Special Relativity and Maxwell's Equations

Richard E. Haskell

July 2003

Computer Science and Engineering Department
Oakland University
Rochester, MI 48309

Table of Contents

Preface

The material in this report was originally written as class notes in 1967 for a course I taught in electric and magnetic fields. The derivation of Maxwell's equations from special relativity and Coulomb's law was developed at that time in collaboration with Dr. Carl T. Case who was then at the Air Force Avionics Laboratory at Wright-Patterson Air Force Base. We had served in the Air Force together between 1963 and 1966 and had become intrigued with the possible limitations of Maxwell's equations based on this derivation. After 1970 I moved on to work in other areas including coherent optics, pattern recognition, microprocessors, and embedded systems. Last year I came across this material when cleaning out my office and decided to reprint it in electronic form and make it available on my web site.

If you are interested in understanding special relativity, then you should read Parts I – III. The derivation of Maxwell's equations from special relativity and Coulomb's law is given in Part IV. If you just want to find out why this topic is so intriguing then skip directly to the discussion in Section 19.

R. E. Haskell
July 2003

I. SPACE AND TIME

1. Reference Frames

We are familiar with using coordinate axes as reference frames to describe the motion of a particle. Thus, in Figure 1 we plot the displacement x of the particle as a function of the time t. We say that any point on this graph represents an event. Thus E_1 is an event which occurs at x_1 at the time t_1 and E_2 is an event which occurs at x_2 at the time t_2. The *world-line* of a particle is the locus of events in the space-time (x-t) graph of Figure 1. The velocity of the particle in Figure 1 is given by $v = \Delta x/\Delta t = \tan\phi$ and is the slope of the world-line.

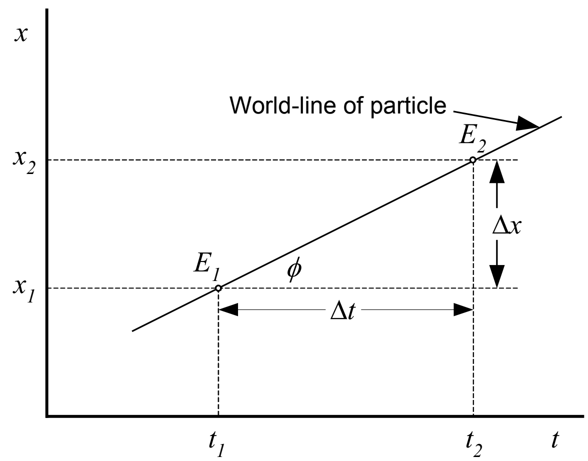

Fig. 1 Definition of a world-line

Now we may ask ourselves the following question: Given the event E_1, is the time t_1 found by dropping a perpendicular to the t-axis or is it found by moving parallel to the x-axis? In Figure 1 it is clear that these two operations are the same so the question may seem unimportant. However, there is no reason other than convenience that our coordinate axes x and t should be orthogonal or perpendicular to each other. For example, we could just as well draw them at an oblique angle as shown in Figures 2 and 3. We now see that we have a choice of how to define our components of the events E_1 and E_2. In Figure 2 we move parallel to the coordinate axes while in Figure 3 we drop perpendiculars to the coordinate axes. It is clear that either method is acceptable and that we can pick the one that is most convenient for any particular purpose. In Figures 2 and 3 the velocity of the particle is still given by $v = \Delta x/\Delta t$, but note that in neither case is this equal to $\tan\phi$ as it was in Figure 1. You may wish to find expressions for the

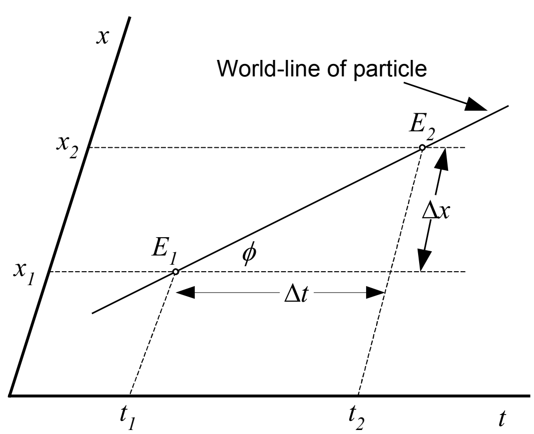

Fig. 2 Moving parallel to the axes

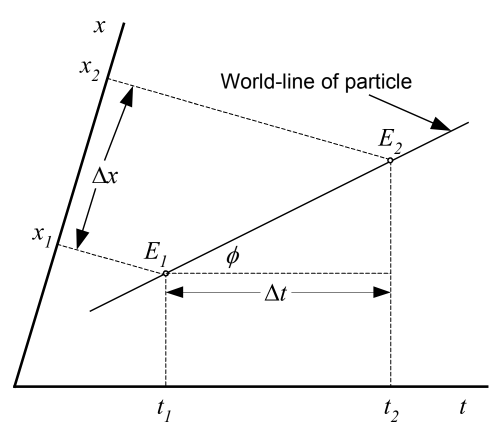

Fig. 3 Moving perpendicular to the axes

velocity of the particle in Figures 2 and 3 in terms ϕ and the angle that the x-axis makes with the t-axis.

2. The Galilean Transformation

We should like to be able to describe how a given event or a series of events appears to two different observers who move at a constant velocity relative to each other. For example, suppose a bird is flying past a moving train in the direction in which the train is moving. How does the motion of the bird appear to an observer on the train and how does it appear to an observer on the ground?

In order to answer this and similar questions we would like to be able to draw a set of space-time axes (x-t) for each of the observers (the train and the earth) in such a way that they could be used to describe a single event (the bird). Let us see how we might be able to do this.

In Figure 1 the world-line which is shown is that of a particle moving with a constant velocity relative to an observer at some fixed value of x. To make these ideas concrete let x-t be the space-time coordinate axes for an observer at rest with respect to the earth. The world-line of such an observer would be a horizontal line as shown in Figure 4 since as time goes on he simply remains at the same value of x, namely x_1.

Now we know from Figure 1 that the world-line of a train which moves relative to the earth with a velocity $v = \Delta x/\Delta t = \tan\phi$ will be a straight line inclined at an angle ϕ as shown in Figure 4. We want to draw a set of space-time axes (called $x^{'} - t$) for an observer at rest on the train. The coordinate axis $x^{'}$ must be such that each point on the world-line of the train has the same value of $x^{'}$. This will be the case if we draw $x^{'}$ as shown in Figure 4 where we have adopted the convention of Figure 3 and locate the coordinate of an event by dropping perpendiculars to the coordinate axes. Since $x^{'} - t$ are coordinates fixed to the train, as time goes on the observer on the train remains at his same value of $x^{'}$, namely $x_1^{'}$. The event E which occurs when the two observers meet is represented by the crossing of the two world-lines. Thus in Figure 4 the observer on the train is just passing by the observer on the ground at time t_1.

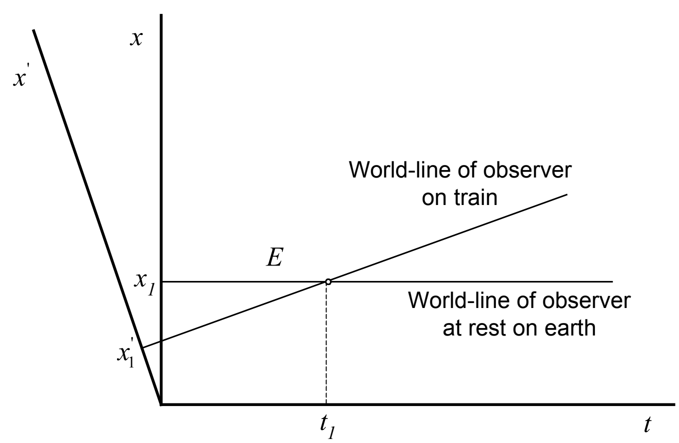

Fig. 4 World-lines of two observers

It is not yet clear how the scale of distance on the x-axis compares with the scale of distance on the $x^{'}$-axis. How can we find out? We need some definition of equivalent lengths. Let us denote the system of space coordinates fixed on the earth by S and those fixed on the train by $S^{'}$. Then let $F^{'}$ be the value of $x^{'}$ at the front of the train and let $B^{'}$ be the value of $x^{'}$ at the back of the train. Then $\Delta x^{'} = F^{'} - B^{'}$ is the length of the train as measured in $S^{'}$ as shown in Figure 5.

We know what we mean by the length of the train in $S^{'}$ since the person on the train can leisurely measure the length by stretching a steel tape from the back of the train to the front. But how can the observer on the ground measure the length of the train? One way might be to have two assistants, one stationed at the front of the train and the other stationed at the back of the train. Each assistant has an identical pistol and at precisely the same time they each fire a bullet straight down into the ground. The observer on the ground then leisurely walks over and measures the distance between the bullet holes in the ground with a steel tape. He then says that this distance is equal to the length of the train. Where is this length in Figure 5?

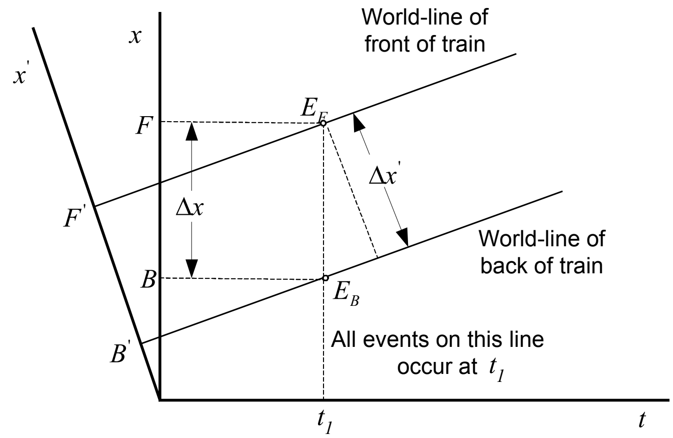

Fig. 5 World-lines of the front and back of a train

The important idea is that the bullets must be fired simultaneously. Suppose they are both fired at time t_1. Then in Figure 5 the event E_B = *firing the bullet at the back of the train* while the event E_F = *firing the bullet at the front of the train*. In the frame of reference fixed to the ground these events are separated by a distance $\Delta x = F - B$. This distance is what the observer on the ground calls the length of the train.

Now if the two observers are to measure the same length for the train (which seems reasonable) we must have $\Delta x = \Delta x^{'}$ in Figure 5. Since these lengths are clearly not equal in Figure 5 we must assign different scales of length to the x and $x^{'}$ axes. This would be cumbersome and we would rather not do it if we can help it. But there is another way out. Since we're finding our coordinates of events by dropping perpendiculars to the coordinate axes, we can make the length of Δx equal to the length of $\Delta x^{'}$ in Figure 5 by tilting the x-axis to the right the same amount as the $x^{'}$-axis is already tilted to the left as shown in Figure 6. It is a simple matter to show that $\Delta x = \Delta x^{'}$ in Figure 6. Thus in Figure 6 the scale lengths along the $x^{'}$ and x axes are the same.

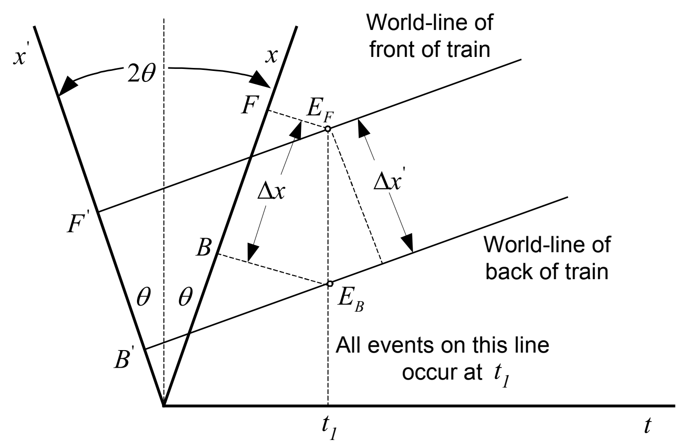

Fig. 6 Symmetric tilting of both x and $x^{'}$ axes

Now since we have tilted the x-axis in Figure 6 it is clear that the velocity of $S^{'}$ (the observer on the train) relative to S (the observer on the ground) will not be $\tan\phi$ as it was in Figure 4. In order to find out what this relative velocity is let's draw the world line of an observer at rest on the train as shown in Figure 7. Note that angle $x_1^{'}E_2x_2$ is equal to 2θ since $x_1^{'}E_2$ is perpendicular to $0x^{'}$ and E_2x_2 is perpendicular to $0x$. Similarly, angle E_2E_1A is equal to θ.

During the time interval $\Delta t = t_2 - t_1$ the event describing the observer at rest on the train moves from E_1 to E_2. In this time interval the position of the train as measured by the observer on the ground changes by an amount $\Delta x = x_2 - x_1$. Therefore, the velocity of the train ($S^{'}$) relative to the ground (S) is $u = \Delta x / \Delta t$. From Figure 7

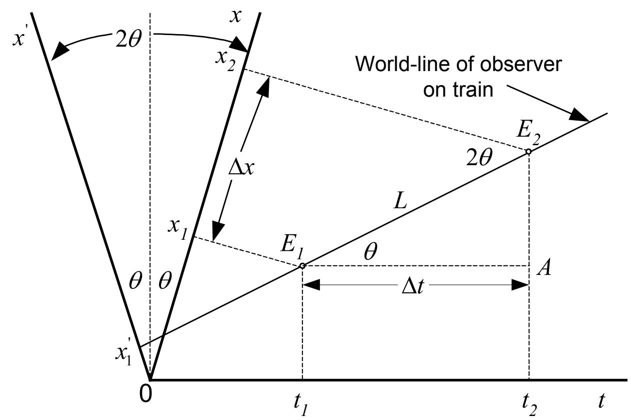

Fig. 7 Calculating the relative speed of the train

$$\Delta x = L \sin 2\theta$$
$$\Delta t = L \cos \theta$$

and since $\sin 2\theta = 2 \sin\theta \cos\theta$ we obtain for the relative velocity u

$$u = \frac{\Delta x}{\Delta t} = \frac{L \sin 2\theta}{L \cos\theta}$$
$$= \frac{2 \sin\theta \cos\theta}{\cos\theta}$$

or

$$u = 2 \sin\theta \tag{1}$$

Figures 6 and 7 show how given events appear to two different observers who are moving relative to each other with a constant velocity. The relationship between the coordinates of the two systems shown in these figures is called the *Galilean transformation*. As an example of this transformation we will now look at how velocities transform from one system to the other.

