THE GEOMETRY OF SPECIAL RELATIVITY

We represent the actual four dimensional geometry with two dimensions, x and ct, where ct is chosen so that the units will be the same on both axis.

As in class, this splits space-time into four regions: future (events we can get to), past (events we can have come from), elsewhere (events we can’t reach or have come from), and light paths.

This four dimensional space-time is called Minkowski space-time. In this geometry we have vectors going from one event to another (points in the space-time are called “events”). What is invariant about a vector is its length (the same in any coordinate system whereas the individual components are not). In ordinary three dimensional space this is given by the Pythagorean Theorem:

\[ ds^2 = dx^2 + dy^2 + dz^2 \]

for two neighboring points. In special relativity we find the invariant to be:

\[ ds^2 = -(cdt)^2 + dx^2 + dy^2 + dz^2 \]

as is readily checked with the Lorentz transformations:

\[ dx' = \gamma \left( x - \frac{v}{c} cdt \right) \]
\[ dy' = dy \]
\[ dz' = dz \]
\[ cdt' = \gamma \left( cdt - \frac{v}{c} dx \right) \]

Then
\[-(ct')^2 + dx'^2 + dy'^2 + dz'^2 = -\gamma^2 \left( c dt - \frac{v}{c} dx \right)^2 + \gamma^2 \left( dx - \frac{v}{c} c dt \right)^2 + dy^2 + dz^2 \]
\[-(ct')^2 + dx'^2 + dy'^2 + dz'^2 = -(ct)^2 \left( \gamma^2 - \gamma^2 \frac{v^2}{c^2} \right) + dx^2 \left( \gamma^2 - \gamma^2 \frac{v^2}{c^2} \right) + (ct) dx \left( 2\gamma^2 \frac{v}{c} - 2\gamma^2 \frac{v}{c} \right) + dy^2 + dz^2 \]
\[-(ct)^2 + dx^2 + dy^2 + dz^2 \]

It is convenient to introduce some new notation. We denote the coordinates by:

\[x^\mu\]
\[\mu = 0, 1, 2, 3\]
\[x^0 = ct\]
\[x^0 = x\]
\[x^1 = x\]
\[x^2 = y\]
\[x^3 = z\]

We also adopt the “summation convention” where repeated up and down indices are to be summed from 0 to 3:

\[x_\mu x^\mu = x_0 x^0 + x_1 x^1 + x_2 x^2 + x_3 x^3\]

Then we can write

\[ds^2 = g_{\mu \nu} x^\mu x^\nu\]

where \(g_{00} = -1, g_{11} = g_{22} = g_{23} = 1, g_{\mu \nu} = 0\) for \(\mu \neq \nu\). In this form \(g_{\mu \nu}\) is called the metric tensor and completely describes the space-time.

**GEOMETRIC ORIGIN OF LORENTZ TRANSFORMATIONS**

Consider two observers with \(S'\) moving down the \(x\) axis of \(S\) with speed \(v\). As seen from \(S\) the situation is as shown:
where
\[
\tan \theta = \frac{x}{ct} = \frac{vt}{ct} = \frac{v}{c} = \beta
\]

Now consider clocks located at the origins of S and S’. They agree at the point where the origins coincide. At time \( t_1 \) a light pulse, showing the time \( t_1 \), leaves S and is received at the origin of S’ when the clock in S’ reads \( t_1’ \).

We must have \( ct_1’ \) proportional to \( ct_1 \) (everything must be linear so that there are no special points in space-time). Thus:
\[
ct_1’ = k(\text{ct}_1)
\]

S’ immediately transmits a pulse, showing the time \( t_1’ \), back to S. Since the laws of nature must be the same for both observers we have:
\[
ct_2 = k(\text{ct}_1’) = k^2ct_1
\]

Now the critical point. What does S think the distance to S’ is at the moment the pulse left S’? The only thing we know for sure is the speed of light. Hence we measure distance by light travel time. Thus:
\[
d = \frac{c(\text{t}_2 - \text{t}_1)}{2}
\]

At what time, as seen by S, was the signal transmitted? From the sketch:
\[
ct_T = \frac{\text{ct}_2 + \text{ct}_1}{2}
\]

But we know the relation between these two since S’ is moving with speed \( v \) down the x axis of S. Hence:
\[
v = \frac{d}{t_T} = \frac{c(t_2 - t_1)}{2\left(\frac{ct_T}{c}\right)}
\]

or

\[
\beta = \frac{k^2ct_1 - ct_1}{k^2ct_1 + ct_1} = \frac{k^2 - 1}{k^2 + 1}
\]

\[
\therefore k^2(1 - \beta) = 1 + \beta \rightarrow k = \left(\frac{1 + \beta}{1 - \beta}\right)^{1/2}
\]

We take the + sign so that the clocks will run in the same direction.

