The Theory of Relativity

L. Knutson

September 4, 2001

1. Introduction

The theory of relativity was first introduced in a publication by Albert Einstein in the year 1905. Einstein was 26 years old at the time and was employed at the patent office in Bern, Switzerland. In spite of the fact that he had little formal training in theoretical physics, Einstein had been working to understand certain conceptual problems having to do with the properties of electromagnetic fields and the propagation of electromagnetic waves. This work eventually led to the relativity theory.

Let us begin by reviewing some of the historical background. The question "what is light" goes back many centuries. One inportant clue to the nature of light is the law of refraction, discovered in 1620 by Snell,

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{1}$$

According to Snell's Law, a beam of light that crosses the boundary between two transparent materials will be deflected in a way that depends on the index of refraction, n, of the two materials. For example, a beam passing from air ($n = 1.00$) into water ($n = 1.33$) will be deflected towards the normal. In 1640 Descartes demonstrated that the law of refraction can be explained by assuming that a beam of light consists of a stream of particles or "corpuscles" which gain velocity as they pass from air into water according to the rule $v \propto n$. In the years that followed, Isaac Newton (1642-1727) became the chief proponent of the particle theory of light, and although most scientists of the day accepted Newton's theory some (Robert Hooke for example) were unconvinced. Among the unconvinced was Christian Huygens (1629-1695), a Dutch scientist and contemporary of Newton, who in 1678 proposed a wave theory of light that was also able to explain Snell's law. In this case one needs to assume that the light waves travel more slowly in water than in air according to the rule $v \propto 1/n$.

The particle *vs* wave issue was settled (at least temporarily) in 1801 when Thomas Young demonstrated that under the appropriate conditions light exhibits interference behavior that, at the time, could only be understood in terms of the wave theory. The wave-like nature of light was confirmed in subsequent experiments by Fresnel and others, who observed and studied a variety of diffraction and interference effects. In 1850, Jean Foucault provided further evidence for the wave theory by demonstrating that light travels more slowly in water than in air, in agreement with the $v \propto 1/n$ rule.

The final confirmation of the wave theory came in the late 19th century with the development by Maxwell (in 1865) of the equations of electricity and magnetism, and with the experiments of Hertz (in 1887) which showed that electromagnetic waves could be produced and detected in the laboratory.

According to Maxwell's equations the electric and magnetic fields must obey the following mathematical rules:

$$\vec{\nabla} \cdot \vec{E} = \frac{1}{\epsilon_0} \rho, \tag{2}$$

$$\vec{\nabla} \cdot \vec{B} = 0, \tag{3}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \tag{4}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t}, \tag{5}$$

where ρ is the charge density and $\vec{J}$ is the current density. While Maxwell's goal had been to construct equations that incorporated the known effects of electricity and magnetism, what he discovered was that his equations allow wave-like electromagnetic fields, apparently capable of propagating through free space. Setting ρ and $\vec{J}$ to zero, one can easily show that the equations have solutions of (for example) the form

$$\begin{aligned} \vec{E} &= E_0 \, \hat{x} \, \mathrm{Re}\left[e^{ik(z-ct)} \right] \\ \vec{B} &= B_0 \, \hat{y} \, \mathrm{Re}\left[e^{ik(z-ct)} \right], \end{aligned} \tag{6}$$

where

$$B_0 = \sqrt{\epsilon_0 \mu_0} \, E_0 \tag{7}$$

and where

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}. \tag{8}$$

Physically, these expressions have the form of a plane wave moving in the $+z$ direction at speed $v = c = 1/\sqrt{\epsilon_0 \mu_0}$. Inserting the numerical values for ϵ_0 and μ_0, one obtains the result

$$\frac{1}{\sqrt{\epsilon_0 \mu_0}} = 2.998 \times 10^8 \text{ m/s}, \tag{9}$$

which is the speed of light.

The discovery by Maxwell that light is an electromagnetic wave was, of course, one of the most important developments in the history of physics. As we shall see, this discovery set the stage for Einstein.

2. Electromagnetic Fields and The Principle of Relativity

Einstein was very much interested in Maxwell's theory of electricity and magnetism, and apparently spent quite a bit of time thinking about various aspects of the theory. One thing Einstein did was to explore the relationships between the electric and magnetic fields that one obtains when collections of charges are viewed from different frames of reference. For example, if the charges are at rest in a particular frame of reference then we obtain only electric fields. However, if the same charge distribution is viewed from a moving frame of reference, there will be magnetic fields as well. The somewhat unsettling result is that if one then calculates the electromagnetic forces, it will turn out that the net force acting on a given charge may be different in different frames of reference!

Problems are also encountered when one thinks about the propagation of electromagnetic waves. Suppose we produce an electromagnetic wave (for example, a pulse from a strobe light) that moves off in the $+z$ direction at speed c, as predicted from Maxwell's equations. Then imagine that we use a spaceship (or a real fast car) travelling at high speed in the same direction as the light pulse in an effort to "catch up" to the wave pulse. According to the usual way of thinking, if the light pulse moves at speed c relative to the original frame, then its speed relative to the moving observer will less than c. For example we would expect that an observer travelling in the $+z$ direction at $\frac{3}{4}c$ would see the light pulse move past him or her at a relative speed of $\frac{1}{4}c$. The problem is that what the moving observer sees, namely an electromagnetic waves travelling at speed $\frac{1}{4}c$, is not consistent with Maxwell's equations.

At this point it is useful to define our frames of reference more carefully (see Fig. 1). Let us think of the frame S as being "at rest" (for example,

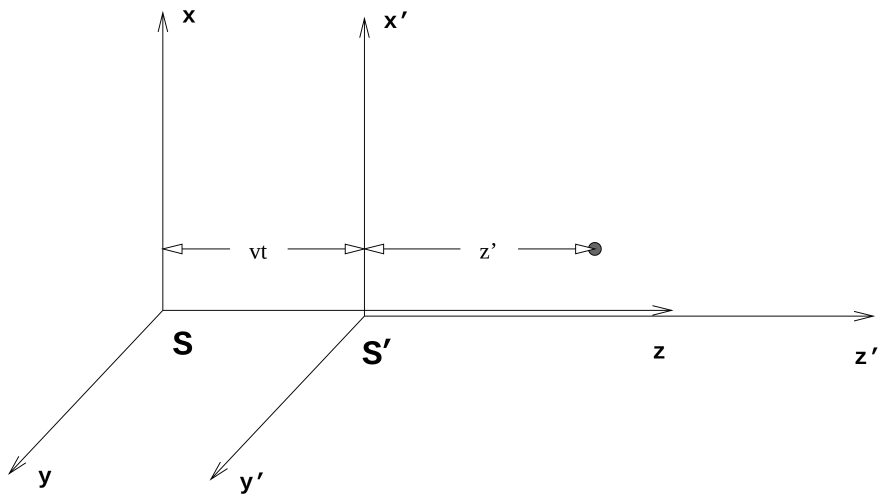

Figure 1: Frames of reference.

relative to the earth), and suppose that the frame S' moves at velocity $\vec{v}$ relative to S. Any "event" that takes place will have space and time coordinates x, y, z, and t in the S frame of reference, and similarly, in S' the save event will have coordinates x', y', z', and t'. For simplicity we take $\vec{v}$ to be along the z axis, and in addition we choose the time coordinates so that $t = t' = 0$ when the origins of S and S' coincide. Then, with the usual assumptions of classical physics, we have (see Fig. 1)

$$x' = x, \qquad y' = y, \qquad z' = z - vt, \qquad t' = t. \tag{10}$$

This set of equations is referred to as the Galilean transformation.

