The Special Theory of Relativity

Grant W. Mason

Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602

INTRODUCTION

The Special Theory of Relativity can seem to be a fantastic, almost unbelievable, subject. In introductory courses, it is often illustrated with thought experiments involving rocket ships or trains moving near the speed of light. The disclaimer is usually made that the effects of relativity are very, very small in ordinary experience, thus giving the impression of an esoteric subject of little practical consequence.

But for many areas of physics, Special Relativity is very important and very practical. Relativistic effects are important in nuclear physics, elementary particle physics, high-temperature plasma physics, beam physics, accelerator technology, astrophysics, space physics, cosmic-ray physics and even sometimes in high-precision atomic physics. But even if it were not so, it would still be true that the education of a physicist is not complete until he or she understands this great revolutionary subject of the twentieth century. The four-vector formulation of Special Relativity can bring one to a deeper understanding and reveal an elegance and structure the one may otherwise not recognize.

GALILEAN TRANSFORMATION

In developing the Newtonian mechanics of mass particles, we have made some assumptions. To be sure, we are reassured by experiment that these assumptions are true within the realm where the experiments have actually been performed. For example, we have taken the masses of the particles to be constants that are independent of the motion of the particle. The mass is a measure of inertia, i.e. a measure of the resistance to acceleration by a given force. Chemistry of the nineteenth century concluded that mass was a strictly conserved quantity and that any motion of the particles did not change this conclusion.

We have also made some assumptions about time and space. Consider a coordinate system with origin O that is fixed in space (the “laboratory frame”) and consider another frame with origin O' moving with constant velocity v along the x-axis of the laboratory frame. For simplicity assume that the x'-axis is parallel to the x-axis. Assume that a clock hangs on the wall of the laboratory and that observers at rest in the laboratory or moving with the primed frame can read time intervals from this clock. This is a second assumption on our part, i.e. that the single clock on the wall of the laboratory serves stationary or moving observers equally as well. Assume that there is a particle of mass m that has position $\mathbf{r}'$ in O' such that its position in O is $\mathbf{r} = \mathbf{r}' + vt$. The particle is not necessarily at rest in O', but we are assuming that if it moves, it moves along the common x and x' direction.

Now, if the particle moves in some small time interval dt that is read from the clock on the wall, we have,

$$dx = dx' + vdt$$

$$dy = dy'$$

$$dz = dz'$$

and, since time is read from the single clock on the wall,

$$dt = dt'.$$

We call this connection between the unprimed and primed coordinates of the particle a *Galilean transformation*. If we divide the first equation by dt, we get a familiar relationship between the speeds, $u_x = dx/dt$ and $u_{x'} = dx'/dt'$, along the x-axis,

$$u_x = u_{x'} + v.$$

When we search for the laws of physics, we seek generalities that serve the observer in the laboratory as well as the observer who moves with the primed frame. Our own experience, perhaps in an airliner moving in a straight line at constant speed, tells us that uniform motion is not observable. By this we mean that we would observe no violation of the laws of motion that we deduced from experiments on the ground when applied to experiments in the airliner. This is very desirable, since one would not want to dignify with the label “law” a relationship that is only valid in a particular frame of reference.

If we differentiate our velocity formula, remembering that v is constant,

$$a_x = a_{x'}.$$

For Newton’s Second Law, this means that if $\mathbf{F} = \mathbf{F}'$, then

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

and,

$$\mathbf{F}' = m\mathbf{a}'.$$

Thus, Newton’s Second Law takes the same form in each of the two frames of reference, and this is made possible, in part, by the assumption that the Galilean transformation is correct as well as the assumption that the forces themselves are equal as seen from the two frames.

To illustrate, imagine a platform on a flatcar on a railroad track that is moving at constant speed on a straight section of track. If a cannon is mounted to the platform in such a way that it fires a projectile exactly vertically, the projectile shares the horizontal speed of the flatcar when it is fired. This horizontal speed of both the platform and the projectile is unchanged (in the absence of air currents) during the motion of the projectile. Hence the projectile falls back on the platform exactly where it was first fired. Newton's Second Law applied to the motion, either from the frame of reference of the platform or from the stationary laboratory frame at the side of the track, predicts this same outcome. The same law is applicable and works in both frames to describe the projectiles's motion, although the motions will appear different to observers in the two frames.

But, consider what happens if our particle is replaced by a photon. If the photon moves along the x-axis, the Galilean transformation tells us that the speed of the photon should be the sum of its speed in the primed system plus the speed of the primed system relative to the unprimed system. It is not so. In a paper published in 1965, researchers at CERN in Geneva produced two photons from the decay of a neutral pion. The pion (which here defines the origin of the primed frame) was moving at 0.999 times the speed of light. The time of flight of these photons was measured over a distance of 80 meters and to within experimental error, the speed of both (moving in opposite directions along the x axis) was equal to the speed of light in vacuum. The motion of the pion made no difference in the speeds of the photons in the laboratory to within a precision of about 130 parts per million. This is a very remarkable result and it means that for photons, at least, something is very wrong with the Galilean transformation, something that was first suspected in 1887 when the Michelson-Morley experiment was first performed.

If you try to fix the Galilean transformation with the simplest linear generalization of the Galilean transformation, one might try,

$$dx' = \alpha dx + \beta dt$$

$$dy' = dy$$

$$dz' = dz$$

and,

$$dt' = \gamma dt + \delta dx.$$

The coefficients $\alpha, \beta, \gamma, \delta$ would presumedly be chosen in such as way as to somehow account for the unexpected outcome of the pion experiment. In doing so, the Galilean transformation would be replaced by a more general one, the relationship between velocities would also be changed, and quite possibly, we would conclude that Newton's Second Law itself as we have known it would have to be replaced with another equation in order to achieve form invariance (so-called *covariance*). Thus, the consequence of the pion experiment is far-reaching.

SPACETIME AND THE LORENTZ TRANSFORMATION

One of the consequences of our proposed new transformation is that our observers can no longer share the common clock on the wall. You can see this particularly in the fourth equation of the set where it is proposed to make dt' equal to a linear combination of dt and dx. So, from henceforth, we will provide observers with their own rulers and clocks which they carry with themselves to make measurements of space and time.

Because of this mixing of space and time, it is helpful to relax the distinction between space and time that is natural to us by beginning to think of a four-dimensional world. The "points" in this world are called *events* and have four coordinates. Three of these are the regular (x,y,z) coordinates of position relative to some suitably chosen origin and the fourth is a time coordinate measured by a clock that belongs to the frame of reference from which the event is observed. The events are "things that happen" at a certain place and at a certain time. Each event can be specified uniquely by a quartet of coordinates, (x,y,z,t) or (x',y',z',t') depending on the frame of reference and origin chosen.

Separations between events can also be specified by a quartet, (dx,dy,dz,dt) and it is these numbers that are related by the new linear transformation that we are seeking. As long as the separations are very small, we can at least assume that the transformation is approximated by a linear transformation and it is for this reason that we concentrate on the transformation for the separations. If we think of the coordinates in the primed system as being some functions of the coordinates in the unprimed system,

$$x'_i = x'_i(x_j,t)$$

$$t' = t'(x_j,t),$$

then, for small displacements and a reasonably behaved transformation, the chain rule of differentiation yields,

$$dx'_i = \frac{\partial x'_i}{\partial x_j}dx_j + \frac{\partial x'_i}{\partial t}dt$$

$$dt' = \frac{\partial t'}{\partial x_j}dx_j + \frac{\partial t'}{\partial t}dt.$$

This is exactly the form of the transformation we seek. By "reasonably behaved transformation," we mean one that is differentiable and has a unique inverse.

Now think of the experiment that we used to illustrate Galilean invariance, i.e. the cannon fired vertically from a horizontally moving platform. Only this time we will use a photon that reflects from a mirror that is overhead and moving with the platform. The primed frame is attached to the platform which we take to be moving from left to right along the direction chosen for the x and x' axes. In the primed frame, the platform is at rest and the photon moves vertically, strikes the mirror at a distance $d\ell$ away, and returns to its point of origin along the same line as the line of its ascent. The line is a series of events which is called a *world line*.

We want to focus on two of the events of this world line of the photon. Event 1 is the event of the photon being produced. Event 2 is the event of the photon striking the platform after being reflected from the mirror. These two events occur at exactly the same place in the primed system, but the time between them is the time it takes for the photon to travel a distance $2d\ell$ at the speed of light, c. Thus, in the primed system,

$$dx' = 0$$

$$dy' = 0$$

$$dz' = 0$$

$$dt' = \frac{2d\ell}{c}.$$

Viewed in the unprimed (laboratory) frame, the photon moves upward and to the right along a straight line, strikes the mirror, then moves downward and to the right before hitting the moving platform. In the unprimed system, the platform has moved a distance vdt. Thus, in the unprimed (laboratory) system,

$$dx = vdt$$

$$dy = 0$$

$$dz = 0$$

$$dt = \frac{2\sqrt{(\frac{vdt}{2})^2 + (d\ell)^2}}{c}.$$

The expression for dt is a straightforward application of Pythagoras' theorem for right triangles.

Exercise: Draw a diagram showing Event 1 and Event 2 and show that the above expressions for dx, dx', dt, dt' are correct.

By assuming that the photon has the same speed in both frames of reference, we are forcing the transformation to describe a world in which the speed of light does not depend on the motion of its source, i.e., the experimental outcome of the pion experiment.

If we substitute these relationships into our proposed transformation, we have,

$$0 = \alpha dx + \beta dt$$

which implies, since $dx = vdt$, that

$$\alpha v + \beta = 0.$$

The fact that $dt' = \frac{2d\ell}{c}$ can be combined with,

$$dt = \frac{2\sqrt{(\frac{vdt}{2})^2 + (d\ell)^2}}{c},$$

to show that,

$$dt' = \sqrt{1 - \frac{v^2}{c^2}}dt.$$

Finally, we can combine this last result with $dt' = \gamma dt + \delta dx$, to show that,

$$\gamma + \delta v = \sqrt{1 - \frac{v^2}{c^2}}.$$

Unfortunately, this gives us just two equations for the four unknown coefficients, $\alpha, \beta, \gamma, \delta$. So let's consider a second experiment to get another set of two equations. This time, hold the mirror stationary and let the photon move exactly vertically up and back in the unprimed system so that for Events 3 and 4 (which are not to be confused with Events 1 and 2 in the quite different first experiment),

$$dx = 0$$

$$dy = 0$$

$$dz = 0$$

$$dt = \frac{2d\ell}{c}$$

and,

$$dx' = -vdt'$$

$$dy' = 0$$

$$dz' = 0$$

$$dt' = \frac{2\sqrt{(\frac{vdt'}{2})^2 + (d\ell)^2}}{c}.$$

The minus sign in the expression for dx' arises because the platform moves from left to right. The reflected photon strikes a point behind the point of origin in the primed system.