Next consider the time interval between when the origins coincided and the moment when the signal was sent from S'.

\[
c\Delta t = \frac{ct_2 + ct_1}{2} = \left(\frac{k^2 + 1}{2}\right)ct_1
\]

\[
c\Delta t = kct_1
\]

\[
\therefore \frac{\Delta t}{\Delta t'} = \frac{k^2 + 1}{2k} = \frac{\left(\frac{1 + \beta}{1 - \beta}\right) + 1}{\left(\frac{1 - \beta}{1 - \beta}\right)^{1/2}} = \frac{1}{\left(\frac{1 - \beta^2}{1 - \beta}\right)^{1/2}} = \frac{1}{\left(1 - \beta^2\right)^{1/2}} = \gamma
\]

Since the clock being observed is in S', this is just our familiar time dilation result:

\[
\Delta t = \gamma \Delta t_0 \quad !!!
\]

Note that \((ct_0)\) is NOT the length in the S system since the length and time scales for S and S’ are different.

Next consider lengths. We first consider where the S’ axis must be as seen from S.
Since the clock at the origin of S’ stays there, the path of S’ must be the ct’ axis. As we have seen above it makes an angle θ with the ct axis, where:

\[ \beta = \tan(\theta) \]

The question is: where is x’? Let it make an angle φ with the x axis as shown. We know that the speed of light must be c in each system. This means that the path of a light ray must be midway between the space and (c*time) axis in EACH system. Hence we must have \( \theta = \phi \).

We now consider how lengths appear to the two observers. Consider a rod of length \( L_0 \) at rest in system S and lying along the x axis with one end at the origin. Note that we denote the length as seen in S by \( L_0 \) since the rod is at rest in S. The rod will trace out a sheet in space-time.

Since we measure distance in terms of the time it takes light to travel to the object, we measure the length \( L_0 \) by sending a light ray to the end of the rod. It strikes the end at point w, and is then reflected back to us. We receive the reflected ray at position Q. It has therefore taken time \( (PQ/c) \) to go out and back. Thus we find:

\[ L_0 = \frac{PQ}{2} \]
Now consider an observer S’ moving with velocity v along the x axis of S. S’ also measures the length of the rod with a light ray. Again both ends must be marked at the same time as seen by S’. However simultaneity is different for S than for S’, as seen in the sketch below.

For S points O and W are simultaneous. For S’ points O and V are simultaneous. Consider the following diagram:

S’ measures L by sending a light pulse from R. It returns to her at point Z. Thus she finds the length to be:

$$L = \frac{Rz}{2}$$

Now we use the k factor found when we looked at time dilation to relate these times.

$$OQ = k_R \cdot (OZ)$$
\[ PO = k_A \cdot (RO) \]

where

\[ k_R = \left( \frac{1 + \beta}{1 - \beta} \right)^{1/2} \quad k_A = \left( \frac{1 - \beta}{1 + \beta} \right)^{1/2} \]

(Note that \( k_A \) is approach and \( k_R \) is receding. Hence \( v \) becomes \(-v\) and \( \beta \) becomes \(-\beta\), giving \( k_A = 1/k_B \)). The requirement of simultaneity gives \( RO = OZ \). Hence:

\[ OQ + PO = \left( \kappa + \frac{1}{\kappa} \right)(OR) = \left( \frac{\kappa^2 + 1}{\kappa} \right) L = \left[ \frac{1 + \beta}{1 - \beta} + 1 \right] L = \frac{2}{(1 - \beta^2)} L \]

But

\[ OQ + PO = 2L_0 \]

or

\[ L = \frac{L_0}{\gamma} \]

as expected.

