Graphical introduction to the special theory of relativity

L. P. Staunton
Department of Physics and Astronomy, Drake University, Des Moines, Iowa 50311

H. van Dam
Institute of Field Physics, Department of Physics, University of North Carolina, Chapel Hill, North Carolina 27514
(Received 30 July 1979; accepted 26 February 1980)

An approach to special relativity employing space-time diagrams and elapsed time intervals recorded by a pair of stationary observers is presented. The conceptual simplicity of the approach does not preclude its use to obtain all of the usual results. The clock paradox as well as a length paradox are discussed as an illustration of the pedagogical technique.

INTRODUCTION

We present here a somewhat different, entirely graphical approach to Einstein's special theory. While there exist a number of simple approaches to the subject,[1] and several excellent texts at an elementary level,[2] there may be some advantage in having available another elementary introduction tailored to a specific audience. We have found, over several years, this particular method of presentation to be a singularly effective one for use in introductory, nonmajor courses in both physics and astronomy. We do not claim this method to be entirely original, but to our knowledge it has not been published, in English,[3] in any detail, nor in a form in which all of the main results of interest are obtained.

Since it is our intention to recommend this approach to our colleagues for use before suitable audiences, we not only outline our method, but present our discussion in the same elementary style used during its presentation. Accordingly, we have relegated to the footnotes discussions of certain points, such as synchronization of clocks in the same inertial frame, which are not central to our development, and which can become tedious when carefully done. Furthermore, we begin with a reassuring foreword to the audience.

In Sec. I we discuss space-time maps, or diagrams. How to actually construct such diagrams from observations, using radar for example, is explained in Sec. IV. Section II contains arguments leading to the proper time formula. The derivation is based upon three postulates whose uses are noted as they occur: First, the homogeneity and isotropy of space and the homogeneity of time; second, the observed fact that light has a unique, finite speed of propagation which is independent of the motion of the source of the light; third, the "relativity" postulate, that the laws of nature should appear to be the same to anyone, whether at rest or in a state of unaccelerated motion. The final part of the derivation, being somewhat more complicated than the rest, is postponed until Sec. VII, so that we may proceed immediately to the more interesting examples. Therefore, in Sec. II we present an incomplete argument, based upon simplicity, for the correct result.

Section III contains a graphical discussion of the twin paradox. In Sec. IV, where the construction of space-time diagrams is discussed, the concept of relative simultaneity of events arises naturally. Section V contains a brief derivation of the Lorentz–Fitzgerald length contraction, which is followed immediately in Sec. VI by a discussion of a length paradox. In the Appendix, the usual Lorentz transformation formulas are obtained, for completeness.

FOREWORD

Einstein's theory of special relativity is a very deep theory which can lead to complicated mathematics in its most complete applications. The introduction presented here makes use of a very simple graphical technique to arrive at the most prominent result, namely, that if a person is in motion relative to us, then his watch keeps a different time from ours.

In using this graphical technique, we will approach the subject from a point of view which relies heavily on our intuitive ideas of how things really should work. Hopefully we can come away with the conviction that Einstein's result is the *natural* conclusion to draw, and that, in a strange and not completely understood way, our older idea of one absolute time for everyone is, in fact, the *unnatural* point of view.

I. SPACE-TIME MAP

Our basic tool will be what is generally called a space-time diagram,[4] a graph on which we draw a "world line" for each person or thing which we wish to discuss. A world line is really just a tracing of each person's path in space which also shows where he was at any particular time.

In Fig. 1 we have the world line for a person who is standing still until at the time (and place) marked A he starts walking to the right. When he arrives at the place (and time) marked B, he again stands still. Note that time runs upward on the diagram so that to trace out the person's history we follow the world line from bottom to top. A vertical path means that the person is standing at the same x value while some time elapses, i.e., that he is stationary. An event must be described by telling not only where it happened, but when, so that event A has a value of the time coordinate t, as well as a value of the space coordinate x, and the same is true of event B.

Of course we really live in a three-dimensional world, so that a space-time diagram should have four dimensions (three spacial and one time), but we cannot draw four-dimensional figures on a two-dimensional sheet of paper. We will just assume that everyone whom we want to observe

 0002-9505/80/100807-11$00.50

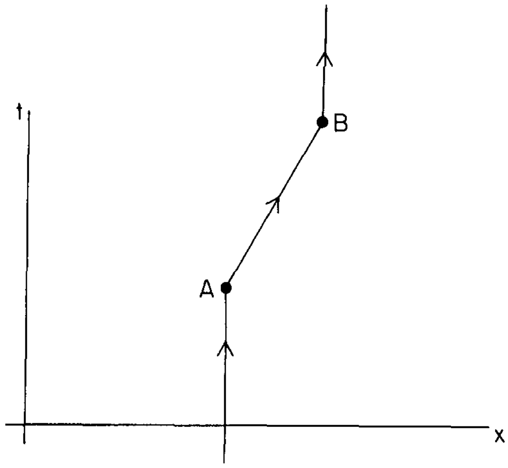

Fig. 1. Space-time map.

will move only in one dimension—to the right or left along the direction we have labeled x.

Our next step should be to choose a scale for our space-time diagrams. We might, for instance, choose to measure time in hours and distance in miles. Then the previous example would now look like Fig. 2. Now we can tell, as well, how far the person traveled D and how long the trip took Δt. The constant speed v at which he walked is given by the familiar result: distance = rate × time, or

$$D = v \times \Delta t. \tag{1}$$

For the example of Fig. 2, 2 miles in 3 h for a speed of $2/3$ miles/h.

We will also add one more element, which has to do with the determination of the times of events. Of course the determination of the distance coordinate x for events such as A and B is easy; one simply consults milestones placed along the path. In practice to determine the time coordinate t, one has a central, accurate clock (represented in Fig. 2 by a church bell) whose ticks are broadcast by radio signals.[5] These radio signals travel at the speed of light, usually denoted c (about 186 000 miles/sec), and are indicated by the dashed lines in Fig. 2.

Unfortunately, these dashed lines represent world lines with a velocity which cannot be distinguished from infinite in Fig. 2, i.e., the lines are drawn horizontal. Because light does have a finite velocity, however, this method for determining the coordinate t is incorrect if one can measure time very accurately, or if large distances (say interplanetary) are involved. Nevertheless, this technique works nicely on this earth and is the reason that in observing everyday events we intuitively feel that time is universal, i.e., may be represented by a horizontal dashed line in Fig. 2 which keeps "marching upward."

In the following discussions, light signals will be crucial, as will be the fact that they travel at a finite, not infinite, speed so that some time is required for them to travel any distance. This effect can be made to show up in our diagrams only if we change the scale rather drastically. This is illustrated in Fig. 3.

In Fig. 3 we have new scales of time and distance chosen so that light signals are drawn at 45°. Time is marked off in seconds while distances are marked off in "light seconds," i.e., the distance light travels in one second (186 000 miles). We have pictured two stationary people. Number 1 sends a light signal to the right. The act of sending out this signal is itself an event which we have marked A. Stationary observer 2 receives this signal, an event which we have marked B. Note that this necessary scale change has reduced the velocities of ordinary persons (a few miles per hour) to values indistinguishable from zero, i.e., in Fig. 3 their world lines are indistinguishable from vertical lines. A world line with a noticeable tilt would represent a velocity which is an appreciable fraction of the speed of light. Our intuitive universal time picture represented by the "upward marching" dashed lines in Fig. 2 has been discarded; in Fig. 3 one would have a series of parallel lines all drawn at 45°. By using space-time diagrams drawn to this new, more physical scale, we can begin to discover one of the results of special relativity.

II. PROPER TIME FORMULA

Let us try to concentrate on the meaning of time itself as it is recorded by the watches of several people (observers).[6] We will consider what happens when three people each record the time interval which elapses between two events, as they see them occur. We label our three people the left man L, the middle man M, and the right man R, arranged as in Fig. 4.

