Relativity minimum for PHY 606

Paulo F. Bedaque
University of Maryland

A. Vectors in Spacetime

We can add time to the three space coordinates and think of physics in *spacetime* instead. This can be down whether or not relativity is relevant for the problem at hand. You certainly drew diagrams of position t versus time in high school, long before hearing about relativity. Those diagrams are a picture of the spacetime! In any case, it seems natural to extend the concept of vectors and tensors to 4 dimensions and have 4-component vectors, ... Four-vectors and four-tensors, are they are called, will be useful only after we make a few modifications in the formalism for three-vectors and three-tensors we discussed before. The reason is that while the laws of physics are valid for any coordinate system in space, regardless of its orientation, we don't expect coordinate systems to be equivalent after a 4-dimensional rotation. Time and space are pretty different after all. There is however, another kind of transformation involving the 4 dimensions of spacetime that *is* indeed a symmetry of Nature. To find out what they are, let us look at the two principles special relativity is based upon

1. The laws of Physics are the same for any inertial observable
2. The speed of light (in vacuum) is the same in any inertial frame

The first principle is reasonable; the second seems crazy as "everybody knows" that objects seem to us to move faster if we ourselves move in the opposite direction. Why should it be different with light? Yet, it is an unavoidable conclusion if you are to believe in the validity of Maxwell's equations in any inertial frame, as Maxwell's equations predict waves moving with speed c but not waves moving at any other speed. Many sources discuss the experimental/theoretical motivations for these postulates and their consequences. Here we will only discuss a formalism (space time 4-tensors) used to calculate things in relativistic theories like electromagnetism.

The location of an event (something that occurs at a certain point in space and time) is determined by four numbers, $x^0 = ct, x^1 = x, x^2 = y$, and $x^3 = z$. We can group this coordinates in a four component object $x^\mu = (x^0, x^1, x^2, x^3)$. The index μ (as well as other greek indices) take the values $\mu = 0, 1, 2$ or 3, as opposed to the latin indices $i, j, k, ..$ that run only from 1 to 3. x^μ specifies the location of an event in space time just like the usual 3-vector $\vec{x} = (x, y, z)$ specifies a spatial location. You should think of four vectors the way you think of regular 3-vectors: a line with an arrow at the end. Except they live on a space with one more dimension. The cartesian components of a 3-vector are dependent on the reference frame one uses. One can say that its components "transform" as the coordinate system is changed.

Now imagine two different events with coordinates x^μ and y^μ "connected by a light ray", that is, such that a light ray passes through both of them. Since light moves with the speed of light, the spatial distance between them should equal the time difference times c

$$\underbrace{(x^0 - y^0)^2}_{(ct)^2} - \underbrace{(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2}_{(\vec{x}-\vec{y})^2} = 0. \tag{1}$$

If we now observe the same light ray in a different frame the coordinates of each event would be different ($x^{'\mu}$ and $y^{'\mu}$) but the fact that they are connected by a light ray is still true. This implies that

$$\underbrace{(x^{'0} - y^{'0})^2}_{(ct')^2} - \underbrace{(x^{'1} - y^{'1})^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2}_{(\vec{x}-\vec{y})^2} = 0. \tag{2}$$

It is not true that either the spatial distance between the two events, or the time difference between them is the same as seen by the two observers. But the combination $(c\Delta t)^2 - \Delta\vec{x}^2$ is invariant. This is similar to the fact that, in three-domensional space, the combination

$$(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2 \tag{3}$$

is invariant. So we define the "length" of a vector in spacetime by

$$|v|^2 = (v^0)^2 - (v^1)^2 - (v^2)^2 - (v^3)^2. \tag{4}$$

The only unusual thing about this is that this "length" is not necessarily positive. This flipped sign in the definition of length will require a few changes when dealing with 4-vectors and 4-tensors and compared to the three dimensional case.

In three dimensions, the length of a vector $|\vec{v}|^2 = (v^1)^2 + (v^2)^2 + (v^3)^2$ is invariant under rotations. What is the set of transformations that keep eq.(3) invariant? Those are the Lorentz transformations. They include regular 3D rotations (keeping v^0 fixed and mixing the spatial components with an orthogonal matrix) and also boosts like

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} \underbrace{\begin{pmatrix} \gamma & 0 & 0 & -\gamma v/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma v/c & 0 & 0 & \gamma \end{pmatrix}}_{L} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}, \quad \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{5}$$

In general, Lorentz transformations are the ones keeping eq.(3) invariant. The coordinate frames related by Lorentz transformations are not only a mathematical device. Real observers moving at constant speed in relation to each other would measure space and time in such a way that the coordinates of an event seen by different observers would be related by a Lorentz transformation. This is not obvious. To prove it we would have to discuss how an observer actually measures distances and time intervals. Again, any book on relativity will explain this essential fact; Einstein original article is scarily understandable. Here, we will assume eq. (5)describes the relation between measurements of different observers and explore the formalism this leads to.

