Chapter 6a. Space-time four-vectors.

The preceding chapters have focused on a description of space in terms of three independent, equivalent coordinates. Here we discuss the addition of time as a fourth coordinate in "space-time." This leads to the consideration of space-time transformations, or Lorentz transformations, which are an extension to four dimensions of rotations in three dimensions. Special relativity is introduced here as a generalization of the invariance of length under rotations in three-space. The transformation of the Maxwell stress tensor under relativistic "boosts" is introduced as an application to electromagnetic theory. A later chapter, intended to follow the study of Maxwell's equations, shows how covariant tensor calculus leads to these equations.

A. The origins of special relativity.

This course is mainly about mathematical methods, so I will completely ignore the rich history of scientific discovery and speculation that lead to the integration of space and time into four-space. I will just go over the most compelling reasons based on modern science for requiring something like special relativity.

1. There are lots of kinds of electromagnetic radiation known: light, radio waves, X-rays, WiFi and microwave ovens. All of these disturbances travel at a certain special speed, \( c = 3 \times 10^8 \text{ m/s} \).
2. Beams of electrons are commonplace. Electrons in radio and television tubes respond like most massive particles, speeding up in response to forces acting on them. However, at particle accelerators such as SLAC, it is observed that as particles approach the speed \( c \), they are harder and harder to speed up. They never quite get up to speed \( c \).
3. Radioactive particles are also commonplace; they decay with a characteristic half life. But, mu mesons which are contained in circular orbit by a magnetic field live longer than when at rest. Furthermore, the light curves of supernovae moving away from us at high velocity are stretched out in time, indicating that decaying iron and nickel isotopes are exhibiting a longer half life, presumably because they are moving with a speed approaching \( c \).

Note that \( c \) is a special property of nature, even for phenomena which have nothing to do with light. However, we always call it "the speed of light."

So - it is not too dumb to propose the following: we have admired the simplicity of coordinate transformations from one frame of reference to a rotated frame. Space coordinates are changed, but time is not involved. However, the phenomena described above suggest that a moving observer sees time differently. We thus suppose that transforming to the point of view of a moving observer involves a special sort of "rotation" in space and time.

B. Four-vectors and invariant proper time.
We will add a time coordinate to the usual 3-component space vector. Time does not have the correct units - but \( ct \) (time multiplied by the speed of light) does. Thus we have a four-vector:

\[
x^{\mu} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}
\]

(6a-1)

[Note about indices. I will use Greek letters such as \( \mu, \nu, \lambda, \) and \( \xi \) (mu, nu, lambda and ksi) to label space-time components, as distinct from \( i, j, k, l \ldots \) for space components. There is also an issue about upper and lower indices, corresponding to contra-variant and co-variant indices. I will try to avoid discussing their difference in detail, leaving that for a more specialized course.] A space-time index takes on one of the four values 0, 1, 2, 3. Thus \( x^1, x^2, \) and \( x^3 \) are just the usual \( x, y, \) and \( z, \) with the fourth coordinate as \( x^0 = ct. \)

Now, what do we do with a four-vector? An important property of a three-space vector is its length. For a vector \( \vec{A} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \), the length \( A = \sqrt{A^2} \) is given by

\[
A^2 = \delta_{ij} A_i A_j \\
= A_1^2 + A_2^2 + A_3^2
\]

(6a-2)

For the special case of the position vector \( \vec{r} \), the scalar length \( r \) is given by

\[
r^2 = \delta_{ij} x_i x_j \\
= x_1^2 + x_2^2 + x_3^2
\]

(6a-3)

We recall that under rotation the components \( x_1, x_2, \) and \( x_3 \) can change, but in such a way that the length \( r \) is invariant. In fact, orthogonal transformations are defined to be just those linear coordinate transformations which leave the length of vectors invariant.

In space-time, the corresponding length is called proper time \( \tau. \) It is defined this way:

\[
(ct)^2 = g_{\mu\nu} x^\mu x^\nu
\]

(6a-4)

where \( g_{\mu\nu} \) is the metric tensor, defined as follows:

\[
g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}
\]

(6a-5)

This gives, in terms of the common variables \( t, x, y, \) and \( z, \)

\[
(ct)^2 = g_{\mu\nu} x^\mu x^\nu = (ct)^2 - x^2 - y^2 - z^2
\]

(6a-6)

This is bit more complicated than the three-space version. For one thing, the inner product is formed using not the Kronecker delta function \( \delta_{ij}, \) but the metric tensor \( g_{\mu\nu}. \) Note that it has two lower indices. A general rule for using four-space indices is the
following: When carrying out a contraction by setting two indices equal, one must be a lower index and the other, an upper index. The Einstein summation convention is of course in force, where a paired index is assumed to be summed, from 0 to 3.