L and R will be just passive, stationary observers, while we will have M send out light signals in both directions at two events A and B which happen to him. For this purpose, we will equip M with a camera flash apparatus constructed so that the light flash is sent in both directions at once, and ask him to flash the light, count off a few seconds on his watch, and then flash it again. In Fig. 4, M is himself also stationary. Now since all light signals "travel" at 45°, the two signals sent, say, to the right are drawn as parallel lines. Then the *elapsed time* as measured by R, Δt_R, is obviously equal to the *elapsed time* according to M, Δt_M. The same is true of the signals sent to the left, so we have the expected result

$$\Delta t_R = \Delta t_L = \Delta t_M, \tag{2}$$

for the case that everyone is standing still. Note that if L and R record the actual times that they *see* the events occur, i.e., when they receive the light signals, they will record later actual times than will M. This is quite natural since some

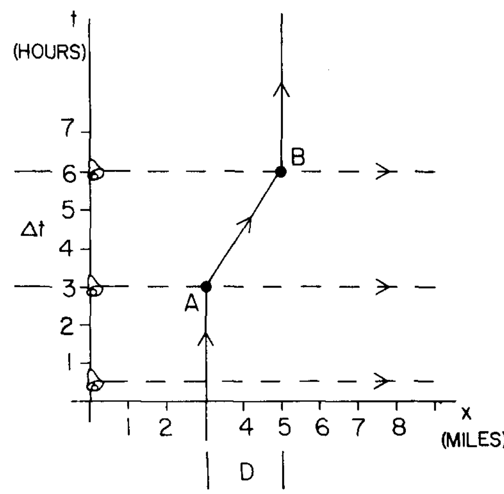

Fig. 2. Space-time map marked in hours and miles. The horizontal dashed lines represent radio signals broadcast from a central clock (church bell).

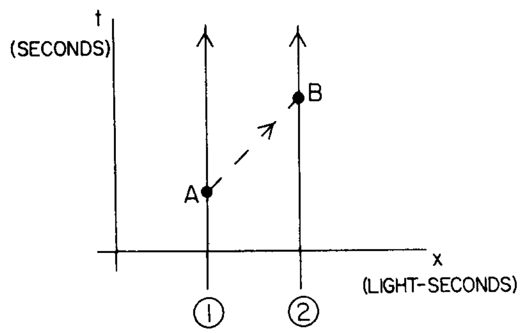

Fig. 3. Space-time diagram marked in seconds and light seconds. To this scale, dashed lines, representing light signals are drawn at 45°.

time passes while the light signals travel. We are concerned only with the number of seconds by which event B is later than event A, and all will agree on this number.

Now for the interesting case. Suppose that L and R remain stationary, but that M is in motion, at a (very rapid) steady speed, to the right, say, while we repeat the process. We draw this situation in Fig. 5.

At event A, M starts his stopwatch and at event B he stops it. He will record this time interval from his own watch, and we will call it a new name, $\Delta\tau$, to remind us that this is the time measured by a moving observer. Note that $\Delta\tau$ is *not* the length of the line segment drawn between the events A and B. Since this segment is not vertical, part of its length as drawn is due to the spacial distance between the events, and the task which will occupy us is to unravel this dependence to determine $\Delta\tau$. Note also that the light signals sent out by M are drawn at 45°. This reflects the fact that *the speed of light is always the same, for everybody, no matter how the light is produced, or by whom.* Naively, we might expect that since M's flash apparatus was traveling to the right, the light signals should travel faster to the right and slower to the left. This is, in fact, not the case.

We would have obtained exactly the same result if we had confiscated M's apparatus, and instead had placed two flash mechanisms at rest along the route and instructed M to flash them as he went flying along. This seemingly strange result is due to the nature of light: it is pure energy and behaves differently from material objects.[7] The flashing of a light merely serves to *deposit* a certain amount of pure energy into space-time *at some event*. Thereafter, of its very nature, this pure energy travels along at the speed of light "away" from that event, not from the person or thing which caused the event. Once we grasp that light naturally *inhabits* space-time, and needs no memory of its prior history, this result becomes understandable.

The independence of the velocity of light upon the motion of its source is one of the central principles upon which the theory of special relativity is based. It is a confirmed experimental fact which was discovered before the turn of the century and puzzled everyone until Einstein published his theory. Basically, Einstein's attitude was, "Look, this is a confirmed fact of nature, so if it doesn't fit with your prejudices, you will just have to change your point of view."

Now look again at Fig. 5. It is obvious that Δt_L is much longer than Δt_R, so that our two stationary observers measure different elapsed times between the events A and B. This is quite natural since, for example, event B happens much closer to R than does event A. Consequently, the light signal from B reaches him much quicker than did the signal from event A, and he measures a small elapsed time between the events *as he sees them.* Exactly the opposite is true for L, so that he measures a much longer elapsed time between the events as he (L) *sees* them.

What then will the moving observer M say? Surely he has the best view of the two events. Since they both happen to him, he is the observer "on the spot." He records an elapsed time by looking at his own watch. This time interval will be called the "proper time," since it is measured by the moving observer who "properly" consults the watch which he is carrying along with him, rather than trying to look at our clocks. Our stationary clocks, even if spread all along his route, would be very difficult for M to use since from his point of view, they are speeding past him as he tries to read them.

We will try to determine a rule, called the T rule, for calculating $\Delta\tau$ if we know the intervals Δt_L and Δt_R measured by our stationary observers:

$$\Delta\tau = T(\Delta t_L, \Delta t_R). \tag{3}$$

We must find out what this function T is. We already know something about it from Fig. 4 and Eq. (2): if $\Delta t_L = \Delta t_R$, then $\Delta\tau = \Delta t_L = \Delta t_R$. If in this case we call Δt the number to which they are all equal, then

$$T(\Delta t, \Delta t) = \Delta t. \tag{4}$$

Also, it should be clear that if we drew another figure exactly like Fig. 5 except that Δt_L and Δt_R were, say, twice as big as in Fig. 5, then $\Delta\tau$ should come out twice as big also. In general, this means that if we replace Δt_R and Δt_L by $s \times (\Delta t_R)$ and $s \times (\Delta t_L)$, where s is any number, then

$$T(s \times \Delta t_L, s \times \Delta t_R) = s \times \Delta\tau = s \times T(\Delta t_L, \Delta t_R), \tag{5}$$

and this principle is termed the "homogeneity of time."

Finally, if we drew Fig. 5 exactly the same except that M moved from R to L instead of from L to R, he should still measure the same elapsed time $\Delta\tau$ on his own watch. This is an expression of the "isotropy of space": a trip from one point to another is exactly the same as one in the opposite direction at the same speed.[8] Therefore,

$$T(\Delta t_R, \Delta t_L) = T(\Delta t_L, \Delta t_R) \tag{6}$$

since the only thing that would change is that the time intervals measured by L and R would just be exchanged.

There are only two *simple* functions T which have the properties (4), (5), and (6). Either

$$T(\Delta t_L, \Delta t_R) = (1/2) \times (\Delta t_L + \Delta t_R) \qquad \text{(Newton)} \tag{7}$$

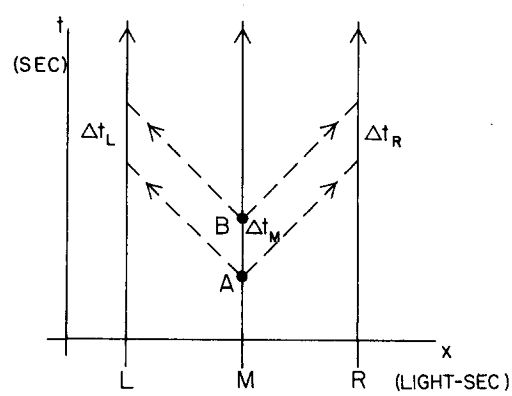

Fig. 4. Space-time diagram with three stationary observers.

