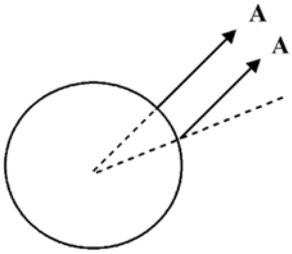
LAST TIME: General tensors, more general definition of components, general Kronecker delta, general quotient theorem, and metric tensor

We want to tackle the problem of tensor differentiation now. What is the problem? Recall that for differentiation in the rectangular Cartesian system, the only variable is the component of the tensor. The basis vectors are constant. However, as we have seen, the vector operators acting on the basis vectors in non-Cartesian systems do not give zero. You can see the problem better by thinking of what it means to take a derivative. Ordinarily, you must think in terms of small displacements and then use a limiting process. Consider what happens when you displace a vector in a curvilinear coordinate system such as a spherical or a cylindrical system. In the top position, vector A has only a radial component, but when it is parallel transported to a new position, it has both radial and tangential components in a cylindrical system. Here, the magnitude and direction of the vector have not changed, but the components have. (Note that if we were to measure in the Cartesian coordinate system, everything would be fine.) Consider the vector given by $\mathbf{A} = A^1\mathbf{e}_1 + A^2\mathbf{e}_2 + A^3\mathbf{e}_3 = A^i\mathbf{e}_i$. Calculate



$$\frac{\partial \mathbf{A}}{\partial x^1} = \frac{\partial(A^1\mathbf{e}_1 + A^2\mathbf{e}_2 + A^3\mathbf{e}_3)}{\partial x^1} = \frac{\partial\left(A^i\mathbf{e}_i\right)}{\partial x^1} = \frac{\partial A^i}{\partial x^1}\mathbf{e}_i + A^i\frac{\partial \mathbf{e}_i}{\partial x^1}.$$

We wish to devise a way to formalize the derivative of the basis vector. The first point to understand is that the derivative of a basis vector will be proportional to another basis vector. This means we may write

$$\Gamma_{ij}^{k}\mathbf{e}_k = \frac{\partial \mathbf{e}_i}{\partial x^j}$$

as the defining equation for Γ_{ij}^{k}, which is known as a Christoffel symbol of the second kind. Here is how the process works. Γ gives the magnitude of one component of the derivative vector. *k* tells which basis vector points in the direction of this component of the derivative vector. *i* tells which basis vector is being considered. *j* tells which coordinate is being varied to change the basis vector.

Suppose we were to find two Christoffel symbols as follows. Consider $\Gamma_{r\phi}^{r}$. This means the change in $\mathbf{e}_r$ in the direction of $\mathbf{e}_r$ as ϕ is changed. If there is no change in $\mathbf{e}_r$ as ϕ is changed, the value is 0. Now look at $\Gamma_{r\phi}^{\phi}$. Now we are asking for the change in $\mathbf{e}_r$ in the direction of $\mathbf{e}_\phi$ as ϕ is changed. Suppose its value is 1/*r*. Then we have

$$\frac{\partial \mathbf{e}_r}{\partial \phi} = (0)\mathbf{e}_r + \left(\frac{1}{r}\right)\mathbf{e}_\phi.$$

We have already seen that we know a great deal if we just know the elements of the metric tensor for a system. It turns out that it is quite possible to derive a relationship between the Christoffel symbols and the metric tensor, but it is a bit tedious. Let me just write the result for you, so we can see its importance.

$$\Gamma_{ij}^{l} = \frac{1}{2} g^{kl} \left[\frac{\partial g_{ik}}{\partial x^{j}} + \frac{\partial g_{jk}}{\partial x^{i}} - \frac{\partial g_{ij}}{\partial x^{k}} \right].$$

