Theory of Relativity
Michaelmas Term 2009: M. Haehnelt

2 Space-Time Approach to Special Relativity

2.1 Tensors

Lorentz transformations leave invariant the squared interval between neighbouring events

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

This is analogous to the invariance of the squared interval

$$dx^2 + dy^2 + dz^2$$

under rotations and translations of orthogonal Cartesian coordinates in Euclidean space.

A Four tensor is an object defined on the 4-space of events

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z$$

which behaves as a **tensor** under a Poincaré transformation.

But what is a tensor?

• Construct a set of real variables $\{x^1,\ x^2,\ \ldots x^N\}$ which we can regard as coordinates of an N-dimensional space V_N.

• A non-singular transformation of these coordinates to a new set of coordinates is a recoordination of the space V_N.

• A tensor is an object in V_N described by a set of M real numbers.

• The components of the tensor (*i.e.* the M numbers) will (generally) change when we transform the coordinates of V_N.

• To qualify as a tensor, **the components must transform in a specific way**.

In N-dimensional space, a tensor of rank r has

$$M = N^r \quad \text{components}$$

$$r = 0, \quad M = 1, \quad \text{defining a scalar}$$

$$r = 1, \quad M = N, \quad \text{defining a vector}$$

$$r = 2, \quad M = N^2 \quad \text{defining a second rank tensor}$$

and so on. So, to understand what a tensor is, we must investigate coordinate transformations.

Coordinate transformations

The space V_N is characterised by coordinates:

$$(x^1,\ x^2,\ x^3\ \ldots x^N),$$

$$(x'^1,\ x'^2,\ x'^3,\ \ldots x'^N).$$

Assume that the transformations from one set of coordinates to another are non-singular and differentiable

$$\frac{\partial x'^\alpha}{\partial x^i} = p_i^\alpha, \quad \frac{\partial x^i}{\partial x'^\alpha} = p_\alpha^i, \quad etc.$$

Definition of Tensors

- **Contravariant tensor**

Has components:	$A^{i,j,\ldots n}$ in x^i system
	$A'^{\alpha,\beta,\ldots\nu}$ in x'^α system
transforms according to	$A'^{\alpha,\beta,\ldots\nu} = A^{i,j,\ldots n}\, p_i^\alpha p_j^\beta \ldots p_n^\nu$

- **Covariant tensor**

Has components:	$A_{i,j,\ldots n}$ in x^i system
	$A'_{\alpha,\beta,\ldots\nu}$ in x'^α system
transforms according to	$A_{\alpha,\beta,\ldots\nu} = A_{i,j,\ldots n}\, p_\alpha^i, p_\beta^j \ldots p_\nu^n$

- **Mixed tensor**

Has components:	$A_{\ell\ldots n}^{i\ldots k}$ in x^i system
transforms according to	$A'^{\alpha\ldots\nu}_{\beta\ldots\mu} = A_{\ell\ldots n}^{i\ldots k}\, p_i^\alpha \ldots p_k^\nu p_\beta^\ell \ldots p_\mu^n.$

Note the following points:

- The components of tensors in the new coordinate system are linear functions of the components in the old coordinate system.

- An object is a tensor if it transforms as a tensor under all non-singular coordinate transformations.

- We can construct objects which behave as tensors only under certain *subgroups* of linear transformations, *e.g.* Lorentz transformations.