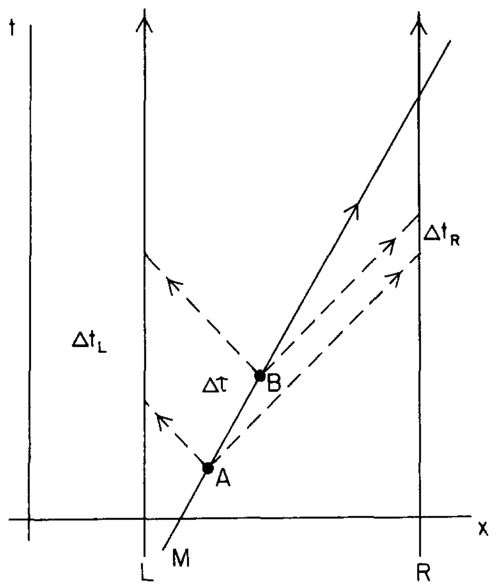

Fig. 5. Space-time diagram with two stationary observers and one in rapid motion. The moving observer's watch records an elapsed time $\Delta\tau$, which is *not* simply the length of the line between A and B.

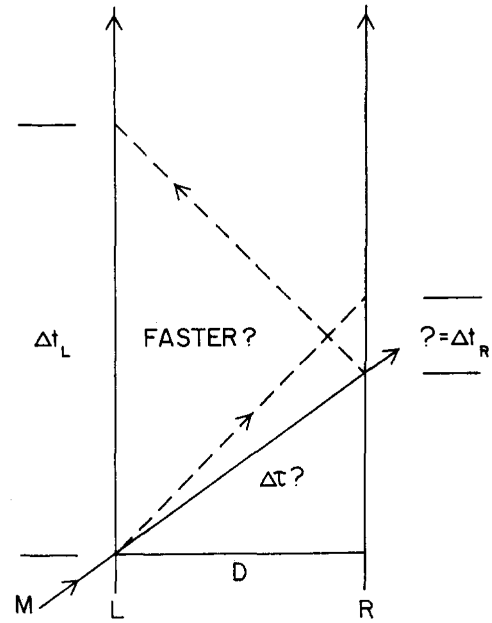

Fig. 6. Hypothetical space-time diagram representing a trip faster than the speed of light.

or

$$T(\Delta t_L, \Delta t_R) = \sqrt{(\Delta t_L) \times (\Delta t_R)} \qquad \text{(Einstein)}. \quad (8)$$

The first choice leads to Newton's concept of a universal, absolute time and has been proven wrong by experiment. The second choice (8) is actually the correct choice, the one that nature makes. If one has the point of view that the laws of nature should be both simple and interesting, then one would expect (8) to be the correct answer because simplicity singles out either (7) or (8), and, as we shall see, (8) has some *very* interesting consequences.

The proof that (8) is the correct rule of nature requires some simple geometry and one further principle called the "relativity postulate," which states that the laws of nature should appear to be the same to anyone, whether they are at rest or in steady motion. Nevertheless, we will put this proof off until Sec. VII and begin right away to investigate some of the more interesting consequences of (8), the proper time formula:

$$\Delta\tau = \sqrt{(\Delta t_L) \times (\Delta t_R)}\,. \quad (9)$$

For example, let's ask whether anyone can ever travel faster than light? Figure 6 shows a *hypothetical* space-time diagram for the following sequence of events: M travels a little faster than light across the distance D between L and R. The two events to be timed are M's departure from L, at which L sends a light signal to R to tell him that M is on the way, and M's arrival at R, at which R sends a light signal back to L to tell him that M has just passed by.

Note first that L sees M go past on his way to R, and a long time Δt_L later receives the light signal from R confirming M's arrival. Nothing unusual here. R, however, has an unsettling experience: he sees M arrive, going past him to the right, *before* he sees him depart (in broad daylight, the light signal from L could be replaced by the natural illumination of the event)! The elapsed time interval between the two events as recorded by R would therefore be negative, while Δt_L is, of course, a positive number. When we now ask how much time has elapsed, according to M, the proper time formula, (9) yields an imaginary answer, since Δt_R is a negative multiplier under the square root. This would of course be hotly disputed by M, who would claim that his watch is no more imaginary than he is! Our conclusion from all of this is, of course, that *both* M and his watch are purely imaginary, and no one can ever travel faster than the speed of light.

Let us consider, a real situation, namely, what happens when M makes his trip from L to R very, very fast, but a bit slower than light. This situation is diagrammed in Fig. 7.

We notice that L sees virtually the same sequence of events and records a long time interval Δt_L. In contrast, however, R now sees events unfold in a reasonable fashion, and records the very short, positive time interval Δt_R from his own observations. Now since L and R are a distance D

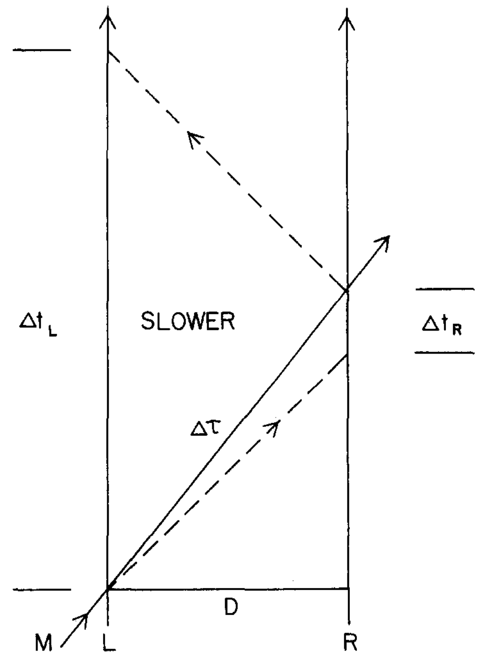

Fig. 7. Space-time diagram representing a real trip at slightly less than the speed of light.

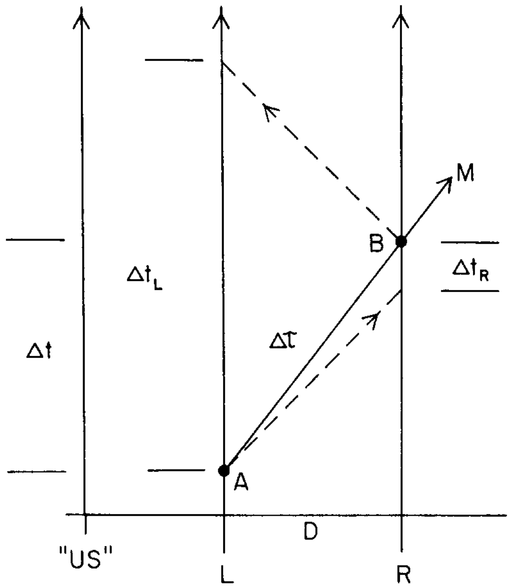

Fig. 8. Comparison of the collective observed trip time Δt to the three observed elapsed times Δt_L, Δt_R, and $\Delta\tau$.

apart, and light signals travel at speed *c*, the time for a light signal to pass from L to R, or vice versa is just D/c. From Fig. 7 we see that the time recorded by L is just twice the travel time for one of the light signals, plus the short time delay measured by R, or

$$\Delta t_L = 2D/c + \Delta t_R. \tag{10}$$

Suppose that the distance between L and R is 7-½ light seconds, for example, and that R measures an elapsed time of 1 sec. Then (10) indicates that $\Delta t_L = 2 \times 7\text{-}\frac{1}{2} + 1 = 16$ sec. The proper time formula (9), however, tells us that according to the watch carried by M, the trip lasts a total of $\sqrt{16 \times 1}$ or only 4 sec. Of course, we should expect M's watch to yield an elapsed time somewhere *between* the long time recorded by L and the short time recorded by R. The result, however, becomes remarkable when we carefully inspect our space-time map and realize that the trip took 1 sec longer (Δt_R) than it took the initial light signal from L to arrive at R, or a total of 8-½ sec!