This means that once you know the metric, you can calculate the Christoffel symbols straightforwardly. Once you know these, you can take the derivatives of tensors that automatically accounts for the change in the basis vectors. When you derived the expressions for the velocity and acceleration in cylindrical and spherical coordinate systems, you were actually doing covariant derivatives the hard way by having to account for the basis vector changes separately. As a short excursion into this area, note that the metric tensor for this system is obtained from the square of the differential length element as $ds^2 = dr^2 + r^2 d\phi^2 + dz^2$. Therefore, the metric tensor is given by

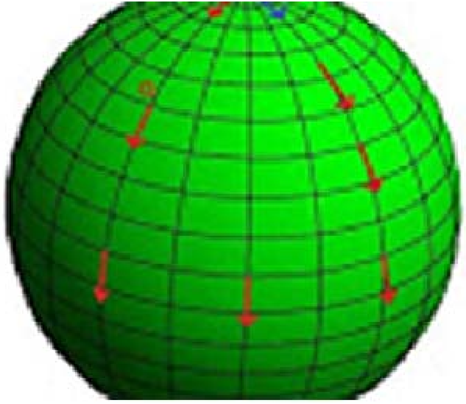
$$g_{ij} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now using $x^1 = r, x^2 = \phi,$ and $x^3 = z,$ we may get the only nonzero Christoffel symbols to be

$$\Gamma_{22}^{1} = -r, \quad \Gamma_{12}^{2} = \Gamma_{21}^{2} = \frac{1}{r}.$$

We are now able to calculate the derivative of a tensor by knowing the Christoffel symbols, which are obtained from the metric tensor, which, in turn, are gotten from the transformation equations. These relationships should not be a surprise, but it is nice to see how everything fits so beautifully.

For a few moments, let's return to the problem of parallel transport. For Cartesian coordinate systems, the problem turns out not to be difficult, and for curvilinear systems, you also saw how to deal with the problem. However, on curved surfaces, for example, the surface of a sphere, things are quite different. The figure shows how parallel transport on a curved surface differs. We are working on the surface of a sphere as our coordinate system. Suppose we create a vector at the North Pole and begin transporting the vector along a meridian keeping it always pointing south. We imagine taking small steps. When we arrive at the equator, we transport the vector along the curved path on the equator, once again keeping it pointing south. When we come to another meridian, we again transport the vector back to the North Pole, again being careful to keep it pointing south. When we arrive back at the North Pole, the vector is not pointing in the same direction it was when we started the process. This result is strictly due to the curvature of the surface on which we are transporting the vector. In general, the definition of parallel transport is that the covariant derivative is zero. I have been advised by most of my colleagues that this is a good place to end this discussion. To proceed further, we would need to derive the expression for



the Riemann curvature tensor, which is the fundamental tensor in general relativity. The vanishing of this tensor is a necessary and sufficient condition for a Euclidean (flat) space.

I want to cover two additional topics concerning special theory of relativity before moving on to a brief discussion of matrices. The first concerns the two different ways in which the transformation equations are treated and the resulting four vectors and four tensors that arise from these different methods. Recall that the Lorentz transformation equations are usually derived by considering how a light pulse emitted at the time the origins of the two reference frames coincide. Using the principles of relativity and the linearity of the transformations, the Lorentz transformation equations become

$$\bar{t} = \frac{t - vx/c^2}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \bar{x} = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \bar{y} = y, \text{and } \bar{z} = z.$$

Note here that the barred reference frame is in relative motion with respect to the unbarred reference frame.
Setting $\beta = \frac{v}{c}$ and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ gives $c\bar{t} = \gamma(ct - \beta x), \bar{x} = \gamma(x - \beta ct), \bar{y} = y,$ and $\bar{z} = z$.