There is one major difference between the length of a three-vector and the length of a four-vector. Because of the minus sign in the metric tensor, the length of the vector can come out to be zero, or even negative. This is a warning that, while time has been introduced as a vector component, it is really still different from the space components.

C. The Lorentz transformation.

The defining property of a three-vector is how it changes when the frame of reference (of the observer) is rotated about an axis. The corresponding change in the frame of reference for a 4-vector is from that of a "stationary" observer to that of an observer moving with velocity \( v \) in a particular direction, usually taken to be the \( x \)-direction. (See figure 6a-1.) What is the transformation which changes space and time coordinates, in such a way as to leave the 4-vector length unchanged? Here it is - the Lorentz transformation.

\[
\Lambda^{\mu}_{\nu} = \begin{pmatrix}
\gamma & -\eta & 0 & 0 \\
-\eta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]  

The Lorentz Transformation \hspace{1cm} (6a-7)

where

\[
\gamma = \frac{1}{\sqrt{1 - \beta^2}} \hspace{1cm} (6a-8a)
\]

and

\[
\eta = \gamma \beta \hspace{1cm} (6a-8b)
\]

These symbols form the language of relativity. The symbol \( \beta \) is often just referred to as the velocity; it is a dimensionless velocity formed by dividing the actual velocity (the relative velocity of the two frames of reference) by the speed of light. The two symbols \( \gamma \) and \( \eta \) are functions of the velocity \( \beta \). They play the role for the Lorentz transformation that \( \cos \theta \) and \( \sin \theta \) play in the rotation \( R_z(\theta) \); instead of \( \cos^2 \theta + \sin^2 \theta = 1 \) we have

\[
\gamma^2 - \eta^2 = 1 \hspace{1cm} (6a-9)
\]

The role of the Lorentz transformation matrix \( \Lambda^{\mu}_{\nu} \) given above is to calculate four-space coordinates in a "moving" coordinate system \( S' \), in terms of those in a "stationary"
system $S$. These systems are illustrated below in figure 6a-1.

![Diagram of two rest-frames, in relative motion.](image)

**Figure 6a-1.** Two rest-frames, in relative motion.

The frame $S'$ is in motion, relative to $S$, with velocity $v$ in the $x$-direction. The Lorentz transformation from $S$ to $S'$ is

$$x'^\mu = \Lambda^\mu_\nu x^\nu = \begin{pmatrix} \gamma & -\eta & 0 & 0 \\ -\eta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \gamma ct - \eta x_1 \\ -\eta ct + \gamma x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

(6a-10)

or, in the form often quoted for the Lorentz transformation,

$$t' = \gamma t - \eta \frac{v}{c^2} x$$

$$x' = \gamma x - \gamma vt$$

$$y' = y$$

$$z' = z$$

(6a-11)

The inverse of this transformation is pretty easy to figure out. It is just obtained by changing $v$ to $-v$. $\gamma$ is unchanged, and $\eta$ changes to $-\eta$, giving

$$\Lambda^{-\mu_\nu} = \begin{pmatrix} \gamma & \eta & 0 & 0 \\ \eta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The Inverse Lorentz Transformation

(6a-12)

and
D. Space-time events.

Some of the most interesting effects in special relativity involve objects in motion - that is, in motion with respect to the observer. The frame of reference in which the object is not in motion is called its **rest frame**. Properties of an object, in space or time, may vary according to the observer's state of motion with respect to the object, and the properties observed in the rest frame are considered to be fundamental properties of the object. For instance, the intrinsic length of an object is that measured in its rest frame. And the time for an object to do something (go to sleep, then wake up, for instance) is properly measured in the object's rest frame. We will now show that time intervals are stretched out ("dilated") if the object is moving, and lengths are shortened ("contracted").

The basis of geometry in three-space consists of points, specified by the coordinates \((x,y,z)\). In four-space we talk instead of **events**, specified by time and position, e.g., for event \(A\),

\[
x_A^\mu = \begin{pmatrix} \gamma t_A \\ x_A \\ y_A \\ z_A \end{pmatrix},
\]

(6a-14)

A point in space can be marked by driving a stake in the ground, or carving your name on a tree. For a four-space version, some authors imagine setting off a small bomb, so that a black mark gives the position, and the sound adds the time. It is of special interest to consider the 4-space displacement between two events. Thus, the displacement from event \(A\) to event \(B\) is

\[
\Delta x^\mu = x_B^\mu - x_A^\mu = \begin{pmatrix} c(t_B - t_A) \\ x_B - x_A \\ y_B - y_A \\ z_B - z_A \end{pmatrix},
\]