The transformation law for the velocity of an object follows directly from Eq. (10). Using the symbols u and u' for the velocities measured relative to S and S' and the definitions

$$\vec{u} = \frac{d\vec{r}}{dt}, \qquad \vec{u}' = \frac{d\vec{r}'}{dt'} \tag{11}$$

we obtain

$$\vec{u}' = \vec{u} - \vec{v}, \tag{12}$$

where this last equation is for $\vec{v}$ of arbitrary direction.

Although the frames S and S' move relative to each other, the two frames are similar in many respects. As seen from S, S' is moving at velocity v, and similarly as seen from S', S is moving at velocity v. The direction of motion is opposite, but other than that there is no distinction. So, if we believe that there is no preferred direction in space, we should think of the two frames as being equivalent. Furthermore, our everyday experiences lead us to believe that uniform, constant velocity motion should not have an effect on the outcome of experiments. These ideas form the basis of the Principle of Relativity. This principle states that **the laws of physics must be the same in all inertial frames of reference**, where an inertial frame is any frame in which objects move in a straight line at constant velocity unless acted upon by an outside force. If the laws of physics are the same in all inertial frames, then the outcomes of all possible experiments should be independent of the overall velocity of the system as a whole.

We are now in a position to understand more clearly why we encountered "problems" with the behavior of the electromagnetic fields. Basically, the difficulty is that that Maxwell's equations do not satisfy the Principle of Relativity. This is probably seen most clearly in the example with the electromagnetic waves. According to Maxwell's equations, light travels at the speed 2.998×10^8 m/s. But according to Eq. (12), if the speed in S is 2.998×10^8 m/s, then the speed in S' will be something else. As a result the electromagnetic fields we would "see" in the moving frame would not correspond to valid solutions of Maxwell's equations. If our analysis is correct, it seems that one needs to use different equations in different reference frames.

So the dilemma that Einstein faced can be summarized as follows: 1) Einstein believed instinctively that the Principle of Relativity had to be correct; 2) Maxwell's equations appear to be inconsistent with the Principle of Relativity; 3) However, Maxwell's equations appear to be correct in the sense that they seem to be in agreement with experiment for a wide range of electromagnetic phenomena. It would seem that there is no way to retain both Maxwell's equations and the Principle of Relativity.

3. The Ether Hypothesis

Given that there is a conceptual problem concerning the propagation of light waves, it is useful to spend a moment or two thinking about how other waves move. Let's use sound waves as an example. In this case, we can easily understand the motion by recognizing that sound waves are simply pressure waves that travel through air, water or some other medium. In air at STP, sound waves propagate at a speed of about 343 m/s, and there is no conceptual problem. In a frame in which the medium is at rest sound waves move at 343 m/s, and in other frames the speed will have a different value, easily found with Eq. (12). There is no contradiction with the Principle of Relativity because the different frames of reference are not equivalent.

Clearly then, one could resolve the electromagnetic wave problem by supposing that light waves also travel through a medium rather than through empty space. In fact, Maxwell himself had postulated the existence of a substance, which he referred to at the "luminiferous ether", to serve as the medium for electromagnetic waves. This was quite natural for Maxwell, since the physicists of the late 19th century liked to view the world from a purely mechanical perspective. By postulating the existence of ether, it was possible to think of the electric and magnetic fields as disturbances transmitted by the ether rather than "action at a distance". If ether exists then we solve the relativity problem by supposing that Maxwell's equations are correct in the ether rest frame and that light propagates at speed c in that frame only.

By 1900, experiments were beginning to cast doubt on the ether hypothesis. One of the relevant experiments concerns the small variation throughout the year of the apparent positions of the stars. This variation arises from the earth's orbital motion and is commonly referred to as "stellar aberration". Suppose we wish to observe a star that lies in the plane of the earth's orbit,

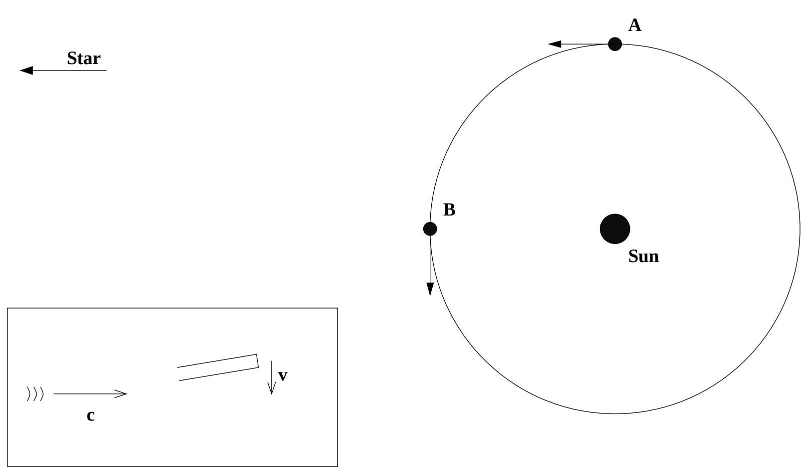

Figure 2: Illustration of stellar aberation. The apparent position of a star in the sky depends on the motion of the earth. When the earth is at A, the telescope must be pointed directly at the star. At point B, the telescope needs to be tilted away from the true position of the star, as illustrated in the inset at the lower left of the figure.

as shown in Fig. 2. When the earth is at point A we are moving towards the star, and therefore to observe the star we need to point our telescope directly at the star's true position. Three months later the earth is at B and our motion is perpendicular to that of the starlight. In this case, we observe the light from the star by tilting our telescope away from the true position of the star by an angle $\theta \simeq v/c$ where v is the earth's orbital speed, $v = 3 \times 10^4$ m/s, so that the starlight neatly "falls" down the axis of the telescope as the telescope moves. If we observe two stars whose true positions in the sky differ by 90°, the apparent angular separation might be greater than 90° in spring and less than 90° in fall.

Stellar aberration was well known to 19th century physicists, having first been observed in 1727 by British astronomer James Bradley. The observation of this effect provided some of the first convincing evidence that the propagation of light was not instantaneous and allowed Bradley to make a fairly accurate determination of c. Of course, this took place well before the wave nature of light had been firmly established.

The observations of stellar aberration are consistent with the wave picture and the ether hypothesis if one simply imagines that the ether is fixed relative to the stars and that the earth moves through the ether as it orbits the sun. However, some complications arise if one takes this picture seriously. For example, the relative motion of the ether should affect the way in which light is refracted by a lens, but this effect was not seen. Various complex explanations were put forward, but in the end these explanations were ruled out in 1887 by the well-known experiment of Michelson and Morley. It seems that there is no easy way to simultaneously explain the Michelson-Morley result, which appears to require the ether to be at rest relative to earth, with the observations of stellar aberration.

The text by McGervey[1] has a nice discussion of the early experiments on the propagation of light and the efforts to reconcile the observations with the ether hypothesis.

4. Einstein's Postulates

The ether of the late 19th century was a peculiar substance. It supposedly permeated all of space, and was capable of responding so quickly that waves could be transmitted at the enormous speed of 3×10^8 m/s. At the same time the ether apparently offered no resistance to the motion of ordinary matter. For example, the motion of the planets had been measured to high precision over many centuries and no effects from ether drag were seen.

Essentially the ether was considered to be unobservable, except through it's effect on the motion of EM waves. In retrospect, it is not so difficult see that the ether was simply an imaginative invention.

Einstein, who had great instincts for physics, never accpeted the ether hypothesis and apparently payed little attention to the ether experiments and their interpretations. Einstein believed that Maxwell's equations were fundamentally correct, and in particular, he accepted the idea that light waves simply propagate through **empty space** at speed c as the equations seem to imply. His experience and instincts also led him to believe that the relativity principle must hold. The difficulty, of course, (as we saw in Section 2) is that these two ideas seem to be incompatable, and it took someone with the brilliance of Einstein to see the way out of the dilemma.