We have arrived at the fascinating result that if a person is in motion with respect to us, then time itself actually evolves more slowly for him than it does for us watching him. Furthermore, the faster he travels, the more exaggerated this effect becomes, since the faster he travels, the smaller becomes the number Δt_R, and this small number is a *multiplier* in the proper time formula (9).

With a little care, we can relate the time interval M measures to his speed and to how long we "know" it took to make the trip. Figure 8 illustrates the point. "We" will make the comparison by arranging with L and R who are at rest along with us to act as our agents. First, we have L and R synchronize their watches with ours.[9] M now makes his rapid journey and we later obtain from L and R not only their measurements Δt_L and Δt_R, but also the time on L's watch when M left and the time on R's watch when he arrived. We subtract these last two numbers to obtain the time interval Δt shown. This is how long we "know" the trip took. M measures of course a small time interval $\Delta\tau$ which we can calculate from Δt_L and Δt_R. We can then report that "our collective watches measured Δt for the trip, while M's measured $\Delta\tau$, so his watch must be running slowly compared to ours."

We can relate $\Delta\tau$ and Δt exactly. The time it takes for the light signal from event B to reach observer L is D/c. Therefore

$$\Delta t_L = \Delta t + D/c.$$

The light signal from event A reaches R in D/c also since it must travel the same distance, so that

$$\Delta t_R = \Delta t - D/c.$$

Then

$$\Delta\tau = \sqrt{(\Delta t_L)(\Delta t_R)}\,,$$

$$\Delta\tau = \sqrt{(\Delta t + D/c)(\Delta t - D/c)},$$

or

$$\Delta\tau = \sqrt{(\Delta t)^2 - D^2/c^2}\,. \tag{11}$$

We remarked earlier that $\Delta\tau$ was not simply the length of the line segment drawn between the events A and B, and that we would seek to unravel its dependence upon the elapsed time between the events, Δt, and the distance between them, D (both according to us). We have just completed this task[10]; however, we can express our result in a simpler and more convenient form as follows. Suppose M travels at the constant speed v, relative to us, in moving between the events A and B. Then since he covers the distance D, in the time Δt, according to us,

$$D = v \times \Delta t.$$

If we substitute this into (11), we obtain

$$\Delta\tau = \sqrt{(\Delta t)^2 - v^2(\Delta t)^2/c^2},$$

$$\Delta\tau = \sqrt{(\Delta t)^2(1 - v^2/c^2)},$$

or as our final result,

$$\Delta\tau = \Delta t\sqrt{1 - v^2/c^2}. \tag{12}$$

This is the exact relativistic result relating the proper time $\Delta\tau$, measured between events *which happen to him*, by an observer moving relative to us at a constant speed v, to the time Δt which we *collectively* observe between the same two events.

As an example, suppose that M's speed is 80% of the speed of light according to us. Then v/c is 0.8, $v^2/c^2 = 0.64$, $1 - 0.64 = 0.36$, and the square root of this number is 0.6. In this case we find that $\Delta\tau = \Delta t \times 0.6$, so that if we observe a time interval Δt of 10 y, say, M will observe only $10 \times 0.6 = 6$ y. Similarly, if M's speed were 99.99% of the speed of light, and we (collectively) observed a trip time of 10 y, only 6 weeks would have passed for M! On the other hand, a car traveling on an interstate will have a typical speed of about 80 *feet*/sec., while the speed of light is about 186 000 *miles*/sec. In this case the ratio v/c is about 0.000 000 009 and the difference between clock rates is so very, very small that measuring it is impossible. That is why our intuition of one absolute time for everyone is not *obviously* wrong to us as we observe earth-bound objects.

III. TWIN "PARADOX"

One of the more interesting applications of our results is the so-called clock, or twin, paradox.[11] Two identical twins choose very different careers, I becoming an accountant, and II an astronaut. II is sent on a mission to investigate a very distant stellar system S. He is to travel at very nearly the speed of light to S, spent 1 y collecting data there, and return, again at very nearly the speed of light. Except for brief periods of time during take-offs and landings, when his rocket will be firing, II will be coasting in space, with no physical sensation of motion, so that our proper time formula applies to his watch, and also to his body, since his heart and metabolism are natural biological clocks.

Now since most of the trip II will be moving at very nearly the speed of light with respect to brother I, who remains on Earth, time will evolve much more slowly for him than for brother I. Therefore if his speed is suitably adjusted, he can return to earth having aged only two or three years to confront a very elderly, long retired accountant, brother I, at the landing field. This effect is quite real, and while it is not likely to occur during our lifetimes, it may well become commonplace sometime in the next few centuries.

Now a true paradox is a puzzle, a set of circumstances which gives rise to two different and mutually contradictory outcomes. The so-called twin paradox, however, is a puzzle which apparently (but not in fact) results in a contradiction.[12] The possible confusion arises when we consider the same sequence of events from the point of view of brother II. During *most* of the trip, he has no physical sensation of motion, and so, according to the relativity postulate, may consider that he is himself at rest while brother I, and the entire Earth, speeds away from him, and returns. Consciously, of course, he "knows" better, but if we concentrate on his watch, or metabolism, our point is a valid one. Consequently, brother II might reason that since brother I is the twin in rapid motion, I's watch, and metabolism, is running slower than his own. He may therefore expect to be met at the landing field by an accountant who is younger than himself. This is the apparent paradox, since there can be only one outcome to the meeting.

That our first analysis is the correct one, and that no contradiction arises, follows from a careful inspection of a space-time map of such a situation. As an additional aid to clarification, let us suppose that the trip begins on their mutual 25th birthday, and that brother I, on earth, broadcasts a happy birthday message to his traveling twin at intervals of one Earth year over the entire time span of the trip. Furthermore, brother II broadcasts a message back to earth upon his arrival at star system S. These circumstances are represented in Fig. 9, where the scale is years, and light years, and II's speed has been chosen to be only 88% of the speed of light, for convenience in drawing the diagram.

From Fig. 9 one may read off that an inhabitant on S will observe the outward bound trip to require 1 y (Δt_R), while the safe arrival signal reaches Earth some 16 y after II's departure (Δt_L). Since the return trip is made at the same speed, brother I on earth, changing roles with S for the return trip, will observe the return to require 1 year also. Adding, then, the approximately one year that II spends at S, I greets his returning brother after about 18 Earth y have elapsed, when I is 43 y of age.

What of brother II? Except for the short periods of acceleration (a day or two each) our proper time rule, Eq. (9), tells us that the outward bound trip requires 4 y of II's time ($\Delta t_L = 16$ y, $\Delta t_R = 1$ y). He remains about 1 y (while ev-

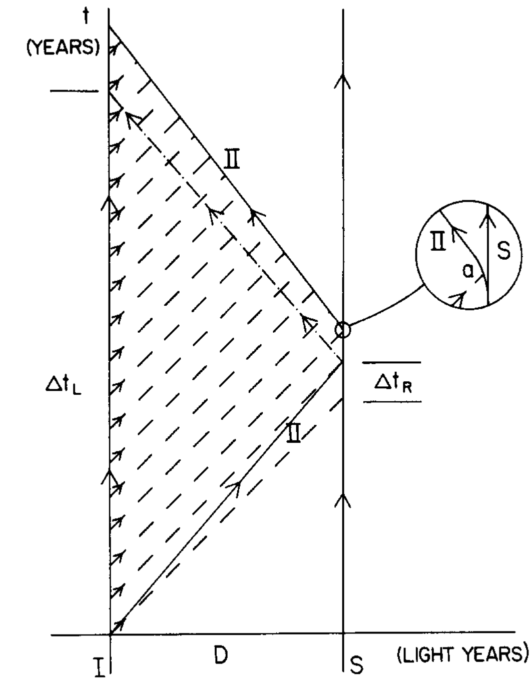

Fig. 9. Graphical representation of the twin paradox. I remains at home while II travels at about 88% of the speed of light a distance $D = 7.5$ light years to star system S. Yearly greetings - - - are broadcast by I, while II sends only a safe arrival message — - —. The inset shows one of the four acceleration stages, marked *a*.

eryone is at rest, so that their clock rates are the same), and requires another 4 y of his own time to return. There are also four very small but mysterious amounts of time which elapse on II's clock while he is accelerating, events represented by small curved portions of his world line. We are not in a position to calculate these times, but since they are very small,[13] we conclude that he returns to earth after about 9 y have elapsed for him, or when he is chronologically, and biologically, only 34 y of age.

