

The Special Theory of Relativity

15-1 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

$$m = \frac{m_0}{\sqrt{1 - v^2/c^2}}, \quad (15.1)$$

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 space ship is drifting along at a uniform speed, all experiments performed in the space ship and all the phenomena in the space ship 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

15-1 The principle of relativity

15-2 The Lorentz transformation

15-3 The Michelson-Morley experiment

15-4 Transformation of time

15-5 The Lorentz contraction

15-6 Simultaneity

15-7 Four-vectors

15-8 Relativistic dynamics

15-9 Equivalence of mass and energy

will appear the same as they would if the system were standing 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 Fig. 15-1. 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,

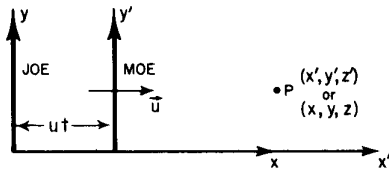


Fig. 15-1. Two coordinate systems in uniform relative motion along their x -axes.

$$\begin{aligned}x' &= x - ut, \\y' &= y, \\z' &= z, \\t' &= t.\end{aligned}\tag{15.2}$$

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 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 15.2, *their form does not remain the same*; therefore, in a moving space ship 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 (15.2) 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 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?

15-2 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:

$$\begin{aligned}x' &= \frac{x - ut}{\sqrt{1 - u^2/c^2}}, \\y' &= y, \\z' &= z, \\t' &= \frac{t - ux/c^2}{\sqrt{1 - u^2/c^2}},\end{aligned}\tag{15.3}$$

namely, Maxwell's equations remain in the same form when this transformation is applied to them! Equations (15.3) 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. (15.1). 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.

15-3 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 Fig. 15-2. 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 con-

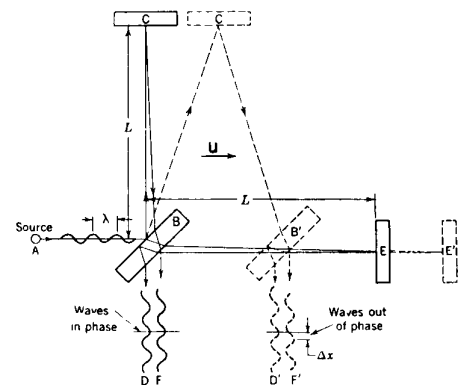


Fig. 15-2. Schematic diagram of the Michelson-Morley experiment.

tinue 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.

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

$$t_1 + t_2 = \frac{2L/c}{1 - u^2/c^2}. \quad (15.4)$$

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

$$2t_3 = \frac{2L}{\sqrt{c^2 - u^2}} = \frac{2L/c}{\sqrt{1 - u^2/c^2}}. \quad (15.5)$$

We are now able to compare the times taken by the two beams of light. In expressions (15.4) and (15.5) 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

$$L_{||} = L_0 \sqrt{1 - u^2/c^2}. \quad (15.6)$$

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. (15.5) is not changed, but the L of Eq. (15.4) must be changed in accordance with Eq. (15.6). When this is done we obtain

$$t_1 + t_2 = \frac{(2L/c) \sqrt{1 - u^2/c^2}}{1 - u^2/c^2} = \frac{2L/c}{\sqrt{1 - u^2/c^2}}. \quad (15.7)$$

Comparing this result with Eq. (15.5), 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.

15-4 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 (15.3). 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 space ship as when calculated by a stationary observer who is watching the space ship. 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. 15.5). In other words, when the outsider sees the man in the space ship 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 space ship 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