The basic starting principles are the following:

1) **Relativity:** The laws of physics are the same in all inertial frames of reference.

2) **Propagation of Light:** The speed of light is the same in all frames of reference, independent of the motion of the source and the observer.

The relativity principle is both straightforward and consistent with our everyday experience. Saying that the laws of physics are the same in all frames is equivalent to saying that there is no way to determine what we might call the "absolute velocity" of a given reference frame. The second postulate is more difficult to accept, since it is inconsistent with the classical velocity transformation given in Eq. (12). According to Einstein's postulate if the "object" we are observing is a light pulse, then both u and u' will have the value c. To someone unfamiliar with the theory of relativity it is difficult to see how Eq. (12) could possibly be wrong since it follows easily from Eq. (10) which in turn follows from the geometry of Fig. 1. The solution is to recognize, as Einstein did, that the properties of space and time are more complex than one ordinarily imagines. In reality, the drawing in Fig. 1 is not a true representation of the geometry of space and time, and the Galilean transformation, Eq. (10), is simply wrong.

5. The Lorentz Transformation

Let us see whether we can find the correct formulas for the transformation from S to S'. Specifically, the goal is to find expressions for the event coordinates x', y', z' and t' in terms of x, y, z and t. In general, x', y' etc.

could be arbitrarily complicated functions of the unprimed coordinates:

$$\begin{aligned} x' &= f(x, y, z, t), \\ y' &= g(x, y, z, t), \\ z' &= h(x, y, z, t), \\ t' &= k(x, y, z, t). \end{aligned} \tag{13}$$

However, there are some simplifications.

First of all, one can argue that the transformation equations must be linear. This follows from the assumption that space and time are homogeneous, or in other words that the laws of physics do not depend on our absolute location in space or time.

To see this result suppose we introduce a second set of reference frames, Σ and Σ', analogous to S and S' except with different space and time origins. We suppose that Σ is at rest relative to S, and so the event coordinates in Σ (which we shall represent by capital letters) differ from those in S by at most a constant:

$$X = x + x_0, \quad Y = y + y_0, \quad Z = z + z_0, \quad T = t + t_0. \tag{14}$$

If S' moves relative to S at velocity $\vec{v}$ and Σ' moves relative to Σ at the same velocity then Σ' will be at rest relative to S' and there will be a corresponding set of equations to relate the coordinates in Σ' to those in S':

$$X' = x' + x'_0, \quad Y' = y' + y'_0, \quad Z' = z' + z'_0, \quad T' = t' + t'_0. \tag{15}$$

Now, by combining Eqs. (13) and (15) we can obtain the transformation from S to Σ'. For example

$$X' = f(x, y, z, t) + x'_0. \tag{16}$$

But if space and time are homogeneous, the transformation for $\Sigma \to \Sigma'$ must be identical to the that for $S \to S'$ (*i.e.* Eq. (13)), and from this we obtain

$$X' = f(x + x_0, y + y_0, z + z_0, t + t_0). \tag{17}$$

The conclusion is that $f(x, y, z, t)$ and $f(x + x_0, y + y_0, z + z_0, t + t_0)$ differ only by a constant. Furthermore, this result must hold for all possible events – i.e., for all x, y, z and t – and this is will be the case if and only if f is a linear function of the coordinates. One can make analogous arguments for

g, h and k, and therefore we have

$$
\begin{aligned}
x' &= a_{11}\,x + a_{12}\,y + a_{13}\,z + a_{14}\,t, \\
y' &= a_{21}\,x + a_{22}\,y + a_{23}\,z + a_{24}\,t, \\
z' &= a_{31}\,x + a_{32}\,y + a_{33}\,z + a_{34}\,t, \\
t' &= a_{41}\,x + a_{42}\,y + a_{43}\,z + a_{44}\,t.
\end{aligned} \tag{18}
$$

The constant terms that would ordinarily be present in arbitrary linear functions have been omitted in Eq. (18) because our coordinates are defined in such a way that $x' = y' = z' = t' = 0$ when $x = y = z = t = 0$.

The next simplification comes from recognizing that the equations for x' and y' cannot be complicated. As S' moves relative to S the z and z' axes remain co-linear and so any event that occurs on the z-axis must also occur on z'. In other words, events with $x = y = 0$ must always have $x' = y' = 0$. Since this will be the case for all z and t, we conclude that a_{13}, a_{14}, a_{23}, and a_{24} are all zero. The coefficient a_{12} must also be zero. This can be seen, for example, by noting that events that occur along the y-axis ($x = z = 0$) at time $t = 0$ must be on the y'-axis in S' (*i.e.*, must have $x' = 0$), since the frames coincide at $t = 0$. The coefficient a_{21} must be zero by an analogous argument, and so we have

$$
x' = a_{11}\,x, \qquad\qquad y' = a_{22}\,y. \tag{19}
$$

According to Einstein's postulates, all inertial frames must be equivalent and thus we may assume that in the inverse transformation x and y should depend only on x' and y', respectively,

$$
x = b_{11}\,x', \qquad\qquad y = b_{22}\,y'. \tag{20}
$$

But the inverse transformation can be obtained by simply solving Eqs. (18) for x and y, and it follows that the coefficients a_{31}, a_{32}, a_{41} and a_{42} in Eq. (18) must be zero.

One additional clue to the form of the transformation can be found by remembering that the origin of S' moves at velocity v in frame S. This means that events with $z = vt$ occur at the origin of S', and so we should get $z' = 0$ whenever $z = vt$.

Combining all of these results we may write Eqs. (18) in the simplified form

$$
\begin{aligned}
x' &= ax, \\
y' &= by, \\
z' &= d(z - vt), \\
ct' &= fz + gct,
\end{aligned} \tag{21}
$$

where we have switched over to a subscript-free notation, and where factors of c have been inserted as needed to make the coefficients a, b, d, f and g dimensionless. Note that at this point the equations are fully consistent with the Galilean transformation, Eq. (10).

To understand what changes are required in the theory of relativity we need to incorporate Einstein's idea that light always propagates at the speed c. Imagine an experiment in which a pulse of light is produced at the common origins of S and S' at time $t = t' = 0$. A photodetector is placed at some arbitrary location in S, and the event we shall be concerned with is the arrival of the wavefront at the detector. If the event has space and time coordinates x, y, z and t, then the wavefront has travelled a distance $s = \sqrt{x^2 + y^2 + z^2}$. By Einstein's second postulate s must be equal to ct and therefore we have

$$x^2 + y^2 + z^2 - c^2t^2 = 0. \tag{22}$$

As seen from S', the light source and the detector are in motion, but this has no effect on the light propagation – the wavefront travels at speed c in S' also, and so if x', y', z' and t' are the event coordinates in S', we must obtain

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0. \tag{23}$$

This obviously places important constraints on the form of the transformation. The condition is that $I' \equiv x'^2 + y'^2 + z'^2 - c^2t'^2$ must be zero whenever $I \equiv x^2 + y^2 + z^2 - c^2t^2 = 0$.

