Relativity minimum for PHY 606

Paulo F. Bedaque
University of Maryland

A. Vectors in space

Who doesn't know about regular three dimensional vectors? We are so used to them that we don't even remember how to define them. We just think of them as a line going from the origin to a certain point in space with an arrow on the tip. And that's a good way of picturing them. We are also familiar with the idea that, *giving a cartesian coordinate system*, the can be represent by three numbers, their coordinates in said system. Of course, if we change the coordinate system, the same vector will have different components. For instance, after a rotation by an angle θ around the z axis the coordinates of a vector changes from (v_x, v_y, v_z) to

$$\begin{pmatrix} v'_x \\ v'_y \\ v'_z \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}. \tag{1}$$

In general, the relation between the components of any 3-vector (representing position, momentum, ...) change as we go from one coordinate system to another related by a (fixed) rotation through an orthogonal matrix (with determinant=+1):

$$\begin{pmatrix} v'_x \\ v'_y \\ v'_z \end{pmatrix} = \underbrace{\mathbf{O}}_{3\times 3 \text{ orthogonal matrix}} \begin{pmatrix} v'_x \\ v'_y \\ v'_z \end{pmatrix}. \tag{2}$$

A better way to represent the equation above is by using indices

$$v'^i = \sum_{j=1}^{3} O^{ij} v^j. \tag{3}$$

We can do a lot of vector algebra and vector calculus without writing their components in any particular basis. That is a very good thing. But not everything so it's useful to learn how to work with components and, for that, you need to learn how to manipulate indices. Basically everything you need to know is contained in the homework 1, problem 1.

We can make a number out of two vectors by taking the scalar product. In terms of coordinates the scalar product is

$$\vec{v}.\vec{w} = \sum_{i=1}^{3} v^i w^i. \tag{4}$$

This combination is special because is an it invariant. By that I mean that, in another coordinate system (related to the original one by a rotation) the components v^i, w^i of the two vectors will be different by the combination $\sum_{i=1}^{3} v^i w^i$ is still the same. Geometrically this is obvious. The scalar product equals the product of the magnitudes of the two vectors and the cosine of the angle between them. But magnitudes and angles are geometrical concepts independent of any coordinate system and so is the scalar product. Thinking only about the components though is not at all obvious that the combination $\sum_{i=1}^{3} v^i w^i$ is invariant. It is good practice to show that it is

$$\sum_{i=1}^{3} v'^i w'^i = \sum_{i=1}^{3} \underbrace{\sum_{j=1}^{3} O^{ij} v^j}_{v'^i} \underbrace{\sum_{k=1}^{3} O^{ik} w^k}_{w'^i} = \sum_{i,j,k=1}^{3} (O^T)^{ki} O^{ij} v^j w^k = \sum_{k,j=1}^{3} \delta^{kj} v^j w^k = \sum_{j=1}^{3} v^j w^j, \tag{5}$$

where $\mathbf{O}^T$ is the transpose of the matrix $\mathbf{O}$ and we use the property $\mathbf{O}^T\mathbf{O} = \mathbb{1}$ of orthogonal matrices.

One may wonder if there are other combinations of the components of two vectors that is also invariant, that is, independent of the coordinate system used. The answer is no. What we can do is to make *three* different combinations

$$u^1 = v^2 w^3 \; - \; v^3 w^2, \tag{6}$$
$$u^2 = v^3 w^1 \; - \; v^1 w^3, \tag{7}$$
$$u^3 = v^1 w^2 \; - \; v^2 w^1, \tag{8}$$
$$\tag{9}$$

that are the components of a vector (usually denoted by $\vec{v} \times \vec{w}$). By that I mean that if the components v^i and w^i transform under rotations as in eq.(3), the components u^i will also transform as in eq.(3). Thee way to prove this is to first write the definition of u^i as

$$u^i = \sum_{j,k=1}^{3} \epsilon^{ijk} v^j w^k, \tag{10}$$

with

$$\epsilon^{ijk} = \begin{cases} 1 \text{ if ijk} = 123 \text{ or cyclic permutation} \\ -1 \text{ if ijk} = 213 \text{ or cyclic permutation} \\ 0 \text{ otherwise} \end{cases} \tag{11}$$

