Bennett; and c) the feature that I did not mention concerning the speed of light–we can't send the signals any faster than the speed of light. These are the only physical limitations on computers that I know of.

10^{-3} - 10^{-4} IN LINEAR DIMENSION
10^{-11} IN VOLUME
10^{-11} IN ENERGY
$10^{-4.5}$ IN TIME

} REDUCTIONS AVAILABLE PER GATE

THEORETICALLY POSSIBLE!

FIGURE 9

If we somehow manage to make an atomic size computer, it would mean (Fig. 9) that the dimension, the *linear* dimension, is a thousand to ten thousand times smaller than those very tiny chips that we have now. It means that the volume of the computer is 100 billionth or 10^{-11} of the present volume, because the volume of the "transistor" is smaller by a factor 10^{-11} than the transistors we make today. The energy requirement for a single switch is also about eleven orders of magnitude smaller than the energy required to switch the transistor today, and the time to make the transitions will be at least ten thousand times faster per step of calculation. So there is plenty of room for improvement in the computer and I leave this to you, practical people who work on computers, as an aim to get to. I underestimated how long it would take for Mr. Ezawa to translate what I said, and I have no more to say that I have prepared for today. Thank you! I will answer questions if you'd like.

Questions and Answers

Q: You mentioned that one bit of information can be stored in one atom, and I wonder if you can store the same amount of information in one quark.

A: Yes. But we don't have control of the quarks and that becomes a really impractical way to deal with things. You

might think that what I am talking about is impractical, but I don't believe so. When I am talking about atoms, I believe that someday we will be able to handle and control them individually. There would be so much energy involved in the quark interactions that they would be very dangerous to handle because of the radioactivity and so on. But the atomic energies that I am talking about are very familiar to us in chemical energies, electrical energies, and those are numbers that are within the realm of reality, I believe, however absurd it may seem at the moment.

Q: You said that the smaller the computing element is the better. But, I think equipment has to be larger, because. . .

A: You mean that your finger is too big to push the buttons? Is that what you mean?

Q: Yes, it is.

A: Of course, you are right. I am talking about internal computers perhaps for robots or other devices. The input and output is something that I didn't discuss, whether the input comes from looking at pictures, hearing voices, or buttons being pushed. I am discussing how the computation is done in principle and not what form the output should take. It is certainly true that the input and the output cannot be reduced in most cases effectively beyond human dimensions. It is already too difficult to push the buttons on some of the computers with our big fingers. But with elaborate computing problems that take hours and hours, they could be done very rapidly on the very small machines with low energy consumption. That's the kind of machine I was thinking of. Not the simple applications of adding two numbers but elaborate calculations.

Q: I would like to know your method to transform the information from one atomic scale element to another atomic scale element. If you will use a quantum mechanical or nat-

ural interaction between the two elements, then such a device will become very close to Nature itself. For example, if we make a computer simulation, a Monte Carlo simulation of a magnet to study critical phenomena, then your atomic scale computer will be very close to the magnet itself. What are your thoughts about that?

A: Yes. All things that we make are Nature. We arrange it in a way to suit our purpose, to make a calculation for a purpose. In a magnet there is some kind of relation, if you wish; there are some kinds of computations going on, just like there are in the solar system, in a way of thinking. But that might not be the calculation we want to make at the moment. What we need to make is a device for which we can change the programs and let it compute the problem that we want to solve, not just its own magnet problem that it likes to solve for itself. I can't use the solar system for a computer unless it just happens that the problem that someone gave me was to find the motion of the planets, in which case all I have to do is to watch. There was an amusing article written as a joke. Far in the future, the "article" appears discussing a new method of making aerodynamical calculations: Instead of using the elaborate computers of the day, the author invents a simple device to blow air past the wing. (He reinvents the wind tunnel!)

Q: I have recently read in a newspaper article that operations of the nerve system in a brain are much slower than present-day computers and the unit in the nerve system is much smaller. Do you think that the computers you have talked about today have something in common with the nerve system in the brain?

A: There is an analogy between the brain and the computer in that there are apparently elements that can switch under the control of others. Nerve impulses controlling or exciting other nerves, in a way that often depends upon whether more

than one impulse comes in–something like an AND or its generalization. What is the amount of energy used in the brain cell for one of these transitions? I don't know the number. The time it takes to make a switching in the brain is very much longer than it is in our computers even today, never mind the fancy business of some future atomic computer, but the brain's interconnection system is much more elaborate. Each nerve is connected to thousands of other nerves, whereas we connect transistors only to two or three others.

Some people look at the activity of the brain in action and see that in many respects it surpasses the computer of today, and in many other respects the computer surpasses ourselves. This inspires people to design machines that can do more. What often happens is that an engineer has an idea of how the brain works (in his opinion) and then designs a machine that behaves that way. This new machine may in fact work very well. But, I must warn you that that does not tell us anything about how the brain actually works, nor is it necessary to ever really know that, in order to make a computer very capable. It is not necessary to understand the way birds flap their wings and how the feathers are designed in order to make a flying machine. It is not necessary to understand the lever system in the legs of a cheetah–an animal that runs fast–in order to make an automobile with wheels that goes very fast. It is therefore not necessary to imitate the behavior of Nature in detail in order to engineer a device which can in many respects surpass Nature's abilities. It is an interesting subject and I like to talk about it.

Your brain is very weak compared to a computer. I will give you a series of numbers, one, three, seven... Or rather, ichi, san, shichi, san, ni, go, ni, go, ichi, hachi, ichi, ni, ku, san, go. Now I want you to repeat them back to me. A computer can take tens of thousands of numbers and give me them back in

reverse, or sum them or do lots of things that we cannot do. On the other hand, if I look at a face, in a glance I can tell you who it is if I know that person, or that I don't know that person. We do not yet know how to make a computer system so that if we give it a pattern of a face it can tell us such information, even if it has seen many faces and you have tried to teach it.

Another interesting example is chess playing machines. It is quite a surprise that we can make machines that play chess better than almost everybody in the room. But they do it by trying many, many possibilities. If he moves here, then I could move here, and he can move there, and so forth. They look at each alternative and choose the best. Computers look at millions of alternatives, but a master chess player, a human, does it differently. He recognizes patterns. He looks at only thirty or forty positions before deciding what move to make. Therefore, although the rules are simpler in Go, machines that play Go are not very good, because in each position there are too many possibilities to move and there are too many things to check and the machines cannot look deeply. Therefore the problem of recognizing patterns and what to do under these circumstances is the thing that the computer engineers (they like to call themselves computer scientists) still find very difficult. It is certainly one of the important things for future computers, perhaps more important than the things I spoke about. Make a machine to play Go effectively!

Q: I think that any method of computation would not be fruitful unless it would give a kind of provision on how to compose such devices or programs. I thought the Fredkin paper on conservative logic was very intriguing, but once I came to think of making a simple program using such devices I came to a halt because thinking out such a program is far more complex than the program itself. I think we could easily get into a kind of infinite regression because the

process of making out a certain program would be more complex than the program itself and in trying to automate the process, the automating program would be much more complex and so on, especially in this case where the program is hard-wired rather than being separated as a software. I think it is fundamental to think of the ways of composition.

A: We have some different experiences. There is no infinite regression: It stops at a certain level of complexity. The machine that Fredkin ultimately is talking about and the one that I was talking about in the quantum mechanical case are both universal computers in the sense that they can be programmed to do various jobs. This is not a hard-wired program. They are no more hard-wired than an ordinary computer that you can put information in–the program is a part of the input–and the machine does the problem that it is assigned to do. It is hard-wired but it is universal like an ordinary computer. These things are very uncertain but I found an algorithm. If you have a program written for an irreversible machine, the ordinary program, then I can convert it to a reversible machine program by a direct translation scheme, which is very inefficient and uses many more steps. Then, in real situations, the number of steps can be much less. But at least I know that I can take a program with $2n$ steps where it is irreversible, convert it to $3n$ steps of a reversible machine. That is many more steps. I did it very inefficiently since I did not try to find the minimum–just one way of doing it. I don't really think that we'll find this regression that you speak of, but you might be right. I am uncertain.

Q: Won't we be sacrificing many of the merits we were expecting of such devices, because those reversible machines run so slow? I am very pessimistic about this point.

A: They run slower, but they are very much smaller. I don't make it reversible unless I need to. There is no point in making the machine reversible unless you are trying very hard to decrease the energy enormously, rather ridiculously, because with only 80 times kT the irreversible machine functions perfectly. That 80 is much less than the present-day 10^9 or 10^{10} kT, so I have at least 10^7 improvement in energy to make, and can still do it with irreversible machines! That's true. That's the right way to go, for the present. I entertain myself intellectually for fun, to ask how far could we go in principle, not in practice, and then I discover that I can go to a fraction of a kT of energy and make the machines microscopic, atomically microscopic. But to do so, I must use the reversible physical laws. Irreversibility comes because the heat is spread over a large number of atoms and can't be gathered back again. When I make the machine very small, unless I allow a cooling element which is lots of atoms, I have to work reversibly. In practice there probably will never come a time when we will be unwilling to tie a little computer to a big piece of lead which contains 10^{10} atoms (which is still very small indeed), making it effectively irreversible. Therefore I agree with you that in practice, for a very long time and perhaps forever, we will use irreversible gates. On the other hand, it is a part of the adventure of science to try to find a limitation in all directions and to stretch the human imagination as far as possible everywhere. Although at every stage it has looked as if such an activity was absurd and useless, it often turns out at least not to be useless.

Q: Are there any limitations from the uncertainty principle? Are there any fundamental limitations on the energy and the clock time in your reversible machine scheme?

A: That was my exact point. There is no further limitation due to quantum mechanics. One must distinguish carefully between the energy lost or consumed irreversibly, the heat generated in the operation of the machine, and the energy content of the moving parts which might be extracted again. There is a relationship between the time and the energy which might be extracted again. But that energy which can be extracted again is not of any importance or concern. It would be like asking whether we should add the mc^2, the rest energy, of all the atoms which are in the device. I only speak of the energy lost times the time, and then there is no limitation. However, it is true that if you want to make a calculation at a certain extremely high speed, you have to supply to the machine parts which move fast and have energy, but that energy is not necessarily lost at each step of the calculation; it coasts through by inertia.

A (to no Q): Could I just say with regard to the question of useless ideas, I'd like to add one more. I waited, if you would ask me, but you didn't. So I will answer it anyway. How would we make a machine of such small dimensions that we have to put the atoms in special places? Today we have no machinery with moving parts whose dimension is extremely small, at the scale of atoms or hundreds of atoms even, but there is no physical limitation in that direction either. There is no reason why, when we lay down the silicon even today, the pieces cannot be made into little islands so that they are movable. We could also arrange small jets so we could squirt the different chemicals on certain locations. We can make machinery which is extremely small. Such machinery will be easy to control by the same kind of computer circuits that we make. Ultimately, for fun again and intellectual pleasure, we could imagine machines as tiny as a few microns

across, with wheels and cables all interconnected by wires, silicon connections, so that the thing as a whole, a very large device, moves not like the awkward motions of our present stiff machines but in the smooth way of the neck of a swan, which after all is a lot of little machines, the cells all interconnected and all controlled in a smooth way. Why can't we do that ourselves?

3

Los Alamos from Below

And now a little something on the lighter side—gems about wisecracker (not to mention safecracker) Feynman getting in and out of trouble at Los Alamos: getting his own private room by seeming to break the no-women-in-the-men's-dormitory rule; outwitting the camp's censors; rubbing shoulders with great men like Robert Oppenheimer, Niels Bohr, and Hans Bethe; and the awesome distinction of being the only man to stare straight at the first atomic blast without protective goggles, an experience that changed Feynman forever.

Professor Hirschfelder's flattering introduction is quite inappropriate for my talk, which is "Los Alamos from Below." What I mean from below is although in my field at the present time I'm a slightly famous man, at the time I was not anybody famous at all. I did not even have a degree when I started to work on my stuff associated with the Manhattan Project.* Many of the other people who tell you about Los Alamos

*The name given to the gargantuan project to build the first atomic bomb, which began in 1942 and culminated with the bombing of Hiroshima and Nagasaki on August 6 and 9, respectively, 1945. The project was spread over the United States, with units at, for example, the University of Chicago; Hanford, Washington; Oak Ridge, Tennessee; and Los Alamos, New Mexico, where the bombs were built, and which was essentially the headquarters of the whole project. *Ed.*

knew somebody up in some higher echelon of governmental organization or something, people who were worried about some big decision. I worried about no big decisions. I was always flittering about underneath somewhere. I wasn't the *absolute* bottom. As it turns out I did sort of get up a few steps, but I wasn't one of the higher people. So I want you to put yourself in a different kind of condition than the introduction said and just imagine this young graduate student who hasn't got his degree yet, who is working on his thesis. I'll start by saying how I got into the project, and then what happened to me. That's all, just what happened to me during the project.

I was working in my office* one day, when Bob Wilson† came in. I was working–[laughter] what the hell, I've got lots funnier yet; what are you laughing at?–Bob Wilson came in and said that he had been funded to do a job that was a secret and he wasn't supposed to tell anybody, but he was going to tell me because he knew that as soon as I knew what he was going to do, I'd see that I had to go along with it. So he told me about the problem of separating different isotopes of uranium. He had to ultimately make a bomb, a process for separating the isotopes of uranium, which was different from the one which was ultimately used, and he wanted to try to develop it. He told me about it and he said there's a meeting. . . I said I didn't want to do it. He said all right, there's a meeting at three o'clock, I'll see you there. I said it's all right you told me the secret because I'm not going to tell anybody, but I'm not going to do it. So I went back to work on my thesis, for about three minutes. Then I began to pace the floor and think about this thing. The Germans had Hitler and the

*at Princeton University.

†Robert Wilson, winner (with Arno Penzias) of the 1978 Physics Nobel Prize for discovering the cosmic microwave background radiation. *Ed.*

possibility of developing an atomic bomb was obvious, and the possibility that they would develop it before we did was very much of a fright. So I decided to go to the meeting at three o'clock. By four o'clock I already had a desk in a room and was trying to calculate whether this particular method was limited by the total amount of current that you can get in an ion beam, and so on. I won't go into the details. But I had a desk, and I had paper, and I'm working hard as I could and as fast as I can. The fellows who were building the apparatus planned to do the experiment right there. And it was like those moving pictures where you see a piece of equipment go bruuuup, bruuuup, bruuuup. Every time I'd look up the thing was getting bigger. And what was happening, of course, was that all the boys had decided to work on this and to stop their research in science. All the science stopped during the war except the little bit that was done in Los Alamos. It was not much science; it was a lot of engineering. And they were robbing their equipment from their research, and all the equipment from different research was being put together to make the new apparatus to do the experiment, to try to separate the isotopes of uranium. I stopped my work also for the same reason. It is true that I did take a six-week vacation after a while from that job and finished writing my thesis. So I did get my degree just before I got to Los Alamos, so I wasn't quite as far down as I led you to believe.

One of the first experiences that was very interesting to me in this project at Princeton was to meet great men. I had never met very many great men before. But there was an evaluation committee that had to decide which way we were going and to try to help us along, and to help us ultimately decide which way we were going to separate the uranium. This evaluation committee had men like Tolman and Smyth and Urey and Rabi and Oppenheimer and so forth on it. And there was

Compton, for example. One of the things I saw was a terrible shock. I would sit there because I understood the theory of the process of what we were doing, and so they'd ask me questions and then we'd discuss it. Then one man would make a point and then Compton, for example, would explain a different point of view, and he would be perfectly right, and it was the right idea, and he said it should be *this* way. Another guy would say well, maybe, there's this possibility we have to consider against it. There's another possibility we have to consider. I'm jumping! He should, Compton, he should say it again, he should say it again! So everybody is disagreeing, it went all the way around the table. So finally at the end Tolman, who's the chairman, says, well, having heard all these arguments, I guess it's true that Compton's argument is the best of all and now we have to go ahead. And it was such a shock to me to see that a committee of men could present a whole lot of ideas, each one thinking of a new facet, and remembering what the other fellow said, having paid attention, and so that at the end the decision is made as to which idea was the best, summing it all together, without having to say it three times, you see? So that was a shock, and these were very great men indeed.

This project was ultimately decided not to be the way that they were going to separate uranium. We were told then that we were going to stop and that there would be starting in Los Alamos, New Mexico, the project that would actually make the bomb and that we would all go out there to make it. There would be experiments that we would have to do, and theoretical work to do. I was in the theoretical work; all the rest of the fellows were in experimental work. The question then was what to do, because we had this hiatus of time since we'd just been told to turn off and Los Alamos wasn't ready yet. Bob Wilson tried to make use of his time by sending me to Chicago to find out all that I could about the bomb and

the problems so that we could start to build in our laboratories equipment, counters of various kinds, and so on that would be useful when we got to Los Alamos. So no time was wasted. I was sent to Chicago with the instructions to go to each group, tell them I was going to work with them, have them tell me about a problem to the extent that I knew enough detail so that I could actually sit down and start to work on the problem, and as soon as I got that far go to another guy and ask for a problem, and that way I would understand the details of everything. It was a very good idea, although my conscience bothered me a little bit. But it turned out accidentally (I was very lucky) that as one of the guys explained a problem I said why don't you do it that way and in a half an hour he had it solved, and they'd been working on it for three months. So, I did something! When I came back from Chicago I described the situation—how much energy was released, what the bomb was going to be like and so forth to these fellows. I remember a friend of mine who worked with me, Paul Olum, a mathematician, came up to me afterwards and said, "When they make a moving picture about this, they'll have the guy coming back from Chicago telling the Princeton men all about the bomb, and he'll be dressed in a suit and carry a briefcase and so on—and you're in dirty shirtsleeves and just telling us all about it." But it's a very serious thing anyway and so he appreciated the difference between the real world and that in the movies.

Well, there still seemed to be a delay and Wilson went to Los Alamos to find out what was holding things up and how they were progressing. When he got there he found that the construction company was working very hard and had finished the theater, and a few other buildings because they understood how, but they hadn't gotten instructions clear on how to build a laboratory—how many pipes for gas, how

much for water–so he simply stood around and decided how much water, how much gas and so on, and told them to start building the laboratories. And he came back to us–we were all ready to go, you see–and Oppenheimer was having some difficulties in discussing some problems with Groves and we were getting impatient. As far as I understand it from the position I was in, Wilson then called Manley in Chicago and they all got together and decided we'd go out there anyway, even if it wasn't ready. So we all went out to Los Alamos before it was ready. We were recruited, by the way, by Oppenheimer and other people and he was very patient with everybody; he paid attention to everybody's problems. He worried about my wife, who had TB, and whether there would be a hospital out there and everything, and it was the first time I met him in such a personal way and he was such a wonderful man. We were told among other things, for example, to be careful. Not to buy our train ticket in Princeton. Because Princeton was a very small train station, and if everybody bought train tickets to Albuquerque, New Mexico, there would be suspicion that something was up. And so everybody bought their tickets somewhere else, except me, because I figured if everybody bought their tickets somewhere else. . . . So when I went to the train station and I said I want to go to Albuquerque, New Mexico, he says, oh, he says, so all this stuff is for *you*! We had been shipping out crates full of counters for weeks and expecting they didn't notice that the address was Albuquerque. So at least I explained why it was that we were shipping out crates–I was going out to Albuquerque.

Well, when we arrived we were ahead of time and the houses for the dormitories and things like that were not ready. In fact, the laboratories weren't quite ready. We were pushing them, we were driving them by coming down ahead

of time. They went crazy at the other end and they rented ranch houses all around in the neighborhood. And we stayed at first in a ranch house and would drive in in the morning. The first morning I drove in was tremendously impressive; the beauty of the scenery, for a person from the East who didn't travel much, was sensational. There are the great cliffs; you've probably seen the pictures, I won't go into it in much detail. These things were high on a mesa and you'd come up from below and see these great cliffs and we were very surprised. The most impressive thing to me was that as I was going up, I said that maybe there were Indians even living here, and the guy who was driving the car just stopped; he stopped the car and walked around the corner and there were Indian caves that you could inspect. So it was really very exciting, in that respect.

When I got to the site the first time, I saw at the gate—you see there was a technical area that was supposed to have a fence around it ultimately, but because they were still building, it was still open. Then there was supposed to be a town and then a *big* fence further out, around the town—my friend Paul Olum, who was my assistant, standing with a clipboard checking the trucks coming in and out and telling them which way to go to deliver the materials in different places. When I went into the laboratory I would meet men I had heard of by seeing their papers in the *Physical Review* and so on. I had never met them before. This is John Williams, they said. A guy comes standing up from a desk which is covered with blueprints, his sleeves all rolled up, and he's standing there by some windows at one of the buildings ordering trucks and things going in different directions to build the things. In other words, we took over the construction company and finished the job. The physicists, in the beginning the experimental physicists particularly, had nothing to do until their build-

ings were ready, and apparatus was ready, so they just built the buildings, or assisted in building the buildings. The theoretical physicists, on the other hand, it was decided that they wouldn't live in the ranch houses, but they would live up at the site because they could start working right away. So we started working immediately, and that meant that we would each get a roll blackboard, you know, on wheels that you'd roll around, and we'd roll it around and Serber would explain to us all the things that they'd thought of in Berkeley about the atomic bomb, and nuclear physics and all these things, and I didn't know very much about it. I had been doing other kinds of things. And so I had to do an awful lot of work. Every day I would study and read, study and read, and it was a very hectic time. I had some luck. All the big shots had by some kind of accident—everybody but Hans Bethe—happened to have left at the same time; like Weisskopf had to go back to fix something at MIT, and Teller was away, just at a certain moment, and what Bethe needed was someone to talk to, to push his ideas against. Well, he came to this little squirt in an office and he starts to argue, to explain his idea. I said, "No, no, you're crazy, it'll go like this." And he said, "Just a moment," and he explained how he's not crazy, that I'm crazy, and we keep on going like this. It turned out that, although I—you see, when I hear about physics I just think about physics and I don't know who I'm talking to and I say the dopiest things like no, no, you're wrong or you're crazy—but it turned out that's exactly what he needed. So I got a notch up on account of that and I ended up as a group leader with four guys under me, which is underneath Bethe.

I had a lot of interesting experiences with Bethe. The first day when he came in, we had an adding machine, a Marchant that you work with your hands, and so he said "Let's see, the pressure"—the formula which he'd been working out involves

the pressure squared—"the pressure is 48; the square of 48. . . ." I reach for the machine; he says it's about 23 hundred. So I plug it out just to find out. He says, "You want to know exactly? It's 2,304." And so it came out 2,304. So I said, "How do you do that?" He says, "Don't you know how to take squares of numbers near 50? If it's near 50, say 3 below, then it's 3 below 25, like 47 squared is 22. And how much is left over is the square of what's residual. For instance, with your 3 less you get 9—2,209 from 47 squared. Very nice, OK?" So we (he was very good at arithmetic) kept on going and a few moments later we had to take the cube root of $2\frac{1}{2}$. Now, to do cube roots there was a little chart that you take that had some trial numbers that you try on the adding machine that the Marchant Company had given us. So (this takes him a little longer, you see) I opened the drawer, take out the chart, and he says "1.35." So I figured there is some way to take cube roots numbers near $2\frac{1}{2}$, but it turns out no. I said, "How do you do that?" He says, "Well," he says "you see the logarithm of 2.5 is so-and-so; you divide by 3 to get the cube root of so and so. Now, the log of 1.3 is this, the log of 1.4 is . . . I interpolate in between." I couldn't have divided anything by three, much less . . . So he knew all his arithmetic and he was very good at it and that was a challenge to me. I kept practicing. We used to have a little contest. Every time we'd have to calculate anything we'd rush to the answer, he and I, and I would win; after several years I began to be able to do it, you know get in there once, maybe one out of four. Of course you'd notice something funny about a number like if you have to multiply 174 by 140, for example. You notice that 173 by 141 which is like the square root of 3, times the square root of 2, which is the square root of 6, which is 245. But you have to notice the numbers, you see, and each guy would notice a different way—we had lots of fun.

Well, when I was first there, as I said, we didn't have the dormitories, and the theoretical physicists had to stay on the site. The first place they put us was in the old school building, from the boys' school which had been there previously. The first place I lived in was a thing called the mechanics lodge; we were all jammed in there in bunk beds and so on and it turned out it wasn't organized very well and Bob Christie and his wife had to go to the bathroom each morning through our bedroom. So that was very uncomfortable.

The next place we moved to was a thing called the Big House, which had a patio all the way around the outside on the second floor where all the beds had been stuck next to each other, all along the wall. And then downstairs was a big chart that told you what your bed number was and which bathroom you changed your clothes in. So under my name it said "Bathroom C," no bed number! As a result of this I was rather annoyed. At last the dormitory is built. I go down to the dormitory place to get rooms assigned, and they say you can pick your room now. I tried to pick one; you know what I did–I looked to see where the girls' dormitory was and I picked one that you could look out across. Later I discovered a big tree was growing right in front of it. But, anyway, I picked this room. They told me that temporarily there would be two people in a room, but that would only be temporary. Two rooms would share a bathroom. It would be double-decker beds, bunks, in there and I didn't want two people in the room. When I first got there, the first night, nobody else was there. Now my wife was sick with TB in Albuquerque, so I had some boxes of stuff of hers. So I opened a box and I took out a little nightgown and I just sort of threw it carelessly. I opened the top bed, I threw the nightgown careless on the top bed. I took out the slippers; I threw some powder on the floor in the bathroom. I just made it look like somebody else was there. OK?

So if the other bed is occupied, nobody else is going to sleep there. OK? So, what happened? Because it's a men's dormitory. Well, I came home that night and my pajamas are folded nicely and put under the pillow, and the slippers put nicely at the bottom of the bed. The lady's pajama is nicely folded and it's been put under the pillow, the bed is all fixed up and made, and the slippers are put down nicely. The powder is cleaned from the bathroom and *nobody* is sleeping up there. I still have the room to myself. So, next night, same thing. When I wake up I mess up the bed up top, I throw the nightgown, and the powder in the bathroom and so on, and I went on like this for four nights until it settled down. Everybody was settled and there was no more danger that they would put a second person in the room. Each night, everything was set very neatly, everything was all right, even though it was a men's dormitory. So, that's what happened in that situation.

I got involved in politics a little bit because there was a thing called the Town Council. Apparently there were certain things the Army people would decide about how the town was supposed to be run, with assistance from some governing board up there that I never knew anything about. But there was all kinds of excitement like there is in any political thing. In particular, there were factions: the housewives' faction, the mechanics' faction, the technical people's faction, and so on. Well, the bachelors and bachelor girls, the people who lived in the dormitory, felt they had to have a faction because a new rule had been promulgated–no women in the men's dorm, for instance. Well, this is absolutely ridiculous. All grown people of course (ha, ha). What kind of nonsense? So we had to have political action. So we decided and we debated and this stuff; you know how it is. And so I was elected to represent the dormitory people, you see, in the Town Council.

After I was in the Town Council for about a year or so, a year and a half, I was talking to Hans Bethe about something. He was up in the governing council during all this time. And I told him this story, I had done this trick one time with my wife's stuff in the upper bed and he starts to laugh. He says, "Ha, that's how you got on the Town Council." Because it turned out that what happened was this. There was a report, a very serious report. The poor woman was shaking, the woman who cleans the rooms in the dormitory had just opened the door and all of a sudden there is trouble–somebody is sleeping with one of the guys! Shaking, she doesn't know what to do. She reports, the charwoman reports to the chief charwoman, the chief charwoman reports to the lieutenant, the lieutenant reports to the major, it goes all the way up, it goes all the way up through to the generals to the governing board–what are they going to do?–they're going to think about it! So in the meantime, what instructions go down, down, through the captains, down through the majors, through the lieutenants, through the chars' chief, through the charwoman?–"Just put things back the way they are, clean 'em up" and see what happens. OK? Next day, report–same thing, brump, bruuuuump, bruuuump. Meantime, for four days, they worried up there, what they're going to do. So they finally promulgated a rule. "No women in the men's dormitory!" And that caused such a *stink* down there. You see, now they had to have all the politics and they elected somebody to represent them. . . .

Now I would like to tell you about the censorship that we had. They decided to do something utterly illegal, which was to censor mail of people inside the United States, in the Continental United States, which they have no right to do. So it had to be set up very delicately, as a voluntary thing. We would all volunteer not to seal our envelopes that we would

send our letters out with. We would accept, it would be all right, that they would open letters coming in to us; that was voluntarily accepted by us. We would leave the outgoing letters open; they would seal them, if they were OK. If they weren't OK in their opinion, in other words they found something that we shouldn't send out, they would send the letter back to us with a note that there was a violation of such and such a paragraph of our "understanding," and so on and so forth. So, very delicately, amongst all these liberal-minded scientific guys agreeing to such a proposition, we finally got the censorship set up. With many rules, such as that we were allowed to comment on the character of the administration if we wanted to, so we could write our senator and tell him we don't like the way things are run, and things like that. So it was all set up and they said that they would notify us if there were any difficulties.

So, the day starts, the first day for censorship. Telephone! Briiing! Me–"What?" "Please come down." I come down. "What's this?" It's a letter from my father. "Well, what is it?" There's lined paper, and there's these lines going out with dots–four dots under, one dot above, two dots under, one dot above, dot under dot. "What's that?" I said, "It's a code." They said, "Yes, it's a code; but what does it say?" I said, "I don't know what it says." They said "Well, what's the key to the code; how do you decipher it?" I said, "Well, I don't know." Then they said, "What's this?" I said, "It's a letter from my wife." "It says TJXYWZ TW1X3. What's this?" I said, "Another code." "What's the key to it?" "I don't know." They said, "You're receiving codes and you don't know the key?" "Precisely," I said, "I have a game. I challenge them to send me a code that I can't decipher, see? So they're making up codes at the other end and they're not going to tell me what the key is and they're sending them in." Now one of the

rules of the censorship was that they aren't going to disturb anything that you would ordinarily do, in the mail. So they said, "Well, you're going to have to tell them please to send the key in with the code." I said, "I don't want to see the key!" They said, "Well, all right, we'll take the key out." So we had that arrangement. OK? All right. Next day, I get a letter from my wife which says "It's very difficult writing because I feel that blank is looking over my shoulder." And in the spot there is nicely eradicated a splotch with ink eradicator. So I went down to the bureau and I said, "You're not supposed to touch the incoming mail if you don't like it, you can tell me but you're not supposed to do anything to it. Just look at it but you're not supposed to take anything out." They said, "Don't be ridiculous; do you think that's the way censors work anyway, with ink eradicator? They cut things out with the scissor." I said OK. So I wrote a letter back to my wife and said, "Did you use ink eradicator in your letter?" She writes back, "No, I didn't use ink eradicator in my letter; it must have been the ––––––," and there's a hole cut out. So I went back to the guy in charge, the major who was supposed to be in charge of all this, and complained. This went on for a few days. I felt I was sort of the representative to get the thing straightened out. He tried to explain to me that these people who were the censors had been taught how to do it, and they didn't understand this new way that we had to be so delicate about. I was trying to be the front, the one with the most experience, I was writing back and forth to my wife every day anyway. So, he said, "What's the matter, don't you think I have good faith, have good will?" I says, "Yes, you have perfectly good will but I don't think you have power." 'Cause you see, this had gone on three or four days. He said "We'll see about that!" He grabs the telephone . . . everything was straightened out. So no more was the letter cut.

However, there were a number of difficulties that arose. For example, one day I got a letter from my wife and a note from the censor that said there was a code enclosed, without the key, and so we removed it. So when I went to see my wife in Albuquerque that day, she says "Well, where's all the stuff?" I said, "What stuff?" She says "Lethage, glycerine, hot dogs, laundry." I said, "Wait a minute, that was a list?" She says, "Yes." "That was a *code,*" I said. They thought it was a code–lethage, glycerine, etc. Then one day I am jiggling around–in the first few weeks all this went on, it was a few weeks before we got each other straightened out, but–I'm piddling around with the adding machine, the computing machine, and I notice something. So, writing every day, I had a lot of things to write–so, it's very peculiar. Notice what happens. If you take one divided by 273 you get 0.004115226337. It's quite acute and then it goes a little cockeyed when your carrying occurs for only about three numbers and then you can see how the 10 10 13 is really equivalent to 114 again, or 115 again, and it keeps on going, and I was explaining that, how nicely it repeated itself after a couple of cycles. I thought it was kind of amusing. Well, I put that in the mail and it comes back to me; it doesn't go through, and there's a little note "Look at Paragraph 17B." I look at paragraph 17B. It says "Letters are to be written only in English, Russian, Spanish, Portuguese, Latin, German, and so forth. Permission to use any other language must be obtained in writing." And then it said "No codes." So I wrote back to the censor a little note included in my letter which said that I feel that of course this cannot be a code, because if you actually *do* divide 273 by 1 you do, in fact, *get* ––––– and I wrote all that there, and therefore there's no more information in the number 1-1-1-1–zero, zero, zero than there is in the number 273 which is hardly any information. And so forth. I therefore asked for permission to write

my letters in Arabic numerals. I like to use Arabic numerals in my letters. So, I got that through all right.

There was always some kind of difficulty with the letters going back and forth. At one time my wife kept insisting on mentioning the fact that she feels uncomfortable writing with the feeling that the censor is looking over [her shoulder]. As a rule we aren't supposed to mention censorship–*we* aren't, but how can they tell her? So they keep sending me a note, "Your wife mentioned censorship." *Certainly,* my wife mentioned censorship, so finally they sent me a note that said "Please inform your wife not to mention censorship in her letters." So I take my letter and I start, "I have been instructed to inform you not to mention censorship in your letters." Phoom, phoooom, it comes right back! So I write, "I have been instructed to inform her not to mention censorship. How in the heck am I going to do it? Furthermore, *why* do I have to instruct her not to mention censorship? You keeping something from me?" It is very interesting that the censor himself has to tell me to tell my wife not to tell me that she's . . . But they had an answer. They said, yes, that they are worried about mail being intercepted on the way from Albuquerque, and that they would find out that there was censorship if they looked in the mail and would she please act much more normal. So I went down the next time to Albuquerque and I talked to her and I said, "Now, look, let's not mention censorship," but we had had so much trouble that we had at last worked out a code, something illegal. We had a code; if I would put a dot at the end of my signature it meant I had had trouble again, and she would move on to the next of the moves that she had concocted. She would sit there all day long because she was ill and she would think of things to do. The last thing that she did was to send me, which she found perfectly legitimately, an advertisement that said "Send your boyfriend a letter on a jig-

saw puzzle. Here are the blanks. We sell you the blank, you write the letter on it, take it all apart, you put it in a little sac and mail it." So I received that one with a note saying "We do not have time to play games. Please instruct your wife to confine herself to ordinary letters!" Well, we were ready with the one more dot. The letter would start, "I hope you remembered to open this letter carefully because I have included the Pepto-Bismol for your stomach as we arranged." It would be a letter full of powder. In the office we expected they would open it quickly, the powder would go all over the floor, they would get all upset because you are not supposed to upset anything, you'd have to gather all this Pepto-Bismol. . . . But we didn't have to use that one. OK?

As a result of all these experiences with the censor, I knew exactly what could get through and what could not get through. Nobody else knew as well as I. And so I made a little money out of all of this by making bets. One day, on the outside fence, I had discovered that workmen who lived still further out and wanted to come in were too lazy to go around through the gate, and so they had cut themselves a hole, some distance along. So I went out the fence, went over to the hole and came in, went out again, and so on, until the guy, the sergeant at the gate begins to wonder what's happening, this guy is always going out and never coming in? And of course his natural reaction was to call the lieutenant and try to put me in jail for doing this. I explained that there was a hole. You see, I was always trying to straighten people out, point out that there was a hole. And so I made a bet with somebody that I could tell where the hole in the fence was, in the mail, and mail it out. And sure enough, I did. And the way I did it was I said—"You should see the way they administer this place"; you see, that's what we were allowed to say. "There's a hole in the fence 71 feet away from such and such

a place, that's this size and that size, that you can walk through." Now, what can they do? They can't say to me that there is no such hole. I mean, what are they going to do, it's their own hard luck that there's such a hole. They should *fix* the hole. So, I got that one through. I also got through a letter that told about how one of the boys who worked in one of my groups had been wakened up in the middle of the night and grilled with lights in front of him by some idiots in the Army, because they found out something about his father or something. I don't know, he was supposed to be a communist. His name was Kamane. He's a famous man now.

Well, there was also some other things. I was always trying to straighten them out, like point out the holes in the fence and so forth, but I was always trying to point these things out in a non-direct manner. And one of the things I wanted to point out was this: that at the very beginning we had terribly important secrets. We'd worked out lots of stuff about uranium, how it worked, and all this stuff was in documents that were in filing cabinets that were made out of wood that had on them little ordinary, common padlocks. Various things made by the shop were on the cabinets, like a rod that would go down and then a padlock to hold it, but it was always just a padlock. Furthermore, you could get the stuff, without even opening the padlock, out of these wooden cabinets; you just tilt it over backwards and the bottom drawer, you know, has a little rod that's supposed to hold it. And there's a hole in the wood underneath; you can pull the papers out from below. So I used to pick the locks all the time and point out that it was very easy to do. And every time we had a meeting of the whole group, everybody together, I would get up and I'd say that we have important secrets and we shouldn't keep them in such things. These were such poor locks. We need better locks. And so one day Teller got up at the meeting, and

said to me, "Well, I don't keep my most important secrets in my filing cabinet; I keep them in my desk drawer. Isn't that better?" I said, "I don't know, I haven't seen your desk drawer." Well, he's sitting near the front of the meeting and I'm sitting further back. So the meeting continues and I sneak out of the meeting and I go down to see his desk drawer. OK? I don't even have to pick the lock on the desk drawer. It turns out that if you put your hand in the back underneath you can pull out the paper like those toilet paper dispensers; you pull out one, it pulls another, it pulls another. . . . I emptied the whole damn drawer, took everything out, and put it away to one side and then went up on the higher floor and came back. The meeting is just ending and everybody is just coming out and I join the crew like this, you see, walking along with it, and run up to catch up with Teller, and say, "Oh, by the way, let me see your desk drawer." So he says, "Certainly," so we walk into his office and he shows me the desk and I look at it and say that looks pretty good to me. I said "Let's see what you have in there." "I'd be very glad to show it to you," he says, putting in the key and opening the drawer, "if you hadn't already seen it yourself." The trouble with playing a trick on a highly intelligent man like Mr. Teller is the *time* it takes him to figure out from the moment that he sees there is something wrong till he understands exactly what happened is too damn small to give you any pleasure!

Well, I had a lot of other fun with the safes but it has nothing to do with Los Alamos, so I won't discuss it further. I want to tell about some of the problems, special problems, that I had that are rather interesting. One thing had to do with the safety of the plant at Oak Ridge. Los Alamos was going to make the bomb, but at Oak Ridge they were trying to separate the isotopes of uranium, uranium 238 and uranium 236, 235, the latter one, which was the explosive one, all right? So, they

were *just* beginning to get infinitesimal amounts from an experimental thing, of 235, and at the same time they were practicing. There was a big plant, they were going to have vats of the stuff, chemicals, and they were going to take the purified stuff and repurify and get it ready for the next stage. You have to purify it in several stages. So they were practicing the chemistry on the one hand and they were just getting a little bit from one of the pieces of apparatus experimentally on the other hand. And they were trying to learn how to assay it, to determine how much uranium 235 there is in it, and we would send them instructions and they never got it right. So finally Segrè* said that the only possible way to get it right was that he'd have to go down there to see what they're doing, to understand why the assay was wrong. The Army people said no, it is our policy to keep all the information of Los Alamos at one place, and that the people in Oak Ridge should not know anything about what it was used for; they just knew what they were trying to do. I mean the higher people knew they were separating uranium, but they didn't know how powerful the bomb was or exactly how it worked or anything. And the people underneath didn't know at *all* what they were doing. And the Army wanted to keep it that way; there was no information going back and forth, but Segrè finally insisted on it, that it was important. They'd never get the assays right, the whole thing would go up in smoke. So Segrè went down to see what they were doing and as he was walking through he saw them wheeling a tank carboy of water, green water; the green water is uranium nitrate. He says, "You're going to handle it like that when it's purified too? Is that what you're going to do?" They said, "Sure, why not?" "Won't it explode?" he says. "Huh?! *Explode!??*" And so the

*Emilio Segrè, winner (with Owen Chamberlain) of the 1959 Physics Nobel Prize for discovering the antiproton. *Ed.*

Army said, "You see, we shouldn't have let any information go across!" Well, it turned out that the Army had realized how much stuff we needed to make a bomb, 20 kilograms or whatever it was, and they realized that that much material would never, purified, would never be in the plant, so there was no danger. But they did not know that the neutrons were enormously more effective when they are slowed down in water. And so in water it takes less than a tenth, no a hundredth, very much less material to make a reaction which makes radioactivity. It doesn't make a big explosion, but it makes radioactivity, it kills people all around and so on. So, it was very dangerous and they had not paid any attention to the safety at all.

So a telegram goes from Oppenheimer to Segrè: Go through the entire plant, notice where all the concentrations are supposed to be, with the process as *they* designed it. We will calculate in the meantime how much material can come together before there's an explosion. And so two groups started working on it. Christie's group worked on water solutions and I worked on dry powder in boxes, my group. And we calculated about how much material. And Christie was going to go down and tell them all at Oak Ridge what the situation was. And so I happily gave all my numbers to Christie, and said, you have all the stuff, and go. Christie got pneumonia; I had to go. I never traveled on an airplane before; I traveled on an airplane. They *strapped* the secrets, with a little thing with a belt, on my back! The airplane in those days was like a bus. You stop off every once in a while except the stations were further apart. You stop off to wait. There's a guy standing there next to me with a keychain, swinging it, saying something like, "It must be *terribly* difficult to fly without a priority on airplanes these days." I couldn't resist. I said, "Well, I don't know, I *have* a priority." A little bit later some generals come aboard and they are going to put out some of us number 3s. It's all right, I'm a

number 2. That passenger probably wrote to his congressman, if he wasn't a congressman himself, saying, what are they doing sending these little kids around with number priorities in the middle of the war? At any rate, I arrived there. The first thing I did was have them take me to the plant and I said nothing; I just looked at everything. I found out that the situation was even worse than Segrè reported because he was confused at the first time through. He noticed certain boxes in big lots but he didn't notice other boxes in another room in a big lot, but it was the same room on the other side. And things like that. So if you have too much stuff together it goes up, you see. So I went through the entire plant and, I have a very bad memory but when I work intensively, I have a good short-term memory and so I could remember all kinds of crazy things like building ninety-two-oh-seven, vat number so and so, and so forth. I went home that night and I went through the whole thing explaining where all the dangers were, what you would have to do to fix this. It's rather easy—you put cadmium in solutions to absorb the neutrons in the water, you separate the boxes so they are not too dense, too much uranium together and so on, according to certain rules. And so I used all the examples, worked out all the examples and how the process of freezing worked. I felt that you couldn't make the plant safe unless you knew how it worked. So the next day there was going to be a big meeting.