(6a-15)

E. The time dilation.

Let us consider two events, \(A\) and \(B\), happening to an object in frame \(S'\). This is the rest frame of the particle, so they both happen at a single point, which we will take to be the origin, \(x' = y' = z' = 0\). Let the first event happen at time \(t = 0\), and the second, at time \(T_0\). So, the four-vector positions of \(A\) and \(B\), in frame \(S'\), are
Vector Spaces in Physics 8/6/2015

\[ x_A^\mu = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad x_B^\mu = \begin{pmatrix} cT_0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \text{ and } \Delta x' = x_B^\mu - x_A^\mu = \begin{pmatrix} cT_0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \]

Now we use the inverse Lorentz transformation to go to the frame \( S \) of the observer:

\[
\Delta x^\nu = \Lambda^{-1 \mu}_{\nu} \Delta x'^\nu = \begin{pmatrix} \gamma & \eta & 0 & 0 \\ \eta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} cT_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma cT_0 \\ \eta cT_0 \\ 0 \\ 0 \end{pmatrix}.
\]

This tells us that the time interval \( T \) observed in \( S \) is a factor of \( \gamma \) greater than the time interval \( T_0 \) observed in the rest frame. That is,

\[ T = \gamma T_0 \]

Example. Suppose a rocket going to Mars travels at relativistic speed \( \beta = 0.1 \), that is, at 10% the speed of light. (This is not actually very practical.) How long would a year of an astronaut's life (observed in her rest frame, moving with the rocket) appear to take?

The time-dilation factor \( \gamma \) is

\[ \gamma = \frac{1}{\sqrt{1 - \beta^2}} = \frac{1}{\sqrt{1 - 0.01}} = 1.005 \]

So, the length of the dilated year (as we see it, not in her rest frame) is

\[ T = \gamma T_0 = 1.005 \text{ years} \]

F. The Lorentz contraction.

According to the Lorentz contraction, fast-moving objects appear shorter than they actually are. Let us see how this works. Suppose that a stick of length \( L_0 \) (if measured in its rest frame \( S' \)) is in fact observed in frame \( S \), where it appears to be moving at velocity \( v \) along the \( x \)-axis. Let events \( A \) and \( B \) be observations of the two ends of the stick in its rest frame, as shown in figure 6a-2 below.
Events $A$ and $B$ are separated by a distance $L_0$, the length of the stick. We do not require them to be at the same time, since the stick is not moving. So we can take $A$ to be at the origin, at time $t'_A = 0$, and event $B$ to be at the end of the stick, at undetermined time $t'_B$. Thus,

$$x''_A = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad x''_B = \begin{pmatrix} ct'_B \\ L_0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{and} \quad \Delta x'' = x''_B - x''_A = \begin{pmatrix} ct'_B \\ L_0 \\ 0 \\ 0 \end{pmatrix}.$$ 

We use the inverse Lorentz transformation to see the length of the stick in frame $S$:

$$\Delta x'' = \Lambda^{-1\mu}_{\nu} \Delta x''^\nu = \begin{pmatrix} \gamma & \eta & 0 & 0 \\ \eta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct'_B \\ L_0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma ct'_B + \eta L_0 \\ \eta ct'_B + \gamma L_0 \\ 0 \\ 0 \end{pmatrix}.$$ 

(6a-18)

We are interested in events $A$ and $B$ which occur at the same time in frame $S$. And with this condition, the separation of events $A$ and $B$ in frame $S$ is the length of the stick, as observed with the stick in motion. That is,

$$\Delta x'' = \begin{pmatrix} \gamma ct'_B + \eta L_0 \\ \eta ct'_B + \gamma L_0 \\ 0 \\ 0 \end{pmatrix}.$$ 

(6a-19)

This gives two equations, for the first two components. The first one can be used to eliminate $t'_B$, giving
\[ t'_B = -\frac{\eta}{\gamma c} L_0, \]

and then the second equation gives
\[
L = \eta t'_B + \gamma L_0
\]
\[
= \eta \left( -\frac{\eta}{\gamma c} \right) L_0 + \gamma L_0
\]
\[
= \frac{L_0}{\gamma} (-\eta^2 + \gamma^2)
\]

or, using \( \gamma^2 - \eta^2 = 1 \),

\[
L = \frac{1}{\gamma} L_0 \quad \text{Lorentz contraction (6a-20)}
\]

G. The Maxwell field tensor.

The electric field \( \vec{E} \) and the magnetic field \( \vec{B} \) each have three components which seem to be related to the directions in space. But how do they fit into relativistic four-space? There are no obvious scalar quantities to provide the fourth component of their four-vectors. Furthermore, the magnetic field has some dubious qualities for a true vector. For one, it is derived from a cross product of vectors, and so does not reverse under the parity transformation, as all true vectors do.

