INVARIANCE OF SPACETIME INTERVALS

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog and include the title or URL of this post in your comment.

Post date: 14 Jan 2021.

If a light beam travels along the x-axis of some observer O_1 we can define the *interval* between the event of the light being emitted at some point (t_a, x_a) and arriving at another point (t_b, x_b). The interval is defined by

$$\Delta s^2 \equiv -(\Delta t)^2 + (\Delta x)^2 \tag{1}$$
$$\Delta x = x_b - x_a \tag{2}$$
$$\Delta t = t_b - t_a \tag{3}$$

or in three dimensions, as

$$\Delta s^2 \equiv -(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \tag{4}$$

Note that the symbol Δs^2 is usually taken as a fundamental quantity and not the square of some other quantity Δs. This is because it is possible for Δs^2 to be negative and we don't want to introduce imaginary numbers.

Going back to one dimension for the moment, (using relativistic units) since $c = 1$, for a light beam we must have $\Delta x / \Delta t = 1$ for any observer, so we must always have for light, in any reference frame

$$\Delta s^2 = 0 \text{ (for light)} \tag{5}$$

In three dimensions, this is still true, since the distance travelled is

$$\Delta r = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \tag{6}$$

and for light, we always have

$$\frac{\Delta r}{\Delta t} = c = 1 \tag{7}$$

The interval as defined above can be applied to any two events and in fact turns out to be an invariant, in the sense that the interval Δs^2 between two events is the same for all inertial frames. Thus Δs^2 in relativity plays the role of Δr^2 in Euclidean space. We see that in relativity both space and time must be included in the calculation of the interval between two events. However, it is not true to say that time is just another dimension, since it is clearly different from the spatial coordinates, given that it occurs with an opposite sign. We should also add here that the definition of Δs^2 as given above is not unique, and some books use the negative of the one given here, so the reader needs to be aware of the convention in any book.

We need to prove that if, for observer O_1, the interval between two events is

$$\Delta s_1^2 \equiv -(\Delta t_1)^2 + (\Delta x_1)^2 + (\Delta y_1)^2 + (\Delta z_1)^2 \tag{8}$$

then for observer O_2:

$$\Delta s_2^2 \equiv -(\Delta t_2)^2 + (\Delta x_2)^2 + (\Delta y_2)^2 + (\Delta z_2)^2 \tag{9}$$

and

$$\Delta s_1^2 = \Delta s_2^2 \tag{10}$$

We can assume that the origins of the two frames coincide, so that $t_1 = x_1 = y_1 = z_1 = 0$ and $t_2 = x_2 = y_2 = z_2 = 0$ both represent the same event. Since the two observers both reside in inertial frames that move at a constant velocity relative to each other, it also seems reasonable to assume that the relation between the two sets of coordinates is linear, in the sense that

$$t_2 = a_0 t_1 + a_1 x_1 + a_2 y_1 + a_3 z_1 \tag{11}$$

where the a_αs are constants, and similar relations hold for the other three coordinates. [These transformations between the two inertial frames turn out to be the Lorentz transformations, but more on this later.]

At this stage, we can introduce some notation which is common in books and papers on relativity. Spacetime coordinates are referred to using superscripts to denote which of the four coordinates we are referring to. Thus we define $(t,x,y,z) = (x^0, x^1, x^2, x^3)$. Note that these superscripts are *not* exponents, and if we do need to raise a coordinate to a power, we enclose the coordinate in parentheses. Thus $(x^1)^2$ is coordinate x^1 squared.

Also, if the superscript is given a Greek letter, as in x^α, this means that it can be any of the four coordinates. However, if we use a Latin letter, as in x^i, this means that it is any of the spatial coordinates. Thus a Greek superscript can be any of the values 0, 1, 2, 3, while a Latin superscript can be one of 1, 2, 3. Although this may look confusing, when we get into longer calculations, the notation does actually make things a lot easier.