Oh, I forgot to say, before I left, Oppenheimer said to me, "Now," he said, "when you go, the following people are technically able down there at Oak Ridge: Mr. Julian Webb, Mr. So and so, and so on. I want you to make sure that these people are at the meeting, that you tell them how the thing, you know, the safety, that they really *understand*—they're in charge." I said, "What if they're not at the meeting, what am I supposed to do?" He said, "Then you should say—*Los*

Alamos cannot accept the responsibility for the safety of the Oak Ridge plant unless!!!" I said, "You mean me, little Richard's, going to go in there and say. . . ?" He says, "Yes, little Richard, you go and do that." I really grew up fast! So when I arrived, sure enough, I arrived there and the meeting was the next day and all these people from the company, the big shots in the company and the technical people that I wanted were there, and the generals and so forth, that were interested in the problems, organizing everything. It was a big meeting about this very serious problem of safety, because the plant would never work. It would have blown up, I swear it would have, if nobody had paid attention. So there was a lieutenant who took care of me. He told me that the colonel said that I shouldn't tell them how the neutrons work and all the details because we want to keep the things separate. Just tell them what to do to keep it safe. I said, in my opinion, it is impossible for them to understand or to obey a bunch of rules if they don't understand, unless they understand how it works. So it's my opinion that it's only going to work if I tell them, and *Los Alamos cannot accept the responsibility for the safety of the Oak Ridge plant unless they are fully informed as to how it works!!* It was great. So he goes to the colonel. "Give me just five minutes," the colonel says. He goes to the window and he stops and thinks and that's what they're very good at. They are good at making decisions. I thought it was very remarkable how a problem of whether or not information as to how the bomb works should be in the Oak Ridge plant had to be decided and could be decided in five minutes. So I have a great deal of respect for these military guys because I never can decide anything very important in any length of time, at all.

So, in five minutes, he says, all right, Mr. Feynman, go ahead. So I sat down and I told them all about neutrons, how they worked, da da, ta ta ta, there are too many neutrons to-

gether, you got to keep the material apart, cadmium absorbs, and slow neutrons are more effective than fast neutrons, and yak yak–all stuff which was elementary primer stuff at Los Alamos, but they had never heard of any of it, so I turned out to be a tremendous genius to them. I was a god coming down from the sky! There was all these phenomena that were not understood and never heard of before, I knew all about it, I could give them facts and numbers and everything else. So, from being rather primitive back there at Los Alamos, I was a super genius at the other end. Well, the result was that they decided, they made little groups, to make their own calculations to learn how to do it. They started to redesign plants. The designers of the plants were there, the construction designers, engineers, chemical engineers for the new plant that was going to handle the separated material were there. And other people were there. And I went away again. They told me to come back in a few months; they were going to redesign their plant for the separation.

So I came back in a few months, a month or so, and Stone and Webster Company, the engineers, had finished the design of the plant and now it was for me to look at the plant. OK? How do you look at a plant that ain't built yet? I don't know. So I go into this room with these fellows. There was always a Lieutenant Zumwalt that was always coming around with me, taking care of me, you know; I had to have an escort everywhere. So he goes with me, he takes me into this room and there are these two engineers and a *looooooong* table, great big long table, tremendous table, covered with a blueprint that's as big as a table; not one blueprint, but a stack of blueprints. I took mechanical drawing when I was in school, but I wasn't too good at reading blueprints. So they start to explain it to me because they thought I was a genius. And they start out, "Mr. Feynman, we would like you to un-

derstand, the plant is so designed, you see one of the things we had to avoid was accumulation." Problems like–there's an evaporator working which is trying to accumulate the stuff; if the valve gets stuck or something like that and they accumulate too much stuff, it'll explode. So they explained to me that this plant is designed so that no *one* valve, if any one valve gets stuck nothing will happen. It needs at least two valves everywhere. So then they explain how it works. The carbon tetrachloride comes in here, the uranium nitrate from here comes in here, it goes up and down, it goes up through the floor, comes up through the pipes, coming up from the second floor, bluuuuurp, from the blueprints, down, up, down, up, very fast talking explaining the very, very complicated chemical plant. I'm completely dazed, worse, I don't know what the symbols on the blueprint mean! There is some kind of a thing that at first I think it's a window. It's a square with a little cross in the middle, all over the damn place. Lines with this damn square, lines with the damn square. I think it's a window; no, it can't be a window, 'cause it ain't always at the edge. I want to ask them what it is. You must have been in a situation like this–you didn't ask them right away, right away it would have been OK. But they've been talking a little bit too long. You hesitated too long. If you ask them now they'll say, what are you wasting my time all this time for? I don't know what to do; I think to myself, often in my life I have been lucky. You are not going to believe this story, but I swear it's absolutely true; it's such sensational luck. I thought what am I going to do, what am I going to *do*????? I got an idea. Maybe it's a valve? So, in order to find out whether it's a valve or not I take my finger and I put it down in the middle of one of the blueprints on page number 3 down in the end and I said, "What happens if this valve gets stuck?" figuring they're going to say, "That's not a valve, sir,

that's a window." So one looks at the other and says, "Well, if that valve gets stuck," and they go up and down on the blueprint, up and down, the other guy up and down, back and forth, back and forth, and they both look at each other and they turn around to me and they open their mouths—"You're absolutely right, sir." So they roll up the blueprints and away they went and we walked out. And Lt. Zumwalt, who had been following me all the way through, said, "You're a genius. I got the idea you were a genius when you went through the plant once and you could tell them about evaporator C-21 in building 90-207 the next morning," he says, "but what you have just done is so *fantastic,* I want to know how, *how* do you do something like that?" I told him, you try to find out whether it's a valve or not.

Well, another kind of problem that I worked on was this. We had to do lots of calculations and we did them on Marchant calculating machines. By the way, just to give you an idea of what Los Alamos was like, we had these Marchant computers. I don't know if you know what they look like, hand calculators with numbers and you push them, and they multiply, divide, add, and so on. Not like they do easy now, but hard; they were mechanical gadgets. And they had to be sent back to the factory to be repaired. We didn't have a special man to do it, which was the standard way, and so they would always be sent to the factory. Pretty soon you were running out of machines. So I and a few other fellows started to take the covers off. We weren't supposed to—the rules "You take the covers off, we cannot be responsible . . ." So we took the covers off and we had a nice series of lessons. Like the first one we took the cover off, there was a shaft with a hole in it and a spring which was hanging this way, and obviously the spring went in the hole—so that was easy. So anyway, we got like a series of lessons, by God, on how to fix them and

we got better and better and we made more and more elaborate repairs. When we got something too complicated we sent it out, back to the factory, but we'd do the easy ones and kept the things going. I also did some typewriters. I ended up doing all the computers; the other fellows quit on me. I did a few typewriters. There was a guy in the machine shop who was better than I was and he took care of typewriters; I took care of adding machines. However, we decided that the big problem was to figure out exactly what happened during the bomb's explosion when you push the stuff in by an explosion and then it goes out again. Exactly what happens, so you can figure out exactly how much energy was released and so on, required much more calculating than we were capable of. And a rather clever fellow by the name of Stanley Frankle realized that it could possibly be done on IBM machines. The IBM company had machines for business purposes, adding machines that are called tabulators for listing sums and a multiplier, just a machine, a big box, you put cards in and it would take two numbers from a card and multiply it and print it on a card. And then there were collators and sorters and so on. So he decided, he'd figured out a nice program. If we got enough of these machines in a room, we would take the cards and put them through a cycle; everybody who does numerical calculations now knows exactly what I'm talking about, but this was kind of a new thing: mass production with machines.

We had done things like this on adding machines. Usually you go one step across yourself, doing everything. But this was different—where you go first to the adder, then we go to the multiplier, then you go to the adder, and so on. So, he designed this thing and ordered the machine from the IBM company, 'cause we realized it was a good way of solving our problems. We found that there was somebody in the Army that had

IBM training. We needed a man to repair them, to keep them going and everything. And they were going to send this fellow, but it was delayed, always delayed. Now, we *always* were in a hurry. I have to explain that—*everything* we did, we tried to do as quickly as possible. In this particular case, we worked out all the numerical steps that we were supposed to, that the machines were supposed to do, multiply this, and then do this, and subtract that. And then we worked out the program, but we didn't have any machine to test it on. So what we did, I arranged, was a room with girls in it, each one had a Marchant. But *she* was the multiplier and *she* was the adder, and this one cubed, and so we had cards, index cards, and all she did was cube this number and send it to the next one. She was imitating the multiplier, the next one was imitating the adder; we went through our cycle, we got all the bugs out. Well, we did it that way. And it turned out that the speed at which we were able to do it—we'd never done mass production calculating and everybody who'd ever calculated before, every single person, did all the steps. But Ford had a good idea, the damn thing works a hell of a lot faster the other way and we got speed with this system that was the predicted speed for the IBM machine, the same. The only difference is that the IBM machines didn't get tired and could work three shifts. But the girls got tired after a while. So, anyway, we got the bugs out during that process and finally the machines arrived, but not the repairman. So, we went down to put them together. And they were one of the most complicated machines of the technology of those days, these computing machines, big things that came partially disassembled with lots of wires and blueprints of what to do. We went down, we put them together, Stan Frankle and I and another fellow, and we had our troubles. Most of the trouble was the big shots coming all the time and saying you're going to break something, going to break

something. We put them together and sometimes they would work, and sometimes they were put together wrong and they didn't work. And so we fiddled around and got them to work. We didn't get them all to work and I was last working on some multiplier, saw a bent part inside and I was afraid to straighten it because it might snap off. They were always telling us we were going to bust it irreversibly. And finally the man from the IBM company came, as a matter of fact, according to schedule, but he came and he fixed the rest that we hadn't got ready, and we got the program going. But he had trouble with the one that I had had trouble with. So after three days he was still working on that last one. I went down, I said, "Oh, I noticed that was bent." He said, "Oh, of course, that's all there is to it!" (Snap)–It was all right. So that was it.

Well, Mr. Frankle started this program and began to suffer from a disease, the computer disease, that anybody who works with computers now knows about. It's a very serious disease and it interferes completely with the work. It was a serious problem that we were trying to do. The disease with computers is you *play* with them. They are so wonderful. You have these x switches that determine, if it's an even number you do this, if it's an odd number you do that, and pretty soon you can do more and more elaborate things if you are clever enough, on one machine. And so after a while it turned out the whole system broke down. He wasn't paying any attention; he wasn't supervising anybody. The system was going very, very slowly. The real problem was that he was sitting in a room figuring out how to make one tabulator automatically print arc-tangent x, and then it would start and it would print columns and then bitsi, bitsi, bitsi and calculate the arc-tangents automatically by integrating as it went along and make a whole table in one operation. Absolutely useless. We *had* tables of arc-tangents. But if you've ever worked with comput-

ers you understand the disease. The *delight* to be able to see how much you can do. But he got the disease for the first time, the poor fellow who invented the thing got the disease.

And so, I was asked to stop working on the stuff I was doing in my group and go down and take over the IBM group. I noticed the disease and I tried to avoid it. And although they did three problems in nine months, I had a very good group. The first problem was that they had never told the fellows—they had selected from all over the country, a thing called Special Engineer Detachment. There were clever boys from high school, who had engineering ability, and the Army collected them together in the Special Engineer Detachment. They sent them up to Los Alamos. They put them in barracks and they would tell them *nothing.* Then they came to work and what they had to do was work on IBM machines, punching holes, numbers that they didn't understand, nobody told them what it was. The thing was going very slowly. I said that the first thing there has to be is that the technical guys know what we're doing. Oppenheimer went and talked to the security people and got special permission. So I had a nice lecture in which I told them what we were doing, and they were all excited. We're fighting a war. We see what it is. They knew what the numbers meant. If the pressure came out higher, that meant there was more energy released and so on and so on. They knew what they were doing. *Complete* transformation! *They* began to invent ways of doing it better. They improved the scheme. They worked at night. They didn't need supervising in the night. They didn't need anything. They understood everything. They invented several of the programs that we used and so forth. So my boys really came through and all that had to be done was to tell them what it was, that's all. It's just, don't tell them, they're punching holes, please. As a result, although it took them nine months

to do three problems before, we did nine problems in *three* months, which is about nearly ten times as fast. But one of the secret ways that we did our problems was this: The problems consisted of a bunch of cards which had to go through a cycle. First add, then multiply, and so it went through the cycle of machines in this room, slowly about as it went around and around. So we figured a way, by taking a different colored set of cards, to put them through a cycle too, but out of phase. We'd do two or three problems at a time. You see this was another problem. While this one was adding, it was multiplying on the other problem. And such managerial schemes, we got many more problems.

Finally, near the end of the war, just before we had to make a test in Alamogordo, the question was, how much energy would be released? We had been calculating the release from various designs but the specific design which was ultimately used we hadn't computed. So Bob Christie came down and said, we would like the results for how this thing is going to work in one month, or some very short time, I don't know, less than that, three weeks. I said, "It's impossible." He said, "Look, you're putting out so and so many problems a week. It takes only two weeks per problem, or three weeks per problem." I said, "I know, it takes much longer to do the problem, but we're doing them in *parallel.* As they go through it takes a long time and there's no way to make it go around faster." So he went out. I began to think–is there a way to make it go around faster? Well, if we did nothing else on the machine, so there was nothing else interfering, and so on and so on. I began to think. I put on the blackboard a challenge–CAN WE DO IT? to the boys. They all respond, yes, we'll work double shifts, we'll work overtime, all this kind of thing, we'll *try* it. We'll *try* it!! And so the rule was, all other problems *out.* Only one problem and just concentrate on this thing. So they started to work.

My wife died in Albuquerque and I had to go down. I borrowed Fuchs' car; he was a friend of mine in the dormitory. He had an automobile. He was using the automobile to take the secrets away, you know, they went down to Santa Fe. He was the spy; I didn't know that. I borrowed his car to go to Albuquerque. The damned thing got three flat tires on the way. I came back from there and I went into the room, because I was supposed to be supervising everything, but I couldn't do it for three days. It was in this *mess,* this big rush to get the answer for the test that was going to be done in the desert. I go into the room and there are three different color cards. There's white cards, there's blue cards, there's yellow cards and I start to say, "Well, you're not supposed to do more than one problem–only one problem!" They said, "Get out, get out, get out. Wait, we'll explain everything." So I waited, and what happened was this. As they went through, sometimes the machine made a mistake or they put a wrong number in; that happened. What we used to have to do was to go back and do that over again. But they noticed this, that the deck represented positions and depth in the machine, in space or something. A mistake made here, in one cycle, only affects the nearby numbers, the next cycle affects the nearby numbers, and so on. It works its way through the pack of cards. If you have fifty cards and you make a mistake at card number 39, it affects 37, 38, and 39. The next card 36, 37, 38, 39, and 40. The next time it spreads like a disease. So they found an error, back a way, and they got an idea. They would only compute a small deck, of ten cards, around the error. And because ten cards could be put through the machine faster than the deck of fifty cards, they would go with this other deck rapidly through while they continued with the fifty cards with the disease spreading. But the other thing was computing faster and they would seal it all up and correct it.

OK? Very clever. That was the way those guys worked, really hard, very clever, to get speed. There was no other way. If they had to stop to try to fix it, we'd have lost out time. We couldn't have got it. That was what they were doing. Of course you know what happened while they were doing that. They found an error in the blue deck. And so they had a yellow deck with a little fewer cards; it was going around faster than the blue deck, you know. Just when they are going crazy, because after they get it straightened out they got to fix the white one, they got to take the other cards out and replace it by the right ones, and continue correctly, and it's rather confusing–you know how those things always are. You don't want to make a mistake. Just at the time when they've got these three decks going, they're trying to seal everything up, the BOSS comes walking in. "Leave us alone," they said, so I left them alone and everything came out; we solved the problem in time and that's the way it was.

I would like to tell you just a few words about some of the people that I met. I was an underling at the beginning. I became a group leader, but I met some very great men–besides the men on the evaluation committee, the men that I met in Los Alamos. And there are so many of them that it's one of my great experiences in life to have met all these wonderful physicists. Men that I had heard of, smaller and larger, but the greatest ones were there also. There was of course Fermi.* He came down once. The first time that he came was from Chicago to consult a little bit, to help us if we had some problems. We had a meeting with him and I had been doing some calculations and gotten some results. The calculations were so elaborate it

*(1901–1954) Winner of the 1938 Nobel Prize in Physics for demonstrating the existence of new radioactive substances produced by neutron irradiation and related work. Fermi was also responsible for the first controlled nuclear reaction at the University of Chicago in December 1942. *Ed.*

was very difficult. Now, usually, I was the expert at this, I could always tell you what the answer was going to look like or when I got it I could explain why. But this thing was so complicated I couldn't explain *why* it was like that. So I said to Fermi that I was doing this problem and I started to calculate–he said, wait, before you tell me the result, let me think. It's going to come out like this (he was right), and it's going to come out like this because of so and so. And there's a perfectly obvious explanation. . . . So *he* was doing what I was supposed to be good at, ten times better. So that was quite a lesson to me.

Then there was Von Neumann, who was the great mathematician. He suggested, I won't go into the things here, some very clever technical observations. We had some very interesting phenomena in the computing of the numbers. The problem looked as if it was unstable and he explained why and so forth. It was very good technical advice. But we used to go for walks often to get rest, like on Sunday or something. We'd walk in the canyons in the neighborhood and we'd often walk with Bethe, Von Neumann, and Bacher. It was a great pleasure. And the one thing that Von Neumann gave me was an idea that he had which was interesting. That you don't have to be responsible for the world that you're in, and so I have developed a very powerful sense of social irresponsibility as a result of Von Neumann's advice. It's made me a very happy man since. But it was Von Neumann who put the seed in which grew now into my *active* irresponsibility!

I also met Niels Bohr.* That was interesting. He came down, his name was Nicholas Baker in those days and he came with Jim Baker, his son, whose name is really Aage†. They came

---

*(1885–1962) Winner of the 1922 Nobel Prize in Physics for his work on the structure of atoms and of the radiation emanating from them. *Ed.*

†Aage Bohr (1922–) winner (with Ben Mattelson and James Rainwater)

from Denmark and they came to visit, and they were *very* famous physicists, as you all know. All the big shot guys, to them, he was even a great god; they were listening to him and so on. And he would talk about things. We were at a meeting and everybody wanted to *see* the great Bohr. So there were a lot of people and I was back in a comer somewhere and we talked about, discussed, the problems of the bomb. That was the first time. He came and he went away and all I could see of him was from between somebody's heads, from the corner. Next time he's due to come, in the morning of the day he's due I get a telephone call. "Hello, Feynman?" "Yes." "This is Jim Baker"; it's his son. "My father and I would like to speak to you." "Me? I'm Feynman, I'm just a . . ." "That's right. OK." So, eight o'clock in the morning, before anybody's awake, I go down to the place. We go into an office in the technical area and he says, "We have been thinking how we could make the bomb more efficient and we think of the following idea." I say, "No, it's not going to work, it's not efficient, blah, blah, blah." So he says, "How about so and so?" I said, "That sounds a little bit better, but it's got this damn fool idea in it." So forth, back and forth. I was always *dumb* about one thing, I never knew who I was talking to. I was always worried about the physics; if the idea looked lousy, I said it looked lousy. If it looked good, I said it looked good. Simple proposition, I've always lived that way. It's nice, it's pleasant, if you can do it. I'm lucky. Just as lucky as I am with that blueprint, I'm lucky in my life that I can do that. So this went on for about two hours of going back and forth over lots of ideas, and then tearing back and forth, arguing. The great Niels always lighting his pipe; perpetually, it always went out. And he

---

of the 1975 Nobel Prize in Physics for their theory of the structure of the atomic nucleus. *Ed.*

talked in a way that was un-understandable. He said, "Mumble, mumble," hard to understand, but his son I could understand better. Finally he said, "Well," he says, lighting his pipe, "I guess we can call in the big shots *now.*" So then they called all the other guys and had a discussion with them. And then the son told me what had happened was—the last time he was there he said to his son—"Remember the name of that little fellow in the back over there? He's the only guy who's not afraid of me, and will say when I've got a crazy idea. So *next* time when we want to discuss ideas, we're not going to be able to do it with these guys who say everything is yes, yes, Dr. Bohr. Get that guy first, we'll talk with him first."

The next thing that happened was, of course, the test after we'd made the calculations. We had to make the test. I was actually at home on a short vacation at that time, I guess because my wife died, and so I got a message from Los Alamos that said, "The baby will be born, is expected on such and such a day." So I flew back, and I *just* arrived on the site while the buses were leaving; I couldn't even get into my room. At Alamogordo we waited out there in the distance; we were 20 miles away. And we had a radio and they were supposed to tell us when the thing was going to go off and so forth. The radio wouldn't work, and we never knew what was happening. But just a few minutes before it was supposed to go off, the radio started to work and they told us there was 20 seconds or something to go. For people who were far away like we were—others were closer, six miles away—they gave out dark glasses that you could watch it with. Dark glasses!! Twenty miles away from the damn thing, you get dark glasses—you couldn't see a damn thing through dark glasses. So then I figured the only thing that could really hurt your eyes—bright light can never hurt your eyes—it's ultraviolet light that does. So I got behind a truck windshield, so the ul-

traviolet can't go through glass, so that would be safe, and so I could *see* the damn thing. Other people were never going to *see* the damn thing. OK. Time comes, and this *trememdous* flash out there, so bright I quickly see this purple splotch on the floor of the truck. I said, "That ain't it. That's an after-image." So I turn back up and I see this white light changing into yellow and then into orange. The clouds form and then they disappear again, the compression and the expansion forms and makes clouds disappear. Then finally, a big ball of orange, the center that was so bright, became a ball of orange that started to rise and billow a little bit and get a little black around the edges and then you see it's a big ball of smoke with flashes on the inside of the fire going out, the heat. I saw all that and all this that I just described in just a moment; took about one minute. It was a series from bright to dark and I had seen it. I am about the only guy that actually looked at the damn thing, the first Trinity Test. Everybody else had dark glasses. The people at six miles couldn't see it because they were all told to lie on the floor with their eyes covered, so nobody saw it. The guys up where I was all had dark glasses. I'm the only guy who saw it with the human eye. Finally, after about a minute and a half, there's suddenly a tremendous noise, *BANG,* and then rumble, like thunder, and that's what convinced me. Nobody had said a word during this whole minute, we were all just watching quietly, but this sound released everybody, released me particularly because the solidity of the sound at that distance meant that it had really worked. The man who was standing next to me said, when the sound went off, "What's that?" I said, "That was the bomb." The man was William Laurence of the *New York Times,* who had come. He was going to write an article that was going to describe the whole situation. I had been the one

who was supposed to have taken him around. It was found that it was too technnical for him.

Later Mr. Smyth of Princeton came and I showed him around Los Alamos. For example, we went into a room and there on the end of a pedestal, a little narrower than that, was a small ball about so big, silver plated—you could put your hand on it, it was warm. It was radioactive; it was plutonium. And we stood at the door of this room talking about it. There was a new element that was made by man that had never existed on the earth before, except for a very short period possibly, at the very beginning. And here it was all isolated and radioactive and had these properties. And we had made it. And so it was very, tremendously valuable, nothing more valuable and so forth and so on. Meanwhile, you know how people do when you talk, you kind of jiggle around the jiggle and so forth. He's kicking the doorstop, you see, and I says, yes, and I said the doorstop is more appropriate than the door. The doorstop was a hemisphere, yellowish metal, gold as a matter of fact. It was a gold hemisphere about so big. What had happened was we needed to do an experiment to see how many neutrons were reflected by different materials in order to save the neutrons so we didn't use so much plutonium. We had tested many different materials. We had tested platinum, we had tested zinc, we had tested brass, we had tested gold. So in making the tests with the gold, we had these pieces of gold and somebody had the clever idea to use that great ball of gold for a doorstop for the door that contained the plutonium, which is quite appropriate.

After the thing went off and we heard about it, there was tremendous excitement at Los Alamos. Everybody had parties, we all ran around. I sat on the end of a jeep and beat drums and so on. Except for one man that I remember. [It] was Bob Wilson, who got me into it in the first place. He's

sitting there moping. I said, "What are you moping about?" He said, "It's a terrible thing that we made." I said, "But you started it, you got us into it." You see, what happened to me, what happened to the rest of us is we *started* for a good reason but then we're working very hard to do something, and to accomplish it, it's a pleasure, it's excitement. And you stop to think, you know, you just stop. After you thought at the beginning, you just stop. So he was the only one who was still thinking about it, at that particular moment. I returned to civilization shortly after that and went to Cornell to teach, and my first impression was a very strange one and I can't understand it anymore but I felt very strongly then. I'd sat in a restaurant in New York, for example, and I looked out at the buildings and how far away, I would think, you know, how much the radius of the Hiroshima bomb damage was and so forth. How far down there was down to 34th Street? All those buildings, all smashed and so on. And I got a very strange feeling. I would go along and I would see people building a bridge. Or they'd be making a new road, and I thought, they're *crazy,* they just don't understand, they don't understand. Why are they making new things, it's so useless? But fortunately, it's been useless for 30 years now isn't it, almost, maybe we'll make 30 years. So I've been wrong for 30 years about its being useless making bridges and I'm glad that those other people were able to go ahead. But my first reaction after I was finished with this thing was it's useless to make anything. Thank you very much.

Question: What about your story about some safe?

Feynman: Well, there's a lot of stories about safes. If you give me ten minutes, I'll tell you three stories about safes. All right? The motivation for me to open the filing cabinet, pick the lock, became my interest in the safety of the whole thing. Somebody had told me how to pick locks. Then they got fil-

ing cabinets which had safe combinations. One of my diseases, one of my things in life is that anything that is secret I try to undo. And so those locks to those filing cabinets, made by the Mosler Lock Company, in which we put our documents after that—everybody had them—they represented a challenge to me. How the hell to open them?! So I worked on them and I worked on them. There's all kinds of stories about how you can feel the numbers and listen to things and so on. That's true; I understand it, very well. For old-fashioned safes. They had a new design so that nothing would be pushing against the wheels while you were trying them. I won't go into the technical details, but none of the old methods would work. I read books by locksmiths. Books by locksmiths always say in the beginning how they opened the locks, the greatest thing, the woman is underwater, the safe is underwater and the woman is drowning or something, and he opened the safe. I don't know, crazy story. And then in the back they tell you how they do it and they don't tell you anything sensible; it doesn't sound like they could really open safes that way. Like *guess* the combination on the basis of the psychology of the person who owns it! So I always figured they'd keep it a secret. Anyway, I kept working. And so like a kind of disease, I kept working on these things until I found out a few things. First I found out how big a range you need to open the combination, how close you have to be. And then I invented a system by which you could try all the combinations that you have to try. Eight thousand as it turned out, because you could be within two of every number. Then it turns out that it's every fifth number out of a hundred and twenty thousand . . . eight thousand combinations. And then I worked out a scheme by which I could try numbers without altering a number that I already set, by correctly moving the wheels, so that I could do it in eight hours, try all the combinations. And

then I discovered still further that—this took me about two years of researching—I had nothing to do up there, you see, and I was fiddling—finally I discovered a way that it's easy to take the numbers, the back two numbers, the last two numbers of the combination off the safe if the safe is open. If the drawer is pulled out you could turn the number and see the bolt go up and play around and find out what makes it, what number it comes back at and stuff like that. With a little trickery you can get the combination off. So I used to practice it like a cardsharp practices cards, you know, all the time, all the time. Quicker and quicker and more and more unobtrusively I would come in and I would talk to some guy and I'd sort of lean against his filing cabinet, just like I'm playing with this watch now; you wouldn't even notice I'm doing anything. I'm not doing anything. I would just play with the dial, that's all, just play with the dial. But I was taking the two numbers off! And I go back to my office and I write the two numbers down. The last two numbers of the three. Now, if you have the last two numbers, it takes a minute to try the first number; there's only twenty possibilities and it's open. OK?

So, I got an excellent reputation for safecracking. They would say to me, "Mr. Schmultz is out of town, we need a document from his safe. Can you open it?" I'd say, "Yes, I can open it; I have to go get my tools" (I don't need any tools). I go to my office and I'd look at the number of his safe. I had the last two numbers. I had everybody's safe numbers in my office. I put a screwdriver in my back pocket, to account for the tool I claimed I needed. I go back to the room and I would close the door. The attitude is that this business about how you open safes is not something that everybody should know because it makes everything very unsafe, it's very dangerous to have everybody know how to do this. So I close the door and then I sit down and I read a magazine, or do some-

thing. I'd average about 20 minutes of doing nothing, and then I'd open it, you see, well, I opened it right away to see that everything was all right and then I'd sit there for 20 minutes to give myself a good reputation that it wasn't too easy, there was no trick to it, no trick to it. And then I'd come out, you know, sweating a bit, and say "It's open. There you are," and so forth. OK?

Also, at one particular moment, I did open a safe purely by accident, and that helped to reinforce my reputation. It was a sensation, it was pure luck, the same kind of luck I had with the blueprints. But after the war was over, I've got to tell you these stories now because after the war was over I went back to Los Alamos to finish some papers and there I did some safe opening which–I could write a safecracker book *better* than any safecracker book. In the beginning it would explain how I opened the safe absolutely cold without knowing the combination, which contained *more* secret things than any safe that's ever been opened. I opened the safe that contained behind it the secret of the atomic bomb, *all* the secrets, the formulas, the rates at which neutrons are liberated from uranium, how much uranium you need to make a bomb, all the theories, all the calculations, the WHOLE DAMN THING!

This is the way it was done. All right? I was trying to write a report. I needed this report. It was a Saturday; I thought everybody worked. I thought it was like Los Alamos *used* to be. So I went down to get it from the library. The library at Los Alamos had all these documents. There was a great vault with a great knob of a different kind I didn't know anything about. Filing cabinets I understood, but I was only an expert on filing cabinets. Not only that but there were guards walking back and forth in front with guns. You can't get that one open, OK? But I think, wait! Old Freddy DeHoffman in the declassification section, he's in charge of declassifying docu-

ments. Which documents now can be declassified? And so he had to run down to the library and back so often, he got tired of it. And he got a brilliant idea. He would get a copy made of every document in the Los Alamos library. And he'd stick it in his file, he had *nine* filing cabinets, one right next to the other in two rooms, *full* of all the documents of Los Alamos and I knew he had that. So I'll go up to DeHoffman and I'll ask him to borrow the documents, from him, he's got a copy. So I went up to his office. The office door is open. It looks like he's coming back, the light is lit; looks like he's coming back any minute. So I wait. And as always when I'm waiting, I diddled the knobs. I tried 10-20-30, didn't work. I tried 20-40-60, didn't work. Tried everything. I'm waiting, nothing to do. Then I begin to think, you know, those locksmith people, I had never been able to figure out how to open them cleverly. Maybe they don't either, maybe all the stuff they're telling me about psychology is right. I'm going to open this one by psychology. First thing, the book says, "The secretary is very nervous that she will forget the combination." She's been told the combination. She might forget and the boss might forget—she has to know. So she nervously writes it somewhere. Where? List of places were a secretary might write combinations. OK? Starts out with, most clever thing, starts right out with—you open the drawer and on the wood along the side of the drawer, the outside, is written carelessly a number, like as if it is an invoice number. That's the combination number. So. It's on the side of the desk. OK? I remembered that, it's in the book. Desk drawer is locked, I picked the lock easy, I opened the lock right away, pull out the drawer, look along the wood—nothing. It's all right, it's all right. There's a lot of papers in the drawer. I fish around among the papers and finally I find it, a nice little piece of paper which has the Greek alphabet. Alpha, beta, gamma,

delta, and so forth, carefully printed. The secretaries have to know how to make those letters and how to call them when they're talking about them, right? So they all had, each one had a copy of the thing. *But*—carelessly scrawled across the top is π is equal to 3.14159. Well, why does she need the numerical value of π, she's not computing anything? So I go up to the safe. Honest, it's honest, right? It's just like in the book. I'm just telling you how it was done. I walk up to the safe. 31-41-59. Doesn't open. 13-14-95—doesn't open. 95-14-13, doesn't open. 14-31, twenty minutes I'm turning π upside down. Nothing happens. So I start walking out of the office and I remember the book about the psychology and I said, you know, but it's true. Psychologically, I'm right. DeHoffman is *just* the kind of a guy to use a mathematical constant for his safe combination. So the other important mathematical constant is *e*. So I walk back to the safe, 27-18-28, click, clock, it opens. I checked by the way, that *all* the combinations were the same. Well, there's another lot of stories about it but it's getting late and that's a good one, so we'll let it go at that.

4

What Is and What Should Be the Role of Scientific Culture in Modern Society

Here is a talk Feynman gave to an audience of scientists at the Galileo Symposium in Italy, in 1964. With frequent acknowledgments and references to the great work and intense anguish of Galileo, Feynman speaks on the effect of science on religion, on society, and on philosophy, and warns that it is our capacity to doubt that will determine the future of civilization.

I am Professor Feynman, in spite of this suit-coat. I usually give lectures in shirtsleeves, but when I started out of the hotel this morning my wife said, "You must wear a suit." I said, "But I usually give lectures in shirtsleeves." She said, "Yes, but this time you don't know what you're talking about so you had better make a good impression. . . ." So, I got a coat.

The Pleasure of Finding Things Out

I am going to talk about the topic that was given me by Professor Bernardini.* I would like to say, at the very beginning, that, in my opinion, to find the proper place of scientific culture in modern society is not to solve the problems of modern society. There are a large number of problems that have nothing much to do with the position of science in society, and it is a dream to think that to simply decide on one aspect of how ideally science and society should be matched is somehow or other to solve all the problems. So, please understand that, although I will suggest some modifications of the relationship, I do not expect these modifications to be the solution to society's problems.

This modern society seems to be threatened by a number of serious threats, and the one that I would like to concentrate on and which will be in fact the central theme, although there will be a lot of subsidiary little items, the central theme of my discussion is that I believe that one of the greatest dangers to modern society is the possible resurgence and expansion of the ideas of thought control; such ideas as Hitler had, or Stalin in his time, or the Catholic religion in the Middle Ages, or the Chinese today. I think that one of the greatest dangers is that this shall increase until it encompasses all of the world.

Now, in discussing the relation of science to the scientific culture of society, the first thing that comes to mind immediately is, of course, the most obvious thing, which is the applications of science. The applications are culture, too. However, I am not going to talk about the applications, but not for any good reasons. I appreciate that all the popular discussions on the subject of the relation of science to society revolve around the applications almost completely, and furthermore that the moral questions that scientists have about

*Chairman of the conference. *Ed.*

the kind of work that they do also usually involve the applications. Nevertheless, I will not speak about them because there are a number of other items which are not spoken about by so many other people, and so for the fun of it I would like to talk in a slightly different direction.

I will, however, say about the applications that, as you all appreciate, science creates a power through its knowledge, a power to do things: You are able to do things after you know something scientifically. But the science does not give instructions with this power as to how to do good against how to do evil. Let us put it a very simple way: There are no instructions along with the power, and the question of applying the science or not is essentially the problem of organizing the applications in a way that doesn't do too much harm and does as much good as possible. But, of course, sometimes people in science try to say it is not their responsibility, because the application is just the power to do; it is independent of what you do with it. But it certainly is in some sense true that to create for mankind the power to control this is good, probably, in spite of the difficulties that he has in trying to figure out how to control the power to do himself good rather than evil.

May I say, too, that although many of us here are physicists, and most of us think of the serious problems of society in terms of physics, I believe most assuredly that the next science to find itself in moral difficulties with its applications is biology, and if the problems of physics relative to science seem difficult, the problems of the development of biological knowledge will be fantastic. These possibilities were hinted at, for example, in the book by Huxley, *Brave New World,* but you can think of a number of things. For example, if energy in the far future can be supplied freely and easily by physics, then it is a matter of mere chemistry to put together the atoms in such a way as to produce food, from energy that the

atoms have conserved, so that you can produce as much food as there are waste products from human beings; and there is therefore a conservation of material and no food problem. There will be serious social problems when we find out how to control heredity, as to what kind of control, good or bad, to use. Suppose that we were to discover the physiological basis of happiness or other feelings, such as the feeling of ambition, and suppose that we could then control whether somebody feels ambitious or does not feel ambitious. And, finally, there is death.

It is one of the most remarkable things that in all of the biological sciences there is no clue as to the necessity of death. If you say we want to make perpetual motion, we have discovered enough laws as we studied physics to see that it is either absolutely impossible or else the laws are wrong. But there is nothing in biology yet found that indicates the inevitability of death. This suggests to me that it is not at all inevitable, and that it is only a matter of time before the biologists discover what it is that is causing us the trouble and that that terrible universal disease or temporariness of the human's body will be cured. Anyhow, you can see that there will be problems of a fantastic magnitude coming from biology.

Now I will talk in a different direction.

Besides the applications there are ideas, and the ideas are of two kinds. One of them is the product of the science itself, that is, a worldview which the science produces. This is in some ways the most beautiful part of the whole thing. Some people think, no, the methods of science are the thing. Well, it depends on whether you like the ends or the means, but the means were to produce some wonderful ends and I will not bore you (well, I wouldn't bore you if I could do it right) with the details. But you all know something about the wonders of science—it isn't a popular audience I am talking

to–so I won't try to make you enthusiastic once again with the facts about the world: the fact that we are all made of atoms, the enormous ranges of time and space that there are, the position of ourselves historically as the result of a remarkable series of evolution. The position of ourselves in the evolutionary sequence; and further, the most remarkable aspect of our scientific worldview is its universality in this sense that although we talk about our being specialists, we are really not. One of the most promising hypotheses in all of biology is that everything the animals do or that living creatures do can be understood in terms of what atoms can do, that is, in terms of physical laws, ultimately, and the perpetual attention to this possibility–so far no exception has been demonstrated–has again and again made suggestions as to how the mechanisms actually occur. So that the fact that our knowledge is in fact universal is something that is not completely appreciated, that the position of the theories are so complete that we hunt for exceptions and we find them very hard to find–in the physics at least–and the great expense of all these machines and so on is to find some exception to what is already known. And, otherwise, that is another aspect of the fact that the world is so wonderful in the sense that stars are made of the same atoms as the cows and as ourselves, and as stones.

From time to time we all try to communicate to our unscientific friends this worldview–and we get into difficulty most often because we get confused in trying to explain to them the latest questions, such as the meaning of the conservation of CP,*

*Conservation of charge and parity, one of the fundamental conservation laws of physics, which says that the total electric charge and parity, an intrinsic symmetry property of subatomic particles, going into an interaction will be the same coming out of that interaction. *Ed.*

whereas they don't know anything about the most preliminary things. For four hundred years since Galileo we have been gathering information about the world which they don't know. Now we are working on something way out, and at the limits of scientific knowledge. And the things that appear in the newspaper and that seem to excite the adult imagination are always those things which they cannot possibly understand, because they haven't learned anything at all of the much more interesting well-known [to scientists] things that people have found out before. It's not the case with children, thank goodness, for a while—at least until they become adults.

I say, and I think you must all know from experience, that people—I mean the average person, the great majority of people, the enormous majority of people—are woefully, pitifully, absolutely ignorant of the science of the world that they live in, and they can stay that way. I don't mean to say the heck with them, what I mean is that they are able to stay that way without it worrying them at all—only mildly—so from time to time when they see CP mentioned in the newspaper they ask what it is. And an interesting question of the relation of science to modern society is just that—why is it possible for people to stay so woefully ignorant and yet reasonably happy in modern society when so much knowledge is unavailable to them?

Incidentally, about knowledge and wonder, Mr. Bernardini said we shouldn't teach wonders but knowledge.

It may be a difference in the meaning of the words. I think we should teach them wonders and that the purpose of knowledge is to appreciate wonders even more. And that the knowledge is just to put into correct framework the wonder that nature is. However, he would probably agree that I just shifted some words around and that meaning trickled into

the conversation. At any rate, I want to answer the question as to why people can remain so woefully ignorant and not get into difficulties in modern society. The answer is that science is irrelevant. And I will explain what I mean in just a minute. It isn't that it has to be, but that we let it be irrelevant to society. I will come back to that point.

The other aspects of science that are important and that have some problem of a relation to society, beside the applications and the actual facts that are discovered, are the ideas and the techniques of scientific investigation: the means, if you will. Because I think that it is hard to understand why the discovery of these means, which seem so self-evident and obvious, weren't discovered earlier; simple ideas which, if you just try them, you see what happens and so forth. It is probably that the human mind evolved from that of an animal; and it evolves in a certain way [such] that it is like any new tool, in that it has its diseases and difficulties. It has its troubles, and one of the troubles is that it gets polluted by its own superstitions, it confuses itself, and the discovery was finally made of a way to keep it sort of in line so that scientists can make a little progress in some direction rather than to go around in circles and force themselves into a hold. And I think that this is, of course, the appropriate time to discuss this matter because the beginnings of this new discovery were at the time of Galileo. These ideas and techniques, of course, you all know. I'll just review them; it's again one of those things that for a lay audience you go into great detail; I just mention them so you appreciate what I am talking about more specifically.

The first is the matter of judging evidence—well, the first thing really is, before you begin you must not know the answer. So you begin by being uncertain as to what the answer is. This is very, very important, so important that I would

like to delay that aspect, and talk about that still further along in my speech. The question of doubt and uncertainty is what is necessary to begin; for if you already know the answer there is no need to gather any evidence about it. Well, being uncertain, the next thing is to look for evidence, and the scientific method is to begin with trials. But another way and a very important one that should not be neglected and that is very vital is to put together ideas to try to enforce a logical consistency among the various things that you know. It is a very valuable thing to try to connect this, what you know, with that, that you know, and try to find out if they are consistent. And the more activity in the direction of trying to put together the ideas of different directions, the better it is.

After we look for the evidence we have to judge the evidence. There are the usual rules about the judging the evidence; it's not right to pick only what you like, but to take all of the evidence, to try to maintain some objectivity about the thing—enough to keep the thing going—not to ultimately depend upon authority. Authority may be a hint as to what the truth is, but is not the source of information. As long as it's possible, we should disregard authority whenever the observations disagree with it. And finally, the recording of results should be done in a disinterested way. That's a very funny phrase which always bothers me—because it means that after the guy's all done with the thing, he doesn't give a darn about the results, but that isn't the point. Disinterest here means that they are not reported in such a way as to try to influence the reader into an idea that's different than what the evidence indicates.

And you all appreciate these various aspects.

Now all this, all these ideas, and all the techniques are in the spirit of Galileo. The man whose birthday we are cele-

brating had a great deal to do with the development and the spreading and, most importantly, the demonstration of the power of these ways of looking at things. In any centennial, or quattro-centennial likewise, one always gets the feeling sooner or later: I wonder if the man were here now and we showed him the world, what he would say. Of course, you say, that's a corny thing to do and you can't do that in a speech, but that's what I am going to do. Suppose Galileo were here and we were to show him the world today and try to make him happy, or see what he finds out. And we would tell him about the questions of evidence, those methods of judging things which he developed. And we would point out that we are still in exactly the same tradition, we follow it exactly—even to the details of making numerical measurements and using those as one of the better tools, in the physics at least. And that the sciences have developed in a very good way directly and continuously from his original ideas, in the same spirit he developed. And as a result there are no more witches and ghosts.