There is another interesting argument indicating that the relativistic transformation properties of the electric and magnetic fields are complicated. Under Lorentz transformation, the four components of the four-vector are re-arranged amongst themselves. But the transformation to a moving coordinate system turns a pure electric field into a combination of electric and magnetic fields. This can be understood in a very rough way from the following observation. An electrostatic field can be produced by a distribution of fixed charges. But if one shifts to a moving coordinate system, the charges are moving, constituting currents, which generate magnetic fields.
A concrete example can make a prediction for the transformation of electric into magnetic fields. Consider two line charges, as shown in figure 6a-3 below. This distribution of stationary charge produces an electrostatic field, as shown. Near the origin, the electric field is in the positive $y$ direction.

**Figure 6a-3.** Two static line-charge distributions, producing an electrostatic electric field. Near the origin, the electric field is in the positive $y$ direction.

distribution of stationary charge produces an electrostatic field, as shown. Near the origin, the electrostatic field is in the $+y$ direction. Now, what does this look like to an observer in system $S'$, moving in the $+x$ direction? This is shown in figure 6a-4. There is
now a magnetic field, in the negative z direction. There is another, more subtle prediction. Because of the Lorentz contraction, the wires appear shorter, and so the charge density on the wires is greater, and the electric field should be stronger.

So, we have this prediction for the transformation of electromagnetic fields: Suppose that there is just an electric field present in \( S \), in the positive y-direction. Then in the \( S' \) frame, the field transformation should produce a magnetic field in the negative z-direction, and a stronger electric field, still in the positive y-direction. Now we will postulate a transformation law for electromagnetic fields, and see if this prediction is fulfilled.

Here is the field-strength tensor of special relativity.
The tensor transformation works just like with three-vectors and rotations $\overline{R}$, except that the Lorentz transformation matrix $\Lambda^\mu_\nu$ plays the role for four-tensors that $\overline{R}$ played for three-tensors. The electromagnetic fields as seen in the moving system are thus

$$F'^{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta}$$

$$= \overline{\Lambda} F \overline{\Lambda}^T$$

$$= \begin{pmatrix}
\gamma & -\eta & 0 & 0 \\
-\eta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{pmatrix}
\begin{pmatrix}
0 & -E_x & -E_y & -E_z \\
E_x & 0 & -B_z & B_y \\
E_y & B_z & 0 & -B_x \\
E_z & -B_y & B_x & 0 \\
\end{pmatrix}
\begin{pmatrix}
\gamma & -\eta & 0 & 0 \\
-\eta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{pmatrix}$$

$$= \begin{pmatrix}
\gamma & -\eta & 0 & 0 \\
-\eta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{pmatrix}
\begin{pmatrix}
\eta E_x & -\gamma E_x & -E_y & -E_z \\
\gamma E_x & -\eta E_x & -B_z & B_y \\
\gamma E_y - \eta B_z & -\eta E_y + \gamma B_z & 0 & -B_x \\
\gamma E_z + \eta B_y & -\eta E_z - \gamma B_y & B_x & 0 \\
\end{pmatrix}$$

$$= \begin{pmatrix}
0 & E_x (-\gamma^2 + \eta^2) & -\gamma E_y + \eta B_z & -\gamma E_z - \eta B_y \\
E_x (\gamma^2 - \eta^2) & 0 & \eta E_y - \gamma B_z & \eta E_z + \gamma B_y \\
\gamma E_y - \eta B_z & -\eta E_y + \gamma B_z & 0 & -B_x \\
\gamma E_z + \eta B_y & -\eta E_z - \gamma B_y & B_x & 0 \\
\end{pmatrix}$$

(6a-22)
Let's try to absorb this result. To start with, note that the matrix is still anti-symmetric; this is a cross check on the algebra. Next, we see that the $x$-components of both $\vec{E}$ and $\vec{B}$, in the direction of the relative velocity, do not change. However, for the transverse fields, they get all scrambled up. We can write the four non-trivial field transformation equations like this:

$$
\begin{align*}
E'_y &= \gamma E_y - \eta B_z \\
E'_z &= \gamma E_z + \eta B_y \\
B'_y &= \gamma B_y + \eta E_z \\
B'_z &= \gamma B_z - \eta E_y
\end{align*}
$$

(6a-24)

We see that the transverse components of both fields get bigger. This is what we predicted for the $\vec{E}$ field. And a bit of the other field, in the other transverse direction, gets added on. Here is another simple check. Take the zero-velocity limit. Do you find that the fields do not change?