3. Velocity Transformation

Let us now consider the problem of a bird flying past the moving train. Since the bird is flying faster than the train its world line must have a positive slope relative to the world line of the moving train. The situation is shown in Figure 8. From Eq. (1) we know that

u = velocity of $S^{'}$ (train) relative to S (ground)
$u = 2\sin\theta$

Also let

$$v = \text{ velocity of bird relative to } S = \Delta x / \Delta t$$
$$v' = \text{ velocity of bird relative to } S' = \Delta x' / \Delta t$$

From Figure 8

$$\Delta x = L \sin(2\theta + \phi)$$
$$\Delta x' = L \sin\phi$$
$$\Delta t = L \cos(\theta + \phi)$$

We can therefore write

$$\begin{aligned}
\Delta x &= L \sin(2\theta + \phi) \\
&= L(\sin 2\theta \cos\phi + \cos 2\theta \sin\phi) \\
&= L\left[2\sin\theta\cos\theta\cos\phi + \left(1 - 2\sin^2\theta\right)\sin\phi\right] \\
&= L\sin\phi + 2\sin\theta L(\cos\theta\cos\phi - \sin\theta\sin\phi) \\
&= L\sin\phi + 2\sin\theta L\cos(\theta + \phi)
\end{aligned}$$

or

$$\Delta x = \Delta x' + u\Delta t$$

which is the equation for the Galilean transformation. It follows that

$$\frac{\Delta x}{\Delta t} = \frac{\Delta x'}{\Delta t} + u \tag{2}$$

or

$$v = v' + u \tag{3}$$

This is, of course, just what you would expect. The velocity of the bird relative to the ground is the velocity of the bird relative to the train plus the velocity of the train relative to the ground. Why all the fuss? You say you could have figured that out in your head much easier. And you're right. If this were the way nature works it certainly would not be worth all the effort to make all these graphs. But this is not the way nature works! We've made a mistake somewhere in our thinking.

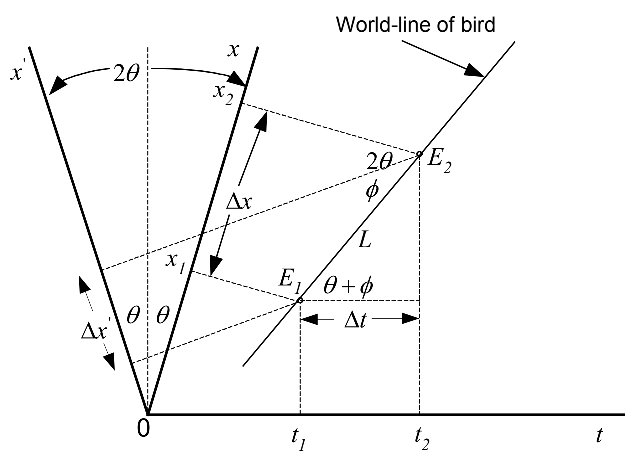

Fig. 8 Calculating the velocity of the bird

The Galilean transformation does not predict everything that we observe experimentally. What then is the proper picture of space and time? We will see in the next section that the graphs we have constructed for the Galilean transformation must be modified. However, our understanding of the use of these graphs will help us considerably in understanding this new theory of relativity. We will find that it is time that is the culprit and that the proper picture of space and time is one in which the asymmetry of Figures 6-8 disappears and space and time are put on an equal basis.

II. RELATIVISTIC KINEMATICS

4. The Principle of Relativity

Newton's laws of motion are stated as holding only with respect to an inertial frame of reference. Such a frame of reference is defined by Newton's first law. That is, an inertial frame of reference is one in which the law of inertia holds. It follows that there are an infinite number of systems of reference (inertial systems) moving uniformly and rectilinearly with respect to each other in which the law of motion ($\boldsymbol{F} = m\boldsymbol{a}$) is of the identical form.

The preceding is a statement of the *principle of relativity* for classical mechanics. It asserts that absolute uniform motion cannot be detected by any experiment of classical mechanics. For example, if you were riding along on a train moving with uniform velocity there is no mechanical experiment that you can do completely within the train that could tell you whether or not the train were moving. As far as you could tell the train might very well be at rest. This is certainly consistent with our everyday experience. Of course if the train accelerates or decelerates we can feel or detect this motion but as long as it is moving uniformly we have no way of detecting the motion.

Now this is a very satisfying principle of physics. It says that it is unnecessary to assume a special or specific frame of reference with respect to which the laws of mechanics are to hold. It says that any of an infinite number of (inertial) frames of reference are equally acceptable and that no one frame of reference is given a special position of importance above any other frame. If this weren't so then we would have to write a different equation of motion for every different frame of reference. It is clearly desirable to have the laws of physics be independent of the particular frame of reference from which the experiment is observed.

We see that for classical mechanics things seem to be in good shape with respect to the principle of relativity. What about the rest of physics? Toward the end of the last century it appeared as though the laws of electromagnetism violated the principle of relativity and as a result physicists thought that it would be possible to detect absolute uniform motion by certain experiments involving the propagation of light which is a form of electromagnetic energy. Many such experiments were carried out, the most famous of which is the Michelson-Morley experiment which was designed to detect the uniform motion of the earth relative to a hypothetical ether which was assumed to be at rest throughout absolute space. However, this experiment was unable to detect the absolute uniform motion of the earth.

There were many other examples of electromagnetic phenomena which were inconsistent with the concept of absolute rest. As a result in 1905 Einstein took as his first postulate that the principle of relativity that we stated above as holding for the laws of mechanics also holds for the laws of electrodynamics and therefore optics. We can therefore state Einstein's first postulate which is called the *Principle of Relativity* as follows: *Absolute uniform motion cannot be detected by any means*. This is to say that the concept of absolute rest and the ether have no meaning.

5. The Nature of Light

The Principle of Relativity described in the previous section does not seem to be a particularly disturbing postulate. As a matter of fact it seems quite reasonable. However, when Einstein stated this principle as his first postulate it did not seem so reasonable. This was because Maxwell's equations of electromagnetism predicted that light would travel with a constant velocity c. The question is -- a velocity c with respect to what? It was thus supposed that it must be with respect to an ether which was at absolute rest in the universe. It then followed from the Galilean transformation that absolute uniform motion with respect to the ether could be detected. As pointed out above all attempts to detect such motion have failed.

In addition to his first postulate of the Principle of Relativity, Einstein stated a second postulate concerning the nature of light. It was that *light is propagated in empty space with a velocity c which is independent of the motion of the source*.

We know that the velocity of some things do depend on the velocity of the source. For example, the velocity of a bullet will appear to travel faster to an observer on the ground if the gun is moving in the direction in which it is fired. On the other hand the velocity of sound does not depend on the velocity of the source but always has the same velocity with respect to the air. Physicists at the end of the last century thought light must act the same way. They believed that the velocity of light was independent of the velocity of the source. After all, the velocity of light must have the same value with respect to the ether. Thus Einstein's second postulate would seem quite natural.

However, we have seen that Einstein's first postulate implies that there is no ether. Thus at first sight there seems to be no way that both postulates can be true. Einstein showed that in order for both postulates to be true we must modify our ideas about the nature of time. Let us remind ourselves that the reason we accept the two postulates by themselves is that they agree with our experience. The combination of the two postulates leads to predictions which at first sight seem quite unlikely. However, many experiments have subsequently shown that these unlikely events do, in fact, occur.

6. The Nature of Time

In order to understand the dilemma of Einstein's two postulates consider the Galilean transformation represented by Figure 8. The velocity of the bird is different when viewed from two different inertial frames (the ground and the train). However, Einstein's postulates state that if we send a light signal from the back of the train to the front of the train then an observer on the train and an observer on the ground must both measure the velocity of the light beam to be c! How can we draw a world-line for the light beam in Figure 8 so this will be so? Only a vertical world-line would have the same velocity in both reference frames but this velocity would be infinite and not c. What else can we do?

The only way that Figure 8 can be modified so the velocity of light will have the same value c in both reference frames is to split the t-axis as shown in Figure 9. If we label the time axes in units of ct and ct' then the world-line of a light signal must have a slope of +1 in both frames of reference. We notice that by splitting the time axis as we have this is accomplished and in each reference frame the velocity of light is given by $\Delta x/\Delta t = \Delta x'/\Delta t' = c$.

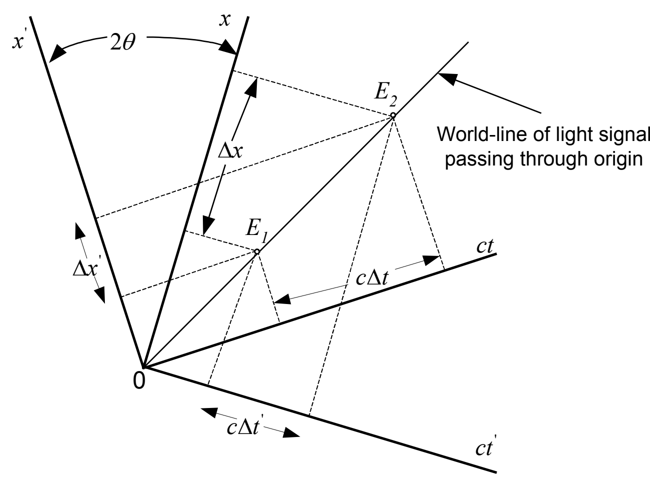

Fig. 9 Splitting the t-axis

This splitting of the time axis is the central result of relativity theory. We must try to understand what it means. Let us first determine what the relative velocity between the reference frames S and S' is in terms of our new picture of space-time given by Figure 9. A particle at rest in S' will have a zero velocity relative to S'. The world-line of such a particle is shown in Figure 10. The velocity of S' relative to S which we denote by u is then the velocity of the particle at rest in S' relative to S. From Figure 10 this velocity will be $u = \Delta x/\Delta t$. Note that we have let $2\theta = \alpha$ and also that the world-line of the particle at rest in S' is parallel to the ct-axis. This is because the axes x' and ct are perpendicular as are the axes x and ct'. From Figure 10 we therefore can write

$$\Delta x = L\sin\alpha$$
$$c\Delta t = L$$

so that

$$u = \frac{\Delta x}{\Delta t} = c\sin\alpha$$

or

$$\sin\alpha = u/c \tag{4}$$

Since $\sin^2\alpha + \cos^2\alpha = 1$ it follows that

$$\cos\alpha = \sqrt{1-\sin^2\alpha}$$
$$\cos\alpha = \sqrt{1-\frac{u^2}{c^2}} \tag{5}$$

We immediately see from Eq. (4) that u must be an appreciable fraction of the speed of light c in order for α to have a significantly large value. For example, if $u = 0.1c$, then $\alpha = 5.7°$. The angle α increases as u increases and approaches $90°$ as u approaches c.

In order to try and understand the nature of time as depicted by Figure 10 let us consider how two observers on a train, one at the front and one at the back, might synchronize their clocks so they will know they read the same time. One way might be for the two persons to meet somewhere on the train and set their clocks to the same time. Then one person moves to the back of the train and the other to the front of the train. Can you think of anything wrong with this method? How do you know that the beating of the clocks remained the same when they were in relative motion? The only direct comparisons of clocks that we can make are when they are at the same place at the same time. Probably the worst thing about the above method of synchronizing clocks from the physicists point of view is that there is no way of testing whether the clocks remained synchronized except by sending time signals once they are at the front and back of the train. This suggests that a better method of synchronizing clocks would be to use time signals to begin with. This could be done in the following way.

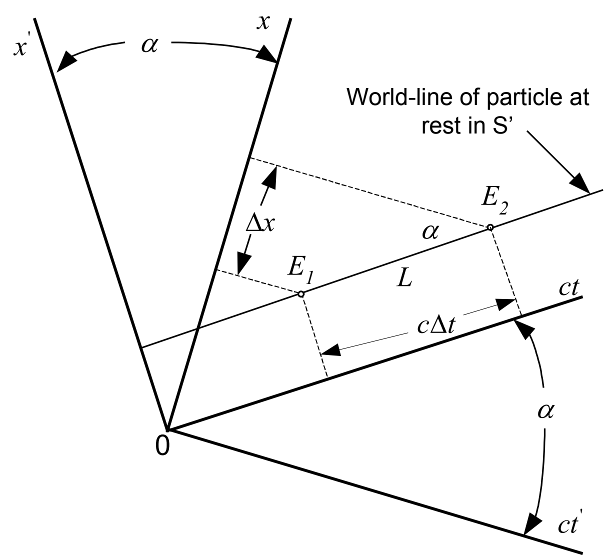

Fig. 10 Finding the relative velocity between S and S'

We measure the length of the train with a steel tape and then place a third person at exactly the center of the train. At a certain time that person explodes a flashbulb which sends a light signal in both directions at the constant velocity c. Each person at the front and back of the train has a clock which automatically starts when the light signal arrives. Now since they know (by Einstein's postulates) that the light will take the same time to travel to the front of the train as it does to travel to the back they are justified in saying that the signals will arrive at the front and back of the train simultaneously and that their clocks will be synchronized. Let us see what this looks like in our figures. Figure 11 shows the world-lines of the front (F'), back (B'), and center (C') of the train. The event E_1 is the exploding of the flashbulb at the center of the train at time t_1'. Event E_2 is the arrival of the light signal at the back of the train and event E_3 is the arrival of the light signal at the front of the train. As advertised these events occur simultaneously to the observer on the train at the time t_2'.

Now how does an observer on the ground in reference frame S describe what is going on? We see immediately from Figure 11 that event E_1 occurs at time t_1, event E_2 occurs at time t_2, and event E_3 occurs at time t_3, all of which are different. We therefore see that events E_2 and E_3 which were simultaneous on the train are not simultaneous when viewed from the ground. Why is this so? Remember that Einstein's postulates say that both observers must measure c for the velocity of light and that it is independent of the velocity of the source. Thus the observer on the ground must measure the light signal to travel at the velocity c in both directions. But since the train is moving with a velocity u with respect to the ground the light signal will clearly arrive at the back of the train, which is moving into the light signal, before it arrives at the front of the train, which is moving away from the light signal. This is exactly what an observer on the ground sees as is shown in Figure 11.

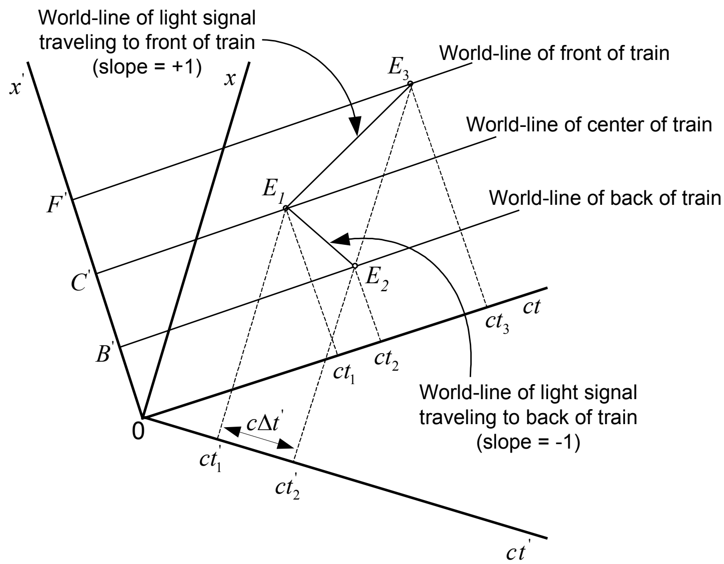

Fig. 11 Showing that simultaneity is relative

Both observers are equally correct in describing the events in Figure 11. All that it means is that the concept of simultaneity is a relative concept. It depends on your frame of reference. *There is no such thing as absolute simultaneity*. If this seems strange to you note that there is now a certain symmetry of space and time as shown in Figure 11. It is not strange to you that two events which occur at the same place to one observer don't occur at the same place to another observer. All the way back in Figure 7 the events E_1 and E_2 occur at the same place on the train but are separated by the distance Δx on the ground. The important thing is that they occur at different times. In a similar way the events E_2 and E_3 which occur at the same time on the train in Figure 11 are separated in time on the ground by the time interval $c\left(t_3 - t_2\right)$. Again the important thing is that these events now occur at different locations.