Examples of tensors

- **Contravariant vector**

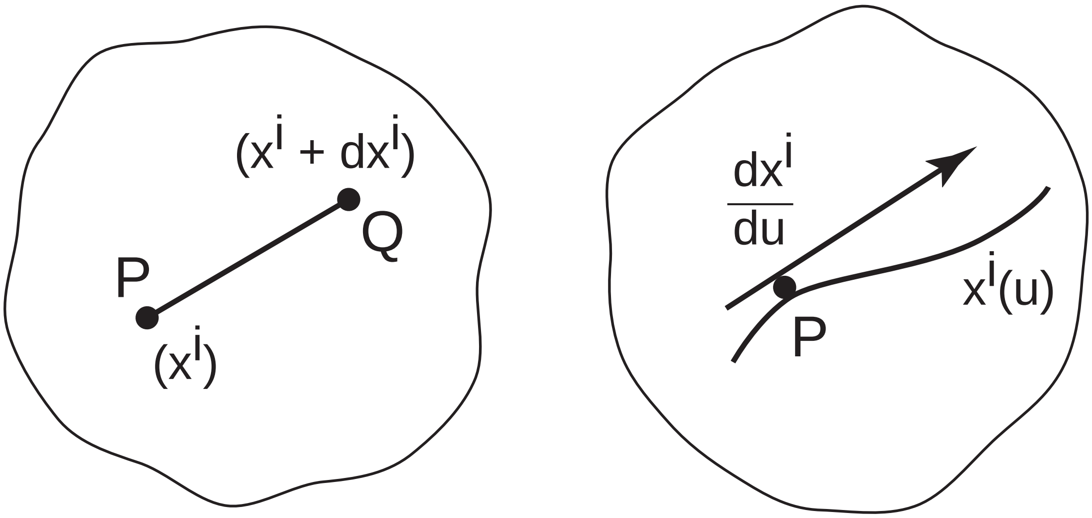

Figure 1:

Consider a curve $x^i(u)$ (u is a parameter that varies along the curve) the tangent vector dx^i/du at some point P is a contravariant vector.

The derivatives of coordinates transform as a contravariant vector since

$$dx'^{\alpha} = \frac{\partial x'^{\alpha}}{\partial x^i} dx^i.$$

- **Covariant Vector**

The gradient of a scalar field $\phi(x^1, \ \ldots, x^n)$ is an example of a covariant vector since the gradient transforms as

$$\frac{\partial \phi'(x')}{\partial x'^{\alpha}} = \frac{\partial \phi(x(x'))}{\partial x^i} \ \frac{\partial x^i}{\partial x'^{\alpha}}.$$

Example of a quantity that is not a tensor

The vector A^i transforms as

$$A'^{\alpha} = A^i \frac{\partial x'^{\alpha}}{\partial x^i}.$$

Now form the derivative and apply the chain rule

$$\frac{\partial A'^{\alpha}}{\partial x'^{\beta}} = \frac{\partial A^i}{\partial x^j}\frac{\partial x^j}{\partial x'^{\beta}}\frac{\partial x'^{\alpha}}{\partial x^i} + A^i \frac{\partial^2 x'^{\alpha}}{\partial x^j \partial x^i}\frac{\partial x^j}{\partial x'^{\beta}}.$$

The first term on the right hand side is exactly what we would expect for the transformation of a rank 2 mixed tensor. However, we see that there is another term that depends on the second derivative $\partial^2 x'^{\alpha}/\partial x^j \partial x^i$. The derivative of a vector field transforms as a tensor only if this second derivative is zero, *i.e.* for the special class of linear coordinate transformations.

Why use tensors?

Suppose that we have two tensors,

$$X_{ij} = Y_{ij}.$$

Apply a transformation of the coordinates,

$$\begin{aligned} X'_{\alpha\beta} &= X_{ij}\frac{\partial x^i}{\partial x'^{\alpha}}\frac{\partial x^j}{\partial x'^{\beta}}, \\ Y'_{\alpha\beta} &= Y_{ij}\frac{\partial x^i}{\partial x'^{\alpha}}\frac{\partial x^j}{\partial x'^{\beta}}, \end{aligned}$$

and so,

$$X'_{\alpha\beta} = Y'_{\alpha\beta}.$$

i.e. a **tensor equation** is **invariant** under all non-singular coordinate transformations and is therefore **coordinate free**.

We would like to formulate the laws of physics so that they are manifestly independent of the coordinate system. The above argument suggests that the laws of physics are formulated as tensorial equations.