It is not hard to show by ordinary algebra that the invariant interval is given by

$$ds^2 = dx^2 + dy^2 + dz^2 - (c\,dt)^2.$$

This invariant interval is the same no matter how we choose to represent the line element from which it is obtained. In one method, usually more common in older books, the line element is represented by a four vector given by

$$d\mathbf{r} = (dx, dy, dz, icdt).$$

One advantage of this approach is that it preserves the notion of an orthogonal transformation that we have already studied at the expense of creating an imaginary component $icdt$. It also turns out that the angle of rotation is not a real angle. One reason this method is used is that the preservation of the orthogonal transformation conforms to much of what we already developed for Cartesian vectors. In other words, $ds^2 = d\mathbf{r} \cdot d\mathbf{r}$. With this approach, there is no need to distinguish between covariant and contravariant four vectors or four tensors. We will define each of these in a moment. The four vector is constructed using the invariance of the scalar product, just as we said that the orthogonal transformation preserved distances. We define a four vector to have components given by $x_o = ict, \quad x_1 = x, \quad x_2 = y, \ and\ x_3 = z$. A four vector is an object that transforms according to the Lorentz transformation. When we are discussing four vectors, it is common to use Greek symbols instead of the usual Latin ones. Therefore, we would write $x_\mu = (x_o, x_1, x_2, x_3)$. We see that the Lorentz transformations may be written in matrix form as

$$\begin{bmatrix} \bar{x}_o \\ \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} \gamma & -i\beta\gamma & 0 & 0 \\ i\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_o \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \Rightarrow \bar{x}_\mu = L_{\mu\nu} x_\nu,$$

where $L_{\mu\nu}$ is the Lorentz transformation matrix.

You can notice that this is an orthogonal matrix which gives makes it an orthogonal transformation. On the other hand, you see that the rotation angle is given by $\tan\theta = i\beta.$ Without getting too far afield, many other four vectors are possible. Because proper time $d\tau$ is a scalar, you can create four velocity and four acceleration. As you develop relativistic electrodynamics next semester, there will be four current (*icp*, **J**), four potential (iV/c, **A**), and other four vectors. Any four vector transforms according to the Lorentz transformation matrix. In general then,

$$A_\mu = L_{\mu\nu} A_\nu$$

represents the transformation of any four vector.

When you begin to think about the transformation of the fields, only six quantities are involved, three components for each vector field. What kind of an object could this be? It cannot be a four vector. The most likely candidate is a four by four antisymmetric tensor of the form

$$F_{\mu\nu} = \begin{bmatrix} 0 & F_{01} & F_{02} & F_{03} \\ -F_{01} & 0 & F_{12} & F_{13} \\ -F_{02} & -F_{12} & 0 & F_{23} \\ -F_{03} & -F_{13} & -F_{23} & 0 \end{bmatrix}.$$

The values for the various *F*'s will be completed by the values for the 6 components of **E** and **B**. In this approach, some of the values with be imaginary. The transformation of the fields is accomplished by allowing the Lorentz matrix to act twice on the electromagnetic field tensor $F_{\mu\nu}$.

Therefore, $\bar{F}_{\alpha\beta} = L_{\alpha\mu} L_{\beta\nu} F_{\mu\nu}$ gives the correct transformation of the fields when the fields are observed from a moving reference frame. It is this use of two Lorentz transformations that make $F_{\mu\nu}$ a four tensor or Lorentz tensor. As you might guess, all of the space and time operators get folded into a single space-time operator. You will see this in greater detail next semester in electrodynamics.

A method that does not use imaginary components uses the notion of contravariant and covariant components of the four vectors and four tensors along with the metric tensor as a way to convert from one to the other as we saw before. Here, the four components of space-time are written as

$$x^o = ct, \qquad x^1 = x, \quad x^2 = y, \text{and} \quad x^3 = z.$$

We have written these components as contravariant components, so we need a way to make sure that the invariant interval remains the same. This is accomplished by recognizing that the metric is given by

$$g_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Using the previous results for contracting, we create the covariant components as $x_\mu = g_{\mu\nu}x^\nu$. The scalar product $\Delta x_\mu \Delta x^\mu = -\Delta(ct)^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$ is the invariant interval as required. Four vectors such as *A* transform as