To see what the consequences are, we substitute from Eq. (21). The result for I' is

$$\begin{aligned} I' = \quad & a^2x^2 + b^2y^2 + [d^2 - f^2]z^2 - [g^2 - \beta^2 d^2]c^2t^2 \\ & -2[fg + \beta d^2]zct, \end{aligned} \tag{24}$$

where

$$\beta \equiv \frac{v}{c}. \tag{25}$$

To use this result we need to remember that we are free to place our photodetector anywhere in space, which means that x, y and z are arbitrary. Given the detector location, t is determined by the speed of light, and the condition is that the resulting I' must be zero. With this in mind, we can easily see that the combination $[fg + \beta d^2]$ must be zero. Detectors placed on the positive z axis and at the corresponding point on the negative z axis will give events with the same time coordinate, and in both cases we must get $I' = 0$. This can only be the case if terms linear in z are absent. Next we note that detectors placed at the same distance s from the origin on the x, y or z axes will also have equal time coordinates. Since we must get the same result, $I' = 0$, in all cases, the coefficients a^2, b^2 and $[d^2 - f^2]$ must all be equal. Furthermore, since $x = ct$ for events corresponding to detectors located on the positive x-axis, it must also be the case that $[g^2 - \beta^2 d^2] = a^2$.

The interesting conclusion we have obtained is that I' can be written in the form

$$I' = a^2x^2 + a^2y^2 + a^2z^2 - a^2c^2t^2 \tag{26}$$

or simply

$$I' = a^2 I. \tag{27}$$

It is important to understand that while Eqs. (22) and (23) hold only for the special "wavefront events" considered above, Eq. (27) is completely general. We have used the wavefront events to obtain information about the transformation coefficients, and since Eq. (27) follows from the resulting constraints on the coefficients it applies equally to all events.

The final step in the argument is to note that the only sensible value for a is 1. As seen from either frame of reference, the other frame is moving, and therefore it would be incongruous to argue that the combination $x^2+y^2+z^2-c^2t^2$ should be larger in one frame than in the other for all possible events. Since all inertial frames are equivalent, the only sensible assumption is that this particular combination has the same value in all frames of reference.

To summarize, we have the following results:

$$a = b = 1, \tag{28}$$

$$fg + \beta d^2 = 0, \tag{29}$$

$$d^2 - f^2 = 1, \tag{30}$$

$$g^2 - \beta^2 d^2 = 1. \tag{31}$$

From this point it is simply a matter of algebra to find the coefficients. First, rearrange Eq. (29) and square to obtain

$$\beta^2 d^4 = f^2 g^2. \tag{32}$$

Then substitute for f^2 from Eq. (30) and for g^2 from Eq. (31). The resulting equation can be solved for d with the result

$$d = \gamma. \tag{33}$$

where we have introduced the shorthand notation

$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}}. \tag{34}$$

Completing the algebra one obtains

$$g = \gamma. \tag{35}$$

and

$$f = -\beta\gamma, \tag{36}$$

Thus, the correct relativistic transformation equations are

$$\begin{aligned} x' &= x, \\ y' &= y, \\ z' &= \gamma(z - \beta ct), \\ ct' &= \gamma(ct - \beta z). \end{aligned} \tag{37}$$

This set of formulas is referred to as the Lorentz transformation.

The inverse transformation is easily obtained by simply inverting these equations. The result is

$$\begin{aligned} x &= x', \\ y &= y', \\ z &= \gamma(z' + \beta ct'), \\ ct &= \gamma(ct' + \beta z'). \end{aligned} \tag{38}$$

Notice that the inverse transformation has the physically reasonable property of being identical to the forward transformation, except for the sign of the velocity.

6. Time Dilation

To illustrate the usefulness of Eq. (37) we shall now use the transformation equations to derive the familiar time dilation formula. Suppose we choose two arbitrary events with space and time coordinates x_1, y_1, z_1, t_1, and x_2, y_2, z_2, t_2 in frame S. From the Lorentz transformation, the elapsed time between the two events in the S' frame will be

$$\Delta t' = t'_2 - t'_1 = \gamma(t_2 - t_1) - \frac{1}{c}\beta\gamma(z_2 - z_1). \tag{39}$$

We see immediately that if the two events occur at the same space point in S, then

$$\Delta t' = \gamma \Delta t. \tag{40}$$

In the same way one can easily demonstrate that if the two events occur at the same space point in S' then

$$\Delta t = \gamma \Delta t'. \tag{41}$$

In general we say that the "proper time" is the time interval between two events measured in a frame in which the events occur at the same space point. What the results given above show is that the time interval measured in any other frame will be longer than the proper time by a factor of γ, where γ is to be calculated using the velocity of the "other" frame relative to the proper-time frame.

Note that the proper time is special in the sense that it can be measured directly with a single clock. The measurement of non-proper time intervals is more complex since it requires the use of sychronized clocks separated in space by a distance $\gamma\beta c\Delta\tau$ (where $\Delta\tau$ is the proper time interval) or equivalently the transmission of information over that distance.

As an example, suppose the two events are the creation and the decay of a π-meson. If the π is at rest in our frame of reference (take that to be S') then the lifetime we measure will be a proper time interval. As seen from any other frame of reference (S) the π will be moving and the measured lifetime will be greater than that measured in the π rest frame by a factor of γ. As usual, γ is calculated using the velocity of S' relative to S, which in this example is just the velocity of the π in S. Experimentally one observes that particles moving at velocity v have longer mean lifetimes, by a factor γ, than corresponding particles at rest in the lab.

7. Four-Vector Notation

Although one can easily work directly with the Lorentz transformation formulas as we have written them in Eq. (37) above, it is convenient to rewrite the equations in matrix form. In our new notation, the quantities x, y, z and t will be expressed as the components of a "4-vector", x_μ, where the index μ takes on the values 1-4. We define

$$x_1 = x, \qquad x_2 = y, \qquad x_3 = z, \qquad x_4 = ict, \tag{42}$$

where the factor $i \equiv \sqrt{-1}$ has been included for reasons to be seen later. With these definitions the transformation of Eq. (37) takes the form

$$\begin{bmatrix} x' \\ y' \\ z' \\ ict' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ ict \end{bmatrix}. \tag{43}$$

If we use the notation $\Gamma_{\mu\nu}$ for the elements of the transformation matrix, then we may write Eq. (43) in the form

$$x'_\mu = \sum_\nu \Gamma_{\mu\nu} x_\nu, \tag{44}$$

or simply

$$\boldsymbol{x'} = \boldsymbol{\Gamma}\, \boldsymbol{x}. \tag{45}$$

The transformation rule in Eq. (44) is analogous in many ways to the transformation that gives the components of an ordinary three-dimensional vector in a frame O' that is rotated relative to some original frame O. Such a transformation would have the form

$$V'_i = \sum_{j=1}^{3} R_{ij} V_j, \tag{46}$$

where R is a 3×3 rotation matrix. Of course, all vectors transform from O to O' according to the same rule, and it follows that we can define a vector to be any set of three quantities that transform, under rotations, according to Eq. (46).

In the same way, we **define** a **4-vector** to be set of four quantities, A_μ with $\mu = $ 1-4, that transform from S to S' in precisely the same way as the space-time coordinates, x_μ,

$$A'_\mu = \sum_\nu \Gamma_{\mu\nu} A_\nu. \tag{47}$$

As we progress through the section on relativity we will learn that it is possible to construct a number of physically interesting relativistic 4-vectors.

8. The Velocity Transformation

In Section 4 we discussed the idea the “common-sense” velocity tranformation given in Eq. (12) must be incorrect if we accept Einstein’s postulates.

Our goal in the present section is to use the Lorentz transformation to find the correct relativistic velocity formula.