Now what about the paradox? We notice that II receives the first of his yearly birthday messages from I upon his arrival at S, some 4 y of his own time after his departure. Of course, II realizes that the signal was not sent three full years late. He knows that he was, after all, some distance from Earth after 1 y and that consequently the message took some time to reach him. Since he has been traveling at 88% of the speed of light, or from his perspective, the Earth has been traveling away from him at that speed, he concludes the earth should have been 0.88 light years away at the time that the signal was sent. *Now the speed of light is the same for everyone, no matter how it is produced,* and from II's viewpoint, he is at rest, when a signal is sent from a rapidly moving Earth at a distance of 0.88 light years. It should have been received, therefore, 0.88 y later. Since a timely message would reach him, then, after 1.88 y, but the actual message arrives after 4 y have elapsed, II concludes that the message was sent much too late, or that I's watch is running much slower than his!

If the tale ended here, we would indeed have a contradiction. Fortunately it does not. II receives the second birthday greeting just as he is departing from S, 1 y after receiving the first message. Notice, however, what II experiences during his return trip. While 4 y of his own time elapse, he receives 16 birthday messages from Earth. With

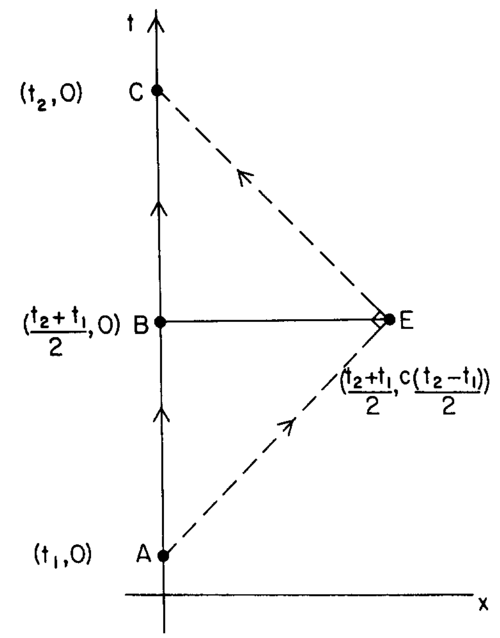

Fig. 10. Construction of a space-time diagram by a single observer O equipped with a clock and radar apparatus.

a jolt he realizes that I is aging much faster than he is, and fully expects to be greeted by an older I upon his arrival.

A little thought leads to the conclusion that the paradox was avoided when II stopped and turned around. During a period of acceleration, such as that marked *a* in the insert to Fig. 9, II experiences tremendous forces which cause him to realize that he is changing his state of motion and that, as a consequence, the clocks of those he has left behind are running at a rate different from his own. Had he continued on his outward journey forever, he could and would observe all other watches to be slower than his own, due to his change of state, but then, of course, he could never return to confront his twin. Having made those subsequent changes of his state of motion which allowed him to return, however, he no longer can apply the relativity postulate to claim that he is, *and has been,* in a state of rest while the rapidly moving Earth turned about and returned to him.

IV. CONSTRUCTION OF SPACE-TIME DIAGRAMS; SIMULTANEITY

Let's digress for a moment into the question of how the t-x maps or space-time diagrams are actually constructed. One observer equipped with an accurate clock and a radar device is all that is necessary (Note: no rulers!). The technique is illustrated in Fig. 10.

The world line of the observer O, who is at rest, is the vertical line, and the location coordinate x is zero on that line. The dashed line is that of one of the radar signals sent out by O which is bounced back by some event E. Now O is aware that radar signals travel at the velocity of light c and that it takes one exactly one exactly as much time to return from an event as it does to go out to one. Since O sends out the signal at event A, at time t_1, and receives the echo at event C, time t_2, he realizes that event B, halfway on his worldline between A and C occurred simultaneously with E. B therefore has the coordinate values $[(t_2+t_1)/2, 0]$ and E has the same *time* coordinate value. Now the total elapsed time for the round trip of the radar signal is the difference t_2-t_1. Traveling at the speed of light, the signal then covered a total distance $c(t_2-t_1)$. Keeping in mind that the signal traveled straight out and came straight back (Fig. 10 is a space-*time* map), O realizes that the distance to the event E is just half the total distance, or $c(t_2-t_1)/2$. Since O himself is at rest at the x-coordinate value zero, E has the x value, $c(t_2-t_1)/2$. Finally, note that as a geometrical figure, triangle AEC is a right triangle, that the line segments AB, BC, and also BE are of equal length, and that the line BE is parallel to the x axis.

Next consider another such space-time map, this time constructed by an astronaut moving at a constant velocity with respect to the original coordinate system. The astronaut uses his own (moving) clock and radar system to construct his $t'-x'$ map. The situation is illustrated in Fig. 11.

The world line of the astronaut, as constructed by O in the original coordinate system, is the tilted line marked t'. The dashed line is the radar reflection of an event E, which is the same event to both observers. (Remember that an event occurs sometime, and someplace according to all observers, although they may disagree as to where and when.) For convenience in making our point, we have arranged in Fig. 11 that O and the astronaut are at the same location at event A, the sending out of the radar signal, and that both observers receive the echo, although at different events, C′ and C.

To the astronaut, an event B′ on his own world line, halfway between A and C′ is simultaneous with the event E. Further, a line B′E, drawn on his own map, will be parallel to his x' axis. In fact, all lines of simultaneity will be parallel to the line B′E, and all lines which mark a fixed location at a certain distance from himself will be parallel to his own world line. Finally, to the astronaut, on his own map, the line segment B′E will be of equal length to both AB′ and B′C′.

These geometrical facts are also true, however, in Fig. 11, which is drawn by O in the original t-x system. This

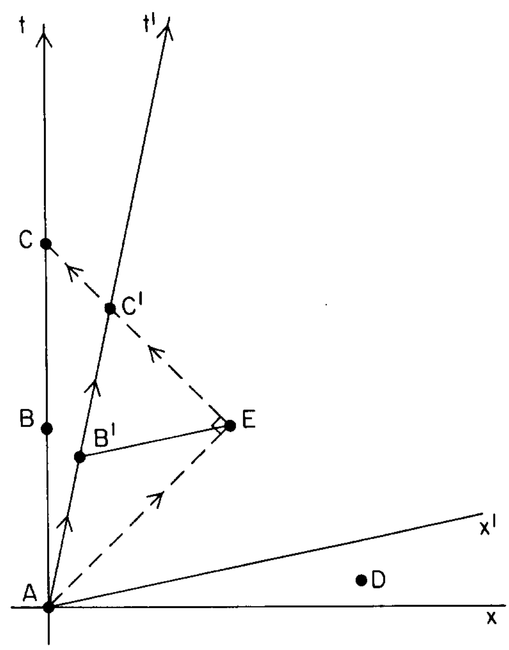

Fig. 11. Space-time axes of a moving observer, constructed by use of his (moving) apparatus, but drawn in the original coordinate system by O. The chronological ordering of events A and D is reversed for the two observers.

follows from the fact that AB′ = B′C′ in the right triangle AEC′. Further, in the same figure, since ∢ B′AE = ∢ B′EA, it follows that the angle which B′E makes with the x axis (and which therefore the x' axis makes with the x axis) is the same angle as that which the t' axis, the world line of the astronaut, makes with the t axis.