$$\bar{A}^\mu = L^\mu_\nu A^\nu,$$

where

$$L^\mu_\nu = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

A Lorentz tensor, say $T^{\mu\nu}$, transforms according to $\bar{T}^{\mu\nu} = L^\mu_\lambda L^\nu_\sigma T^{\lambda\sigma}$. In this case, the electromagnetic field tensor is still a 4 by 4 antisymmetric tensor, but all of the components are real. The transformation law is exactly the same as before. For every scalar product of four vectors, the covariant component will have the opposite sign as the contravariant component. Now, however, the transformation matrix is no longer orthogonal.

The second is what I call the geometry of special relativity. Having taught this subject many times, it sometimes strikes me that many students do not fully comprehend the dramatic difference in the geometry involved in the Lorentz transformation and ordinary transformations. I do not intend to derive the Lorentz transformation, but rather, I will use the results to illustrate my points and show you how to use the Minkowski diagrams to understand better what is going on. Assume that we have our two inertial reference frames with a primed frame moving at velocity *v* along the *x* direction with respect to the other.

The Minkowski diagram uses one orthogonal coordinate system and one nonorthogonal (skewed) coordinate system to represent events in spacetime. To understand better the Minkowski diagram, consider the Lorentz transformation equations for relative motion in the x direction given by

$$x' = \gamma(x - vt),\ y' = y,\ z' = z,\ \text{and}\ t' = \gamma[t - (v/c^2)x].$$

Here, $\gamma = [1 - (v/c)^2]^{-1/2}$, and it is assumed that the primed coordinate system moves along the positive x direction with speed v. The inverse transformation equations are obtained by simply interchanging the primed and unprimed coordinates and changing v to –v, thereby yielding

$$x = \gamma(x' + vt'),\ y = y',\ z = z',\ \text{and}\ t = \gamma[t' + (v/c^2)x'].$$

Most authors use *ct* as the variable on the vertical axis and *x* as the variable on the horizontal axis, causing the transformation and inverse transformations equations to become

$$x' = \gamma [x - \beta (ct)],\ y' = y,\ z' = z,\ and\ ct' = \gamma [(ct) - \beta x]$$

and

$$x = \gamma [x' + \beta (ct')],\ y = y',\ z = z',\ and\ ct = \gamma [(ct') + \beta x'].$$

To construct the ct' - x' axis with respect to the ct – x axis, simply set $x' = 0$ and solve for ct as a function of x and set $ct' = 0$ and solve for x as a function of ct. Doing so gives the following equations for $x' = 0$ and $ct' = 0$. For the ct'-axis, $ct = \beta^{-1}x$, and for the x'-axis, $ct = \beta x$. Defining the axes is only part of the problem. The invariant interval $s^2 = (ct)^2 - x^2$ must be incorporated by using the hyperbolas with values of s = 1, 2, 3, 4, …, n. Figure 1 shows the construction of the $ct - x$ system and the ct' - x' systems with the hyperbolas overlayed to show unit length in each system. Figure 2 shows the two systems with the hyperbolas removed and the unit lengths in each system clearly labeled. As the value of β increases, the ct' and x' axes approach one another, making unit length in the ct' - x' system appear much longer than unit length in the $ct - x$ system. Figures 3 and 4 show the construction of length contraction and time dilation, respectively.

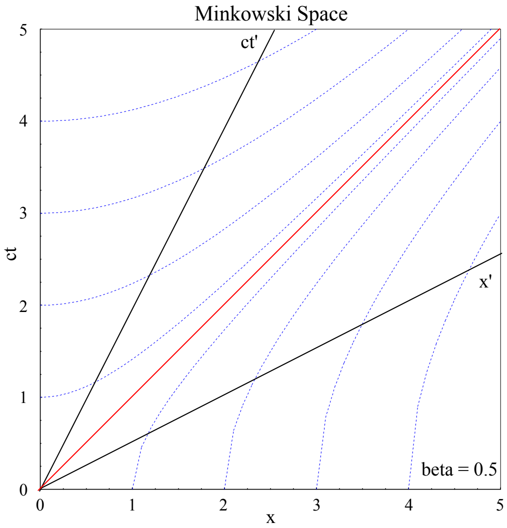


Figure 1 Minkowski spacetime diagram with hyperbolas overlayed.