Finally, let's consider the example above. We predicted that transforming a positive $E_y$ would give a negative $B_z$. Look at the fourth equation above: that is just what happens. I. I. Rabi would love it.

Note on units: The form of $F^{\mu \nu}$ given is in Gaussian units. To use SI units, replace $E_i$ by $E_i / c$.

**Problems**

**Problem 6a-1.** The algebra of special relativity leans heavily on the following dimensionless symbols:

$$
\begin{align*}
\beta &\equiv \frac{v}{c} \\
\gamma &\equiv \frac{1}{\sqrt{1 - \beta^2}} \\
\eta &\equiv \gamma \beta
\end{align*}
$$

Here $v$ represents the velocity of one frame of reference with respect to the other.

(a) What are the limiting values of these three symbols as the velocity $v$ approaches the speed of light $c$?

(b) Calculate the value of 

$$\gamma^2 - \eta^2$$

The result should be independent of the velocity.
(c) The inverse of the Lorentz boost

\[ \Lambda^{\mu}_{\nu} = \begin{pmatrix} \gamma & -\eta & 0 & 0 \\ -\eta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \]

is obtained by reversing the velocity \( v \), which causes the change \( \beta \rightarrow -\beta \). The result is given above, in equation (6a-12). Carry out the matrix multiplication to demonstrate that this works; that is, show by direct calculation that

\[ \Lambda \Lambda^{-1} = I \]

**Problem 6a-2.** The magnitude of the 4-position vector,

\[ (ct)^2 - x^2 - y^2 - z^2, \]

should be invariant under Lorentz transformation. Using the relations given in equation (6a-11) above, calculate

\[ (ct')^2 - x'^2 - y'^2 - z'^2 \]

and see if the invariance works out.

**Problem 6a-3.** The energy of an object at rest is (famously) given by \( E = mc^2 \), where \( m \) (as always in our discussions) represents the object's rest mass. And, in the object's rest frame, its four-momentum

\[ p^\mu = \begin{pmatrix} E \\ p_x c \\ p_y c \\ p_z c \end{pmatrix} \] (general case)

becomes

\[ p^\mu = \begin{pmatrix} mc^2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \] (in the particle's rest frame).

(a) Set the invariant length-squared of the first expression above equal to the invariant length-squared of the second one. Solve for the object's energy \( E \), in terms of its momentum \( p \) and its rest-mass \( m \). (The speed of light, \( c \), will be there too.)

(b) Multiply the second four-vector by the Lorentz-transformation matrix \( \Lambda^{-1} \) (that is, transform it to a frame of reference moving backwards along the \( x \)-axis, with velocity \(-\beta \) ) and use the result to derive expressions for the object's energy and momentum as a function of the relativistic velocity of the particle.

**Problem 6a-4.** The nearest star to our sun is about 3 light-years away. That is, something traveling from Earth at speed \( v = c \) would take 3 years to get to the star, according to observers in the Earth frame of reference. Consider a rocket, carrying an
astronaut, traveling to the star from Earth at speed $\beta$. The time $T$ to get there would be

$$T = \frac{3 \text{ years}}{\beta}.$$

(a) How fast would the rocket have to travel, in m/sec, to get to the star within a reasonable life expectancy of the astronaut, say $T_0 = 50$ years? (Start by calculating the value of $\beta$.) Note: here you can approximate $\gamma \approx 1$, so astronaut time and Earth time will be about the same.

(b) Answer the same question about travel to Andromeda, $1,500,000$ light-years away, also in $50$ years of astronaut time. Note: in calculating the travel time in the Earth frame you can approximate $\beta \approx 1$, so that the travel time in the Earth frame is about $1,500,000$ years.

**Problem 6a-5.** Recent studies of distant supernovae played a central role in the discovery of dark energy. One researcher (Prof. Gerson Goldhaber) pointed out that the relativistic time dilation should be observable for the most distant galaxies, which are moving close to the speed of light. This is because the decrease in the brightness of a supernova over the first $100$ days after the initial explosion is due to the decaying of the isotope Fe$^{58}$, which decays with a half-life of $20$ days. According to the theory of special relativity, this half-life should appear longer to us. (The half-life is the time for the decay rate to decrease by a factor of 1/2.)

If the galaxy is moving with speed $\beta = 0.8$, how long should it take for its light to decrease in intensity by a factor of 1/2? By a factor of 1/128?