7. Time Dilation

Suppose the person at the back of the train in the caboose has a clock which he uses to keep time. As he passes a person *A* on the ground who has his or her own clock, a photograph is taken which shows both clocks. They happen to read the same value. Now down the track a certain distance away there is another person *B* who has a clock which he has previously synchronized with *A* by using time signals as explained in the previous section. As the caboose passes this second person another photo is taken which shows both the clock of person B and the clock in the caboose. The clock in the caboose reads less than the clock of person B. What has happened?

Let us follow this sequence of events on our diagrams. The world- line of the person in the caboose is shown in Figure 12. Event E_1 is the caboose passing person *A* and event E_2 is the caboose passing person *B*. The time between these two events is measured to be $\Delta t'$ to the person on the train, while it is measured to be Δt to the observers on the ground. From Figure 12 we see that

$$\Delta t' = \Delta t \cos\alpha$$

or, from Eq. (5)

$$\Delta t = \frac{\Delta t'}{\cos\alpha} = \frac{\Delta t'}{\sqrt{1-\frac{u^2}{c^2}}} \tag{6}$$

Therefore Δt is greater than $\Delta t'$. How can we understand this time dilation of the time interval between events E_1 and E_2 as observed by observers on the ground?

First let us review exactly what has happened. The clock in the primed system (i.e., the clock in the caboose) has moved from *A* to *B* and has recorded the time interval $\Delta t'$. This time interval is measured by one and the same clock. Such time which is measured by a single clock is called *proper time*. On the other hand it required two different clocks which were separated in space to measure the time interval Δt. This kind of time is called non-proper, or *coordinate time*. It is always the case that the shortest time interval is shown by the clock which measures proper time. Let us try to uncover the source of this dilation.

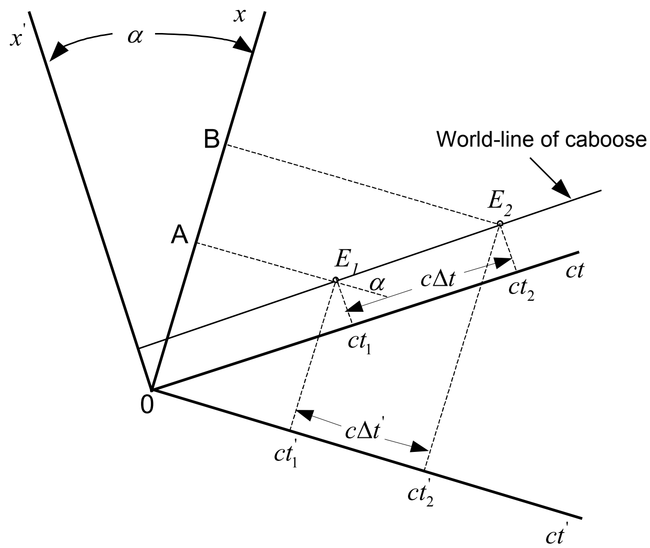

Fig. 12 World-line of caboose

In order to measure the time interval Δt on the ground we had to synchronize the clocks at *A* and *B*. We found previously that we could do this by exploding a flashbulb at a point *C* exactly half-way between *A* and *B*. If the clocks at *A* and *B* both start automatically just as the light signal reaches each one then we say that the clocks at *A* and *B* are synchronized. Let's suppose that the light signal reaches *A* (and therefore *B*) just as the caboose is passing *A*. This situation is shown in Figure 13. Again, the event E_1 is the caboose passing *A* while the event E_2 is the caboose passing *B* as in Figure 12. The event E_3 is the exploding of the flashbulb at *C*. The arrival of this signal at *B* is event E_3 while the arrival of this signal at *A* is E_1. Events E_1 and E_3 occur simultaneously at time

t_1 in the reference frame of the earth. However, as we already know these events are not simultaneous in the reference frame of the moving caboose. In this frame of reference the light signal arrives at *B* (at time $t_3^{'}$) before it arrives at *A* (at time $t_1^{'}$). Therefore, from the point of view of the caboose, *B* starts his or her clock before he or she "ought to", i.e., before the caboose reaches *A* at time *t′*. Therefore, it is not surprising that when *A* and *B* compare their clocks they get a longer time interval between events E_1 and E_2 than does the person on the caboose. However, it's not quite that simple. For if the person on the caboose tried to correct this "error" by starting his or her clock at $t_3^{'}$ instead of $t_1^{'}$ he or she would find that $t_2^{'} - t_3^{'}$ is *greater* than $\Delta t = t_2 - t_1$ by the same amount that Δt was previously greater than $\Delta t^{'} = t_2^{'} - t_1^{'}$. Why? Because Δt is now the proper time since it measures the time between events E_3 and E_2 with the single clock at *B*. However, the time $t_3^{'}$ could only be determined by someone on the train who happened to be passing *B* at the instant it was receiving the light signal from *C*. Perhaps the person *L* in the locomotive at the front of the train just passes *B* at E_3 as shown in Figure 13. This person *L* can note the time $t_3^{'}$ on his or her clock and later compare it with the person in the caboose who measures $t_2^{'}$ when passing *B*. But remember these two clocks must have previously been synchronized, but of course they will not appear synchronized to observers on the ground. Thus the time interval $t_2^{'} - t_3^{'}$ is an improper time interval since it requires two different clocks (one in the locomotive and one in the caboose) to measure. It is, therefore, longer than the corresponding proper time interval Δt.

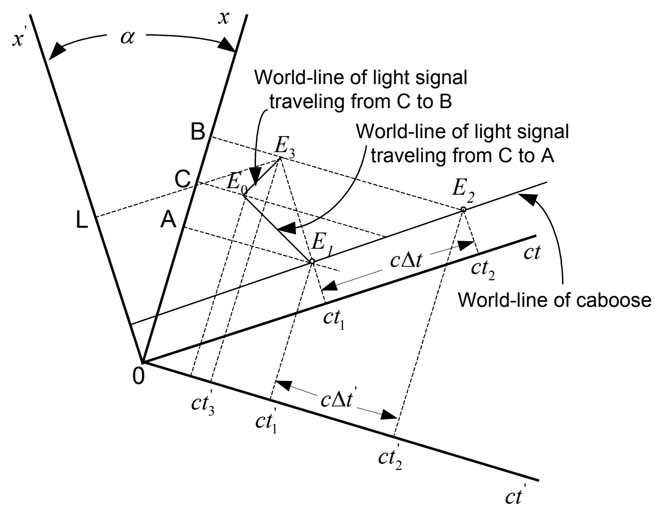

Fig. 13 Understanding time dilation

We therefore see that time dilation is a direct consequence of the fact that simultaneity is only a relative concept. This in turn is a direct consequence of the fact that the velocity of light measures the same in all inertial frames. As a result it is impossible to synchronize clocks which are in relative motion.

8. Length Contraction

Closely associated with the idea of time dilation is the phenomenon of length contraction. Suppose the people on the train measure the length of the train with a steel tape and find it to be $\Delta x^{'} = F^{'} - B^{'}$ as shown in Figure 14. How can observers on the ground determine the length of the train? You will recall in Section 2 that this length was measured by firing bullets simultaneously from the front and back of the train into the ground and then measuring the distance between the bullet holes. But that was before we knew that simultaneity is only relative. Events which appear to be simultaneous on the train will not be simultaneous to observers on the ground. What we generally mean when we talk about the length of the train as measured by observers on the ground is the distance between two observers on the ground one of whom is beside the front of the train and the second of whom is simultaneous beside the back of the train. This is, of course, simultaneous in the reference frame of the ground. We can therefore not use the method of firing bullets simultaneously from the train. What can we do?

We can have an observer on the ground who notes the time on his or her clock t_1 as the back of the train passes. We label this event E_1 and designate the location of this person by B. Now where is the front of the train at this same time t_1? We don't know ahead of time so we must station observers all along the rack each one of whom is a known distance from B and has a clock which has previously been synchronized with the clock at B. Then each one records the time at which the front of the train passes by. They later get together and compare notes. One of these many observers will have recorded the same time t_1 that B recorded when the back of the train passed by B.

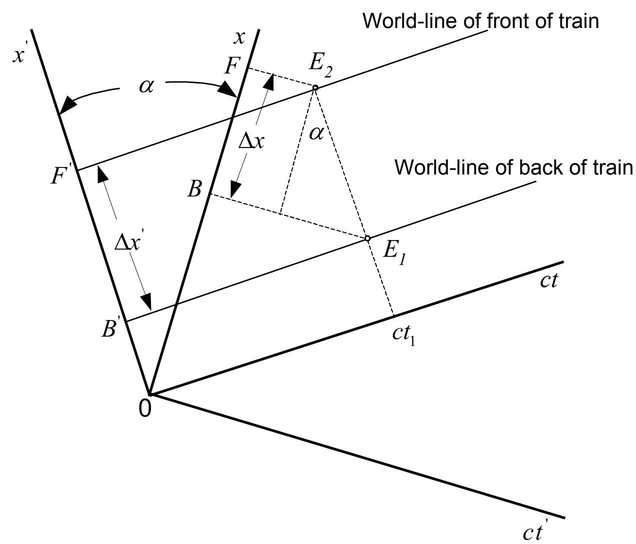

Fig. 14 Understanding length contraction

This observer was at the location F and we define the length of the train as measured on the ground to be $\Delta x = F - B$ as shown in Figure 14. From this figure we immediately see that

$$\Delta x = \Delta x^{'} \cos\alpha$$

$$\Delta x = \Delta x^{'}\sqrt{1-\frac{u^2}{c^2}} \tag{8}$$

so that the length of the train is shorter when measured by observers on the ground than it is when measured in the train. This result is clearly another consequence of relative simultaneity and is closely related to time dilation.

Now Δx and $\Delta x'$ are lengths measured in the direction of the relative motion *u*. What about lengths such as Δy and $\Delta y'$ measured perpendicular to the relative motion. It should be apparent that these lengths will be the same. In order to see this consider Figure 15 in which a light signal is sent out from point *C* just as *C'* (the center of the front of the train) passes *C*. If line *AB* is perpendicular to the direction of motion it is clear that *A* and *B* will receive the light signal simultaneously in the earth's reference frame. But it is also clear that A' and B' on the train will also receive the light signal simultaneously. Thus, if A' and B' fire bullets into the ground simultaneously (in their frame) this will also be simultaneous in the earth's frame so that if *A* and *B* go over and measure the distance between the holes in the ground they will measure the same width *w* of the train as A' and B' measure on the train. Thus if the relative motion *u* is in the *x*-direction then $\Delta y' = \Delta y$ and $\Delta z' = \Delta z$.

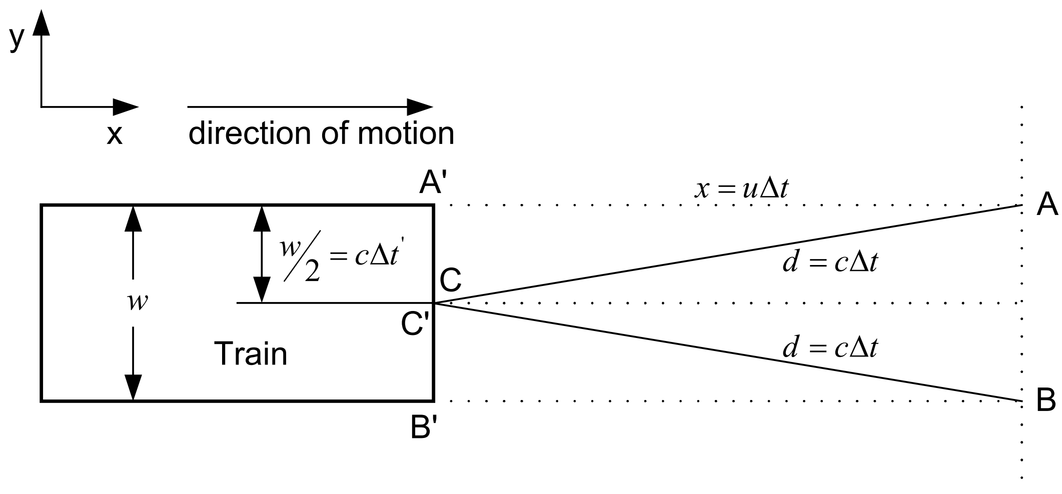

Fig. 15 Lengths perpendicular to the direction of motion remain unchanged

Of course time intervals are still measured differently in the two reference frames. Thus, although *A* and *B* receives the light signals simultaneously and so do A' and B', the two sets of observers differ on how long they say it takes the light to reach them from the source since both must measure the velocity of light to be *c*. From Figure 15 we see that to observers on the ground

$$\frac{w}{2} = \sqrt{d^2 - x^2} = \Delta t\sqrt{c^2 - u^2} = c\Delta t\sqrt{1 - \frac{u^2}{c^2}}$$

while on the train $\frac{w}{2} = c\Delta t'$. Therefore $\Delta t' = \Delta t\sqrt{1 - \frac{u^2}{c^2}}$ just as was found in Section 7.

9. The Lorentz Transformation

In Sections 2 and 3 we obtained diagrams which represented the so-called Galilean transformation. This transformation is shown in Figure 8 for the case of a bird flying past a moving train and is given by Equation (3). However, we have now seen that Figure 8 is really not an accurate picture of space and time but must be replaced by Figure 16. From this figure we can write

$$\Delta x = L\sin(\phi+\alpha) \qquad (9)$$

$$\Delta x' = L\sin\phi \qquad (10)$$

$$c\Delta t = L\cos\phi \qquad (11)$$

$$c\Delta t' = L\cos(\phi+\alpha) \qquad (12)$$

From Eqs. (5) and (6) recall that

$$\sin\alpha = \frac{u}{c} \qquad (13)$$

and let

$$\gamma = \frac{1}{\cos\alpha} = \frac{1}{\sqrt{1-\dfrac{u^2}{c^2}}} \qquad (14)$$

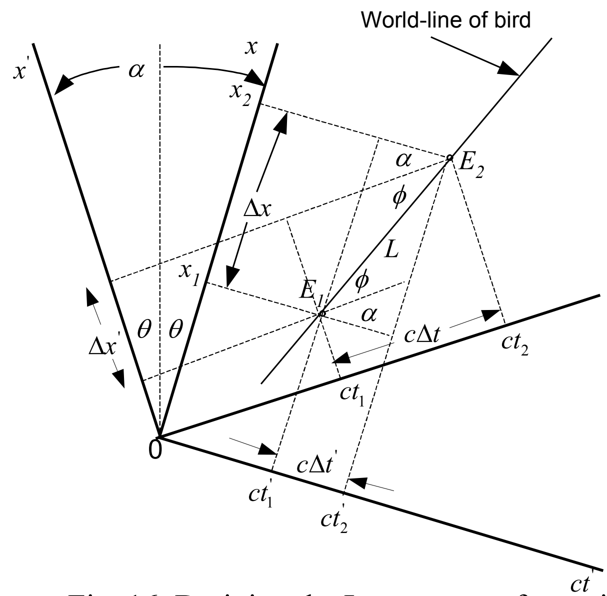

Fig. 16 Deriving the Lorentz transformation

Then from Eqs. (9), (10), (11) and (14) it follows that

$$\begin{aligned}\Delta x &= L(\sin\phi\cos\alpha + \cos\phi\sin\alpha)\\ &= \Delta x'\cos\alpha + c\Delta t\sin\alpha\\ &= \frac{\Delta x'}{\gamma} + u\Delta t\end{aligned}$$

from which

$$\Delta x' = \gamma(\Delta x - u\Delta t) \qquad (15)$$

Also, from Eqs. (12), (11), (10), (14) and (13) it follows that

$$\begin{aligned}c\Delta t' &= L(\cos\phi\cos\alpha - \sin\phi\sin\alpha)\\ &= c\Delta t\cos\alpha - \Delta x'\sin\alpha\\ &= \frac{c\Delta t}{\gamma} - \frac{u}{c}\Delta x'\end{aligned}$$

from which

$$\Delta t = \gamma\left(\Delta t^{'} + \frac{u}{c^2}\Delta x^{'}\right) \tag{16}$$

Since there is complete symmetry between the observation in the two frames of reference except that the sign of the relative velocity u changes we can obtain Δx in terms of $\Delta x^{'}$ and $\Delta t^{'}$ by simply interchanging the primed and unprimed quantities in Eq. (15) and changing the sign of u. Thus

$$\Delta x = \gamma\left(\Delta x^{'} + u\Delta t^{'}\right) \tag{17}$$

This equation can be verified by substituting Eq. (16) in (15) and solving for Δx.