2.2 Space-time

2.2.1 The metric

The space of events in special relativity is endowed with a metric

$$ds^2 = g_{\mu\nu} dx^{\mu} dx^{\nu}$$

where $g_{\mu\nu}$ is constant under Lorentz transformations:

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

2.2.2 Scalar product

The metric tensor $g_{\mu\nu}$ defines the scalar product of two four vectors,

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} &= g_{\mu\nu} A^\mu B^\nu \\ &= A^0 B^0 - A^1 B^1 - A^2 B^2 - A^3 B^3. \end{aligned}$$

and (given our definition of tensors) is invariant under any coordinate transformation.

2.2.3 Raising and lowering indices

The metric tensor $g_{\mu\nu}$ can be used to raise and lower indices, since we can write our scalar product

$$\mathbf{A} \cdot \mathbf{B} = g_{\mu\nu} A^\mu B^\nu,$$

as

$$\mathbf{A} \cdot \mathbf{B} = A_\nu B^\nu,$$

where

$$A_\nu = g_{\mu\nu} A^\mu$$

is the covariant form of the vector A^μ. So, $g_{\mu\nu}$ acts to *contract* the index μ of A^μ.

Now, we can define a tensor

$$g^{\mu\nu}$$

that will *raise* the index of a vector:

$$\begin{aligned} A^\mu &= g^{\mu\nu} A_\nu \\ A^\mu &= g^{\mu\nu} g_{\lambda\nu} A^\lambda. \end{aligned}$$

Evidently $g^{\mu\nu}$ must satisfy

$$g^{\mu\nu} g_{\lambda\nu} = \delta^\mu_\lambda,$$

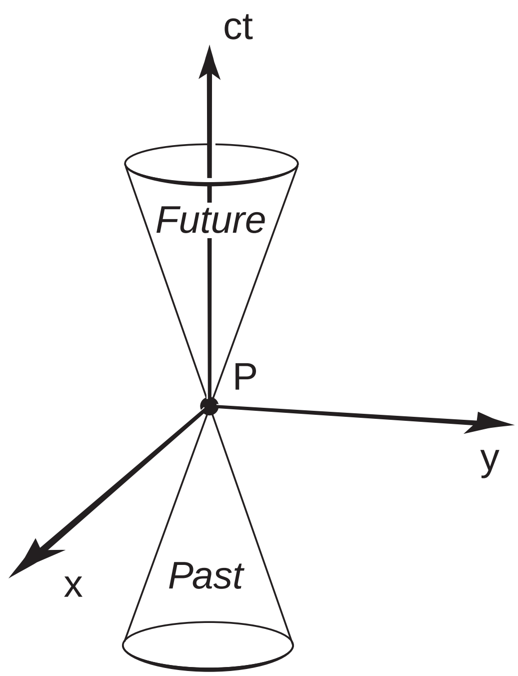

Figure 2:

where δ^μ_λ in the Kronecker δ,

$$\delta^\mu_\lambda = \begin{cases} 1 & \mu = \lambda, \\ \\ 0 & \mu \neq \lambda, \end{cases}$$

i.e. $g^{\mu\nu}$ is the inverse of $g_{\mu\nu}$. Now in Special Relativity the metric tensor $g_{\mu\nu}$ is diagonal and equal to

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} := (\eta_{\mu\nu}),$$

and so the inverse $g^{\mu\nu}$ is identical,

$$(g^{\mu\nu}) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} := (\eta_{\mu\nu}).$$

We can use the metric tensors $g_{\mu\nu}$ and $g^{\mu\nu}$ to raise and lower indices of tensors of arbitrary rank. For example, we can produce a mixed tensor from the second rank tensor $B^{\mu\lambda}$ as follows

$$B^{\mu}_{\nu} = g_{\nu\lambda} B^{\mu\lambda}.$$

Now, returning to the discussion of space-time, consider an event P at (x^α) and all events in its neighbourhood $(x^\alpha + \Delta x^\alpha)$ as shown in Figure 2 above.

We can group the neighbouring events into classes according to whether the (invariant) metric interval

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$$

is positive, zero or negative.

• $\Delta s^2 = 0$ - defines a light cone, *i.e.* a spherical wavefront converging on P. Events lying on the light-cone are termed 'light-like' or 'null'.

• $\Delta s^2 > 0$ - events with $\Delta s^2 > 0$ lie inside the light-cone and are termed 'time-like'.