Now to the proof of the invariance of the interval. Since we are assuming the relation between coordinate systems is linear, we can write the interval Δs_2^2 in terms of the coordinates in O_1's system as

$$\Delta s_2^2 = \sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} M_{\alpha\beta}(\Delta x_1^\alpha)(\Delta x_1^\beta) \tag{12}$$

where the $M_{\alpha\beta}$ are (possibly) functions of the relative velocity **v**. This follows since the most general function we can get when calculating the square of a linear expression is a quadratic function, and that's what we've written here. Because (Δx_1^α) and (Δx_2^β) appear symmetrically, if $\alpha \neq \beta$, the two terms involving (Δx_1^α) and (Δx_2^β) together have the coefficient $M_{\alpha\beta} + M_{\beta\alpha}$. Without any loss of generality, we can therefore take $M_{\alpha\beta} = M_{\beta\alpha}$ and define the coefficient of the two terms involving (Δx_1^α) and (Δx_2^β) as $2M_{\alpha\beta}$.

This transformation must be valid for all intervals, so for the particular case of $\Delta s_1^2 = \Delta s_2^2 = 0$ (remember that if an interval is observed to be zero in one frame, it is zero in all frames, due to the assumed constancy of the speed of light in all frames) we can use 6 above to write

$$\Delta r_1 = \Delta t_1 \equiv \Delta x_1^0 \tag{13}$$

We can substitute this into 12 to replace Δt by Δr and we get

$$0 = M_{00}(\Delta r_1)^2 + 2\Delta r_1 \sum_{i=1}^{3} M_{0i}\Delta x_1^i + \sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}(\Delta x_1^i)(\Delta x_1^j) \tag{14}$$

This equation must be true for *all* spatial separations that give the same value for Δr_1, so we can replace Δx_1^i by $-\Delta x_1^i$ to get

$$\begin{aligned} 0 &= M_{00}(\Delta r_1)^2 - 2\Delta r_1 \sum_{i=1}^{3} M_{0i}\Delta x_1^i + \sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}(-\Delta x_1^i)(-\Delta x_1^j) && (15)\\ &= M_{00}(\Delta r_1)^2 - 2\Delta r_1 \sum_{i=1}^{3} M_{0i}\Delta x_1^i + \sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}(\Delta x_1^i)(\Delta x_1^j) && (16) \end{aligned}$$

Note that the last line is the same as 14 except that the middle term has the opposite sign. Therefore, if we subtract the two equations, the first and last terms cancel and we get

$$\begin{aligned} 4\Delta r_1 \sum_{i=1}^{3} M_{0i}\Delta x_1^i &= 0 && (17)\\ \sum_{i=1}^{3} M_{0i}\Delta x_1^i &= 0 && (18) \end{aligned}$$

We can again use the arbitrariness of the Δx_1^is and suppose that $\Delta x_1^1 = \Delta x_1^2 = 0$; $\Delta x_1^3 = \Delta r$. This gives $M_{03} = 0$. Similar cases show that $M_{01} = M_{02} = M_{03} = 0$. We can therefore rewrite 14 as

$$0 = M_{00}(\Delta r_1)^2 + \sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}(\Delta x_1^i)(\Delta x_1^j) \tag{19}$$

What can we do with the double sum term? Suppose we consider again the case where $\Delta x_1^1 = \Delta x_1^2 = 0$; $\Delta x_1^3 = \Delta r_1$. This gives us

$$0 = M_{00}(\Delta r_1)^2 + M_{33}(\Delta r_1)^2 \tag{20}$$

from which we conclude

$$M_{33} = -M_{00} \tag{21}$$

Similarly, we obtain

$$M_{11} = M_{22} = M_{33} = -M_{00} \tag{22}$$

Now suppose $\Delta x_1^1 \neq 0$; $\Delta x_1^2 \neq 0$; $\Delta x_1^3 = 0$. Then