In a similar manner from Eq. (16) we can immediately write

$$\Delta t^{'} = \gamma\left(\Delta t - \frac{u}{c^2}\Delta x\right) \tag{18}$$

which can also be verified by substituting Eq. (15) in (16) and solving for $\Delta t^{'}$. Eqns. (15) - (18) are called the Lorentz transformation which can be summarized as follows:

$$\boxed{\begin{aligned} \Delta x &= \gamma\left(\Delta x^{'} + u\Delta t^{'}\right) \\ \Delta t &= \gamma\left(\Delta t^{'} + \frac{u}{c^2}\Delta x^{'}\right) \\ \Delta x^{'} &= \gamma\left(\Delta x - u\Delta t\right) \\ \Delta t^{'} &= \gamma\left(\Delta t - \frac{u}{c^2}\Delta x\right) \end{aligned}} \tag{19}$$

10. Relativistic Velocity Transformation

In Section 3 we found that the Galilean transformation gave rise to a velocity transformation of the form (see Eq. 4)

$$v = v' + u \tag{20}$$

However, we see from Eq. (19) that the Lorentz transformation will produce a different velocity transformation law. In particular

$$\frac{\Delta x}{\Delta t} = \frac{\gamma\left(\Delta x' + u\Delta t'\right)}{\gamma\left(\Delta t' + \frac{u}{c^2}\Delta x'\right)}$$

If we divide numerator and denominator by $\Delta t'$ we obtain

$$\frac{\Delta x}{\Delta t} = \frac{\left(\Delta x'/\Delta t' + u\right)}{\left(1 + \frac{u\Delta x'}{c^2\Delta t'}\right)}$$

or

$$v = \frac{v' + u}{1 + \frac{uv'}{c^2}} \tag{21}$$

Similarly, from Eq. (19)

$$\frac{\Delta x'}{\Delta t'} = \frac{\gamma\left(\Delta x - u\Delta t\right)}{\gamma\left(\Delta t - \frac{u}{c^2}\Delta x\right)} = \frac{\Delta x/\Delta t - u}{1 - \frac{u\Delta x}{c^2\Delta t}}$$

or

$$v' = \frac{v - u}{1 - \frac{uv}{c^2}} \tag{22}$$

Eqs. (21) and (22) are the velocity transformation laws for the Lorentz transformation. The first thing to notice is that they reduce to the Galilean transformation law, Eq. (20), for low relative velocities, $u << c$.

Let's suppose the train has a speed relative to the earth of $u = 0.6c$, a very high speed! Suppose also that a bird flies past the train with a velocity relative to the train of $v' = 0.8c$, a very fast bird. How fast is the bird flying relative to the ground? Our old

Galilean transformation Eq. (18) would tell us that $v = 0.8c + 0.6c = 1.4c$, faster than the speed of light. However, the Lorentz transformation predicts by Eq. (21) that

$$v = \frac{0.8c + 0.6c}{1 + \frac{(0.6c)(0.8c)}{c^2}} = \frac{1.4c}{1.48} = 0.946c$$

which is still less than the speed of light.

Let the train be moving with a velocity $u = kc$ (k<1) relative to the earth. A light signal which travels in the train frame of reference with a velocity $v' = c$ will by Eq. (21) travel in the earth frame of reference with a velocity

$$v = \frac{c + kc}{1 + \frac{kc^2}{c^2}} = \frac{1+k}{1+k}c = c$$

as it should by Einstein's postulates.

Now the preceding examples show that Eqs. (21) and (22) predict that if an object is traveling less than the speed of light in one reference frame it will travel less than the speed of light in all other frames which move relative to the first with velocities less than c. But is it possible for objects to travel faster than the speed of light in the first place? Let us consider what this would have to mean by referring to Figure 17. Suppose in the earth's reference frame a flashbulb is exploded at point A at time t_1. This is event E_1. Event E_2 is the arrival of the light signal at point B at time t_2. An observer on a fast moving train would say that event E_1 occurred at time t_1' and event E_2 occurred at time t_2'. Now suppose that at the instant the flashbulb exploded a gun was fired which was able to fire a bullet at twice the speed of light in the earth's frame of reference. It would, therefore, arrive at point B at the time t_3. This is event E_3 in Figure 17. But to the observer on the moving train this event occurs at time t_3'

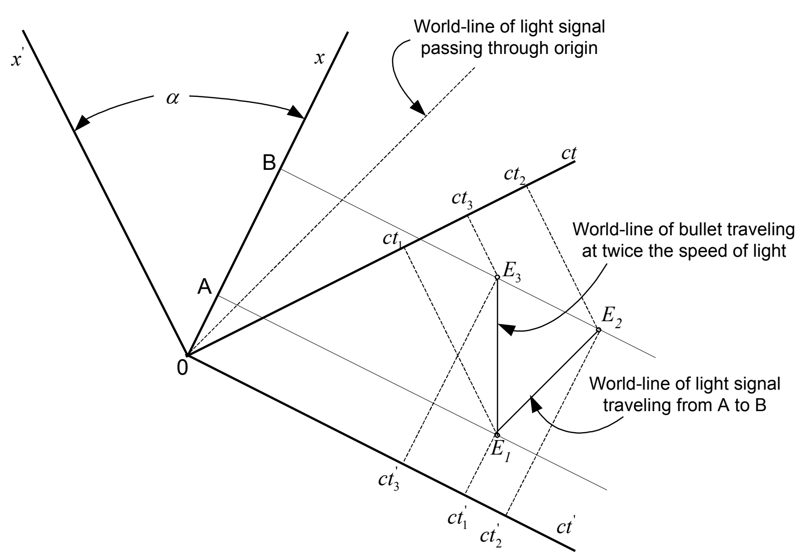

Fig. 17 A bullet traveling at twice the speed of light

which is less than $t_1^{'}$. That is, to the observer on the train the bullet arrives at *B* before it was fired at *A*! We might be able to accept this if there were no way for observers in the primed system to prevent the shooting. But if they can have bullets that travel faster than *c* in their frame they can prevent the shooting. To see this, consider Figure 18. At time t_1 *A* fires a bullet (event E_1) at twice the speed of light and kills *B* at time t_3 (event E_3). Observer $B^{'}$ who is just passing by at time $t_3^{'}$ observes the killing and scoops up the dead body of B. Observer $B^{'}$ immediately sends a message on a bullet traveling faster than *c* to $A^{'}$ telling of the shooting. This message arrives at $A^{'}$ at time $t_4^{'}$ (event E_4). However, at this time $A^{'}$ is just passing *A* and since the time in the reference frame of *A* is t_4 which is less than t_1, then *A* hasn't yet fired the shot, so $A^{'}$ disarms *A*. But $B^{'}$ has *B*'s dead body! In order to avoid this kind of serious contradiction we must conclude that the bullets, or anything else, cannot travel faster than the speed of light in any reference frame.

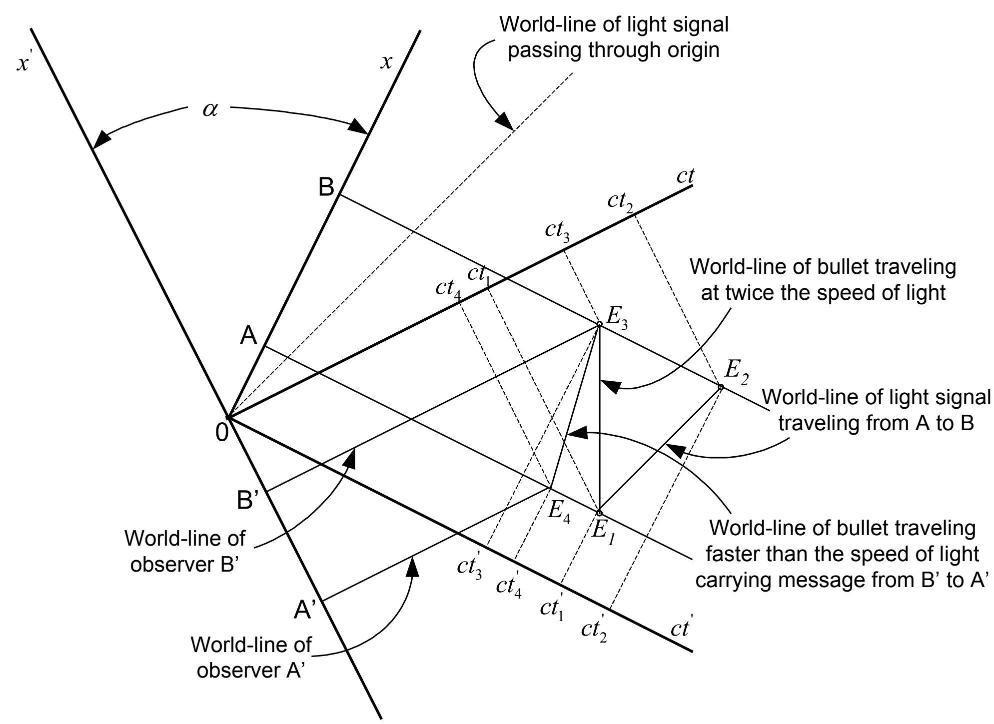

Fig. 18 Illustrating logical contradiction when speed of light is exceeded

11. Vector Representation of the Lorentz Transformation

The Lorentz transformation was found in Section 9 to be given by Eq. (19) when the relative velocity u was in the x direction. For this case we also found that $\Delta y' = \Delta y$ and $\Delta z' = \Delta z$. If the origins of the two coordinate systems coincide at $t = t' = 0$ then we can write the Lorentz transformation as

$$\boxed{\begin{aligned} x' &= \gamma(x - ut) \\ y' &= y \\ z' &= z \\ t' &= \gamma\left(t - \frac{ux}{c^2}\right) \end{aligned}} \tag{23}$$

$$\boxed{\begin{aligned} x &= \gamma(x' + ut') \\ t &= \gamma\left(t' + \frac{ux'}{c^2}\right) \end{aligned}} \tag{24}$$

where

$$\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}$$

Let us consider the general case in which the relative motion can be in an arbitrary direction as shown in Figure 19. The speed of S' relative to S is u and α_i is a unit vector in the direction of motion ($\alpha_i \alpha_i = 1$).[1]

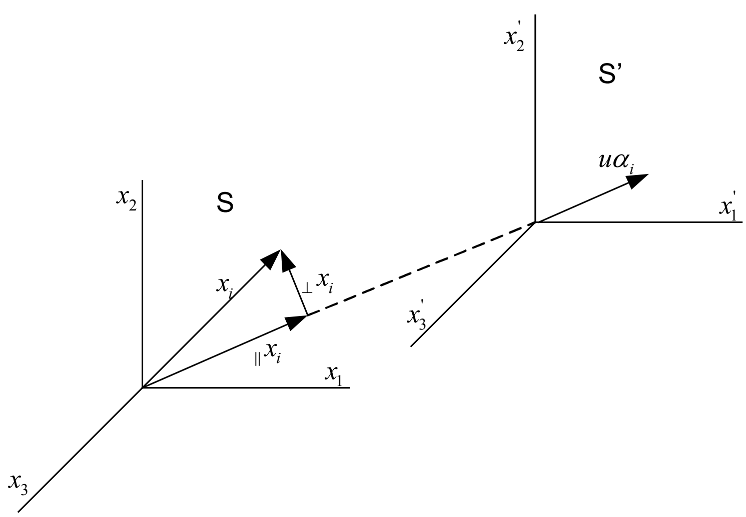

Fig. 19 Reference frames for arbitrary direction of motion

---

[1] α_i (i =1,3) are the three components of the vector $\boldsymbol{\alpha}$. Repeated indices are summed from 1 to 3; i.e., $\alpha_i \alpha_i = \alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1$

Let

$$\beta_i = \alpha_i \frac{u}{c}$$

then

$$\beta^2 = \beta_i \beta_i = \frac{u^2}{c^2}$$

and

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}$$

If $x_i =$ the displacement vector in S

$x_i' =$ the displacement vector in S'

${}_{\parallel}x_i =$ the component of x_i parallel to the direction of motion

${}_{\perp}x_i =$ the component of x_i perpendicular to the direction of motion

then

$$ {}_{\parallel}x_i = x_j \alpha_j \alpha_i \tag{25}$$

$$ {}_{\perp}x_i = x_i - {}_{\parallel}x_i \tag{26}$$

and similarly for ${}_{\parallel}x_i'$ and ${}_{\perp}x_i'$.