• $\Delta s^2 < 0$ - events with $\Delta s^2 < 0$ lie outside light-cone and are termed 'space-like'.

What is the significance of these classifications? This is made clear in the following diagram:

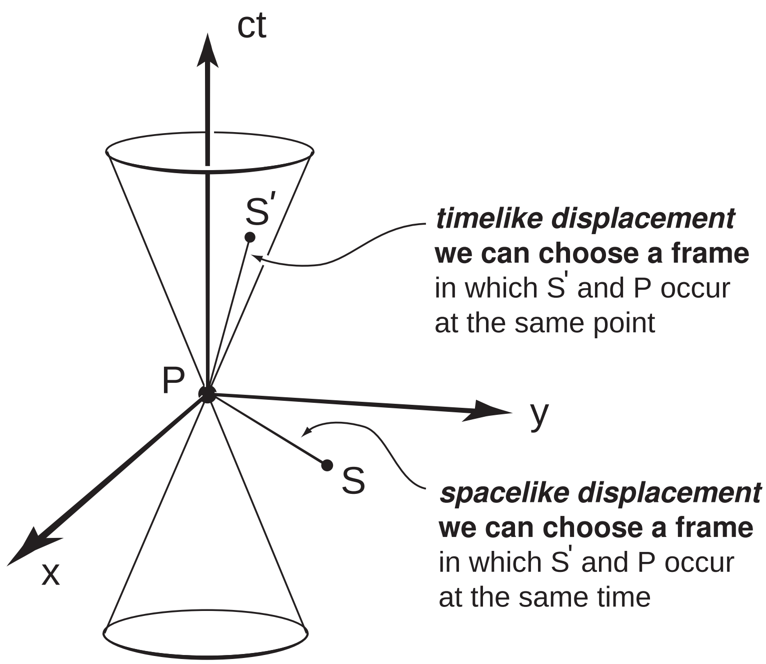

Figure 3:

This shows two events S' and S. S' lies within the light cone of P and so is displaced by a timelike interval. If the interval is timelike, we can find an inertial frame in which the events S' and P appear at the same point in space.

The events P and S are separated by a spacelike interval. For these we can choose an inertial frame in which the events S and P appear to occur at the same time.

These points become obvious if we view Lorentz transformations as rotations, as will be discussed in the next section.

2.3 Lorentz transformations as rotations

The Lorentz transformation between our two standard frames S and S' is

$$\left.\begin{aligned} x' &= \gamma(x - vt) \\ t' &= \gamma(t - vx/c^2) \end{aligned}\right\} \qquad (1)$$

which we have seen preserves the line-interval

$$x'^2 - c^2t'^2 = x^2 - c^2t^2.$$

We can write (1) as

$$\left.\begin{array}{rcl} x' &=& x\ \cosh\psi + ct\ \sinh\psi, \\ \\ ct' &=& x\ \sinh\psi + ct\ \cosh\psi, \end{array}\right\} \tag{2}$$

and it straightforward to show that this parameterization preserves the invariance of the line interval,

$$\begin{array}{rcl} x'^2 - ct'^2 &=& x^2\cosh^2\psi + c^2t^2\sinh^2\psi + 2ctx\sinh\psi\cosh\psi \\ &-& x^2\sinh^2\psi - c^2t^2\cosh^2\psi - 2ctx\sinh\psi\ \cosh\psi \\ &=& x^2 - c^2t^2. \end{array}$$

Obviously, equivalence of equations (1) and (2) requires

$$\begin{array}{rcl} \cosh\psi &=& \gamma, \\ \sinh\psi &=& -\gamma v/c, \end{array}$$

which satisfies the constraint equation

$$\cosh^2\psi - \sinh^2\psi = \gamma^2\left(1 - v^2/c^2\right) = 1.$$

We can rewrite (1) in a form suggestive of a rotation

$$\begin{pmatrix} x' \\ \\ ct' \end{pmatrix} = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \\ \sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} x \\ \\ ct \end{pmatrix} \tag{3}$$

with inverse

$$\begin{pmatrix} x \\ \\ ct \end{pmatrix} = \begin{pmatrix} \cosh\psi & -\sinh\psi \\ \\ -\sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} x' \\ \\ ct' \end{pmatrix}.$$