Exercise: Draw a diagram showing Event 3 and Event 4 and show that the above expressions for dx, dx', dt, dt' are correct.

Because $dx = 0$, the transformation is simplified and we use the dx' and dt' equations to show, respectively,

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

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

Combining with the results from the first experiment,

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

$$\delta = \frac{\sqrt{1 - \frac{v^2}{c^2}} - \gamma}{v} = \frac{-v}{c^2\sqrt{1 - \frac{v^2}{c^2}}}.$$

We may now write down our transformation as

$$dx' = \gamma(dx - vdt)$$

$$dy' = 0$$
$$dz' = 0$$
$$dt' = \gamma(dt - \frac{vdx}{c^2}).$$

We call this transformation the *special or restricted Lorentz transformation*. It is "restricted" because we have constrained all of our motions to be along the x, x' axes. Later we will relax this restriction and derive the full Lorentz transformation in three dimensions.

The special Lorentz transformation can be inverted simply by taking the view of an observer in the primed system. From her perspective, the primed system is at rest and the unprimed system moves from right to left with speed $-v$. Thus,

$$dx = \gamma(dx' + vdt')$$
$$dy = 0$$
$$dz = 0$$
$$dt = \gamma(dt' + \frac{vdx'}{c^2}).$$

Inertial frames are frames in which there is no physical consequence of assuming that the frame is at rest or in uniform motion and this assumption is now built into our Lorentz transformation. Observe also that in the limit as $v/c \to 0$, we have $\gamma \to 1$ and the Lorentz transformation becomes the Galilean transformation. This is as it should be because the Galilean transformation serves very well for physics when speeds are much less than the speed of light. The speeds in the pion experiment were all at or near the speed of light. We have also assumed throughout that the speed of light is the same for observers in either the primed or unprimed system. The Lorentz transformation thus embodies two essential postulates which can be thought of as the foundation of the Special Theory of Relativity,

1. **The Postulate of Special Relativity**: Every "law of physics" that holds in any reference frame must equally well hold in any reference frame that moves in a straight line at constant speed relative to the first.
2. **Postulate of the Speed of Light**: The speed of light in vacuum is the same for observers at rest or moving relative to one another in a straight line at constant speed. It is an absolute invariant.

The Lorentz transformation has a singularity at $v \to c$. But, what would be the consequences of velocities that exceed the speed of light? Specifically, we focus on the transfer of information at superluminary speeds. Imagine two events, A and B, such that A precedes B. We may think of A being a "cause" of B and that information generated at the event A must propagate to B before B can happen. For example, event A might be the creation of a muon and event B might be its decay. The muon itself carries the information from event A to event B. For these two events $dt_{AB} > 0$. The Lorentz transformation tells us that in a primed system moving relative to the first with speed v,

$$dt'_{AB} = \gamma(dt_{AB} - \frac{vdx_{AB}}{c^2}).$$

Since $dx_{AB} = udt_{AB}$, we may write,

$$dt'_{AB} = \gamma dt_{AB}(1 - \frac{vu}{c^2}).$$

If there is no restriction on the magnitude of u at which the information is carried from event A to event B, i.e. $u > c$, the sign of dt'_{AB} may be made negative. Then, in the primed frame, the effect precedes its cause. We therefore add a third postulate to insure causality,

3. **Postulate of Causality**: Information is not propagated at greater than the speed of light.

CONSEQUENCES OF THE LORENTZ TRANSFORMATION

There are several consequences that immediately follow from the Lorentz transformation.

1. **Non-Simultaneity of Events**: The first consequence has to do with what we mean when we say that two events happen "at the same time." Consider a frame of reference in which two events occur. We make the following operational definition: Two events which generate a light signal are said to be *simultaneous* if an observer located midway between the spatial positions of the two events receives the light signals coincidentally. (An "operational definition" is one made in terms of an experiment. Thus experiment and not semantics can be invoked to decide whether the events are simultaneous.)
   Now, if two events that occur at different places separated by dx in the unprimed (laboratory) frame are simultaneous, then $dt = 0$. Thus, the Lorentz transformation tells us,
   $$dt' = \gamma(dt - \frac{vdx}{c^2}) = -\frac{\gamma vdx}{c^2} \neq 0.$$
   Two events that are simultaneous in one frame are not simultaneous in a frame moving relative to the first with speed v.
   Think of an unprimed system at rest and a primed stystem moving from left to right. Think of two event at points A and B (to the right of A) that emit light signals just as two points A' and B' are lined up side-by-side with A and B. These events are such that an observer halfway between A and B at point C receives a light signal from each coincidentally. We have purposely created two events that are simultaneous in the unprimed system according to the operational definition. But the observer at C' (halfway between A' and B' in the primed system and moving with it) moves forward after the light signals are emitted so that the light has less of a distance to travel to meet her. The speed of light is unaffected by the motion of the unprimed system and the signal from B' arrives first followed by the signal from A'. The observer in the primed system judges that the events are not simultaneous, just as the Lorentz transformation predicted.
2. Lorentz Contraction: A second immediate consequence of the Lorentz transformation is called the Lorentz contraction. Let us imagine a separation dx' that marks

the length of a stick that lies along the x' axis and is at rest in the primed system. As the stick moves by, an observer in the unprimed system wants to measure the length of the stick with his own ruler. One way to accomplish this is for the unprimed observer to mark the positions of the two ends of the moving stick on his stationary ruler as the stick flies by. The markings of the two ends are two events which the unprimed observer must make simultaneous for his measurement to make sense. Certainly he does not want to mark the front end first, then allow some time to elapse before marking the position of the trailing end. So, in the unprimed system, $dt = 0$. The Lorentz transformation then tells us,

$$dx' = \gamma(dx - vdt) = \gamma dx$$

i.e.,

$$dx = dx'\sqrt{1 - \frac{v^2}{c^2}}.$$

The measured length, dx, is always shorter than dx'. This we call the *Lorentz Contraction*. The length, dx', measured in the rest system of the stick (primed) is called the *proper length* of the stick.

To the primed observer, the Lorentz transformation yields,

$$dt' = -\frac{\gamma v dx}{c^2} = -\frac{\gamma v}{c^2}(\frac{dx'}{\gamma}) = -\frac{vdx'}{c^2} \neq 0.$$

The primed observer judges that the primed observer's measurement of the length of the stick is "incorrect" because, from the primed observer's point of view, the unprimed observer did not mark the ends of the stick simultaneously. Indeed, the negative sign for dt' means that from the primed point of view, the trailing end of the stick (A') was measured after the forward edge (B') so that $t_{B'} - t_{A'} < 0$, i.e. $t_{B'} < t_{A'}$.

3. **Time Dilation**: Finally, the Lorentz transformation gives us a result called *time dilation*. The tick of a clock is an event. If the clock is at rest in the primed system, two ticks of the clock occur at the same place and $dx' = 0$. The Lorentz transformation then tells us,

$$dt = \gamma(dt' + \frac{vdx'}{c^2}) = \gamma dt'$$

$$dt = \frac{dt'}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

The time interval measured in the rest frame of the clock is called the *proper time*. The temporal interval dt is always greater than dt' if the clock moves relative to the unprimed frame so that the moving clock is judged by the unprimed observer to be running slow.

Time dilation is very commonly observed for radioactively unstable particles moving at high speeds. Muons are unstable and decay to an electron, a neutrino and an antineutrino. The proper lifetime of a muon is 2.2 microseconds. Muons moving near the speed of light have a dilated mean lifetime in the laboratory. For muons in cosmic-ray showers, the dilation factor, γ, may be 10^3 or more.

4. **Addition Law for Velocities**: Having replaced the Galilean transformation with a more general Lorentz transformation, we must ask what relationship replaces the Galilean rule for adding velocities, $u'_x = u_x + v$. Assume that the primed and unprimed frames are oriented so that the special Lorentz transformation holds. Then,

$$u_x = \frac{dx}{dt} = \frac{\gamma(dx' + vdt')}{\gamma(dt' + \frac{vdx'}{c^2})} = \frac{\frac{dx'}{dt'} + v}{1 + \frac{v}{c^2}\frac{dx'}{dt'}}.$$

Thus, we have the new addition formulas for velocities,

$$u_x = \frac{u'_x + v}{1 + \frac{u'_x v}{c^2}}$$

and, similarly,

$$u_y = \frac{1}{\gamma}\frac{u'_y}{1 + \frac{u'_x v}{c^2}}$$

$$u_z = \frac{1}{\gamma}\frac{u'_z}{1 + \frac{u'_x v}{c^2}}.$$

Observe that as u'_x and v both approach the speed of light, u_x nevertheless remains less than the speed of light. For $u'_x << c$ and $v << c$, the formulas are essentially the Galilean formulas.

Exercise: Demonstrate that the formulas for u_y and u_z are correct.

Notice again the emphasis in each of these examples on the concept of an “event.” Once the Lorentz transformation is defined, the consequences of Special Relativity follow by systematically identifying events and their spatial and temporal separations. Notice also that one must carefully differentiate between what the primed and the unprimed observer measures for these separations. If one is not careful in making the distinctions between the viewpoints of the primed and unprimed frames, Special Relativity can sometimes seem paradoxical. Many of these apparent paradoxes can be resolved by focussing on the spatial and temporal separation of events as was done above.

Exercise: Consider the following. Perhaps you have seen a pole vaulter carrying a pole as he runs down the ramp to approach the crossbar. The pole is held horizonal. Imagine an unprimed (laboratory) observer watching such a pole vaulter approach a small shed equipped with garage doors, both open, at the front and back. The pole vaulter is approaching the front of the shed. Let us imagine that the proper length of the pole and of the shed is 10 meters, but that the vaulter can run very, very fast. To the unprimed observer, the moving pole is Lorentz contracted to half its length, but the shed is at rest and is not contracted. So, the five-meter contracted pole fits nicely inside the shed and the garage doors could be closed quickly, trapping the vaulter and pole inside. However, from the vaulter's (primed) point of view, the pole is at rest and has its proper length of 10 meters. The shed is moving relative to the vaulter and has a Lorentz-contracted length of 5 meters. The 10-meter pole can never fit inside the 5-meter shed. How then can relativity be correct? What two events might you wish to focus on? Are these events simultaneous? How is the paradox resolved?