Actually I say [that the quantitative method works very well] in science, but that is in fact almost a definition of science today; the sciences that Galileo was worried about, the physics, mechanics and such things, have of course developed, but the same techniques worked in biology, in history, geology, anthropology, and so on. We know a great deal about the past history of man, the past history of animals, and of the earth, through very similar techniques. With somewhat similar success, but not quite as complete because of the difficulties, the same systems work in economics. But there are places where only lip service is paid to the forms—in which many people just go through the motions. I would be ashamed to tell Mr. Galileo, but it doesn't really work very well, for example, in the social sciences. For example, my own

personal experience–as you will realize, there is an awful lot of studying of the methods of education going on, particularly of the teaching of arithmetic–but if you try to find out what is really known about what is the better way to teach arithmetic than some other way, you will discover that there is an enormous number of studies and a great deal of statistics, but they are all disconnected from one another and they are mixtures of anecdotes, uncontrolled experiments, and very poorly controlled experiments, so that there is very little information as a result.

And now finally, as I'd like to show Galileo our world, I must show him something with a great deal of shame. If we look away from the science and look at the world around us, we find out something rather pitiful: that the environment that we live in is so actively, intensely unscientific. Galileo could say: "I noticed that Jupiter was a ball with moons and not a god in the sky. Tell me, what happened to the astrologers?" Well, they print their results in the newspapers, in the United States at least, in every daily paper every day. Why do we still have astrologers? Why can someone write a book like *Worlds in Collision* by somebody with a name beginning with a "V," it's a Russian name? Huh? Vininkowski?* And how did it become popular? What is all this nonsense about Mary Brody, or something? I don't know, that was crazy stuff. There is always some crazy stuff. There is an infinite amount of crazy stuff, which, put another way, is that the environment is actively, intensely unscientific. There is talk about telepathy still, although it's dying out. There is faith-healing galore, all over. There is a whole religion of faith-healing. There's a miracle at Lourdes where healing goes on. Now, it might be true that as-

*Actually it was Immanuel Velikovsky: *Worlds in Collision* (Doubleday, New York, 1950). *Ed.*

trology is right. It might be true that if you go to the dentist on the day that Mars is at right angles to Venus, that it is better than if you go on a different day. It might be true that you can be cured by the miracle of Lourdes. But if it is true it ought to be investigated. Why? To improve it. If it is true then maybe we can find out if the stars do influence life; that we could make the system more powerful by investigating statistically, scientifically judging the evidence objectively, more carefully. If the healing process works at Lourdes, the question is how far from the site of the miracle can the person, who is ill, stand? Have they in fact made a mistake and the back row is really not working? Or is it working so well that there is plenty of room for more people to be arranged near the place of the miracle? Or is it possible, as it is with the saints which have recently been created in the United States–there is a saint who cured leukemia apparently indirectly–that ribbons that are touched to the sheet of the sick person (the ribbon having previously touched some relic of the saint) increase the cure of leukemia–the question is, is it gradually being diluted? You may laugh, but if you believe in the truth of the healing, then you are responsible to investigate it, to improve its efficiency and to make it satisfactory instead of cheating. For example, it may turn out that after a hundred touches it doesn't work anymore. Now it's also possible that the results of this investigation have other consequences, namely, that nothing is there.

And another thing that bothers me, I might as well mention, are the things that the theologians in modern times can discuss, without feeling ashamed of themselves. There are many things that they can discuss that they need not feel ashamed of themselves, but some of the things that go on in the conferences on religion, and the decisions that have to be made, are ridiculous in modern times. I would like to explain

that one of the difficulties, and one of the reasons why this can keep going, is that it is not realized what a profound modification of our worldview would result, if just one example of one of these things would really work. The whole idea, if you could establish the truth, not of the whole idea of astrology but just one little item, could have a fantastic influence on our understanding of the world. And so the reasons we laugh a little bit is that we are so confident of our view of the world that we are sure they aren't going to contribute anything. On the other hand, why don't we get rid of it? I will come to why we don't get rid of it in a minute, because science is irrelevant [to astrology], as I said before.

Now I am going to mention still another thing which is a little more doubtful, but still I believe that in the judging of evidence, the reporting of evidence and so on, there is a kind of responsibility which the scientists feel toward each other which you can represent as a kind of morality. What's the right way and the wrong way to report results? Disinterestedly, so that the other man is free to understand precisely what you are saying, and as nearly as possible not covering it with your desires. That this is a useful thing, that this is a thing which helps each of us to understand each other, in fact to develop in a way that isn't personally in our own interest, but for the general development of ideas, is a very valuable thing. And so there is, if you will, a kind of scientific morality. I believe, hopelessly, that this morality should be extended much more widely; this idea, this kind of scientific morality, that such things as propaganda should be a dirty word. That a description of a country made by the people of another country should describe that country in a disinterested way. What a miracle–that's worse than a miracle at Lourdes! Advertising, for example, is an example of a scientifically immoral description of the products. This immoral-

ity is so extensive that one gets so used to it in ordinary life, that you do not appreciate that it is a bad thing. And I think that one of the important reasons to increase the contact of scientists with the rest of society is to explain, and to kind of wake them up to this permanent attrition of cleverness of the mind that comes from not having information, or [not] having information always in a form which is interesting.

There are other things in which scientific methods would be of some value; they are perfectly obvious but they get more and more difficult to discuss–such things as making decisions. I do not mean that it should be done scientifically, such as [the way] in the United States that the Rand Company sits down and makes arithmetical calculations. That reminds me of my sophomore days at college in which, in discussing women, we discovered that by using electrical terminology–impedance, reluctance, resistance–that we had a deeper understanding of the situation. The other thing that gives a scientific man the creeps in the world today are the methods of choosing leaders–in every nation. Today, for example, in the United States, the two political parties have decided to employ public relations men, that is, advertising men, who are trained in the necessary methods of telling the truth and lying in order to develop a product. This wasn't the original idea. They are supposed to discuss situations and not just make up slogans. It's true, if you look in history, however, that choosing political leaders in the United States has been on many different occasions based on slogans. (I am sure that each party now has million-dollar bank accounts and there are going to be some very clever slogans.) But I can't do a sum-up of all this stuff now.

I kept saying that the science was irrelevant. That sounds strange and I would like to come back to it. Of course it is relevant, because of the fact that it is relevant to astrology; be-

cause if we understand the world the way we do, we cannot understand how the astrological phenomena can take place. And so that is relevant. But for people who believe in astrology there is no relevance, because the scientist never bothers to argue with them. The people who believe in faith healing have not to worry about science at all, because nobody argues with them. You don't have to learn science if you don't feel like it. So you can forget the whole business if it is too much mental strain, which it usually is. Why can you forget the whole business? Because we don't do anything about it. I believe that we must attack these things in which we do not believe. Not attack by the method of cutting off the heads of the people, but attack in the sense of discuss. I believe that we should demand that people try in their own minds to obtain for themselves a more consistent picture of their own world; that they not permit themselves the luxury of having their brain cut in four pieces or two pieces even, and on one side they believe this and on the other side they believe that, but never try to compare the two points of view. Because we have learned that, by trying to put the points of view that we have in our head together and comparing one to the other, we make some progress in understanding and in appreciating where we are and what we are. And I believe that science has remained irrelevant because we wait until somebody asks us questions or until we are invited to give a speech on Einstein's theory to people who don't understand Newtonian mechanics, but we never are invited to give an attack on faith healing, or on astrology—on what is the scientific view of astrology today.

I think that we must mainly write some articles. Now what would happen? The person who believes in astrology will have to learn some astronomy. The person who believes in faith healing might have to learn some medicine, because of

the arguments going back and forth; and some biology. In other words, it will be necessary that science become relevant. The remark which I read somewhere, that science is all right so long as it doesn't attack religion, was the clue that I needed to understand the problem. As long as it doesn't attack religion it need not be paid attention to and nobody has to learn anything. So it can be cut off from modern society except for its applications, and thus be isolated. And then we have this terrible struggle to try to explain things to people who have no reason to want to know. But if they want to defend their own point of view, they will have to learn what yours is a little bit. So I suggest, maybe incorrectly and perhaps wrongly, that we are too polite. There was in the past an era of conversation on these matters. It was felt by the church that Galileo's views attacked the church. It is not felt by the church today that the scientific views attack the church. Nobody is worrying about it. Nobody attacks; I mean, nobody writes trying to explain the inconsistencies between the theological views and the scientific views held by different people today—or even the inconsistencies sometimes held by the same scientist between his religious and scientific beliefs.

Now the next subject, and the last main subject that I want to talk about, is the one I really consider the most important and the most serious. And that has to do with the question of uncertainty and doubt. A scientist is never certain. We all know that. We know that all our statements are approximate statements with different degrees of certainty; that when a statement is made, the question is not whether it is true or false but rather how likely it is to be true or false. "Does God exist?" "When put in the questional form, how likely is it?" It makes such a terrifying transformation of the religious point of view, and that is why the religious point of view is unscientific. We must discuss each question within the uncertain-

ties that are allowed. And as evidence grows it increases the probability perhaps that some idea is right, or decreases it. But it never makes absolutely certain one way or the other. Now we have found that this is of paramount importance in order to progress. We absolutely must leave room for doubt or there is no progress and there is no learning. There is no learning without having to pose a question. And a question requires doubt. People search for certainty. But there *is* no certainty. People are terrified—how can you live *and not know*? It is not odd at all. You only think you know, as a matter of fact. And most of your actions are based on incomplete knowledge and you really don't know what it is all about, or what the purpose of the world is, or know a great deal of other things. It is possible to live and not know.

Now the freedom to doubt, which is absolutely essential for the development of the sciences, was born from a struggle with the constituted authorities of the time who had a solution to every problem, namely, the church. Galileo is a symbol of that struggle—one of the most important strugglers. And although Galileo himself apparently was forced to recant, nobody takes the confession seriously. We do not feel that we should follow Galileo in this way and that we should all recant. In fact, we consider the recantation as a foolishness—that the church asked for such a foolishness that we see again and again; and we feel sympathetic to Galileo as we feel sympathetic to the musicians and the artists of the Soviet Union who had to recant, and fortunately in apparently somewhat fewer numbers in recent times. But the recantation is a meaningless thing, no matter how cleverly it is organized. It is perfectly obvious to people from the outside that it is nothing to consider, and that Galileo's recantation is not something that we need to discuss as demonstrating anything about Galileo, except that perhaps he was an old man and

that the church was very powerful. The fact that Galileo was right is not essential to this discussion. The fact that he was trying to be suppressed is, of course.

We are all saddened when we look at the world and see what few accomplishments we have made, compared to what we feel are the potentialities of human beings. People in the past, in the nightmare of their times, had dreams for the future. And now that the future has materialized we see that in many ways the dreams have been surpassed, but in still more ways many of our dreams of today are very much the dreams of people of the past. There have, in the past, been great enthusiasms for one or another's method of solving a problem. One was that education should become universal, for then all men would become Voltaires, and then we would have everything straightened out. Universal education is probably a good thing, but you could teach bad as well as good–you [could] teach falsehood as well as truth. The communication between nations as it develops through a technical development of science should certainly improve the relations between nations. Well, it depends what you communicate. You can communicate truth and you can communicate lies. You can communicate threats or kindnesses. There was a great hope that the applied sciences would free man of his physical struggles, and particularly in medicine it seems, for example, that all is to the good. Yes, but while we are talking, scientists are working in hidden secret laboratories trying to develop, as best they can, diseases which the other man can't cure. Perhaps today we have the dream that economic satiation of all men is the solution to the problem. I mean everybody should have enough stuff. I don't mean, of course, that we shouldn't try to do that. I don't mean, by what I'm saying, that we should not educate, or that we should not communicate, or that we shouldn't produce economic satiation.

But that this is the solution all by itself, of all problems, is doubtful. Because in those places where we have a certain degree of economic satisfaction, we have a whole host of new problems, or probably old problems that just look a little different because we happen to know enough about history.

So today we are not very well off, we don't see that we have done too well. Men, philosophers of all ages, have tried to find the secret of existence, the meaning of it all. Because if they could find the real meaning of life, then all this human effort, all this wonderful potentiality of human beings, could then be moved in the correct direction and we would march forward with great success. So therefore we tried these different ideas. But the question of the meaning of the whole world, of life, and of human beings, and so on, has been answered very many times by very many people. Unfortunately all the answers are different; and the people with one answer look with horror at the actions and behavior of the people with another answer. Horror, because they see the terrible things that are done; the way man is being pushed into a blind alley by this rigid view of the meaning of the world. In fact, it is really perhaps by the fantastic size of the horror that it becomes clear how great are the potentialities of human beings, and it is possibly this which makes us hope that if we could move things in the right direction, things would be much better.

What then is the meaning of the whole world? We do not know what the meaning of existence is. We say, as the result of studying all of the views that we have had before, we find that we do not know the meaning of existence; but in saying that we do not know the meaning of existence, we have probably found the open channel—if we will allow only that, as we progress, we leave open opportunities for alternatives, that we do not become enthusiastic for the fact, the knowl-

edge, the absolute truth, but remain always uncertain–[that we] "hazard it." The English, who have developed their government in this direction, call it "muddling through," and although a rather silly, stupid sounding thing, it is the most scientific way of progressing. To decide upon the answer is not scientific. In order to make progress, one must leave the door to the unknown ajar–ajar only. We are only at the beginning of the development of the human race; of the development of the human mind, of intelligent life–we have years and years in the future. It is our responsibility not to give the answer today as to what it is all about, to drive everybody down in that direction and to say: "This is a solution to it all." Because we will be chained then to the limits of our present imagination. We will only be able to do those things that we think today are the things to do. Whereas, if we leave always some room for doubt, some room for discussion, and proceed in a way analogous to the sciences, then this difficulty will not arise. I believe, therefore, that although it is not the case today, that there may some day come a time, I should hope, when it will be fully appreciated that the power of government should be limited; that governments ought not to be empowered to decide the validity of scientific theories, that that is a ridiculous thing for them to try to do; that they are not to decide the various descriptions of history or of economic theory or of philosophy. Only in this way can the real possibilities of the future human race be ultimately developed.

5

There's Plenty of Room at the Bottom

In this famous talk to the American Physical Society on December 29, 1959, at Caltech, Feynman, the "father of nanotechnology," expounds, decades ahead of his time, on the future of miniaturization: how to put the entire Encyclopaedia Brittanica on the head of a pin, the drastic reduction in size of both biological and inanimate objects, and the problems of lubricating machines smaller than the period at the end of this sentence. Feynman makes his famous wager, challenging young scientists to construct a working motor no more than 1/64 of an inch on all sides.

An Invitation to Enter a New Field of Physics

I imagine experimental physicists must often look with envy at men like Kamerlingh-Onnes,* who discovered a field like low temperature, which seems to be bottomless and in which

*Heike Kamerlingh-Onnes (1853–1926), winner of the 1913 Physics Nobel Prize for investigations of the properties of matter at low temperatures, which led to the production of liquid helium. *Ed.*

one can go down and down. Such a man is then a leader and has some temporary monopoly in a scientific adventure. Percy Bridgman,* in designing a way to obtain higher pressures, opened up another new field and was able to move into it and to lead us all along. The development of ever higher vacuum was a continuing development of the same kind.

I would like to describe a field in which little has been done, but in which an enormous amount can be done in principle. This field is not quite the same as the others in that it will not tell us much of fundamental physics (in the sense of, "What are the strange particles?"), but it is more like solid-state physics in the sense that it might tell us much of great interest about the strange phenomena that occur in complex situations. Furthermore, a point that is most important is that it would have an enormous number of technical applications.

What I want to talk about is the problem of manipulating and controlling things on a small scale.

As soon as I mention this, people tell me about miniaturization, and how far it has progressed today. They tell me about electric motors that are the size of the nail on your small finger. And there is a device on the market, they tell me, by which you can write the Lord's Prayer on the head of a pin. But that's nothing; that's the most primitive, halting step in the direction I intend to discuss. It is a staggeringly small world that is below. In the year 2000, when they look back at this age, they will wonder why it was not until the year 1960 that anybody began seriously to move in this direction.

Why cannot we write the entire 24 volumes of the Encyclopaedia Brittanica *on the head of a pin?*

---

*(1882–1961) Winner of the 1946 Physics Nobel Prize for his invention of an apparatus for producing extremely high pressures, and further work in high pressure physics. *Ed.*

Let's see what would be involved. The head of a pin is a sixteenth of an inch across. If you magnify it by 25,000 diameters, the area of the head of the pin is then equal to the area of all the pages of the *Encyclopaedia Brittanica.* Therefore, all it is necessary to do is to reduce in size all the writing in the *Encyclopaedia* by 25,000 times. Is that possible? The resolving power of the eye is about 1/120 of an inch–that is roughly the diameter of one of the little dots on the fine half-tone reproductions in the *Encyclopaedia.* This, when you demagnify it by 25,000 times, is still 80 angstroms* in diameter–32 atoms across, in an ordinary metal. In other words, one of those dots still would contain in its area 1,000 atoms. So, each dot can easily be adjusted in size as required by the photoengraving, and there is no question that there is enough room on the head of a pin to put all of the *Encyclopaedia Brittanica.*

Furthermore, it can be read if it is so written. Let's imagine that it is written in raised letters of metal; that is, where the black is in the *Encyclopaedia,* we have raised letters of metal that are actually 1/25,000 of their ordinary size. How would we read it?

If we had something written in such a way, we could read it using techniques in common use today. (They will undoubtedly find a better way when we do actually have it written, but to make my point conservatively I shall just take techniques we know today.) We would press the metal into a plastic material and make a mold of it, then peel the plastic off very carefully, evaporate silica into the plastic to get a very thin film, then shadow it by evaporating gold at an angle against the silica so that all the little letters will appear clearly, dissolve the plastic away from the silica film, and then look through it with an electron microscope!

---

*One angstrom = one ten-billionth of a meter. *Ed.*

There is no question that if the thing were reduced by 25,000 times in the form of raised letters on the pin, it would be easy for us to read it today. Furthermore, there is no question that we would find it easy to make copies of the master; we would just need to press the same metal plate again into plastic and we would have another copy.

How Do We Write Small?

The next question is: How do we *write* it? We have no standard technique to do this now. But let's argue that it is not as difficult as it first appears to be. We can reverse the lenses of the electron microscope in order to demagnify as well as magnify. A source of ions, sent through the microscope lenses in reverse, could be focused to a very small spot. We could write with that spot like we write in a TV cathode ray oscilloscope, by going across in lines, and having an adjustment which determines the amount of material which is going to be deposited as we scan in lines.

This method might be very slow because of space charge limitations. There will be more rapid methods. We could first make, perhaps by some photo process, a screen which has holes in it in the form of the letters. Then we would strike an arc behind the holes and draw metallic ions through the holes; then we could again use our system of lenses and make a small image in the form of ions, which would deposit the metal on the pin.

A simpler way might be this (though I am not sure it would work): We take light and, through an optical microscope running backwards, we focus it onto a very small photoelectric screen. Then electrons come away from the screen where the light is shining. These electrons are focused down in size by the electron microscope lenses to impinge directly upon the

surface of the metal. Will such a beam etch away the metal if it is run long enough? I don't know. If it doesn't work for a metal surface, it must be possible to find some surface with which to coat the original pin so that, where the electrons bombard, a change is made which we could recognize later.

There is no intensity problem in these devices–not what you are used to in magnification, where you have to take a few electrons and spread them over a bigger and bigger screen; it is just the opposite. The light which we get from a page is concentrated onto a very small area so it is very intense. The few electrons which come from the photoelectric screen are demagnified down to a very tiny area so that, again, they are very intense. I don't know why this hasn't been done yet!

That's the *Encyclopaedia Brittanica* on the head of a pin, but let's consider all the books in the world. The Library of Congress has approximately 9 million volumes; the British Museum Library has 5 million volumes; there are also 5 million volumes in the National Library in France. Undoubtedly there are duplications, so let us say that there are some 24 million volumes of interest in the world.

What would happen if I print all this down at the scale we have been discussing? How much space would it take? It would take, of course, the area of about a million pinheads because, instead of there being just the 24 volumes of the *Encyclopaedia*, there are 24 million volumes. The million pinheads can be put in a square of a thousand pins on a side, or an area of about 3 square yards. That is to say, the silica replica with the paper-thin backing of plastic, with which we have made the copies, with all this information, is on an area of approximately the size of 35 pages of the *Encyclopaedia*. That is about half as many pages as there are in this magazine. All of the information which all of mankind has ever recorded in books can be carried around in a pamphlet in

your hand—and not written in code, but a simple reproduction of the original pictures, engravings, and everything else on a small scale without loss of resolution.

What would our librarian at Caltech say, as she runs all over from one building to another, if I tell her that, ten years from now, all of the information that she is struggling to keep track of—120,000 volumes, stacked from the floor to the ceiling, drawers full of cards, storage rooms full of the older books—can be kept on just one library card! When the University of Brazil, for example, finds that their library is burned, we can send them a copy of every book in our library by striking off a copy from the master plate in a few hours and mailing it in an envelope no bigger or heavier than any other ordinary air mail letter.

Now, the name of this talk is "There Is *Plenty* of Room at the Bottom"—not just "There Is Room at the Bottom." What I have demonstrated is that there is room—that you can decrease the size of things in a practical way. I now want to show that there is *plenty* of room. I will not now discuss how we are going to do it, but only what is possible in principle—in other words, what is possible according to the laws of physics. I am not inventing anti-gravity, which is possible someday only if the laws are not what we think. I am telling you what could be done if the laws *are* what we think; we are not doing it simply because we haven't yet gotten around to it.

Information on a Small Scale

Suppose that, instead of trying to reproduce the pictures and all the information directly in its present form, we write only the information content in a code of dots and dashes, or something like that, to represent the various letters. Each letter represents six or seven "bits" of information; that is, you

need only about six or seven dots or dashes for each letter. Now, instead of writing everything, as I did before, on the *surface* of the head of a pin, I am going to use the interior of the material as well.

Let us represent a dot by a small spot of one metal, the next dash by an adjacent spot of another metal, and so on. Suppose, to be conservative, that a bit of information is going to require a little cube of atoms 5 times 5 times 5–that is, 125 atoms. Perhaps we need a hundred and some odd atoms to make sure that the information is not lost through diffusion, or through some other process.

I have estimated how many letters there are in the *Encyclopaedia*, and I have assumed that each of my 24 million books is as big as an *Encyclopaedia* volume, and have calculated, then, how many bits of information there are (10^{15}). For each bit I allow 100 atoms. And it turns out that all of the information that man has carefully accumulated in all the books in the world can be written in this form in a cube of material one two-hundredths of an inch wide–which is the barest piece of dust that can be made out by the human eye. So there is *plenty* of room at the bottom! Don't tell me about microfilm!

This fact–that enormous amounts of information can be carried in an exceedingly small space–is, of course, well known to the biologists, and resolves the mystery which existed before we understood all this clearly, of how it could be that, in the tiniest cell, all of the information for the organization of a complex creature such as ourselves can be stored. All this information–whether we have brown eyes, or whether we think at all, or that in the embryo the jawbone should first develop with a little hole in the side so that later a nerve can grow through it–all this information is contained in a very tiny fraction of the cell in the form of long-chain

DNA molecules in which approximately 50 atoms are used for one bit of information about the cell.

Better Electron Microscopes

If I have written in a code, with 5 times 5 times 5 atoms to a bit, the question is: How could I read it today? The electron microscope is not quite good enough; with the greatest care and effort, it can only resolve about 10 angstroms. I would like to try and impress upon you while I am talking about all of these things on a small scale, the importance of improving the electron microscope by a hundred times. It is not impossible; it is not against the laws of diffraction of the electron. The wave length of the electron in such a microscope is only 1/20 of an angstrom. So it should be possible to see the individual atoms. What good would it be to see individual atoms distinctly?

We have friends in other fields—in biology, for instance. We physicists often look at them and say, "You know the reason you fellows are making so little progress?" (Actually I don't know any field where they are making more rapid progress than they are in biology today.) "You should use more mathematics, like we do." They could answer us—but they're polite, so I'll answer for them: "What you should do in order for *us* to make more rapid progress is to make the electron microscope 100 times better."

What are the most central and fundamental problems of biology today? They are questions like: What is the sequence of bases in the DNA? What happens when you have a mutation? How is the base order in the DNA connected to the order of amino acids in the protein? What is the structure of the RNA; is it single-chain or double-chain, and how is it related in its order of bases to the DNA? What is

the organization of the microsomes? How are proteins synthesized? Where does the RNA go? How does it sit? Where do the proteins sit? Where do the amino acids go in? In photosynthesis, where is the chlorophyll; how is it arranged; where are the carotenoids involved in this thing? What is the system of the conversion of light into chemical energy?

It is very easy to answer many of these fundamental biological questions; you just *look at the thing*! You will see the order of bases in the chain; you will see the structure of the microsome. Unfortunately, the present microscope sees at a scale which is just a bit too crude. Make the microscope one hundred times more powerful, and many problems of biology would be made very much easier. I exaggerate, of course, but the biologists would surely be very thankful to you—and they would prefer that to the criticism that they should use more mathematics.

The theory of chemical processes today is based on theoretical physics. In this sense, physics supplies the foundation of chemistry. But chemistry also has analysis. If you have a strange substance and you want to know what it is, you go through a long and complicated process of chemical analysis. You can analyze almost anything today, so I am a little late with my idea. But if the physicists wanted to, they could also dig under the chemists in the problem of chemical analysis. It would be very easy to make an analysis of any complicated chemical substance; all one would have to do would be to look at it and see where the atoms are. The only trouble is that the electron microscope is one hundred times too poor. (Later, I would like to ask the question: Can the physicists do something about the third problem of chemistry—namely, synthesis? Is there a *physical* way to synthesize any chemical substance?)

The reason the electron microscope is so poor is that the f-value of the lenses is only 1 part to 1,000; you don't have a big enough numerical aperture. And I know that there are theorems which prove that it is impossible, with axially symmetrical stationary field lenses, to produce an f-value any bigger than so and so; and therefore the resolving power at the present time is at its theoretical maximum. But in every theorem there are assumptions. Why must the field be symmetrical? I put this out as a challenge: Is there no way to make the electron microscope more powerful?

The Marvelous Biological System

The biological example of writing information on a small scale has inspired me to think of something that should be possible. Biology is not simply writing information; it is *doing something* about it. A biological system can be exceedingly small. Many of the cells are very tiny, but they are very active; they manufacture various substances; they walk around; they wiggle; and they do all kinds of marvelous things–all on a very small scale. Also, they store information. Consider the possibility that we too can make a thing very small which does what we want–that we can manufacture an object that maneuvers at that level!

There may even be an economic point to this business of making things very small. Let me remind you of some of the problems of computing machines. In computers we have to store an enormous amount of information. The kind of writing that I was mentioning before, in which I had everything down as a distribution of metal, is permanent. Much more interesting to a computer is a way of writing, erasing, and writing something else. (This is usually because we don't want to waste the material on which we have just written. Yet if we

could write it in a very small space, it wouldn't make any difference; it could just be thrown away after it was read. It doesn't cost very much for the material.)

Miniaturizing the Computer

I don't know how to do this on a small scale in a practical way, but I do know that computing machines are very large; they fill rooms. Why can't we make them very small, make them of little wires, little elements—and by little, I mean *little.* For instance, the wires should be 10 or 100 atoms in diameter, and the circuits should be a few thousand angstroms across. Everybody who has analyzed the logical theory of computers has come to the conclusion that the possibilities of computers are very interesting—if they could be made to be more complicated by several orders of magnitude. If they had millions of times as many elements, they could make judgments. They would have time to calculate what is the best way to make the calculation that they are about to make. They could select the method of analysis which, from their experience, is better than the one that we would give to them. And in many other ways, they would have new qualitative features.

If I look at your face I immediately recognize that I have seen it before. (Actually, my friends will say I have chosen an unfortunate example here for the subject of this illustration. At least I recognize that it is a *man* and not an *apple.*) Yet there is no machine which, with that speed, can take a picture of a face and say even that it is a man; and much less that it is the same man that you showed it before—unless it is exactly the same picture. If the face is changed; if I am closer to the face; if I am further from the face; if the light changes—I recognize it anyway. Now, this little computer I carry in my head is eas-

ily able to do that. The computers that we build are not able to do that. The number of elements in this bone box of mine are enormously greater than the number of elements in our "wonderful" computers. But our mechanical computers are too big; the elements in this box are microscopic. I want to make some that are *sub*microscopic.

If we wanted to make a computer that had all these marvelous extra qualitative abilities, we would have to make it, perhaps, the size of the Pentagon. This has several disadvantages. First, it requires too much material; there may not be enough germanium in the world for all the transistors which would have to be put into this enormous thing. There is also the problem of heat generation and power consumption; TVA would be needed to run the computer. But an even more practical difficulty is that the computer would be limited to a certain speed. Because of its large size, there is finite time required to get the information from one place to another. The information cannot go any faster than the speed of light—so, ultimately, when our computers get faster and faster and more and more elaborate, we will have to make them smaller and smaller.

But there is plenty of room to make them smaller. There is nothing that I can see in the physical laws that says the computer elements cannot be made enormously smaller than they are now. In fact, there may be certain advantages.

Miniaturization by Evaporation

How can we make such a device? What kind of manufacturing processes would we use? One possibility we might consider, since we have talked about writing by putting atoms down in a certain arrangement, would be to evaporate the material, then evaporate the insulator next to it. Then, for the

next layer, evaporate another position of a wire, another insulator, and so on. So, you simply evaporate until you have a block of stuff which has the elements–coils and condensers, transistors and so on–of exceedingly fine dimensions.

But I would like to discuss, just for amusement, that there are other possibilities. Why can't we manufacture these small computers somewhat like we manufacture the big ones? Why can't we drill holes, cut things, solder things, stamp things out, mold different shapes all at an infinitesimal level? What are the limitations as to how small a thing has to be before you can no longer mold it? How many times when you are working on something frustratingly tiny, like your wife's wrist-watch, have you said to yourself, "If I could only train an ant to do this!" What I would like to suggest is the possibility of training an ant to train a mite to do this. What are the possibilities of small but movable machines? They may or may not be useful, but they surely would be fun to make.

Consider any machine–for example, an automobile–and ask about the problems of making an infinitesimal machine like it. Suppose, in the particular design of the automobile, we need a certain precision of the parts; we need an accuracy, let's suppose, of 4/10,000 of an inch. If things are more inaccurate than that in the shape of the cylinder and so on, it isn't going to work very well. If I make the thing too small, I have to worry about the size of the atoms; I can't make a circle of "balls" so to speak, if the circle is too small. So, if I make the error, corresponding to 4/10,000 of an inch, correspond to an error of 10 atoms, it turns out that I can reduce the dimensions of an automobile 4,000 times, approximately–so that it is 1 mm across. Obviously, if you redesign the car so that it would work with a much larger tolerance, which is not at all impossible, then you could make a much smaller device.

It is interesting to consider what the problems are in such small machines. Firstly, with parts stressed to the same degree, the forces go as the area you are reducing, so that things like weight and inertia are of relatively no importance. The strength of material, in other words, is very much greater in proportion. The stresses and expansion of the flywheel from centrifugal force, for example, would be the same proportion only if the rotational speed is increased in the same proportion as we decrease the size. On the other hand, the metals that we use have a grain structure, and this would be very annoying at small scale because the material is not homogeneous. Plastics and glass and things of this amorphous nature are very much more homogeneous, and so we would have to make our machines out of such materials.

There are problems associated with the electrical part of the system—with the copper wires and the magnetic parts. The magnetic properties on a very small scale are not the same as on a large scale; there is the "domain" problem involved. A big magnet made of millions of domains can only be made on a small scale with one domain. The electrical equipment won't simply be scaled down; it has to be redesigned. But I can see no reason why it can't be redesigned to work again.

Problems of Lubrication

Lubrication involves some interesting points. The effective viscosity of oil would be higher and higher in proportion as we went down (and if we increase the speed as much as we can). If we don't increase the speed so much, and change from oil to kerosene or some other fluid, the problem is not so bad. But actually we may not have to lubricate at all! We have a lot of extra force. Let the bearings run dry; they won't

run hot because the heat escapes away from such a small device very, very rapidly. This rapid heat loss would prevent the gasoline from exploding, so an internal combustion engine is impossible. Other chemical reactions, liberating energy when cold, can be used. Probably an external supply of electrical power would be most convenient for such small machines.

What would be the utility of such machines? Who knows? Of course, a small automobile would only be useful for the mites to drive around in, and I suppose our Christian interests don't go that far. However, we did note the possibility of the manufacture of small elements for computers in completely automatic factories, containing lathes and other machine tools at the very small level. The small lathe would not have to be exactly like our big lathe. I leave to your imagination the improvement of the design to take full advantage of the properties of things on a small scale, and in such a way that the fully automatic aspect would be easiest to manage.

A friend of mine (Albert R. Hibbs)* suggests a very interesting possibility for relatively small machines. He says that, although it is a very wild idea, it would be interesting in surgery if you could swallow the surgeon. You put the mechanical surgeon inside the blood vessel and it goes into the heart and "looks" around. (Of course the information has to be fed out.) It finds out which valve is the faulty one and takes a little knife and slices it out. Other small machines might be permanently incorporated in the body to assist some inadequately-functioning organ.

Now comes the interesting question: How do we make such a tiny mechanism? I leave that to you. However, let me suggest one weird possibility. You know, in the atomic energy

*A student and later a colleague of Feynman. *Ed.*

plants they have materials and machines that they can't handle directly because they have become radioactive. To unscrew nuts and put on bolts and so on, they have a set of master and slave hands, so that by operating a set of levers here, you control the "hands" there, and can turn them this way and that so you can handle things quite nicely.

Most of these devices are actually made rather simply, in that there is a particular cable, like a marionette string, that goes directly from the controls to the "hands." But, of course, things also have been made using servo motors, so that the connection between the one thing and the other is electrical rather than mechanical. When you turn the levers, they turn a servo motor, and it changes the electrical currents in the wires, which repositions a motor at the other end.

Now, I want to build much the same device—a master-slave system which operates electrically. But I want the slaves to be made especially carefully by modern large-scale machinists so that they are one-fourth the scale of the "hands" that you ordinarily maneuver. So you have a scheme by which you can do things at one-quarter scale anyway—the little servo motors with little hands play with little nuts and bolts; they drill little holes; they are four times smaller. Aha! So I manufacture a quarter-size lathe; I manufacture quarter-size tools; and I make, at the one-quarter scale, still another set of hands again relatively one-quarter size! This is one-sixteenth size, from my point of view. And after I finish doing this I wire directly from my large-scale system, through transformers perhaps, to the one-sixteenth-size servo motors. Thus I can now manipulate the one-sixteenth-size hands.

Well, you get the principle from there on. It is rather a difficult program, but it is a possibility. You might say that one can go much farther in one step than from one to four. Of course, this has all to be designed very carefully and it is not

necessary simply to make it like hands. If you thought of it very carefully, you could probably arrive at a much better system for doing such things.

If you work through a pantograph, even today, you can get much more than a factor of four in even one step. But you can't work directly through a pantograph which makes a smaller pantograph which then makes a smaller pantograph–because of the looseness of the holes and the irregularities of construction. The end of the pantograph wiggles with a relatively greater irregularity than the irregularity with which you move your hands. In going down this scale, I would find the end of the pantograph on the end of the pantograph on the end of the pantograph shaking so badly that it wasn't doing anything sensible at all.

At each stage, it is necessary to improve the precision of the apparatus. If, for instance, having made a small lathe with a pantograph, we find its lead screw irregular–more irregular than the large-scale one–we could lap the lead screw against breakable nuts that you can reverse in the usual way back and forth until this lead screw is, at its scale, as accurate as our original lead screws, at our scale.

We can make flats by rubbing unflat surfaces in triplicates together–in three pairs–and the flats then become flatter than the thing you started with. Thus, it is not impossible to improve precision on a small scale by the correct operations. So, when we build this stuff, it is necessary at each step to improve the accuracy of the equipment by working for a while down there, making accurate lead screws, Johansen blocks, and all the other materials which we use in accurate machine work at the higher level. We have to stop at each level and manufacture all the stuff to go to the next level–a very long and very difficult program. Perhaps you can figure a better way than that to get down to small scale more rapidly.

Yet, after all this, you have just got one little baby lathe four thousand times smaller than usual. But we were thinking of making an enormous computer, which we were going to build by drilling holes on this lathe to make little washers for the computer. How many washers can you manufacture on this one lathe?

A Hundred Tiny Hands

When I make my first set of slave "hands" at one-fourth scale, I am going to make ten sets. I make ten sets of "hands," and I wire them to my original levers so they each do exactly the same thing at the same time in parallel. Now, when I am making my new devices one-quarter again as small, I let each one manufacture ten copies, so that I would have a hundred "hands" at the 1/16 size.

Where am I going to put the million lathes that I am going to have? Why, there is nothing to it; the volume is much less than that of even one full-scale lathe. For instance, if I made a billion little lathes, each 1/4000 of the scale of a regular lathe, there are plenty of materials and space available because in the billion little ones there is less than 2 percent of the materials in one big lathe. It doesn't cost anything for materials, you see. So I want to build a billion tiny factories, models of each other, which are manufacturing simultaneously, drilling holes, stamping parts, and so on.

As we go down in size, there are a number of interesting problems that arise. All things do not simply scale down in proportion. There is the problem that materials stick together by the molecular (Van der Waals*) attractions. It would be

---

*Van der Waals forces: Weak attractive forces between atoms or molecules. Johannes Diderik Van der Waals (1837–1923) received the 1910 Nobel Prize in Physics for his work on the equation of state for gases and liquids. *Ed.*

like this: After you have made a part and you unscrew the nut from a bolt, it isn't going to fall down because the gravity isn't appreciable; it would even be hard to get it off the bolt. It would be like those old movies of a man with his hands full of molasses, trying to get rid of a glass of water. There will be several problems of this nature that we will have to be ready to design for.

Rearranging the Atoms

But I am not afraid to consider the final question as to whether, ultimately–in the great future–we can arrange the atoms the way we want; the very *atoms,* all the way down! What would happen if we could arrange the atoms one by one the way we want them (within reason, of course; you can't put them so that they are chemically unstable, for example)?

Up to now, we have been content to dig in the ground to find minerals. We heat them and we do things on a large scale with them, and we hope to get a pure substance with just so much impurity, and so on. But we must always accept some atomic arrangement that nature gives us. We haven't got anything, say, with a "checkerboard" arrangement, with the impurity atoms exactly arranged 1,000 angstroms apart, or in some other particular pattern.

What could we do with layered structures with just the right layers? What would the properties of materials be if we could really arrange the atoms the way we want them? They would be very interesting to investigate theoretically. I can't see exactly what would happen, but I can hardly doubt that when we have some *control* of the arrangement of things on a small scale, we will get an enormously greater range of possible properties that substances can have, and of different things that we can do.

Consider, for example, a piece of material in which we make little coils and condensers (or their solid-state analogs) 1,000 or 10,000 angstroms in a circuit, one right next to the other, over a large area, with little antennas sticking out at the other end–a whole series of circuits. Is it possible, for example, to emit light from a whole set of antennas, like we emit radio waves from an organized set of antennas to beam the radio programs to Europe? The same thing would be to beam light out in a definite direction with very high intensity. (Perhaps such a beam is not very useful technically or economically.)

I have thought about some of the problems of building electric circuits on a small scale, and the problem of resistance is serious. If you build a corresponding circuit on a small scale, its natural frequency goes up, since the wave length goes down as the scale; but the skin depth only decreases with the square root of the scale ratio, and so resistive problems are of increasing difficulty. Possibly we can beat resistance through the use of superconductivity if the frequency is not too high, or by other tricks.

Atoms in a Small World

When we get to the very, very small world–say, circuits of seven atoms–we have a lot of new things that would happen that represent completely new opportunities for design. Atoms on a small scale behave like *nothing* on a large scale, for they satisfy the laws of quantum mechanics. So, as we go down and fiddle around with the atoms down there, we are working with different laws, and we can expect to do different things. We can manufacture in different ways. We can use, not just circuits, but some system involving the quantized energy levels, or the interactions of quantized spins, etc.

Another thing we will notice is that, if we go down far enough, all of our devices can be mass produced so that they are absolutely perfect copies of one another. We cannot build two large machines so that the dimensions are exactly the same. But if your machine is only 100 atoms high, you only have to get it correct to one-half of one percent to make sure the other machine is exactly the same size–namely, 100 atoms high!

At the atomic level, we have new kinds of forces and new kinds of possibilities, new kinds of effects. The problems of manufacture and reproduction of materials will be quite different. I am, as I said, inspired by the biological phenomena in which chemical forces are used in repetitious fashion to produce all kinds of weird effects (one of which is the author). The principles of physics, as far as I can see, do not speak against the possibility of maneuvering things atom by atom. It is not an attempt to violate any laws; it is something, in principle, that can be done; but in practice, it has not been done because we are too big.

Ultimately, we can do chemical synthesis. A chemist comes to us and says, "Look, I want a molecule that has the atoms arranged thus and so; make me that molecule." The chemist does a mysterious thing when he wants to make a molecule. He sees that it has got that ring, so he mixes this and that, and he shakes it, and he fiddles around. And, at the end of a difficult process, he usually does succeed in synthesizing what he wants. By the time I get my devices working, so that we can do it by physics, he will have figured out how to synthesize absolutely anything, so that this will really be useless.

But it is interesting that it would be, in principle, possible (I think) for a physicist to synthesize any chemical substance that the chemist writes down. Give the orders and the physicist synthesizes it. How? Put the atoms down where the

chemist says, and so you make the substance. The problems of chemistry and biology can be greatly helped if our ability to see what we are doing, and to do things on an atomic level, is ultimately developed–a development which I think cannot be avoided. Now, you might say, "Who should do this and why should they do it?" Well, I pointed out a few of the economic applications, but I know that the reason that you would do it might be just for fun. But have some fun! Let's have a competition between laboratories. Let one laboratory make a tiny motor which it sends to another lab which sends it back with a thing that fits inside the shaft of the first motor.

High School Competition

Just for the fun of it, and in order to get kids interested in this field, I would propose that someone who has some contact with the high schools think of making some kind of high school competition. After all, we haven't even started in this field, and even the kids can write smaller than has ever been written before. They could have competition in high schools. The Los Angeles high school could send a pin to the Venice high school on which it says, "How's this?" They get the pin back, and in the dot of the "i" it says, "Not so hot."

Perhaps this doesn't excite you to do it, and only economics will do so. Then I want to do something; but I can't do it at the present moment, because I haven't prepared the ground. It is my intention to offer a prize of $1,000 to the first guy who can take the information on the page of a book and put it on an area 1/25,000 smaller in linear scale in such manner that it can be read by an electron microscope.

And I want to offer another prize–if I can figure out how to phrase it so that I don't get into a mess of arguments about definitions–of another $1,000 to the first guy who makes an

operating electric motor–a rotating electric motor which can be controlled from the outside and, not counting the lead-in wires, is only 1/64 inch cube.

I do not expect that such prizes will have to wait very long for claimants.

Ultimately Feynman had to make good on both challenges. The following is from the overview to Feynman and Computation, *edited by Anthony J.G. Hey (Perseus, Reading, MA, 1998), reprinted with permission.* Ed.