Now, let's recall the relations for rotations in Cartesian coordinates,

$$\begin{array}{rcl} x &=& x'\cos\theta - y'\sin\theta, \\ y &=& y'\cos\theta + x'\sin\theta. \end{array}$$

or, equivalently, in matrix notation

$$\begin{pmatrix} x \\ \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ \\ y' \end{pmatrix}.$$

In fact, the Lorentz transformation can be made to look identical to a rotation of Cartesian coordinates if we make the transformations:

$$\begin{array}{rcl} \psi &=& i\theta, \\ x_0 &=& ict. \end{array}$$

So, if we use imaginary time, a Lorentz transformation corresponds to a rotation of the axes by an angle θ

$$\begin{pmatrix} x \\ \\ x_0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ \\ x'_0 \end{pmatrix}.$$

Some of the older text books on Special Relativity use an imaginary time coordinate. By using imaginary time, the line-element of SR becomes identical to that of 4 dimension Cartesian space, and as we have seen above, Lorentz transformations are equivalent to rotations of these Cartesian coordinates. However, imaginary time is not particularly intuitive and so it seems a high price to pay to work in obscure coordinates to gain a smidgen of mathematical simplicity. Furthermore, in General Relativity there is no way of avoiding the full machinery of non-Euclidean geometry, so generally we gain no simplicity in GR by using imaginary time. For these reasons, it is extremely rare to find anybody using imaginary time nowadays.[1]

If a Lorentz transformation in the coordinates (ct, x, y, z) does not correspond to a rotation of Cartesian coordinates, what does it correspond to? This is straightforward to see from equations (1):

$$\begin{aligned} x' &= \gamma(x - vt), \\ t' &= \gamma(t - vx/c^2). \end{aligned}$$

The line $t' = 0$ corresponds to $ct = vx/c$, and so lines of constant time t' look like this in the (ct, x) plane:

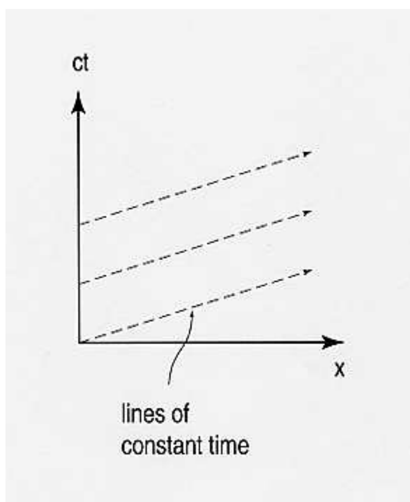

Figure 4:

[1] An exception is an approach to quantum gravity called Euclidean quantum gravity, which utilises imaginary time.

Similarly, the line $x' = 0$ corresponds to $ct = cx/v$, and so lines of constant x' lie parallel to the $x' = 0$ line in the (ct, x) plane:

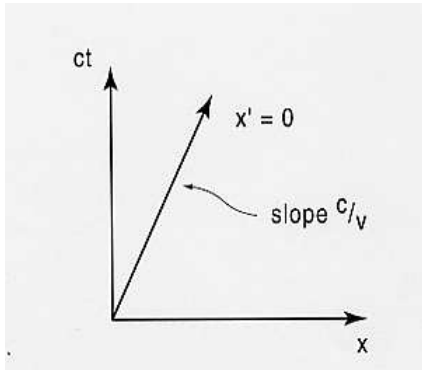

Figure 5:

Hence the transformation looks like a squishing of the axes:

The bigger the relative velocity between the two frames, the greater the squishing of the primed frame's axes. As the relative velocity between the two frames tends towards the speed of light, both the t' and x' axes get squished closer to the $x = ct$ null path of a light ray.

Once you have understood the squishing effect of a Lorentz transformation, it is easy to visualize the time dilation and Lorentz contraction effects in Figure 6.