$$(\Delta r_1)^2 = (\Delta x_1^1)^2 + (\Delta x_1^2)^2 \tag{23}$$

$$0 = M_{00}(\Delta r_1)^2 + M_{11}(\Delta x_1^1)^2 + M_{22}(\Delta x_1^2)^2 + 2M_{12}(\Delta x_1^1)(\Delta x_1^2) \tag{24}$$

$$= M_{00}(\Delta r_1)^2 - M_{00}((\Delta x_1^1)^2 + (\Delta x_1^2)^2) + 2M_{12}(\Delta x_1^1)(\Delta x_1^2) \tag{25}$$

$$= M_{00}(\Delta r_1)^2 - M_{00}(\Delta r_1)^2 + 2M_{12}(\Delta x_1^1)(\Delta x_1^2) \tag{26}$$

$$= 2M_{12}(\Delta x_1^1)(\Delta x_1^2) \tag{27}$$

from which we conclude

$$M_{12} = 0 \tag{28}$$

By other combinations of non-zero Δx_1^is, we conclude that all the M_{ij}s for $i \neq j$ are zero, so we get finally

$$M_{ij} = -M_{00}\delta_{ij} \tag{29}$$

where δ_{ij} is the Kronecker delta ($\delta_{ij} = 0$ if $i \neq j$; $\delta_{ij} = 1$ if $i = j$). Putting all this together, and remembering that these results for $M_{\alpha\beta}$ are valid for *all* transformations even though they were derived for the special case of $\Delta s = 0$, we get

$$\Delta s_2^2 = -M_{00}(-(\Delta t_1)^2 + (\Delta x_1)^2 + (\Delta y_1)^2 + (\Delta z_1)^2) \quad (30)$$
$$= -M_{00}\Delta s_1^2 \quad (31)$$

That is, the intervals must transform by a simple multiplicative factor which may depend on the relative velocity.

To complete the proof, we would like to show that $-M_{00} = 1$, so that $\Delta s_2^2 = \Delta s_1^2$ and the two intervals are equal. We can write $-M_{00}$ as some function $\phi(\boldsymbol{v})$ which may depend on the relative velocity $\boldsymbol{v}$ between the two frames.

Suppose observer O_1 places a metre stick along the y axis (that is, perpendicular to the relative motion of the two frames). We've argued earlier that both observers see the length of the stick as 1 m, and that if one observer makes measurements of the two ends of the stick at the same time in his frame, then the two events that define these measurements are also simultaneous to the other observer. In other words, the interval Δs_1^2 between the two measurements seen by O_1 is the same as the interval Δs_2^2 seen by O_2. This argument doesn't depend on the direction of $\boldsymbol{v}$, provided $\boldsymbol{v}$ is perpendicular to the metre stick. Thus in this case, since $\Delta s_2^2 = \Delta s_1^2$ is known to be true, we must have $\phi(\boldsymbol{v}) = 1$. Thus we have the general result that, for any two events, the spacetime interval between them is the same in all inertial frames:

$$\boxed{\Delta s_2^2 = \Delta s_1^2} \quad (32)$$

You might be bothered a bit by this argument, as it seems to imply that we can say that intervals are equal only for events that are spatially separated by a distance perpendicular to the relative motion. I think the point is that the relation 31 was derived without any assumption about the relative positioning of the events, and is therefore true for *all* pairs of events. We merely used the specific case of a metre stick oriented perpendicular to $\boldsymbol{v}$ to determine $-M_{00}$.

In any case, it is possible to verify 32 directly once we have the Lorentz transformations.

PINGBACKS

Pingback: Spacetime diagrams - two observers
Pingback: Lorentz transformations
Pingback: Four-vectors - basics
Pingback: Invariant hyperbolas
Pingback: Geodesics - paths of longest proper time

Pingback: Four-velocity - an example
Pingback: Metric tensor under Lorentz transformation