THREE-VECTORS

The results that we have derived thus far from the Lorentz transformation are, for the most part, treated in every introduction to Special Relativity in about the way that we have done here. We wish now to develop a very elegant and powerful formalism that will allow us to understand Special Relativity much more deeply and to derive additional important consequences of the theory. To see what we are doing and why, it is probably worthwhile to revisit some ideas and formalism associated with ordinary vectors.

The concept of a vector is often first introduced to represent physical quantities that have a magnitude and a direction. We sometimes come to think of a vector as being defined by the characteristics of magnitude and direction. We often represent a vector by an arrow with the direction of the arrow giving the direction of the vector and the length of the arrow scaled to represent the magnitude. We will now define a vector in a more general and abstract way, which, nevertheless incorporates the characteristics of magnitude and direction.

On a piece of paper, draw a vector **a** which reaches upward and to the right. After drawing **a**, draw an x_1-axis horizontally and an x_2-axis vertically. Observe that the vector existed independently before the coordinate system was drawn, but once you draw the axes, the vector acquires components in the coordinate system. In a three-dimensional coordinate system, these might be represented as (a_1, a_2, a_3).

Now add a second coordinate system. Imagine x'_1 originally along x_1, and x'_2 along x_2. Then rotate the primed system through an angle θ counterclockwise around a common x_3, x'_3-axis imagined extending perpendicularly out of the page. The vector **a** has components in the primed system, (a'_1, a'_2, a'_3) as well, but they are not the same as the components in the unprimed system. Indeed,

$$a'_1 = a_1 \cos\theta + a_2 \sin\theta$$

$$a'_2 = -a_1 \sin\theta + a_2 \cos\theta$$

$$a'_3 = a_3.$$

Since this is a very special rotation, we shall refer to it as a *limited or special rotation.*

Exercise: Draw a diagram and derive the special rotation transformation above.

The triplet of numbers that represents the vector depends on your choice of orientation of the coordinate system. We might write the relationship among the components in matrix form,

$$\begin{pmatrix} a'_1 \\ a'_2 \\ a'_3 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

The matrix,

$$R = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is called the *rotation transformation matrix* for the special rotation. It connects the unprimed and primed components of a vector when the coordinate systems is rotated in the special way described above. In a completely general rotation of the primed systemstem relative to the unprimed system, we can represent the matrix by,

$$R = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix}.$$

We can also represent the transformation in the much more economical form,

$$a'_i = \sum_{j=1}^{3} R_{ij}a_j \equiv R_{ij}a_j.$$

In what follows, a repeated index, such as j in this example, will imply a sum on that index. Once the sum has been carried out, j disappears altogether from the expression and is called a *dummy index*. The index could, of course, be replaced by any other symbol, i.e., $R_{ij}a_j \equiv R_{ik}a_k$. The matrix R for a general rotation has a number of important generic characteristics which we can only summarize here:

1. The transformation R is an *orthogonal transformation*. This means that the transpose of the matrix R is the inverse of R, i.e.,

   $$R^{-1} = R^t$$

   $$R^tR = RR^t = I.$$

   These inverse relations may also be expressed,

   $$R_{ij}R_{ik} = \delta_{jk}$$

   $$R_{ji}R_{ki} = \delta_{jk},$$

   where δ_{jk} is the Kronecker delta function.
2. The determinant of R = +1. This feature is characteristic of a *proper* orthogonal transformation and includes rotations. If a single coordinate axis is inverted, the determinant is -1 and the transformation, although orthogonal, is said to be *improper*.

A *scalar* under rotations is a single function of position in space. Its value at a particular point in space is independent of the coordinate systems chosen, although the value of the scalar may be some function of the coordinates and may be a different function in different coordinate systems. Temperature is a classic example of a scalar function under rotations.

We can define a *3-vector* to be a triplet of components that transforms under rotation of the coordinate system according to the specific pattern:

$$a'_i = R_{ij} a_j.$$

This definition includes such familiar vectors as displacement, velocity and acceleration.

We can define a *3-tensor* to be a nonet of components that transforms under rotation of the coordinate system according to the specific pattern:

$$A'_{ij} = R_{ik} R_{jl} A_{kl}.$$

Some physical quantities, such as the dielectric tensor, are tensor quantities. Tensors of higher order with more than two subscripts can be defined in an analogous fashion. In this way, scalars, vectors and tensors are defined by the patterns or their transformation under rotations.

A scalar which has the same functional form in different coordinate systems is said to be an *invariant*. The scalar dot product of two vectors is an example of an invariant,

$$\mathbf{a} \cdot \mathbf{b} = a'_i b'_i = R_{ij} a_j R_{ik} b_k = R_{ij} R_{ik} a_j b_k = \delta_{jk} a_j b_k = a_k b_k.$$

Thus, the dot product has the same form and value in either coordinate system $(a'_i b'_i$ or $a_i b_i)$. It is an invariant. Since the dot product of a vector with itself is the square of its magnitude, the invariance property says that the length of a vector does not change when one simply rotates the coordinate system.

LORENTZ TRANSFORMATIONS

The quartet (dx, dy, dz, cdt) are said to be components of a *four-vector.* Any quartet of variables that transforms under a special Lorentz transformation in the same way as this quartet transforms is likewise a four-vector under Lorentz transformations. We may write the transformation as,

$$dx' = \gamma(dx - vdt)$$

$$dy' = 0$$

$$dz' = 0$$

$$dt' = \gamma(dt - \frac{vdx}{c^2}).$$

However, to develop the formalism, it helps to make some notational changes. Let us define,

$$\begin{pmatrix} dx \\ dy \\ dz \\ cdt \end{pmatrix} \equiv \begin{pmatrix} dx^1 \\ dx^2 \\ dx^3 \\ dx^4 \end{pmatrix}.$$

We can adapt our summation convention on repeated indices to the Lorentz transformation if Greek letters are used when the range of the indices is four rather than three. We may then write our special Lorentz transformation as,

$$\begin{pmatrix} dx'^1 \\ dx'^2 \\ dx'^3 \\ dx'^4 \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\frac{v}{c}\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{v}{c}\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} dx^1 \\ dx^2 \\ dx^3 \\ dx^4 \end{pmatrix}.$$

We might also write this as,

$$dx^{\lambda'} = L^{\lambda'}_{\mu} dx^{\mu}.$$

This is the pattern that defines a 4-vector under Lorentz transformations.

The inverse Lorentz transformation can be read from

$$dx = \gamma(dx' + vdt')$$

$$dy = 0$$

$$dz = 0$$

$$dt = \gamma(dt' + \frac{vdx'}{c^2}),$$

namely,

$$L^{-1} = \begin{pmatrix} \gamma & 0 & 0 & +\frac{v}{c}\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ +\frac{v}{c}\gamma & 0 & 0 & \gamma \end{pmatrix}.$$

It follows that,

$$LL^{-1} = L^{-1}L = I.$$

All of this is done in terms of the special Lorentz transformation. However, we can include our discussion of 3-vectors into the 4-vector formalism by simply expanding R by one dimension, i.e., the rotation transformation becomes a special case of the Lorentz transformation. For example, the matrix for the special rotation transformation is generalized to,

$$R \rightarrow \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

We can generate more general Lorentz transformations by building them from an ordered sequence of special Lorentz transformations (including special rotations). For example, consider a primed frame whose origin does not move along the x-axis, but rather moves along a line in the $x-y$ plane. The line is inclined by an angle θ to the x-axis. Thus, unlike the special Lorentz transformation, the velocity of the primed system has two components, $v_x = v\cos\theta$ and $v_y = v\sin\theta$. However, assume that the x'-axis is still parallel to the x-axis and that the y'-axis is still parallel to the y-axis.

Begin with the primed and unprimed systems both at rest with respective axes parallel. First, rotate the primed system by an angle θ. Then give the primed system a *boost* to give it a speed v relative to the unprimed system. Finally, rotate the x'-axis back through an angle $-\theta$ to get it parallel to the x-axis again. The overall Lorentz transformation is a combination of these three,

$$L = R_{-\theta} L_v R_\theta,$$

or, $L =$

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & 0 & 0 & -\frac{v}{c}\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{v}{c}\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1+(\gamma-1)\frac{v_x^2}{c^2} & (\gamma-1)\frac{v_x v_y}{c^2} & 0 & -\frac{v_x}{c}\gamma \\ (\gamma-1)\frac{v_x v_y}{c^2} & 1+(\gamma-1)\frac{v_y^2}{c^2} & 0 & -\frac{v_y}{c}\gamma \\ 0 & 0 & 1 & 0 \\ -\frac{v_x}{c}\gamma & -\frac{v_y}{c}\gamma & 0 & \gamma \end{pmatrix}$$

In a similar fashion or simply by extending the patterns that are evident from our example, one can generate the *general Lorentz transformation without rotation*,

$$L = \begin{pmatrix} 1+(\gamma-1)\frac{v_x^2}{c^2} & (\gamma-1)\frac{v_x v_y}{c^2} & (\gamma-1)\frac{v_x v_z}{c^2} & -\frac{v_x}{c}\gamma \\ (\gamma-1)\frac{v_x v_y}{c^2} & 1+(\gamma-1)\frac{v_y^2}{c^2} & (\gamma-1)\frac{v_y v_z}{c^2} & -\frac{v_y}{c}\gamma \\ (\gamma-1)\frac{v_x v_z}{c^2} & (\gamma-1)\frac{v_y v_z}{c^2} & 1+(\gamma-1)\frac{v_z^2}{c^2} & -\frac{v_z}{c}\gamma \\ -\frac{v_x}{c}\gamma & -\frac{v_y}{c}\gamma & -\frac{v_z}{c}\gamma & \gamma \end{pmatrix}.$$

This may also be written in 3-vector notation as,

$$d\mathbf{r}' = d\mathbf{r} + \mathbf{v}[(\gamma-1)\frac{d\mathbf{r}\cdot\mathbf{v}}{v^2} - \gamma dt]$$

$$dt' = \gamma[dt - \frac{d\mathbf{r}\cdot\mathbf{v}}{c^2}].$$

Exercise: Use the general Lorentz transformation without rotation to derive the expression for $d\mathbf{r}'$ in 3-vector notation and the expression for dt'.

Exercise: How do ordinary velocities transform under the general Lorentz transformation without rotation?