He paid out on both–the first, less than a year later, to Bill McLellan, a Caltech alumnus, for a miniature motor which satisfied the specifications but which was somewhat of a disappointment to Feynman in that it required no new technical advances. Feynman gave an updated version of his talk in 1983 to the Jet Propulsion Laboratory. He predicted 'that with today's technology we can easily . . . construct motors a fortieth of that size in each dimension, 64,000 times smaller than . . . McLellan's motor, and we can make thousands of them at a time.'

It was not for another 26 years that he had to pay out on the second prize, this time to a Stanford graduate student named Tom Newman. The scale of Feynman's challenge was equivalent to writing all twenty-four volumes of the *Encyclopaedia Brittanica* on the head of a pin: Newman calculated that each individual letter would be only about fifty atoms wide. Using electron-beam lithography when his thesis advisor was out of town, he was eventually able to write the first page of Charles Dickens' *A Tale of Two Cities* at 1/25,000 reduction in scale. Feynman's paper is often credited with starting the field of nanotechnology and there are now regular 'Feynman Nanotechnology Prize' competitions.

6

The Value of Science

Of all its many values, the greatest must be the freedom to doubt.

In Hawaii, Feynman learns a lesson in humility while touring a Buddhist temple: "To every man is given the key to the gates of heaven; the same key opens the gates of hell." This is one of Feynman's most eloquent pieces, reflecting on science's relevance to the human experience and vice versa. He also gives a lesson to fellow scientists on their responsibility to the future of civilization.

From time to time, people suggest to me that scientists ought to give more consideration to social problems–especially that they should be more responsible in considering the impact of science upon society. This same suggestion must be made to many other scientists, and it seems to be generally believed that if the scientists would only look at these very difficult social problems and not spend so much time fooling with the less vital scientific ones, great success would come of it.

It seems to me that we do think about these problems from time to time, but we don't put full-time effort into them–the

reason being that we know we don't have any magic formula for solving problems, that social problems are very much harder than scientific ones, and that we usually don't get anywhere when we do think about them.

I believe that a scientist looking at nonscientific problems is just as dumb as the next guy–and when he talks about a nonscientific matter, he will sound as naive as anyone untrained in the matter. Since the question of the value of science is not a scientific subject, this discussion is dedicated to proving my point–by example.

The first way in which science is of value is familiar to everyone. It is that scientific knowledge enables us to do all kinds of things and to make all kinds of things. Of course if we make good things, it is not only to the credit of science; it is also to the credit of the moral choice which led us to good work. Scientific knowledge is an enabling power to do either good or bad–but it does not carry instructions on how to use it. Such power has evident value–even though the power may be negated by what one does.

I learned a way of expressing this common human problem on a trip to Honolulu. In a Buddhist temple there, the man in charge explained a little bit about the Buddhist religion for tourists, and then ended his talk by telling them he had something to say to them that they would *never* forget–and I have never forgotten it. It was a proverb of the Buddhist religion:

"To every man is given the key to the gates of heaven; the same key opens the gates of hell."

What, then, is the value of the key to heaven? It is true that if we lack clear instructions that determine which is the gate to heaven and which the gate to hell, the key may be a dangerous object to use, but it obviously has value. How can we enter heaven without it?

The instructions, also, would be of no value without the key. So it is evident that, in spite of the fact that science could produce enormous horror in the world, it is of value because it *can* produce *something*.

Another value of science is the fun called intellectual enjoyment which some people get from reading and learning and thinking about it, and which others get from working in it. This is a very real and important point and one which is not considered enough by those who tell us it is our social responsibility to reflect on the impact of science on society.

Is this mere personal enjoyment of value to society as a whole? No! But it is also a responsibility to consider the value of society itself. Is it, in the last analysis, to arrange things so that people can enjoy things? If so, the enjoyment of science is as important as anything else.

But I would like *not* to underestimate the value of the worldview which is the result of scientific effort. We have been led to imagine all sorts of things infinitely more marvelous than the imaginings of poets and dreamers of the past. It shows that the imagination of nature is far, far greater than the imagination of man. For instance, how much more remarkable it is for us all to be stuck—half of us upside down—by a mysterious attraction, to a spinning ball that has been swinging in space for billions of years, than to be carried on the back of an elephant supported on a tortoise swimming in a bottomless sea.

I have thought about these things so many times alone that I hope you will excuse me if I remind you of some thoughts that I am sure you have all had—or this type of thought—which no one could ever have had in the past, because people then didn't have the information we have about the world today.

For instance, I stand at the seashore, alone, and start to think. There are the rushing waves . . . mountains of molecules, each stupidly minding its own business . . . trillions apart . . . yet forming white surf in unison.

Ages on ages . . . before any eyes could see . . . year after year . . . thunderously pounding the shore as now. For whom, for what? . . . on a dead planet, with no life to entertain.

Never at rest . . . tortured by energy . . . wasted prodigiously by the sun . . . poured into space. A mite makes the sea roar.

Deep in the sea, all molecules repeat the patterns of one another till complex new ones are formed. They make others like themselves . . . and a new dance starts.

Growing in size and complexity . . . living things, masses of atoms, DNA, protein . . . dancing a pattern ever more intricate.

Out of the cradle onto the dry land . . . here it is standing . . . atoms with consciousness . . . matter with curiosity.

Stands at the sea . . . wonders at wondering . . . I . . . a universe of atoms . . . an atom in the universe.

The Grand Adventure

The same thrill, the same awe and mystery, come again and again when we look at any problem deeply enough. With more knowledge comes deeper, more wonderful mystery, luring one on to penetrate deeper still. Never concerned that the answer may prove disappointing, but with pleasure and confidence we turn over each new stone to find unimagined strangeness leading on to more wonderful questions and mysteries—certainly a grand adventure!

It is true that few unscientific people have this particular type of religious experience. Our poets do not write about it; our artists do not try to portray this remarkable thing. I don't

know why. Is nobody inspired by our present picture of the universe? The value of science remains unsung by singers, so you are reduced to hearing–not a song or a poem, but an evening lecture about it. This is not yet a scientific age.

Perhaps one of the reasons is that you have to know how to read the music. For instance, the scientific article says, perhaps, something like this: "The radioactive phosphorus content of the cerebrum of the rat decreases to one-half in a period of two weeks." Now, what does that mean?

It means that phosphorus that is in the brain of a rat (and also in mine, and yours) is not the same phosphorus as it was two weeks ago, but that all of the atoms that are in the brain are being replaced, and the ones that were there before have gone away.

So what is this mind, what are these atoms with consciousness? Last week's potatoes! That is what now can *remember* what was going on in my mind a year ago–a mind which has long ago been replaced.

That is what it means when one discovers how long it takes for the atoms of the brain to be replaced by other atoms, to note that the thing which I call my individuality is only a pattern or dance. The atoms come into my brain, dance a dance, then go out; always new atoms but always doing the same dance, remembering what the dance was yesterday.

The Remarkable Idea

When we read about this in the newspaper, it says, "The scientist says that this discovery may have importance in the cure of cancer." The paper is only interested in the use of the idea, not the idea itself. Hardly anyone can understand the importance of an idea, it is so remarkable. Except that, possibly, some children catch on. And when a child catches on to

an idea like that, we have a scientist. These ideas do filter down (in spite of all the conversation about TV replacing thinking), and lots of kids get the spirit–and when they have the spirit you have a scientist. It's too late for them to get the spirit when they are in our universities, so we must attempt to explain these ideas to children.

I would now like to turn to a third value that science has. It is a little more indirect, but not much. The scientist has a lot of experience with ignorance and doubt and uncertainty, and this experience is of very great importance, I think. When a scientist doesn't know the answer to a problem, he is ignorant. When he has a hunch as to what the result is, he is uncertain. And when he is pretty darn sure of what the result is going to be, he is in some doubt. We have found it of paramount importance that in order to progress we must recognize the ignorance and leave room for doubt. Scientific knowledge is a body of statements of varying degrees of certainty–some most unsure, some nearly sure, none *absolutely* certain.

Now, we scientists are used to this, and we take it for granted that it is perfectly consistent to be unsure–that it is possible to live and *not* know. But I don't know whether everyone realizes that this is true. Our freedom to doubt was born of a struggle against authority in the early days of science. It was a very deep and strong struggle. Permit us to question–to doubt, that's all–not to be sure. And I think it is important that we do not forget the importance of this struggle and thus perhaps lose what we have gained. Here lies a responsibility to society.

We are all sad when we think of the wondrous potentialities human beings seem to have, as contrasted with their small accomplishments. Again and again people have thought that we could do much better. They of the past saw in the nightmare of their times a dream for the future. We, of

their future, see that their dreams, in certain ways surpassed, have in many ways remained dreams. The hopes for the future today are, in good share, those of yesterday.

Education, for Good and Evil

Once some thought that the possibilities people had were not developed because most of those people were ignorant. With education universal, could all men be Voltaires? Bad can be taught at least as efficiently as good. Education is a strong force, but for either good or evil.

Communications between nations must promote understanding: So went another dream. But the machines of communication can be channeled or choked. What is communicated can be truth or lie. Communication is a strong force also, but for either good or bad.

The applied sciences should free men of material problems at least. Medicine controls diseases. And the record here seems all to the good. Yet there are men patiently working to create great plagues and poisons. They are to be used in warfare tomorrow.

Nearly everybody dislikes war. Our dream today is peace. In peace, man can develop best the enormous possibilities he seems to have. But maybe future men will find that peace, too, can be good and bad. Perhaps peaceful men will drink out of boredom. Then perhaps drink will become the great problem which seems to keep man from getting all he thinks he should out of his abilities.

Clearly, peace is a great force, as is sobriety, as are material power, communication, education, honesty, and the ideals of many dreamers.

We have more of these forces to control than did the ancients. And maybe we are doing a little better than most of

them could do. But what we ought to be able to do seems gigantic compared with our confused accomplishments.

Why is this? Why can't we conquer ourselves?

Because we find that even great forces and abilities do not seem to carry with them clear instructions on how to use them. As an example, the great accumulation of understanding as to how the physical world behaves only convinces one that this behavior seems to have a kind of meaninglessness. The sciences do not directly teach good and bad.

Through all ages men have tried to fathom the meaning of life. They have realized that if some direction or meaning could be given to our actions, great human forces would be unleashed. So, very many answers must have been given to the question of the meaning of it all. But they have been of all different sorts, and the proponents of one answer have looked with horror at the actions of the believers in another. Horror, because from a disagreeing point of view all the great potentialities of the race were being channeled into a false and confining blind alley. In fact, it is from the history of the enormous monstrosities created by false belief that philosophers have realized the apparently infinite and wondrous capacities of human beings. The dream is to find the open channel.

What, then, is the meaning of it all? What can we say to dispel the mystery of existence?

If we take everything into account, not only what the ancients knew, but all of what we know today that they didn't know, then I think that we must frankly admit that *we do not know.*

But, in admitting this, we have probably found the open channel.

This is not a new idea; this is the idea of the age of reason. This is the philosophy that guided the men who made the democracy that we live under. The idea that no one really

knew how to run a government led to the idea that we should arrange a system by which new ideas could be developed, tried out, tossed out, more new ideas brought in; a trial and error system. This method was a result of the fact that science was already showing itself to be a successful venture at the end of the 18th century. Even then it was clear to socially minded people that the openness of the possibilities was an opportunity, and that doubt and discussion were essential to progress into the unknown. If we want to solve a problem that we have never solved before, we must leave the door to the unknown ajar.

Our Responsibility as Scientists

We are at the very beginning of time for the human race. It is not unreasonable that we grapple with problems. There are tens of thousands of years in the future. Our responsibility is to do what we can, learn what we can, improve the solutions and pass them on. It is our responsibility to leave the men of the future a free hand. In the impetuous youth of humanity, we can make grave errors that can stunt our growth for a long time. This we will do if we say we have the answers now, so young and ignorant; if we suppress all discussion, all criticism, saying, "This is it, boys, man is saved!" and thus doom man for a long time to the chains of authority, confined to the limits of our present imagination. It has been done so many times before.

It is our responsibility as scientists, knowing the great progress and great value of a satisfactory philosophy of ignorance, the great progress that is the fruit of freedom of thought, to proclaim the value of this freedom, to teach how doubt is not to be feared but welcomed and discussed, and to demand this freedom as our duty to all coming generations.

7

Richard P. Feynman's Minority Report to the Space Shuttle *Challenger* Inquiry

When the Space Shuttle Challenger *exploded shortly after its launch on January 28, 1986, six professional astronauts and one schoolteacher were tragically killed. The nation was devastated, and NASA was shaken out of its complacency, brought on by years of successful—or at least nonlethal—space missions. A commission was formed, led by Secretary of State William P. Rogers and composed of politicians, astronauts, military men, and one scientist, to investigate the cause of the accident and to recommend steps to prevent such a disaster from ever happening again. The fact that Richard Feynman was that one scientist may have made the difference between answering the question of why the* Challenger *failed and eternal mystery. Feynman was gutsier than most men, not afraid to jet all over the country to talk to the men on the ground, the engineers who had recognized the fact that propaganda was taking the lead over care and safety in the shuttle program. His report, which was perceived by the Commission as embarrassing to NASA, was almost suppressed by*

the Commission, but Feynman fought to have it included; it was relegated to an appendix. When the Commission held a live press conference to answer questions, Feynman did his now-famous tabletop experiment with one of the shuttle's gaskets, or O-rings, and a cup of ice water. It dramatically proved that those key gaskets had failed because the warning of the engineers that it was too cold outside to go ahead with the launch went unheeded by managers eager to impress their bosses with the punctuality of their mission schedule. Here is that historic report.

Introduction

It appears that there are enormous differences of opinion as to the probability of a failure with loss of vehicle and of human life. The estimates range from roughly 1 in 100 to 1 in 100,000. The higher figures come from working engineers, and the very low figures from management. What are the causes and consequences of this lack of agreement? Since 1 part in 100,000 would imply that one could put a Shuttle up each day for 300 years expecting to lose only one, we could more properly ask "What is the cause of management's fantastic faith in the machinery?"

We have also found that certification criteria used in Flight Readiness Reviews often develop a gradually decreasing strictness. The argument that the same risk was flown before without failure is often accepted as an argument for the safety of accepting it again. Because of this, obvious weaknesses are accepted again and again, sometimes without a sufficiently serious attempt to remedy them, or to delay a flight because of their continued presence.

There are several sources of information. There are published criteria for certification, including a history of modifications in the form of waivers and deviations. In addition, the records of the Flight Readiness Reviews for each flight document the arguments used to accept the risks of the flight. Information was obtained from the direct testimony and the reports of the range safety officer, Louis J. Ullian, with respect to the history of success of solid fuel rockets. There was a further study by him (as chairman of the launch abort safety panel [LASP]) in an attempt to determine the risks involved in possible accidents leading to radioactive contamination from attempting to fly a plutonium power supply (RTG) for future planetary missions. The NASA study of the same question is also available. For the history of the Space Shuttle Main Engines, interviews with management and engineers at Marshall, and informal interviews with engineers at Rocketdyne, were made. An independent (Caltech) mechanical engineer who consulted for NASA about engines was also interviewed informally. A visit to Johnson was made to gather information on the reliability of the avionics (computers, sensors, and effectors). Finally there is a report, "A Review of Certification Practices Potentially Applicable to Man-rated Reusable Rocket Engines," prepared at the Jet Propulsion Laboratory by N. Moore, et al., in February 1986, for NASA Headquarters, Office of Space Flight. It deals with the methods used by the FAA and the military to certify their gas turbine and rocket engines. These authors were also interviewed informally.

Solid Fuel Rockets (SRB)

An estimate of the reliability of solid fuel rockets was made by the range safety officer, by studying the experience of all

previous rocket flights. Out of a total of nearly 2,900 flights, 121 failed (1 in 25). This includes, however, what may be called early errors, rockets flown for the first few times in which design errors are discovered and fixed. A more reasonable figure for the mature rockets might be 1 in 50. With special care in the selection of the parts and in inspection, a figure of below 1 in 100 might be achieved but 1 in 1,000 is probably not attainable with today's technology. (Since there are two rockets on the Shuttle, these rocket failure rates must be doubled to get Shuttle failure rates from Solid Rocket Booster failure.)

NASA officials argue that the figure is much lower. They point out that these figures are for unmanned rockets but since the Shuttle is a manned vehicle "the probability of mission success is necessarily very close to 1.0." It is not very clear what this phrase means. Does it mean it is close to 1 or that it ought to be close to 1? They go on to explain "Historically this extremely high degree of mission success has given rise to a difference in philosophy between manned space flight programs and unmanned programs; i.e., numerical probability usage versus engineering judgement." (These quotations are from "Space Shuttle Data for Planetary Mission RTG Safety Analysis," pages 3-1, 3-2, February 15, 1985, NASA, JSC.) It is true that if the probability of failure was as low as 1 in 100,000 it would take an inordinate number of tests to determine it (for you would get nothing but a string of perfect flights from which no precise figure, other than that the probability is likely less than the number of such flights in the string so far). But, if the real probability is not so small, flights would show troubles, near failures, and possibly actual failure with a reasonable estimate. In fact, previous NASA experience had shown, on occasion, just such difficulties, near accidents, and accidents, all giving warning that

the probability of flight failure was not so very small. The inconsistency of the argument not to determine reliability through historical experience, as the range safety officer did, is that NASA also appeals to history, beginning "Historically this high degree of mission success . . ." Finally, if we are to replace standard numerical probability usage with engineering judgment, why do we find such an enormous disparity between the management estimate and the judgment of the engineers? It would appear that, for whatever purpose, be it for internal or external consumption, the management of NASA exaggerates the reliability of its product, to the point of fantasy.

The history of the certification and Flight Readiness Reviews will not be repeated here. (See other part of Commission reports.) The phenomenon of accepting for flight seals that had shown erosion and blow-by in previous flights is very clear. The *Challenger* flight is an excellent example. There are several references to flights that had gone before. The acceptance and success of these flights is taken as evidence of safety. But erosion and blow-by are not what the design expected. They are warnings that something is wrong. The equipment is not operating as expected, and therefore there is a danger that it can operate with even wider deviations in this unexpected and not thoroughly understood way. The fact that this danger did not lead to a catastrophe before is no guarantee that it will not the next time, unless it is completely understood. When playing Russian roulette the fact that the first shot got off safely is little comfort for the next. The origin and consequences of the erosion and blow-by were not understood. They did not occur equally on all flights and all joints; sometimes more, and sometimes less. Why not sometime, when whatever conditions determined it were right, still more, leading to catastrophe?

In spite of these variations from case to case, officials behaved as if they understood it, giving apparently logical arguments to each other often depending on the "success" of previous flights. For example, in determining if flight 51-L was safe to fly in the face of ring erosion in flight 51-C, it was noted that the erosion depth was only one-third of the radius. It had been noted in an experiment cutting the ring that cutting it as deep as one radius was necessary before the ring failed. Instead of being very concerned that variations of poorly understood conditions might reasonably create a deeper erosion this time, it was asserted, there was "a safety factor of three." This is a strange use of the engineer's term "safety factor." If a bridge is built to withstand a certain load without the beams permanently deforming, cracking, or breaking, it may be designed for the materials used to actually stand up under three times the load. This "safety factor" is to allow for uncertain excesses of load, or unknown extra loads, or weaknesses in the material that might have unexpected flaws, etc. If now the expected load comes on to the new bridge and a crack appears in a beam, this is a failure of the design. There was no safety factor at all; even though the bridge did not actually collapse because the crack only went one-third of the way through the beam. The O-rings of the Solid Rocket Boosters were not designed to erode. Erosion was a clue that something was wrong. Erosion was not something from which safety can be inferred.

There was no way, without full understanding, that one could have confidence that conditions the next time might not produce erosion three times more severe than the time before. Nevertheless, officials fooled themselves into thinking they had such understanding and confidence, in spite of the peculiar variations from case to case. A mathematical model was made to calculate erosion. This was a model based not on

physical understanding but on empirical curve fitting. To be more detailed, it was supposed a stream of hot gas impinged on the O-ring material, and the heat was determined at the point of stagnation (so far, with reasonable physical, thermodynamic laws). But to determine how much rubber eroded it was assumed this depended only on this heat by a formula suggested by data on a similar material. A logarithmic plot suggested a straight line, so it was supposed that the erosion varied as the .58 power of the heat, the .58 being determined by a nearest fit. At any rate, adjusting some other numbers, it was determined that the model agreed with the erosion (to a depth of one-third the radius of the ring). There is nothing much so wrong with this as believing the answer! Uncertainties appear everywhere. How strong the gas stream might be was unpredictable, it depended on holes formed in the putty. Blow-by showed that the ring might fail even though not, or only partially, eroded through. The empirical formula was known to be uncertain, for it did not go directly through the very data points by which it was determined. There were a cloud of points some twice above, and some twice below the fitted curve, so erosions twice predicted were reasonable from that cause alone. Similar uncertainties surrounded the other constants in the formula, etc., etc. When using a mathematical model, careful attention must be given to uncertainties in the model.

Liquid Fuel Engine (SSME)

During the flight of 51-L the three Space Shuttle Main Engines all worked perfectly, even, at the last moment, beginning to shut down the engines as the fuel supply began to fail. The question arises, however, as to whether, had it failed, and we were to investigate it in as much detail as we did the

Solid Rocket Booster, we would find a similar lack of attention to faults and a deteriorating reliability. In other words, were the organization weaknesses that contributed to the accident confined to the Solid Rocket Booster sector or were they a more general characteristic of NASA? To that end the Space Shuttle Main Engines and the avionics were both investigated. No similar study of the Orbiter or the External Tank was made.

The engine is a much more complicated structure than the Solid Rocket Booster, and a great deal more detailed engineering goes into it. Generally, the engineering seems to be of high quality and apparently considerable attention is paid to deficiencies and faults found in operation.

The usual way that such engines are designed (for military or civilian aircraft) may be called the component system, or bottom-up design. First it is necessary to thoroughly understand the properties and limitations of the materials to be used (for turbine blades, for example), and tests are begun in experimental rigs to determine those. With this knowledge larger component parts (such as bearings) are designed and tested individually. As deficiencies and design errors are noted they are corrected and verified with further testing. Since one tests only parts at a time, these tests and modifications are not overly expensive. Finally one works up to the final design of the entire engine, to the necessary specifications. There is a good chance, by this time, that the engine will generally succeed, or that any failures are easily isolated and analyzed because the failure modes, limitations of materials, etc., are so well understood. There is a very good chance that the modifications to the engine to get around the final difficulties are not very hard to make, for most of the serious problems have already been discovered and dealt with in the earlier, less expensive, stages of the process.

The Space Shuttle Main Engine was handled in a different manner, top down, we might say. The engine was designed and put together all at once with relatively little detailed preliminary study of the material and components. Then when troubles are found in the bearings, turbine blades, coolant pipes, etc., it is more expensive and difficult to discover the causes and make changes. For example, cracks have been found in the turbine blades of the high pressure oxygen turbopump. Are they caused by flaws in the material, the effect of the oxygen atmosphere on properties of the material, the thermal stresses of startup or shutdown, the vibration and stresses of steady running, or mainly at some resonance at certain speeds, etc.? How long can we run from crack initiation to crack failure, and how does this depend on power level? Using the completed engine as a test bed to resolve such questions is extremely expensive. One does not wish to lose entire engines in order to find out where and how failure occurs. Yet, an accurate knowledge of this information is essential to acquire a confidence in the engine reliability in use. Without detailed understanding, confidence cannot be attained.

A further disadvantage of the top-down method is that, if an understanding of a fault is obtained, a simple fix, such as a new shape for the turbine housing, may be impossible to implement without a redesign of the entire engine.

The Space Shuttle Main Engine is a very remarkable machine. It has a greater ratio of thrust to weight than any previous engine. It is built at the edge of, or outside of, previous engineering experience. Therefore, as expected, many different kinds of flaws and difficulties have turned up. Because, unfortunately, it was built in the top-down manner, they are difficult to find and to fix. The design aim of a lifetime of 55 mission equivalent firings (27,000 seconds of operation, ei-

ther in a mission of 500 seconds, or on a test stand) has not been obtained. The engine now requires very frequent maintenance and replacement of important parts, such as turbopumps, bearings, sheet metal housings, etc. The high-pressure fuel turbopump had to be replaced every three or four mission equivalents (although that may have been fixed, now) and the high-pressure oxygen turbopump every five or six. This is at most ten percent of the original specification. But our main concern here is the determination of reliability.

In a total of about 250,000 seconds of operation, the engines have failed seriously perhaps 16 times. Engineering pays close attention to these failings and tries to remedy them as quickly as possible. This it does by test studies on special rigs experimentally designed for the flaw in question, by careful inspection of the engine for suggestive clues (like cracks), and by considerable study and analysis. In this way, in spite of the difficulties of top-down design, through hard work many of the problems have apparently been solved.

A list of some of the problems follows. Those followed by an asterisk (*) are probably solved:

Turbine blade cracks in high pressure fuel turbopumps (HPFTP). (May have been solved.)
Turbine blade cracks in high pressure oxygen turbopumps (HPOTP).
Augmented Spark Igniter (ASI) line rupture.*
Purge check valve failure.*
ASI chamber erosion.*
HPFTP turbine sheet metal cracking.
HPFTP coolant liner failure.*
Main combustion chamber outlet elbow failure.*
Main combustion chamber inlet elbow weld offset.*
HPOTP subsynchronous whirl.*

Flight acceleration safety cutoff system (partial failure in a redundant system).*
Bearing spalling (partially solved).
A vibration at 4,000 Hertz making some engines inoperable, etc.

Many of these solved problems are the early difficulties of a new design, for 13 of them occurred in the first 125,000 seconds and only three in the second 125,000 seconds. Naturally, one can never be sure that all the bugs are out, and, for some, the fix may not have addressed the true cause. Thus, it is not unreasonable to guess there may be at least one surprise in the next 250,000 seconds, a probability of 1/500 per engine per mission. On a mission there are three engines, but some accidents would possibly be contained, and only affect one engine. The system can abort with only two engines. Therefore let us say that the unknown surprises do not, even of themselves, permit us to guess that the probability of mission failure due to the Space Shuttle Main Engine is less than 1/500. To this we must add the chance of failure from known, but as yet unsolved, problems (those without the asterisk in the list above). These we discuss below. (Engineers at Rocketdyne, the manufacturer, estimate the total probability as 1/10,000. Engineers at Marshall estimate it as 1/300, while NASA management, to whom these engineers report, claims it is 1/100,000. An independent engineer consulting for NASA thought 1 or 2 per 100 a reasonable estimate.)

The history of the certification principles for these engines is confusing and difficult to explain. Initially the rule seems to have been that two sample engines must each have had twice the time operating without failure, as the operating time of the engine to be certified (rule of 2x). At least that is the FAA practice, and NASA seems to have adopted it, originally

expecting the certified time to be 10 missions (hence 20 missions for each sample). Obviously the best engines to use for comparison would be those of greatest total (flight plus test) operating time–the so-called "fleet leaders." But what if a third sample and several others fail in a short time? Surely we will not be safe because two were unusual in lasting longer. The short time might be more representative of the real possibilities, and in the spirit of the safety factor of 2, we should only operate at half the time of the short-lived samples.

The slow shift toward decreasing safety factor can be seen in many examples. We take that of the HPFTP turbine blades. First of all the idea of testing an entire engine was abandoned. Each engine number has had many important parts (like the turbopumps themselves) replaced at frequent intervals, so that the rule must be shifted from engines to components. We accept an HPFTP for a certification time if two samples have each run successfully for twice that time (and of course, as a practical matter, no longer insisting that this time be as large as 10 missions). But what is "successfull"? The FAA calls a turbine blade crack a failure, in order, in practice, to really provide a safety factor greater than 2. There is some time that an engine can run between the time a crack originally starts until the time it has grown large enough to fracture. (The FAA is contemplating new rules that take this extra safety time into account, but only if it is very carefully analyzed through known models within a known range of experience and with materials thoroughly tested. None of these conditions apply to the Space Shuttle Main Engine.)

Cracks were found in many second stage HPFTP turbine blades. In one case three were found after 1,900 seconds, while in another they were not found after 4,200 seconds, although usually these longer runs showed cracks. To follow

this story further we shall have to realize that the stress depends a great deal on the power level. The *Challenger* flight was to be at, and previous flights had been at, a power level called 104% of rated power level during most of the time the engines were operating. Judging from some material data it is supposed that at the level 104% of rated power level, the time to crack is about twice that at 109% or full power level (FPL). Future flights were to be at this level because of heavier payloads, and many tests were made at this level. Therefore dividing time at 104% by 2, we obtain units called equivalent full power level (EFPL). (Obviously, some uncertainty is introduced by that, but it has not been studied.) The earliest cracks mentioned above occurred at 1,375 EFPL.

Now the certification rule becomes "limit all second stage blades to a maximum of 1,375 seconds EFPL." If one objects that the safety factor of 2 is lost, it is pointed out that the one turbine ran for 3,800 seconds EFPL without cracks, and half of this is 1,900 so we are being more conservative. We have fooled ourselves in three ways. First we have only one sample, and it is not the fleet leader, for the other two samples of 3,800 or more seconds had 17 cracked blades between them. (There are 59 blades in the engine.) Next we have abandoned the 2x rule and substituted equal time. And finally, 1,375 is where we did see a crack. We can say that no crack had been found below 1,375, but the last time we looked and saw no cracks was 1,100 seconds EFPL. We do not know when the crack formed between these times; for example, cracks may have formed at 1,150 seconds EFPL. (Approximately 2/3 of the blade sets tested in excess of 1,375 seconds EFPL had cracks. Some recent experiments have, indeed, shown cracks as early as 1,150 seconds.) It was important to keep the number high, for the *Challenger* was to fly an engine very close to the limit by the time the flight was over.

Finally it is claimed that the criteria are not abandoned, and the system is safe, by giving up the FAA convention that there should be no cracks, and considering only a completely fractured blade a failure. With this definition no engine has yet failed. The idea is that since there is sufficient time for a crack to grow to fracture we can insure that all is safe by inspecting all blades for cracks. If they are found, replace them, and if none are found we have enough time for a safe mission. This makes the crack problem not a flight safety problem, but merely a maintenance problem.

This may in fact be true. But how well do we know that cracks always grow slowly enough that no fracture can occur in a mission? Three engines have run for long times with a few cracked blades (about 3,000 seconds EFPL) with no blades broken off.

But a fix for this cracking may have been found. By changing the blade shape, shot-peening the surface, and covering with insulation to exclude thermal shock, the blades have not cracked so far.

A very similar story appears in the history of certification of the HPOTP, but we shall not give the details here.

It is evident, in summary, that the Flight Readiness Reviews and certification rules show a deterioration for some of the problems of the Space Shuttle Main Engine that is closely analogous to the deterioration seen in the rules for the Solid Rocket Booster.

Avionics

By "avionics" is meant the computer system on the Orbiter as well as its input sensors and output actuators. At first we will restrict ourselves to the computers proper and not be concerned with the reliability of the input information from

the sensors of temperature, pressure, etc., or with whether the computer output is faithfully followed by the actuators of rocket firings, mechanical controls, displays to astronauts, etc.

The computing system is very elaborate, having over 250,000 lines of code. It is responsible, among many other things, for the automatic control of the entire ascent to orbit, and for the descent until well into the atmosphere (below Mach 1) once one button is pushed deciding the landing site desired. It would be possible to make the entire landing automatically (except that the landing gear lowering signal is expressly left out of computer control, and must be provided by the pilot, ostensibly for safety reasons) but such an entirely automatic landing is probably not as safe as a pilot controlled landing. During orbital flight it is used in the control of payloads, in displaying information to the astronauts, and the exchange of information to the ground. It is evident that the safety of flight requires guaranteed accuracy of this elaborate system of computer hardware and software.

In brief, the hardware reliability is ensured by having four essentially independent identical computer systems. Where possible each sensor also has multiple copies, usually four, and each copy feeds all four of the computer lines. If the inputs from the sensors disagree, depending on circumstances, certain averages, or a majority selection is used as the effective input. The algorithm used by each of the four computers is exactly the same, so their inputs (since each sees all copies of the sensors) are the same. Therefore at each step the results in each computer should be identical. From time to time they are compared, but because they might operate at slightly different speeds a system of stopping and waiting at specified times is instituted before each comparison is made. If one of the computers disagrees, or is too late in having its answer ready, the three which do agree are assumed to be correct and

the errant computer is taken completely out of the system. If, now, another computer fails, as judged by the agreement of the other two, it is taken out of the system, and the rest of the flight canceled, and descent to the landing site is instituted, controlled by the two remaining computers. It is seen that this is a redundant system since the failure of only one computer does not affect the mission. Finally, as an extra feature of safety, there is a fifth independent computer, whose memory is loaded with only the programs for ascent and descent, and which is capable of controlling the descent if there is a failure of more than two of the computers of the main line of four.

There is not enough room in the memory of the main line computers for all the programs of ascent, descent, and payload programs in flight, so the memory is loaded about four times from tapes, by the astronauts.

Because of the enormous effort required to replace the software for such an elaborate system, and for checking a new system out, no change has been made in the hardware since the system began about fifteen years ago. The actual hardware is obsolete; for example, the memories are of the old ferrite core type. It is becoming more difficult to find manufacturers to supply such old-fashioned computers reliably and of high quality. Modern computers are very much more reliable, can run much faster, simplifying circuits, and allowing more to be done, and would not require so much loading of memory, for their memories are much larger.

The software is checked very carefully in a bottom-up fashion. First, each new line of code is checked, then sections of codes or modules with special function are verified. The scope is increased step by step until the new changes are incorporated into a complete system and checked. This complete output is considered the final product, newly released.

But completely independently there is an independent verification group, that takes an adversary attitude to the software development group, and tests and verifies the software as if it were a customer of a delivered product. There is additional verification in using the new programs in simulators, etc. A discovery of an error during the verification testing is considered very serious, and its origin studied very carefully to avoid such mistakes in the future. Such unexpected errors have been found only about six times in all the programming and program changing (for new or altered payloads) that have been done. The principle that is followed is that all the verification is not an aspect of program safety, it is merely a test of that safety, in a non-catastrophic verification. Flight safety is to be judged solely on how well the programs do in the verification tests. A failure here generates considerable concern.

To summarize, then, the computer software checking system and attitude is of highest quality. There appears to be no process of gradually fooling oneself while degrading standards so characteristic of the Solid Rocket Booster or Space Shuttle Main Engine safety systems. To be sure, there have been recent suggestions by management to curtail such elaborate and expensive tests as being unnecessary at this late date in Shuttle history. This must be resisted for it does not appreciate the mutual subtle influences, and sources of error generated by even small changes of one part of a program on another. There are perpetual requests for changes as new payloads and new demands and modifications are suggested by the users. Changes are expensive because they require extensive testing. The proper way to save money is to curtail the number of requested changes, not the quality of testing for each.

One might add that the elaborate system could be very much improved by more modern hardware and program-

ming techniques. Any outside competition would have all the advantages of starting over, and whether that is a good idea for NASA now should be carefully considered.

Finally, returning to the sensors and actuators of the avionics system, we find that the attitude to system failure and reliability is not nearly as good as for the computer system. For example, a difficulty was found with certain temperature sensors sometimes failing. Yet 18 months later the same sensors were still being used, still sometimes failing, until a launch had to be scrubbed because two of them failed at the same time. Even on a succeeding flight this unreliable sensor was used again. Again reaction control systems, the rocket jets used for reorienting and control in flight, still are somewhat unreliable. There is considerable redundancy, but a long history of failures, none of which has yet been extensive enough to seriously affect a flight. The action of the jets is checked by sensors, and if they fail to fire, the computers choose another jet to fire. But they are not designed to fail, and the problem should be solved.

Conclusions

If a reasonable launch schedule is to be maintained, engineering often cannot be done fast enough to keep up with the expectations of originally conservative certification criteria designed to guarantee a very safe vehicle. In these situations, subtly, and often with apparently logical arguments, the criteria are altered so that flights may still be certified in time. They therefore fly in a relatively unsafe condition, with a chance of failure of the order of a percent (it is difficult to be more accurate).

Official management, on the other hand, claims to believe the probability of failure is a thousand times less. One reason

for this may be an attempt to assure the government of NASA perfection and success in order to ensure the supply of funds. The other may be that they sincerely believe it to be true, demonstrating an almost incredible lack of communication between themselves and their working engineers.

In any event this has had very unfortunate consequences, the most serious of which is to encourage ordinary citizens to fly in such a dangerous machine, as if it had attained the safety of an ordinary airliner. The astronauts, like test pilots, should know their risks, and we honor them for their courage. Who can doubt that McAuliffe was equally a person of great courage, who was closer to an awareness of the true risk than NASA management would have us believe?

Let us make recommendations to ensure that NASA officials deal in a world of reality in understanding technological weaknesses and imperfections well enough to be actively trying to eliminate them. They must live in reality in comparing the costs and utility of the Shuttle to other methods of entering space. And they must be realistic in making contracts, in estimating costs, and the difficulty of the projects. Only realistic flight schedules should be proposed, schedules that have a reasonable chance of being met. If in this way the government would not support them, then so be it. NASA owes it to the citizens from whom it asks support to be frank, honest, and informative, so that these citizens can make the wisest decisions for the use of their limited resources.

For a successful technology, reality must take precedence over public relations, for nature cannot be fooled.

8

What Is Science?

What is science? It is common sense! Or is it? In April 1966 the master teacher delivered an address to the National Science Teachers' Association in which he gave his fellow teachers lessons on how to teach their students to think like a scientist and how to view the world with curiosity, open-mindedness, and, above all, doubt. This talk is also a tribute to the enormous influence Feynman's father—a uniforms salesman—had on Feynman's way of looking at the world.

I thank Mr. DeRose for the opportunity to join you science teachers. I also am a science teacher. I have too much experience only in teaching graduate students in physics, and as a result of that experience I know that I don't know how to teach.

I am sure that you who are real teachers working at the bottom level of this hierarchy of teachers, instructors of teachers, experts on curricula, also are sure that you, too, don't know how to do it; otherwise you wouldn't bother to come to the Convention.

The subject "What Is Science?" is not my choice. It was Mr. DeRose's subject. But I would like to say that I think that "What Is Science?" is not at all equivalent to "how to teach

science," and I must call that to your attention for two reasons. In the first place, from the way that I am preparing to give this lecture, it may seem that I am trying to tell you how to teach science—I am not at all in any way, because I don't know anything about small children. I have one, so I know that I don't know. The other is I think that most of you (because there is so much talk and so many papers and so many experts in the field) have some kind of a feeling of lack of self-confidence. In some way you are always being lectured on how things are not going too well and how you should learn to teach better. I am not going to berate you for the bad works you are doing and indicate how it can definitely be improved; that is not my intention.

As a matter of fact, we have very good students coming into Caltech, and during the years we found them getting better and better. Now how it is done, I don't know. I wonder if you know. I don't want to interfere with the system; it's very good.

Only two days ago we had a conference in which we decided that we don't have to teach a course in elementary quantum mechanics in the graduate school anymore. When I was a student, they didn't even have a course in quantum mechanics in the graduate school it was considered too difficult a subject. When I first started to teach, we had one. Now we teach it to undergraduates. We discover now that we don't have to have elementary quantum mechanics for graduates from other schools. Why is it getting pushed down? Because we are able to teach better in the university, and that is because the students coming up are better trained.

What is science? Of course you all must know, if you teach it. That's common sense. What can I say? If you don't know, every teacher's edition of every textbook gives a complete discussion of the subject. There is some kind of distorted distilla-

tion and watered-down and mixed-up words of Francis Bacon from some centuries ago, words which then were supposed to be the deep philosophy of science. But one of the greatest experimental scientists of the time who was really doing something, William Harvey,* said that what Bacon said science was, was the science that a lord chancellor would do. He spoke of making observations, but omitted the vital factor of judgment about what to observe and what to pay attention to.

And so what science is, is not what the philosophers have said it is and certainly not what the teacher editions say it is. What it is, is a problem which I set for myself after I said I would give this talk.

After some time I was reminded of a little poem.

> A centipede was happy quite, until a toad in fun
> Said, "Pray, which leg comes after which?"
> This raised his doubts to such a pitch
> He fell distracted in the ditch
> Not knowing how to run.

All my life, I have been doing science and known what it was, but what I have come to tell you–which foot comes after which–I am unable to do, and furthermore, I am worried by the analogy with the poem, that when I go home I will no longer be able to do any research.

There have been a lot of attempts by the various press reporters to get some kind of a capsule of this talk; I prepared it only a little time ago, so it was impossible; but I can see them all rushing out now to write some sort of headline which says: "The Professor Called the President of NSTA a Toad."

---

*Harvey (1578–1657) discovered the body's circulatory system. *Ed.*

Under these circumstances of the difficulty of the subject, and my dislike of philosophical exposition, I will present it in a very unusual way. I am just going to tell you how I learned what science is. That's a little bit childish. I learned it as a child. I have had it in my blood from the beginning. And I would like to tell you how it got in. This sounds as though I am trying to tell you how to teach, but that is not my intention. I'm going to tell you what science is like by how I learned what science is like.

My father did it to me. When my mother was carrying me, it is reported–I am not directly aware of the conversation–my father said that "if it's a boy, he'll be a scientist." How did he do it? He never told me I should be a scientist. He was not a scientist; he was a businessman, a sales manager of a uniform company, but he read about science and loved it.

When I was very young–the earliest story I know–when I still ate in a high chair, my father would play a game with me after dinner. He had bought a whole lot of old rectangular bathroom floor tiles from someplace in Long Island City. We set them up on end, one next to the other, and I was allowed to push the end one and watch the whole thing go down. So far so good.

Next, the game improved. The tiles were different colors. I must put one white, two blues, one white, two blues, and another white and then two blues–I may want to put another blue, but it must be a white. You recognize already the usual insidious cleverness; first delight him in play, and then slowly inject material of educational value!

Well, my mother, who is a much more feeling woman, began to realize the insidiousness of his efforts and said, "Mel, please let the poor child put a blue tile if he wants to." My father said, "No, I want him to pay attention to patterns. It is the only thing I can do that is mathematics at this earli-

est level." If I were giving a talk on "what is mathematics?" I would have already answered you. Mathematics is looking for patterns. (The fact is that this education had some effect. We had a direct experimental test at the time I got to kindergarten. We had weaving in those days. They've taken it out; it's too difficult for children. We used to weave colored paper through vertical strips and make patterns. The kindergarten teacher was so amazed that she sent a special letter home to report that this child was very unusual, because he seemed to be able to figure out ahead of time what pattern he was going to get, and made amazingly intricate patterns. So the tile game did do something to me.)

I would like to report other evidence that mathematics is only patterns. When I was at Cornell, I was rather fascinated by the student body, which seems to me was a dilute mixture of some sensible people in a big mass of dumb people studying home economics, etc., including lots of girls. I used to sit in the cafeteria with the students and eat and try to overhear their conversations and see if there was one intelligent word coming out. You can imagine my surprise when I discovered a tremendous thing, it seemed to me.

I listened to a conversation between two girls, and one was explaining that if you want to make a straight line, you see, you go over a certain number to the right for each row you go up, that is, if you go over each time the same amount when you go up a row, you make a straight line. A deep principle of analytic geometry! It went on. I was rather amazed. I didn't realize the female mind was capable of understanding analytic geometry.