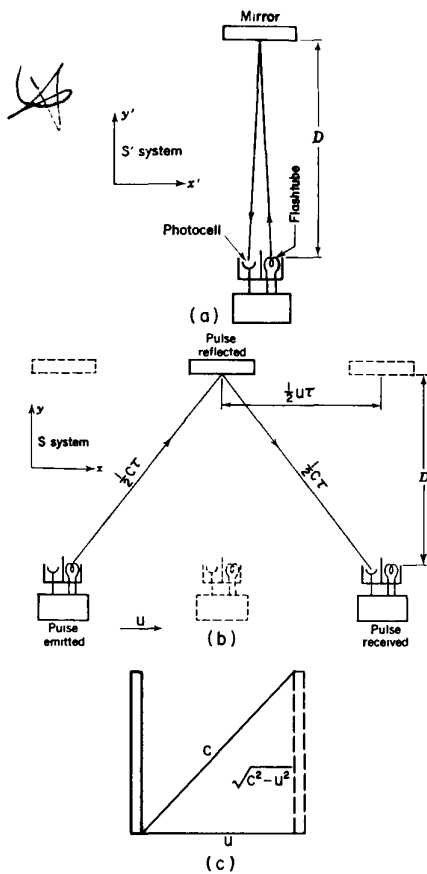


Fig. 15-3. (a) A "light clock" at rest in the S' system. (b) The same clock, moving through the S system. (c) Illustration of the diagonal path taken by the light beam in a moving "light clock."

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 space ship, 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 space ship 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 sidewise 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 Fig. 15-3, the distance the 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 space ship, 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 space ship. 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 space ship, 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 $2 \mu\text{sec}$, from our point of view they live considerably longer—enough longer that they may reach the earth. The factor by 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.

15-5 The Lorentz contraction

Now let us return to the Lorentz transformation (15.3) 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.

15-6 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 separated 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 space ship (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 space ship 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!

15-7 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 11 for a rotation of coordinates. We then had

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta, \\ y' &= y \cos \theta - x \sin \theta, \end{aligned} \tag{15.8}$$

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

$$x'^2 + y'^2 + z'^2 - c^2 t'^2 = x^2 + y^2 + z^2 - c^2 t^2. \tag{15.9}$$

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. (15.9) 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*.

15-8 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. 15.1). 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

$$\mathbf{p} = m\mathbf{v} = \frac{m_0\mathbf{v}}{\sqrt{1 - v^2/c^2}}. \quad (15.10)$$

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 (15.10), 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 (15.10), 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. 15.10)—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 $\frac{1}{4}$ the thickness of a piece of paper! When the electrons are going that fast their masses are enormous, but their speed cannot exceed the speed of light.

* 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.

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 $m_0/\sqrt{1-v^2/c^2} = m_0(1-v^2/c^2)^{-1/2}$ in a power series, using the binomial theorem. We get

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

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

$$m \cong m_0 + \frac{1}{2}m_0v^2\left(\frac{1}{c^2}\right) \quad (15.11)$$

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 $\Delta m = \Delta(\text{K.E.})/c^2$.

15-9 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 (15.1), if we say that the mass is equal to the total energy content divided by c^2 . If Eq. (15.11) is multiplied by c^2 the result is

$$mc^2 = m_0c^2 + \frac{1}{2}m_0v^2 + \dots \quad (15.12)$$

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 (15.1) 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

$$\frac{dE}{dt} = \mathbf{F} \cdot \mathbf{v}. \quad (15.13)$$

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. (15.13) becomes

$$\frac{d(mc^2)}{dt} = \mathbf{v} \cdot \frac{d(mv)}{dt}. \quad (15.14)$$

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

$$c^2(2m) \frac{dm}{dt} = 2mv \frac{d(mv)}{dt}. \quad (15.15)$$

We need to get rid of the derivatives, which can be accomplished by integrating

both sides. The quantity $(2m) dm/dt$ can be recognized as the time derivative of m^2 , and $(2mv) \cdot d(mv)/dt$ is the time derivative of $(mv)^2$. So, Eq. (15.15) is the same as

$$c^2 \frac{d(m^2)}{dt} = \frac{d(m^2v^2)}{dt}. \quad (15.16)$$

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

$$m^2c^2 = m^2v^2 + C. \quad (15.17)$$

We need to define the constant C more explicitly. Since Eq. (15.17) 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. (15.17) gives

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

We can now use this value of C in Eq. (15.17), which becomes

$$m^2c^2 = m^2v^2 + m_0^2c^2. \quad (15.18)$$

Dividing by c^2 and rearranging terms gives

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

from which we get

$$m = m_0/\sqrt{1 - v^2/c^2}. \quad (15.19)$$

This is the formula (15.1), and is exactly what is necessary for the agreement between mass and energy in Eq. (15.12).

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 experiments 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.