The subsuming of rotation transformations into the larger class of Lorentz transformations can be made more natural in the following way. Define $\tanh\theta = v/c$. The function θ is called the *rapidity*. Since, $-1 < \tanh\theta < 1$, this definition is consistent with all possible frame speeds less than the speed of light. Using,

$$\cosh^2\theta - \sinh^2\theta = 1$$

$$\tanh\theta = \frac{\sinh\theta}{\cosh\theta},$$

it follows that,

$$\cosh\theta = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \gamma$$

$$\sinh\theta = \gamma\frac{v}{c}.$$

We may then write the special Lorentz transformation as,

$$dx' = dx\cosh\theta - cdt\sinh\theta$$

$$cdt' = -dx\sinh\theta + cdt\cosh\theta,$$

or,

$$\begin{pmatrix} dx'^1 \\ dx'^4 \end{pmatrix} = \begin{pmatrix} \cosh\theta & -\sinh\theta \\ -\sinh\theta & \cosh\theta \end{pmatrix} \begin{pmatrix} dx^1 \\ dx^4 \end{pmatrix}.$$

For special Lorentz transformations, the transformation matrix has the appearance of a rotation transformation matrix with hyperbolic functions replacing their trigonometric counterparts.

The rapidities have an important characteristic that makes them very useful. Consider a primed frame, a double primed frame and an unprimed frame that are originally identical. Then, boost the primed frame and double primed frame together to a velocity v' with respect to the first. Finally, boost the double primed frame to a velocity v'' with respect to the primed frame. The overall Lorentz transformation that relates the double primed frame to the unprimed frame is obtained,

$$L = L''L' = \begin{pmatrix} \cosh\theta'' & -\sinh\theta'' \\ -\sinh\theta'' & \cosh\theta'' \end{pmatrix} \begin{pmatrix} \cosh\theta' & -\sinh\theta' \\ -\sinh\theta' & \cosh\theta' \end{pmatrix}$$

$$= \begin{pmatrix} \cosh\theta''\cosh\theta' + \sinh\theta''\sinh\theta' & -\cosh\theta''\sinh\theta' - \cosh\theta'\sinh\theta'' \\ -\cosh\theta''\sinh\theta' - \cosh\theta'\sinh\theta'' & \cosh\theta''\cosh\theta' + \sinh\theta''\sinh\theta' \end{pmatrix}.$$

Thus,

$$L = \begin{pmatrix} \cosh\theta & -\sinh\theta \\ -\sinh\theta & \cosh\theta \end{pmatrix} = \begin{pmatrix} \cosh(\theta'+\theta'') & -\sinh(\theta'+\theta'') \\ -\sinh(\theta'+\theta'') & \cosh(\theta'+\theta'') \end{pmatrix}.$$

Thus, rapidities have a much simpler addition property than relativistic velocities. Rapidities add like Galilean velocities!

A BRIEF TENSORIAL TUTORIAL

We noted in passing that the special Lorentz transformation,

$$dx' = \gamma(dx - vdt)$$

$$dy' = dy$$

$$dz' = dz$$

$$dt' = \gamma(dt - \frac{vdx}{c^2}),$$

can be compared to the chain rule result,

$$dx'_i = \frac{\partial x'_i}{\partial x_j} dx_j + \frac{\partial x'_i}{\partial t} dt$$

$$dt' = \frac{\partial t'}{\partial x_j} dx_j + \frac{\partial t'}{\partial t} dt.$$

By comparison of the two sets of expressions and defining $dx_1 = dx^1, dx_2 = dx^2, dx_3 = dx^3, cdt = dx^4$ we may write the Lorentz transformation as, $dx^{\mu'} = L^{\mu'}_{\nu} dx^{\nu}$. There is an implied sum on the repeated index, ν. A four-vector is defined to be a quartet that transforms in this way. In fact, this pattern must be distinguished from a closely related pattern, so we will call it the *contravariant* transformation pattern. We will distinguish quartets that transform according to this pattern by using superscripts and we will refer to the quartet as the *contravariant components* of the four-vector. The positioning of the primes tells us that the elements of the matrix L are given by,

$$L^{\mu'}_{\nu} = \frac{\partial x^{\mu'}}{\partial x^{\nu}}.$$

Now, we could have done the transformation the other way,

$$dx = \gamma(dx' + vdt')$$

$$dy = dy'$$

$$dz = dz'$$

$$dt = \gamma(dt' + \frac{vdx'}{c^2}),$$

and, the chain rule,

$$dx_i = \frac{\partial x_i}{\partial x'_j} dx'_j + \frac{\partial x_i}{\partial t'} dt'$$

$$dt = \frac{\partial t}{\partial x'_j} dx'_j + \frac{\partial t}{\partial t'} dt'.$$

If we compare these expressions, we obtain the elements of the inverse Lorentz transformation, L^{-1}

$$(L^{-1})^{\mu}_{\nu'} = \frac{\partial x^{\mu}}{\partial x'^{\nu}}.$$

We shall usually let the position of the primes alone indicate that this is the inverse matrix and write, $LL^{-1} = I$ and $L^{-1}L = I$ as

$$L^{\mu'}_{\nu} L^{\nu}_{\lambda'} = \delta^{\mu}_{\lambda},$$

$$L^{\mu}_{\nu'}L^{\nu'}_{\lambda} = \delta^{\mu}_{\lambda}.$$

The symbol, δ^{μ}_{λ}, is the Kronecker delta and it represents the elements of the identity matrix. Note again, that a repeated index indicates a sum. In virtually all cases in expressions that follow, the repeated index will appear once as a subscript and once as a superscript as it does here. This is part of the pattern that gives elegance to the formalism.

The chain rule yields another very important result. Think of a function, $f = f(x_i,t)$. Then,

$$\frac{\partial f}{\partial x'_i} = \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial x'_i} + \frac{\partial f}{\partial t}\frac{\partial t}{\partial x'_i} = \left[\frac{\partial x_j}{\partial x'_i}\frac{\partial}{\partial x_j} + \frac{\partial t}{\partial x'_i}\frac{\partial}{\partial t}\right] f.$$

We may think of this relationship defining a transformation property for the derivative operation, namely,

$$\frac{\partial}{\partial x'^{\mu}} = L^{\lambda}_{\mu'}\frac{\partial}{\partial x^{\lambda}}.$$

Observe that this is not the contravariant pattern (the primes on L are in the wrong place.) Instead, we will write this expression as,

$$\partial_{\mu'} = L^{\lambda}_{\mu'}\partial_{\lambda}.$$

This transformation pattern is called the *covariant* transformation pattern. Quartets which transform in this way are said to be the *covariant components* of a four-vector and they are labelled with subscripts rather than superscripts.

Similarly, 4-tensors have contravariant and covariant components as well. They transform according to the patterns:

$$A'^{\alpha\beta} = L^{\alpha'}_{\mu}L^{\beta'}_{\nu}A^{\mu\nu}$$

$$A'_{\alpha\beta} = L^{\mu}_{\alpha'}L^{\nu}_{\beta'}A_{\mu\nu}.$$

Exercise: Demonstrate explicitly that

$$F^{\mu\nu}U_{\nu} = A^{\mu}$$

is a form-invariant equation under Lorentz transformations.

RELATIVISTIC FOUR-VECTORS

Four-vectors play a very important role in relativity for the following reasons:

1. All four-vectors share the same Lorentz transformation pattern from one frame to another. If you know a thing is a four-vector, you immediately know how it transforms from frame to frame.
2. The generalized scalar dot product of two four-vectors is an invariant. If you know the magnitude in one frame, you know it immediately in all frames. This property is useful if the invariant is very simply evaluated in one frame, but much more complicated in another.

3. If we wish our “laws of physics” to have the same form (covariance) in all inertial frames, we must look for laws that express relationships among four-vectors. The Lorentz transformation will transform such relationships so that they have the same form in all admissable frames.

At this point in our treatment, however, we have encountered only one four-vector, (dx^1, dx^2, dx^3, cdt). We shall see below how to create other useful four-vectors beginning with this one.

If we define $d\mathbf{s}$ to be a separation between events, we can generalize a result from differential geometry, namely,

$$ds^2 = g_{ij} dq^i dq^j \to ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta.$$

This pattern to create $d\mathbf{s} \cdot d\mathbf{s}$ serves as a generalization of a scalar dot product. Scalar dot products are scalar invariants under rotations. In four-space, the thing which is invariant under the special Lorentz transformation is,

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2 = dx'^2 + dy'^2 + dz'^2 - c^2 dt'^2.$$

You can verify that this is true by direct substitution of the special Lorentz transformation.

Exercise: Demonstrate the invariance of ds^2 under the special Lorentz transformation.

Therefore, to make the formalism work, we must define our *metric tensor* (in Cartesian coordinates),

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Then we can write,

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta$$

in analogy to the form taken from differential geometry. Indeed, we can use curvilinear coordinates rather than Cartesian coordinates. For spherical coordinates, the metric tensor would be,

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & r^2 & 0 & 0 \\ 0 & 0 & r^2 \sin^2\theta & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

However, one familiar characteristic of ds^2 that we give up by this generalization is that, unlike the case in three-dimensional differential geometry, we may (and usually do) have $ds^2 < 0$! The function ds^2 is an invariant. It has the same form, $ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta$, and same value in all admissable frames.

Observe that the inner product of two four-vectors can be written,

$$g_{\mu\nu} A^\mu B^\nu = A^\mu B_\mu = A_\nu B^\nu.$$

The practical consequence of the metric tensor for Cartesian four-vectors is to associate a negative sign with the fourth-component term in the sum,

$$g_{\mu\nu}A^{\mu}B^{\nu} = A^1B^1 + A^2B^2 + A^3B^3 - A^4B^4 = A_1B_1 + A_2B_2 + A_3B_3 - A_4B_4.$$

In the rest frame of a moving clock,

$$d\mathbf{s} = (0,0,0,cdt') \equiv (0,0,0,cd\tau).$$

Thus, $ds^2 = -c^2d\tau^2$ is the invariant "squared length" of $d\mathbf{s}$. The squared proper time, $d\tau^2$, being equal to the invariant ds^2 divided by a constant of nature, c, is also an invariant. If we attach a frame of reference to a moving particle, its velocity in the unprimed system and the velocity of the primed frame are the same, $u = v$. In the laboratory frame, $d\mathbf{s} = (dx,0,0,cdt)$, so that

$$ds^2 = dx^2 - c^2dt^2 = -c^2dt^2(1 - \frac{1}{c^2}(\frac{dx}{dt})^2) = -\frac{c^2dt^2}{\gamma^2}.$$

By the invariance property, the value in the laboratory frame is the value in the rest frame, so,

$$-c^2d\tau^2 = -\frac{c^2dt^2}{\gamma^2},$$

or,

$$d\tau = \frac{dt}{\gamma}.$$

We have thus recovered the result called "time dilation."