We define velocity in the usual way. To determine the velocity $\vec{u}$ of the object in frame S we note the position of the object at two times t_1 and t_2. The components of the average velocity for this time interval are then given by

$$u_x = \frac{\Delta x}{\Delta t}, \qquad u_y = \frac{\Delta y}{\Delta t}, \qquad u_z = \frac{\Delta z}{\Delta t}. \tag{48}$$

where $\Delta x = x_2 - x_1$, $\Delta y = y_2 - y_1$, $\Delta z = z_2 - z_1$ and $\Delta t = t_2 - t_1$. In the S' frame the velocity components will be different. To find $\vec{u}'$ we consider the two measurements in S to be two events and we use those same two events to determine the velocity components in S'. What we want then is

$$u'_x = \frac{\Delta x'}{\Delta t'}, \qquad u'_y = \frac{\Delta y'}{\Delta t'}, \qquad u'_z = \frac{\Delta z'}{\Delta t'}. \tag{49}$$

where $\Delta x' = x'_2 - x'_1$, *etc.*, and where the primed coordinates are related to the unprimed ones by the usual Lorentz transformation, Eq. (37) or (43).

Since the Lorentz transformation is linear, the quantity $\boldsymbol{\Delta x} \equiv \boldsymbol{x_2} - \boldsymbol{x_1}$ is a 4-vector and therefore transforms according to Eq. (47). The velocity components of Eq. (49) are then easily expressed in terms of the unprimed $\boldsymbol{\Delta x}$ components: for example

$$u'_x = \frac{\Delta x}{\gamma\,(\Delta t - \frac{1}{c}\beta \Delta z)}. \tag{50}$$

Dividing both the numerator and denominator by Δt we obtain

$$u'_x = \frac{u_x}{\gamma\,(1 - \beta u_z/c)}. \tag{51}$$

Similar manipulations give

$$u'_y = \frac{u_y}{\gamma\,(1 - \beta u_z/c)}, \tag{52}$$

and

$$u'_z = \frac{u_z - \beta c}{(1 - \beta u_z/c)}. \tag{53}$$

At this point is should be obvious that Eqs. (51)–(53) are valid for instantaneous velocities as well as for average velocities since the details are unchanged if one imagines the space and time intervals to be infinitesmal.

Compared with Galilean formula given in Eq. (12), the relativistic expression for the velocity is somewhat more complex. Notice, however, that if

v and u_z are less than $0.1c$, the relativistic result will differ from the classical one by no more than about 1%.

It would be interesting and convenient if it turned out that u_x, u_y and u_z transformed as the first three components of a 4-vector. However, we can easily see that this is not the case; for any 4-vector the transformation must be of the form $A_1 = A_1'$, . . . , and so we would have needed $u_x = u_x'$.

It turns out that there is a quantity, closely related to $\vec{u}$, that does transform as a 4-vector. Notice that in the definitions (48) and (49) the numerators, Δx, Δy, and Δz, are 4-vector components, and so the added complexity of Eqs. (51)–(53) arises from the fact that $\Delta t \neq \Delta t'$. It also follows that one can obtain a "4-velocity" by making a definition in which the time denominator is a quantity that has the same value in all frames. The obvious choice is to use the elapsed *proper time*, $\Delta\tau$, and define

$$U_\mu = \frac{\Delta x_\mu}{\Delta \tau}. \tag{54}$$

At this point it is useful to remember that the events we are concerned with here are the two measurements of the object's position, and it follows that the proper time is the time interval measured in the object's rest frame. So effectively we have introduced a third frame of reference, O, and our definition of the 4-velocity is odd in the sense that the space intervals are measured in one frame (S or S') and the time interval is measured in another (O). Of course, if the velocities are small compared to c then the time intervals are nearly equal in all frames and $U_1 \simeq u_x$, etc.

More generally, there is a simple relationship between $\vec{u}$ and $\boldsymbol{U}$. As seen from frame S, the object (and therefore frame O) moves at velocity $\vec{u}$ and it follows that

$$\Delta t = \gamma \Delta \tau \tag{55}$$

where the γ in this equation is to be calculated using the velocity $\vec{u}$. For the space components of $\boldsymbol{U}$ we then find

$$U_1 = \frac{\Delta x_1}{\Delta \tau} = \gamma \frac{\Delta x}{\Delta t} = \gamma u_x, \tag{56}$$

and finally, recalling that $\Delta x_4 = ic\Delta t$, we obtain

$$\boldsymbol{U} = \begin{bmatrix} \gamma u_x \\ \gamma u_y \\ \gamma u_z \\ i\gamma c \end{bmatrix}. \tag{57}$$

This last result gives us an alternate method of finding $\vec{u}'$ if $\vec{u}$ is known. One can first construct the 4-velocity $\boldsymbol{U}$, then use the $\boldsymbol{\Gamma}$ matrix to transform to the S' frame, and finally extract $\vec{u}'$ from $\boldsymbol{U'}$ using, for example, $u_x' = icU_1'/U_4'$, etc.

9. Properties of 4-Vectors

There are many parallels between the properties of relativistic 4-vectors and familiar 3-dimensional vectors of classical physics. Suppose $\vec{A}$ and $\vec{B}$ are vectors, and a, b and c are scalars. As we all know, the basic laws of classical physics have forms such as $\vec{A} = \vec{B}$ or $a = b$. We may have laws of the form $\vec{A} \cdot \vec{B} = c$, but we never have laws of the form $A_x = b$. The reason, of course, is that the value of A_x depends on the orientation of the coordinate system we adopt whereas b does not, and we require that the laws of physics should not depend on the choice of coordinates. The equation $\vec{A} \cdot \vec{B} = c$ is acceptable because $\vec{A} \cdot \vec{B}$ is independent of the choice of coordinates.

Many of these ideas carry over to 4-vectors. According to Einstein's postulates, the laws of physics must be the same in all inertial frames. We say **the fundamental laws must be covariant under the Lorentz transformation**, which means that the basic equations must be identical in S and S'. If one can identify 4-vectors, it is straightforward to construct equations that satisfy this condition. For example, if $\boldsymbol{A}$ and $\boldsymbol{B}$ are 4-vectors, then the set of equations

$$A_\mu = kB_\mu, \qquad \mu = 1,4 \tag{58}$$

where k is a constant, will satisfy the Lorentz covariance condition.

Ordinary vectors have the property that the dot product is a scalar quantity, and one can easily demonstrate that 4-vectors have an analogous property: namely,

$$\sum_\mu A_\mu B_\mu = \sum_\mu A'_\mu B'_\mu. \tag{59}$$

We say that the quantity $\sum A_\mu B_\mu$ is an "invariant" since it has the same value in all inertial frames. This means that equations of the form

$$\sum_\mu A_\mu B_\mu = constant \tag{60}$$

also satisfy the Lorentz covariance condition.

10. The Momentum 4-Vector

One of our goals in the study of relativity is to understand how the laws of classical mechanics, and in particular the laws of dynamics, need to be modified in light of Einstein's postulates. In classical physics the motion of a particle is described by Newton's second law, $\vec{F} = m\vec{a}$, and it should come as no surprise that this equation is no longer valid in Einstein's theory. In the present section we will take an important step towards understanding relativistic dynamics by introducing a new quantity, the momentum 4-vector.

In classical mechanics the momemtum of a particle traveling at velocity $\vec{u}$ is $\vec{p} = m\vec{u}$. We therefore wish to define the relativistic momentum in such a way that $\boldsymbol{p}$ is a 4-vector having the property $p_1 \to mu_x$, $p_2 \to mu_y$, and $p_3 \to mu_z$ in the limit $u \to 0$. We can satisfy these requirements by making use of the 4-velocity $\boldsymbol{U}$ of Eq. (57) and defining

$$p_\mu = mU_\mu, \tag{61}$$

so that

$$\boldsymbol{p} = \begin{bmatrix} \gamma m u_x \\ \gamma m u_y \\ \gamma m u_z \\ i\gamma m c \end{bmatrix}. \tag{62}$$

Since we want $\boldsymbol{p}$ to be a 4-vector, m must be a Lorentz invariant, and so the m in our definition is understood to be the rest mass of the particle.