She went on and said, "Suppose you have another line coming in from the other side and you want to figure out where they are going to intersect." Suppose on one line you go over two to the right for every one you go up, and the other line

goes over three to the right for every one that it goes up, and they start twenty steps apart, etc.–I was flabbergasted. She figured out where the intersection was! It turned out that one girl was explaining to the other how to knit argyle socks.

I, therefore, did learn a lesson: The female mind is capable of understanding analytic geometry. Those people who have for years been insisting (in the face of all obvious evidence to the contrary) that the male and female are equal and capable of rational thought may have something. The difficulty may just be that we have never yet discovered a way to communicate with the female mind. If it is done in the right way, you may be able to get something out of it.

Now I will go on with my own experience as a youngster in mathematics.

Another thing that my father told me–and I can't quite explain it, because it was more an emotion than a telling–was that the ratio of the circumference to the diameter of all circles was always the same, no matter what the size. That didn't seem to me too unobvious, but the ratio had some marvelous property. That was a wonderful number, a deep number, pi.* There was a mystery about this number that I didn't quite understand as a youth, but this was a great thing, and the result was that I looked for π everywhere.

When I was learning later in school how to make the decimals for fractions, and how to make $3\frac{1}{8}$, I wrote 3.125, and thinking I recognized a friend wrote that it equals π, the ratio of circumference to diameter of a circle. The teacher corrected it to 3.1416.

I illustrate these things to show an influence. The idea that there is a mystery, that there is a wonder about the number was important to me, not what the number was.

---

*That is, the Greek lowercase letter π.

Very much later when I was doing experiments in the laboratory–I mean my own home laboratory–fiddling around–no, excuse me, I didn't do experiments, I never did; I just fiddled around. I made radios and gadgets. I fiddled around. Gradually through books and manuals I began to discover there were formulas applicable to electricity in relating the current and resistance, and so on. One day, looking at the formulas in some book or other, I discovered a formula for the frequency of a resonant circuit which was $2\pi\sqrt{LC}$ where L is the inductance and C the capacitance of the circuit. And there was π, and where was the circle? You laugh, but I was very serious then. π was a thing with circles, and here is π coming out of an electric circuit, where [it stood for] the circle. Do you who laughed know how that π comes about?

I have to love the thing. I have to look for it. I have to think about it. And then I realized, of course, that the coils are made in circles. About a half year later, I found another book which gave the inductance of round coils and square coils, and there were other π's in these formulas. I began to think about it again, and I realized that the π did not come from the circular coils. I understand it better now; but in my heart I still don't quite know where that circle is, where that π comes from. [...]

I would like to say a word or two–may I interrupt my little tale–about words and definitions, because it is necessary to learn the words. It is not science. That doesn't mean just because it is not science that we don't have to teach the words. We are not talking about what to teach; we are talking about what science is. It is not science to know how to change centigrade to Fahrenheit. It's necessary, but it is not exactly science. In the same sense, if you were discussing what art is,

you wouldn't say art is the knowledge of the fact that a 3-B pencil is softer than a 2-H pencil. It's a distinct difference. That doesn't mean an art teacher shouldn't teach that, or that an artist gets along very well if he doesn't know that. (Actually you can find out in a minute by trying it; but that's a scientific way that art teachers may not think of explaining.)

In order to talk to each other, we have to have words, and that's all right. It's a good idea to try to see the difference, and it's a good idea to know when we are teaching the tools of science, such as words, and when we are teaching science itself.

To make my point still clearer, I shall pick out a certain science book to criticize unfavorably, which is unfair, because I am sure that with little ingenuity, I can find equally unfavorable things to say about others.

There is a first-grade science book which, in the first lesson of the first grade, begins in an unfortunate manner to teach science, because it starts off on the wrong idea of what science is. There is a picture of a dog, a windable toy dog, and a hand comes to the winder, and then the dog is able to move. Under the last picture, it says "What makes it move?" Later on, there is a picture of a real dog and the question "What makes it move?" Then there is a picture of a motor bike and the question "What makes it move?" and so on.

I thought at first they were getting ready to tell what science was going to be about: physics, biology, chemistry. But that wasn't it. The answer was in the teacher's edition of the book; the answer I was trying to learn is that "energy makes it move."

Now energy is a very subtle concept. It is very, very difficult to get right. What I mean by that is that it is not easy to understand energy well enough to use it right, so that you can deduce something correctly using the energy idea. It is beyond the first grade. It would be equally well to say that "God makes

it move," or "spirit makes it move," or "movability makes it move." (In fact equally well to say "energy makes it stop.")

Look at it this way: That's only the definition of energy. It should be reversed. We might say when something can move that it has energy in it, but not "what makes it move is energy." This is a very subtle difference. It's the same with this inertia proposition. Perhaps I can make the difference a little clearer this way:

If you ask a child what makes the toy dog move; if you ask an ordinary human being what makes a toy dog move, that is what you should think about. The answer is that you wound up the spring; it tries to unwind and pushes the gear around. What a good way to begin a science course. Take apart the toy; see how it works. See the cleverness of the gears; see the ratchets. Learn something about the toy, the way the toy is put together, the ingenuity of people, devising the ratchets and other things. That's good. The question is fine. The answer is a little unfortunate, because what they were trying to do is teach a definition of energy. But nothing whatever is learned.

Suppose a student would say, "I don't think energy makes it move." Where does the discussion go from there?

I finally figured out a way to test whether you have taught an idea or you have only taught a definition. Test it this way: You say, "Without using the new word which you have just learned, try to rephrase what you have just learned in your own language." "Without using the word 'energy,' tell me what you know now about the dog's motion." You cannot. So you learned nothing except the definition. You learned nothing about science. That may be all right. You may not want to learn something about science right away. You have to learn definitions. But for the very first lesson is that not possibly destructive?

I think, for lesson number one, to learn a mystic formula for answering questions is very bad. The book has some others—"gravity makes it fall"; "the soles of your shoes wear out because of friction." Shoe leather wears out because it rubs against the sidewalk and the little notches and bumps on the sidewalk grab pieces and pull them off. To simply say it is because of friction is sad, because it's not science.

My father dealt a little bit with energy and used the term after I got a little bit of the idea about it. What he would have done I know, because he did in fact essentially the same thing—though not the same example of the toy dog. He would say, "It moves because the sun is shining," if he wanted to give the same lesson. I would say "No. What has that to do with the sun shining? It moved because I wound up the springs."

"And why, my friend, are you able to move to wind up this spring?"

"I eat."

"What, my friend, do you eat?"

"I eat plants."

"And how do they grow?"

"They grow because the sun is shining."

And it is the same with the dog. What about gasoline? Accumulated energy of the sun which is captured by plants and preserved in the ground. Other examples all end with the sun. And so the same idea about the world that our textbook is driving at is phrased in a very exciting way. All the things that we see that are moving are moving because the sun is shining. It does explain the relationship of one source of energy to another, and it can be denied by the child. He could say, "I don't think it is on account of the sun shining," and you can start a discussion. So there is a difference. (Later I could challenge him with the tides, and what makes the earth turn, and have my hand on mystery again.)

That is just an example of the difference between definitions (which are necessary) and science. The only objection in this particular case was that it was the first lesson. It must certainly come in later, telling you what energy is, but not to such a simple question as "What makes a dog move?" A child should be given a child's answer. "Open it up; let's look at it."

During walks in the woods with my father, I learned a great deal. In the case of birds, for example: Instead of naming them, my father would say, "Look, notice that the bird is always pecking in its feathers. It pecks a lot in its feathers. Why do you think it pecks the feathers?"

I guessed it's because the feathers are ruffled, and he's trying to straighten them out. He said "Okay, when would the feathers get ruffled, or how would they get ruffled?"

"When he flies. When he walks around, it's okay; but when he flies it ruffles the feathers."

Then he would say, "You would guess then when the bird just landed he would have to peck more at his feathers than after he has straightened them out and has been walking around the ground for a while. Okay; let's look."

So we would look, and we would watch, and it turned out, as far as I could make out, that the bird pecked about as much and as often no matter how long he was walking on the ground and not just directly after flight.

So my guess was wrong, and I couldn't guess the right reason. My father revealed the reason.

It is that the birds have lice. There is a little flake that comes off the feather, my father taught me, stuff that can be eaten, and the louse eats it. And then on the louse, there is a little bit of wax in the joints between the sections of the leg that oozes out, and there is a mite that lives in there that can eat that wax. Now the mite has such a good source of food that it doesn't digest it too well, so from the rear end there

comes a liquid that has too much sugar, and in that sugar lives a tiny creature, etc.

The facts are not correct. The spirit is correct. First I learned about parasitism, one on the other, on the other, on the other.

Second, he went on to say that in the world whenever there is any source of something that could be eaten to make life go, some form of life finds a way to make use of that source; and that each little bit of leftover stuff is eaten by something.

Now the point of this is that the result of observation, even if I were unable to come to the ultimate conclusion, was a wonderful piece of gold, with a marvelous result. It was something marvelous.

Suppose I were told to observe, to make a list, to write down, to do this, to look, and when I wrote my list down, it was filed with 130 other lists in the back of a notebook. I would learn that the result of observation is relatively dull, that nothing much comes of it.

I think it is very important—at least it was to me—that if you are going to teach people to make observations, you should show that something wonderful can come from them. I learned then what science was about. It was patience. If you looked, and you watched, and you paid attention, you got a great reward from it (although possibly not every time). As a result, when I became a more mature man, I would painstakingly, hour after hour, for years, work on problems—sometimes many years, sometimes shorter times—many of them failing, lots of stuff going into the wastebasket; but every once in a while there was the gold of a new understanding that I had learned to expect when I was a kid, the result of observation. For I did not learn that observation was not worthwhile.

Incidentally, in the forest we learned other things. We would go for walks and see all the regular things, and talk about many things; about the growing plants, the struggle of the trees for light, how they try to get as high as they can, and to solve the problem of getting water higher than 35 or 40 feet, the little plants on the ground that look for the little bits of light that come through, all that growth, and so forth.

One day after we had seen all this, my father took me to the forest again and said, "In all this time we have been looking at the forest, we have only seen half of what is going on, exactly half."

I said, "What do you mean?"

He said, "We have been looking at how all these things grow; but for each bit of growth, there must be the same amount of decay, otherwise the materials would be consumed forever. Dead trees would lie there having used up all the stuff from the air, and the ground, and it wouldn't get back into the ground or the air, and nothing else could grow, because there is no material available. There must be for each bit of growth exactly the same amount of decay."

There then followed many walks in the woods during which we broke up old stumps, saw funny bugs and funguses growing—he couldn't show me bacteria, but we saw the softening effects, and so on. I saw the forest as a process of the constant turning of materials.

There were many such things, description of things, in odd ways. He often started to talk about a thing like this: "Suppose a man from Mars were to come down and look at the world." It's a very good way to look at the world. For example, when I was playing with my electric trains, he told me that there is a great wheel being turned by water which is connected by filaments of copper, which spread out and spread out and spread out in all directions; and then there are little wheels, and all

those little wheels turn when the big wheel turns. The relation between them is only that there is copper and iron, nothing else, no moving parts. You turn one wheel here, and all the little wheels all over the place turn, and your train is one of them. It was a wonderful world my father told me about. [...]

What science is, I think, may be something like this: There was on this planet an evolution of life to the stage that there were evolved animals, which are intelligent. I don't mean just human beings, but animals which play and which can learn something from experience (like cats). But at this stage each animal would have to learn from its own experience. They gradually develop, until some animal could learn from experience more rapidly and could even learn from another's experience by watching, or one could show the other, or he saw what the other one did. So there came a possibility that all might learn it, but the transmission was inefficient and they would die, and maybe the one who learned it died, too, before he could pass it on to others.

The question is, is it possible to learn more rapidly what somebody learned from some accident than the rate at which the thing is being forgotten, either because of bad memory or because of the death of the learner or inventors?

So there came a time, perhaps, when for some species the rate at which learning was increased reached such a pitch that suddenly a completely new thing happened; things could be learned by one animal, passed on to another, and another, fast enough that it was not lost to the race. Thus became possible an accumulation of knowledge of the race.

This has been called time-binding. I don't know who first called it this. At any rate, we have here some samples of those animals, sitting here trying to bind one experience to another, each one trying to learn from the other.

This phenomenon of having a memory for the race, of having an accumulated knowledge passable from one generation to another, was new in the world. But it had a disease in it. It was possible to pass on mistaken ideas. It was possible to pass on ideas which were not profitable for the race. The race has ideas, but they are not necessarily profitable.

So there came a time in which the ideas, although accumulated very slowly, were all accumulations not only of practical and useful things, but great accumulations of all types of prejudices, and strange and odd beliefs.

Then a way of avoiding the disease was discovered. This is to doubt that what is being passed from the past is in fact true, and to try to find out *ab initiio,* again from experience, what the situation is, rather than trusting the experience of the past in the form in which it is passed down. And that is what science is: the result of the discovery that it is worthwhile rechecking by new direct experience, and not necessarily trusting the race experience from the past. I see it that way. That is my best definition.

I would like to remind you all of things that you know very well in order to give you a little enthusiasm. In religion, the moral lessons are taught, but they are not just taught once—you are inspired again and again, and I think it is necessary to inspire again and again, and to remember the value of science for children, for grown-ups, and everybody else, in several ways; not only so that we will become better citizens, more able to control nature and so on. There are other things.

There is the value of the worldview created by science. There is the beauty and the wonder of the world that is discovered through the results of these new experiences. That is to say, the wonders of the content which I just reminded you of; that things move because the sun is shining, which is a deep idea, very strange and wonderful. (Yet, not everything

moves because the sun is shining. The earth rotates independent of the sun shining, and the nuclear reactions recently produced energy on the earth, a new source. Probably volcanoes are generally [powered by] a source different from the shining sun.)

The world looks so different after learning science. For example, the trees are made of air, primarily. When they are burned, they go back to air, and in the flaming heat is released the flaming heat of the sun which was bound in to convert the air into trees, and in the ash is the small remnant of the part which did not come from air, that came from the solid earth, instead.

These are beautiful things, and the content of science is wonderfully full of them. They are very inspiring, and they can be used to inspire others.

Another of the qualities of science is that it teaches the value of rational thought, as well as the importance of freedom of thought; the positive results that come from doubting that the lessons are all true. You must here distinguish—especially in teaching—the science from the forms or procedures that are sometimes used in developing science. It is easy to say, "We write, experiment, and observe, and do this or that." You can copy that form exactly. But great religions are dissipated by following form without remembering the direct content of the teaching of the great leaders. In the same way it is possible to follow form and call it science but it is pseudoscience. In this way we all suffer from the kind of tyranny we have today in the many institutions that have come under the influence of pseudoscientific advisers.

We have many studies in teaching, for example, in which people make observations and they make lists and they do statistics, but they do not thereby become established sci-

ence, established knowledge. They are merely an imitative form of science–like the South Sea Islanders making airfields, radio towers, out of wood, expecting a great airplane to arrive. They even build wooden airplanes of the same shape as they see in the foreigners' airfields around them, but strangely, they don't fly. The result of this pseudoscientific imitation is to produce experts, which many of you are–experts. You teachers who are really teaching children at the bottom of the heap, maybe you can doubt the experts once in a while. Learn from science that you *must* doubt the experts. As a matter of fact, I can also define science another way: Science is the belief in the ignorance of experts.

When someone says science teaches such and such, he is using the word incorrectly. Science doesn't teach it; experience teaches it. If they say to you science has shown such and such, you might ask, "How does science show it–how did the scientists find out–how, what, where?" Not science has shown, but this experiment, this effect, has shown. And you have as much right as anyone else, upon hearing about the experiments (but we must listen to *all* the evidence), to judge whether a reusable conclusion has been arrived at.

In a field which is so complicated that true science is not yet able to get anywhere, we have to rely on a kind of old-fashioned wisdom, a king of definite straightforwardness. I am trying to inspire the teacher at the bottom to have some hope, and some self-confidence in common sense, and natural intelligence. The experts who are leading you may be wrong.

I have probably ruined the system, and the students that are coming into Caltech no longer will be any good. I think we live in an unscientific age in which almost all the buffeting of communications and television words, books, and so on are unscientific. That doesn't mean they are bad, but they

are unscientific. As a result, there is a considerable amount of intellectual tyranny in the name of science.

Finally, a man cannot live beyond the grave. Each generation that discovers something from its experience must pass that on, but it must pass that on with a delicate balance of respect and disrespect, so that the race (now that it is aware of the disease to which it is liable) does not inflict its errors too rigidly on its youth, but it does pass on the accumulated wisdom, plus the wisdom that it may not be wisdom.

It is necessary to teach both to accept and to reject the past with a kind of balance that takes considerable skill. Science alone of all the subjects contains within itself the lesson of the danger of belief in the infallibility of the greatest teachers of the preceding generation.

So carry on. Thank you.

9

The Smartest Man in the World

Here is that wonderful 1979 interview of Feynman by Omni magazine. This is Feynman on what he knows and loves best–physics–and what he loves least, philosophy. ("Philosophers should learn to laugh at themselves.") Here Feynman discusses the work that earned him the Nobel Prize, quantum electrodynamics (QED); he then goes on to cosmology, quarks, and those pesky infinities that gum up so many equations.

"I think the theory is simply a way to sweep the difficulties under the rug," Richard Feynman said. "I am, of course, not sure of that." It sounds like the kind of criticism, ritually tempered, that comes from the audience after a controversial paper is presented at a scientific conference. But Feynman was at the podium, delivering a Nobel Prize winner's address. The theory he was questioning, quantum electrodynamics, has recently been called "the most precise ever devised"; its predictions are routinely verified to within one part in a million. When Feynman, Julian Schwinger, and Sin-Itiro Tomon-

aga independently developed it in the 1940s, their colleagues hailed it as "the great cleanup": a resolution of long-standing problems, and a rigorous fusion of the century's two great ideas in physics, relativity and quantum mechanics.

Feynman has combined theoretical brilliance and irreverent skepticism throughout his career. In 1942, after taking his doctorate at Princeton with John Wheeler, he was tapped for the Manhattan Project. At Los Alamos, he was a twenty-five-year-old whiz kid, awed neither by the titans of physics around him (Niels Bohr, Enrico Fermi, Hans Bethe) nor by the top-secret urgency of the project. The security staff was unnerved by his facility at opening safes—sometimes by listening to the tiny movements of the lock mechanism, sometimes by guessing which physical constant the safe's user had chosen as the combination. (Feynman hasn't changed since then; many of his students at Caltech have acquired safe-cracking skills along with their physics.)

After the war, Feynman worked at Cornell University. There, as he recounts in this interview, Bethe was the catalyst for his ideas on resolving "the problem of the infinities." The precise energy levels of electrons in hydrogen atoms, and the forces between the electrons (moving so rapidly that relativistic changes had to be taken into account), had already been the subject of pioneering work for three decades. Every electron, theory asserted, was surrounded by transient "virtual particles" which its mass-energy summoned up from vacuum; those particles in turn summoned up others—and the result was a mathematical cascade which predicted an infinite charge for every electron. Tomonaga had suggested a way around the problem in 1943, and his ideas became known just as Feynman at Cornell and Schwinger at Harvard were making the same crucial

step. All three shared the Nobel Prize for Physics in 1965. By then, Feynman's mathematical tools, the "Feynman integrals," and the diagrams he had invented to trace particle interactions were part of the equipment of every theoretical physicist. Mathematician Stanislaw Ulam, another Los Alamos veteran, cites the Feynman diagrams as "a notation that can push thoughts in directions that may prove useful or even novel and decisive." The idea of particles that travel backward in time, for example, is a natural outgrowth of that notation.

In 1950, Feynman moved to Caltech, in Pasadena. His accent is still unmistakably the transplanted New Yorker's, but Southern California seems the appropriate habitat for him: Among the "Feynman stories" his colleagues tell, his fondness for Las Vegas and nightlife in general looms large. "My wife couldn't believe I'd actually accept an invitation to give a speech where I'd have to wear a tuxedo," he says. "I did change my mind a couple of times." In the preface to *The Feynman Lectures on Physics,* widely used as a college text since they were collected and published in 1963, he appears with a maniacal grin, playing a conga drum. (On the bongos, it is said, he can play ten beats with one hand against eleven with the other; try it, and you may decide that quantum electrodynamics is easier.)

Among Feynman's other achievements are his contribution to understanding the phase changes of super-cooled helium, and his work with Caltech colleague Murray Gell-Mann* on the theory of beta decay of atomic nuclei. Both subjects are still far from final resolution, he points out; indeed, he does

---

*(1929–) Winner of the 1969 Nobel Prize in Physics for his contributions and discoveries concerning the classification of elementary particles and their interactions. In 1954 Gell-Mann and G. Zweig introduced the concept of quarks. *Ed.*

not hesitate to call quantum electrodynamics itself a "swindle" that leaves important logical questions unanswered. What kind of man can do work of that caliber while nursing the most penetrating doubts? Read on and find out.

Omni: To someone looking at high-energy physics from the outside, its goal seems to be to find the ultimate constituents of matter. It seems a quest we can trace back to the Greeks' atom, the "indivisible" particle. But with the big accelerators, you get fragments that are more massive than the particles you started with, and maybe quarks that can never be separated. What does that do to the quest?

Feynman: I don't think that ever was the quest. Physicists are trying to find out *how nature behaves;* they may talk carelessly about some "ultimate particle" because that's the way nature looks at a given moment, but . . . Suppose people are exploring a new continent, OK? They see water coming along the ground, they've seen that before, and they call it "rivers." So they say they're exploring to find the headwaters, they go upriver, and sure enough, there they are, it's all going very well. But lo and behold, when they get up far enough they find the whole system's different: There's a great big lake, or springs, or the rivers run in a circle. You might say, "Aha! They've failed!" but not at all! The *real* reason they were doing it was to explore the land. If it turned out not to be headwaters, they might be slightly embarrassed at their carelessness in explaining themselves, but no more than that. As long as it looks like the way things are built is wheels within wheels, then you're looking for the innermost wheel–but it might not be that way, in which case you're looking for whatever the hell it is that you find!

Omni: But surely you must have some guess about what you'll find; there are bound to be ridges and valleys and so on. . . ?

Feynman: Yeah, but what if when you get there it's all clouds? You can expect certain things, you can work out theorems about the topology of watersheds, but what if you find a kind of mist, maybe, with things coagulating out of it, with no way to distinguish the land from the air? The whole idea you started with is gone! That's the kind of exciting thing that happens from time to time. One is presumptuous if one says, "We're going to find the ultimate particle, or the unified field laws," or "*the*" anything. If it turns out surprising, the scientist is even more delighted. You think he's going to say, "Oh, it's not like I expected, there's no ultimate particle, I don't want to explore it"? No, he's going to say, "What the hell *is* it, then?"

Omni: You'd rather see that happen?

Feynman: Rather doesn't make any difference: I get what I get. You can't say it's *always* going to be surprising, either; a few years ago I was very skeptical about the gauge theories,* partly because I expected the strong nuclear interaction to be more different from electrodynamics than it now looks. I was expecting mist, and now it looks like ridges and valleys after all.

Omni: Are physical theories going to keep getting more abstract and mathematical? Could there be today a theorist like Faraday in the early nineteenth century, not mathematically sophisticated but with a very powerful intuition about physics?

Feynman: I'd say the odds are strongly against it. For one thing, you need the math just to understand what's been done so far. Beyond that, the behavior of subnuclear systems is so strange compared to the ones the brain evolved to deal

---

*Theories in particle physics that describe the various interactions between subatomic particles. *Ed.*

with that the analysis *has* to be very abstract: To understand ice, you have to understand things that are themselves very unlike ice. Faraday's models were mechanical–springs and wires and tense bands in space–and his images were from basic geometry. I think we've understood all we can from that point of view; what we've found in this century is different enough, obscure enough, that further progress will require a lot of math.

Omni: Does that limit the number of people who can contribute, or even understand what's being done?

Feynman: Or else somebody will develop a way of thinking about the problems so that we can understand them more easily. Maybe they'll just teach it earlier and earlier. You know, it's not true that what is called "abstruse" math is so difficult. Take something like computer programming, and the careful logic needed for that–the kind of thinking that mama and papa would have said was only for professors. Well, now it's part of a lot of daily activities, it's a way to make a living; their children get interested and get hold of a computer and they're doing the most crazy, wonderful things!

Omni: . . . with ads for programming schools on every matchbook!

Feynman: Right. I don't believe in the idea that there are a few peculiar people capable of understanding math, and the rest of the world is normal. Math is a human discovery, and it's no more complicated than humans can understand. I had a calculus book once that said, "What one fool can do, another can." What we've been able to work out about nature may look abstract and threatening to someone who hasn't studied it, but it was fools who did it, and in the next generation, all the fools will understand it.

There's a tendency to pomposity in all this, to make it all deep and profound. My son is taking a course in philosophy, and last night we were looking at something by Spinoza—and there was the most childish reasoning! There were all these Attributes, and Substances, all this meaningless chewing around, and we started to laugh. Now, how could we do that? Here's this great Dutch philosopher, and we're laughing at him. It's because there was no excuse for it! In that same period there was Newton, there was Harvey studying the circulation of the blood, there were people with methods of analysis by which progress was being made! You can take every one of Spinoza's propositions, and take the contrary propositions, and look at the world—and you can't tell which is right. Sure, people were awed because he had the courage to take on these great questions, but it doesn't do any good to have the courage if you can't get anywhere with the question.

Omni: In your published lectures, the philosopher's comments on science come in for some lumps. . .

Feynman: It isn't the philosophy that gets me, it's the pomposity. If they'd just *laugh* at themselves! If they'd just say, "I think it's like this, but von Leipzig thought it was like that, and he had a good shot at it, too." If they'd explain that this is their best guess. . . But so few of them do; instead, they seize on the possibility that there may not be any ultimate fundamental particle, and say that you should stop work and ponder with great profundity. "You haven't thought deeply enough, first let me define the world for you." Well, I'm going to investigate it *without* defining it!

Omni: How do you know which problem is the right size to attack?

Feynman: When I was in high school, I had this notion that you could take the importance of the problem and mul-

tiply by your chance of solving it. You know how a technically minded kid is, he likes the idea of optimizing everything . . . anyway, if you can get the right combination of those factors, you don't spend your life getting nowhere with a profound problem, or solving lots of small problems that others could do just as well.

Omni: Let's take the problem that won the Nobel Prize for you, Schwinger, and Tomonaga. Three different approaches: Was that problem especially ripe for solution?

Feynman: Well, quantum electrodynamics had been invented in the late 1920s by Dirac and others, just after quantum mechanics itself. They had it fundamentally correct, but when you went to calculate answers you ran into complicated equations that were very hard to solve. You could get a good first-order approximation, but when you tried to refine it with corrections these infinite quantities started to crop up. Everybody knew that for twenty years; it was in the back of all the books on quantum theory.

Then we got the results of experiments by Lamb* and Rutherford† on the shifts in energy of the electron in hydrogen atoms. Until then, the rough prediction had been good enough, but now you had a very precise number: 1060 megacycles or whatever. And everybody said dammit, this problem has to be solved . . . they'd known the theory had problems, but now there was this very precise figure.

So Hans Bethe took this figure and made some estimates of how you could avoid the infinities by subtracting this effect from that effect, so the quantities that would tend to go to in-

---

*Willis Lamb (1913–), winner of the 1955 Nobel Prize in Physics for his discoveries concerning the fine structure of the hydrogen spectrum. *Ed.*

†Lord Ernest Rutherford, 1st Baron Rutherford of Nelson (1871–1937), winner of the 1908 Nobel Prize in Chemistry for his investigations into the disintegration of the elements, and the chemistry of radioactive substances. *Ed.*

finity were stopped short, and they'd probably stop in this order of magnitude, and he came out with something around 1000 megacycles. I remember, he'd invited a bunch of people to a party at his house, at Cornell, but he'd been called away to do some consulting. He called up during the party and told me he'd figured this out on the train. When he came back he gave a lecture on it, and showed how this cut-off procedure avoided the infinities, but was still very *ad hoc* and confusing. He said it would be good if someone could show how it could be cleaned up. I went up to him afterwards and said, "Oh, that's easy, I can do that." See, I'd started to get ideas on this when I was a senior at MIT. I'd even cooked up an answer then–wrong, of course. See, this is where Schwinger and Tomonaga and I came in, in developing a way to turn this kind of procedure into solid analysis–technically, to maintain relativistic invariance all the way through. Tomonaga had already suggested how it could be done, and at this same time Schwinger was developing his own way.

So I went to Bethe with my way of doing it. The funny thing was, I didn't know how to do the simplest practical problems in this area–I should have learned long before, but I'd been busy playing with my own theory–so I didn't know how to find out if my ideas worked. We did it together on the blackboard, and it was wrong. Even worse than before. I went home and thought and thought, and decided I had to learn to solve examples. So I did, and I went back to Bethe and we tried it and it worked! We've never been able to figure out what went wrong the first time . . . some dumb mistake.

Omni: How far had it set you back?

Feynman: Not much: maybe a month. It did me good, because I reviewed what I'd done and convinced myself that it had to work, and that these diagrams I'd invented to keep things straight were really OK.

Omni: Did you realize at that time that they'd be called "Feynman diagrams," that they'd be in the books?

Feynman: No, not–I do remember one moment. I was in my pajamas, working on the floor with papers all around me, these funny-looking diagrams of blobs with lines sticking out. I said to myself, wouldn't it be funny if these diagrams really are useful, and other people start using them, and *Physical Review* has to print these silly pictures? Of course, I couldn't foresee–in the first place, I had no idea how many of these pictures there'd be in *Physical Review,* and in the second place, it never occurred to me that with everybody using them, they wouldn't look funny anymore. . .

[At this point the interview adjourned to Professor Feynman's office, where the tape recorder refused to start again. The cord, power switch, "record" button, all were in order; then Feynman suggested taking the tape cassette out and putting it in again.]

Feynman: There. See, you just have to know about the world. Physicists know about the world.

Omni: Take it apart and put it back together?

Feynman: Right. There's always a little dirt, or infinity, or something.

Omni: Let's follow that up. In your lectures, you say that our physical theories do well at uniting various classes of phenomena, and then X-rays or mesons or the like show up; "There are always many threads hanging out in all directions." What are some of the loose threads you see in physics today?

Feynman: Well, there are the masses of the particles: The gauge theories give beautiful patterns for the interactions, but not for the masses, and we need to understand this irregular set of numbers. In the strong nuclear interaction, we have this

theory of colored* quarks and gluons, very precise and completely stated, but with very few hard predictions. It's technically very difficult to get a sharp test of the theory, and that's a challenge. I feel passionately that that's a loose thread; while there's no evidence in conflict with the theory, we're not likely to make much progress until we can check hard predictions with hard numbers.

Omni: What about cosmology? Dirac's suggestion that the fundamental constants change with time, or the idea that physical law was different at the instant of the Big Bang?

Feynman: That would open up a lot of questions. So far, physics has tried to find laws and constants without asking where they came from, but we may be approaching the point where we'll be forced to consider history.

Omni: Do you have any guesses on that?

Feynman: No.

Omni: None at all? No leaning either way?

Feynman: No, really. That's the way I am about almost everything. Earlier, you didn't ask whether I thought that there's a fundamental particle, or whether it's all mist; I would have told you that I haven't the slightest idea. Now, in order to work hard on something, you have to get yourself believing that the answer's over *there,* so you'll dig hard there, right? So you temporarily prejudice or predispose yourself–but all the time, in the back of your mind, you're laughing. Forget what you hear about science without prejudice. Here, in an interview, talking about the Big Bang, I have no prejudices–but when I'm working, I have a lot of them.

---

*"Color" is actually a name physicists gave to a certain property of quarks and gluons not because they have any actual color, but for want of a better name for a new property of elementary particles. *Ed.*

Omni: Prejudices in favor of . . . what? Symmetry, simplicity. . . ?

Feynman: In favor of my mood of the day. One day I'll be convinced there's a certain type of symmetry that everybody believes in, the next day I'll try to figure out the consequences if it's not, and everybody's crazy but me. But the thing that's unusual about good scientists is that while they're doing whatever they're doing, they're not so sure of themselves as others usually are. They can live with steady doubt, think "maybe it's so" and act on that, all the time knowing it's only "maybe." Many people find that difficult; they think it means detachment or coldness. It's not coldness! It's a much deeper and warmer understanding, and it means you can be digging somewhere where you're temporarily convinced you'll find the answer, and somebody comes up and says, "Have you seen what they're coming up with over there?", and you look up and say "*Jeez! I'm in the wrong place!*" It happens all the time.

Omni: There's another thing that seems to happen a lot in modern physics: the discovery of applications for kinds of mathematics that were previously "pure," such as matrix algebra or group therapy. Are physicists more receptive now than they used to be? Is the time lag less?

Feynman: There never was any time lag. Take Hamilton's* quaternions: the physicists threw away most of this very powerful mathematical system, and kept only the part–the mathematically almost trivial part–that became vector analysis. But when the whole power of quaternions *was* needed, for quantum mechanics, Pauli† re-invented the system on the

*Sir William Rowan Hamilton (1805–1865), Irish mathematician who invented quaternions, an alternate construct to tensor and vector analysis. *Ed.*

†Wolfgang Pauli (1900–1958), winner of the 1945 Nobel Prize in Physics for his discovery of exclusion principle. *Ed.*

spot in a new form. Now, you can look back and say that Pauli's spin matrices and operators were nothing but Hamilton's quaternions . . . but even if physicists had kept the system in mind for ninety years, it wouldn't have made more than a few weeks' difference.

Say you've got a disease, Werner's granulomatosis or whatever, and you look it up in a medical reference book. You may well find that you then know more about it than your doctor does, although he spent all that time in medical school . . . you see? It's much easier to learn about some special, restricted topic than a whole field. The mathematicians are exploring in all directions, and it's quicker for a physicist to catch up on what he needs than to try to keep up with everything that might conceivably be useful. The problem I was mentioning earlier, the difficulties with the equations in the quark theories—it's the physicists' problem, and we're going to solve it, and maybe when we solve it we'll be doing mathematics. It's a marvelous fact, and one I don't understand, that the mathematicians had investigated groups and so on before they turned up in physics—but in regard to the speed of progress in physics, I don't think it's all that significant.

Omni: One more question from your lectures: you say there that "the next great era of awakening of human intellect may well produce a method of understanding the *qualitative* content of equations." What do you mean by that?

Feynman: In that passage I was talking about the Schrödinger* equation. Now, you can get from that equation to atoms bonding in molecules, chemical valences—but when

*Erwin Schrödinger (1887–1961), winner (with P.A.M. Dirac) of the 1933 Nobel Prize in Physics for the discovery of new, productive forms of atomic theory. *Ed.*

you look at the equation, you can see nothing of the wealth of phenomena that the chemists know about; or the idea that quarks are permanently bound so you can't get a free quark—maybe you can and maybe you can't, but the point is that when you look at the equations that supposedly describe quark behavior, you can't see why it should be so. Look at the equations for the atomic and molecular forces in water, and you can't see the way water behaves; you can't see turbulence.

Omni: That leaves the people with questions about turbulence—the meteorologists and oceanographers and geologists and airplane designers—kind of up the creek, doesn't it?

Feynman: Absolutely. And it might be one of those up-the-creek people who'll get so frustrated he'll figure it out, and at that point he'll be doing physics. With turbulence, it's not just a case of physical theory being able to handle only simple cases—we can't do *any*. We have no good fundamental theory at all.

Omni: Maybe it's the way the textbooks are written, but few people outside science appear to know just how quickly real, complicated physical problems get out of hand as far as theory is concerned.

Feynman: That's very bad education. The lesson you learn as you grow older in physics is that what we can do is a very small fraction of what there is. Our theories are really very limited.

Omni: Do physicists vary greatly in their ability to see the qualitative consequences of an equation?

Feynman: Oh, yes—but nobody is very good at it. Dirac said that to *understand* a physical problem means to be able to see the answer without solving equations. Maybe he exaggerated; maybe solving equations is experience you need to gain understanding—but until you do understand, you're just solving equations.

Omni: As a teacher, what can you do to encourage that ability?

Feynman: I don't know. I have no way to judge the degree to which I'm getting across to my students.

Omni: Will a historian of science someday trace the careers of your students as others have done with the students of Rutherford and Niels Bohr and Fermi?

Feynman: I doubt it. I'm disappointed with my students all the time. I'm not a teacher who knows what he's doing.

Omni: But you can trace influences the other way, say, the influence on you of Hans Bethe or John Wheeler. . . ?

Feynman: Sure. But I don't know the effect *I'm* having. Maybe it's just my character: I don't know. I'm not a psychologist or sociologist, I don't know how to understand people, including myself. You ask, how can this guy teach, how can he be motivated if he doesn't know what he's doing? As a matter of fact, I love to teach. I like to think of new ways of looking at things as I explain them, to make them clearer–but maybe I'm not making them clearer. Probably what I'm doing is entertaining myself.

I've learned how to live without knowing. I don't have to be sure I'm succeeding, and as I said before about science, I think my life is fuller because I realize that I don't know what I'm doing. I'm delighted with the width of the world!

Omni: As we came back to the office, you stopped to discuss a lecture on color vision you'll be giving. That's pretty far from fundamental physics, isn't it? Wouldn't a physiologist say you were "poaching"?

Feynman: Physiology? It has to be physiology? Look, give me a little time and I'll give a lecture on anything in physiology. I'd be delighted to study it and find out all about it, because I can *guarantee* you it would be very interesting. I don't know anything, but I do know that *everything is interesting* if you go into it deeply enough.

My son is like that, too, although he's much wider in his interests than I was at his age. He's interested in magic, in computer programming, in the history of the early church, in topology—oh, he's going to have a terrible time, there are so many interesting things. We like to sit down and talk about how different things could be from what we expected; take the Viking landers on Mars, for example, we were trying to think how many ways there could be life that they *couldn't* find with that equipment. Yeah, he's a lot like me, so at least I've passed on this idea that everything is interesting to at least one other person.

Of course, I don't know if that's a good thing or not. . . . You see?

10

Cargo Cult Science: Some Remarks on Science, Pseudoscience, and Learning How to Not Fool Yourself

The 1974 Caltech Commencement Address

Question: *What do witch doctors, ESP, South Sea Islanders, rhinoceros horns, and Wesson Oil have to do with college graduation?* Answer: *They're all examples the crafty Feynman uses to convince departing graduates that honesty in science is more rewarding than all the kudos and temporary successes in the world. In this address to Caltech's class of 1974, Feynman gives a lesson in scientific integrity in the face of peer pressure and glowering funding agencies.*

During the Middle Ages there were all kinds of crazy ideas, such as that a piece of rhinoceros horn would increase potency. (Another crazy idea of the Middle Ages is these hats we have on today—which is too loose in my case.) Then a

method was discovered for separating the ideas–which was to try one to see if it worked, and if it didn't work, to eliminate it. This method became organized, of course, into science. And it developed very well, so that we are now in the scientific age. It is such a scientific age, in fact, that we have difficulty in understanding how witch doctors could *ever* have existed, when nothing that they proposed ever really worked–or very little of it did.

But even today I meet lots of people who sooner or later get me into a conversation about UFOs, or astrology, or some form of mysticism, expanded consciousness, new types of awareness, ESP, and so forth. And I've concluded that it's *not* a scientific world.

Most people believe so many wonderful things that I decided to investigate why they did. And what has been referred to as my curiosity for investigation has landed me in a difficulty where I found so much junk to talk about that I can't do it in this talk. I'm overwhelmed. First I started out by investigating various ideas of mysticism, and mystic experiences. I went into isolation tanks (they're dark and quiet and you float in Epsom salts) and got many hours of hallucinations, so I know something about that. Then I went to Esalen, which is a hotbed of this kind of thought (it's a wonderful place; you should go visit there). Then I became overwhelmed. I didn't realize how *much* there was.

I was sitting, for example, in a hot bath and there's another guy and a girl in the bath. He says to the girl, "I'm learning massage and I wonder if I could practice on you?" She says OK, so she gets up on a table and he starts off on her foot–working on her big toe and pushing it around. Then he turns to what is apparently his instructor, and says, "I feel a kind of dent. Is that the pituitary?" And she says, "No, that's not the way it feels." I say, "You're a hell of a long way from the pi-

tuitary, man." And they both looked at me—I had blown my cover, you see—and she said, "It's reflexology." So I closed my eyes and appeared to be meditating.

That's just an example of the kind of things that overwhelm me. I also looked into extrasensory perception and PSI phenomena, and the latest craze there was Uri Geller, a man who is supposed to be able to bend keys by rubbing them with his finger. So I went to his hotel room, on his invitation, to see a demonstration of both mind reading and bending keys. He didn't do any mind reading that succeeded; nobody can read my mind, I guess. And my boy held a key and Geller rubbed it, and nothing happened. Then he told us it works better under water, and so you can picture all of us standing in the bathroom with the water turned on and the key under it, and him rubbing the key with his finger. Nothing happened. So I was unable to investigate that phenomenon.

But then I began to think, what else is there that we believe? (And I thought then about the witch doctors, and how easy it would have been to check on them by noticing that nothing really worked.) So I found things that even *more* people believe, such as that we have some knowledge of how to educate. There are big schools of reading methods and mathematics methods, and so forth, but if you notice, you'll see the reading scores keep going down—or hardly going up—in spite of the fact that we continually use these same people to improve the methods. *There's* a witch doctor remedy that doesn't work. It ought to be looked into; how do they know that their method should work? Another example is how to treat criminals. We obviously have made no progress—lots of theory, but no progress—in decreasing the amount of crime by the method that we use to handle criminals.

Yet these things are said to be scientific. We study them. And I think ordinary people with commonsense ideas are in-

timidated by this pseudoscience. A teacher who has some good idea of how to teach her children to read is forced by the school system to do it some other way—or is even fooled by the school system into thinking that her method is not necessarily a good one. Or a parent of bad boys, after disciplining them in one way or another, feels guilty for the rest of her life because she didn't do "the right thing," according to the experts.

So we really ought to look into theories that don't work, and science that isn't science.

I tried to find a principle for discovering more of these kinds of things, and came up with the following system. Anytime you find yourself in a conversation at a cocktail party in which you do not feel uncomfortable that the hostess might come around and say, "Why are you fellows talking shop?" or that your wife will come around and say, "Why are you flirting again?"—then you can be sure you are talking about something about which nobody knows anything.

Using this method, I discovered a few more topics that I had forgotten—among them the efficacy of various forms of psychotherapy. So I began to investigate through the library, and so on, and I have so much to tell you that I can't do it all. I will have to limit myself to just a few little things. I'll concentrate on the things more people believe in. Maybe I will give a series of speeches next year on all these subjects. It will take a long time.

I think the educational and psychological studies I mentioned are examples of what I would like to call Cargo Cult Science. In the South Seas there is a Cargo Cult of people. During the war they saw airplanes land with lots of good materials, and they want the same thing to happen now. So they've arranged to make things like runways, to put fires along the sides of the runways, to make a wooden hut for a

man to sit in, with two wooden pieces on his head like headphones and bars of bamboo sticking out like antennas–he's the controller–and they wait for the airplanes to land. They're doing everything right. The form is perfect. It looks exactly the way it looked before. But it doesn't work. No airplanes land. So I call these things Cargo Cult Science, because they follow all the apparent precepts and forms of scientific investigation, but they're missing something essential, because the planes don't land.