An inspection of Fig. 11 should then make it obvious that whether or not two events occurred simultaneously is a matter of opinion, the answer depending on which observer is consulted. The observer O would claim that B′ occurred before E, while the astronaut would find them to be simultaneous.

Similarly, the astronaut would find that B occurred after E, while O would consider them simultaneous. Which observer is correct, you might ask? The proper response is that both are correct; simultaneity is entirely relative. Finally, if we ask each observer to simply rank in proper chronological order event A, and the isolated event D shown in Fig. 11, we will obtain conflicting answers. Since A lies on O's x axis, while D lies above it, O will rank them: first A and then D; in contrast, while A lies also on the astronaut's x' axis, but D lies below it, he will rank them: first D and then A. Again, both are correct from their relative points of view.

Surely, you might conclude, disagreements of this sort can lead to paradoxical situations. Indeed, after we obtain one more result, we will investigate one such situation, the "pole vaulter paradox," in Sec. VI.

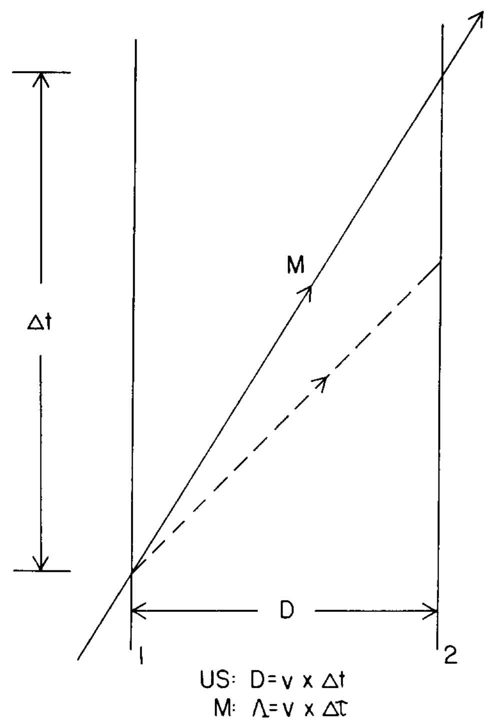

Fig. 12. M moves at speed v relative to us over a measured distance D. From his perspective, the markers 1 and 2 whiz by him at the same speed v, but in the opposite direction. M measures a distance Λ between the markers.

V. LENGTH CONTRACTION

We have already seen that a watch carried by a moving observer runs at a rate different from our own. Since time is not an absolute for everyone, it should not be surprising that space, or distance, is also relative.

We can develop this result most easily by setting out two objects a measured distance D apart and asking a moving observer M how far apart they are. The experiment is shown in Fig. 12. We determine the elapsed time Δt for M to pass between the markers 1 and 2 in the usual way. Consequently, we can assign to M a speed v, relative to us, such that $D = v \times \Delta t$. M, of course, measures on his own watch a smaller elapsed time $\Delta\tau$, given by (12).

Now applying the relativity postulate, we realize that from M's point of view, he is at rest while markers 1 and 2 pass him by at the same speed v, but in the opposite direction. M therefore assigns them a distance apart Λ, given by the *same law of nature,* distance = rate × time, or $\Lambda = v \times \Delta\tau$. Notice that M uses, of course, the elapsed time $\Delta\tau$ on his own watch. Since this is a smaller elapsed time than we measure, the distance Λ which he obtains is smaller than our D.

We can make this precise. Since $\Delta\tau$ is related to Δt by (12), M obtains the result

$$\Lambda = v\Delta\tau = v\Delta t \sqrt{1 - v^2/c^2}\,.$$

But the distance we obtain, D, is just $v\Delta t$. Substituting this quantity, we obtain the result

$$\Lambda = D\sqrt{1 - v^2/c^2}, \qquad (13)$$

relating the length Λ of an object measured while it is in motion, along its length at speed v, to the length D obtained by measurements made by those for whom it is at rest. This result is commonly known as the Lorentz–Fitzgerald length contraction.

As an example, consider an arrow 10 in. long. If this arrow could be shot past us at 80% of the speed of light, we would measure its length, as it passed us, to be only 6 in. Of course, at the ordinary speeds of our everyday experience, we do not notice this effect, since it is then very, very small.

Finally, let's consider that our result (13) should not really be surprising. Since we have established that our initial prejudice in favor of one absolute time rate must be abandoned, we should have no reason to expect that our similar initial prejudice in favor of an absolute measure of distance can be retained.

VI. POLE VAULTER PARADOX

An interesting application of the length contraction result is the so-called pole vaulter paradox.[14] We give a 20-ft-long vaulting pole to a swift runner, tell him to go some distance away and come running back, holding the pole horizontally, toward a barn 30 ft deep. The front door of the barn is open, while the back door is closed. We reassure him that it is our intention to open the back door just before the leading edge of his pole arrives there so that no harm will come to him in the process. Furthermore, we tell him that, as an experiment, we plan to close the front door just after he has passed.

Since the barn is fully 10 ft deeper than the pole is long, we reason that there will be a short span of time when the pole will be completely enclosed within the barn. Accordingly, we plan to *first* close the front door behind him, and *then* to open the back door in front of him, and we tell him this also. For safety we will station an observer at each door

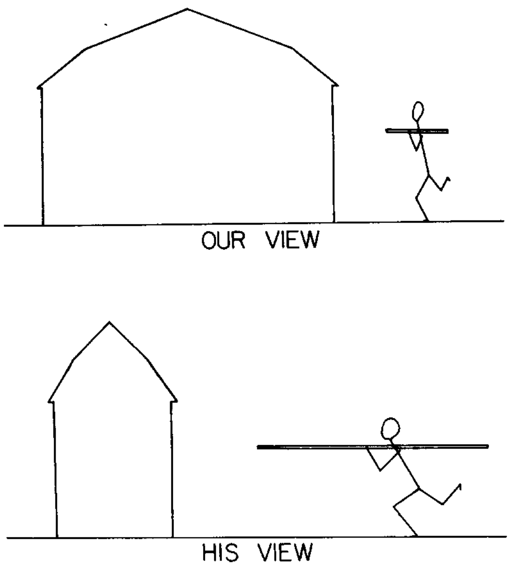

Fig. 13. Pole vaulter paradox. Top view: we observe a runner, carrying a 2-ft-long pole, approaching an open barn 30 ft deep. Bottom view: the runner, carrying a 20-ft-long pole, observes an open barn 3 ft deep to be approaching him.

and equip them with door mechanisms which operate almost instantaneously.

Although he has some misgivings over the arrangements, the runner finally agrees to the plan, and comes running as fast as he can, about 25 miles/h. The experiment, of course, succeeds as designed, with the runner as well as the vaulting pole being completely enclosed within the barn for a short period of time.

Having gained his confidence, we now propose to him a second experiment. We show him a remotely controlled flying platform, and ask him to stand on it, holding the vaulting pole in the same horizontal position, while we repeat the procedure at 99.5% of the speed of light. Being completely ignorant of the theory of relativity, he agrees.

Now at 99.5% of the speed of light, (13) indicates that a moving object will be contracted in the direction of its motion to $1/10$ of its natural length. As the experiment begins, therefore, our forward observers inform us, via radio, that a very thin runner, holding a pole just 2 ft long is approaching our 30-ft-deep barn.[15] Therefore, we have no reason to alter our plans, and will close the front and *then* open the back door, in order that for a brief span of time, the runner will be enclosed. Figure 13 illustrates the situation.

Unhappily, however, the runner, *who has no physical sensation of motion,* observes a barn only 3 ft deep rushing toward him while he is simply standing on a platform holding a 20-ft pole in a horizontal position. Belatedly, he recalls that we plan to *first* close the front door behind him and *then* to open the back door before the vaulting pole reaches it. Since a 20-ft pole cannot possibly be enclosed in a 3-ft space, he concludes that he is about to be sacrificed to science.