The relativistic momentum is an exceedingly important quantity. As we shall see later, it plays a role in equations of relativistic dynamics. In addition to that, one finds that, with the above definition, the **total momentum is conserved** in both particle collisions and decay processes.

Let us focus for the moment on the conservation law. Our definition of the momentum is certainly a "reasonable" one, and we know that the first three components of $\boldsymbol{p}$ (summed over all participating particles) will be conserved in the limit of low velocities. So it is plausible that $\boldsymbol{p}_{total}$ might be conserved in relativistic processes, but there is certainly no simple proof that this must be the case.

On the other hand, there are good reasons to expect that momentum will conserved in collisions if we accept the results of the previous sections. In particular, what we shall demonstrate below is that momentum is, in fact, rigorously conserved in certain simple collision processes involving equal mass particles. Subsequently, we will assume that $\boldsymbol{p}_{total}$ is conserved in all situations.

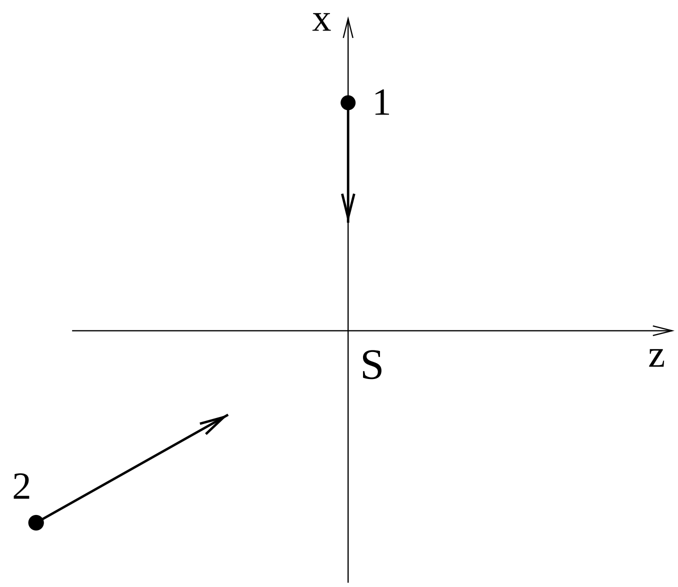

Figure 3: Inelastic collision of two identical masses

Consider the collision of two identical objects or particles of equal mass m. We assume that the two objects are moving as shown in Fig. 3 and that when they collide, they stick together forming some composite particle. The goal is to find the final velocity of this resulting blob.

Let particle 1 be moving in the $-x$ direction with some arbitrary momentum $p_x = -p_0$. We assume that particle 2 has $p_x = +p_0$ and some non-zero p_z. The two particles have equal rest masses, but move at different velocities so $\gamma_1 \neq \gamma_2$. The initial momentum 4-vectors can then be written in the form

$$\boldsymbol{p}_1 = \begin{bmatrix} -p_0 \\ 0 \\ 0 \\ i\gamma_1 mc \end{bmatrix} \qquad \boldsymbol{p}_2 = \begin{bmatrix} +p_0 \\ 0 \\ \gamma_2 m u_z \\ i\gamma_2 mc \end{bmatrix}. \tag{63}$$

Let us first find the final velocity of the blob assuming that momentum is conserved. Initially, the total momentum is just $\boldsymbol{p}_1 + \boldsymbol{p}_2$. After the collision we have a blob of mass M moving with some velocity $\vec{v}$, and so the final momentum can be written in the form of Eq. (62). Equating the initial and final momenta we obtain

$$\boldsymbol{p}_{total} = \begin{bmatrix} 0 \\ 0 \\ \gamma_2 m u_z \\ i(\gamma_1 + \gamma_2) mc \end{bmatrix} = \begin{bmatrix} \gamma_f M v_x \\ \gamma_f M v_y \\ \gamma_f M v_z \\ i\gamma_f M c \end{bmatrix}. \tag{64}$$

From this equation we can simply read off the final velocity. We have $v_x = v_y = 0$ and from the ratio ip_3/p_4 we obtain the simple result

$$\frac{v_z}{c} = \left(\frac{\gamma_2}{\gamma_1 + \gamma_2}\right) \frac{u_z}{c}. \tag{65}$$

We shall now demonstrate that this is the correct answer by solving the problem another way. In this case we transform to a new frame of reference, S', in which the colliding particles have equal and opposite momenta – *i.e.*, to the "center-of-momentum" frame. In this frame the particles have equal and opposite velocities, which means that there is no preferred direction. It then follows from symmetry that when the particles collide and stick, they come to rest. So if the frame S' is traveling at velocity $\vec{v}$ relative to S, as seen from S the final blob will have this same velocity $\vec{v}$.

This means that the problem is reduced to finding the velocity of the c.m. frame. Applying the Lorentz transformation to the momentum 4-vectors of Eq. (64) we find the momenta in an arbitrary frame S' are

$$\boldsymbol{p}_1' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{bmatrix} \begin{bmatrix} -p_0 \\ 0 \\ 0 \\ i\gamma_1 mc \end{bmatrix} = \begin{bmatrix} -p_0 \\ 0 \\ -\beta\gamma\gamma_1 mc \\ i\gamma\gamma_1 mc \end{bmatrix}. \tag{66}$$

and

$$\boldsymbol{p}_2' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{bmatrix} \begin{bmatrix} p_0 \\ 0 \\ \gamma_2 m u_z \\ i\gamma_2 mc \end{bmatrix} = \begin{bmatrix} p_0 \\ 0 \\ \gamma\gamma_2 m u_z - \beta\gamma\gamma_2 mc \\ i\gamma\gamma_2 mc - i\beta\gamma\gamma_2 m u_z \end{bmatrix}. \tag{67}$$

To make the momenta equal and opposite we therefore need

$$\beta\gamma\gamma_1 mc = \gamma\gamma_2 m u_z - \beta\gamma\gamma_2 mc, \tag{68}$$

which gives

$$\beta = \frac{\gamma_2}{\gamma_1 + \gamma_2} \frac{u_z}{c}. \tag{69}$$

So in this particular example, where we can find the correct final velocity by exploiting the symmetry of the collision, we see that momentum is, in fact, conserved. From this point on we shall simply assume that momentum is conserved in all reaction and decay processes. Of course, agreement with experiment is the real test of any theory in physics, and in the present case, many years of experimentation confirm the assumption that total momentum is conserved.

11. Energy and the Equation of Motion

So far we have not said anything about the 4^{th} component of the momentum. This component of $\boldsymbol{p}$ is also conserved in collisions and decay processes. In fact, as we shall see shortly, p_4 is proportional to the energy.