Now it behooves me, of course, to tell you what they're missing. But it would be just about as difficult to explain to the South Sea Islanders how they have to arrange things so that they get some wealth in their system. It is not something simple like telling them how to improve the shapes of the earphones. But there is *one* feature I notice that is generally missing in Cargo Cult Science. That is the idea that we all hope you have learned in studying science in school–we never explicitly say what this *is,* but just hope that you catch on by all the examples of scientific investigation. It is interesting, therefore, to bring it out now and speak of it explicitly. It's a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty–a kind of leaning over backwards. For example, if you're doing an experiment, you should report everything that you think might make it invalid–not only what you think is right about it: other causes that could possibly explain your results; and things you thought of that you've eliminated by some other experiment, and how they worked–to make sure the other fellow can tell they have been eliminated.

Details that could throw doubt on your interpretation must be given, if you know them. You must do the best you can–if you know anything at all wrong, or possibly wrong–to explain it. If you make a theory, for example, and advertise it,

or put it out, then you must also put down all the facts that disagree with it, as well as those that agree with it. There is also a more subtle problem. When you have put a lot of ideas together to make an elaborate theory, you want to make sure, when explaining what it fits, that those things it fits are not just the things that gave you the idea for the theory; but that the finished theory makes something else come out right, in addition.

In summary, the idea is to try to give *all* of the information to help others to judge the value of your contribution; not just the information that leads to judgment in one particular direction or another.

The easiest way to explain this idea is to contrast it, for example, with advertising. Last night I heard that Wesson Oil doesn't soak through food. Well, that's true. It's not dishonest; but the thing I'm talking about is not just a matter of not being dishonest, it's a matter of scientific integrity, which is another level. The fact that should be added to that advertising statement is that *no* oils soak through food, if operated at a certain temperature. If operated at another temperature, they *all* will–including Wesson Oil. So it's the implication which has been conveyed, not the fact, which is true, and the difference is what we have to deal with.

We've learned from experience that the truth will out. Other experimenters will repeat your experiment and find out whether you were wrong or right. Nature's phenomena will agree or they'll disagree with your theory. And, although you may gain some temporary fame and excitement, you will not gain a good reputation as a scientist if you haven't tried to be very careful in this kind of work. And it's this type of integrity, this kind of care not to fool yourself, that is missing to a large extent in much of the research in Cargo Cult Science.

A great deal of their difficulty is, of course, the difficulty of the subject and the inapplicability of the scientific method to the subject. Nevertheless, it should be remarked that this is not the only difficulty. That's *why* the planes don't land–but they don't land.

We have learned a lot from experience about how to handle some of the ways we fool ourselves. One example: Millikan measured the charge on an electron by an experiment with falling oil drops and got an answer which we now know not to be quite right. It's a little bit off, because he had the incorrect value for the viscosity of air. It's interesting to look at the history of measurements of the charge of the electron, after Millikan. If you plot them as a function of time, you find that one is a little bigger than Millikan's, and the next one's a little bit bigger than that, and the next one's a little bit bigger than that, until finally they settle down to a number which is higher.

Why didn't they discover that the new number was higher right away? It's a thing that scientists are ashamed of–this history–because it's apparent that people did things like this: When they got a number that was too high above Millikan's, they thought something must be wrong–and they would look for and find a reason why something might be wrong. When they got a number closer to Millikan's value, they didn't look so hard. And so they eliminated the numbers that were too far off, and did other things like that. We've learned those tricks nowadays, and now we don't have that kind of a disease.

But this long history of learning how to not fool ourselves–of having utter scientific integrity–is, I'm sorry to say, something that we haven't specifically included in any particular course that I know of. We just hope you've caught on by osmosis.

The first principle is that you must not fool yourself–and you are the easiest person to fool. So you have to be very careful about that. After you've not fooled yourself, it's easy not to fool other scientists. You just have to be honest in a conventional way after that.

I would like to add something that's not essential to the scientist, but something I kind of believe, which is that you should not fool the layman when you're talking as a scientist. I am not trying to tell you what to do about cheating on your wife, or fooling your girlfriend, or something like that, when you're not trying to be a scientist, but just trying to be an ordinary human being. We'll leave those problems up to you and your rabbi. I'm talking about a specific, extra type of integrity that is not lying, but bending over backwards to show how you're maybe wrong, that you ought to do when acting as a scientist. And this is our responsibility as scientists, certainly to other scientists, and I think to laymen.

For example, I was a little surprised when I was talking to a friend who was going to go on the radio. He does work on cosmology and astronomy, and he wondered how he would explain what the applications of this work were. "Well," I said, "there aren't any." He said, "Yes, but then we won't get support for more research of this kind." *I* think that's kind of dishonest. If you're representing yourself as a scientist, then you should explain to the layman what you're doing–and if they don't want to support you under these circumstances, then that's their decision.

One example of the principle is this: If you've made up your mind to test a theory, or you want to explain some idea, you should always decide to publish it whichever way it comes out. If we only publish results of a certain kind, we can make the argument look good. We must publish *both* kinds of results. For example–let's take advertising again–suppose

some particular cigarette has some particular property, like low nicotine. It's published widely by the company that this means it is good for you–they don't say, for instance, that the tars are a different proportion, or that something else is the matter with the cigarette. In other words, publication probability depends upon the answer. That should not be done.

I say that's also important in giving certain types of government advice. Supposing a senator asked you for advice about whether drilling a hole should be done in his state; and you decide it would be better in some other state. If you don't publish such a result, it seems to me you're not giving scientific advice. You're being used. If your answer happens to come out in the direction the government or the politicians like, they can use it as an argument in their favor; if it comes out the other way, they don't publish it at all. That's not giving scientific advice.

Other kinds of errors are more characteristic of poor science. When I was at Cornell, I often talked to the people in the psychology department. One of the students told me she wanted to do an experiment that went something like this–I don't remember it in detail, but it had been found by others that under certain circumstances, X, rats did something, A. She was curious as to whether, if she changed the circumstances to Y, they would still do A. So her proposal was to do the experiment under circumstances Y and see if they still did A.

I explained to her that it was necessary first to repeat in her laboratory the experiment of the other person–to do it under condition X to see if she could also get result A–and then change to Y and see if A changed. Then she would know that the real difference was the thing she thought she had under control.

She was very delighted with this new idea, and went to her professor. And his reply was, no, you cannot do that, because

the experiment has already been done and you would be wasting time. This was in about 1935 or so, and it seems to have been the general policy then to not try to repeat psychological experiments, but only to change the conditions and see what happens.

Nowadays there's a certain danger of the same thing happening, even in the famous field of physics. I was shocked to hear of an experiment done at the big accelerator at the National Accelerator Laboratory, where a person used deuterium. In order to compare his heavy hydrogen results to what might happen with light hydrogen, he had to use data from someone else's experiment on light hydrogen, which was done on different apparatus. When asked why, he said it was because he couldn't get time on the program (because there's so little time and it's such expensive apparatus) to do the experiment with light hydrogen on this apparatus because there wouldn't be any new result. And so the men in charge of programs at NAL are so anxious for new results, in order to get more money to keep the thing going for public relations purposes, they are destroying—possibly—the value of the experiments themselves, which is the whole purpose of the thing. It is often hard for the experimenters there to complete their work as their scientific integrity demands.

All experiments in psychology are not of this type, however. For example, there have been many experiments running rats through all kinds of mazes, and so on—with little clear result. But in 1937 a man named Young did a very interesting one. He had a long corridor with doors all along one side where the rats came in, and doors along the other side where the food was. He wanted to see if he could train the rats to go in at the third door down from wherever he started them off. No. The rats went immediately to the door where the food had been the time before.

The question was, how did the rats know, because the corridor was so beautifully built and so uniform, that this was the same door as before? Obviously there was something about the door that was different from the other doors. So he painted the doors very carefully, arranging the textures on the faces of the doors exactly the same. Still the rats could tell. Then he thought maybe the rats were smelling the food, so he used chemicals to change the smell after each run. Still the rats could tell. Then he realized the rats might be able to tell by seeing the lights and the arrangement in the laboratory like any commonsense person. So he covered the corridor, and still the rats could tell.

He finally found that they could tell by the way the floor sounded when they ran over it. And he could only fix that by putting his corridor in sand. So he covered one after another of all possible clues and finally was able to fool the rats so that they had to learn to go in the third door. If he relaxed any of his conditions, the rats could tell.

Now, from a scientific standpoint, that is an A-Number-1 experiment. That is the experiment that makes rat-running experiments sensible, because it uncovers the clues that the rat is really using–not what you think it's using. And that is the experiment that tells exactly what conditions you have to use in order to be careful and control everything in an experiment with rat-running.

I looked into the subsequent history of this research. The next experiment, and the one after that, never referred to Mr. Young. They never used any of his criteria of putting the corridor on sand, or being very careful. They just went right on running rats in the same old way, and paid no attention to the great discoveries of Mr. Young, and his papers are not referred to, because he didn't discover anything about the rats. In fact, he discovered *all* the things you have to do to dis-

cover something about rats. But not paying attention to experiments like that is a characteristic of Cargo Cult Science.

Another example is the ESP experiments of Mr. Rhine, and other people. As various people have made criticisms–and they themselves have made criticisms of their own experiments–they improve the techniques so that the effects are smaller, and smaller, and smaller until they gradually disappear. All the parapsychologists are looking for some experiment that can be repeated–that you can do again and get the same effect–statistically, even. They run a million rats–no, it's people this time–they do a lot of things and get a certain statistical effect. Next time they try it they don't get it anymore. And now you find a man saying that it is an irrelevant demand to expect a repeatable experiment. This is *science*?

This man also speaks about a new institution, in a talk in which he was resigning as Director of the Institute of Parapsychology. And, in telling people what to do next, he says that one of the things they have to do is be sure they only train students who have shown their ability to get PSI results to an acceptable extent–not to waste their time on those ambitious and interested students who get only chance results. It is very dangerous to have such a policy in teaching–to teach students only how to get certain results, rather than how to do an experiment with scientific integrity.

So I wish to you–I have no more time, so I have just one wish for you–the good luck to be somewhere where you are free to maintain the kind of integrity I have described, and where you do not feel forced by a need to maintain your position in the organization, or financial support, or so on, to lose your integrity. May you have that freedom. May I also give you one last bit of advice: Never say that you'll give a talk unless you know clearly what you're going to talk about and more or less what you're going to say.

11

It's as Simple as One, Two, Three

An uproarious tale of Feynman the precocious student experimenting—with himself, his socks, his typewriter, and his fellow students—to solve the mysteries of counting and of time.

When I was a kid growing up in Far Rockaway, I had a friend named Bernie Walker. We both had "labs" at home, and we would do various "experiments." One time, we were discussing something—we must have been eleven or twelve at the time—and I said, "But thinking is nothing but talking to yourself inside."

"Oh, yeah?" Bernie said. "Do you know the crazy shape of the crankshaft in a car?"

"Yeah, what of it?"

"Good. Now, tell me: How did you describe it when you were talking to yourself?"

So I learned from Bernie that thoughts can be visual as well as verbal.

Later on, in college, I became interested in dreams. I wondered how things could look so real, just as if light were hitting the retina of the eye, while the eyes are closed: Are the

nerve cells on the retina actually being stimulated in some other way–by the brain itself, perhaps–or does the brain have a "judgment department" that gets slopped up during dreaming? I never got satisfactory answers to such questions from psychology, even though I became very interested in how the brain works. Instead, there was all this business about interpreting dreams, and so on.

When I was in graduate school at Princeton a kind of dumb psychology paper came out that stirred up a lot of discussion. The author had decided that the thing controlling the "time sense" in the brain is a chemical reaction involving iron. I thought to myself, "Now, how the hell could he figure that?"

Well, the way he did it was, his wife had a chronic fever which went up and down a lot. Somehow he got the idea to test her sense of time. He had her count seconds to herself (without looking at a clock), and checked how long it took her to count up to 60. He had her counting–the poor woman–all during the day: When her fever went up, he found she counted quicker; when her fever went down, she counted slower. Therefore, he thought, the thing that governed the "time sense" in the brain must be running faster when she's got fever than when she hasn't got fever.

Being a very "scientific" guy, the psychologist knew that the rate of a chemical reaction varies with the surrounding temperature by a certain formula that depends on the energy of the reaction. He measured the differences in speed of his wife's counting, and determined how much the temperature changed the speed. Then he tried to find a chemical reaction whose rates varied with temperature in the same amounts as his wife's counting did. He found that iron reactions fit the pattern best. So he deduced that his wife's sense of time was governed by a chemical reaction in her body involving iron.

Well, it all seemed like a lot of baloney to me—there were so many things that could go wrong in his long chain of reasoning. But it *was* an interesting question: What *does* determine the "time sense"? When you're trying to count at an even rate, what does that rate depend on? And what could you do to yourself to change it?

I decided to investigate. I started by counting seconds—without looking at a clock, of course—up to 60 in a slow, steady rhythm: 1, 2, 3, 4, 5. . . . When I got to 60, only 48 seconds had gone by, but that didn't bother me: The problem was not to count for exactly one minute, but to count at a standard rate. The next time I counted to 60, 49 seconds had passed. The next time, 48. Then 47, 48, 49, 48, 49. . . . So I found I could count at a pretty standard rate.

Now, if I just sat there, without counting, and waited until I thought a minute had gone by, it was very irregular—complete variations. So I found it's very poor to estimate a minute by sheer guessing. But by counting, I could get very accurate.

Now that I knew I could count at a standard rate, the next question was—what affects the rate?

Maybe it has something to do with the heart rate. So I began to run up and down the stairs, up and down, to get my heart beating fast. Then I'd run into my room, throw myself down on the bed, and count up to 60.

I also tried running up and down the stairs and counting to myself *while* I was running up and down.

The other guys saw me running up and down the stairs, and laughed. "What are you doing?"

I couldn't answer them—which made me realize I couldn't talk while I was counting to myself—and kept right on running up and down the stairs, looking like an idiot.

(The guys at the graduate college were used to me looking like an idiot. On another occasion, for example, a guy came into my room–I had forgotten to lock the door during the "experiment"–and found me in a chair wearing my heavy sheepskin coat, leaning out of the wide-open window in the dead of winter, holding a pot in one hand and stirring with the other. "Don't bother me! Don't bother me!" I said. I was stirring Jell-O and watching it closely: I had gotten curious as to whether Jell-O would coagulate in the cold if you kept it moving all the time.)

Anyway, after trying every combination of running up and down the stairs and lying on the bed, surprise! The heart rate had no effect. And since I got very hot running up and down the stairs, I figured temperature had nothing to do with it either (although I must have known that your temperature doesn't really go up when you exercise). In fact, I couldn't find anything that affected my rate of counting.

Running up and down stairs got pretty boring, so I started counting while I did things I had to do anyway. For instance, when I put out the laundry, I had to fill out a form saying how many shirts I had, how many pants, and so on. I found I could write down "3" in front of "pants" or "4" in front of "shirts," but I couldn't count my socks. There were too many of them: I'm already using my "counting machine"–36, 37, 38–and here are all these socks in front of me–39, 40, 41. . . . How do I count the socks?

I found I could arrange them in geometrical patterns–like a square, for example: a pair of socks in this corner, a pair in that one; a pair over here, and a pair over there–eight socks.

I continued this game of counting by patterns, and found I could count the lines in a newspaper article by grouping the lines into patterns of 3, 3, 3, and 1 to get 10; then 3 of those patterns, 3 of those patterns, 3 of those patterns, and 1 of those

patterns made 100. I went right down the newspaper like that. After I had finished counting up to 60, I knew where I was in the patterns and could say, "I'm up to 60, and there are 113 lines." I found that I could even *read* the articles while I counted to 60, and it didn't affect the rate! In fact, I could do anything while counting to myself–except talk out loud, of course.

What about typing–copying words out of a book? I found that I could do that, too, but here my time was affected. I was excited: Finally, I've found something that appears to affect my counting rate! I investigated it more.

I would go along, typing the simple words rather fast, counting to myself 19, 20, 21, typing along, counting 27, 28, 29, typing along, until–What the hell is that word? Oh, yeah–and then continue counting 30, 31, 32, and so on. When I'd get to 60, I'd be late.

After some introspection and further observation, I realized what must have happened: I would interrupt my counting when I got to a difficult word that "needed more brains," so to speak. My counting rate wasn't slowing down; rather, the counting itself was being held up temporarily from time to time. Counting to 60 had become so automatic that I didn't even notice the interruptions at first.

The next morning, over breakfast, I reported the results of all these experiments to the other guys at the table. I told them all the things I could do while counting to myself, and said the only thing I absolutely could not do while counting to myself was talk.

One of the guys, a fella named John Tukey, said, "I don't believe you can read, and I don't see why you can't talk. I'll bet you I can talk while counting to myself, and I'll bet you you can't read."

So I gave a demonstration: They gave me a book and I read it for a while, counting to myself. When I reached 60 I said,

"Now!"–48 seconds, my regular time. Then I told them what I had read.

Tukey was amazed. After we checked him a few times to see what his regular time was, he started talking: "Mary had a little lamb; I can say anything I want to, it doesn't make any difference; I don't know what's bothering you"–blah, blah, blah, and finally, "Okay!" He hit his time right on the nose! I couldn't believe it!

We talked about it awhile, and we discovered something. It turned out that Tukey was counting in a different way: He was visualizing a tape with numbers on it going by. He would say, "Mary had a little lamb," and he would *watch* it! Well, now it was clear: He's "looking" at his tape going by, so he can't read, and I'm "talking" to myself when I'm counting, so I can't speak!

After that discovery, I tried to figure out a way of reading out loud while counting–something neither of us could do. I figured I'd have to use a part of my brain that wouldn't interfere with the seeing or speaking departments, so I decided to use my fingers, since that involved the sense of touch.

I soon succeeded in counting with my fingers and reading out loud. But I wanted the whole process to be mental, and not rely on any physical activity. So I tried to imagine the feeling of my fingers moving while I was reading out loud.

I never succeeded. I figured that was because I hadn't practiced enough, but it might be impossible: I've never met anybody who can do it.

By that experience Tukey and I discovered that what goes on in different people's heads when they *think* they're doing the same thing–something as simple as *counting*–is different for different people. And we discovered that you can externally and objectively test how the brain works: You don't have to ask a person how he counts and rely on his own ob-

servations of himself; instead, you observe what he can and can't do while he counts. The test is absolute. There's no way to beat it; no way to fake it.

It's natural to explain an idea in terms of what you already have in your head. Concepts are piled on top of each other: This idea is taught in terms of that idea, and that idea is taught in terms of another idea, which comes from counting, which can be so different for different people!

I often think about that, especially when I'm teaching some esoteric technique such as integrating Bessel functions. When I see equations, I see the letters in colors–I don't know why. As I'm talking, I see vague pictures of Bessel functions from Jahnke and Emde's book, with light-tan *j*s, slightly violet-bluish *n*s, and dark brown *x*s flying around. And I wonder what the hell it must look like to the students.

12

Richard Feynman Builds a Universe

In a previously unpublished interview made under the auspices of the American Association for the Advancement of Science, Feynman reminisces about his life in science: his terrifying first lecture to a Nobel laureate-packed room; the invitation to work on the first atomic bomb and his reaction; cargo-cult science; and that fateful predawn wake-up call from a journalist informing him that he'd just won the Nobel prize. Feynman's answer: "You could have told me that in the morning."

NARRATOR:

Mel Feynman was a salesman for a uniform company in New York City. On May 11, 1918, he welcomed the birth of his son Richard. Forty-seven years later, Richard Feynman received the Nobel Prize for Physics. In many ways, Mel Feynman had a lot to do with that accomplishment, as Richard Feynman relates.

FEYNMAN:

Well, before I was born, he [my father] said to my mother that "this boy is going to be a scientist." You can't say things like that in front of women's lib these days, but that is what

they said in those days. But he never told me to be a scientist....I learned to appreciate things I had known. There was never any pressure....Later when I got older, he'd take me for walks in the woods and show me the animals and birds and so on...tell me about the stars and the atoms and everything else. He'd tell me what it was about them that was so interesting. He had an attitude about the world and the way to look at it which I found was deeply scientific for a man who had no direct scientific training.

NARRATOR:

Richard Feynman is now professor of physics at the California Institute of Technology in Pasadena, where he has been since 1950. Part of his time he spends teaching and another part he devotes to theorizing about the tiny fragments of matter from which our universe is built. Throughout his career, his sometimes poetic imagination has carried him into many exotic areas: the mathematics involved in creating an atomic bomb, the genetics of a simple virus, and the properties of helium at extremely low temperatures. His Nobel prize-winning work toward developing the theory of quantum electrodynamics helped solve many physical problems more directly and more efficiently than had ever been possible. But again, what set that long train of accomplishments in motion were long walks in the woods with his father.

FEYNMAN:

He had ways of looking at things. He used to say, "Suppose we were Martians and we came down to the earth and then we would see these strange creatures doing things; what would we think? For instance," he would say, "to take an example, suppose that we never went to sleep. We're Martians, but we have a consciousness that works all the time, and we find these creatures who for eight hours every day stop and close their eyes and become more or less inert. We'd have an

interesting question to ask them. We'd say, 'How does it feel to do that all the time? What happens to your ideas? You're running along very well, you're thinking clearly—and what happens? Do they suddenly stop? Or do they go more and more slowly and stop, or exactly how do you turn off thoughts?'" Then later I thought about that a lot and I did experiments when I was in college to try to find out the answer to that—what happened to your thoughts when you went to sleep.

NARRATOR:

In his early days, Dr. Feynman planned to be an electrical engineer, to get his hands into physics and make it do useful things for him and the world around him. It didn't take him long to realize that he was really more interested in what made things work, in the theoretical and mathematical principles that underlie the operation of the universe itself. His mind became his laboratory.

FEYNMAN:

When I was young, what I call the laboratory was just a place to fiddle around, make radios and gadgets and photocells and whatnot. I was very shocked when I discovered what they call a laboratory in a university. That's a place where you are supposed to measure something very seriously. I never measured a damn thing in my laboratory. I just fiddled around and made things. That was the kind of lab I had when I was young and I thought entirely that way. I thought that was the way I was going to go. Well, in that lab, I had to solve certain problems. I used to repair radios. I had to, for example, get some resistance to put in line with some voltmeters so it would run in different scales. Things like that. So I began to find these formulas, electrical formulas, and a friend of mine had a book with electrical formulas in it and [it] had relations between the resistors. It had things like, the power is

the square of the current times the voltage. The voltage divided by the current is the resistance and all; it had six or seven formulas. It seemed to me that they were all related, they really weren't all independent, that one could come from the other. And so, I got to fiddling about and I understood from the algebra I had been learning in school how to do it. I realized that mathematics was somehow important in this business.

So I got more and more interested in the mathematical business associated with physics. In addition, mathematics by itself had a great appeal for me. I loved it all my life. [...]

NARRATOR:

After graduation from the Massachusetts Institute of Technology, Richard Feynman moved approximately 400 miles southwest to Princeton University, where he would eventually get his Ph.D. It was there, at the age of 24 that he gave his first formal lecture. It was a very eventful lecture, as it turned out.

FEYNMAN:

When I was an undergraduate I worked with Professor Wheeler* as a research assistant, and we had worked out together a new theory about how light worked, how the interaction between atoms in different places worked; and it was at that time an apparently interesting theory. So Professor Wigner†, who was in charge of the seminars there, suggested that we give a seminar on it, and Professor Wheeler said that

---

*John Archibald Wheeler (1911–), physicist, best known to the general public for having coined the term "black hole." *Ed.*

†Eugene P. Wigner (1902–1995), 1963 Nobel Prize in Physics, for his contributions to the theory of the atomic nucleus and elementary particles, through his work on symmetry principles. *Ed.*

since I was a young man and hadn't given seminars before, it would be a good opportunity to learn how to do it. So this was the first technical talk that I ever gave.

I started to prepare the thing. Then Wigner came to me and said that he thought the work was important enough that he'd made special invitations to the seminar to Professor Pauli, who was a great professor of physics visiting from Zurich; to Professor von Neumann, the world's greatest mathematician; to Henry Norris Russell, the famous astronomer; and to Albert Einstein, who was living near there. I must have turned absolutely white or something because he said to me, "Now don't get nervous about it, don't be worried about it. First of all, if Professor Russell falls asleep, don't feel bad, because he always falls asleep at lectures. When Professor Pauli nods as you go along, don't feel good, because he always nods, he has palsy," and so on. That kind of calmed me down a bit, but I was still worried. So Professor Wheeler promised me that he would answer all the questions and all that I would do would be to give the lecture.

So I remember coming in–you can imagine that first time, it was like going through fire. I had written all the equations on the blackboard way ahead of time so that all the blackboards were full of equations. People don't want so many equations . . . they want to understand the ideas better. And then I remember getting up to talk and there were these great men in the audience and it was frightening. And I can still see my own hands as I pulled out the papers from the envelope that I had them in. They were shaking. As soon as I got the paper out and started to talk, something happened to me which has always happened since and which is a wonderful thing. If I'm talking physics, I love the thing, I think only about physics, I don't worry where I am; I don't worry about anything. And everything went very easily. I simply explained

the whole business as best I could. I didn't think about who was there. I was thinking only about the problem I was explaining. And then at the end when the question time came, I had nothing to worry about because Professor Wheeler was going to answer them. Professor Pauli stood up—he was sitting next to Professor Einstein. He said, "I do not think this theory can be right because of this and this and that and the other thing and so forth, don't you agree, Professor Einstein?" Einstein said, "No-o-o-o," and that was the nicest no I ever heard.

NARRATOR:

It was at Princeton that Richard Feynman learned that even if he lived his entire life in the world of mathematics and theoretical physics, there was another world out there that would insist on making very practical demands of him. In those years the world was at war, and the United States had just begun to work on the atomic bomb.

FEYNMAN:

Just about that time, Bob Wilson came into my room to tell me about a project he was starting that had to do with making uranium for atomic bombs. He said there was a meeting at 3:00 and it was a secret, but he knew that when I knew what the secret was I would have to go along with it, so there was no harm in telling me. I said, "You made a mistake in telling me the secret. I am not going along with you. I'm just going back and do my work—back to working on my thesis." He went out of the room saying "We're going to have a meeting at 3:00." That [happened] in the morning. I started to pace the floor and think about the consequences of the bomb being in the hands of the Germans and all that stuff and decided that it was very exciting and important to do. So I was at the meeting at 3:00 and I stopped working on my degree.

The problem was that you had to separate the isotopes of uranium in order to make a bomb. The uranium came in two

isotopes and U_{235} was the reactive one and you wanted to separate them. Wilson had invented a scheme for doing the separation–making a beam of ions and bunching the ions–the velocity of the two isotopes at the same energy is slightly different. So if you make little lumps and they go down a long tube, one gets ahead of the other and you can separate them that way. That's the plan he had. I was theoretical by that time. What I was originally set to do was find out if the device as it was designed was at all practical; could it be done at all? There were a lot of questions about space charge limitations and so on and I deduced that it could be done.

NARRATOR:

Even though Feynman deduced that Wilson's method for separating uranium isotopes was indeed theoretically possible, another method was eventually used to produce uranium-235 for the atomic bomb. Nevertheless, there was still plenty for Richard Feynman and his high-level theorizing to do at the main laboratory in Los Alamos, New Mexico, charged with developing the bomb. After the war, he joined the staff of the Laboratory of Nuclear Studies at Cornell University. Today, he has mixed emotions about the work he did toward making the atomic bomb possible. Had he done the right thing or the wrong thing?

FEYNMAN:

No, I don't think that I was wrong exactly at the time I made the decision. I thought about it and I think correctly that it was very dangerous if the Nazis got it. There was, however, I think, an error in my thought in that after the Germans were defeated–that was much later, three or four years later–we were working very hard. I didn't stop; I didn't even consider that the motive for originally doing it was no longer there. And that's one thing I did learn, that if you have some reason for doing something that's very strong and you start

working at it, you must look around every once in a while and find out if the original motives are still right. At the time I made the decision, I think that was right, but to continue without thinking about it may have been wrong. I don't know what would have happened if I had thought about it. I may have decided to continue anyway, I don't know. But the point of not thinking about it when the original conditions that made [me make] the original decision had changed, that's a mistake.

NARRATOR:

After five stimulating years at Cornell, Dr. Feynman, like many other easterners before and after him, was lured to California and the equally stimulating surroundings of the California Institute of Technology. And there were other reasons.

FEYNMAN:

First of all, the weather is no good in Ithaca. Secondly, I kind of like going out to nightclubs and stuff like that.

Bob Bacher invited me to come out here to give a series of lectures on some work I had developed at Cornell. So I gave the lecture and then he said, "May I lend you my car?" I enjoyed that and I took his car and every night I went to Hollywood and the Sunset Strip and hung around there and had a good time, and that mixture of good weather and a wider horizon than is available in a small town in upper New York State is what finally convinced me to come here. It wasn't very hard. It wasn't a mistake. There was another decision that wasn't a mistake.

NARRATOR:

On the California Institute of Technology faculty, Dr. Feynman serves as Richard Chace Tolman Professor of Theoretical Physics. In 1954 he received the Albert Einstein Award, and in 1962 the Atomic Energy Commission gave him the E. O. Laurence Award for "especially meritorious

contributions to the development, use or control of atomic energy." Finally, in 1965, he received the greatest scientific award of all, the Nobel Prize. He shared it with Sin-Itiro Tomonaga of Japan and Julian Schwinger of Harvard. For Dr. Feynman, the Nobel Prize was a rude awakening.

FEYNMAN:

The telephone rang, the guy said [he was] from some broadcasting company. I was very annoyed to be awakened. That was my natural reaction. You know, you're half awake and you're annoyed. So the guy says, "We'd like to inform you that you've won the Nobel Prize." And I'm thinking to myself–I'm still annoyed, see–it didn't register. So I said, "You could have told me that in the morning." So he says, "I thought you'd like to know." Well, I said I was asleep and put the telephone back. My wife said, "What was that?" and I said, "I won the Nobel Prize." And she said, "Go on, you're kidding me." I've often tried to fool her but I can never fool her. Every time I try to fool her she sees through me, so this time she was wrong. She thought I was kidding. She thought it was some student, some drunken student or something. So she didn't believe me. But when the second telephone call came ten minutes later from another newspaper, I said to that fellow, "Yes, I've already heard it, leave me alone." Then I took the receiver off the hook and I thought I'd just go back to sleep and by 8:00 I'd put the receiver back on the hook. I couldn't go back to sleep and my wife couldn't either. I got up and walked around, and finally I put the receiver back and started to answer the phone.

Some short time after that, I had a ride in a taxi somewhere and the taxi driver is talking and I'm talking and I'm telling him my problems about how these guys ask me and I don't know how to explain it. He says, "I heard an interview of yours. I saw it on TV. The guy says to you, 'Would you please

explain what you did to get the prize in two minutes.' And you tried to do it and you're crazy. You know what I would have said? 'Hell, man, if I could have told you in two minutes, I wouldn't be worth the Nobel Prize.'" So that's the answer I gave after that. When somebody asks me, I always tell them, look, if I could explain it that easily, it wouldn't have been worth the Nobel Prize. It's not really fair, but it's kind of a fun answer.

NARRATOR:

As mentioned earlier, Dr. Feynman received the Nobel Prize for his contributions to developing a theory that would define the newly emerging field of quantum electrodynamics. It is, as Dr. Feynman puts it, "the theory of everything else." It does not apply to nuclear energy or the forces of gravity, but it does apply to the interaction of electrons with particles of light called photons. It underlies the way electricity flows, the phenomenon of magnetism, and the way that X-rays are produced and interact with other forms of matter. The "quantum" in quantum electrodynamics recognizes a theory of the mid-twenties which states that the electrons surrounding the nucleus of every atom are limited to certain quantum states or energy levels. They can exist only at those levels and nowhere in between. These quantized energy levels are determined by the intensity of the light that falls on the atom, among other things.

FEYNMAN:

One of the biggest and most important tools of theoretical physics is the wastebasket. You have to know when to leave it alone, hmmm? In fact, I learned almost everything I know about electricity, magnetism, and quantum mechanics and everything else in attempting to develop that theory. And what I got a Nobel Prize for ultimately was that in 1947, the regular people's theory, the ordinary theory which I was trying to fix by changing it, was in some trouble so I was trying to fix

it, but Bethe had found out that if you do just the right things, if you kind of forget some things and don't forget other things, do it just right, you can get the right answers compared to experiment and he made some suggestions to me. And I knew so much about electrodynamics by this time from having tried this crazy theory and written it in some 655 different forms that I knew how to do what he wanted, how to control and organize this calculation in a very smooth and convenient way and have powerful methods to do it. In other words, I used the stuff, the machinery which I had developed to evolve my own theory on the old theory–sounds like the obvious thing to do, but I didn't think of it for years–and found out it was extremely powerful by that time and I could do things by the old theory much faster than anyone else had before.

NARRATOR:

In addition to a lot of other things, Dr. Feynman's theory of quantum electrodynamics provides new insights into understanding the forces that hold matter together. It also adds a little bit more to what we know about the properties of the infinitesimally small, short-lived particles from which everything else in the universe is composed. As physicists have probed deeper and deeper into the structure of nature, they have found that what once seemed very simple may be very complex and what once seemed very complex may be very simple. Their tools are the high-powered atom smashers that can fracture atomic particles into smaller and smaller fragments.

FEYNMAN:

When we start out, we look at matter and we see many different phenomena–winds and waves and moon and all that kind of stuff. And we try to reorganize it. Is the motion of the wind like the motion of the waves and so forth? Gradually we find that many, many things are similar. It's not as big a vari-

ety as we think. We get all the phenomena and we get the principles underneath, and one of the most useful principles seemed to be the idea that things are made of other things. We found, for example, that all matter was made out of atoms, and then a large amount is understood as long as you understand the properties of atoms. And at first the atoms are supposed to be simple, but it turns out that in order to explain all the varieties, the phenomena of matter, the atoms have to be more complicated, and that there are 92 atoms. In fact, there are many more, because they have different weights. Then to understand the variety of the properties of atoms is the next problem. And we find that we can understand that if we make out that the atoms themselves are made of constituents—in this particular case, the nucleus around which the electrons go. And that all the different atoms are just different numbers of electrons. It's a beautifully unifying system that works.

All the different atoms are just the same thing with different numbers of electrons. However, the nuclei then differ. And so we start to study the nuclei. And there was a great variety as soon as we started experiments hitting nuclei together—Rutherford and so forth. From 1914 on, they discovered that they were complicated at first. But then it was realized that they could be understood if they're made of constituents, too. They are made of protons and neutrons. And they interact with some force that holds them together. In order to understand the nuclei, we have to understand that force a little bit better. Incidentally, in the case of atoms there was also a force; that's an electrical force and that we understood. So besides electrons there was also the electrical force, which we represent by photons of light. The light and electrical force are integrated into one thing called photons, so the outside world, so to speak, outside the nucleus is electrons and photons. And the theory of the behavior of elec-

trons is quantum electrodynamics and that's what I got the Nobel Prize for working on.

But now we go into the nucleus and find that they could be made of protons and neutrons, but there's this strange force. Trying to understand that force is the next problem. And various suggestions that there might be other particles were made by Yukawa,* and so we did experiments hitting the protons and neutrons together with higher energy and indeed new things came out, just like when you hit electrons together with high enough energy, photons came out. So we have these new things coming out. They were mesons. So it looked like Yukawa was right. We continued the experiment. And then what happened to us was that we got a tremendous variety of particles; not just one kind of photon, you see, but we hit photons and neutrons together and we got over 400 different kinds of particles–lambda particles and sigma particles. They're all different. And π mesons and K mesons and so on. Well, we incidentally also made muons, but they have nothing to do with neutrons and protons apparently. At least no more than electrons do. That's a strange extra part that we don't understand where it goes. It's just like an electron but heavier. So we have electrons and muons out here which don't interact strongly with these other things. These other things we call strongly interacting particles, or hadrons. And they include protons and neutrons and all the things you get immediately when you hit them together very hard. So now the problem is to try and represent the properties of all these particles in some organizing fashion. And that's a great game and we're all working on it. It's called high-energy physics or fundamental particle physics. It used to be called fundamen-

*Hideki Yukawa (1907–1981), winner of the 1949 Nobel Prize in Physics for predicting the existence of mesons. *Ed.*

tal particle physics, but nobody can believe that 400 different constituents are fundamental. Another possibility is that they are themselves made of some deeper constituent. And that seems to be a reasonable possibility. And so it turns out that a theory has been invented—the theory of quarks; that certain of these things like the proton, for instance, or neutron, are made of three objects called quarks.

NARRATOR:

No one has yet seen a quark, which is too bad, because they may represent the fundamental building block for all the other more complicated atoms and molecules that make up the universe. The name was chosen for no particular reason by Dr. Feynman's colleague, Murray Gell-Mann, some years ago. Somewhat to Dr. Gell-Mann's surprise, the Irish novel writer James Joyce had already anticipated that name thirty years earlier in his book, *Finnegan's Wake*. The key phrase was "three quarks for Muster Mark." This was even a bigger coincidence since, as Dr. Feynman explained, the quarks that make up the particles of the universe seem to come in threes. In the search for quarks, physicists are knocking protons and neutrons together at such high energies with the hope that they will break apart into their quark components in the process.

FEYNMAN:

All true, and one of the things that's been holding up the quark theory was that it's obviously cockeyed, because if the things were made of quarks, if we hit two protons together, we ought to produce three quarks sometimes. It turns out that in this quark model that we are talking about, the quarks carry very peculiar electrical charges. All the particles in the world that we know contain integral charges. Usually one electric charge plus or minus or nothing. But the theory of quarks has it that the quarks carry charges like minus a third or plus two

thirds of an electric charge. And if such a particle exists, it would be obvious, because the number of bubbles it would leave in a bubble chamber when it made a track would be much [smaller]. Say you had a charge of a third; then it whips up one ninth as many atoms as it turns out–the square–along the track, so there would be one ninth as many bubbles along the track as you would get for an ordinary particle. And that's obvious; if you see a lightly drawn track, there's something wrong. And they've looked and looked for such a track, and they haven't found them yet. So that's one of the serious problems. That's the excitement. Are we on the right track or are we walking around in the utter darkness when the answer is way over here to the right, or are we smelling it closely and just haven't quite got it right? And if we just get it right, we'll suddenly understand why that experiment looks different.

NARRATOR:

And what if these high-powered experiments with atom smashers and bubble chambers do show that the world is made of quarks? Will we ever be able to see them in a practical way?

FEYNMAN:

Well, for the problem of understanding the hadrons and the muons and so on, I can see at the present time no practical applications at all, or virtually none. In the past many people have said that they could see no applications and then later they found applications. Many people would promise under those circumstances that something's bound to be useful. However, to be honest–I mean he looks foolish; saying there will never be anything useful is obviously a foolish thing to do. So I'm going to be foolish and say these damn things will never have any application, as far as I can tell. I'm too dumb to see it. All right? So why do you do it? Applications aren't the only thing in the world. It's interesting in un-

derstanding what the world is made of. It's the same interest, the curiosity of man that makes him build telescopes. What is the use of discovering the age of the universe? Or what are these quasars that are exploding at long distances? I mean what's the use of all that astronomy? There isn't any. Nonetheless, it's interesting. So it's the same kind of exploration of our world that I'm following and it's curiosity that I'm satisfying. If human curiosity represents a need, the attempt to satisfy curiosity, then this is practical in the sense that it is that. That's the way I would look at it at the present time. I would not put out any promise that it would be practical in some economic sense.

NARRATOR:

As for science itself and what it means to all of us, Dr. Feynman says he is reluctant to philosophize on the subject. Nevertheless, that does not prevent him from coming up with some interesting and provocative ideas about what he believes science is and what it is not.

FEYNMAN:

Well, I'll say it is the same as it always was from the day it began. It's the pursuit of understanding of some subject or some thing based on the principle that what happens in nature is true and is the judge of the validity of any theory about it. If Lysenko says that you cut off rats' tails for 500 generations, then the new rats that are born will not have tails. (I don't know if he says that or not. Let's say Mr. Jones says that.) Then if you try it and it doesn't work, we know that it isn't true. That principle, the separation of the true from the false by experiment or experience, that principle and the resultant body of knowledge which is consistent with that principle, that is science.

To science we also bring, besides the experiment, a tremendous amount of human intellectual attempt at generalization.

So it's not merely a collection of all those things which just happen to be true in experiments. It's not just a collection of facts about what happens when you cut off [rats'] tails because it would be much too much for us to hold in our heads. We've found a large number of generalizations. For example, if it's true of rats and cats, we say it's true of mammals. Then we discover if it's true of other animals; then we discover it's true of plants, and finally it becomes a property of life to a certain extent that we don't inherit as an acquired characteristic. It's not exactly true, actually, absolutely. We later discovered experiments that show that cells can carry information through the mitochondria or something so that we modify it as we go along. But as all the principles must be as wide as possible, must be as general as possible, and still be in complete accord with experiment, that's the challenge.

You see, the problem of obtaining facts from experience—it sounds very, very simple. You just try it and see. But man is a weak character and it turns out to be much more difficult than you think to just try it and see. For instance, you take education. Some guy comes along and he sees the way people teach mathematics. And he says, "I have a better idea. I'll make a toy computer and teach them with it." So he tries it with a group of children, he hasn't got a lot of children, maybe somebody gives him a class to try it with. He loves what he's doing. He's excited. He understands completely what his thing is. The kids know that it's something new, so they're all excited. They learn very, very well and they learn the regular arithmetic better than the other kids did. So you make a test—they learn arithmetic. Then this is registered as a fact—that the teaching of arithmetic can be improved by this method. But it's not a fact, because one of the conditions of the experiment was that the particular man who invented it was doing the teaching. What you really want to know is, if

you just had this method described in a book to an average teacher (and you have to have average teachers; there are teachers all over the world and there must be many who are average), who then gets this book then tries to teach it with the method described, will it be better or not? In other words, what happens is that you get all kinds of statements of fact about education, about sociology, even psychology–all kinds of things which are, I'd say, pseudoscience. They've done statistics which they say they've done very carefully. They've done experiments which are not really controlled experiments. [The results] aren't really repeatable in controlled experiments. And they report all this stuff. Because science which is done carefully has been successful; by doing something like that, they think that they get some honor. I have an example.