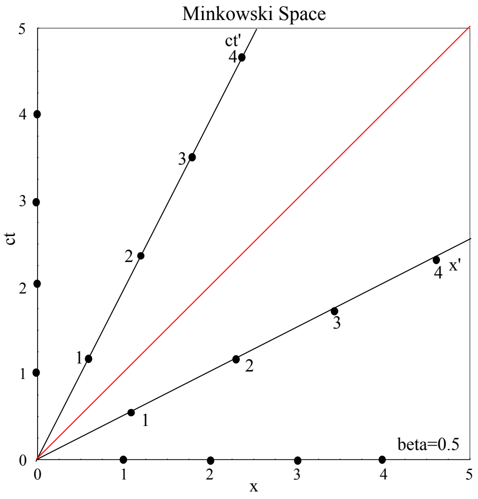


Figure 2. Unit lengths in ct' - x' and ct – x systems clearly labeled for β = 0.5

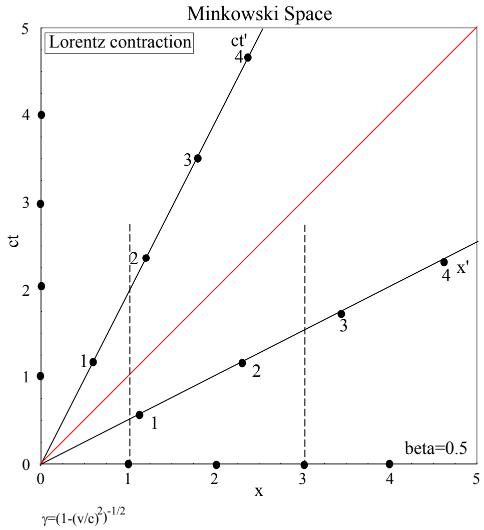


Figure 3. Length contraction. An object 2 units long in $ct - x$ appears to be longer in the ct' - x', but is clearly shorter because of the appearance of unit lengths.

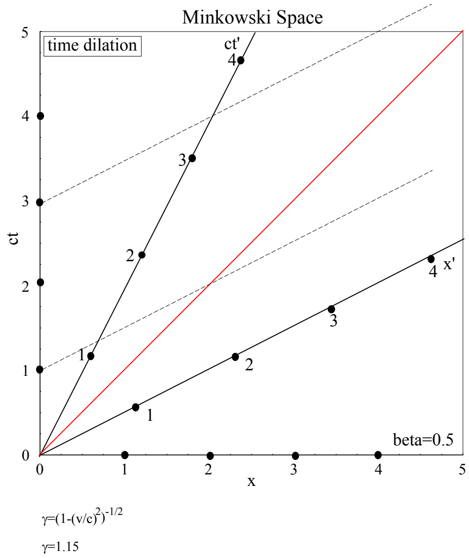


Figure 4. Time Dilation. A value of $ct = 2$ in $ct - x$ appears larger in ct' - x', but not by as much as it appears.

It is clear from these examples that the Minkowski diagram is useful for visualizing time dilation and length contraction, but its major shortcoming is the difference between unit length appearance in the two systems. Without introducing a scaling factor, it is not possible to measure the two lengths or two time intervals directly to compare them. In fact, the reason for this problem is that the Minkowski diagram does exactly what we tell students not to do – give preference to one frame of reference over another. We start with one of the reference frames represented by an orthogonal system, and this causes the hyperbolas that define unit lengths to have a different appearance in the two frames of reference.

NEXT TIME: Matrices in physics