Consider the two curves

$$\begin{aligned} c^2t^2 - x^2 &= a^2, \\ c^2t^2 - x^2 &= -b^2. \end{aligned}$$

These curves define hyperbolae as shown in the diagram below:

The hyperbolae define Lorentz invariant intervals and so can be used to set the scales (in metres and seconds) of the tilted axes (t', x') shown in Figure 6.

For example, consider the two events A and B plotted in Figure 7.

Event A has coordinates $x = 0$ and $ct = a$ since it sits on the hyperbola defined by $c^2t^2 - x^2 = a^2$.

Event B has coordinates $x' = 0$ and $ct' = a$ since it sits on the (same) hyperbola defined by $(ct')^2 - (x')^2 = a^2$. A standard unit of time (lets say a second) in the primed frame is clearly longer in the unprimed frame and so one can see geometrically how time dilation comes

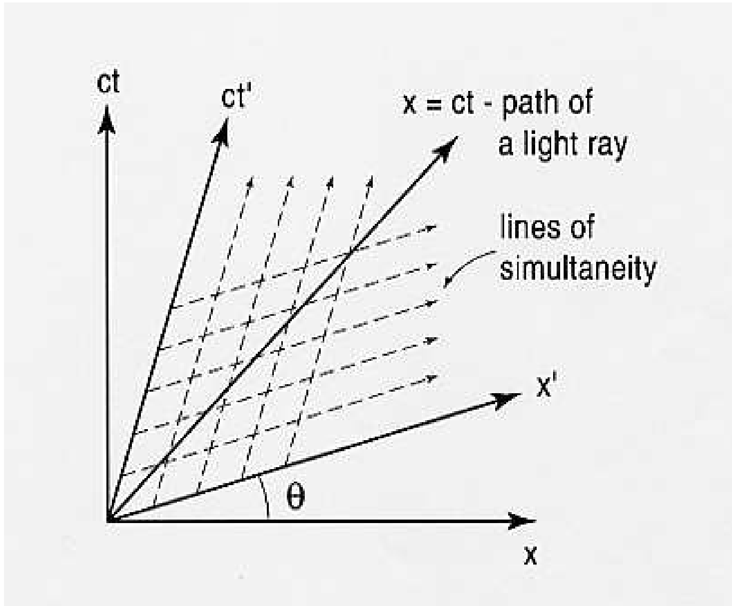

Figure 6:

about. Explicitly, the event B has coordinates ($x' = 0$, ct'), but in the unprimed frame it has coordinates

$$(ct, x = vt),$$

hence

$$c^2t^2 - v^2t^2 = c^2t'^2$$

i.e.,

$$t = \frac{t'}{(1 - v^2/c^2)^{1/2}},$$

as expected from the time-dilation formula derived in Section 1.5.

We can analyse length contraction in a similar way, by calibrating the x' axis of Figure 6 using the spacelike hyperbola $c^2t^2 - x^2 = -b^2$. For example, consider a rod of standard length L in the S' frame. One end of the rod is assumed to be a the origin of the S' coordinate system ($x' = 0$, $t' = 0$).