A four-vector divided by a scalar invariant is a four-vector because it has the transformation pattern of a four-vector. We can create a new four-vector by dividing $d\mathbf{s}$ by $d\tau$. The resulting four-vector is called the *four-velocity* and has components,

$$U^{\mu} = \frac{dx^{\mu}}{d\tau}.$$

We then have,

$$\frac{dx^1}{d\tau} = \frac{dx^1}{dt/\gamma} = \gamma\frac{dx^1}{dt} = \gamma u_x,$$

$$\frac{dx^4}{d\tau} = \frac{cdt}{dt/\gamma} = \gamma c.$$

We might also write the four-velocity as,

$$U^{\mu} = (\gamma u_x, \gamma u_y, \gamma u_z, \gamma c) = (\gamma\mathbf{u}, \gamma c) = (\hat{\mathbf{n}}c\sinh\theta, c\cosh\theta),$$

where $\hat{\mathbf{n}}$ is a unit vector tangent to the particle motion in three-space.

Because U is a four-vector, we immediately know that it transforms from frame to frame according to the pattern, $U^{\mu'} = L^{\mu'}_{\nu}U^{\nu}$. We also know that its "squared length,"

$g_{\mu\nu}U^{\mu}U^{\nu}$, is an invariant. Its value is most easily obtained in the rest system of the particle where $u = 0$ and $\gamma = 1$. Then, $U = (0,0,0,c)$ and $g_{\mu\nu}U^{\mu}U^{\nu} = -c^2$.

Like proper time, the proper length and the proper mass are invariants. We may multiply U by the proper mass, m_0, to create yet another four-vector, the *four-momentum*,$P^{\mu} = m_0U^{\mu} = (m_0\gamma\mathbf{u}, m_0\gamma c)$. We immediately know that this four-vector transforms according to the usual pattern and we know that it has an invariant squared length,

$$g_{\mu\nu}P^{\mu}P^{\nu} = m_0^2(g_{\mu\nu}U^{\mu}U^{\nu}) = -m_0^2c^2.$$

It is customary to combine the factor of γ and the proper mass to create a *relativistic mass*, $m \equiv \gamma m_0$ and a *relativistic momentum*, $\mathbf{p} \equiv \gamma m_0\mathbf{u} \equiv m\mathbf{u}$. We may also define a *relativistic energy*, $E \equiv \gamma m_0c^2 \equiv mc^2$. Then, the four-momentum in an arbitrary frame can be written, $P^{\mu} = (\mathbf{p}, E/c)$. Using the invariance property of the squared length, we obtain the useful relationship,

$$p^2 - \frac{E^2}{c^2} = -m_0^2c^2,$$

or,

$$E^2 = p^2c^2 + m_0^2c^4.$$

Exercise: Derive this latter result, taking care to show that the algebraic signs are handled correctly.

We may create a *four-acceleration* by differentiating the four-velocity with respect to proper time, $A^{\mu} = \dot{U}^{\mu}$. To maintain form invariance (covariance) under Lorentz transformations, the laws of physics must be expressed as relationships among four-vectors and, possibly, four-tensors. With m_0A^{μ}, we have a four-vector with which we can build a covariant form of Newton's Second Law. But, first, we need to develop four-vector forces.

We can obtain another important four-vector from a different line of reasoning. Consider a plane electromagnetic wave propagating in space. We may represent the fields of the electromagnetic wave as,

$$\mathbf{A} = \mathbf{A}_0 \cos(\mathbf{k}\cdot\mathbf{x} - \omega t).$$

A represents either the electric field, **E**, or the magnetic field, **B**, associated with the electromagnetic wave. The vector **k** is the *wave number* defined by

$$\mathbf{k} = \frac{2\pi}{\lambda}\hat{\mathbf{n}},$$

where $\hat{\mathbf{n}}$ is a unit vector in the direction of propagation and λ is the wavelength of the wave. The magnitude of **k** is the number of waves in 2π units of length. Since it depends on a measured length, the Lorentz contraction implies that **k** is not a relativistic invariant. Similarly, the frequency, ω, of the wave is a certain number of radians in a measured period of time and because of time dilation is also not a relativistic invariant.

The quantity $\mathbf{k}\cdot\mathbf{x} - \omega t$ is called the *phase* of the wave. It is a dimensionless quantity and is unaffected by either Lorentz contraction or time dilation. The phase is a function of a position and a time. We might think of an observer making an observation of the wave at $(\mathbf{x}, t)$ thus defining an event in space time. The primed coordinates of this same event are $(\mathbf{x}', t')$. The phase of the wave determines whether the observers are observing the peak, the trough, or some other part of the cycle of the wave. But, since they will both agree that they are observing the same part of the wave cycle, the phase must be the same for both. The phase is a relativistic invariant, which means that,

$$\mathbf{k}' \cdot \mathbf{x}' - \omega' t' = \mathbf{k} \cdot \mathbf{x} - \omega t.$$

For simplicity, consider a special Lorentz transformation and an electromagnetic wave propagating along the common x, x'-axes. Then,

$$k_x' dx' - \omega' dt' = k_x' [\gamma(dx - v dt)] - \omega' [\gamma(dt - \frac{v dx}{c^2})]$$

$$= \gamma(k_x' + \frac{v\omega'}{c^2}) dx - \gamma(\omega' + k_x' v) dt$$
$$= k_x dx - \omega dt.$$

If the phase is an invariant, we must have,

$$k_x = \gamma(k_x' + \frac{v\omega'}{c^2})$$

$$\frac{\omega}{c} = \gamma(\frac{\omega'}{c} + \frac{k_x' v}{c}).$$

But, this is just the pattern for the special Lorentz transformation of a four-vector, $k^\mu = (\mathbf{k}, \omega/c)$. We will call this four-vector the *four-wavenumber*.

Two important results follow immediately from the Lorentz transformation of this four-vector. Imagine electromagnetic waves that are observed approaching an observer at an angle, α, in the x, y-plane. Electromagnetic waves have an especially simple *dispersion relation*, that relates the wave number and the frequency, namely $k = \omega/c$. Because the four-wavenumber is a four-vector, we know immediately how it transforms,

$$\begin{pmatrix} k' \cos\alpha' \\ k' \sin\alpha' \\ 0 \\ \omega'/c \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\frac{v}{c}\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{v}{c}\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} k \cos\alpha \\ k \sin\alpha \\ 0 \\ \omega/c \end{pmatrix}.$$

In particular,

$$k' \cos\alpha' = \gamma k \cos\alpha - \frac{v\gamma}{c}\frac{\omega}{c}$$

$$\frac{\omega'}{c} = -\frac{v}{c}\gamma k \cos\alpha + \gamma\frac{\omega}{c}.$$

Using the dispersion relationship for electromagnetic waves, the second of these can be written,

$$\omega' = \gamma\omega(1 - \frac{v}{c}\cos\alpha).$$

This relationship is the *relativistic Doppler shift*. It tells what the frequency of electromagnetic waves is in a primed frame if those same waves are observed with frequency, ω, in the laboratory and arriving at angle, α.

The second important relationship comes from dividing the first of the transformation equations by the second and using the fact that $k' = \omega'/c$ and $k = \omega/c$. We have then,

$$\cos\alpha' = \frac{\cos\alpha - v/c}{1 - (v/c)\cos\alpha}.$$

This is the *relativistic aberation formula* and gives the angle of approach, α', in the primed system in terms of the angle in the unprimed system.

THE FOUR-VECTOR FORMULATION OF MAXWELL'S EQUATIONS

To develop a Lorentz form-invariant replacement for Newton's Second Law, we need a four-vector expression for force. Of the four fundamental forces, the strong and weak forces are quantum mechanical and do not fall in the domain of phenomena described by Newton's Second Law, although the quantum field theories that treat them are usually expressed in relativistically covariant form. Gravity, on the other hand, is treated in an extension of the Special Theory of Relativity called *General Relativity*. The laws of General Relativity are expressed in covariant form, but the treatment goes beyond what we can do here.

We are left with the electromagnetic force. It is the force exerted on charged particles by electric and magnetic fields. In nonrelativistic mechanics, it is expressed as a 3-vector, called the *Lorentz Force*,

$$\mathbf{F} = q(\mathbf{E} + \frac{1}{c}\mathbf{u} \times \mathbf{B}).$$

Since the force depends on a combination of electric and magnetic fields, as well as the charge, q, and velocity, $\mathbf{u}$, of a particle, it is a somewhat more difficult task than we have yet encountered to express the electromagnetic force in four-vector form.

Maxwell's equations express the relationship between electromagnetic fields and the sources (charge and current) that generate them. For our purposes, it is not necessary to understand many of the applications of Maxwell's equations to electromagentic phenomena. What we shall show is that Maxwell's equations are already fully consistent with Special Relativity. We do this by showing how Maxwell's equations can be written down in terms of four-vectors and four-tensors, thus ensuring that the equations have exactly the same form in all coordinate systems of Special Relativity. We shall also show how the Lorentz Force is expressed as part of a four-vector. With this force as an example, we will then show how Newton's Second Law and Newtonian dynamics are incorporated into Special Relativity. Finally, we will conclude with some examples.

The sources of the electromagnetic field are stationary or moving charges. Here we must add another postulate to our theory of Special Relativity:

4. **Postulate of Charge Invariance**: The electric charge on an elementary particle is an absolute invariant.

The charge density, on the other hand, is the amount of charge per unit volume,

$$\rho = \frac{Nq}{Adx}$$

where, N is the total number of charged particles, q is the charge on each particle, $A = dydz$ is an area that is invariant under the special Lorentz transformation, and dx is a differential of distance along the common x, x'-axes. For a special Lorentz transformation, $dy = dy'$ and $dz = dz'$ so that $A = A'$. However, $dx = \gamma(dx' + vdt')$ so that charge density is not relativistically invariant. Neverthless, like proper length, proper time and proper mass, the proper density, ρ_0, is a relativistic scalar and can be used to create a new four-vector (called the *four-current*),

$$J^\mu = \rho_0 U^\mu = (\rho_0\gamma\mathbf{u}, \rho_0\gamma c).$$

As we have done to create relativistic mass, it is customary to define the *relativistic charge density* as $\rho \equiv \gamma\rho_0$ so that $J^\mu = (\rho\mathbf{u}, \rho c) = (\mathbf{j}, \rho c)$. Thus, we have defined the *relativistic current*, $\mathbf{j}$, with units of charge per unit area per unit time. In this way we have created a four-vector that describes the sources of the electromagnetic field.