Let us begin by considering what happens if we apply a force $\vec{F}$ to some particle of rest mass m. According to classical mechanics the resulting acceleration can be found from Newton's second law,

$$\vec{F} = m\vec{a} = m\frac{d\vec{v}}{dt}, \tag{70}$$

where we have adopted a more conventional notation in which we use $\vec{v}$ rather than $\vec{u}$ for the particle velocity. Now it should be clear that Eq. (70) can no longer be correct, since this equation implies that if a constant force is applied to an object, the velocity will increase indefinitely with no upper bound. This would allow velocities to exceed c, which is not consistent with the equations of relativity.[1]

We would obtain a somewhat more reasonable equation of motion if we were to suppose that the mass of a particle increases with increasing velocity according to the rule $m_{rel} = \gamma m$, and then simply use this relativistic mass in Eq. (70) in place of the rest mass. However, this is still not the best choice. Instead we rewrite the classical equation of motion in the form

$$\vec{F} = \frac{d\vec{p}}{dt}, \tag{71}$$

[1] There are a number of ways in which this can be seen. For example, one can show from Eqs. (51)–(53) velocities greater than c can not be obtained by velocity addition. Also we know that γ becomes imaginary for $v > c$, which means that most of our equations would become meaningless. Finally we see that the momentum, $\vec{p} = \gamma m\vec{v}$, approaches infinity as $v \to c$ which suggests that particles with nonzero rest mass may never reach the speed of light.

and generalize by replacing the classical momentum, $\vec{p} = m\vec{v}$, by the relativistic momentum $\vec{p} = \gamma m\vec{v}$:

$$\vec{F} = \frac{d}{dt}\gamma m\vec{v}. \tag{72}$$

The advantage of adopting this equation of motion, or equivalently this definition of $\vec{F}$, is that when one has a system of interacting particles with no external forces, the total momentum of the system will be conserved if the mutual interactions satisfy Newton's Third Law, $\vec{F}_{ij} = -\vec{F}_{ji}$. Turning the logic around, if total momentum is conserved (which we assume to be the case), then the definition of $\vec{F}$ given in Eq. (72) will give forces that satisfy Newton's Third Law.

Let us now look more closely at the meaning of p_4. We begin with the easily demonstrated result,

$$\sum_{\mu} p_\mu p_\mu = -m^2c^2. \tag{73}$$

As the particle accelerates, the individual components of $\boldsymbol{p}$ change, but according to Eq. (73), $\sum p_\mu p_\mu$ remains constant. Thus, by taking the time derivative and using Eq. (62) we obtain

$$2\gamma\, m\, \vec{v}\cdot\frac{d\vec{p}}{dt} + 2i\gamma\, m\, c\,\frac{dp_4}{dt} = 0, \tag{74}$$

which gives

$$\frac{dp_4}{dt} = \frac{i}{c}\,\vec{F}\cdot\vec{v}. \tag{75}$$

This equation says that in the time interval dt, p_4 will change by an amount

$$dp_4 = \frac{i}{c}\,\vec{F}\cdot\vec{v}\,dt = \frac{i}{c}\,\vec{F}\cdot d\vec{s} \tag{76}$$

where $d\vec{s}$ is the net displacement in dt. But $\vec{F}\cdot d\vec{s}$ is the work done on the particle in time dt, and therefore, if we assume that the work-energy theorem of classical mechanics (*work done = gain in energy*) carries over to relativity, we have

$$dp_4 = \frac{i}{c}\,dE. \tag{77}$$

This says that, except for an overall arbitrary integration constant, p_4 is i/c times the energy. Taking the integration constant to be zero, we obtain

$$p_4 = \frac{i}{c}\,E, \tag{78}$$

and then, from Eq. (62) we have

$$E = \gamma mc^2. \tag{79}$$

The two great conservation laws of classical physics, conservation of energy and conservation of momentum, have in relativity become a single law, conservation of the four-momentum.

An additional useful relationship can be obtained by substituting Eq. (78) into Eq. (73). Rearranging the terms, we find the result

$$E^2 = p^2c^2 + m^2c^4. \tag{80}$$

12. Relativistic Dynamics

In classical mechanics the motion of particles is governed by the equation $\vec{F} = m\vec{a}$. As we have already noted, the correct relativistic generalization of Newton's Second Law is given in Eq. (72),

$$\vec{F} = \frac{d\vec{p}}{dt} = \frac{d}{dt}\gamma m\vec{v}. \tag{81}$$

One could think of this equation as a definition of what we mean by $\vec{F}$. However, this perspective is potentially missleading, since one might conclude that any definition of $\vec{F}$ is equally reasonable. It is important to remember all the fundamental equations of physics, including the law of motion, must be covariant under the Lorentz transformation. In the present context the relativity principle requires that $\vec{F}$ and $d\vec{p}/dt$ transform in the same way.

Let us now make use of this rule to determine how forces we observe in nature must transform if the covariance condition is to be satisfied. First we must understand that the components of $\Delta\vec{p}/\Delta t$ (for example $\Delta p_x/\Delta t$) are *not* Lorentz invariants, and it follows that the force components (F_x for example) will be different in different Lorentz frames.

To see how the force components transform we note that the set of quantities Δp_x, Δp_y, Δp_z and $i\Delta E/c$ transform as a 4-vector, since $\boldsymbol{\Delta p}$ is just the difference of two 4-vectors; *i.e.*

$$\boldsymbol{\Delta p} = \boldsymbol{p_2} - \boldsymbol{p_1}, \tag{82}$$

where $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ are the momentum 4-vectors of the particle at times t_1 and t_2 respectively. It follows that if $\Delta\tau$ is the elapsed proper time between the two measurements of $\boldsymbol{p}$, the quantities $\Delta p_\mu/\Delta\tau$ comprise a 4-vector. Then, recalling that $\Delta\tau$ is the elapsed time in the particle rest frame we have

$$\frac{\Delta p_\mu}{\Delta\tau} = \frac{\Delta t}{\Delta\tau}\frac{\Delta p_\mu}{\Delta t} = \gamma\frac{\Delta p_\mu}{\Delta t} \tag{83}$$

(where γ is calculated with the particle's velocity), and it follows that

$$\boldsymbol{\gamma\frac{dp}{dt}} \equiv \begin{bmatrix} \gamma\, dp_x/dt \\ \gamma\, dp_y/dt \\ \gamma\, dp_z/dt \\ i\frac{\gamma}{c}\, dE/dt \end{bmatrix} \tag{84}$$

is a 4-vector.

To make use of this result we multiply both sides of Eq. (81) by γ. We then conclude that it must be possible to construct a 4-vector whose first three components are γF_x, γF_y and γF_z. To find the corresponding fourth component we need a quantity that matches the fourth component of $\boldsymbol{\gamma\frac{dp}{dt}}$ from Eq. (84). According to Eq. (75), the appropriate choice is $i\frac{\gamma}{c}\,\vec{F}\cdot\vec{v}$. Thus we are led to the construction

$$\boldsymbol{K} \equiv \begin{bmatrix} \gamma\, F_x \\ \gamma\, F_y \\ \gamma\, F_z \\ i\frac{\gamma}{c}\,\vec{F}\cdot\vec{v} \end{bmatrix}. \tag{85}$$

This quantity is commonly referred to as the Minkowski Force.

So our conclusions are as follows. The basic equation of motion can be written in the form

$$\boldsymbol{K = \gamma\frac{dp}{dt}}. \tag{86}$$

where $\boldsymbol{K}$ and $\boldsymbol{\gamma\frac{dp}{dt}}$ are defined in Eqs. (85) and (84) respectively. Upon canceling the common factors of γ, the first three lines of Eq. (86) give the usual equation of motion, Eq. (81), while the fourth equality is the statement of the work-energy theorem

$$dE = \vec{F}\cdot d\vec{s}. \tag{87}$$

In order to satisfy the principle of relativity, the quantity $\boldsymbol{K}$ must transform as a 4-vector – if this condition is satisfied, Eq. (86) is obviously covariant.

13. Electricity and Magnetism

Since the theory of relativity was initially formulated to address concerns about the nature of electromagnetism, it is fitting that we should finally return to the subject of how the theory of electricity and magnetism can be formulated in a way that is fully consistent with the ideas of relativity. One of our specific goals in this section will be to demonstrate that electromagnetic forces obey the principles outlined in the preceeding paragraphs.

Classically, the electromagnetic force on a charged particle moving in an electromagnetic field is

$$\vec{F} = q\,(\vec{E} + \vec{v}\times\vec{B}), \tag{88}$$

and this same force law carries over in relativity. So, we can find the transformation law for $\vec{F}$, if we first understand how the fields transform. The field transformation will be somewhat complicated as we can see from the following example.