The other end of the rod defines an event C with coordinates

$$\begin{aligned} t' &= 0, \\ x' &= L, \end{aligned}$$

hence, the hyperbolae passing through point C obeys the equation

$$c^2t^2 - x^2 = -L^2.$$

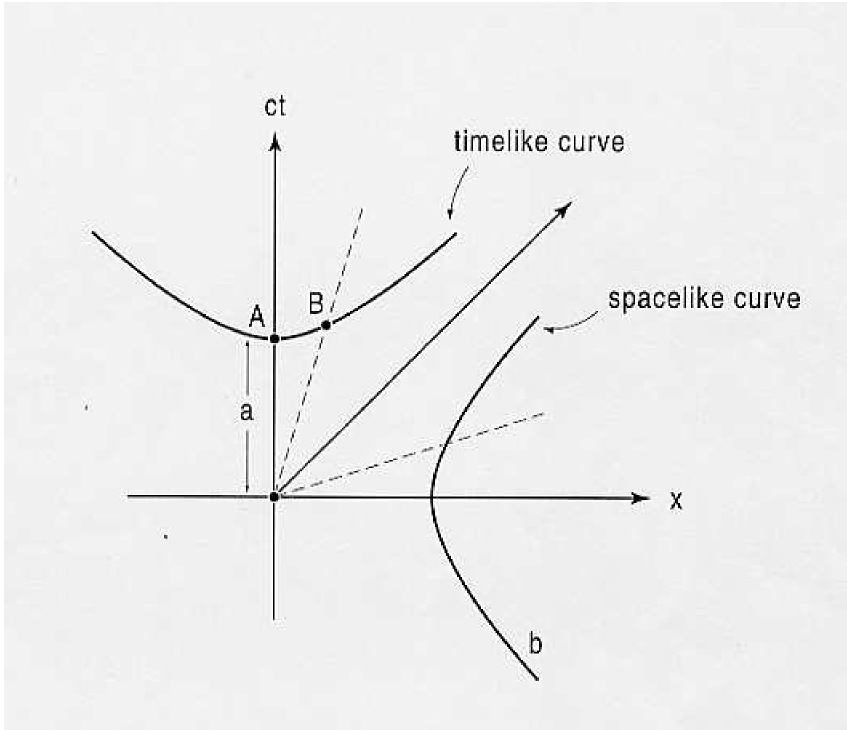

Figure 7:

The x coordinate (x_C) of event C is therefore given by the equation,

$$\frac{v^2}{c^2}x_C^2 - x_C^2 = -L^2$$

i.e.

$$x_C = \gamma L.$$

The time in frame S assigned to event C is therefore

$$t_C = \frac{v}{c^2}x_C = \frac{v}{c^2}\gamma L.$$

Now the length of the rod as seen in frame S is not equal to OC, since the events O and C correspond to different times. The length of the rod in S must be defined at a *fixed time* in frame S (*i.e.* on a constant time hypersurface in S). For example, the length of OB in Figure 8 gives the length of the rod at the fixed time $t = 0$. We therefore need to calculate the x coordinate of event B (x_B). But if the rod is moving at speed v relative to frame S, the events B and C must satisfy

$$\frac{x_C - x_B}{t_C - t_B} = v,$$

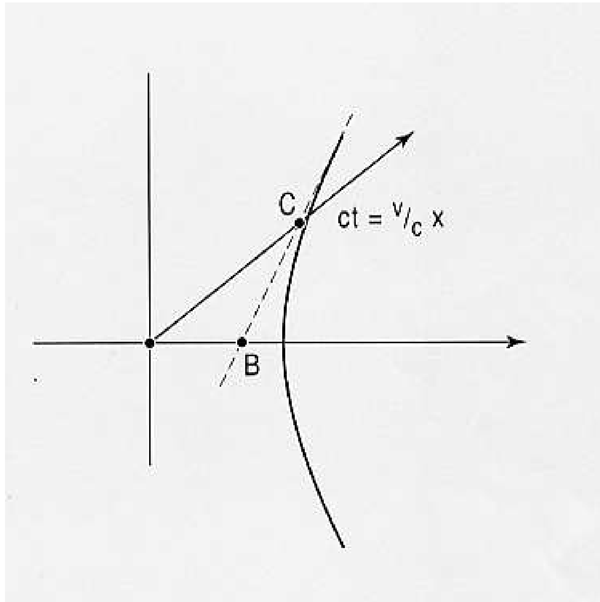

Figure 8:

and so we can calculate x_B when $t_B = 0$ and hence the length of the rod as seen in S:

$$\begin{aligned} x_B &= x_C - vt_C \\ &= \gamma L - \frac{v^2}{c^2}\gamma L = \left(1 - \frac{v^2}{c^2}\right)^{1/2} L, \end{aligned}$$

which is the familiar Lorentz contraction formula. It is easy to see visually from Figures 6 and 8 how Lorentz contraction works – note that as the speed of the rod approaches the speed of light, the x' axis approaches the null ray $x = ct$ and the length of the rod (given by the length of OB) tends to zero as expected from the length contraction formula.