The electromagnetic fields, **E** and **B**, do not individually extend naturally to become parts of four-vectors. Indeed, it is one of the surprising insights of Special Relativity that the six components of these 3-vectors are elements of a single electromagnetic field tensor. In a sense, electric and magnetic fields are not different things, but rather equivalent parts of the same thing. What one observer reckons to be a pure electric field, is reckoned by another observer in a different frame to be a combination of electric and magnetic fields!

This time we are going to do it backwards. We are not going to start with electromagnetic experiments and develop equations to describe them, which, after much work are to be shown to be relativistically consistent. Rather, we are going to do a very remarkable thing. We are going to assume as a basic premise that the laws of physics must be expressed in relativistically form-invariant equations and let that assumption tell us essentially what the laws of electrodynamics must be. We will have to make a few "judicious choices" along the way so that the laws of physics come out in their familiar form, but these are more convention than anything else.

We begin by assuming that there exists a four-vector, $A^\mu = (\mathbf{A}, \phi/c)$. The factor $1/c$ is not essential, but including it here makes it easier to connect our results with the familiar equations of electrodynamics. The four-vector A is generated in some way by the sources that we have described by the four-current, J. The four-vector A has only four components, so it alone cannot incorporate the six components of **E** and **B**. Therefore, we must use one of the methods we have identified for generating four-tensors from four-vectors.

We can generate a four-tensor by differentiating a four-vector. For example, $T_{\mu\nu} = \partial_\mu A_\nu$, but the result is a tensor that has sixteen components. We only need six slots for our electromagnetic fields. However, we could add two tensors and create another tensor. Consider, $T_{\mu\nu} = \partial_\mu A_\nu + \partial_\nu A_\mu$. This tensor is symmetric because $T_{\mu\nu} = T_{\nu\mu}$. Symmetric tensors have ten independent components which you can see if you write down a 4×4 symmetric matrix of the elements. But then we have ten slots for our six components of electromagnetic field. Finally, consider, $T_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. This tensor is anti-symmetric because $T_{\mu\nu} = -T_{\nu\mu}$. This means that the four diagonal elements of its matrix must be zero. The anti-symmetric tensor has six independent elements and that is just what we need! So, we define the covariant components of an antisymmetric electromagnetic field tensor,

$$T_{\mu\nu} = \begin{pmatrix} 0 & B_3 & -B_2 & E_1 \\ -B_3 & 0 & B_1 & E_2 \\ B_2 & -B_1 & 0 & E_3 \\ -E_1 & -E_2 & -E_3 & 0 \end{pmatrix}.$$

Arranging the electromagnetic fields in just this way is one of the "judicious choices" mentioned above. We have done it so that the four-vector equation, $T_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, is equivalent to the two vector equations,

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}$$

$$\mathbf{B} = \nabla \times \mathbf{A}.$$

These relationships are familiar in electrodynamics. Ordinarily, they are used to define the *vector potential*, $\mathbf{A}$, and scalar potential, ϕ, by defining their relationships to the more familiar electrical and magnetic fields. In our approach here, we introduced the four-vector $A^\mu = (\mathbf{A}, \phi/c)$ first. By arranging the components of the electric and magnetic fields in the field tensor as we did, we identify that our assumed $\mathbf{A}$ is the vector potential of conventional electrodynamics and that ϕ is the scalar potential of conventional electrodynamics.

We can also "raise the indices" to create the contravariant components of the electromagnetic field tensor, $T^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} T_{\alpha\beta}$. Then,

$$T^{\mu\nu} = \begin{pmatrix} 0 & B_3 & -B_2 & -E_1 \\ -B_3 & 0 & B_1 & -E_2 \\ B_2 & -B_1 & 0 & -E_3 \\ E_1 & E_2 & E_3 & 0 \end{pmatrix}.$$

Exercise: Derive the expression for $T^{\mu\nu}$ by raising the indices of $T_{\alpha\beta}$.

With a little patience you can show that the following two form-invariant equations,

$$\partial_\mu T^{\nu\mu} = \frac{4\pi}{c}\rho_0 U^\nu$$

$$\partial_\mu T_{\nu\lambda} + \partial_\lambda T_{\mu\nu} + \partial_\nu T_{\lambda\mu} = 0$$

are completely equivalent to Maxwell's Equations in Gaussian units,

$$\nabla \cdot \mathbf{E} = 4\pi\rho$$

$$\nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j}$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0.$$

The first two of Maxwell's equations have sources on the right as does the first of the two relativistic equations. The second two of Maxwell's equations and the second of the relativistic equations are equal to zero on the right.

Exercise: Demonstrate how $\nabla \cdot \mathbf{B} = 0$ follows from the four-vector form of Maxwell's equations.

It follows that the Lorentz force, being a function of fields and velocity, can be written as some combination of the field tensor and the four-velocity. It can be readily verified that the expression,

$$K^{\mu} = \frac{q}{c} T^{\mu\nu} U_{\nu} = (\gamma \mathbf{F}, \frac{1}{c}\gamma \mathbf{F} \cdot \mathbf{u}),$$

is the four-vector that contains the Lorentz force, **F**. The four-force, K^{μ}, is sometimes called the *Minkowski force*. Therefore, the relativistic version of Newton's Second Law in those instances where the force is electromagnetic must be,

$$\frac{q}{c} T^{\mu\nu} U_{\nu} = m_0 \frac{dU^{\mu}}{d\tau}.$$

Exercise: Verify the relationship between the Minkowski force and the Lorentz force by expanding $T^{\mu\nu}U_{\nu}$.

RELATIVISTIC PARTICLE MECHANICS

Special Relativity deals with frames of reference that move relative to one another at constant velocity. Occasionally, the newcomer to the subject will have the mistaken idea that Special Relativity only deals with constant velocities, while General Relativity is the extension of Special Relativity to deal with accelerated motions. This is a misconception. General Relativity is an extension of Special Relativity to deal with the force of gravity. The only restriction of Special Relativity is that motion, including accelerated motion, be described from a frame which is at rest or in uniform motion. Many of the applications of Special Relativity are to the acceleration of charged particles in electromagnetic fields, either in accelerators or elsewhere in the cosmos.

Consider a charged particle moving with velocity **u** in the laboratory frame. We may write the four-vector law of motion as,

$$K^{\mu} = (\gamma \mathbf{F}, \frac{1}{c}\gamma \mathbf{F} \cdot \mathbf{u}) = m_0 \frac{dU^{\mu}}{d\tau}.$$

Since $d\tau = dt/\gamma$, we may write this as,

$$\left(\gamma \mathbf{F}, \frac{1}{c}\gamma \mathbf{F} \cdot \mathbf{u}\right) = m_0 \gamma \left(\frac{d}{dt}(\gamma \mathbf{u}), \frac{d}{dt}(\gamma c)\right).$$

Thus,

$$\left(\mathbf{F}, \frac{1}{c}\mathbf{F}\cdot\mathbf{u}\right) = \left(\frac{d}{dt}(\gamma m_0\mathbf{u}), \frac{d}{dt}(\gamma m_0 c)\right) = \left(\frac{d\mathbf{p}}{dt}, \frac{1}{c}\frac{dE}{dt}\right),$$

where, **p** is the relativistic momentum and E is the relativistic energy of the particle. We may separate the single four-equation into two,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}$$

$$\mathbf{F}\cdot\mathbf{u} = \frac{dE}{dt}.$$

The first is a relativistic statement of Newton's Second Law, except that **p** is the relativistic momentum rather than the Newtonian momentum. The second of these two equations is a statement that the power, $\mathbf{F}\cdot\mathbf{u}$, delivered to the particle by the force, **F**, equals the rate of increase of the total relativistic energy of the particle. The total relativistic energy, $E = \gamma m_0 c^2$, is the sum of the kinetic energy of the particle plus its rest mass energy, $m_0 c^2$.

There is a peculiar consequence of the relativistic Second Law.

Exercise: Show that

$$\frac{d\gamma}{dt} = \frac{\gamma^3}{c^2}u\frac{du}{dt}.$$

Exercise: Show that

$$\gamma + \gamma^3\frac{u^2}{c^2} = \gamma^3.$$

We have then that,

$$\frac{d}{dt}(\gamma\mathbf{u}) = \gamma\frac{d\mathbf{u}}{dt} + \mathbf{u}\frac{d\gamma}{dt} = \gamma\mathbf{a} + \mathbf{u}\frac{\gamma^3}{c^2}u\frac{du}{dt} = \gamma(\mathbf{a}_\perp + \mathbf{a}_\|) + \mathbf{u}\frac{\gamma^3}{c^2}u\frac{du}{dt},$$

where $\mathbf{a}_\perp$ is the component of acceleration perpendicular to the direction of motion and $\mathbf{a}_\|$ is the component of acceleration parallel to the direction of motion.

Now, $\mathbf{u}\cdot\mathbf{u} = u^2$. If we differentiate,

$$\mathbf{u}\cdot\mathbf{a} = u\frac{du}{dt} = ua_\|.$$

Because **u** and $\mathbf{a}_\|$ are in the same direction, we may use that fact that $\mathbf{u}a_\| = u\mathbf{a}_\|$ to write,

$$\frac{d}{dt}(\gamma\mathbf{u}) = \gamma\mathbf{a}_\perp + (\gamma + \frac{\gamma^3 u^2}{c^2})\mathbf{a}_\| = \gamma\mathbf{a}_\perp + \gamma^3\mathbf{a}_\|.$$

Thus,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = m_0\gamma\mathbf{a}_\perp + m_0\gamma^3\mathbf{a}_\| \equiv m_\perp\mathbf{a}_\perp + m_\|\mathbf{a}_\|.$$

The particle behaves as if it has two different masses, one for parallel acceleration and one for perpendicular acceleration! It is more difficult to accelerate a particle by an electromagnetic force along its direction of motion than it is to accelerate it perpendicular to its motion.

THE TWIN PARADOX

Imagine twins, Albert and Henri, who part one day for separate, extended voyages in spacetime. Let Albert stay behind and remain at the origin of his stationary, unprimed reference frame, but let Henri move away in a rocket ship which defines the origin of his primed frame. After some time, the two meet again and compare their ages (clocks). Let us assume that they meet where they first start with the origins of the two systems once again coincident. The parting and the reuniting define two events in space time for which $dx = dx' = 0$. Thus, the invariant $ds^2 = -c^2 d\tau^2$. But which clock has actually measured the proper time? From Albert's point of view, his clock has remained stationary and has measured the proper time, while Henri's clock has moved and measured a dilated time. But from Henri's point of view, it is Henri's clock that has remained stationary, while Albert and his clock have moved and have measured a dilated time. Do their clocks agree or not? How can they agree if each thinks the other has measured a dilated time relative to his own? How can they disagree if there is symmetry in their points of view? We have here an apparent paradox, often called the *twin paradox*.