The possible paradox, of course, is that there can be only one outcome to the experiment: either the runner meets his doom, as he expects, or no collision occurs, as we expect. Perhaps you would like to take a few moments to decide for yourself what will occur and why.

Once again, the resolution of the apparent paradox follows from a careful inspection of a space time diagram. In Fig. 14, the two events of interest are event A, the closing of the front door behind the pole, and event B, the opening of the rear door just before the leading edge of the pole arrives.

Since we have designed the experiment, and fully expect the pole to be briefly enclosed within the barn, that is indeed what we observe. In our space-time map of the experiment, event A occurs before event B. Now since we observe no collision to occur, the runner, of course, must observe the same thing. His explanation of this unexpected outcome is that, in fact, we opened the rear door first, so that his vaulting pole could pass safely beyond the back of the barn, long before we closed the front door behind him, i.e., the chronological order of the events on his (t'-x') space-time map is first B and then A. The fact that simultaneity is relative, as is the temporal order of certain events, is the resolution of the "paradox."

VII. PROOF OF THE PROPER TIME FORMULA

Now that we have shown that our proper time rule,

$$\Delta\tau = T(\Delta t_L, \Delta t_R) = \sqrt{(\Delta t_L)(\Delta t_R)}, \qquad (14)$$

has some very interesting consequences, we would like to end by completing our demonstration begun in Sec. II, that it is the correct law of nature. For this purpose it is necessary to use two moving observers, M_L and M_R, traveling together at the same speed, as well as our usual pair of stationary observers, L and R. Their world lines are shown in Fig. 15.

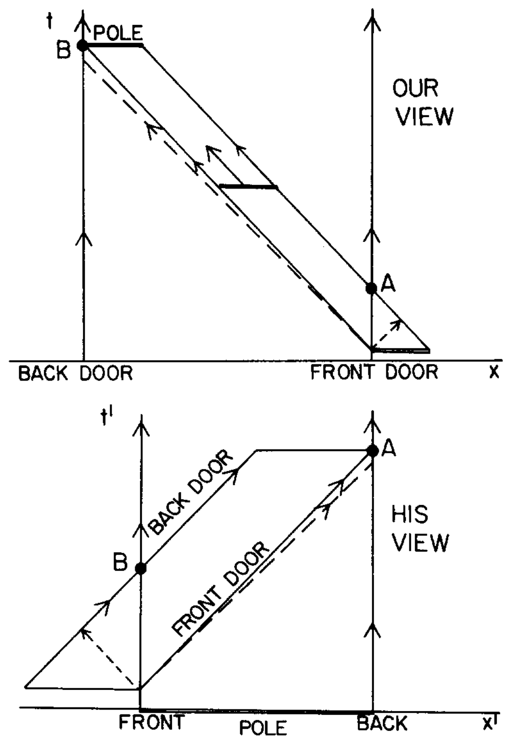

Fig. 14. Space-time diagrams of the temporal sequence of two events: Event A, the front door is closed behind the pole, and event B, the rear door is opened before the pole. Top view: we observe A to precede B, so that the pole is enclosed. Bottom view: the runner observes B to precede A, so that some part of the pole always protruded from the barn. For clarity, light signals sent at the beginning of the experiment are also shown.

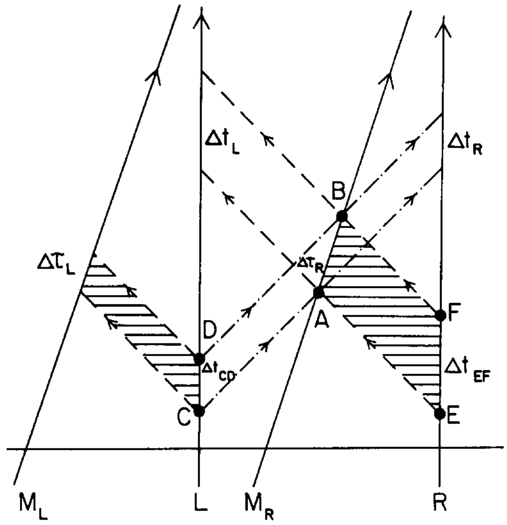

Fig. 15. World lines of two moving observers, M_L and M_R, as well as those of L and R. Use of the relativity postulate, that the laws of nature must be the same for M_L and M_R as they are for L and R, leads to the correct proper time formula.

Ignore, momentarily, the cross hatching in the diagram, and notice that $\Delta\tau_R$ represents the proper elapsed time between the events A and B, which were shown in Fig. 5 of Sec. II. We now consider these events to be those of the arrival of light signals sent out by L at events C and D. $\Delta\tau_L$ represents the time interval between C and D as observed by the second moving observer, M_L.[16] Further, we imagine that R sends out light signals at events E and F so that they arrive exactly at events A and B. This may in practice be hard to arrange exactly, but it is perfectly possible.

Now we have the (as yet unspecified) T rule

$$\Delta\tau_R = T(\Delta t_L, \Delta t_R), \qquad (15)$$

which allows the calculation of the proper time between two events, as witnessed by a moving observer, from the two observed elapsed times Δt_L and Δt_R. At this point we must invoke the principle of relativity: the laws of nature must be the same for observers moving with a constant velocity. After all, M_L and M_R have no physical sensation of motion, and to them, it is L and R who are moving (to the left).

From the perspective of M_L and M_R, L, who is a moving observer, sends out light signals from events C and D, which occur over a *proper* elapsed time interval Δt_{CD}, as measured by L. M_L and M_R, *stationary* observers, record the intervals $\Delta\tau_L$ and $\Delta\tau_R$. Using the same law of nature, the T rule, M_L and M_R would therefore calculate

$$\Delta t_{CD} = T(\Delta\tau_L, \Delta\tau_R). \qquad (16)$$

Now from Fig. 15 it should be clear that the elapsed time interval Δt_{CD} is numerically equal to the interval Δt_R measured by R. Equation (16) therefore may be written

$$\Delta t_R = T(\Delta\tau_L, \Delta\tau_R). \qquad (17)$$

From the same figure, it should be equally clear that Δt_{EF}, as measured by R, is numerically equal to the interval Δt_L. Now focus on the two cross-hatched regions, and observe that as geometrical figures, their boundaries are exactly parallel, one to the other. It follows that the ratios of their similar sides must then be equal[17] or

$$\Delta\tau_L/\Delta\tau_R = \Delta t_{CD}/\Delta t_{EF}.$$

Since $\Delta t_{CD} = \Delta t_R$ and $\Delta t_{EF} = \Delta t_L$, we have, then

$$\Delta\tau_L/\Delta\tau_R = \Delta t_R/\Delta t_L$$

or

$$\Delta\tau_L = \Delta\tau_R(\Delta t_R/\Delta t_L).$$

We substitute this result for $\Delta\tau_L$ in (17) to obtain

$$\Delta t_R = T((\Delta\tau_R)(\Delta t_R)/\Delta t_L, \Delta\tau_R),$$

and then multiply the final $\Delta\tau_R$ above by $\Delta t_L/\Delta t_L$, which is numerically equal to 1, to obtain

$$\Delta t_R = T((\Delta\tau_R/\Delta t_L)\Delta t_R, (\Delta\tau_R/\Delta t_L)\Delta t_L). \qquad (18)$$

Recall now the principle of the "homogeneity of time," as expressed in (5), and identify s as $\Delta\tau_R/\Delta t_L$. It then follows that

$$\Delta t_R = (\Delta\tau_R/\Delta t_L)T(\Delta t_R, \Delta t_L)$$

or

$$(\Delta t_L)(\Delta t_R) = \Delta\tau_R T(\Delta t_R, \Delta t_L). \qquad (19)$$

But the principle of the "isotropy of space," (6), indicates that the T rule on the right-hand side of (19) is the same T rule as in (15). It follows then that the right-hand side of (19) is just the square of our T rule, or

$$(\Delta t_L)(\Delta t_r) = [T(\Delta t_L, \Delta t_R)]^2.$$

Therefore,

$$T(\Delta t_L, \Delta t_R) = \sqrt{(\Delta t_L)(\Delta t_R)}.$$

APPENDIX

For completeness we demonstrate here how our method can be utilized to obtain the usual Lorentz transformation formulas. The majority of the necessary material has been presented in Sec. IV. Additionally, we require the use of (12) and (13).

