LORENTZ TRANSFORMATIONS AS A LINEAR MAP

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog.

When we derived the Lorentz transformations, we assumed that they formed a linear map from one set of coordinates to another. That is, we assumed that, for any two vectors $\mathbf{X}$ and $\mathbf{Y}$ and real number k, the following two relations hold for the Lorentz transformation L:

$$L(\mathbf{X}+\mathbf{Y}) = L(\mathbf{X})+L(\mathbf{Y}) \tag{1}$$
$$L(k\mathbf{X}) = kL(\mathbf{X}) \tag{2}$$

The first condition is called *additivity* and the second *homogeneity*. These assumptions allow us to write the Lorentz transformation as a matrix (assuming that the transformed frame is moving parallel to the x axis of the other frame with speed v)

$$L = \begin{pmatrix} a & b & 0 & 0 \\ d & e & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3}$$

for some functions a, b, d and e that turn out to be functions of v.

Here we have a closer look at this assumption of linearity. In particular, if we make less restrictive assumptions about L can we derive the linearity condition? To this end, let us assume that L maps straight lines into straight lines, is invertible (that is, we can transform in both directions between the two frames), continuous (so that if two source vectors are infintesimally close to each other, then so are the respective transformed vectors) and transforms the origin in one frame into the origin in the other. The assumption about mapping straight lines is reasonable since if an object moves at a constant speed in one frame, its world line in that frame is a straight line, and if we transform to another frame moving at a constant velocity relative to the first frame, then we would expect the object to be moving at a constant speed relative to the new frame as well, so its world line is still straight. The assumption about transforming the origin means that the two frames coincide at one event which we define as the origin in the two frames.

A number of consequences follow from these assumptions. First, the requirement that an inverse transformation exists means that parallel lines in one frame are transformed into parallel lines in the other. Since we can always orient the frames so that their x axes coincide and their respective y and z axes are parallel, we need consider only the transformation in two dimensions x and t. If L transformed two parallel lines into lines that are not parallel, then, since we're in two dimensions, these two transformed lines must intersect somewhere. That would mean that for that intersection point, it is impossible to define an inverse transformation, since two distinct points (one on each parallel line) in the first frame are mapped into a single point in the second frame. There is no way we could determine which point in the first frame gave rise to the point in the second, so there is no unique inverse.

Given that fact, suppose we now consider a parallelogram in the first system, and put one corner of the parallelogram at the origin. Define vectors $\mathbf{X}$ and $\mathbf{Y}$ to be the vectors along the two edges that meet at the origin. These two vectors also define the directions of the other two edges of the parallelogram, since opposite sides are parallel. Since parallel lines transform into other parallel lines, L must transform the parallelogram into another parallelogram, and since it also transforms the origin into the other origin, that parallelogram will also have one corner at the origin. The transformed vectors are $L(\mathbf{X})$ and $L(\mathbf{Y})$.

In the first frame, the diagonal of the parallelogram is the vector $\mathbf{X} + \mathbf{Y}$, while in the second frame, it is $L(\mathbf{X}) + L(\mathbf{Y})$. However, the transformed diagonal must also be $L(\mathbf{X} + \mathbf{Y})$, so we get the additivity condition

$$L(\mathbf{X} + \mathbf{Y}) = L(\mathbf{X}) + L(\mathbf{Y}) \tag{4}$$

Note that the assumption that one origin transforms into the other origin is essential here. If the transformation did a translation of coordinates (by, for example, shifting everything one unit to the right in the x direction), the additivity condition would not apply. For example, suppose $\mathbf{X} = (0,0)$ and $\mathbf{Y} = (1,0)$ and the map transforms all vectors by +1 in the x direction. Then $\mathbf{X} + \mathbf{Y} = (1,0)$ and $L(\mathbf{X} + \mathbf{Y}) = (2,0)$. However, $L(\mathbf{X}) = (1,0)$ and $L(\mathbf{Y}) = (2,0)$, so that $L(\mathbf{X}) + L(\mathbf{Y}) = (3,0) \neq L(\mathbf{X} + \mathbf{Y})$.