Lorentz transformations do not preserve $\sum_{\mu=0}^{3} v^\mu v^\mu$. In order to get an invariant we define the covariant components of a vector (denoted by lower indices) as $v_\mu = (v^0, -v^1, -v^2, -v^3)$. Then $\sum_{\mu=0}^{3} v^\mu v_\mu$ is an invariant. (Sometimes we say "covariant" and "contravariant" components of the same vector, implying they are the same object presented in two different ways. Other times we say "covariant vector" or "contravariant vector", implying they are two different objects. But both ways of saying mean exactly the same thing. I know, this is confusing, but I was not the one who invented it.) The regular components of vectors (also known as contravariant components) transform as

$$\boxed{v^\mu = \sum_{\nu=0}^{3} L^\mu{}_\nu v^\nu}. \tag{6}$$

We can find how the covariant components v_μ change under a change of basis by demanding that $\sum_{\mu=0}^{3} v^\mu v^\mu$ remains the same. We find

$$\boxed{v'_\mu = \sum_{\nu=0}^{3} v_\nu (L^{-1})^\nu{}_\mu}. \tag{7}$$

It is easy to check now that, with these transformation laws the combination $\sum_{\mu=0}^{3} v^\mu w_\mu$ is indeed invariant [1]

$$v'_\mu w'^\mu = (L^{-1})^\beta{}_\mu v_\beta L^\mu{}_\alpha w^\alpha = \underbrace{(L^{-1})^\beta{}_\mu L^\mu{}_\alpha}_{\delta^\beta{}_\alpha} v_\beta w^\alpha = \delta^\beta{}_\alpha v_\beta w^\alpha = v_\alpha w^\alpha. \tag{8}$$

Similarly, we can have 4-tensors with any number of upper and lower indices. Each index transforms like the contra- or co-variant components of a vector. For instance, a rank-two mixed (covariant and contravariant) tensor transforms as

$$T'^\mu_\nu = (L^{-1})^\beta{}_\nu L^\mu{}_\alpha T^\alpha_\beta. \tag{9}$$

A sum over any two repeated index (as in $v^\mu v_\mu$) leads to another quantity transforming like a tensor.

Just like tree-tensors, we can multiply any number of tensors of any rank and sum over pair of repeated indices and we will get another tensor. The only difference between 3- and 4-tensors is that the two indices contracted (summed over) need to be one covariant (downstairs) and one contravariant (upstairs).

---

[1] At this point we revert to the Einstein convention and drop the summations signs. Any pair of identical indices, one upstairs and one downstairs, are assumed to be summed over.

We can find the covariant components of a vector from its contravariant components as

$$\boxed{v_\mu = g_{\mu\nu} v^\nu} \tag{10}$$

where the symbol $g_{\mu\nu}$ (called the "metric") stands for:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{11}$$

It turns out that $g_{\mu\nu}$ is an invariant tensor, that is, it had the same components in any orthonormal basis [2] This can be verified by using explicit expressions for the Lorentz transformations like eq. (5).

$g_{\mu\nu}$ can also be used to relate co- and contravariant components of a higher rank tensor tensor

$$T_{\mu\nu} = g_{\mu\alpha} g_{\nu\beta} T^{\alpha\beta}. \tag{12}$$

The inverse matrix, $g^{\mu\nu}$ (with the same components as $g_{\mu\nu}$) can be used to raise indices:

$$v^\mu = g^{\mu\nu} v_\nu. \tag{13}$$

since $g_{\mu\nu} g^{\nu\lambda} = \delta^\lambda_\mu$. Using the $g_{\mu\nu}$ tensor we can give a general characterization of Lorentz transformations. By definition they keep eq. (3) invariant, so

$$g_{\mu\nu} v'^\mu w'^\nu = g_{\mu\nu} L^\mu{}_\alpha L^\nu{}_\beta v^\alpha w^\beta = g_{\mu\nu} v^\mu w^\nu \tag{14}$$

for any v and w so

$$g_{\mu\nu} L^\mu{}_\alpha L^\nu{}_\beta = g_{\alpha\beta}. \tag{15}$$

Notice that the same calculation in 3D would have $\delta_{\mu\nu}$ instead of $g_{\mu\nu}$ and the relation above would be equivalent to the statement that $LL^T = \mathbb{1}$.

We saw before that the space derivatives transform under rotations as a vector. The spacetime derivatives transform a *covariant* vector

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} = \frac{\partial (L^{-1})^\nu{}_\alpha x'^\alpha}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} = (L^{-1})^\nu{}_\mu \frac{\partial}{\partial x^\nu}. \tag{16}$$

Derivatives are so frequent that we will use the notation

$$\partial_\mu = \frac{\partial}{\partial x^\mu}. \tag{17}$$

Finally, just like in 3D we had the invariant tensor ϵ^{ijk}, in 4D Minkowski space we have the invariant tensor $\epsilon^{\mu\nu\lambda\rho}$. The proof that this is an invariant tensor is essentially the same as in 3D.

Maybe now it would be a good idea to summarize everything in the Prime Directive:

One can make tensors by multiplying any number of tensors of any rank, including the invariant tensors $\delta^\nu_\mu, g_{\mu\nu}, g^{\mu\nu}$ and $\epsilon^{\mu\nu\lambda\rho}$, and contracting co- and contravariant indices pairwise. There is no other way.

Everything else is not Lorentz invariant so it cannot describe Nature.

[2] In Minkowski space an orthonormal basis is one where the four vectors are orthogonal, one has magnitude 1 and three have magnitude -1.