Suppose we have an infinite line of charge at rest in the S frame. This line of charge will give rise to an electric field which we can easily calculate. As seen from S', the charges will be in motion, and the resulting current will produce a magnetic field not present in S. We conclude that the field transformation law must mix electric and magnetic fields. Since there are six field components altogether, we can anticipate that the field transformation will involve something more complicated than 4-vectors.

As it turns out, it is easiest to work initially with potentials rather than fields. Given the scalar potential, ϕ and the vector potential, $\vec{A}$, the fields are obtained according to the rules

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial\vec{A}}{\partial t}\,; \qquad\qquad \vec{B} = \vec{\nabla}\times\vec{A}. \tag{89}$$

Furthermore, it turns out that the potentials *do* comprise a 4-vector:

$$\boldsymbol{A} = \begin{bmatrix} A_x \\ A_y \\ A_z \\ i\,\frac{\phi}{c} \end{bmatrix}. \tag{90}$$

For now, we simply accept this without proof.

We now use Eq. (89) to work out the field components. For example,

$$E_x = -\frac{\partial\phi}{\partial x} - \frac{\partial A_x}{\partial t}, \tag{91}$$

which, with the substitutions $x = x_1$, $ict = x_4$, $A_x = A_1$ and $i\phi/c = A_4$, becomes

$$\frac{i}{c}E_x = \frac{\partial A_1}{\partial x_4} - \frac{\partial A_4}{\partial x_1}. \tag{92}$$

With similar manipulations, all of the electric and magnetic field components can be written in an analogous forms. Thus, it is useful introduce an object that we shall refer to as the "electromagnetic field tensor". The EM field tensor is a 4×4 matrix of quantities, defined by

$$T_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu}. \tag{93}$$

The conventional electromagnetic field components are given in terms of these new quantities by

$$\begin{array}{rrr} \frac{i}{c}E_x = T_{41} & \frac{i}{c}E_y = T_{42} & \frac{i}{c}E_z = T_{43} \\ B_x = T_{23} & B_y = T_{31} & B_z = T_{12}, \end{array} \tag{94}$$

and since we have $T_{\nu\mu} = -T_{\mu\nu}$, the full field tensor is

$$\boldsymbol{T} = \begin{bmatrix} 0 & B_z & -B_y & -\frac{i}{c}E_x \\ -B_z & 0 & B_x & -\frac{i}{c}E_y \\ B_y & -B_x & 0 & -\frac{i}{c}E_z \\ \frac{i}{c}E_x & \frac{i}{c}E_y & \frac{i}{c}E_z & 0 \end{bmatrix}. \tag{95}$$

Since each element of the field matrix is constructed of parts that involve the derivative of one component of a 4-vector with respect to a component of a second 4-vector, the transformation rule for the field tensor should be easy to find.

To simplify somewhat, assume that $\boldsymbol{Q}$ is any 4-vector, and suppose that we define a tensor

$$G_{\mu\nu} \equiv \frac{\partial Q_\mu}{\partial x_\nu}. \tag{96}$$

Our goal is to find the transformation that gives $G'_{\mu\nu}$ (a given component of $\boldsymbol{G}$ in the S' frame) in terms of the quantities $G_{\rho\lambda}$ (the components of $\boldsymbol{G}$ in the S). Now from the definition $\boldsymbol{G}$ and the transformation rule for 4-vectors, Eq. (47), we have

$$G'_{\mu\nu} = \frac{\partial Q'_\mu}{\partial x'_\nu} = \frac{\partial}{\partial x'_\nu} \sum_\rho \Gamma_{\mu\rho} Q_\rho = \sum_\rho \Gamma_{\mu\rho} \frac{\partial Q_\rho}{\partial x'_\nu}. \tag{97}$$

To relate this expression to that given in Eq. (96) we need to think of Q_ρ as a function of the un-primed coordinates. Then, from the chain rule we have

$$\frac{\partial Q_\rho}{\partial x'_\nu} = \sum_\lambda \frac{\partial Q_\rho}{\partial x_\lambda} \cdot \frac{\partial x_\lambda}{\partial x'_\nu} \tag{98}$$

But $\boldsymbol{x}' = \boldsymbol{\Gamma}\boldsymbol{x}$ and similarly $\boldsymbol{x} = \boldsymbol{\Gamma}^{-1}\boldsymbol{x}'$ where $\boldsymbol{\Gamma}^{-1}$ is the transformation that takes us from S' to S, and so we have

$$x_\lambda = \sum_\sigma \Gamma^{-1}_{\lambda\sigma}\, x'_\sigma \tag{99}$$

which gives

$$\frac{\partial x_\lambda}{\partial x'_\nu} = \Gamma^{-1}_{\lambda\nu} = \Gamma_{\nu\lambda}, \tag{100}$$

where the last step follows from the fact that $\boldsymbol{\Gamma}^{-1}$ is just the transpose of $\boldsymbol{\Gamma}$. Combining these results we obtain

$$G'_{\mu\nu} = \sum_{\rho,\lambda} \Gamma_{\mu\rho}\Gamma_{\nu\lambda}G_{\rho\lambda} \tag{101}$$

or in shorthand notation

$$\boldsymbol{G}' = \boldsymbol{\Gamma}\,\boldsymbol{G}\,\boldsymbol{\Gamma}^{-1}. \tag{102}$$

It is not too difficult to see that the electromagnetic field tensor must transform in the same way,

$$\boldsymbol{T}' = \boldsymbol{\Gamma}\,\boldsymbol{T}\,\boldsymbol{\Gamma}^{-1}. \tag{103}$$

The final step is to demonstrate that the Minkowski force constructed with Eqs. (85) and (88) is a 4-vector. Working out the individual components of $\boldsymbol{K}$ we obtain

$$K_1 = \gamma q\,(E_x + v_y B_z - v_z B_y), \tag{104}$$

$$K_2 = \gamma q\,(E_y + v_z B_x - v_x B_z), \tag{105}$$

$$K_3 = \gamma q\,(E_z + v_x B_y - v_y B_x), \tag{106}$$

and

$$K_4 = i\frac{\gamma}{c}q\,(v_x E_x + v_y E_y + v_z E_z), \tag{107}$$

where we have made use of the fact that the magnetic force is perpendicular to $\vec{v}$. Inspecting these results, we see that $\boldsymbol{K}$ can be obtained by contracting $\boldsymbol{T}$ with the relativistic 4-velocity, $\boldsymbol{U}$, given in Eq. (57):

$$K_\mu = q\sum_\nu T_{\mu\nu}U_\nu. \tag{108}$$

We now easliy demonstrate that this quantity is a 4-vector:

$$\boldsymbol{K}' = q\boldsymbol{T}'\boldsymbol{U}' = q\,\left(\boldsymbol{\Gamma}\boldsymbol{T}\boldsymbol{\Gamma}^{-1}\right)(\boldsymbol{\Gamma}\boldsymbol{U}) = q\,\boldsymbol{\Gamma}\boldsymbol{T}\boldsymbol{\Gamma}^{-1}\boldsymbol{\Gamma}\,\boldsymbol{U} = q\,\boldsymbol{\Gamma}\boldsymbol{T}\,\boldsymbol{U} = \boldsymbol{\Gamma}\boldsymbol{K}, \tag{109}$$

where we have made use of the fact that $\boldsymbol{\Gamma}^{-1}\boldsymbol{\Gamma}$ is the unit matrix.

References

[1] J.D. McGervey, *Introduction to Modern Physics*, (Academic Press, New York, 1983).

[2] R. Resnick, *Introduction to Special Relativity*, (Wiley, New York, 1968).