Now suppose we have a sequence of vectors $\mathbf{Y}_j$ and that this sequence converges to $\mathbf{X}$ as $j \to \infty$. For large enough j, $\mathbf{Y}_j$ will be arbitrarily close to $\mathbf{X}$, so by the assumption of continuity of L, the sequence of transformed vectors must also converge. That is, $L(\mathbf{Y}_j) \to L(\mathbf{X})$ as $j \to \infty$. So we get

$$L(\mathbf{X}+\mathbf{Y}_j) \rightarrow L(\mathbf{X}+\mathbf{X}) \tag{5}$$

$$= L(2\mathbf{X}) \tag{6}$$

$$L(\mathbf{X}+\mathbf{Y}_j) \rightarrow 2L(\mathbf{X}) \tag{7}$$

where in the last line we used the additivity property just proved. We can extend this argument by induction, since we've just proved the anchor step. That is, we can assume that for some integer m, $L((m-1)\mathbf{X}+\mathbf{Y}_j) \rightarrow mL(\mathbf{X})$ and $L((m-1)\mathbf{X}+\mathbf{Y}_j) \rightarrow L(m\mathbf{X})$. Then

$$L(m\mathbf{X}+\mathbf{Y}_j) = L((m-1)\mathbf{X}+\mathbf{Y}_j+\mathbf{X}) \tag{8}$$

$$= L((m-1)\mathbf{X}+\mathbf{Y}_j)+L(\mathbf{X}) \tag{9}$$

$$\rightarrow mL(\mathbf{X})+L(\mathbf{X}) \tag{10}$$

$$= (m+1)L(\mathbf{X}) \tag{11}$$

The second line uses additivity, and the third line uses the inductive assumption. Since $L(m\mathbf{X}+\mathbf{Y}_j) \rightarrow L((m+1)\mathbf{X})$, we have shown that homogeneity applies for positive integers, that is, for m a positive integer:

$$L(m\mathbf{X}) = mL(\mathbf{X}) \tag{12}$$

We can extend this to negative integers by noting that in the proof of additivity above, we could equally well have chosen the other diagonal of the parallelogram which is given by the vector $\mathbf{X}-\mathbf{Y}$, and this would give use the result that

$$L(\mathbf{X}-\mathbf{Y}) = L(\mathbf{X})-L(\mathbf{Y}) \tag{13}$$

Therefore, $L(-m\mathbf{X}) - L(\mathbf{0}-m\mathbf{X}) = L(\mathbf{0}) - L(m\mathbf{X}) = \mathbf{0} - mL(\mathbf{X}) = -mL(\mathbf{X})$, since one origin is mapped into the other.

Next we can prove that homogeneity applies for any rational number $q = p/n$: $L(q\mathbf{X}) = qL(\mathbf{X})$. If we let $\mathbf{X} = n\mathbf{Z}$ then

$$L(q\mathbf{X}) = L((p/n)(n\mathbf{Z})) \tag{14}$$

$$= L(p\mathbf{Z}) \tag{15}$$

$$= pL(\mathbf{Z}) \tag{16}$$

$$= \frac{p}{n}nL(\mathbf{Z}) \tag{17}$$

$$= \frac{p}{n}L(n\mathbf{Z}) \tag{18}$$

$$= qL(\mathbf{X}) \tag{19}$$

Finally, we need to prove homegeneity for all real numbers k. It is possible to construct a sequence of rational numbers that converges on any irrational number, since we can write any irrational number in decimal form as an integer followed by a non-repeating fractional part, so we can take as our sequence the set of rational numbers q_j which retains j decimal places in the expansion of the irrational number. This sequence must converge to k as $j \to \infty$. For example, the sequence of rational numbers 3, 3.1, 3.14, 3.141, 3.1415... converges to π if we keep adding on an extra decimal place at each stage.

In that case, we can use the result above for rational numbers together with the continuity of the map to say that

$$
\begin{align}
L(q_j \mathbf{X}) \quad &= \quad q_j L(\mathbf{X}) \tag{20} \\
&\to \quad k L(\mathbf{X}) \tag{21}
\end{align}
$$

and

$$
L(q_j \mathbf{X}) \to L(k\mathbf{X}) \tag{22}
$$

Thus we get the final result of homogeneity over all real numbers:

$$
L(k\mathbf{X}) = k L(\mathbf{X}) \tag{23}
$$

PINGBACKS

Pingback: Lorentz transformations
Pingback: Lorentz transformations: derivation from symmetry