The Lorentz transformation given by Eq. (23) can then be written as

$$\boxed{\begin{aligned} {}_{\parallel}x_i' &= \gamma\left({}_{\parallel}x_i - u\alpha_i t\right) \\ {}_{\perp}x_i' &= {}_{\perp}x_i \\ t' &= \gamma\left(t - \frac{u}{c^2}\alpha_j x_j\right) \end{aligned}} \tag{27}$$

We can therefore write

$$\begin{aligned} x_i' &= {}_{\parallel}x_i' + {}_{\perp}x_i' \\ &= \gamma\left({}_{\parallel}x_i - u\alpha_i t\right) + {}_{\perp}x_i \\ &= \gamma\, {}_{\parallel}x_i - \gamma u \alpha_i t + x_i - {}_{\parallel}x_i \\ &= (\gamma - 1) x_j \alpha_j \alpha_i - \gamma u \alpha_i t + x_i \\ &= \left[\delta_{ij} + (\gamma - 1)\alpha_i \alpha_j\right] x_j - \gamma u \alpha_i t \end{aligned}$$

In the last step δ_{ij} is the Kronecker delta that is equal to 1 when $i = j$ and is equal to 0 when $i \neq j$; therefore, $\delta_{ij} x_j = \delta_{i1} x_1 + \delta_{i2} x_2 + \delta_{i3} x_3 = x_i$. Thus, the Lorentz transformation can be written in the vector form

$$\boxed{\begin{aligned} x_i' &= \left[\delta_{ij} + (\gamma - 1)\alpha_i \alpha_j\right] x_j - \gamma u \alpha_i t \\ x_i &= \left[\delta_{ij} + (\gamma - 1)\alpha_i \alpha_j\right] x_j' + \gamma u \alpha_i t' \\ t' &= \gamma\left(t - \frac{u}{c^2}\alpha_j x_j\right) \\ t &= \gamma\left(t' + \frac{u}{c^2}\alpha_j x_j'\right) \end{aligned}} \tag{28}$$

We can find the general velocity transformation by noting that

$$\begin{aligned} dx_i' &= \left[\delta_{ij} + (\gamma - 1)\alpha_i \alpha_j\right] dx_j - \gamma u \alpha_i dt \\ dt' &= \gamma\left(dt - \frac{u}{c^2}\alpha_j dx_j\right) \end{aligned}$$

Therefore

$$v_i' = \frac{dx_i'}{dt'} = \frac{\left[\delta_{ij} + (\gamma - 1)\alpha_i \alpha_j\right] v_j - \gamma u \alpha_i}{\gamma\left(1 - \frac{u}{c^2}\alpha_j v_j\right)} \tag{29}$$

and

$$v_i = \frac{dx_i}{dt} = \frac{\left[\delta_{ij} + (\gamma - 1)\alpha_i \alpha_j\right] v_j' + \gamma u \alpha_i}{\gamma\left(1 + \frac{u}{c^2}\alpha_j v_j'\right)} \tag{30}$$

Note that if u is in the x_1 direction; i.e., $\alpha_i = (1, 0, 0)$ then (29) reduces to

$$v_1' = \frac{\left[1 + (\gamma - 1)\right] v_1 - \gamma u}{\gamma\left(1 - \frac{u v_1}{c^2}\right)} = \frac{v_1 - u}{1 - \frac{u v_1}{c^2}} \tag{31}$$

which agrees with Eq. (22). In addition

$$v_2' = \frac{v_2}{\gamma\left(1 - \frac{u v_1}{c^2}\right)} = \frac{v_2\sqrt{1 - u^2/c^2}}{1 - \frac{u v_1}{c^2}} \tag{32}$$

III. RELATIVISTIC DYNAMICS

12. Relativistic Momentum

We have seen that the principle of relativity demands that the equations of physics have the same form in all inertial frames of reference. We must therefore re-examine the laws of dynamics in terms of the relativistic kinematics developed in the preceding sections. For low velocities these laws must reduce to Newton's three laws of motion.

The first law essentially defines an inertial frame of reference. That is, it is one in which, in the absence of all external influence (i.e. forces), a body will remain in a state of rest or uniform motion in a straight line. The third law is a statement about forces, namely, that if body A acts on body B with a force $\mathbf{F}_{BA}$ then body B acts on body A with an equal and opposite force $\mathbf{F}_{AB} = -\mathbf{F}_{BA}$. Let us accept these laws as holding in the relativistic case.

Newton's second law of motion states that in an inertial frame of reference as defined by the first law the net force acting on a body is equal to the rate of change of momentum $\mathbf{p}$ where $\mathbf{p} = m\mathbf{v}$, the product of the mass times the velocity of the body. Now since this law involves quantities like velocity and time which we have seen have special relativistic transformation laws we must examine this second law with care.

Let us assume that for the relativistic case we can write the second law in the form $\mathbf{F} = d\mathbf{p}/dt$ and try to determine what $\mathbf{p}$ must be in order to be consistent with the Lorentz transformation.

Consider two isolated bodies A and B which are only interacting with each other. If $\mathbf{F}_{BA}$ is the force acting on A due to B and $\mathbf{F}_{BA}$ is the force acting on B due to A then by Newton's third law and our assumption about the form of the second law we can write

$$\mathbf{F}_{AB} = \frac{d\mathbf{p}_A}{dt} = -\mathbf{F}_{BA} = -\frac{d\mathbf{p}_B}{dt}$$

or

$$\frac{d}{dt}\left(\mathbf{p}_A + \mathbf{p}_B\right) = 0$$

so that the total momentum of the two bodies is conserved. Since it must be conserved in all coordinate systems we can write

$$\frac{d}{dt}\left(\mathbf{p}_A + \mathbf{p}_B\right) = \frac{d}{dt'}\left(\mathbf{p}'_A + \mathbf{p}'_B\right) \tag{33}$$

where the primed system is any inertial frame of reference.

Now let us suppose that body A is at rest (at least instantaneously) in the unprimed system and that the primed system is moving in the $+x_1$ direction with a speed u. That is, $\alpha_i = (1,0,0)$. Further let the velocity of body B be $(v_i)_B = (v_B, 0, 0)$ in the unprimed system and be $(v_i^{'}) = (v_B^{'}, 0, 0)$ in the primed system. From (31) these velocities are then related by the expression

$$v_B^{'} = \frac{v_B - u}{1 - \frac{uv_B}{c^2}} \tag{34}$$

Also, since $(v_i)_A = 0$ then $(v_i)_A = (v_A^{'}, 0, 0)$ where $v_A^{'} = -u$.

We are looking for an expression for $\mathbf{p}$ which will be some function of the velocity of a body and will reduce to $m\mathbf{v}$ for low velocities. We will therefore assume that $\mathbf{p}$ is in the direction of $\mathbf{v}$ and so for the one dimensional case being considered $\mathbf{p} = (p, 0, 0)$ and (33) can be written as

$$\frac{d}{dt}\left[p(v_A) + p(v_B)\right] = \frac{d}{dt^{'}}\left[p(v_A^{'}) + p(v_B^{'})\right] \tag{35}$$

Since $v_A = 0$ it follows that $p(v_A) = 0$ since for low velocities $p(v_A) = m_A v_A$. Also

$$\frac{dp(v_A^{'})}{dt} = \frac{dp(v_A^{'})}{dv_A^{'}}\frac{dv_A^{'}}{dt^{'}} = 0$$

since $v_A^{'} = -u =$ constant. Eq. (35) then reduces to

$$\frac{dp(v_B)}{dt} = \frac{dp(v_B^{'})}{dt^{'}}$$

or

$$\frac{dp(v_B)}{dv_B}\frac{dv_B}{dt} = \frac{dp(v_B^{'})}{dv_B^{'}}\frac{dv_B^{'}}{dt}\frac{dt}{dt^{'}} \tag{36}$$

From (34)

$$\frac{dv_B^{'}}{dt} = \frac{\left(1 - \frac{uv_B}{c^2}\right)\frac{dv_B}{dt} + \left(v_B - u\right)\frac{u}{c^2}\frac{dv_B}{dt}}{\left(1 - \frac{uv_B}{c^2}\right)^2}$$

$$\frac{dv_B^{'}}{dt} = \frac{1}{\gamma^2\left(1 - \frac{uv_B}{c^2}\right)^2}\frac{dv_B}{dt} \tag{37}$$

where $\gamma = \sqrt{1 - u^2/c^2}$. Since $t^{'} = \gamma\left(t - \frac{u}{c^2}x_B\right)$ then

$$\frac{dt'}{dt} = \gamma\left(1 - \frac{uv_B}{c^2}\right) \tag{38}$$

Substituting (37) and (38) into (36) we obtain

$$\frac{dp(v_B)}{dv_B} = \frac{1}{\gamma^3\left(1 - \frac{uv_B}{c^2}\right)^3}\frac{dp(v_B')}{dv_B'} \tag{39}$$

Since (39) must hold for any v_B, let $v_B = v = u$ so that $v_B' = 0$. B Then (39) can be written as

$$\frac{dp(v_B)}{dv_B} = \frac{K_0}{\left(1 - \frac{v^2}{c^2}\right)^{3/2}} \tag{40}$$

where

$$K_0 = \left.\frac{dp(v_B')}{dv_B'}\right|_{v_B'=0}$$

Integrating (40) we obtain

$$p(v) = K_0 \frac{v}{\sqrt{1 - v^2/c^2}} + const$$

Since $p(0) = 0$ and $p(v) = mv$ for $v/c << 1$ the constant of integration is zero and $K_0 = m$. Therefore, the relativistic expression for momentum is

$$p(v) = \frac{mv}{\sqrt{1 - v^2/c^2}} \tag{41}$$

For arbitrary directions of motion the relativistic momentum can be written in the vector form

$$p_i = \frac{mv_i}{\sqrt{1 - v^2/c^2}} \tag{42}$$

where $v^2 = v_i v_i$.

The expression for the relativistic momentum is often written in the non-relativistic form

$$p_i = m_R v_i$$

where

$$m_R = \frac{m}{\sqrt{1 - v^2/c^2}}$$

is called the relativistic mass (and m the rest mass) and increases with velocity. This introduction of a relativistic mass that increases with velocity in order to force the expression for momentum to be $m_R v_i$ is sometimes useful but not essential. It will not be used in the following sections so the mass m which appears will always be the rest mass.

13. Relativistic Energy

In the previous section we found that the force acting on a particle is given by $F_i = dp_i/dt$ where the momentum p_i is given by

$$p_i = \frac{mv_i}{\sqrt{1 - v^2/c^2}} \tag{43}$$

The work dW done by the force F_i in moving the particle through a distance dx_i is by definition

$$\begin{aligned} dW &= F_i dx_i \\ &= \frac{dp_i}{dt} dx_i \\ &= dp_i \frac{dx_i}{dt} \\ &= v_i dp_i \end{aligned} \tag{44}$$

From (43)

$$\begin{aligned} dp_i &= \frac{\left(1 - v^2/c^2\right)^{1/2} m dv_i + m v_i \frac{1}{2}\left(1 - v^2/c^2\right)^{-1/2} \frac{2v}{c^2} dv}{\left(1 - v^2/c^2\right)} \\ &= \frac{m}{\left(1 - v^2/c^2\right)^{1/2}} dv_i + \frac{m v v_i}{c^2 \left(1 - v^2/c^2\right)^{3/2}} dv \end{aligned} \tag{45}$$

Since $v^2 = v_i v_i$, $vdv = v_i dv_i$ and from (45) we can write

$$\begin{aligned} v_i dp_i &= \frac{mvdv}{\left(1 - v^2/c^2\right)^{3/2}} \left(1 - \frac{v^2}{c^2} + \frac{v^2}{c^2}\right) \\ &= \frac{mvdv}{\left(1 - v^2/c^2\right)^{3/2}} = dW \end{aligned} \tag{46}$$

Let us define the kinetic energy T of the particle, as in classical mechanics, as the work done in bringing the particle from rest to a velocity v. Then from (46) we have that

$$\begin{aligned} T &= \int_0^v dw \\ &= \int_0^v \frac{mvdv}{\left(1-v^2/c^2\right)^{3/2}} \\ &= \left.\frac{mc^2}{\sqrt{1-v^2/c^2}}\right|_0^v \end{aligned}$$

$$T = mc^2\left(\frac{1}{\sqrt{1-v^2/c^2}}-1\right) \tag{47}$$

Note that for $v/c << 1$ this equation reduces to

$$T = mc^2\left(1+\frac{v^2}{2c^2}+\frac{3v^4}{8c^4}+\ldots-1\right)$$

$$T = \frac{1}{2}mv^2$$

as in classical mechanics.

Now the conservation of momentum discussed in Section 12 has another important consequence. Consider the inelastic collision shown in Figure 20. In the unprimed system two equal masses approach each other with velocities v and $-v$. They collide and stick together so that after the collision the total mass M is at rest. (This is required because of the symmetry of the situation). The statement of the conservation of momentum for this case is thus

$$\frac{mv}{\sqrt{1-v^2/c^2}}-\frac{mv}{\sqrt{1-v^2/c^2}}=0$$

Now let us view this same inelastic collision from a reference frame moving to the right with a velocity $u = v$. Before the collision one mass is at rest in this frame while the other has a velocity

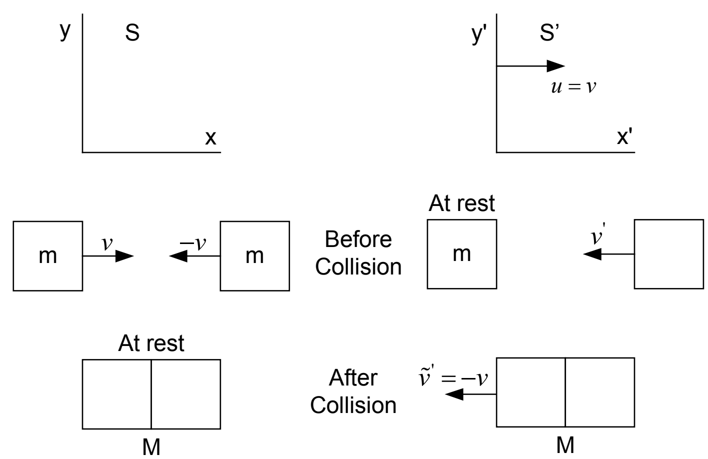

Fig. 20 Geometry of an inelastic collision

$$v^{'} = \frac{-v-u}{1+\frac{uv}{c^2}} = -\frac{2v}{1+\frac{v^2}{c^2}} \tag{48}$$

After the collision the resultant mass M will be moving with a velocity $\tilde{v}^{'} = -u = -v$. The statement of the conservation of momentum in this frame is thus

$$\frac{mv^{'}}{\sqrt{1-v^{'2}/c^2}} = \frac{M\tilde{v}^{'}}{\sqrt{1-\tilde{v}^{'2}/c^2}}$$

or

$$\frac{-m2v}{\left(1+\frac{v^2}{c^2}\right)\sqrt{1-\frac{4v^2/c^2}{\left(1+\frac{v^2}{c^2}\right)^2}}} = \frac{-Mv}{\sqrt{1-v^2/c^2}}$$

or

$$\frac{2m}{\sqrt{\left(1-v^2/c^2\right)^2}} = \frac{M}{\sqrt{1-v^2/c^2}}$$

from which

$$M = \frac{2m}{\sqrt{\left(1-v^2/c^2\right)}} \tag{49}$$

We therefore see that the total mass of the system after the collision M is greater than the total mass of the system before the collision $2m$. We also note that the system loses kinetic energy as a result of the collision. This suggests that mass is a form of energy and that the loss in kinetic energy accounts for the increase in mass. If this is the case we can define the total energy of a particle E to be equal to the sum of the kinetic energy plus the mass energy. The statement of the conservation of energy would then read (*kinetic energy* + *mass energy*) before collision = (*kinetic energy* + *mass energy*) after collision. We expect the mass energy to be some function of the mass $f(m)$. Using (47) and (49) and applying the conservation of energy to the unprimed frame we can write

$$2mc^2\left(\frac{1}{\sqrt{1-v^2/c^2}}-1\right) + f(2m) = f(M)$$

or

$$\frac{2mc^2}{\sqrt{1-v^2/c^2}} - 2mc^2 = f\left(\frac{2m}{\sqrt{\left(1-v^2/c^2\right)}}\right) - f(2m)$$