As an exercise in understanding the graphs we are using we derive this result in a different way. The two measurements (\( S \) and \( S' \)) are illustrated in the following diagram:
S uses rays $y_1$ and $y_3$ for the length measurement, and gets the result:

$$L_0 = (AO + AQ)/2$$

S’ uses the rays $y_0 + y_2$ and gets the result:

$$L = (RO + OE)/2$$

We choose:

$$A = (ct_A,0) \quad \quad R = (ct_A,\beta ct_A) \quad \text{with } \tan(\theta) = \beta$$

Note that the $x$ component of $R$ is fixed by the geometry. All lines are at $45^\circ$ with respect to the $x$ axis since they are light rays. Note that this also means that triangles OAR and OWV are congruent since all angles are equal and $OA = OW$ ( $y_1$ is a light path). Because both S and S’ must make their measurements of both ends simultaneously we must have:

$$W = (0,L_0) \quad \quad V = (L_0 \beta ,L_0)$$

The equations of $y_2$ and $y_3$ are:

$$y_2 = -x + a_2$$
$$y_3 = -x + a_3$$

But since $y_2$ must go through $(L_0 \beta ,L_0)$ and $y_3$ must go through $(0,L_0)$ we get:

$$y_2 = -x + L_0(1 + \beta)$$
$$y_3 = -x + L_0$$

Now we need points $Q$ and $E$. $Q$ is the intersection of lines:

$$x = 0 \text{ and } y_3$$

while $E$ is the intersection of lines:

$$x = \beta y \text{ and } y_2$$

Thus for $Q$ we have:

$$y = L_0 \rightarrow Q = (L_0,0)$$

For $E$ we get:
\[ \frac{x}{\beta} = -x + L_0 (1 + \beta) \]

\[ x(\beta + 1) = L_0 (\beta + 1) \beta \rightarrow x = L_0 \beta, \ y = L_0 \]

\[ : \ E = (L_0, L_0 \beta) \]

Thus the y coordinates of Q and E are the same, as are the y components of A and R (by construction). We can then use our time dilation results to get:

\[ (AO + OQ) = \gamma (RO + OE) \]

Thus:

\[ L_0 = \gamma L \]

as above!

**LORENTZ TRANSFORMATIONS**

Consider an event E as viewed by S and S’. The event of course is unchanged by who looks. Only the numbers used to describe the event change. The situation is shown below.

Where

- J is the ct’ component of E
- I is the ct component of E
- F is the x’ component of E
- D is the x component of E
Now consider triangles OHJ and EFG. They are congruent since all angles are equal and OJ = FE. In the same way triangles EKF and OFC are congruent. Then since ED = OI we have CF = HI. Because of our length contraction result we have OF = (OC/γ). (Recall that what caused the effect was the requirement that the measurements in each system be simultaneous. But this is also true here, since O and F are simultaneous in S’ while O and C are simultaneous in S.) But OF is x’. Hence \( x' = (OC/\gamma) \). We also have:

\[
HJ = (OH) \tan \theta = (OH)\beta
\]

\[
x = OC + CD = OC + HJ = OC + \beta(OH)
\]

\[
= \gamma(OF) + \beta(OH) = \gamma x' + \gamma \beta t' = \gamma(x' + \beta t')
\]

Next consider ct.

\[
ct = OI = OH + HI = OH + CF = OH + (OC) \tan \theta = OH + \beta(OC)
\]

\[
\therefore ct = \gamma ct' + \gamma \beta x' = \gamma(ct' + \beta x')
\]

But these are exactly the Lorentz transformation results we derived algebraically!

**S AND S’ SCALES**

We have noticed throughout that the scale for S’ is different than that for S. This is why, for example, that a time that appeared longer for S’ in our sketches had a ct’ component which was actually smaller than ct. We can now find the S’ scales by using the invariant:

\[
ds^2 = -(cdt)^2 + dx^2 + dy^2 + dz^2
\]

\[
s^2 = -(ct)^2 + x^2 + y^2 + z^2
\]

Suppose \( y = z = 0 \). Then for a given value of \( s^2 \) we have the equation of a hyperbola:

\[
x^2 - (ct)^2 = s^2
\]
But along the ct axis we have

\[ t = \frac{1}{c} \left( s^2 \right)^{1/2} \]

For possible particle motions \( s^2 < 0 \). Hence

\[ t = \frac{1}{c} |s| \]

Thus \( t \) is the time shown on the clock at the origin of \( S \) as seen by the observer \( S \). We call this the “proper time”. Suppose \(|s| = c\). Then \( t = 1 \) sec. Now consider \( S' \).