The paradox is resolved in the following way. First, the viewpoints of the twins are not necessarily symmetric. If the two are to be reunited, at least one of the pair must turn around, i.e., be accelerated during at least part of the journey. Special Relativity is a description of motion from a frame of reference that is at rest or in uniform motion. Thus, the unprimed frame of Albert is an admissable frame for using the relationships of Special Relativity, but Henri's accelerated frame is not during the period of acceleration. Thus the symmetry is broken. It is true that $ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta = g'_{\alpha\beta} dx'^\alpha dx'^\beta$, but the metric tensor in the primed frame is no longer a function of constant values of γ and v. Indeed, without some thought, it is not obvious what it should actually be.

Nevertheless, we can calculate a comparison of Albert's elapsed time relative to the proper time measured on Henri's clock if we use Special Relativity in Albert's frame. Let us assume that Henri's rocket ship is accelerated by some kind of constant force, $\mathbf{F} = m_0 \mathbf{g}$. Newton's relativistic Second Law is then,

$$\mathbf{F} = m_0 \mathbf{g} = \frac{d}{dt}(\gamma m_0 \mathbf{u}).$$

If we assume that the acceleration is one-dimensional along the common x, x'-axes, we have

$$g = \frac{d}{dt}(\gamma u).$$

In this instance we can make good use of the rapidity variable by using,

$$\sinh\theta = \gamma \frac{u}{c}$$

$$\cosh\theta = \gamma$$

$$\tanh\theta = \frac{u}{c}.$$

Then,

$$g = \frac{d}{dt}(\gamma u) = \frac{1}{\gamma}\frac{d}{d\tau}(\gamma u) = \frac{c}{\cosh\theta}\frac{d}{d\tau}(\cosh\theta \tanh\theta) = c\frac{d\theta}{d\tau}.$$

Assuming a boundary condition of $\theta = 0$ at $\tau = 0$, we can integrate to obtain,

$$\theta = \frac{g\tau}{c}.$$

The four-velocity yields,

$$U^{\mu} = \frac{dx^{\mu}}{d\tau} = \left(\frac{dx}{d\tau}, 0, 0, c\frac{dt}{d\tau}\right) = (c\sinh\theta, 0, 0, c\cosh\theta),$$

i.e.,

$$\frac{dx}{d\tau} = c\sinh(\frac{g\tau}{c})$$

$$\frac{dt}{d\tau} = \cosh(\frac{g\tau}{c}).$$

Each can be integrated using appropriate boundary conditions to yield,

$$x_a = \frac{c^2}{g}\left(\cosh(\frac{g\tau}{c}) - 1\right)$$

$$t_a = \frac{c}{g}\sinh(\frac{g\tau}{c}).$$

Observe that the function for t_a is an even function of g, so that it serves as well for acceleration and deceleration.

If Henri accelerates to a terminal coasting speed, u_c, we have during the coasting period,

$$\frac{d}{dt}(\gamma u) = 0$$

from which $\gamma u = \text{constant} = \gamma_c u_c$. Then,

$$\sinh\theta_c = \frac{\gamma_c u_c}{c}$$

$$\cosh\theta_c = \gamma_c.$$

Again,

$$U^{\mu} = \left(\frac{dx}{d\tau}, 0, 0, c\frac{dt}{d\tau}\right) = (c\sinh\theta_c, 0, 0, c\cosh\theta_c),$$

from which,

$$x_c = c\tau\sinh(\frac{\gamma_c u_c}{c})$$

$$t_c = \tau\cosh(\frac{\gamma_c u_c}{c}).$$

By combining periods of acceleration, coasting, and deceleration, one can demonstrate that when Henri returns to meet with Albert, his (proper) time and age will be less than that of Albert who stayed behind.

Exercise: If g is chosen for comfort to be the local acceleration of gravity (which in truly unusual units is 1.03 light years/year/year), how far would Henri travel in 22 years of his own time? How much time will elapse on earth if Henri makes a round trip by accelerating for 6 years (proper time), then decelerating for six years and then reversing the process to come back? How much time will he add to the elapsed time on earth if he coasts for a year (proper time) on the way out and again for a year on the way back?

CONSERVATION LAWS

The force law,

$$K^\mu = m_0 \frac{dU^\mu}{d\tau}$$

tells us that for a particle,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}$$

$$\mathbf{F} \cdot \mathbf{u} = \frac{dE}{dt}.$$

In the absence of force, the four-velocity is a constant of the motion, i.e., $\mathbf{p}$ and E are constants of the motion. If u is modest in magnitude relative to c, we can expand γ,

$$\frac{1}{\sqrt{1-u^2/c^2}} = 1 + \frac{1}{2}\frac{u^2}{c^2} + \frac{3}{8}\frac{u^4}{c^4} + \ldots.$$

Hence,

$$E = m_0 c^2 + \frac{1}{2} m_0 u^2 + \frac{3}{8}\frac{m_0 u^4}{c^2} + \ldots.$$

The first term exists even when $u = 0$. We call $m_0 c^2$ the *rest energy* of the particle. The remaining terms, beginning with $1/2 m_0 u^2$, are associated with motion and together form the *relativistic kinetic energy*. Because relativistic kinetic energy is not directly identifiable with a component of a four-vector, it usually plays a less important role in relativistic mechanics than the relativistic momentum and the total relativistic energy.

It is possible within Special Relativity to have a particle for which $m_0 = 0$ if that particle moves at the speed of light. In such a case, both the numerator and the denominator of γm_0 become zero. However, if we let this happen in such a way that the total energy of the particle is finite, we have,

$$\gamma m_0 c^2 \rightarrow E$$

$$\gamma m_0 \rightarrow E/c^2$$

$$\mathbf{p} = \gamma m_0 u \hat{\mathbf{n}} \rightarrow (E/c^2)(c)\hat{\mathbf{n}} = E/c\hat{\mathbf{n}}.$$

The four momentum then becomes, $P^\mu = (E/c\hat{\mathbf{n}}, E/c)$. The photon of quantum mechanics is such a massless particle for which the total energy is $\hbar\omega$. The four-momentum of a photon is

$$P^\mu = (\frac{\hbar\omega}{c}\hat{\mathbf{n}}, \frac{\hbar\omega}{c}).$$

For a system of n particles, we define the total relativistic momentum of the system to be the sum of the relativistic momenta of the individual particles of the system,

$$\mathbf{P} \equiv \sum_{i=1}^{n} \mathbf{p}_i$$

and the total relativistic energy, E_t, to be the sum of the relativistic energies of the individual particles. If Newton's Third Law holds relativistically, this means that **P** and E_t are conserved in the absence of external forces.

The center of mass system is defined to be that (primed) system for which $\mathbf{P}' = 0$. The total relativistic three-momentum **P** and E_t/c are the elements of a four-vector and transform according to the pattern of four-vectors. This means that,

$$E_t = \gamma(E_t' + \mathbf{v} \cdot \mathbf{P}').$$

If we choose the primed system to be the center-of-mass system, $\mathbf{P}' = 0$, we have $E_t = \gamma E_t'$. If we now think of our system of particles as a body made up of the constituent particles, we can give it a proper mass, M_0,

$$E_t = \gamma M_0 c^2 = \gamma E_t',$$

from which,

$$M_0 = \sum_{i=1}^{n} \left(\frac{m_{0i}}{\sqrt{1 - u_i'^2/c^2}} + \frac{V_i'}{c^2} \right) = \sum_{i=1}^{n} m_{0i} c^2 + \sum_{i=1}^{n} \frac{T_i'}{c^2} + \sum_{i=1}^{n} \frac{V_i'}{c^2}.$$

Here, T_i' is the relativistic kinetic energy of the particles in the center-of-mass frame and V_i' represents the potential energies that the particles might have as a result of their interaction. Clearly, the proper mass of the body is more than the sum of the rest masses of the constituent particles. If the kinetic energy of the particles is increased by heating, for example, the mass M_0 of the body increases! Similarly, if the potential energy of the particles decreases as a result of their interaction, the mass of the body might even be less than the sum of the rest masses by an amount Δm called the *mass defect*. In any case, it is clear that mass and energy are not different things, but manifestations of the same thing, *mass-energy*.

USES OF INVARIANTS

Conservation laws tell us that certain quantities are the same at different times. Invariant quantities tell us that certain quantities are the same in different frames. Sometimes the two can work together advantageously.

Imagine a proton accelerator that collides protons on stationary targets. Imagine that the intent of the collision is to collide a proton on a proton target to produce a proton-antiproton pair from the kinetic energy of the accelerated proton. We want to know how much energy the initial proton must have if an proton and antiproton exist after collison in addition to the original two protons. In other words, we want to know the *threshold energy* for the process.

In this case we are not concerned with the dynamics of the interaction itself. We are just interested in a relationship between quantities before the collision (beam energy) and conditions after the collision (four masses). This is something about which conservation laws tell us. Experiments of this type are carried out in the laboratory frame, but they are usually most easily analyzed in the center-of-mass frame. Invariants tell us something about relationships between frames.

Both in the laboratory frame and in the center-of-mass frame, we have a "before" four-momentum and an "after" four-momentum, i.e., four different four-momentum vectors. The "before" four-momentum in the laboratory frame consists of the 3-momentum of the single moving proton as well as the total relativistic energies of the two protons,

$$P^{\mu}_{b,lab} = \left(p_{1b}, 0, 0, \frac{E_{p1} + m_p c^2}{c} \right).$$

The "after" four-momentum (for threshold) in the center-of-mass frame consists of four stationary particles clustered together. The four-momentum is,

$$P^{\mu}_{a,cm} = \left(0, 0, 0, \frac{4 m_p c^2}{c} \right).$$

Since there are no external forces on the system, four-momentum is conserved: $P^{\mu}_{b,cm} = P^{\mu}_{a,cm}$. Thus we have,

$$g_{\mu\nu} P^{\mu}_{b,lab} P^{\nu}_{b,lab} = g_{\mu\nu} P^{\mu}_{b,cm} P^{\nu}_{b,cm} = g_{\mu\nu} P^{\mu}_{a,cm} P^{\nu}_{a,cm}.$$

The first equality follows from invariance and the second follows from the conservation law. Skipping the intermediate step,

$$p_{1b}^2 - \frac{(E_{1b} + m_p c^2)^2}{c^2} = -\frac{(4 m_p c^2)^2}{c^2}.$$

However, we may also use the fact that the four-momentum of the original beam particle alone forms an invariant connecting the laboratory and the rest frame of that single particle,

$$p_{1b}^2 - \frac{E_{1b}^2}{c^2} = -\frac{m_p^2 c^4}{c^2}.$$

Combining these two relationships, we have

$$E_{1b} = 7 m_p c^2.$$

If we add the rest energy of the stationary target, we have $8 m_p c^2$ in the laboratory frame, of which $4 m_p c^2$ is available in the center-of-mass frame to produce particles.