We note that this equation will be satisfied if the mass energy $f(m) = mc^2$. Thus the total energy of a particle will be

$$\begin{aligned} E &= T + mc^2 \\ &= mc^2\left(\frac{1}{\sqrt{1-v^2/c^2}} - 1\right) + mc^2 \end{aligned}$$

or

$$E = \frac{mc^2}{\sqrt{1-v^2/c^2}} \tag{50}$$

There are several important relations between the momentum of a particle given by (43) and the energy of a particle given by (50). Since

$$p^2 = \frac{m^2v^2}{1-v^2/c^2} \quad \text{and} \quad E^2 = \frac{m^2c^4}{1-v^2/c^2}$$

it follows that

$$E^2 = \frac{p^2c^4}{v^2} \quad \text{or} \quad \frac{p^2}{E^2} = \frac{v^2}{c^4}$$

from which

$$\frac{p_i}{E} = \frac{v_i}{c^2} \tag{51}$$

Also note that

$$p^2c^2 = \frac{m^2v^2c^4}{c^2 - v^2}$$

so that

$$\begin{aligned} p^2c^2 + m^2c^4 &= m^2c^4\left(\frac{c^2}{c^2-v^2}\right) \\ &= \frac{m^2c^4}{1-\dfrac{v^2}{c^2}} = E^2 \end{aligned}$$

or

$$E^2 = p^2c^2 + m^2c^4 \tag{52}$$

From (52) we can write

$$2E\frac{dE}{dt} = 2p_ic^2\frac{dp_i}{dt}$$

or

$$\frac{dE}{dt} = \frac{p_i}{E} c^2 \frac{dp_i}{dt}$$

Using (48) and the relation $F_i = dp_i / dt$ we obtain

$$\frac{dE}{dt} = v_i F_i \tag{53}$$

14. Transformation of Energy and Momentum

The expressions (see Eqs. (43) and (50))

$$p_i = \frac{mv_i}{\sqrt{1 - v^2/c^2}} \tag{54}$$

and

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \tag{55}$$

are the momentum and energy of a particle as measured in a certain reference frame (namely, one in which the velocity of the particle is v_i). An observer in a different reference frame will measure a different momentum and energy for this particle since he or she will measure a different velocity. In this new primed reference frame the momentum and energy are

$$p_i^{'} = \frac{mv_i^{'}}{\sqrt{1 - v^{'2}/c^2}} \tag{56}$$

and

$$E^{'} = \frac{mc^2}{\sqrt{1 - v^{'2}/c^2}} \tag{57}$$

where $v_i^{'}$ is related to v_i by (29). In order to relate $E^{'}$ and $p_i^{'}$ directly to E and p_i we first need an expression for

$$\sqrt{1 - v^{'2}/c^2}$$

in terms of the unprimed quantities. If we let $\alpha_j v_j = v_{\|}$ then from (29) we can write

$$
\begin{aligned}
v'^2 &= v_i' v_i' \\
&= \frac{\left\{\left[\delta_{ij} + (\gamma-1)\alpha_i\alpha_j\right]v_j - \gamma u\alpha_i\right\}\left\{\left[\delta_{ik} + (\gamma-1)\alpha_i\alpha_k\right]v_k - \gamma u\alpha_i\right\}}{\gamma^2\left(1-\frac{u}{c^2}v_\parallel\right)^2} \\
&= \frac{\left[\delta_{jk} + 2(\gamma-1)\alpha_j\alpha_k + (\gamma-1)^2\alpha_j\alpha_k\right]v_j v_k + \gamma^2u^2 - 2\gamma^2 u\alpha_i\alpha_i}{\gamma^2\left(1-\frac{u}{c^2}v_\parallel\right)^2} \\
&= \frac{v^2 + (\gamma^2-1)v_\parallel^2 + \gamma^2u^2 - 2\gamma^2 uv_\parallel}{\gamma^2\left(1-\frac{u}{c^2}v_\parallel\right)^2} \\
&= \frac{\left(1-\frac{u^2}{c^2}\right)v^2 + \frac{u^2}{c^2}v_\parallel^2 + u^2 - 2uv_\parallel}{\left(1-\frac{uv_\parallel}{c^2}\right)^2}
\end{aligned}
$$

Therefore

$$
\begin{aligned}
1-\frac{v'^2}{c^2} &= \frac{c^2\left(1-\frac{2uv_\parallel}{c^2}+\frac{u^2v_\parallel^2}{c^4}\right) - \left(1-\frac{u^2}{c^2}\right)v^2 - \frac{u^2}{c^2}v_\parallel^2 - u^2 + 2uv_\parallel}{c^2\left(1-\frac{uv_\parallel}{c^2}\right)^2} \\
&= \frac{c^2\left(1-\frac{u^2}{c^2}\right) - \left(1-\frac{u^2}{c^2}\right)v^2}{c^2\left(1-\frac{uv_\parallel}{c^2}\right)^2} \\
&= \frac{\left(1-\frac{u^2}{c^2}\right)\left(1-\frac{v^2}{c^2}\right)}{\left(1-\frac{uv_\parallel}{c^2}\right)^2}
\end{aligned}
$$

so that

$$
\sqrt{1-v'^2/c^2} = \frac{\sqrt{1-\frac{u^2}{c^2}}\sqrt{1-\frac{v^2}{c^2}}}{\left(1-\frac{uv_\parallel}{c^2}\right)} \tag{58}
$$

Using (58) we can write the energy El given by (57) as

$$E' = \frac{\left(1-\frac{uv_{\|}}{c^2}\right)mc^2}{\sqrt{1-\frac{u^2}{c^2}}\sqrt{1-\frac{v^2}{c^2}}}$$

$$= \frac{mc^2}{\sqrt{1-\frac{u^2}{c^2}}\sqrt{1-\frac{v^2}{c^2}}} - \frac{umv_{\|}}{\sqrt{1-\frac{u^2}{c^2}}\sqrt{1-\frac{v^2}{c^2}}}$$

$$= \frac{E}{\sqrt{1-\frac{u^2}{c^2}}} - \frac{u\alpha_j p_j}{\sqrt{1-\frac{u^2}{c^2}}}$$

The energy therefore transforms according to the relation

$$E' = \gamma\left(E - u\alpha_j p_j\right) \tag{59}$$

or its inverse relation

$$E = \gamma\left(E' + u\alpha_j p'_j\right) \tag{60}$$

Using (29) and (58) we can write the momentum p'_i

$$p'_i = \frac{mv'_i}{\sqrt{1-v'^2/c^2}}$$

$$= \frac{\left(1-\frac{uv_{\|}}{c^2}\right)m}{\sqrt{1-\frac{u^2}{c^2}}\sqrt{1-\frac{v^2}{c^2}}}\left[\frac{\left[\delta_{ij}+(\gamma-1)\alpha_i\alpha_j\right]v_j - \gamma u\alpha_i}{\gamma\left(1-\frac{uv_{\|}}{c^2}\right)}\right]$$

$$= \frac{\left[\delta_{ij}+(\gamma-1)\alpha_i\alpha_j\right]mv_j}{\sqrt{1-\frac{v^2}{c^2}}} - \frac{\gamma u\alpha_i mc^2}{c^2\sqrt{1-\frac{v^2}{c^2}}}$$

or,

$$p'_i = \left[\delta_{ij}+(\gamma-1)\alpha_i\alpha_j\right]p_j - \frac{\gamma u}{c^2}\alpha_i E \tag{61}$$

with its inverse relation

$$p_i = \left[\delta_{ij} + (\gamma - 1)\alpha_i\alpha_j\right] p_j^{'} - \frac{\gamma u}{c^2}\alpha_i E^{'} \tag{62}$$

Equations (59) - (62) are the transformation relations for momentum and energy.

15. Transformation Law for Force

The equation of motion of a particle in the unprimed frame is

$$F_i = \frac{dp_i}{dt} \tag{63}$$

and in the primed frame is

$$F_i^{'} = \frac{dp_i^{'}}{dt} \tag{64}$$

We wish to relate F_i and $F_i^{'}$ directly.
From (61) we can write

$$dp_i^{'} = \left[\delta_{ij} + (\gamma - 1)\alpha_i\alpha_j\right] dp_j - \frac{\gamma u}{c^2}\alpha_i dE \tag{65}$$

and from (28)

$$dt^{'} = \gamma\left(dt - \frac{u}{c^2}\alpha_k dx_k\right) \tag{66}$$

Thus

$$F_i^{'} = \frac{dp_i^{'}}{dt^{'}} = \frac{\left[\delta_{ij} + (\gamma - 1)\alpha_i\alpha_j\right] F_j - \frac{\gamma u}{c^2}\alpha_i \frac{dE}{dt}}{\gamma\left(1 - \frac{u}{c^2}\alpha_k v_k\right)}$$

or, using (53)

$$F_i^{'} = \frac{\left[\delta_{ij} + (\gamma - 1)\alpha_i\alpha_j\right]F_j - \frac{\gamma u}{c^2}\alpha_i v_j F_j}{\gamma\left(1 - \frac{u}{c^2}\alpha_k v_k\right)}$$

or

$$F_i^{'} = \frac{1}{\gamma\left(1 - \frac{uv_{\|}}{c^2}\right)}\left\{F_i + \left[(\gamma - 1)\alpha_i\alpha_j - \frac{\gamma u\alpha_i v_j}{c^2}\right]F_j\right\} \tag{67}$$

Equation (67) gives the force measured in the primed system in terms of the force in the unprimed system and the velocity of the particle in the *unprimed* system. We could obtain the inverse relation by interchanging primed and unprimed quantities and changing the sign of *u*. That would give us the force in the unprimed system in terms of the force in the primed system and the velocity of the particle in the *primed* system. However, it turns out to be convenient to have an expression for the force in the unprimed system in terms of the force in the primed system and the velocity of the particle in the *unprimed* system. This expression can be obtained simply by solving Eq. (67) for the unprimed force F_i. Thus from (67) we can write

$$F_i = \gamma\left(1 - \frac{uv_{\|}}{c^2}\right)F_i^{'} - \left[(\gamma - 1)\alpha_i\alpha_j - \frac{\gamma u}{c^2}\alpha_i v_j\right]F_j \tag{68}$$

Multiplying (65) through by α_i we obtain

$$\alpha_i F_i = \gamma\left(1 - \frac{uv_{\|}}{c^2}\right)\alpha_i F_i^{'} - (\gamma - 1)\alpha_j F_j + \frac{\gamma u}{c^2}v_j F_j$$

or

$$\alpha_j F_j = \left(1 - \frac{uv_{\|}}{c^2}\right)\alpha_j F_j^{'} + \frac{u}{c^2}v_j F_j \tag{69}$$

Substituting (69) into (68) we obtain

$$F_i = \gamma\left(1 - \frac{uv_{\|}}{c^2}\right)F_i^{'} - (\gamma - 1)\left[\left(1 - \frac{uv_{\|}}{c^2}\right)\alpha_i\alpha_j F_j^{'} + \frac{u}{c^2}\alpha_i v_j F_j\right] + \frac{\gamma u}{c^2}\alpha_i v_j F_j$$

$$F_i = \gamma\left(1 - \frac{uv_{\|}}{c^2}\right)F_i^{'} - (\gamma - 1)\left(1 - \frac{uv_{\|}}{c^2}\right)\alpha_i\alpha_j F_j^{'} + \frac{u}{c^2}\alpha_i v_j F_j \tag{70}$$

Multiplying (70) through by v_i we obtain

$$v_i F_i = \left(1 - \frac{uv_{\|}}{c^2}\right)\left[\gamma v_i F_i^{'} - (\gamma - 1)\alpha_i \alpha_j v_i F_j^{'}\right] + \frac{u}{c^2}\alpha_i v_i v_j F_j$$

$$v_j F_j \left(1 - \frac{uv_{\|}}{c^2}\right) = \left(1 - \frac{uv_{\|}}{c^2}\right)\left[\gamma v_j F_j^{'} - (\gamma - 1)\alpha_k v_k \alpha_j F_j^{'}\right]$$

$$v_j F_j = \gamma v_j F_j^{'} - (\gamma - 1) v_{\|} \alpha_j F_j^{'} \tag{71}$$

Substituting (71) into (70) we obtain

$$F_i = \gamma\left(1 - \frac{uv_{\|}}{c^2}\right) F_i^{'} - (\gamma - 1)\left(1 - \frac{uv_{\|}}{c^2}\right)\alpha_i \alpha_j F_j^{'} + \frac{u}{c^2}\gamma \alpha_i v_j F_j^{'} - \frac{u}{c^2}(\gamma - 1) v_{\|} \alpha_i \alpha_j F_j^{'}$$

$$F_i = \gamma F_i^{'} - (\gamma - 1)\alpha_i \alpha_j F_j^{'} - \frac{\gamma u}{c^2}\alpha_j v_j F_i^{'} + \frac{\gamma u}{c^2}\alpha_i v_j F_j^{'}$$

from which

$$F_i = \left[\gamma \delta_{ij} - (\gamma - 1)\alpha_i \alpha_j\right] F_j^{'} + \frac{\gamma u}{c^2}\left(\alpha_i F_j^{'} - \alpha_j F_i^{'}\right) v_j \tag{72}$$

This is the equation we sought giving the force in the unprimed system F_i in terms of the force in the primed system $F_i^{'}$ and the velocity in the unprimed system v_j. This equation will be used to derive the Lorentz force for charged particles in Part IV.

IV. MAXWELL'S EQUATIONS

16. Coulomb's Law

Consider a charge Q located at $\mathbf{r}_1$ and a second charge q located at $\mathbf{r}_2$. If $\hat{\mathbf{r}}_{12}$ is a unit vector in the direction of $\mathbf{r}_2 - \mathbf{r}_1$ and $r_{12} = |\mathbf{r}_2 - \mathbf{r}_1|$ then Coulomb's law states that the force on q due to Q is given by

$$\mathbf{F} = \frac{qQ}{4\pi\varepsilon_0}\frac{\hat{\mathbf{r}}_{12}}{r_{12}^2} \tag{73}$$

where $\varepsilon_0 = 8.85\times10^{-12}\ coul^2/m^2 Nt$ is the permittivity of free space.

If we define the electric field $\mathbf{E}$ due to Q to be the force per unit charge on q then from (73) we can write

$$\mathbf{E} = \frac{\mathbf{F}}{q} = \frac{Q}{4\pi\varepsilon_0}\frac{\hat{\mathbf{r}}_{12}}{r_{12}^2} \tag{74}$$

We can think of the force on q as being proportional to $\mathbf{E}$, and $\mathbf{E}$ is proportional to Q and is inversely proportional to the square of r_{12}. If we consider a sphere to be centered at Q then we can picture the electric field vector as shown in Figure 21 where the number of electric field lines is proportional to Q and the force on a test charge q will be proportional to the density of the electric field lines per unit area. This force must decrease as $1/r^2$ inasmuch as the area of a sphere increases as r^2 and the number of electric field lines is constant for a given value of Q.