We wish to obtain an expression for the values of the coordinates (t, x) of an event in terms of the values (t', x') assigned to the event by a moving observer. The geometry is that of Fig. 16, where t and x are marked t_E and x_E. The astronaut considers event B, on his own world line, to be simultaneous with event E, and so assigns the value $(t', 0')$ to event B.

We will suppose that at event 0, both observers set their respective clocks to zero, so that 0 serves as the mutual origin of the two coordinate systems. (Note that E does *not* lie on a light signal line from 0.) Therefore t' represents the *elapsed* time between events 0 and B according to the moving observer. We assign event B the values t_B and x_B, and from (12) we have that $t' = t_B\sqrt{1 - v^2/c^2}$, or

$$t_B = t'/(1 - v^2/c^2)^{1/2}.$$

Further, from our perspective the astronaut moving at speed v has traveled a distance vt_B to arrive at event B. Therefore $x_B = vt_B$, or

$$x_B = vt'/(1 - v^2/c^2)^{1/2}.$$

Now the astronaut at event B knows that the (simultaneous) event E is located a distance x' from himself. We would measure the same distance at that moment to be D, however, and according to (13), $x' = D(\sqrt{1 - v^2/c^2})$, or

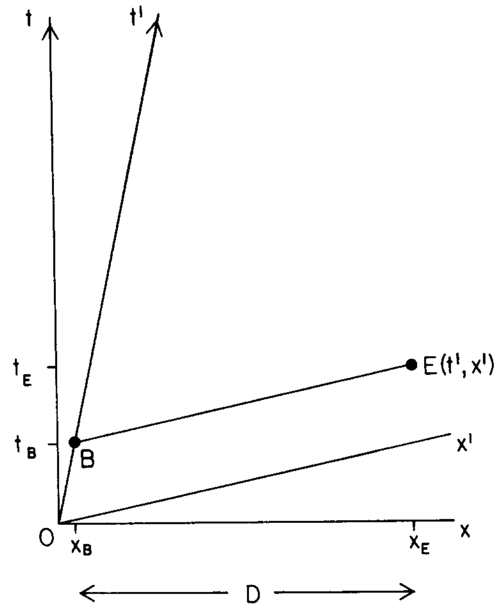

Fig. 16. Coordinates (t', x') of event E, assigned by a moving observer, may be related to (t_E, x_E) via consideration of the event B.

$$D = x'/(1 - v^2/c^2)^{1/2}.$$

The location of event E, according to us, is then $x = D + x_B$, or

$$x = \frac{x'}{(1 - v^2/c^2)^{1/2}} + \frac{vt'}{(1 - v^2/c^2)^{1/2}}. \quad \text{(A1)}$$

Recalling that the t' axis makes the same angle with the t axis as the x' axis makes with the x axis, we realize that to obtain the complementary expression for t in terms of t' and x', we need only replace t' in (A1) by x'/c (c sets the scale of our diagram) and $x(x')$ by $ct(ct')$.
Therefore,

$$ct = \frac{ct'}{(1 - v^2/c^2)^{1/2}} + \frac{vx'/c}{(1 - v^2/c^2)^{1/2}}. \quad \text{(A2)}$$

An alternative procedure is to use the relativity postulate. To the astronaut it is we who are in motion, with a speed $(-v)$. He would therefore apply to us the *same law* (A1) to write

$$x' = \frac{x}{(1 - v^2/c^2)^{1/2}} - \frac{vt}{(1 - v^2/c^2)^{1/2}}. \quad \text{(A3)}$$

The results (A1) and (A3) may then be used together to obtain (A2), as well as the result

$$ct' = \frac{ct}{(1 - v^2/c^2)^{1/2}} - \frac{vx/c}{(1 - v^2/c^2)^{1/2}}. \quad \text{(A4)}$$

[1]For a review of the different methods, see J. Rekveld, Am. J. Phys. **37**, 716 (1969).

[2]See, for instance, E. F. Taylor and J. A. Wheeler, *Spacetime Physics* (Freeman, San Francisco, 1966); A. P. French, *Special Relativity* (Nelson, London, 1968); N. D. Merrin, *Space and Time in Special Relativity* (McGraw-Hill, New York, 1969).

[3]H. van Dam, Ned. Tijdschr. Natuurkde. **37**, 431 (1971).

[4]H. Minkowski, Phys. Z. **10**, 104 (1909), translated in *The Principle of Relativity* (*A Collection of Papers*) (Methuen, London, 1923, reproduced by Dover, New York, 1952).

[5]In the United States accurate time signals are broadcast continuously over WWV at 2.5, 5, 10, 15, 20, and 25 Mc shortwave.

[6]We will depend, however, on space-time (t-x) maps. This suggests the question of whether these can be constructed independently. In fact, this can be accomplished by one observer equipped with a watch and a radar apparatus as is discussed in Sec. IV.

[7]To be more precise, light has zero rest mass. Its behavior is to be contrasted to that of an object which has a nonzero rest mass. The famous equation $E = mc^2$ states that mass may be converted into energy, or vice versa, but the behavior of the two modes is nonetheless quite different.

[8]Strictly speaking, the isotropy of space means that in three spacial dimensions, space appears the same in any direction from a given point.

[9]The actual process of synchronization of the watches of several observers who are at rest relative to one another is somewhat complicated, involving the reflection back and forth of light signals. See Taylor and Wheeler, Ref. 2. The details do not concern us here.

[10]The sophisticated reader will recognize that the length of the line segment as drawn is $[(\Delta t)^2 + D^2/c^2]^{1/2}$, which differs from Eq. (11) by the change in sign between the terms. This sign change means that while geometry in ordinary three-dimensional space is Euclidean, that of space-time is not. See Ref. 2.

[11]Compare the books mentioned in Ref. 2.

[12]It has often been said that nature cannot support a paradox. To be more precise, it is our theories of nature, which must provide a description of natural phenomena which cannot lead to paradoxes. Indeed, since a set of natural circumstances cannot give rise to contradictory outcomes, any theory which predicts contradictions must be abandoned.

[13]When accelerations are involved, the calculations require techniques which are beyond the level of our development. The result is that time passes even slower for a traveler during these maneuvers. For a calculation involving *only* special relativity, see C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), pp. 163–167.

[14]See p. 70 of Taylor and Wheeler, Ref. 2. For similar paradoxes, see W. Rindler, Am. J. Phys. **29**, 365 (1961); R. Shaw, *ibid.* **30**, 72 (1962).

[15]The question of what those who are waiting by the barn actually *see* is an interesting one also. The contraction of the pole would hardly be apparent, due to the natural foreshortening of a pole pointed almost directly at the observer as it approaches. The same is true of the runner, who is facing the observer. In fact, the contraction of a moving object along its direction of motion can never be *seen* by a single observer, even as the object passes him by, but must rather be *measured* via a set of events. See, for example, Taylor and Wheeler, Ref. 2.

[16]Note that $\Delta\tau_L$ and $\Delta\tau_R$ are *not* the lengths of the line segments drawn. However, $\Delta\tau_L$ and $\Delta\tau_R$ are certainly *proportional* to the lengths as drawn, and by the *same* factor for both segments, since M_L and M_R are traveling at the same speed relative to L and R.

[17]While $\Delta\tau_L$ and $\Delta\tau_R$ are not the lengths of the sides in question as drawn, they are proportional to the lengths of these line segments by the same factor. Therefore, they are in the same proportion.