In the Solomon Islands, as many people know, the natives didn't understand the airplanes which came down during the war and brought all kinds of goodies for the soldiers. So now they have airplane cults. They make artificial landing strips and they make fires along the landing strips to imitate the lights and this poor native sits in a wooden box he's built with wooden earphones with bamboo sticks going up to represent the antenna and turning his head back and forth, and they have radar domes made of wood and all kinds of things hoping to lure the airplanes to give goods to them. They're imitating the action. It's just what the other guy did. Well, a hell of a lot of our modern activity in many, many fields is that kind of science. Just like aviation. That's a science. The science of education, for example, is no science at all. It's a lot of work. It takes a lot of work to carve those things out, those wooden airplanes. But it doesn't mean that they are actually finding out something. Penology, prison reform–to understand why people do crimes; look at the world–we under-

stand it more and more with our modern understanding of these things. More about education, more about crime; the scores on the tests are going down and there's more people in prison; young people are committing crimes, we just don't understand it at all. It just isn't working, to discover things about these things by using the scientific method in the type of imitation which they are using now. Now whether the scientific method would work in these fields if we knew how to do it, I don't know. It's particularly weak in this way. There may be some other method. For example, to listen to the ideas of the past and the experience of people for a long time might be a good idea. It's only a good idea not to pay attention to the past when you have another independent source of information that you've decided to follow. But you've got to watch out who it is you're following if you're going to [ignore] the wisdom of the people who've looked at this thing and thought about it and unscientifically came to a conclusion. They have no less right to be right than you have to be right in modern times; to equally unscientifically come to a conclusion.

Well, how's that? Am I doing okay as a philosopher?

NARRATOR:

In this edition of the Future for Science–a taped series of interviews with Nobel laureates–you've heard Dr. Richard Feynman of the California Institute of Technology. The series has been prepared under the auspices of the American Association for the Advancement of Science.

13

The Relation of Science and Religion

In a kind of thought experiment, Feynman takes the various points of view of an imaginary panel to represent the thinking of scientists and spiritualists and discusses the points of agreement and of disagreement between science and religion, anticipating, by two decades, the current active debate between these two fundamentally different ways of searching for truth. Among other questions, he wonders whether atheists can have morals based on what science tells them, in the way that spiritualists can have morals based on their belief in God—an unusually philosophical topic for pragmatic Feynman.

In this age of specialization, men who thoroughly know one field are often incompetent to discuss another. The great problems of the relations between one and another aspect of human activity have for this reason been discussed less and less in public. When we look at the past great debates on these subjects, we feel jealous of those times, for we should have liked the excitement of such argument. The old problems, such as the relation of science and religion, are still with us, and I believe present as difficult dilemmas as ever, but

they are not often publicly discussed because of the limitations of specialization.

But I have been interested in this problem for a long time and would like to discuss it. In view of my very evident lack of knowledge and understanding of religion (a lack which will grow more apparent as we proceed), I will organize the discussion in this way: I will suppose that not one man but a group of men are discussing the problem, that the group consists of specialists in many fields—the various sciences, the various religions and so on—and that we are going to discuss the problem from various sides, like a panel. Each is to give his point of view, which may be molded and modified by the later discussion. Further, I imagine that someone has been chosen by lot to be the first to present his views, and I am he so chosen.

I would start by presenting the panel with a problem: A young man, brought up in a religious family, studies a science, and as a result he comes to doubt—and perhaps later to disbelieve in—his father's God. Now, this is not an isolated example; it happens time and time again. Although I have no statistics on this, I believe that many scientists—in fact, I actually believe that more than half of the scientists—really disbelieve in their father's God; that is, they don't believe in a God in a conventional sense.

Now, since the belief in a God is a central feature of religion, this problem that I have selected points up most strongly the problem of the relation of science and religion. Why does this young man come to disbelieve?

The first answer we might hear is very simple: You see, he is taught by scientists, and (as I have just pointed out) they are all atheists at heart, so the evil is spread from one to another. But if you can entertain this view, I think you know less of science than I know of religion.

Another answer may be that a little knowledge is dangerous; this young man has learned a little bit and thinks he knows it all, but soon he will grow out of this sophomoric sophistication and come to realize that the world is more complicated, and he will begin again to understand that there must be a God.

I don't think it is necessary that he come out of it. There are many scientists—men who hope to call themselves mature—who still don't believe in God. In fact, as I would like to explain later, the answer is not that the young man thinks he knows it all—it is the exact opposite.

A third answer you might get is that this young man really doesn't understand science correctly. I do not believe that science can disprove the existence of God; I think that is impossible. And if it is impossible, is not a belief in science and in a God—an ordinary God of religion—a consistent possibility?

Yes, it is consistent. Despite the fact that I said that more than half of the scientists don't believe in God, many scientists *do* believe in both science and God, in a perfectly consistent way. But this consistency, although possible, is not easy to attain, and I would like to try to discuss two things: Why it is not easy to attain, and whether it is worth attempting to attain it.

When I say "believe in God," of course, it is always a puzzle—what is God? What I mean is the kind of personal God, characteristic of the Western religions, to whom you pray and who has something to do with creating the universe and guiding you in morals.

For the student, when he learns about science, there are two sources of difficulty in trying to weld science and religion together. The first source of difficulty is this—that it is imperative in science to doubt; it is absolutely necessary, for progress in science, to have uncertainty as a fundamental

part of your inner nature. To make progress in understanding, we must remain modest and allow that we do not know. Nothing is certain or proved beyond all doubt. You investigate for curiosity, because it is *unknown,* not because you know the answer. And as you develop more information in the sciences, it is not that you are finding out the truth, but that you are finding out that this or that is more or less likely.

That is, if we investigate further, we find that the statements of science are not of what is true and what is not true, but statements of what is known to different degrees of certainty: "It is very much more likely that so and so is true than that it is not true"; or "such and such is almost certain but there is still a little bit of doubt"; or—at the other extreme—"well, we really don't know." Every one of the concepts of science is on a scale graduated somewhere between, but at neither end of, absolute falsity or absolute truth.

It is necessary, I believe, to accept this idea, not only for science, but also for other things; it is of great value to acknowledge ignorance. It is a fact that when we make decisions in our life, we don't necessarily know that we are making them correctly; we only think that we are doing the best we can—and that is what we should do.

Attitude of Uncertainty

I think that when we know that we actually do live in uncertainty, then we ought to admit it; it is of great value to realize that we do not know the answers to different questions. This attitude of mind—this attitude of uncertainty—is vital to the scientist, and it is this attitude of mind which the student must first acquire. It becomes a habit of thought. Once acquired, one cannot retreat from it anymore.

What happens, then, is that the young man begins to doubt everything because he cannot have it as absolute truth. So the question changes a little bit from "Is there a God?" to "How sure is it that there is a God?" This very subtle change is a great stroke and represents a parting of the ways between science and religion. I do not believe a real scientist can ever believe in the same way again. Although there are scientists who believe in God, I do not believe that they think of God in the same way as religious people do. If they are consistent with their science, I think that they say something like this to themselves: "I am almost certain there is a God. The doubt is very small." That is quite different from saying, "I know that there is a God." I do not believe that a scientist can ever obtain that view–that really religious understanding, that real knowledge that there is a God–that absolute certainty which religious people have.

Of course this process of doubt does not always start by attacking the question of the existence of God. Usually special tenets, such as the question of an afterlife, or details of the religious doctrine, such as details of Christ's life, come under scrutiny first. It is more interesting, however, to go right into the central problem in a frank way, and to discuss the more extreme view which doubts the existence of God.

Once the question has been removed from the absolute, and gets to sliding on the scale of uncertainty, it may end up in very different positions. In many cases it comes out very close to being certain. But on the other hand, for some, the net result of close scrutiny of the theory his father held of God may be the claim that it is almost certainly wrong.

Belief in God—and the Facts of Science

That brings us to the second difficulty our student has in trying to weld science and religion: Why does it often end up

that the belief in God—at least, the God of the religious type—is considered to be very unreasonable, very unlikely? I think that the answer has to do with the scientific things—the facts or partial facts—that the man learns.

For instance, the size of the universe is very impressive, with us on a tiny particle whirling around the sun, among a hundred thousand million suns in this galaxy, itself among a billion galaxies.

Again, there is the close relation of biological man to the animals, and of one form of life to another. Man is a latecomer in a vast evolving drama; can the rest be but a scaffolding for his creation?

Yet again, there are the atoms of which all appears to be constructed, following immutable laws. Nothing can escape it; the stars are made of the same stuff, and the animals are made of the same stuff, but in such complexity as to mysteriously appear alive—like man himself.

It is a great adventure to contemplate the universe beyond man, to think of what it means without man—as it was for the great part of its long history, and as it is in the great majority of places. When this objective view is finally attained, and the mystery and majesty of matter are appreciated, to then turn the objective eye back on man viewed as matter, to see life as part of the universal mystery of greatest depth, is to sense an experience which is rarely described. It usually ends in laughter, delight in the futility of trying to understand. These scientific views end in awe and mystery, lost at the edge in uncertainty, but they appear to be so deep and so impressive that the theory that it is all arranged simply as a stage for God to watch man's struggle for good and evil seems to be inadequate.

So let us suppose that this is the case of our particular student, and the conviction grows so that he believes that indi-

vidual prayer, for example, is not heard. (I am not trying to disprove the reality of God; I am trying to give you some idea of–some sympathy for–the reasons why many come to think that prayer is meaningless.) Of course, as a result of this doubt, the pattern of doubting is turned next to ethical problems, because, in the religion which he learned, moral problems were connected with the word of God, and if the God doesn't exist, what is his word? But rather surprisingly, I think, the moral problems ultimately come out relatively unscathed; at first perhaps the student may decide that a few little things were wrong, but he often reverses his opinion later, and ends with no fundamentally different moral view.

There seems to be a kind of independence in these ideas. In the end, it is possible to doubt the divinity of Christ, and yet to believe firmly that it is a good thing to do unto your neighbor as you would have him do unto you. It is possible to have both these views at the same time; and I would say that I hope you will find that my atheistic scientific colleagues often carry themselves well in society.

Communism and the Scientific Viewpoint

I would like to remark, in passing, since the word "atheism" is so closely connected with "communism," that the communist views are the antithesis of the scientific, in the sense that in communism the answers are given to all the questions–political questions as well as moral ones–without discussion and without doubt. The scientific viewpoint is the exact opposite of this; that is, all questions must be doubted and discussed; we must argue everything out–observe things, check them, and so change them. The democratic government is much closer to this idea, because there is discussion and a chance of modification. One doesn't launch the ship in

a definite direction. It is true that if you have a tyranny of ideas, so that you know exactly what has to be true, you act very decisively, and it looks good–for a while. But soon the ship is heading in the wrong direction, and no one can modify the direction anymore. So the uncertainties of life in a democracy are, I think, much more consistent with science.

Although science makes some impact on many religious ideas, it does not affect the moral content. Religion has many aspects; it answers all kinds of questions. First, for example, it answers questions about what things are, where they come from, what man is, what God is–the properties of God, and so on. Let me call this the metaphysical aspect of religion. It also tells us another thing–how to behave. Leave out of this the idea of how to behave in certain ceremonies, and what rites to perform; I mean it tells us how to behave in life in general, in a moral way. It gives answers to moral questions; it gives a moral and ethical code. Let me call this the ethical aspect of religion.

Now, we know that, even with moral values granted, human beings are very weak; they must be reminded of the moral values in order that they may be able to follow their consciences. It is not simply a matter of having a right conscience; it is also a question of maintaining strength to do what you know is right. And it is necessary that religion give strength and comfort and the inspiration to follow these moral views. This is the inspirational aspect of religion. It gives inspiration not only for moral conduct–it gives inspiration for the arts and for all kinds of great thoughts and actions as well.

Interconnections

These three aspects of religion are interconnected, and it is generally felt, in view of this close integration of ideas, that to

attack one feature of the system is to attack the whole structure. The three aspects are connected more or less as follows: The moral aspect, the moral code, is the word of God—which involves us in a metaphysical question. Then the inspiration comes because one is working the will of God; one is for God; partly one feels that one is with God. And this is a great inspiration because it brings one's actions in contact with the universe at large.

So these three things are very well interconnected. The difficulty is this: that science occasionally conflicts with the first of the three categories—the metaphysical aspect of religion. For instance, in the past there was an argument about whether the earth was the center of the universe—whether the earth moved around the sun or stayed still. The result of all this was a terrible strife and difficulty, but it was finally resolved—with religion retreating in this particular case. More recently there was a conflict over the question of whether man has animal ancestry.

The result in many of these situations is a retreat of the religious metaphysical view, but nevertheless, there is no collapse of the religion. And further, there seems to be no appreciable or fundamental change in the moral view.

After all, the earth moves around the sun—isn't it best to turn the other cheek? Does it make any difference whether the earth is standing still or moving around the sun? We can expect conflict again. Science is developing and new things will be found out which will be in disagreement with the present-day metaphysical theory of certain religions. In fact, even with all the past retreats of religion, there is still real conflict for particular individuals when they learn about the science and they have heard about the religion. The thing has not been integrated very well; there are real conflicts here—and yet morals are not affected.

As a matter of fact, the conflict is doubly difficult in this metaphysical region. Firstly, the facts may be in conflict, but even if the facts were not in conflict, the attitude is different. The spirit of uncertainty in science is an attitude toward the metaphysical questions that is quite different from the certainty and faith that is demanded in religion. There is definitely a conflict, I believe—both in fact and in spirit—over the metaphysical aspects of religion.

In my opinion, it is not possible for religion to find a set of metaphysical ideas which will be guaranteed not to get into conflicts with an ever-advancing and always-changing science which is going into an unknown. We don't know how to answer the questions; it is impossible to find an answer which someday will not be found to be wrong. The difficulty arises because science and religion are both trying to answer questions in the same realm here.

Science and Moral Questions

On the other hand, I don't believe that a real conflict with science will arise in the ethical aspect, because I believe that moral questions are outside of the scientific realm.

Let me give three or four arguments to show why I believe this. In the first place, there have been conflicts in the past between the scientific and the religious view about the metaphysical aspect and, nevertheless, the older moral views did not collapse, did not change.

Second, there are good men who practice Christian ethics and who do not believe in the divinity of Christ. They find themselves in no inconsistency here.

Thirdly, although I believe that from time to time scientific evidence is found which may be partially interpreted as giving some evidence of some particular aspect of the life of Christ,

for example, or of other religious metaphysical ideas, it seems to me that there is no scientific evidence bearing on the Golden Rule. It seems to me that that is somehow different.

Now, let's see if I can make a little philosophical explanation as to why it is different—how science cannot affect the fundamental basis of morals.

The typical human problem, and one whose answer religion aims to supply, is always of the following form: Should I do this? Should we do this? Should the government do this? To answer this question we can resolve it into two parts: First—If I do this, what will happen?—and second—Do I want that to happen? What would come of it of value—of good?

Now a question of the form: If I do this, what will happen? is strictly scientific. As a matter of fact, science can be defined as a method for, and a body of information obtained by, trying to answer only questions which can be put into the form: If I do this, what will happen? The technique of it, fundamentally, is: Try it and see. Then you put together a large amount of information from such experiences. All scientists will agree that a question—any question, philosophical or other—which cannot be put into the form that can be tested by experiment (or, in simple terms, that cannot be put into the form: If I do this, what will happen?) is not a scientific question; it is outside the realm of science.

I claim that whether you want something to happen or not—what value there is in the result, and how you judge the value of the result (which is the other end of the question: Should I do this?), must lie outside of science because it is not a question that you can answer only by knowing what happens; you still have to *judge* what happens—in a moral way. So, for this theoretical reason I think that there is a complete consistency between the moral view—or the ethical aspect of religion—and scientific information.

Turning to the third aspect of religion—the inspirational aspect—brings me to the central question that I would like to present to this imaginary panel. The source of inspiration today—for strength and for comfort—in any religion is very closely knit with the metaphysical aspect; that is, the inspiration comes from working for God, for obeying his will, feeling one with God. Emotional ties to the moral code—based in this manner—begin to be severely weakened when doubt, even a small amount of doubt, is expressed as to the existence of God; so when the belief in God becomes uncertain, this particular method of obtaining inspiration fails.

I don't know the answer to this central problem—the problem of maintaining the real value of religion, as a source of strength and of courage to most men, while, at the same time, not requiring an absolute faith in the metaphysical aspects.

The Heritages of Western Civilization

Western civilization, it seems to me, stands by two great heritages. One is the scientific spirit of adventure—the adventure into the unknown, an unknown which must be recognized as being unknown in order to be explored; the demand that the unanswerable mysteries of the universe remain unanswered; the attitude that all is uncertain; to summarize it—the humility of the intellect. The other great heritage is Christian ethics—the basis of action on love, the brotherhood of all men, the value of the individual—the humility of the spirit.

These two heritages are logically, thoroughly consistent. But logic is not all; one needs one's heart to follow an idea. If people are going back to religion, what are they going back to? Is the modern church a place to give comfort to a man who doubts God—more, one who disbelieves in God? Is the modern church a place to give comfort and encouragement

to the value of such doubts? So far, have we not drawn strength and comfort to maintain the one or the other of these consistent heritages in a way which attacks the values of the other? Is this unavoidable? How can we draw inspiration to support these two pillars of Western civilization so that they may stand together in full vigor, mutually unafraid? Is this not the central problem of our time?

I put it up to the panel for discussion.

Intro

It's interesting that some people find science so easy and others find it kind of dull and difficult especially kids you know, some of them are just eat it up and I don't know why it is, it's the same perhaps for all subjects.

For instance lot of people love music and I could never carry a tune and I lose a great deal a pleasure out of that and I think that people lose a lot of pleasure who find the science dull In the case of science, I think that one of the things that make it very difficult is that it takes a lot of imagination It's very hard to imagine all the crazy things that things really are like

Jiggling Atoms

Nothing's really as it seems, we used to get you know hot and cold and all that hot and cold is the speeds that the atoms are jiggling if they jiggle more it corresponds to hotter and colder is jiggling less So if you have a bunch of atom, like a cup of coffee or something, sitting on a table and the atoms are jiggling a great deal and they bounce against the cup and the cup then gets shaking and the atoms in the cup shake and the bounce against each other and the heat heats the cup and heats everything else and then hot things spread that heat to other by mere contact because the atoms that are jiggling a lot in the hot thing shake the ones that are jiggling only a little bit in the cold thing so that the hot (heat we say) goes into the cold thing, it spreads but what is spreading is just jiggling, an irregular motion, but it is easy to kind of understand.

It brings up another thing that's kind of curious:

that I say that things jiggle and if you're used to balls bouncing you know they slow up and stop after a while but we have to imagine with the atoms a perfect elasticity they never lose any energy every time they bounce they keep on bouncing all the time they don't lose anything they're perpetually moving. And that the things that happen when we say something loses energy if a ball comes down and bounces, it shakes irregularly some of the atoms in the floor and when it comes up again, it leaves some of the atoms moving, they jiggling

So as it bounces, it is passing its extra energies, its extra motions to little patches on the floor each time it rebounces and it loses a little heat each time until it settles down, we say as the falling motion stops But what's left is the floor is shaking more than it was before

and the atoms in the ball are shaking more than they were before that the organized motion of all these atoms moving the same way falling down and the quiet floor, is now transformed into a ball sitting on the ground But all the motion is still there in the form of energy of motion

in the form of the jiggling of the floor which is a little bit warmer (unbelievable!).

But anybody who's hammered a great deal on something knows that it's true that if you pound something and hit it a lot you can feel the temperature difference it heats up it heats up simply because you're jiggling it this picture of atoms is a beautiful one that you can keep looking at all kinds of things this way you see a little drop of water a tiny drop and the atoms attract each other they like to be next to each other they want as many partners as they can get

Now the guys at the surface have only partners on one side here in the air on the other side so they're trying to get in. And you can imagine this team of people, this team in people, all moving very fast all try (to get) to have as many partners as possible and the guys on the edge are very unhappy and nervous and they keep pounding in trying to get in, and that makes a tight ball instead of a flat and that's what you know surface tension

when you realize when you see how sometimes a water drop sits like this on a table then you start to imagine why it's like that because everybody is trying to get in to the water and at the same time while all this is happening, other atoms leaving the surface and the water drop is slowly disappearing I find myself trying to imagine all kinds of things all the time and I get a kick out of it like a runner gets a kick out of sweating I GET A KICK of thinking about these things! I can't stop I mean if you may I could talk forever If you could cool off the water so that the jiggling is less and less, it jiggles slower and slower then the atoms get stuck in a place, they like to be with their friend there's force of attraction and they get packed together, they're not rolling over each other, they're in a nice pattern like oranges in a crate in a nice organized pattern, all just jiggling in place but not having enough motion to get loose of their own place and to break the structure down And that what I'm describing is a solid, it's ice, it has a structure. If you held the atom in one end in a certain position all the rest are lining up in a position sticking out and it’s solid at the end whereas if you heat that harder then they begin to get loose and roll all over each other and that's the liquid. And if you heat that still harder, then they bounce still harder and they simply bounce apart from each other and they're just individuals, I said atoms, these are really little groups of atoms: molecule which come flying and hit and although they have a tendency to stick, they're moving too fast their hands don't grab so to speak, as they pass and they fly apart again and this is the gas we call steam. You can get all kinds of understanding. When I was a kid with this "air", I was always interested in. I've noticed that when I pumped up my tires on the bicycle you can learn a lot by having a bicycle I'd pump up the tires that the pump would get hot and that also understand we see as the pump handle comes down and the atoms are coming up against it and bouncing off and it's moving in the ones that are coming off have a bigger speed than the ones that are coming in so that as it comes down and each time they collide it speeds them up and so they're hotter when you compress the gas it heats and when you pull the piston back out then the atoms which are coming faster than the piston feel receiving or sort of a give it gives and it comes out with less energy it's like going up against something which is soft and yielding it go boom boom and it loses so as you pull the piston out and the atoms are hit they lose their speed and they cool off and gases are cool when they expand and the fun of it is that all these things which you see or you notice in the world about it the pump heats the gas and the gas cools when it expands or the steam evaporates until you cover the cover and all these things you can understand from these simple pictures and that's kind of a lot of fun to think about I don't want to take this stuff seriously I think we should just have fun imagining it not worry about there's no teacher going to ask you questions at the end otherwise it's a horrible subject

Fire

the atoms like each other the different degrees oxygen for instance in the air would like to be next to carbon and if they're getting near each other they snap together if they're not too close though they repel and they go apart so they don't know that they could snap together it's just as if you had a ball that was trying to climb a hill and there was a hole it could go into like a volcano hole a deep one it's rolling along it doesn't go down in the deep hole because if it starts to climb the hill and then rolls away again but if you made it go fast enough it'll fall into the hole and so if it's have something like wood in oxygen there's carbon in the wood from a tree and the oxygen

comes and hits it carbon but not hot enough it just goes away again the air is always coming nothing's happening if you can get it faster by heating it up somehow somewhere somehow get it started a few of them come fast they go over the top so to speak they come close enough to the carbon and snap in and that gives a lot of jiggly motion which might hit some other atoms making those go faster so they can climb up and bump against other carbon atoms and they jiggle and they make other jiggle, and you get an horrible catastrophe which is one after the other all these things are going faster and faster and snapping in and the whole thing is changing that catastrophe is a fire it's just a way of looking at it and these things are happening they perpetual once it gets started it keeps on going the heat makes the other atoms capable of reaching to make more heat to make other atoms and so on so this terrible snapping is producing a lot of jiggling and if I put with all lack activity of the atoms there and I put a cup of coffee over that mess of wood that's doing this it's going to get a lot of jiggling so that's what the heat of the fire is and then of course uh you see what's happening when you start thinking, just go on and on wonder how did it get started why is it that the wood's been sitting around all this time with the oxygen all this time and it didn't do this earlier or something where did I get this from？ well it came from a tree and the substance of a tree is carbon, where did that come from? that comes from the air it's carbon dioxide from the air people look at trees and they think it comes out of the ground the plants grow out of the ground but if you ask where the substance comes from you find out where do they come from the trees come out of the air? they surely come out of you know they come out of the air the carbon dioxide in the air goes into the tree and it changes it kicking out the oxygen and uh pushing the oxygen away from the carbon and leaving the carbon substance with water water comes out of the ground you see only it had to get in there it came out of the air didn't it it came down from the sky so in fact most of a tree almost all of the tree is out of the ground I'm sorry it's out of the air there's a little bit from the ground some minerals and so forth now of course I told you the oxygen and we… oxygen carbon stick together very tight how is that the tree is so smart to take the carbon dioxide which is carbon and oxygen nicely combined and undo that so easy? Ah! Life! Life has some mysterious force! No! The sun is shining, and this sunlight comes down and knocks this oxygen away from the carbon, so it takes sunlight to get the plant to work!
and so the sun, all the time, is doing the work of separating the oxygen away from the carbon the oxygen is some kind of terrible by-product, which it spits back into the air
and leaving the carbon and water and stuff to make the substance of the tree and then we take the substance of the tree to get the fireplace and there's all the oxygen made by these trees and all the carbons would much prefer to be close together again and once you let the heat to get it started it continues and make an awful lot of activity while it's going back together again
and all those nice light and everything comes out and everything is being undone you're going from carbon and oxygen back to carbon dioxide and the light and heat that's coming out, that's the light and heat of the sun that went in so it's sort of stored sun that is coming out when you burned a log
next question: how is the sun so jiggly, so hot?
I gotta stop somewhere; I leave you something to imagine

Rubber Bands
most elastic things like steel springs and so on is nothing but this electrical thing pulling back

you pull the atoms up a little bit apart when you bend something and then they try to come back together again but rubber bands work on a different principle there there's some long molecules like chains and other little ones that are shaking all the time that are bombarding them these chains and the chains are all kind of kinky and knockabout in shape when you pull open the rubber band the strings get straighter but these strings are being bombarded on the side by these other atoms trying to shorten them by kinking them so it pulls back it's trying to pull back and it's pulling back only because of the heat so if you heat a rubber band it'll pull strong more strongly for instance if you hang a weight with a rubber band put a little match to it it's kind of fun to watch it rise because heats want and there's another thing you can check that this idea is right that is heat that drives a rubber band if you pull the band out just like when we push the piston and the gas if you pull the band out this tightening string hitting those molecules makes them move faster so it's warmer and if you take the band and let it in then the molecules hitting the strings which sort of give as the thing hits it they give in to the soft like and they lose energy when they hit these retiring band, string so it cools and there is a little way you can do this you're not very sensitive it's a small effect and if you take a fairly wide rubber band and put it between your lips and pull it out you'll certainly notice it's hotter and if you then hold it out and let it in you'll notice it's cooler at least you'll notice a certain difference in whether what happens when you expand it and when you contract it and that's i've always found rubber bands fascinating to think that when they're sitting on an old package of papers for a long time holding those papers together it's done by a perpetual pounding pounding pounding and the atoms that get these chains to hold it, try to kink them and try kink them year after year well rubber bands don't last that long but anyhow for a long time trying to hold this whole thing together.

The world is a dynamic mess of jiggling things if you look at it right And if you magnify, you will hardly see a little thing anymore because everything is jiggling in its own pattern, and there's a lot of little balls It's lucky that we have such a large scale of view of everything that we can see these as "things"without having worry about all these little atoms all the time.

Magnets

If you get hold of two magnet and you push them you can feel this pushing between them. Turn around the other way and they slam together. Now, what is it, the feeling between those two magnets? what do you mean "what's the feeling between two magnets when you hold them"? well, there's something there, isn't it? I mean that the sensation that they're something there when you push the two magnets together. Listen to my question what is the meaning when you say that there's there's a feeling of course you feel it. Now what do you want to know?
What I want to know is what's going on, between these two bits of matter?
magnets repel each other Well then, what does that mean or why are they doing that or how are they doing that? I'm not saying... That's a perfectly reasonable question. Of course it's a reasonab... it's an excellent question. Okay? But the problem that you are asking, you see, when you ask why something happens
how does a person answers "why something happens?" for example
Aunt Minnie is at the hospital. Why? Because she slip. she went out and she slipped on the ice and broke her hip

that satisfies people. It satisfies but wouldn't satisfy someone who came from another planet knew nothing about things first you understand "Why when you break your hip you go to the hospital? "How do you get to the hospital when the hip is broken"
well because her husband seeing that she had her hip broken called the hospital up and send somebody to get her all that is understood by people now when you explain a "Why"
you have to be in some framework that you allow something to be true Otherwise you are perpetually asking why Why did the husband call up the hospital? Because the husband is interested in his wife's welfare. Not always some husbands aren't interested in their wives' welfare when they are drunk and angry So you begin to get very interesting understanding of the world and all its complications in order to… if you try to follow anything up you go deeper and deeper in various directions. for example you could go “why did she slip upon the ice?" Well ice is lippery everybody knows that no problem
But you ask "why is ice slippery?" that's kind of curious, ice is extremely slippery it's very interesting
You say "How does it work?" you could either say "I'm satisfied that you have answered me ice is slippery, that explains it"
or you could go on and say "Why is ice slippery?” and then you're involved with something because there are not many things as slippery as ice It's very hard to get greasy stuff, but that's a sort of wet slimy
But a solid that is so slippery? Because it is in the case of ice
when you stand on it (they say), momentarily the pressure melts the ice a little bit so you get a sort of instantaneous water surface on which you are slipping
Why on ice and not on other things? Because ice expands… water expands when it freezes so the pressure tries to undo the expansion and melts it. It is capable about melting it But other substances contract when they're freezing
and when you push them they're just satisfied to be solid why does water expand when it freezes and another substance don't expand when they freeze all right? I'm not answering the question but I am telling you how difficult a "why" question is.
You have to know what it is that you are permitted to understand and allow to be understood and known and what it is you're not you'll notice in this example that the more I ask why, it gets interesting after all that's my idea that the deeper thing is the more interesting in it and you can even go further and say "Why did she fall down when she slip?"
That has to do with gravity and involves all other planets, and everything else never mind, it goes on and on!
Now when you ask for example "Why two magnets repel?" there are many different levels.
It depends on whether you are a student of physics or an ordinary person who doesn’t know anything or not. If you are somebody that doesn't know anything about all I can say is that it is the magnetic force that makes things repel And that you are feeling that force. You see, that is very strange because I don't feel any kind of force like that in other circumstances. When you turn them in the other way they attract there's a very analogous force, electrical force which is the same kind of a question and you say that's also very weird but you're not at all disturbed by the fact that when you put your hand on the chair, it pushes you back.
But we have find that looking at it that it is the same force as a matter of fact the electrical force (not magnetic exactly in that case)

but it is the same electric repulsions that are involved in keeping you finger away from the chair because everything is made out... it is electrical force in minor, microscopic details there are other forces involved, but they are connected to electrical force It turns out that the magnetic and the electric forces with which I wish to explain these things this repulsion in the first place is what ultimately is the deeper thing that we have to start that we can start with to explain many other things that looks like they were... Everybody would just accept them.
You know you cannot put your hand through the chair, that's taken for granted. But you can't put your hand through the chair when you look at it more closely: "Why?"
it involves these same repulsive forces that appear in magnets The situation is then to have to explain is:
"why in the magnet it goes over a bigger distance than ordinarily?" There it has to do with the fact that in iron, all the electrons are spinning in the same direction, they all get lined up and they magnify the effect of the force until it's large enough at a distance that you can feel it
but it's a force which is present all the time and very common and is in a basic force or almost I mean I can go a a little further back if I went more technical but in the early level, I just have to tell you that is going to be one of the thing you will have to take as an element in the world the existence of magnetic repulsion or electrical or magnetic attraction I can't explain that attraction in terms of anything else that is familiar to you.
For example, if we say that the magnets attract like if they were connected by rubber bands I would be cheating you because they are not connected by rubber bands; I shouldn't be in trouble you'll soon ask me about the nature of the bands. And secondly, if you are curious enough you will ask me "why rubber bands tend to pull back together again?"
I would end up explaining that in terms of electrical forces which are the very things I try to use the rubber band to explain so I have cheated very badly you see So I'm not going to be able to give you an answer to "why magnets attract each other" except to tell you that they do and to tell you that's one of the elements in the world among different forces:
there are electrical forces, magnetic forces, gravitational forces and others, and those are some of the parts If you are a student, I could go further and tell you that the magnetic forces are related to the electrical forces very intimately that the relationship between the gravity forces and the electrical forces remains unknown and so on. But I really can't do a good job, any job of explaining magnetic forces in terms of something else that you are more familiar with, because I don't understand it in terms of anything else that you are more familiar with.

Electricity

The stuff of fantasizing in looking at the world, imagining things, which really isn't fantasizing because you just try to imagine the way it really is, comes up handy sometimes. The other day I was at the dentist, he was getting ready with this electric drill to make holes and I thought I'd better think of something fast or else it's gonna hurt. And then I thought about this little motor going around and what was that make it turn? And what was going on? and what's going on is that there's a dam some distance away here and water going over the dam turns a great big wheel, alright and this wheel is connected with long thin pieces of copper
which split up into other pieces of copper and split up and spread all over the city
and then they're connected back to another little gadget and makes wheels turn

all the wheels in the city are turning, because this thing turns. If this thing stops, all the wheels stop. If it starts again, they all start again. And I think it's kind of a marvelous thing of nature. It's extremely curious that phenomenon. I like to think about a lot, because all it is, it's copper and iron. see, sometimes we think it's man-made generator very complicated, the phenomenon is a result of somet special something that we've made. But it's nature doing it, and it's just iron and copper; if you just take a big long loop of copper, and add iron at each end and move the piece of iron here, the other iron move at the other piece. And if you get it down to the NOTHING here just moving piece of iron in a loop a copper and see other piece of iron move
you realize what a fantastic mystery nature is!
you don't even need the iron you could if you at least get this pump prime primed and started by jiggling copper strands around fast enough knotting them and unknotting them and so forth you can get other copper strands move at the other end, over a long connection.
And what is it? It's only copper! And motion!
We're so used to circumstances in which these electrical phenomena are all canceled out
Everything is sort of neutral: pushing and pulling, it's all very dull. But nature has these wonderful things
Magnetic forces and electrical forces when you comb your hair with your comb and you get some strange condition
so you put it in front of a piece a paper, that lifts up the paper, the paper jiggles at a distance far away
that in fact turns out that that is the thing that's deeper inside of everything than the things we're used to
We're used to forces that only act directly, right? you push with your finger, it only acts directly, but then you have to imagine what it is that's pushing with the finger here's this finger is made of little balls of atoms.
And it has got another bunch of atoms that are pushing it. At that little space between those atoms. And that pushing is going through that space.
And the only thing that happens with the comb and the paper is that the circumstances have a reason which makes it possible to see those forces go through a bigger distance than just the short distance between the atoms
What it is they have charges like electrons, that are both the same
they repel each other with a force. They are very tiny parts, they are piece of the atoms and they repel each other with a force which is enormous
it's inversely as the square of the distance just like gravity is inverse to the square of the distance but gravity is attractive whereas this one is repulsive
and for two electrons the gravity is so weak compared to the electricity
electricity is so much more enormous than the gravity I can't express because I don’t know the name of the numbers
it's one with thirty eight or forty zeros after the one Bigger is electricity!
It's so enormous, that if I were all electrons... well, the number is too big!
There's also however for electrical thing other kind of charges, positive charges example of protons are positive, they're inside the nucleus of the atoms and the attract electrons
Opposite charges attracts, alike charges repel So you have to imagine enormous forces where likes are trying to get away from likes, and unlikes are trying to get near the opposite

What would happen if you had a lot of them? they'd be all the likes would collect with unlike, they attract each other
and they'd get an intimate mixture of pluses and minuses all on top of each other, very close together
You wouldn't have a lot of pluses anywhere, because they repel each other They're all being compensated with minus very close,
and you get these little knots of plus and minus the reason that the knots don't get smaller and smaller is because they are particles and there are quantum mechanical effects
that we won't discuss that don't make they can't get any smaller than a certain size So you get these little lumps which are balls, they are the atoms
The atoms are positive and negative charges and they neutralize they cancel their charges as nearly as they can
and because of these force are so big, it ends up nowhere, with very little left
because they're so big they cancels out, there's always so exactly the same pluses and minuses in any normal material
When you comb your hair, it rubs just a little bit extra off just a few extra minuses say here, and somewhere else a few extra pluses
But the forces are so big that just the extra ones which make a force that we can see, that seems to get over a long range
and that we find mysterious and that we need an explanation for
and we try to find an explanation for it in terms of ideas like forces that are inside of rubber bands, or steel bars and twisted things
We would like to have some kind of puller, at a distance because we're used to it that we don't get any push until we're touching
but the fact is that the reason we don't get any push until we're touching is the same force as you see at a long distance only it's come down to short
because the pluses and minuses have cancelled out so well that you don't feel anything until it gets very very close
When it gets close enough of course it makes a difference which is plus and which is minus and where they are and they repel each other
so it's kind of fun to imagine that this intimate mixture of highly attractive opposites
which are so strong that they cancel out the effects and it's only sometimes, when you have an excess of one kind or another that you get this mysterious electrical force
And how can I explain the mysterious electrical forces in any other way? Why should I try to explain it in terms of something like jelly or other things which are made?
And I understand the other way around in terms of strong, long distance forces which are all canceled out
So it's the electrical forces in fact, and the magnetic forces in fact that we have to accept as the base reality
in which we are going to explain all the other things. So again it turns out it's hard to understand, you have to do a lot of imagining
that the real world has as its base, a force that acts at long distance
that we haven't got much experience with that force we have peculiar phenomena here and there, but ordinarily

we don't have much experience with that force is simply because that's what requires explanation
That's what requires imagination. The long distance force we haven't other picture for
And in the example of the generator for instance
what happens is that the electrons which are part of an atom they're pushed by the motion of the copper wires
wonderful to think that if you push a few here, and they get too close together so they push the others because they repel at a long distance
so it's not just like water which repel at a short distance but it's a wonderful fluid which repel at a long distance
and the effects therefore can go very quickly through the wire there is a little concentration you go ZINNNG through the wire all over the city at once.
And you can use that stuff to make signals, you can push a few electrons here and there by talking in a telephone,
at the other end of the line, a long line of copper across the city the electrons respond
because of this very rapid interactions over these long distances to what you're saying in this room
and they discovered experimentally the existence of these long forces and that this rapid motion action and so forth
was a tremendous thing for human beings I think that the discovery of electricity and magnetism and the electromagnetic effects
which are finally worked out, the full equations were worked out by Maxwell in 1873
probably the most fundamental transformation, the most remarkable thing in history
the biggest change in history
Mirror and Train puzzles
I went to a scientific school-MIT, and then fraternity when you first join
they try to keep you from being feeling that you're too smart
by giving you what looked like simple questions to try to figure out what actually happens and it's like training for imagination you know, it's kind of fun
and I'd tell you some of them that I remember I learned them of course once you learn them
the next time somebody comes along with this wonderful puzzle you look at them kind of quietly you wait two or three seconds or five seconds
to show ways that you were thinking and then you come up with this answer to astonish your friends
but the fact was of course that you were trained by your fraternity brothers as to how to answer these things early on
one of the questions we used to we got was the problem about the mirror it's an old-fashioned it's an old problem
You look in a mirror, and let's say you part your hair in the right side
and you look in the mirror and the image has got its hair part on the left side so the image is left to right mixed up it's not top and bottom mixed up
because the top of the head of the image is on the top and the bottom of the feet are at the bottom and the question is how does the mirror know to get the left and right mixed up and not the up and down

you get a better idea of the problem if you think of lying down and looking at the mirror
all right your hair is still on the left side and now the left and right was the up and down Whereas the up and down which look okay was the left and right before
and the mirror somehow figured out what you are gonna do when you're looking at it so what to describe in a sort of symmetrical way what the mirror does
that it doesn't look lopsided and it takes left and mixes it up with right and doesn't do the same with up and down
and after a lot of fiddling you gradually read we can worked out the answer to that one
you see if you wave this hand then the hand in the mirror that waves is the right opposite at it
The hand on the east is the hand on the east and the hand on the west is the hand on the west, and the hand that head that up is up and the feet that is down is down.
Everything is really all right! but what's wrong is if this is north
your nose is to the north of the back of your head, but in the image the nose is to the south of the back of the head
so what happens really in the image is neither the left and the right mixed nor the top and the bottom, but the front and back had been reversed, you see
that is just the nose of the thing is on the wrong side of the head if you want it, all alright? Now ordinarily when we think of the image we think at it as of another person
and we think the normal way that another person would get on that condition over there It's a psychological thing We don't think of the idea that the person has been squashed and pushed backward with his nose and his head
because that's not what ordinarily happens to people A person gets to look like you looks in the mirror by walking around and facing you
and because people when they walk around don't turn their head for their feet we leave that part alone
but they get their right and left hand swung about you see when they turn around so we say that it's left and right interchanged
but really the symmetrical way is along the axis of the mirror that thing get interchangeable but that's kind of an easy one
a harder one and very entertaining was "what keeps a train on the track?"
And of course the answer is, as everyone thinks: the flanges on the wheels you know the wheels have some kind of flange on them
but that's not the answer. Because flange is just safety devices if the flanges rub against the tracks you hear a terrible squealing
they're just in case the real mechanism doesn't work there's another problem with trains that's connected to it
now people all know this about their automobile that when you go around the corner
the outside wheels have to go further than the inside wheels and if the wheels were connected on a solid shaft
you couldn't do that you can't turn the outside wheels further than the inside wheels and so the shaft is broken in the middle with a gear system it's called a differential
did you ever see the differential on a railroad train? no you look at those wheels under a freight car
and there are the two wheels and there's a solid steel rod going from one wheel to the other

there's nothing that one turns the same as the other so now how does it go around the corner, a curve? when the outside wheel has to go further than the inside wheel
and the answer is that the wheels are flanged like this I mean not flange they're cones this way that is they're a little fatter closer to the train and a little thinner further out if you look closely you'll see they've got this beveled edge
and it's all very simple when they go around the curve, they slide out on the track a bit
So that this wheel travels on a fatter pond a bigger diameter and this on a smaller diameter so when they both turn one turn this swings further than the other
and that's what keeps it on the track also the same way suppose the train's running along on this thing, on the track
and the track's here and here the two wheels are exactly balanced and it's nice and even suppose accidentally it gets a bump or something and slides out this way
then this wheel is on a bigger circumference than this one but they're on a solid shaft
so when it turns once around it carries this wheel forward relative to the other and steers the train back on the track
of course if it gets too far off on the other side it goes back and forth and it stays on the track because the wheels are tapered and the flange is safety
well we had a lot of stuff like that that we had to learn you know that would get straightened out before we could become full-fledged members of the fraternity
Seeing Things
If I'm sitting next to a swimming pool and somebody dives in and she's not too pretty, so I can think of something else
I think of the waves and things that have formed in the water and when there's lot's of people that have dived in the pool, there's a very great choppiness of all these waves all over the water, and to think that it's possible, maybe, that in those waves there a clue as to what's happening in the pool,
that some sort of insect or something with sufficient cleverness, could sit in the corner of the pool and just be disturbed by the waves,
and by the nature of the irregularities and bumping of the waves have figured out who jumped in, where and when and what's happening all over the pool.
And that's what we're doing when we're looking at something: the light that comes out is waves just like in the swimming pool, except in 3 dimensions instead of 2 dimensions of the pool and it's going in all directions
and we have a 8'th of an inch black hole into which these things go, which is particularly sensitive to parts of the wave that are coming in a particular direction
and it's not particularly sensitive when they're coming in at the wrong angle, which we say is from the corner of our eye, and if we want to get more information from the corner of our eye, we swivel this ball about so that the hole moves from place to place.
Then, it's quite wonderful that we figure out so easy;
that's really because the light waves are easier and the waves in water are a little bit more complicated; it would have been harder for the bug than for us, but it's the same idea,
to figure out what the thing is that we're looking at at a distance,
and it's really kind-of incredible because when I'm looking at you, someone standing to my left can see somebody who's standing at my right;