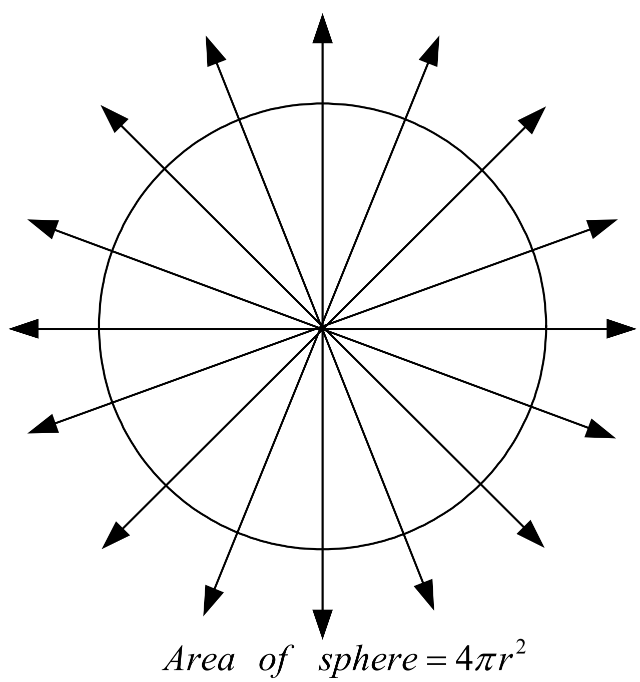

Fig. 21 Electric field of a point charge

More generally we can consider a charge density ρ such that the total charge Q within a volume V is given by

$$Q = \oint\!\!\!\oint\!\!\!\oint_V \rho\, dV$$

The electric flux Φ_E on a closed surface S is defined as

$$\Phi_E = \oiint_S \mathbf{E} \bullet \mathbf{n} dA \tag{75}$$

where **n** is an outward unit vector normal to the area dA. Gauss's law states that this flux is equal to $1/\varepsilon_0$ times the total charge in the volume V enclosed by S. We can then write Gauss's law as

$$\Phi_E = \oiint_S \mathbf{E} \bullet \mathbf{n} dA = \frac{1}{\varepsilon_0} \oiiint_V \rho dV \tag{76}$$

Using the divergence theorem we can write (76) as

$$\oiint_S \mathbf{E} \bullet \mathbf{n} dA = \oiiint_V \nabla \bullet \mathbf{E} dV = \frac{1}{\varepsilon_0} \oiiint_V \rho dV \tag{77}$$

or

$$\oiiint_V \left(\nabla \bullet \mathbf{E} - \frac{\rho}{\varepsilon_0} \right) dV = 0$$

from which

$$\nabla \bullet \mathbf{E} = \frac{\rho}{\varepsilon_0} \tag{78}$$

Let charge Q be at the origin and define the scalar potential

$$\psi = \frac{Q}{4\pi\varepsilon_0 r} \tag{79}$$

If we take the gradient of ψ in spherical coordinates we can write

$$\nabla \psi = \frac{Q\hat{\mathbf{r}}}{4\pi\varepsilon_0} \frac{d}{dr}\left(\frac{1}{r}\right) = -\frac{Q\hat{\mathbf{r}}}{4\pi\varepsilon_0 r^2} \tag{80}$$

Comparing (74) and (80) we see that we can write E as

$$\mathbf{E} = -\nabla \psi \tag{81}$$

Because **E** can be written as the negative gradient of a scalar, it follows that

$$\nabla \times \mathbf{E} = 0 \tag{82}$$

since the curl of a gradient is always zero. You can verify this by writing the components of the curl as[2]

$$\varepsilon_{ijk} \frac{\partial E_k}{\partial x_j} = \varepsilon_{ijk} E_{k,j} \tag{83}$$

and the components of the gradient as

$$\frac{\partial \psi}{\partial x_k} = \psi_{,k} \tag{84}$$

Then writing the components of (81) as $E_k = -\psi_{,k}$ we can write the curl of the gradient from (83) as

$$\varepsilon_{ijk} E_{k,j} = -\varepsilon_{ijk} \psi_{,kj} = 0 \tag{85}$$

which is easily seen to be zero by summing the repeated indices *j* and *k* from 1 to 3 and noting that $\psi_{,jk} = \psi_{,kj}$ for all *j* and *k*.

We therefore see from Coulomb's law that when the charges are at rest the electric field E satisfies the electrostatic equations (82) and (78) which can be written in component form as

$$\varepsilon_{ijk} \frac{\partial E_k}{\partial x_j} = 0 \tag{86}$$

and

$$\frac{\partial E_i}{\partial x_i} = \frac{\rho}{\varepsilon_0} \tag{87}$$

17. The Lorentz Force

Let a collection of source charges with a charge density ρ' be at rest in a reference frame S' which has velocity $u_i = u\alpha_i$ and moves uniformly relative to another reference frame *S*. This charge density gives rise to an electrostatic field E_i' in the frame S' and will produce a force $F_i' = q' E_i'$ on a test charge q'. The force on a test charge *q* as measured by observers in reference frame *S* can be found by using the force transformation given by (72).

---

[2] ε_{ijk} is the alternating unit tensor that is equal to +1 if i, j, k are in cyclic order, -1 if i, j, k are in noncyclic order, and 0 if any two subscripts are repeated. We also use the comma notation $\frac{\partial}{\partial x_j} = _{,j}$.

Two cases can be distinguished. If the velocity v_i of the test charge is zero, the electric field E_i measured in *S* is defined by the force relation $F_i = qE_i$. Assuming that charge is invariant (i.e., that $q = q^{'}$) and setting $F_i = qE_i$, $F_j^{'} = q^{'}E_j^{'}$ and $v_j = 0$ in (72), one finds that E_i is given by

$$E_i = \left[\gamma\delta_{ij} - (\gamma - 1)\alpha_i\alpha_j\right]E_j^{'} \tag{88}$$

If the test charge *q* is now allowed to move with a velocity v_i then (72) becomes

$$F_i = \left[\gamma\delta_{ij} - (\gamma - 1)\alpha_i\alpha_j\right]qE_j^{'} + \frac{q\gamma u}{c^2}\left(\alpha_i E_j^{'} - \alpha_j E_i^{'}\right)v_j \tag{89}$$

which, using (88), can be written as

$$F_i = qE_i + qC_{ij}v_j \tag{90}$$

where

$$C_{ij} = \frac{\gamma u}{c^2}\left(\alpha_i E_j^{'} - \alpha_j E_i^{'}\right) \tag{91}$$

Since C_{ij} is anti-symmetric, an axial vector B_i, called the magnetic flux density, can be defined by the relation[3]

$$C_{ij} = \varepsilon_{ijk}B_k \tag{92}$$

Therefore, (90) can be written as

$$F_i = q\left(E_i + \varepsilon_{ijk}v_j B_k\right) \tag{93}$$

or

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v}\times\mathbf{B}\right) \tag{94}$$

which is the Lorentz force. Note that the introduction of the magnetic flux vector is a convenient, not a necessary, step.

A variety of relationships exists between the field vectors $E_i^{'}$, E_i, and B_i. These relations will be used in Section 18 to aid in the derivation of Maxwell's equations. By multiplying (91) by ε_{rij} and using (92), we can write[4]

---

[3] Note that $C_{ij} = \varepsilon_{ijk}B_k = \varepsilon_{ij1}B_1 + \varepsilon_{ij2}B_2 + \varepsilon_{ij3}B_3 = \begin{pmatrix} 0 & B_3 & -B_2 \\ -B_3 & 0 & B_1 \\ B_2 & -B_1 & 0 \end{pmatrix}$

$$\varepsilon_{rij}\varepsilon_{ijk}B_k = \frac{\gamma u}{c^2}\left(\varepsilon_{rij}\alpha_i E_j^{'} - \varepsilon_{rij}\alpha_j E_i^{'}\right)$$

$$\varepsilon_{ijr}\varepsilon_{ijk}B_k = \frac{\gamma u}{c^2}\left(\varepsilon_{rij}\alpha_i E_j^{'} + \varepsilon_{rji}\alpha_j E_i^{'}\right)$$

$$\left(\delta_{jj}\delta_{rk} - \delta_{jk}\delta_{rj}\right)B_k = \frac{\gamma u}{c^2}\left(\varepsilon_{rij}\alpha_i E_j^{'} + \varepsilon_{rij}\alpha_i E_j^{'}\right)$$

$$\left(3\delta_{rk} - \delta_{rk}\right)B_k = \frac{2\gamma u}{c^2}\varepsilon_{rij}\alpha_i E_j^{'}$$

$$B_r = \frac{\gamma u}{c^2}\varepsilon_{rij}\alpha_i E_j^{'} \tag{95}$$

which is the same as

$$\mathbf{B} = \frac{\gamma}{c^2}\left(\mathbf{u}\times\mathbf{E}^{'}\right) \tag{96}$$

Multiplying (88) by the unit vector α_i, leads to

$$\alpha_i E_i = \alpha_i E_i^{'} \tag{97}$$

and thus the component of E in the direction of motion is invariant. Also multiplying (88) by $\varepsilon_{rsi}\alpha_s$ leads to[5]

$$\varepsilon_{rsi}\alpha_s E_i = \gamma\varepsilon_{rsj}\alpha_s E_j^{'} \tag{98}$$

and thus the component of **E** perpendicular to the direction of motion is larger than the corresponding component of $\mathbf{E}^{'}$ by the factor γ.

From (95) and (98) an expression relating B_i and E_i can be written as

$$B_i = \frac{u}{c^2}\varepsilon_{ijk}\alpha_j E_k \tag{99}$$

Multiplying (91) by α_j and using (92) and (97), we can write

$$\alpha_j C_{ij} = \frac{\gamma u}{c^2}\left(\alpha_j\alpha_i E_j^{'} - \alpha_j\alpha_j E_i^{'}\right)$$

---

[4] In step 3 we use the general relationship $\varepsilon_{ijk}\varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}$

[5] Note that $\varepsilon_{rsi}\alpha_s\alpha_i = 0$

$$\alpha_j \varepsilon_{ijk} B_k = \frac{\gamma u}{c^2}\left(\alpha_j \alpha_i E_j - \alpha_j \alpha_j E_i^{'}\right)$$

from which

$$E_i^{'} = \alpha_i \alpha_j E_j - \frac{c^2}{\gamma u}\varepsilon_{ijk}\alpha_j B_k \tag{100}$$

Using (97), Eq. (88) can also be written as

$$E_i^{'} = \frac{1}{\gamma}E_i + \frac{(\gamma-1)}{\gamma}\alpha_i \alpha_j E_j \tag{101}$$

Equating the right-hand sides of (85) and (86), we can write

$$\alpha_i \alpha_j E_j - \frac{c^2}{\gamma u}\varepsilon_{ijk}\alpha_j B_k = \frac{1}{\gamma}E_i + \frac{(\gamma-1)}{\gamma}\alpha_i \alpha_j E_j$$

$$-\frac{c^2}{u}\varepsilon_{ijk}\alpha_j B_k = E_i - \alpha_i \alpha_j E_j$$

$$\alpha_i \alpha_j E_j = E_i + \frac{c^2}{u}\varepsilon_{ijk}\alpha_j B_k \tag{102}$$

Substituting (102) into (100), and using $u^2/c^2 = \left(\gamma^2 - 1\right)/\gamma^2$ we obtain

$$E_i^{'} = E_i + \frac{c^2}{u}\varepsilon_{ijk}\alpha_j B_k - \frac{c^2}{\gamma u}\varepsilon_{ijk}\alpha_j B_k$$

$$E_i^{'} = E_i + \frac{c^2}{u}\left(\frac{\gamma-1}{\gamma}\right)\varepsilon_{ijk}\alpha_j B_k$$

$$E_i^{'} = E_i + \frac{\gamma}{\gamma+1}u\varepsilon_{ijk}\alpha_j B_k \tag{103}$$

18. Maxwell's Equations

Since $E_i^{'}$ is a static field in the $S^{'}$ reference frame, it satisfies the electrostatic equations (see Eqs. (86) and (87))

$$\varepsilon_{ijk}\frac{\partial E_k^{'}}{\partial x_j^{'}} = 0 \tag{104}$$

and

$$\frac{\partial E_i^{'}}{\partial x_i^{'}} = \frac{\rho^{'}}{\varepsilon_0} \tag{105}$$

In this section it will be shown that these two equations transform into the four Maxwell equations which describe the time and spatial variations of E_i and B_i in the S frame.

First it is necessary to relate time and spatial variations in the two frames of reference. Consider some function $f\left(x_i^{'}, t^{'}\right)$. From the Lorentz transformations given in (28) we see that both $x_i^{'}$ and $t^{'}$ will be functions of x and t. Using (28) we can write

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= \frac{\partial f}{\partial x_j^{'}} \frac{\partial x_j^{'}}{\partial x_i} + \frac{\partial f}{\partial t^{'}} \frac{\partial t^{'}}{\partial x_i} \\ &= \left[\delta_{ij} + (\gamma - 1)\alpha_i \alpha_j \right] \frac{\partial f}{\partial x_j^{'}} - \frac{\gamma u}{c^2} \alpha_i \frac{\partial f}{\partial t^{'}} \end{aligned} \tag{106}$$

and

$$\begin{aligned} \frac{\partial f}{\partial t} &= \frac{\partial f}{\partial t^{'}} \frac{\partial t^{'}}{\partial t} + \frac{\partial f}{\partial x_j^{'}} \frac{\partial x_j^{'}}{\partial t} \\ &= \gamma \frac{\partial f}{\partial t^{'}} - \gamma u \alpha_j \frac{\partial f}{\partial x_j^{'}} \end{aligned} \tag{107}$$

In a similar way we can write

$$\begin{aligned} \frac{\partial f}{\partial x_i^{'}} &= \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial x_i^{'}} + \frac{\partial f}{\partial t} \frac{\partial t}{\partial x_i^{'}} \\ &= \left[\delta_{ij} + (\gamma - 1)\alpha_i \alpha_j \right] \frac{\partial f}{\partial x_j} + \frac{\gamma u}{c^2} \alpha_i \frac{\partial f}{\partial t} \end{aligned} \tag{108}$$

and

$$\begin{aligned} \frac{\partial f}{\partial t^{'}} &= \frac{\partial f}{\partial t} \frac{\partial t}{\partial t^{'}} + \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial t^{'}} \\ &= \gamma \frac{\partial f}{\partial t} + \gamma u \alpha_j \frac{\partial f}{\partial x_j} \end{aligned} \tag{109}$$

When f is a static field in $S^{'}$ [i.e., $f \neq f\left(t^{'}\right)$], then (106) and (107) reduce to

$$\frac{\partial f}{\partial x_i} = \left[\delta_{ij} + (\gamma - 1)\alpha_i \alpha_j \right] \frac{\partial f}{\partial x_j^{'}} \tag{110}$$

and

$$\alpha_j \frac{\partial f}{\partial x_j^{'}} = -\frac{1}{\gamma u} \frac{\partial f}{\partial t} \tag{111}$$

and (109) reduces to

$$\frac{\partial f}{\partial t} = -u\alpha_j \frac{\partial f}{\partial x_j} \tag{112}$$

Substituting (112) into (108) we obtain

$$\begin{aligned}
\frac{\partial f}{\partial x_i^{'}} &= \left[\delta_{ij} + (\gamma-1)\alpha_i\alpha_j\right]\frac{\partial f}{\partial x_j} - \frac{\gamma u^2}{c^2}\alpha_i\alpha_j\frac{\partial f}{\partial x_j} \\
\frac{\partial f}{\partial x_i^{'}} &= \left[\delta_{ij} - \frac{(\gamma-1)}{\gamma}\alpha_i\alpha_j\right]\frac{\partial f}{\partial x_j}
\end{aligned} \tag{113}$$

where $u^2/c^2 = \left(\gamma^2-1\right)/\gamma^2$ was used in the last step.