The Levi-Civita symbol ϵ^{ijk} satisfies the relation

$$O^{il} O^{jm} O^{kn} \epsilon^{lmn} = \epsilon^{ijk}. \tag{12}$$

The relation above can be shown by observing that the the left-hand side of eq.(12) is antisymmetric under a permutation of two of the indices (ijk), the defining relation of the right-hand side. The proportionality constant can be found to be 1 by looking a the value of both sides when $ijk = 123$. The right-hand side equals 1; the left-hand side is the determinant of **O** that is also 1[1].

The we have

$$u'^i = \sum_{j,k=1}^{3} \epsilon^{ijk} v'^j w'^k = \sum_{j,k,l,m=1}^{3} \epsilon^{ijk} O^{jl} v^l O^{km} w^m = (O^T)^{ji} \epsilon^{jlm} v^l w^m = O^{ij} u^j. \tag{13}$$

One more observation: the components of the∇ operator also transform as a vector:

$$\frac{\partial}{\partial x'^i} = \sum_{j=1}^{3} \frac{\partial x^j}{\partial x'^i} \frac{\partial}{\partial x^j} = \sum_{j,k=1}^{3} \frac{\partial (O^T)^{jk} x'^k}{\partial x'^i} \frac{\partial}{\partial x^j} = \sum_{j=1}^{3} (O^T)^{ji} \frac{\partial}{\partial x^j} = O^{ij} \frac{\partial}{\partial x^j}. \tag{14}$$

That means that the combination $\frac{\partial}{\partial x^i} v^i = \nabla.\vec{v}$ is an invariant and that $\epsilon^{ijk} \frac{\partial}{\partial x^j} v^k = (\nabla \times \vec{v})^i$ are the components of a vector.

Besides vectors (objects with 3 components) we also find objects with more components. They are called *tensors*. Examples that may be familiar to you are the inertia tensor of a rigid body or the quadrupole moment of a charge distribution. In terms of its components they are objects with more than one index. Under a coordinate transformation the components change as

$$T'^{\overbrace{i \cdots j}^{n\ indices}} = \underbrace{O^{ik} \cdots O^{jl}}_{n\ O\ matrices} T^{k \cdots l} \tag{15}$$

Out of tensors and vecotrs we can make other tensors and vectors by summing over an index, for instance

$$T^{ij} \; = \; v^i w^j \; (tensor), \tag{16}$$
$$\sum_{i=1}^{3} T^{iij} \; = \; a^j \; (vector), \tag{17}$$
$$\sum_{i,j,k=1}^{3} M^{ijk} T^{ij} v^k \; = \; a \; (scalar). \tag{18}$$

[1] This proof is sketchy; you were invited to fill in the details in homework 1.

The proof that all these combinations indeed transform as tensors, vectors or scalar is similar to the proof given above that the scalar product is invariant.

The importance of all this is that a lot of physical laws are invariant under rotations, that is, are equally valid in any coordinate system regardless of its orientation. But if we write these laws in terms of components, how can we know they are invariant? The answer is that the only rotation invariant laws are the ones that can be written in terms of vector with indices summed over. So

$$\nabla.\vec{B} = \sum_{i=1}^{3} \frac{\partial}{\partial x^i} B^i = 0, \tag{19}$$

can be a law of physics while

$$\frac{\partial}{\partial x} B^y + \frac{\partial}{\partial y} B^z + \frac{\partial}{\partial z} B^x = 0 \tag{20}$$

cannot (at least if the laws are rotationally invariant).