It’s path will intersect the hyperbola at point \( A \). Since \( s^2 \) is an invariant we will have:

\[ s^2 = - (ct')^2 + x'^2 \]

and since at \( x' = 0 \) at \( A \), we have:

\[ t' = \frac{|s|}{c} \]

This gives the scale in \( S' \) as seen from \( S \). In other words, \( OA \) is one in \( S' \), whereas \( OB \) is one in \( S \).
Clearly we can do the same thing when \( s^2 > 0 \). Then the hyperbola looks like:

For \( s^2 = 1 \), OB is one in S while OA is ne in \( S' \).

**EXAMPLE**

As an example of the usefulness of the geometric approach we consider the following situation. A rocket ship approaches the earth at speed \( v \). The ship has length \( L \) as seen by the crew. There are beacons at the front and rear of the ship which flash once per second simultaneously as seen by the crew. We want to answer the following questions on both approach and recession (after it passes earth). Because of the difference in simultaneity for observers on the ship compared to observers on the earth, the beacons will not appear to flash at the same time as seen from the earth. In each second, which will the earth observer think flashed first, and by how much? Also which pulse will arrive first at the earth and by how much? Note that these are different questions since the earth observer can take account of the length of the ship when he decides which flashed first. The situation on approach is shown in the following figure:
From our length contraction results we have:

\[ CA = \gamma L_0 \]

Don’t be confused by the seeming notation confusion. \( L_0 \) is just the name we gave the length of the ship in \( S' \). What matters is that the measurements of \( C \) and \( G \) are simultaneous in \( S' \) while those of \( C \) and \( A \) are simultaneous in \( S \). We then have:

\[ \gamma c t_F - \gamma c t_B = BF = AG = \gamma L_0 \tan(\theta) = \gamma \beta L_0 \]

Hence the earth observer thinks the back beacon flashed first by \( \gamma \beta L_0 \).

Next we ask which pulse arrived first, and by how much. For this the situation is shown below:

We know the coordinates of \( C \) and \( G \). They are:

\[ C = (ct_B, \beta ct_B) \quad \quad \quad G = (ct_F, \beta ct_B + \gamma L_0) \]

The time required for the light to travel from \( C \) to \( H \) is:

\[ t_1 = (0 - \beta ct_B)/c \]

Thus it arrives at the earth \((x = 0)\) at:

\[ ct_{BA} = ct_B - \beta ct_B = ct_B(1 - \beta) \]

The time required to go from \( G \) to \( K \) is:
\[ t_2 = (0 - \beta ct_B - \gamma L_0)/c \]

Thus it arrives at:

\[ ct_{FA} = ct_F - (\beta ct_B + \gamma L_0) \]

Thus:

\[ t_{BA} - t_{FA} = t_B - t_F + \gamma L_0 = -\gamma \beta L_0/c + \gamma L_0/c = (\gamma L_0/c)(1 - \beta) \]

Thus the front pulse arrives first as expected from the diagram.

Following the same procedure after the ship passes earth we find the same result for time of flash, but that the back beacon pulse arrives first by:

\[ (\gamma L_0/c)(1 + \beta) \]

\[ ct_F - ct_B = \gamma L_0 \beta \]

\[ H = [(ct)_B + \beta ct_B, 0] \]

\[ K = [(ct)_F + \beta ct_B + \gamma L_0, 0] \]

\[ \therefore ct_{FA} - ct_{BA} = K_{ct} - H_{ct} = (ct)_F + \beta ct_B + \gamma L_0 - (ct_B) - \beta ct_B = (ct_F) - (ct_B) + \gamma L_0 (\beta + 1) \]

\[ \therefore t_{FA} - t_{BA} = \frac{\gamma L_0}{c} (\beta + 1) \]
Hence back arrives first by

\[ \frac{\gamma L_0}{c} (\beta + 1) \]

which again is to be expected from the diagrams. These results can obviously also be obtained algebraically, but only with considerable care.