Consider first B_i and form the divergence of B_i using (110) and (95). Thus

$$\begin{aligned}
\frac{\partial B_i}{\partial x_i} &= \left[\delta_{ir} + (\gamma-1)\alpha_i\alpha_r\right]\frac{\gamma u}{c^2}\varepsilon_{ijk}\alpha_j\frac{\partial E_k^{'}}{\partial x_r^{'}} \\
\frac{\partial B_i}{\partial x_i} &= -\frac{\gamma u}{c^2}\alpha_j\varepsilon_{jik}\frac{\partial E_k^{'}}{\partial x_i^{'}} + \frac{\gamma(\gamma-1)u}{c^2}\varepsilon_{ijk}\alpha_j\alpha_i\alpha_r\frac{\partial E_k^{'}}{\partial x_r^{'}} \\
\frac{\partial B_i}{\partial x_i} &= 0
\end{aligned} \tag{114}$$

since the first term is zero by (104) and in the second term $\varepsilon_{ijk}\alpha_j\alpha_i = 0$.

Therefore, the divergence of B_i is zero in the S frame because the curl of $E_i^{'}$ is zero in the $S^{'}$ frame.

Now form the curl of $E_i^{'}$ as in (104) and use (113) and (112) to write

$$\begin{aligned}
\varepsilon_{ijk}\frac{\partial E_k^{'}}{\partial x_j^{'}} = 0 &= \varepsilon_{ijk}\left[\delta_{js} - \frac{(\gamma-1)}{\gamma}\alpha_j\alpha_s\right]\frac{\partial E_k^{'}}{\partial x_s} \\
&= \varepsilon_{ijk}\left[\frac{\partial E_k^{'}}{\partial x_j} + \frac{(\gamma-1)}{\gamma u}\alpha_j\frac{\partial E_k^{'}}{\partial t}\right]
\end{aligned} \tag{115}$$

Substituting (103) for $E_k^{'}$ in the first term of (115) and using (95) for B_i in the second term of (114) we can write

$$\varepsilon_{ijk}\frac{\partial E_k}{\partial x_j}+\frac{\gamma}{\gamma+1}u\varepsilon_{ijk}\varepsilon_{krs}\alpha_r\frac{\partial B_s}{\partial x_j}+\frac{c^2(\gamma-1)}{\gamma^2u^2}\frac{\partial B_i}{\partial t}=0$$

$$\varepsilon_{ijk}\frac{\partial E_k}{\partial x_j}+\frac{\gamma}{\gamma+1}u\left(\delta_{ir}\delta_{js}-\delta_{is}\delta_{jr}\right)\alpha_r\frac{\partial B_s}{\partial x_j}+\frac{c^2(\gamma-1)}{\gamma^2u^2}\frac{\partial B_i}{\partial t}=0$$

$$\varepsilon_{ijk}\frac{\partial E_k}{\partial x_j}+\frac{\gamma}{\gamma+1}u\left(\alpha_i\frac{\partial B_j}{\partial x_j}-\alpha_j\frac{\partial B_i}{\partial x_j}\right)+\frac{1}{(\gamma+1)}\frac{\partial B_i}{\partial t}=0 \qquad (116)$$

where $u^2/c^2=\left(\gamma^2-1\right)/\gamma^2$ was used in the last step. The first term in the parentheses in (116) is zero by (114) and the second term in parentheses can be changed to a time derivative using (112). Equation (116) then reduces to

$$\varepsilon_{ijk}\frac{\partial E_k}{\partial x_j}+\frac{\gamma}{\gamma+1}\frac{\partial B_i}{\partial t}+\frac{1}{(\gamma+1)}\frac{\partial B_i}{\partial t}=0$$

from which

$$\varepsilon_{ijk}\frac{\partial E_k}{\partial x_j}=-\frac{\partial B_i}{\partial t} \qquad (117)$$

which is Maxwell's equation for the curl of E. That is, (117) is the same as

$$\nabla\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t} \qquad (118)$$

Equation (114) is the same as

$$\nabla\bullet\mathbf{B}=0 \qquad (119)$$

Equations (118) and (119) are Maxwell's homogeneous equations.

Both Eqs. (119) and (118) for the divergence of **B** and the curl of **E** have been derived from (104) for the curl of $E_i^{'}$. In order to derive other equations by transforming (105), it is necessary to determine how the charge density $\rho^{'}$ transforms under a Lorentz transformation. Since it has been postulated that charge is invariant, then

$$\rho^{'}dV^{'}=\rho dV \qquad (120)$$

where

$$dV=dx_1^{\perp}dx_2^{\perp}dx^{\parallel} \text{ and } dV'=dx_1^{'\perp}dx_2^{'\perp}dx^{'\parallel} \qquad (121)$$

Then

$$dx_1^{'\perp}=dx_1^{\perp} \text{ and } dx_2^{'\perp}=dx_2^{\perp} \qquad (122)$$

and

$$x^{'\parallel}=\gamma(x^{\parallel}-ut)$$
$$dx^{'\parallel}=\gamma dx^{\parallel} \qquad (123)$$

Substituting (121), (122), and (123) into (120) we find that

$$\rho^{'} = \frac{\rho}{\gamma} \tag{124}$$

Using_(113) and (124) one can write (105) as

$$\frac{\partial E_i^{'}}{\partial x_i} - \frac{(\gamma-1)}{\gamma}\alpha_i\alpha_j\frac{\partial E_i^{'}}{\partial x_j} = \frac{\rho}{\varepsilon_0\gamma} \tag{125}$$

Substituting (100) for $E_i^{'}$ in (125), we can write

$$\alpha_i\alpha_j\frac{\partial E_j}{\partial x_i} - \frac{c^2}{\gamma u}\varepsilon_{ijk}\alpha_j\frac{\partial B_k}{\partial x_i} - \frac{(\gamma-1)}{\gamma}\alpha_i\alpha_j\alpha_i\alpha_k\frac{\partial E_k}{\partial x_j} + \frac{(\gamma-1)}{\gamma}\alpha_i\alpha_j\frac{c^2}{\gamma u}\varepsilon_{irs}\alpha_r\frac{\partial B_s}{\partial x_j} = \frac{\rho}{\varepsilon_0\gamma}$$

$$\alpha_i\alpha_j\frac{\partial E_j}{\partial x_i}\left(1-\frac{(\gamma-1)}{\gamma}\right) - \frac{c^2}{\gamma u}\alpha_j\varepsilon_{jki}\frac{\partial B_k}{\partial x_i} = \frac{\rho}{\varepsilon_0\gamma} \tag{126}$$

which, by using (112), can be written as

$$-\frac{1}{u}\alpha_j\frac{\partial E_j}{\partial t}\frac{1}{\gamma} - \frac{c^2}{\gamma u}\alpha_j\varepsilon_{jki}\frac{\partial B_k}{\partial x_i} = \frac{\rho}{\varepsilon_0\gamma}$$

$$-\alpha_i\varepsilon_0\frac{\partial E_i}{\partial t} + \varepsilon_0 c^2\alpha_i\varepsilon_{ijk}\frac{\partial B_k}{\partial x_j} = \rho u \tag{127}$$

Multiplying the right-hand side of (127) by $\alpha_i\alpha_i = 1$ and defining the current density of the source charges as $J_i = \rho u \alpha_i$, the permeability as $\mu_0 = 1/\varepsilon_0 c^2$, the magnetic field intensity as $H_k = B_k/\mu_0$, and the electric flux density as $D_i = \varepsilon_0 E_i$, we can write (127) as

$$\alpha_i\left(-J_i - \frac{\partial D_i}{\partial t} + \varepsilon_{ijk}\frac{\partial H_k}{\partial x_j}\right) = 0 \tag{128}$$

Since J_i is in the direction of α_i, the total vector in parentheses cannot be perpendicular to α_i. Therefore, the expression in parenthesis in (128) must vanish and

$$\varepsilon_{ijk}\frac{\partial H_k}{\partial x_j} = J_i + \frac{\partial D_i}{\partial t} \tag{129}$$

which is Maxwell's equation for the curl of **H** and is the same as

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{130}$$

Now, if we return to (125) and substitute (101) for $E_i^{'}$, we obtain

$$\frac{\partial E_i}{\partial x_i} + (\gamma - 1)\alpha_i \alpha_j \frac{\partial E_j}{\partial x_i} - \frac{(\gamma - 1)}{\gamma} \alpha_i \alpha_j \frac{\partial E_i}{\partial x_j} - \frac{(\gamma - 1)^2}{\gamma} \alpha_i \alpha_j \alpha_i \alpha_k \frac{\partial E_k}{\partial x_j} = \frac{\rho}{\varepsilon_0}$$

$$\frac{\partial E_i}{\partial x_i} + \alpha_i \alpha_j \frac{\partial E_j}{\partial x_i} \left[(\gamma - 1) - \frac{(\gamma - 1)}{\gamma} - \frac{(\gamma - 1)^2}{\gamma} \right] = \frac{\rho}{\varepsilon_0} \tag{131}$$

The term in the square bracket in (131) is equal to zero so that (131) reduces to

$$\frac{\partial E_i}{\partial x_i} = \frac{\rho}{\varepsilon_0}$$

which is Maxwell's equation for the divergence of **E** and is the same as

$$\nabla \bullet \mathbf{E} = \frac{\rho}{\varepsilon_0} \tag{132}$$

We have therefore derived the four Maxwell equations given by (118), (119), (130), and (132) directly from special relativity and Coulomb’s law. These equations are summarized in (133).

$$\boxed{\begin{aligned} &\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \\ &\nabla \bullet \mathbf{B} = 0 \\ &\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \\ &\nabla \bullet \mathbf{E} = \frac{\rho}{\varepsilon_0} \end{aligned}} \tag{133}$$

19. Discussion

In the preceding sections we have derived Maxwell’s equations directly from special relativity and Coulomb’s law. R. S. Elliott published a similar derivation in 1966 [1,2]. The graphical methods of showing the Lorentz transformation described in Parts I and II were first given by Brehme in 1962 [3]. Good paperbacks on special relativity include those by Helliwell [4], Ney [5], and Born [6].

In his classic book on electromagnetic theory, Stratton states, "A vast wealth of experimental evidence accumulated over the past century leads one to believe that large-scale electromagnetic phenomena are governed by Maxwell's equations."[7] Indeed the evidence appears to be overwhelming since Maxwell's equations with singular success have predicted electrostatic, magnetostatic, and electromagnetic phenomena associated with a wide range of technological applications. However, certain aspects of electromagnetic theory, which deal with the radiation emitted by accelerating charges, remain troublesome. In discussing Maxwell's theory of electromagnetism, Feynman has written, "this tremendous edifice, which is such a beautiful success in explaining so many phenomena, ultimately falls on its face . . .. The classical theory of electromagnetism is an unsatisfactory theory all by itself. There are difficulties associated with the ideas of Maxwell's theory which are not solved by and not directly associated with quantum mechanics...when electromagnetism is joined to quantum mechanics, the difficulties remain" [8]. In the discussion of radiation damping, Jackson states that ". . .a completely satisfactory treatment of the reactive effects of radiation does not exist" [9].

In this report we have shown that the electric and magnetic fields measured in an unprimed reference frame satisfy Maxwell's equations when the source charges giving rise to the fields are moving uniformly with respect to the unprimed frame. It is to be expected, therefore, that Maxwell's equations will be valid for fields resulting from unaccelerated source charges. Note that in the derivation given above, the time-variation of the field vectors arises from the spatial variation of the source charges in the primed reference frame. Of course, Maxwell's equations are frequently applied in cases where the source charges do accelerate; here they predict that accelerating charges radiate electromagnetic waves. Based upon the derivation given above, is it justified to apply Maxwell's equation in cases where the source charges are accelerated? Elliott [2] answers this question in the affirmative by arguing that the most general time-varying spatial distributions of current and charge density can be Fourier synthesized from static charge distributions in all Lorentzian frames. Some of these frames will have velocities greater than c. Elliott states that this is mathematically admissible since these are fictitious charge distributions that are added to synthesize real time-varying charges that do not have velocities in excess of c. However, is it admissible to do things mathematically, which cannot, in principle, be done physically--what is the consequence of assuming that some frame of reference has a velocity greater than the speed of light? The consequence is that a frame of reference always exists in which causality is violated. Therefore, if one insists on applying Maxwell's equations to systems involving accelerated charges, then one should expect to encounter causality problems. Hence, it is not surprising that attempts to explain radiation reaction effects using Maxwell's theory lead to results that violate causality [10,11].

If the basic cause of these is the fact that one should not expect Maxwell's equations to be exactly valid for accelerating charges, the derivation carried out in Sections 16 - 18 should give some clue as to how the appropriate equations might be found. If the source charges are at rest, then the electric field is introduced to describe the force exerted on a test charge also at rest. However, if the source charges are in uniform motion, the magnetic field is introduced to describe that additional force exerted on a test charge that is in uniform motion. Might it be convenient to introduce additional field vectors to describe source and test charges in non-uniform motion? For instance, might a field vector that arises from source charges moving with a constant acceleration be introduced to describe the additional force exerted on a uniformly accelerating test charge? Could the form of this field vector be determined if it were known how forces transform between reference frames which are accelerating with respect to each other? Would additional fields and forces arise for each additional time derivative of velocity of the source and test charges, and would the equations describing these fields be a set of coupled equations with an additional equation being added for each additional time derivative of velocity of the source charges considered?

The failure of classical electromagnetic theory mentioned above arises from the fact that the Lorentz force does not predict a radiation reaction force, while Maxwell's equations predict that accelerated charges lose energy by radiation. This loss of energy should be accounted for by some additional force acting upon the accelerated charge. From the point of view outlined in the previous paragraph, an accelerated charge would feel an additional force, but the origin of the force would be accelerating source charges located elsewhere. Such a description leads to the conclusion that an accelerated point charge in otherwise charge-free space would not radiate electromagnetic energy. This, of course, is not a new idea and is, in fact, one of the assumptions used by Wheeler and Feynman [12] in their description of radiation reaction.

The derivation in Part IV shows Maxwell's equations to hold under conditions of special relativity; however the above discussion indicates that they cannot be used with complete confidence for problems involving accelerated charges. Therefore, the question of whether something is missing from Maxwell's equations remains.

References

1. R. S. Elliott, IEEE Spectrum 3, 140, 1966.
2. R. S. Elliott, *Electromagnetics*, McGraw-Hill, New York, 1966.
3. R. W. Brehme, Am. J. Phys. 30, 489 (1962)
4. T. M. Helliwell, *Introduction to Special Relativity*, Allyn and Bacon, 1966
5. E. P. Ney, *Electromagnetism and Relativity*, Harper and Row, 1962.
6. M. Born, *Einstein's Theory of Relativity*, Dover, 1962.
7. J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, p. 1.
8. R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison-Wesley, Reading, Massachusetts, 1964, Vol. II, p. 28-1.
9. J. D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, Inc., New York, 1962, p. 579.
10. See, e.g., J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys., **17**, 157, 1945.
11. Jackson, *op. cit*., pp. 597-599.
12. Wheeler and Faynman *op. cit*., p. 160.