B. Vectors in Spacetime

We can add time to the three space coordinates and think of physics in *spacetime* instead. It seems natural to extend the concept of vectors and tensors to 4 dimensions and have 4-component vectors, ... This will be useful only after a few modifications. The reason is that while the laws of physics are valid for any coordinate system in space, regardless of its orientation, but 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 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 but, if you are to believe in the validity of Maxwell's equations in any inertial frame, it's unavoidable as Maxwell's equations predict waves moving with speed c but not any other velocity. Many places 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

$$(x^0 - y^0)^2 - \underbrace{(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2}_{(\vec{x}-\vec{y})^2} = 0. \tag{21}$$

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

$$(x^{'0} - y^{'0})^2 - \underbrace{(x^{'1} - y^{'1})^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2}_{(\vec{x}-\vec{y})^2} = 0. \tag{22}$$

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

$$(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2 \tag{23}$$

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

$$|v|^2 = (v^0)^2 - (v^1)^2 - (v^2)^2 - (v^3)^2. \tag{24}$$

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.(23) 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

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} \underbrace{\begin{pmatrix} \gamma & 0 & 0 & -\gamma v/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma v/c & 0 & 0 & \gamma \end{pmatrix}}_{L} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}, \quad \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{25}$$

In general, Lorentz transformations are the one keeping eq.(23) 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.

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. The regular components of vectors (also known as contravariant components) transform as

$$v^\mu = \sum_{\nu=0}^{3} L^\mu_\nu v^\nu, \tag{26}$$

while the covariant components transform as

$$v_\mu = \sum_{\nu=0}^{3} (L^{-1})^\nu_\mu v_\nu. \tag{27}$$

It is easy to check now that the combination $\sum_{\mu=0}^{3} v^\mu w_\mu$ is invariant

$$v'_\mu w'^\mu = (L^{-1})^\beta_\mu v_\beta L^\mu_\alpha w^\alpha = \delta^\beta_\alpha v_\beta w^\alpha = v_\alpha w^\alpha. \tag{28}$$

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,

$$T'^\mu_\nu = (L^{-1})^\beta_\nu L^\mu_\alpha T^\alpha_\beta. \tag{29}$$

A sum over any two repeated index (as in $\sum_{\mu=0}^{3} v^\mu v_\mu$) leads to another quantity transforming like a tensor. This operation is so common that there's a convention of omitting the summation sign when one upper and one lower index is repeated:

$$v_\mu w^\mu \text{ means } \sum_{\mu=0}^{3} v_\mu w^\mu. \tag{30}$$

It is convenient to define a tensor $g_{\mu\nu}$ with components

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{31}$$

in any reference frame (convince yourself this is really a tensor, namely, that under the transformation (29) its components won't change). It can be used to relate co- and contravariant components of a vector or tensor

$$T_{\mu\nu} = g_{\mu\alpha} g_{\nu\beta} T^{\alpha\beta}. \tag{32}$$

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

$$v^{\mu} = g^{\mu\nu} v_{\nu}. \tag{33}$$

Using the $g_{\mu\nu}$ tensor we can give a general characterization of Lorentz transformations. By definition they keep eq (23) invariant, so

$$g_{\mu\nu} v'^{\mu} w'^{\nu} = g_{\mu\nu} L^{\mu}_{\alpha} L^{\nu}_{\beta} v^{\alpha} w^{\beta} = g_{\mu\nu} v^{\mu} w^{\nu} \tag{34}$$

for any v and w so

$$g_{\mu\nu} L^{\mu}_{\alpha} L^{\nu}_{\beta} = g_{\alpha\beta}. \tag{35}$$

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 a *covariant* vector

$$\frac{\partial}{\partial x'^{\mu}} = \frac{\partial x^{\nu}}{\partial x'^{\mu}} \frac{\partial}{\partial x^{\nu}} = \frac{\partial (L^{-1})^{\nu}_{\alpha} x'^{\alpha}}{\partial x'^{\mu}} \frac{\partial}{\partial x^{\nu}} = (L^{-1})^{\nu}_{\mu} \frac{\partial}{\partial x^{\nu}}. \tag{36}$$

Derivatives are so frequent that we will use the notation

$$\partial_{\mu} = \frac{\partial}{\partial x'^{\mu}}. \tag{37}$$

Lorentz invariant equations are written in terms o 4-tensors with upper indices "contracted" (summed over) to lower indices. Everything else is not Lorentz invariant so it cannot describe Nature.