that is, the light could be going right across this way, the waves are going this way, the waves are going this way, the waves are going this way, it's just a complete network.
Now, it's easy to just think of them as arrows passing each other, but that's not the way it is, because all of this is something shaking -it's called the electric field,
but we don't have to bother with what it is- it's just like the water height is going up and down. So there's some quantity shaking about here
and the combination of motions that's so elaborate and complicated that the result is to produce an influence which makes me see you.
At the same time, completely undisturbed by the fact that there are influences that represent the other guy seeing him on this side.
So that there's this TREMENDOUS mess of waves all over in space which we call
which is the light bouncing around the room and going from one thing to the other, because of course most of the room doesn't have 8'th inch black holes. It's not interested in that light,
but the light is there anyway, and it bounces off this, and it bounces off that, and all this is going on, and yet we can sort it out with this instrument.
But beside all that, you see, those waves that I was talking about in the water, maybe they're so big - some of them - and then there's slower swashes which are longer, and shorter.
Perhaps that animal is making it's study only using waves between this length and that length, so it turns out that the eye is only using waves between this length and that length,
except those two lengths are 100,000'th of an inch - 100,000'th of an inch big,
and what about the slower swashes? The waves that go more slowly, that have a longer distance from crest to trough.
Those represent heat. We feel those, but our eye doesn't see them focused very well, we don't in fact at all.
The shorter waves are blue, the longer waves are red. But when it gets longer than that then we call them infrared.
And all this is in there at the same time. That's the heat. Pit viper that get down here in the desert, they have a very little thing that they can see longer waves
and pick up mice, which are radiating their heat in the longer waves but their body heat by looking at them with this eye, which is the pit of the pit viper.
But we can't, we are not able to do that. And then these waves get longer and longer, and all through the same space,
all these things are going on at the same time, so that in this space, there's not only my vision of you,
but information from Moscow Radio that's being broadcasted at present moment, and the seeing of somebody from Peru.
All the radio waves are just the same kind of waves, only they are longer waves. And there's the radar, from the airplane which is looking at the ground to figure out where it is, which is coming to the room at the same time.
Plus X-rays, cosmic rays and all of these other things that are the same kind of waves, EXACTLY the same kind of waves, but shorter, faster or longer, slower.
It is exactly the same thing. So this big field, this - this area of irregular motions of this electric field, this vibration, contains this tremendous information,

and it's ALL REALLY there, that's what gets you. If you don't believe it, then you pick a piece of wire and connect it to a box
and in the wire the electrons would be pushed back and forth by this electric field, swashing just at the right speed for the certain kind of long waves,
and you turn some knobs on the box to get the swashing just right, and you hear Radio Moscow!
Then you know that it was there. How else did it get there? It was there all the time. It is only when you turn on the radio that you notice it.
But that all these things are going through the room at the same time which everybody knows, but you gotta stop and think about it to really get the pleasure about the complexity - the INCONCEIVABLE nature of nature.
Big Numbers
When we were talking about the atoms, one of the trouble we have with the atoms is that they are so tiny and it is so hard to imagine the scale.
The atoms are in size compared to an apple is the same scale as an apple is compared to the size of the earth.
That's kind of a hard think to take, and you have to go through all these things all the time and people find these numbers inconceivable and I do too
And the only thing you do is just change your scale you just think of small balls but you don't try to know exactly how small they are too often
or you get kind of a bit nutty, alright? but in astronomy you have the same thing in reverse because the distances to these stars are so enormous You know that light goes so fast, and it only takes few seconds to go to the moon and back
or it goes around the earth in seven and a half times in a second and it goes for years... two years, three years before it gets to the nearest other star that there is to us!
But all our stars are in nearby galaxies, a big mess of stars which is called a galaxy, a group but our galaxy is (what is it) some hundred thousand light years, like 100 000 years
And then there's another patch of stars It takes a million years for the light to get here going at this enormous rate
and you just go crazy trying to make too real that distance. You have to do everything in proportion. That's easy
say that galaxies are little patches of stars and they're ten times as far apart as they are big So that's an easy picture. But you just go to a different scale, that's easier
once in a while you try to come back to earth scale to discuss the galaxies, but it's kind of hard
The number of stars we see at night is only about five thousand. but the number of stars in our galaxy, the telescope have shown when you improve the instruments
Oh! We look at a galaxy, we look at the stars, all the light that we see, the little tiny influence, spread from the stars over this enormous distance of what three light-years for the nearest stars
on! on! on! this light from the star is spreading, the wavefront's getting wider and wider
weaker and weaker, weaker and weaker out into all of space and finally the tiny fraction that comes in one square eighth of an inch little black hole
and does something to me so I know it's there! Well to know a little bit more about, I'd rather gather a little more of this tiny fraction of the front of light

and so I make a big telescope which is a kind of funnel. The light that comes over this big area-200 inches in cross-is very carefully organized
so it is all concentrated back so it can go through pupil. Actually it's better to photograph, and nowadays they use photocells which are better instruments
but anyway the idea of a telescope is to focus the light from a bigger area into a smaller area so that we see things that are weaker, less light
and in that way we find there's a very large number of stars in the galaxy.
There's so many that if you try to name them, one in a second, all of the stars in our galaxy
I don't mean all the stars in the universe just this galaxy here it takes three thousand years! And yet that's not a very big number
Because if those stars were to drop one dollar bill on the earth during a year each star dropping one dollar bill
they might take care of the deficit, which is suggested for the budget of the United States!
So you see what kind of numbers we have to deal with! Anyway I think that the numbers are problems in astronomy, the size and numbers
the best thing to do is to relax and enjoy the tininess of us and the enormity of the rest of the universe
Of course, if you're feeling depressed by that, you can always look at it the other way and think how big you are compared to the atoms and the parts of atoms then you're an enormous universe to those atoms
so you can sort of stand in the middle and enjoy everything both ways
But the great part of astronomy is the imagination that is necessary to guess what kind of structures
what kind of things can be happening to produce the light and the effect of the light of the stars that we do see
and I could take an example, a historical example
many times in science, by using imagination you imagine something which could be according to all the known knowledge and the laws
and you don't know whether it is yet or not And that's very interesting, there is a creative imagination you'd like to call it
not just imagining thing that are relatively easy, but something different And to take an example of, a star as we understand it
ordinary stars like the Sun, which is just a big ball of gas, of hydrogen that's burning up the hydrogen and so forth and it's an enormous mass of gas
and it's held together by gravity you don't to always understand gravity as a curved space good enough for the purpose that a force inversely square the distance When the things are closer together, the force is stronger, and it pulls everything together
By the way that's why the world is round: because the globe of earth is pulled together as much as possible
and if it had a great mountain and an irregularity like a bump so it would be pulled in by gravity and it all gets smooth
Rocks aren't strong enough to hold a bump much bigger than a few miles and Mount Everest is our biggest bump

But on the moon where the gravity is less, the bumps are higher; the mountains are bigger on the moon. Anyway, to get back to the star, it's all held together by gravity
and it's got a nuclear fuel which we've haven't been talking about that's burning up the hydrogen and generating energy which keeps things going
And after a while, it would use the fuel a lot people began to think about what would happen then and it would be possible to just be gas sort of hanging around held together by gravity but quiet
but another possibility was to think if I push the stuff together closer, the gravity is stronger, would hold it together.
Well if you push a little bit together, the pressure increases when you push the gas together, there are more atoms and they pound on it
so the pressure is higher but the gravity is stronger and it turns out the pressure wins so it would just come out again
if you're pushing a star like that, it oscillates and there are some stars that are oscillating and vibrating and so on
but it turns out if you keep on analyzing you push it together very far to the incredible concentration
that the whole mass of the sun is down to the size of the earth or smaller
then it turns all the nuclear matter all the nuclei of the atoms are all stuck next to each other tight
Spaces where the electrons are all squashed out and it comes out when you get to THAT far, the gravity is strong enough to overpower the pressure again
even though the pressure has got to be enormous, the gravity has to be even more enormous and the thing will stay steady at a different size
and be nothing but a neutron, a nuclear matter, nothing solid, nuclear matter
and this possibility was worked out by Oppenheimer and Volkov, it's called a neutron star
And people waited to see if there were any such neutron stars for years until recently they found these strange pulsars which emit flashes of radio waves later they found light
which can go 30 times a second for instance the fastest ones or maybe 10 times a second or one a second
And at first, that's very mysterious; you are used to stars being big and slow how can anything in a star move in a thirtieth of a second?
Well these things are very small neutron stars and they're spinning very fast.
For reason not yet understood, they are emitting a beam, a beam of radio waves like a search light in an airport or something those things that go around boop boop boop
so we get the flashes tick tick tick, that fast to imagine a star the mass of the sun, doing something, turning so fast 30 times a second
another one of these big numbers, hard to conceive imaginary things okay?
and the whole idea that there could be a star of such enormous density that a teaspoon would weigh so much
of the matter if you put it on the earth surface it is so heavy that that it will just plough right to the center of the earth!
and things like that, it took a lot of imagination: it comes out the mathematics and the analysis of all this helps you to make sure you are not making a mistake

and it turns out that such a star is possible, and it turned out a little bit later they do exist and that's a good example of how imagination is a useful thing
and produces a guessing ahead of time and how we make advances by using it
Besides, the very difficult thing of imagining all the things that might be up there to explain the things we see
in the case of astronomy, we have a large number of things we see that we have not yet quite clearly got the imagination to see what it is that's producing them
Quasars are very powerful sources of light and radio waves from very great distances
and we see them because they are so bright. The exact cause of their sources is gradually been recently understood
in terms of another nutty concept of imagination: the black hole
which is something that comes from following the logic of gravity of Einstein to its ultimate working out the consequences in crazy circumstances Suppose you had an amount of matter so great
that the gravity force is so much that even light trying to get out falls back
Nothing can go faster that light, and nothing could escape. You couldn't see it!
how would you get there? If you have a large amount of matter to start with, it could fall together and get into this condition that no longer could the light come out. So you would have this thing which continues to attract things to it
things would go in and nothing would come out. That is called the black hole. and you say well how can a black hole which is absorbing everything make all this energy that we see
Is that an explanation of a quasar? Actually, it may well be Because the things that are falling in don't go pluck in but go around, falling in by swirling
then as they are falling irregularly and so forth, and in the fast motions that it produces they go down this whirlpool
they generate a lot of energy and friction and so forth, and different kind of effects magnetic and electric effects that could make the jets of matter that come out of the quasar and the radio galaxies
in ways that are not really understood. We don't have a real picture why there are jets of radio waves, matter emitting radio waves in galaxies
there are galaxies which great jets coming out with big clouds of matter on each side which are emitting radio waves
so there's some kind of a source in there it sort of gets wound up and shoots these jets of matter out with tremendous energy
And it's guessed that maybe it's a black hole somehow or other
and the somehow or other is the challenge of the imagination Which has not yet been answered. by anybody, with any great confidence.
Ways of Thinking
You ask me if an ordinary person, by studying hard, would get to be able to imagine these things, like I imagine
Of course! I was an ordinary person who had studied hard. There are no miracle people.
It just happen they got interested in these things and they learned all these stuffs.
There are just people There's no talent, special, miracle ability to understand quantum mechanics or a miracle ability to imagine electromagnetic fields

that comes without practicing and reading and learning and study so if you say it take an ordinary person who's willing to devote a great deal of time
and study and work and thinking and mathematics and time, then he has become a scientist.
When I'm actually doing my own things and working in a high, deep and esoteric stuff that I worry about
I don't think I can describe very well what it's like
First of all, it's like asking a centipede which leg comes after which it happens quickly and I'm not exactly sure what flashes and stuff go in the head
But I know it's a crazy mixture of partial equations, partial solving in equations,
then having some sort of picture of what is happening that the equation is saying it's happening but they're not that well separated as the words I'm using and it's a kind of a nut nutty thing
it's very hard to describe, and I don't know that it does any good to describe And there's something that struck me, it's very curious:
I suspect that what goes on in every man's head might be very very different, the actual imagery or semi-imagery which comes
and when we are talking to each other at these high and complicated levels and we think we are speaking very well, that we are communicating
but what we are really doing is having a some kind of big translation scheme going on for translating what this fellow says into our images, which are very different.
I found that out because in the very lowest level I wouldn't go into much details but I got interested in...
Well I was doing some experiments and I was trying to figure out something about our time sense
and so what I would do is trying to count to a minute actually say I'd count to 48 then it would be one minute
so I calibrate myself and I would count a minute in 48 think I was counting seconds but it's close enough
and then it turns out if you repeat that you can do very accurately when you get to 48 or 47 or 49, not far off, you're very close to a minute.
And I was trying to find out what affected that time sense and whether I could do anything at the same time I was counting
and I found that I could do many things I could, there were some things that not
For example, I had great difficulty… I was in the university, I had to get my laundry ready, and I was putting the socks out
and I had to make a list "how many socks", there were something like 6 or 8 socks and I couldn't count them
because the counting machine was being used, and I couldn't count them until I found that I could put them in a pattern and recognize the number
and so I learned a way after practicing by which I could count the line of type in a newspaper and see them in groups
three, three, three, three, one that's a group of ten, three, three, three, one without saying the numbers just seeing the groupings
I could therefore count the line of types I was practicing in the newspaper the same time I was counting internally the seconds

so I could do this fantastic trick of saying "forty-eight, that's one minute and there are sixty seven lines of type" you see!
It was quite wonderful and I discovered many things I could read while I was..
no, excuse me, yes, I could read perfectly alright while I was counting and get an idea of what it was about
But I couldn't speak, I couldn’t say anything. Because of course I was sort of trying to speak to myself, inside, I would say "one, two, three" or sort of in the head.
Then I went down to the breakfast, and there was John Tukey was a mathematician at Princeton in the same time
and we had many discussions, and I was telling him about these experiments and what I could do and he says "that's absurd!" he says
He said "I don't see why you have any difficulty talking whatsoever, And I can't possibly believe that you could read"
so I couldn't believe all this but we calibrated him it was 52 for him to get to 60 seconds or whatever I don't remember the numbers now
and then he said "alright, what do you want me to say? Mary had a little lamb. I can speak about anything, blah blah blah, blah blah,
52! that's one minute". He was right. And I couldn't possibly do that And he wanted me to read, because he couldn’t possibly believe it.
and then we compared note, and it turned out that when he thought of counting what he did inside his head when he counted was he saw a tape with numbers it went "clink, clink clink"
the tape would change with numbers printed on it, he could see well since it's sort of an optical system that he was using, and not voice.
He could speak as much as he wanted but if he had to read, then he couldn't look at his clock!
Whereas for me it was in the other way. And that's where I discovered, at least in this very simple operation of counting
the great difference in what goes on in the head when people think they are doing the same thing
And so it struck me therefore, if that is already true at the most elementary level
That when we learn mathematics and Bessel functions, and the exponential and the electric field and all these things
that the imageries and the method by which we are storing it all and the way we think about it could be really, if we get to each other's head, entirely different And in fact, while somebody sometimes has a great deal of difficulty to understanding a point which you see as obvious, and vice versa
it's maybe because it's a little hard to translate what you just said into his particular framework and so on
Now I'm talking like a psychologist, and you know I know nothing about this!
Suppose that little things behaved very differently that anything that was big, anything that you're familiar with
because you see as the animal evolves and so on, as brain evolves it gets used to handling the brain is designed for ordinary circumstances
but if the gut particles in the deep inner workings where by some other rules and some other characters

they behave differently they were very different than anything on a large scale then there would some kind of difficulty in understanding and imagining reality.
And that difficulty, we are in. The behavior of things on the small scale is so fantastic! It is so wonderfully different!
so marvelously different that ANYTHING that behaves on a large scale. You said "electrons act like wave", no they don't exactly,
"they act like particles", no, they don't exactly, "they act like a kind of a fog around the nucleus", no they don't exactly.
And if you want to get a clear, sharp picture of an atom, so that you can tell exactly how it's going to behave correctly
and have a good image in other words, really good image of reality I don't know how to do it. Because that image has to be mathematic:
we have a mathematical expression, a strange mathematics, I don't understand how it is but we can write mathematical expressions and calculate what the thing is going to do
without actually being able to picture it It would something like a computer in which you put certain numbers in
and you have a formula for what time the car will arrive at different destination and the thing does the arithmetic to figure out what time the car arrives at the different destinations
but cannot picture the car. It is just doing the arithmetic So we know how to do the arithmetic, but we cannot picture the car
It's not a 100% because for certain approximate situations, certain kind of approximate pictures work
that it's simply a fog around the nucleus that when you squeeze it, it repels you it's very good for understanding the stiffness of certain material
that it's a wave which does this and that is very good for some other phenomena alright So when you're working with certain particular aspects of the behavior of atoms
for instance when I was talking about temperature and so forth, that it's just little balls it's good enough and it gives a very nice picture of temperature
but if you ask more specific question and you get down to questions like "how is that when you cool helium down, even to absolute zero where it's not supposed to be any motion
it's a perfect fluid and it has no resistance and it flows perfectly, and it isn't freezing"
Well if you want to get a picture of atoms as all of that in it, I can't do it But I can explain why the helium behaves as it does, by taking my equations
and seeing that the consequences of them is that the helium would behave as it is observed to behave
So we know that we have the theory right, but we haven't got the pictures that would go with the theory
And it's that because we haven't caught on the right picture
or it's because there aren't any right pictures for people who have to make pictures out of things that are familiar to them
Well let's suppose it's the last one, that there's no right picture in terms of things that are familiar to them
Is it possible then to develop a familiarity with those things that are not familiar on hand, by studying

by learning the properties of atoms and quantum mechanics, by practicing with the equations until it becomes a kind of second nature
just like it's a second nature to know that two balls came towards each other, they smash into bits
You don't say "the two balls when they come toward each other turn blue". You know what they do
So the question is whether you can get to know what things do without... better that we do today as the generations develop, will they invent ways of teaching so that the new people will learn tricky ways in looking at things
and be so trained, so well trained, that they won't have our troubles, with the atom picturing.
There's still a school of thought that cannot believe that the atomic behaviors is so different than large scale behaviors
I think that's a deep prejudice, it's a prejudice of being so used to large scale behaviors and they're always seeking to find to waiting for the day that we discover
underneath the quantum mechanics there's some mundane, ordinary balls hitting or particles moving and so on
I think they're gonna be defeated I think Nature's imagination is so much greater than man's, she's never gonna let us relax!

THE WORLD FROM ANOTHER POINT OF VIEW
I don't know it's hard to make up a very crazy one day witches or something like that you tell about what people used to
believe in witches and of course nobody believes in witches now and you say how could they believe in witches then you're turning around you say let's see
what which is do we believe in though what ceremonies do we do every morning we brush our teeth
what is the evidence that the brushing of teeth does us any good in cavities you start wondering are we all imagine
if it the as the Earth turns on the orbit there's an edge between light and dark
and along that edge all the people along that edge who are doing the same ritual
for no good reason just like in the Middle Ages they had other rituals and you've got a picture this Perpetual line
of toothbrushes going around the Earth it's to take the world from another point of view now it may be May well be
that brushing teeth is a very good thing because it gets rid of cavities and you can ask you can find out whether it does
or it doesn't by trying to find out now you can ask your dentist he says of course and you say how an Evidence I have not found the evidence
from dentists because they just learned it in school now I'm not trying to argue that it's good or bad to brush teeth
when I'm trying to argue for is to think about thing from a new point of view
thank you
see I have had in my life a number of pleasant experiences when the earliest one when I was a kid I invented a

problem for myself to some of the powers of the integers and in trying to get the funnest part I developed a certain set of numbers that I for formula for which I couldn't get and I discovered later those were known
as the Bernoulli numbers and discovered in 1739 so I was up to 1739 when I was
about 14 you see and then a little later I discovered something I'd find out I just they invented a thing called which
we now call uh operated calculus and that was invented
in 1890 something you see I was gradually I was inventing things that came later and later but the moment when
I began to realize that I was now working on something new was what I read
about Quantum electrodynamics at the time and I read a book and I learned about it for example I read the Iraq's
book and he had these problems that nobody knew how to solve it would described there I couldn't understand
the book very well because I really wasn't up to it but there in the last paragraph at the end of the book it said
some new ideas are here needed and so there I was some new ideas were
needed okay so I started to think of new ideas
Richard Feynman Nobel Prize winner and his son Carl stepped gingerly down the
wet cobbles of Milbank high in the Yorkshire pen eyes Feynman professor of physics at the California Institute of Technology retreats to this remote Village near his wife's home for a special purpose it's
here he finds the time and Solitude to sift the ideas that have made him the most feared and original mind in modern
physics
Feynman is in the Forefront of one of the oldest and most intriguing games of Hide and Seek in science finding the
ultimate constituents of the world in this search Feynman is a celebrated Maverick who is encouraged by his father
a New York Clothing salesman to confront conventional wisdom
once said that all the kids were all walking in little parties with their fathers in the woods
then the next Monday we were playing in a field and the kids said to me say what's that
bird what's the name of do you know the name of that person I'm the slightest idea he said well it's a brown-throated crush
he says your father doesn't teach you anything but my father had already taught me about the names of birds he
once we walked and he says that's a brown sort of thrush he says know what the name of that bird it's a Brown tour of the trash
in German it's called the freaking plague in Chinese it's called a king long pong
in Japanese and so on and when you know all the names in every language of that bird you know nothing but absolutely nothing about the bird then we would go on and talk about the pecking in the feathers so I had learned
already that names though constitute knowledge is to knowing the name of something

that's caused me a certain trouble since because I refuse to learn the name of anything so when someone comes in and
says you got the explanation for the fitzclillon experiment I says what what what's that he says you know they take
long lived came as and disintegrates into two pies oh oh yes now I know but I never
know the names of things what he forgot to tell me was that the knowing the names of things is useful if you want to talk to somebody else
so you're telling what you're talking about but the basic principle of knowing about something rather than just knowing its name is something that you stuck to
is it yes of course it's we have to learn these are kind of disciplines in the field of science that you have to
learn that to know when you know and when you don't know and what it is you know and what it is you don't know
it's uh you got to be very careful not to confuse yourself
mold your methods of thinking the way you looked at the world well we had a lot of little games
like you would say at the dinner table you'd think of some little problem and if they suppose we were you were a
martian you were the Martians and we came down to this earth and we'd look at it from the outside but now I can't
explain exactly what he meant but there's a way of looking something Anew as if you never saw it before for the
first time and asking questions about it as if you were different for instance if
you would ask later I did some a little amusing research for a paper in college on sleep but it started with a question
of his kind suppose you were a martian who never slept they didn't have sleep you didn't have to sleep and you came
down to this earth and you saw these people had this funny puppy that every day for a certain amount of time had to
lie down and become unconscious and then the natural question would be how does it feel to get unconscious what happens
to you ideas run along and suddenly they stop or do they just run more and more
slowly but what happens to your ideas how does it feel to become unconscious so I tried to answer the question what
happens when you become unconscious but you find that these days you still when you're faced with a particularly
difficult problem when you're absolutely stuck you tend to say let's look at it like a martian would look at it sometimes there are lots of things that
people did for example Maxwell Put the equations together with the Faraday he
formulated the equations mathematically with some model in his hand and then the rack uh got his answer
by just writing and guessing an equation and other people got there is like in
relativity got the idea by looking at principles of symmetry now all these methods at Heisenberg got

his quantum mechanics but I think he only talk about the things that you can measure now all of these ideas we should only
talk about things that we can measure try to Define things in terms of only things you measure let's formulate the equation mathematically or let's guess
the equation or all these things are tried all the time look for symmetries all that stuff is tried all that stuff
when we're going against the problem we do all that that's very useful but we all know that that's what we learned in
the physics classes how to do that but the new problem where we're stuck we're stuck because all those methods
don't work if any of those methods would have worked we would have gone through there so when we get stuck in a certain place
it's a place where history will not repeat herself and that's what makes it even more exciting because whatever
we're going to look at at the other the method and the trick and the way it's going to look is going to be very
different than anything that we've seen before because we've used all the methods from before so uh therefore a thing like the history
of the idea is an accident of how things actually happen and if I want to turn the history around to try to get a new
way of looking at it it doesn't make any difference it's I I don't care the only thing that
the real test in physics is experiment the history is fundamentally irrelevant
the most enduring Legacy from his father was not just learning to question the physical world but an enthusiasm for the
inquiry which at 54 Feynman still shares today it has to do with curiosity it has
to do with people wondering what makes something do something and then to discover that if he's trying to get
answers that they're related to each other the things that make the wind make the waves and the motion of water is
like the motion of theirs like the motion of sand the fact that things have common features turns out more and more
Universal what we're looking for is how everything works and how everything is what makes everything work
and uh what happens first in the history is we discover the things that are on the face
of it and then gradually that we ask more questions and then we dig in a little
deeper into things that we can just make we need to do a little more complicated experiment to find out about it but it's
curiosity as to where we are what we are it's very much more exciting to discover we're on a ball half of the stick and
upside down it's spinning around in space it's a mysterious Force which holds us on it's going around a great big glob of gas that's burning by a fuel
by a fire that's completely different than the fire any fire we can make well now we can make that fire nuclear fire
no but that's much more exciting story to many people

then the tales which other people used to make up who worried about the universe that we were living on the back
of a turtle or something like that they were wonderful stories but the truth is so much more remarkable and so what's
the pleasure in physics is that to me is that as it's revealed the truth is so remarkable so amazing
and I can't I have this disease and many other people who have studied far enough
to begin to understand a little of how things work are fascinated by it and this Fascination drives them on to such an
extent that they've been able to convince governments and so on to keep supporting them in this investigation that the race
is making into its own environment as a theoretical physicist Feynman doesn't
have a laboratory and he finds family relaxation helps him to concentrate in recent years he's been concerned with
the long-assed almost childlike question what are things really made of what
makes up the world we see around us have we at last come to the foundation stone from which we can make anything a tree a
human being or must we go on looking at smaller and smaller pieces and going deeper and deeper into a bottomless pit
Feynman is trying to knit together our scattered knowledge of the smallest pieces of matter to see whether they fit
a pattern the problem although fundamentally important to all branches of science seems far removed from
everyday reality the world is strange the whole universe is very strange but see when you look at the details and you
find out that the rules are very simple of the game the mechanical Rules by which you can figure out exactly what's
going to happen when the situation is simple it's again this chess game isn't if you were in just a corner where only
a few pieces are involved you can work out exactly what should happen and you can always do that when there's
only a few pieces and so you know you understand it and yet in the real game there's so it's
so many pieces you can't figure out what's going to happen so there was a kind of hierarchy of different
complexities it's hard to believe it's incredible in fact most people don't
believe that the behavior of say me Juan yak yak and you naughty and all this
stuff is the result over lots and lots of items all obeying these very simple rules
come out that it evolves into such a creature that the billion years of life
with its experiences has produced the thing with prongs that stick out like this and so on
the real there's such a lot in the world there's so much distance between the
fundamental rules and the final phenomena that it's almost unbelievable that the final variety of phenomena can
come from such a steady operation of such simple rules but you've had to build the most complex

scaffolding to find out the simple rules but it is not complicated it's just a
lot of it and if you start at the beginning which nobody wants to do I mean you come in to me now as in an
interview and you're asking me about the latest discoveries that have made nobody ever asked about a simple ordinary
phenomenon in a street you know like what about those colors or something I don't have a nice interview explain all
about the colors butterfly wings whole big deal don't care about that you want the big final results then it's going to
be complicated because I am at the end of a 400 years a very effective method
of finding things out about the world in the search for the ground rules of the physical world John Dalton worked out a
comprehensive explanation over 150 years ago he assumed that everything we see is
made out of tiny atoms that they are immutable and indestructible that atoms of different chemical elements like lead
or copper have different weights too small to be observed the atoms combine with each other to form complicated
molecules and vast collections of these molecules are recognizable to us as tables trees or whatever
but in the final analysis atoms would be the smallest constituents of matter ultimate and unchangeable
at the turn of the century we evolved our present picture of the atom light electrons surrounding a heavier Central
core or nucleus once the atom was shown to be destructible attention turned to
the nucleus and during the 30s it was found that bombarding one nucleus with another led to a release of energy and
the breaking up of the nuclei this process which takes place in nuclear accelerators is photographed in a liquid
bubble chamber take a liquid liquid hydrogen under some other liquid and expand it so there's
relate to boil low temperature and your decrease the pressure is ready to boil and it has to form bubbles somewhere and it's any little piece of dirt or any
little disturbance that will form a bubble in that condition if a particle comes flying through from some machine
it leaves a track it tears up the atoms along with electrons are locked off the atoms along its track
and we can't see that but when the gas tried to expand when the liquid tries to boil the bubbles form around these
charged particles which are left so it leaves a string of Bubbles and then forward then you can take a picture of
the bubbles so simple as picture would be the other machine had made fast particles the particles go through and
you see a string of bubbles but if the particle on the way through hit the nucleus of another atom then you see a

string of Bubbles and a kind of a y if it made it's recoil plus some other thing instead of a wire makes a more
complicated track three or four coming out and then one of them going along and going into two then you know that some
particle went along and disintegrated now these things are going nearly at the speed of light and so if you can see a
short distance a few centimeters that's corresponds to a tenth of a billionth of
a second that is if a track comes out it goes along here and then bifurcates into two you know you made a product which
integrate into two in less than a 10 billionth of a second so you see it's not very difficult to
find out about these things with the right with clever techniques since the wall with evidence from bubble
chamber photographs like this physicists have explored the nucleus of the atom the results have been spectacular and
confusing the harder the nuclei were bombarded against each other the more they disintegrated into even tinier
particles until literally hundreds were known in the last 10 years some order has been
made out of seeming chaos by arranging the particles into patterns each pattern has eight or ten members
related by nuclear properties like Spin and Mass
to the physicist patterns like this imply the possibility of even smaller particles not yet identified but already
named e to the question of what makes up the physical world then lies in the
understanding of the nature of these nuclear patterns we're getting close because we have another little theories
by which we can understand these patterns one picture which describes what
particles you're going to find rather well is that all these particles are made of
out of something else which we happen to call quarks and our Quark is an object
which comes in three varieties it's either a a type B type or c-type Quark okay and that the particles that we find
are two big classes in one class we can understand as being made out of three quarks
and depending on the different proportions how many A's B's and C's and how they're moving around each other
if we count how many states we would get from putting three objects together
could be made in so many ways at 27 different ways each one being three we find groups
of particles in groups of 27 analogously and so on a little more complicated but more subtle but it's like that and then
when we allow for their motion around each other we find the higher energy states analogous to the way that that we
ought to get and even semi-quantitatively there seems to be a relation between the states the rates at which
one turn into another so it looks like they're made out of just three quarks then there's this other class of

particles which are called mesons the first class were called barriers the words aren't going to do any good but the other class
we have to understand as being made of a quark one quark and one antiquark an
antiquark is a negative product with all the numbers all the charge properties the exact opposite of a clock so we make
a quark an antiquark put those together we understand the Meson State put three quarks together we understand all the
others so we have made a really great progress in
analyzing these patterns so much salt that it looks very much as if to me at
least that we're very close to understanding this part of physics this strongly interacting system but what's the main barrier still to well the quarks have well the main barriers we don't understand it
quantitative we don't know exactly the laws I mean we do things like I'm just talking to you on a little bit more
carefully counting how many states we should get inside but we don't know exactly how they move and exactly what
holds them together and so and so on well so there are a number of paradoxes with this Quark picture
this picture helps to give us the behavior at low energies of the what kinds of particles to expect but then you'd expect that a particle would be made out of only three parts but we've done some experiments with
very high energy uh it's hitting a proton with an electron which can only
be interpreted by supposing that the number of particles inside is really infinite
if there are particles inside it can't be done with just three you can calculate that doesn't come out
right so there's a difficulty furthermore the idea that there just be three particles
is self-contradict it's contradictory to the ideas of Relativity and so on which imply the existence of particles and
anti-particles and when there are three there should be possible for the forces to produce pairs of particle anti-particle in various numbers so it should be not just three but many more so the infinity is not a
paradox by itself the three is more of a paradox why is it so simple why can't we
get away and understand so much for just three when there should be an infinite number probably in there both theoretically and experimentally
another thing uh it's a little technical but very paradoxical is that we've had a rule
back for atoms that no two electrons can occupy the same state it's called the Exclusion Principle and we thought we
understood that that was necessary according to quantum mechanics of Relativity though it has to be
and with the quarks we find the exact opposite rule two particles tend to occupy the same state the exact opposite
seems to be contradictory with principles there are ways of escaping this all the
time only by complicating the picture but the simplest picture just three which explains everything is

self-contradictory furthermore some people suppose that maybe these
quarks could come apart that would mean the prediction of new states which consists of only one Quark say if there
were such a state it would have to have a charge of one-third the normal charges of our objects for example or two-thirds
and uh we don't find experimentally any such
particles not everybody's looking for them but it looks as if they exist at all they
have to be extremely heavy then the problem is very good if they're extremely heavy how compared to a proton
say how is it when you put three of them together you get a light object that's not heavy like the proton there are technical ways
of erasing it but they're always complicated every the situation is as it always is
when we're near the answer it looks much simpler that it has any right to be
and we have to understand that Simplicity and why we think it must be more complicated our minds are
complicated somehow just like the the orbits of the planets which were supposed to be circles which
looked simple and they were experimentally they were in circles so they made circles on circles on circles on circles and more and more complicated
it turns out it was really much simpler it was a faucet versus a square the distance
which made Ellipsis inside but different way of formulating entirely which was beautiful
so now we have our Wheels within Wheels it looks simple in nature is no doubt
simpler than all our thoughts about it now and the question is what way do we have
to think about it so that we understand its Simplicity that's where we stand now
on holiday in the pennines Richard Feynman is paid a neighborly visit by yorkshireman Sir Fred Hoyle the
astronomer cosmologist and science fiction writer at First Sight there seems little in
common between the study of galaxies and nebulae billions of miles in diameter and millions of light years old and
nuclear physics where particles exist for only a million millionth of a second
but the formation of stars and galaxies is determined on a massive scale by the behavior of the very nuclear particles
Feynman studies Hoyle and Feynman share an interest in the foundations of physics and
exchanging ideas in the local pub is always as profitable as it is enjoyable did you think you agree that the quasis
are in real trouble that the very big redshifts I think so all right I've had this uneasy feeling now for about five
years it looked crazy for a while but it's like not enough of evidence all the time
as well which one makes a new problem yeah every piece of evidence is the same problem in the same sense if there were any cause for a redshift as big as that
other than recession we'd be all right that's right but in the present physical laws there doesn't seem to be any place
and at the same time the same kind of laws predict the kind of peculiar phenomenon in the black holes which

really confusing yeah and it could be that either the gravity is wrong or one of the physical laws are wrong too
um what's the law that's involved because I'm not arguing at the moment the physical laws are wrong I mean you
would agree that one has to push it through along these lines yeah the best way to progress I always think maybe is
to try to be as conservative that's what wheeler always said to try to be as conservative about the physical laws as
possible and explaining the phenomena if you're continuously fail then you gradually realize you've got to change
something but we start up they've got to change something there's so many ways of changing and you don't know how the
it's most likely you don't have to change anything most of the time we succeed ultimately and explaining these damn things in terms of the known laws
but it's the cases that fail of the interesting one yeah yeah the chat with the uh under the single lamp in the street yes uh where passerby says what you're looking for he says I'm looking for my key and they search for
it for a few minutes and at the end of the minute these minutes they pass it by I said are you sure you lost it here and
the man said not at all but unless I lost it here I'll never find it because the light's better yeah and yeah we work with Alexa better yeah once
I was thinking by analogy that there was a time in the 1900s when the thought that the properties of substances were
not physics for example they would be numbers we would find a series of numbers the index of refraction that was
physics but the number for the index that glass had an index of 1.543 and so
on that salt had another index that those numbers were the properties of substances would come from chemistry or
something but that there was it was that time it was considered a different branch then when the quantum mechanical
understanding of the atoms was a vowel then we could calculate all these properties and we realize that all these numbers were really part of physics and
so copies of substances became a bridge in physics whereas previously it was a sort of chemical brand and I wondered by
analogy you know it's always worked by analogy what today do we not consider part of physics which may ultimately be
part of physics I see and they didn't realize immediately something
we consider at the present moment most people consider that we study the laws of physics that is how things go given a
certain condition how the things behave after that but how did they get into that condition
is considered another problem in other words we are given the conditions the circumstances and then it evolves
you're studying the laws yeah it's as though we're doing a chess game again and we're working on the rules but we're

not worrying about how the pieces are supposed to be set up on the board in the first place that's not our business that's the business of History how the
world evolved it's astronomical history history cosmology how the the universe
exploded or through steady state or whatever it was it's not our business it's interesting that in many other
Sciences there's a historical question like in geology the question how did the Earth evolve to the present condition
in biology how do the various species evolve and they get to be the way they are but the one field which has is not
admitted any evolutionary question is physics here are the laws we say here
are the laws today yeah how did they get that way in time we don't even think of it we think of well that is that way
from Forever it's always been like that the same laws and we try to explain the universe that way so it might turn out
that they're not the same all the time and that there is a historical evolutionary question but how do you see
it going yeah it's not it's hard to speculate is it is it a continuous change or is it something that depends
on big you're the best you would I think differently I think of the possibilities but I'm afraid to put things in
when I see oh but it's the dark I always think of the dark it's too big
for me to guess at because yeah it's a guess it's not much use in guessing particular things but but uh you're
different than I would like to discuss with you sometimes how do you do that because I'm really a little afraid to
make specifications I don't know why you kind of grow up I
don't know I'm afraid to make specific guesses because the moment I'm making that guess I can see seven other alternatives
and so since I see these other Alternatives I don't know which one to to piddle with well I don't like to
spend a lot of energy my choice is is very simple I I don't set any requirement that the answer be right
it's just what I'm interested that's the difference yeah if I'm interested in it I'm trying
to find out how neat you could be but how nature is see how that's right you see I don't think you've ever find
it I said your idea is to find out what nature could be yeah but what I what I think is
yeah even if it's wrong common ground is enthusiastically explored but is it only
shared experience and knowledge that forms a bond between working scientists and separates them from us the
interested Layman or even the artist and his scientific fields of becoming so
specialized and they're so varied are you really saying that you have more in common with say a paleontologist or someone in a branch of science very far removed from yours and you would with a playwright or

a poet absolutely especially if he's a good paleontologist because he's a good paleontologist he's
not just looking at old rocks he's looking at the history of the earth he's looking when he stands and he looks at
his own fingers and he knows it's got five bumps and he thinks of how did it evolve with five months he got the same
as Wales and so on and we keep talking about the importance of the fact that
the thumb of poses then we can start discussing is it really so important to the thumb Expo or is that language that
has been involved the system is symbol then or the size of the brain this is a paleontologist I can talk about this
stuff that's close to his field dolphins have bigger brains than we are they have a signaling system and you get
interested in that and you start to discuss all that they know about Davos and you complain that the way the United
States Navy has been doing its experiments is not right and we ought to find out more about that and they go on and on you talk those are things of the
day they're just as good but you can go on and on with I talk to a playwright or
something I I find because I don't look go to plays or something I don't find that easy
to talk to them that I don't get much out of it I was going to say this is because you
can talk to scientists in other fields presumably because you read the scientific magazines and hear the
scientific gossip rather than you know because we don't have to have magazines or gossip we think originally we think
of a new idea we talk to each other and we try to look at something from a new point of view
and we Delight each other in a new point of view and when you're talking to somebody else who's trying to think of something new
different and he thinks he's thought about the whales or the Dolphins and he had some
little thing he's thought of that's a little different than the thing that you've thought of and so when you're talking
back and forth he's excited by yours point of view about dreams
and you're excited by his little observation that he has made about dreams if he has happened to have
thought about that so the point is and our backgrounds Give us a slightly
different point of view I mean a scientific way like I specialize in physics to say he specializes in paleontology so his his information on
dreams might be more deeper more evolutionary for example he might Molly can't we don't have way of
telling I suppose about the evolution of dreams but he might know for example about other animals he might have
thought about whether other animals dream and what the signs are and other things that I hadn't thought of I can't
make it up now because I'm not the paleontologist but I believe that yes I find always that a good man uh I take it all back

I take it all back good man I talked to Good Men in other fields
there's certain kinds of men in every field that I can talk to as well as I can talk
to a good scientist I met a historian or writer of history from France once
and I had a marvelous conversation with it his name was Andrea
and then I met an artist Robert Irwin who's a very important
artist in Los Angeles in modern art and I could talk to him at the same depth of
excitement so I take it all back if you give me the right man in any field I can talk to him I know
what the condition is that he did whatever he did as far as he can go
that he studied every aspect of it as far as he has stretched himself to the end he's not a dilettante in any way and
so he talked deep as far as he can go and he therefore he's up against Mysteries all the way around the edge
and or and we can talk about mystery and all that's what we haven't come
we're talking a bit about these Fallout periods when things are getting very
painful after discussing working problems it isn't at all that Feynman and Hoyle should Savor that most
thrilling pleasure of all the moment of Revelation you try all sorts of things and you're hopeful about trying it have
you had a moment when in a complicated problem where quite suddenly the thing comes into your head and you're almost
sure you've got to be right oh yes I mean this is
and then you try to figure out what the conditions were of that moment that you can do it again for example I worked out
the theory of helium woods and suddenly saw everything they were struggling struggling for two years and suddenly
saw everything remember everything about it by the way it's psychologically funny you can
remember the color of the paper you were writing on has that through yeah and then you wonder what's the
psychological condition well I know at that particular time I simply looked up and I said wait a minute it can't be quite that difficult it must
be very easy I'll stand back and I'll just treat it very lightly I'll just tap it and it'll say and there it was
so how many times since then I'm walking on the beach and I said well look it can't be so complicated
there's nothing happens you know the lights are great but uh but the secret way
is that missing bit in the brain isn't it right something suddenly lights up and yeah and I have no idea I've thought
about it because some uh some may have suggested I think about that because if I can only figure out the formula for
how what condition to be in to get good ideas yeah and be much more efficient and more happy you know so I wasn't paid
attention to what the condition is and have never found any in our relations with anything by the way it's the Delight it's
absolutely you just go absolutely wild that drives how does it last for today

it's not very short it's a very big moment yeah yeah and then there were lesser
Pleasures yeah as you work on both things and more people notice it than the highest about it on the high fig for
about three days it's like a supernova I suppose yes that's better yeah but uh I was going to say that it's the the hope of that kind of gold
it keeps you going that can keep you going through these doldrums yeah yeah and that I think what I learned when I
was a child for my father was that if you did work a little bit at
these things there would be a time which should get there yeah yeah and I had to learn that first yeah I'd never been
able to do it yeah and then afterwards you wonder why the
devil was not so stupid that I didn't see this that's not only two of you it's true of history of the history of the
science you can always look at it to the moments history and wonder why they hadn't thought of it 20 years earlier or
10 years there depending on the case it's because we're dubbed somehow it's more mysterious that it just means that
however good you may get comparatively compared to uh
and we're still very bad at it absolutely yeah yeah we're doing the best we can yeah it's very good yeah
this is depressing and sobering thought