How much energy would we have in the center-of-mass frame if we accelerated both protons to an energy of $7 m_p c^2$ and figured out a way to collide them head-on. In this case, the laboratory frame becomes the center-of-mass frame and all $14m_pc^2$ is available to produce particles, i.e. 3.5 times as much as before. In terms of four-vectors,

$$P^{\mu}_{b,lab} = \left(0,0,0,\frac{2E_{1b}}{c}\right)$$

$$P^{\mu}_{a,cm} = \left(0,0,0,\frac{E_{cm}}{c}\right),$$

from which $E_{cm} = 14m_pc^2$. For this reason, many of the modern accelerators are designed as colliders rather than stationary target accelerators, although there are other tradeoffs, including the rate of collision events, that also must be considered.

As another example, consider the decay of a neutron. A free neutron decays into a proton, an electron and an antineutrino with a proper mean lifetime of 14.8 minutes. Suppose we want to know the maximum possible energy of the ejected electron. We shall suppose that in this case the antineutrino carries away negligible energy and momentum. In this special circumstance, it is as if the antineutrino doesn't exist.

Let the neutron be at rest, so that the laboratory frame and the center-of-mass frame are the same. In the absence of external forces, the four-momentum of the system is conserved,

$$(0,0,0,m_nc)_{b,cm} = (\mathbf{p}_e + \mathbf{p}_p,(E_e+E_p)/c)_{a,cm} = (p_e - p_p,0,0,(E_e+E_p)/c)_{a,cm}$$

and has an invariant scalar product. Thus, we have,

$$p_e = p_p$$

$$m_nc^2 = E_e + E_p$$

$$(p_e - p_p)^2 - \frac{(E_e+E_p)^2}{c^2} = -m_n^2c^2.$$

The four-momenta of the electron and proton taken separately have invariant scalar products,

$$p_e^2 - (\frac{E_e}{c})^2 = -m_e^2c^2$$

$$p_p^2 - (\frac{E_p}{c})^2 = -m_p^2c^2.$$

Combining these conservation and invariance relationships,

$$E_{e,max} = \frac{(m_n^2 - m_p^2 + m_e^2)c^2}{2m_n}.$$

Since the neutron was assumed to be at rest, we can take this result to be a center-of-mass result. If the neutron were moving along the x-axis of the laboratory with velocity v, we could transform the result back to the laboratory frame to obtain a result for the maximum energy of an electron resulting from the decay of a moving neutron.

Let us imagine now that the electron is emitted at an angle θ relative to the x-axis in the laboratory system. We want to know what $E_{e,max}$ is in the laboratory as a function of θ. The value for $E_{e,max}$ already obtained in the center-of-mass frame now becomes $E'_{e,max}$. Since energy is an element of a four-vector, we know how it transforms,

$$\frac{E'_e}{c} = \gamma\left(\frac{E_e}{c} - \frac{v}{c^2}p_e\cos\theta\right)$$

or,

$$\frac{E_e}{c} = \gamma\left(\frac{E'_e}{c} + \frac{v}{c^2}p'_e\cos\theta'\right).$$

Because we want an answer in terms of the laboratory angle, we use the former of the two. We eliminate p_e by using the invariant scalar product relationship for the electron,

$$p_e^2 - (E_e^2/c^2) = -m_e^2c^2.$$

With p_e eliminated, the transformation becomes a quadratic equation in E_e, for which, after some algebraic simplification,

$$E_{e,max} = \frac{E'_{e,max} + \frac{v}{c}\cos\theta\sqrt{E'^2_{e,max} - \gamma^2 m_e^2 c^4\left(1 - \frac{v^2}{c^2}\cos^2\theta\right)}}{\gamma\left(1 - \frac{v^2}{c^2}\cos^2\theta\right)}.$$

As a final example of this type, let us consider a phenomenon called *Compton scattering*. A photon, originally moving along the x-axis scatters from a stationary electron. As a result, the electron recoils at an angle ϕ relative to the x-axis and the photon, with an altered wavelength and energy, is scattered to an angle θ relative to the x-axis. We are interested in the change in wavelength of the photon as a function of the angle of scatter, θ.

In this instance, nothing is gained by working in the center-of-mass frame and all calculations are done directly in the laboratory frame. We have a four-momentum for the original photon before the scattering,

$$P^\mu_{b,\omega} = \left(\frac{\hbar\omega}{c}, 0, 0, \frac{\hbar\omega}{c}\right),$$

for which $g_{\mu\nu}P^\mu_{b,\omega}P^\nu_{b,\omega} = 0$. There is also a total four-momentum before the scattering,

$$P^\mu_{b,tot} = P^\mu_{b,e} + P^\mu_{b,\omega} = \left(\frac{\hbar\omega}{c}, 0, 0, \frac{\hbar\omega}{c} + m_e c\right).$$

There is a total four-momentum after the scattering, $P^\mu_{a,e} + P^\mu_{a,\omega^*}$,

$$P^\mu_{a,tot} = \left(\frac{\hbar\omega^*}{c}\cos\theta + p_e\cos\phi, \frac{\hbar\omega^*}{c}\sin\theta - p_e\sin\phi, 0, \frac{\hbar\omega^*}{c} + \frac{E_e}{c}\right).$$

The scalar product of each with itself is an invariant. Conservation of four-momentum tells us that $P^{\mu}_{b,tot} = P^{\mu}_{a,tot}$, so that,

$$g_{\mu\nu} P^{\mu}_{b,tot} P^{\nu}_{b,tot} = \left(\frac{\hbar\omega}{c}\right)^2 - \left(\frac{\hbar\omega}{c} + m_e c\right)^2 = g_{\mu\nu} P^{\mu}_{a,tot} P^{\nu}_{a,tot}.$$

We simplify the "after" expression by systematically eliminating four-vectors that contain information about the scattered electron for which we are not presently concerned. Hence,

$$\begin{aligned} g_{\mu\nu} P^{\mu}_{a,tot} P^{\nu}_{a,tot} &= g_{\mu\nu}(P^{\mu}_{a,e} + P^{\mu}_{a,\omega^*})(P^{\nu}_{a,e} + P^{\nu}_{a,\omega^*}) \\ &= g_{\mu\nu} P^{\mu}_{a,e} P^{\nu}_{a,e} + g_{\mu\nu} P^{\mu}_{a,\omega^*} P^{\nu}_{a,\omega^*} + 2 g_{\mu\nu} P^{\mu}_{a,e} P^{\nu}_{a,\omega^*} \\ &= -m_e^2 c^2 + 0 + 2 g_{\mu\nu}(P^{\mu}_{a,tot} - P^{\mu}_{a,\omega^*}) P^{\nu}_{a,\omega^*} \\ &= -m_e^2 c^2 + 2 g_{\mu\nu} P^{\mu}_{b,tot} P^{\nu}_{a,\omega^*} - 0 \\ &= -m_e^2 c^2 + 2\left(\frac{\hbar\omega}{c}\frac{\hbar\omega^*}{c}\cos\theta - (\frac{\hbar\omega}{c} + m_e c)(\frac{\hbar\omega^*}{c})\right). \end{aligned}$$

After some simplifying algebra and using,

$$\hbar\omega = \frac{2\pi\hbar c}{\lambda},$$

we obtain the desired Compton scattering result,

$$\Delta\lambda = \frac{2\pi\hbar}{m_e c}(1 - \cos\theta).$$

There are other kinds of important invariants. The dynamics of an interaction of colliding particles is characterized by a *cross-section*. Conceptually, the cross-section is an area around a target seen by the approaching particle, such that if the approaching particle passes within this area, some process will take place. For classical Rutherford scattering, alpha particles approach gold nuclei and are scattered by Coulomb repulsion. Surrounding the gold nucleus like a halo with radius s is a strip of area, $d\sigma = 2\pi s ds$ in the plane containing the gold nucleus but perpendicular to the motion of the alpha particle. If the alpha particle passes through this area, it will be scattered into an angle between ϕ and $\phi + d\phi$. The area $d\sigma$ is said to be the *differential cross-section* for this particular process. Similar cross-sections are defined for other nuclear processes.

Since $d\sigma$ represents an area perpendicular to the direction of motion of the approaching particle, taken to be along the x-axis, the cross-section is invariant under the special Lorentz transformation. It provides a simple bridge to connect measurements in the laboratory and analysis in the center-of-mass system.

Variants of the differential cross-section are also used. For Rutherford scattering, the differential cross-section per unit solid angle is defined,

$$\frac{d\sigma}{d\Omega} = \frac{2\pi s ds}{2\pi \sin\phi d\phi} = \frac{s}{\sin\phi}\frac{ds}{d\phi},$$

but this cross-section is not Lorentz invariant because the angle ϕ is not Lorentz invariant. For relativistic purposes it is sometimes convenient to note that a four-volume,

$$(dxdydzcdt) = (\frac{dx'}{\gamma}dy'dz'(\gamma cdt')) = dx'dy'dz'cdt'$$

is invariant. Since four-vectors have the same transformation properties, it follows that,

$$dp_x dp_y dp_x \frac{E}{c}$$

is an invariant volume in four-momentum space. Thus the cross-section,

$$\frac{1}{E}\frac{d\sigma}{d\mathbf{p}}$$

is an invariant. It is a cross-section for a process that produces a particle into a certain volume of four-momentum space.

The proper mass of a system of particles is an invariant with another use. Suppose that a proton collides with a proton, producing a shower of pions. One suspects that some of the pions result from the decay of a very short-lived particle that is also produced in the collision. Let us imagine that the short-lived particle is expected to decay into three pions. The detector has sufficient resolution to allow one to reckon the relativistic momenta and energies of the pions, but not enough to even detect the suspected particle.

The four-momentum invariant of the suspected particle in its own frame, $P_M^\mu = (0,0,0,M_0c)$, has a value $-M_0^2c^2$. If one takes the four-momenta of the pions, three at a time, forms

$$P_t^\mu = ((\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3), (E_1 + E_2 + E_3)/c)$$

and computes $g_{\mu\nu}P_t^\mu P_t^\nu$, and finds that the values cluster around $-M_0^2c^2$, one may take it as compelling evidence for the short-lived existence of the suspected particle.

Thus, invariants are used in a variety of ways to simplify relativistic calculations.