FOUR-VECTORS - BASICS

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: 17 Jan 2021.

In relativity, a *four-vector* is a vector with four components. A general four-vector $\vec{A}$ is denoted by a letter with an arrow on top, and its four components are defined as

$$\vec{A} \overset{\mathcal{O}}{\rightarrow} \left(A^0, A^1, A^2, A^3\right) \tag{1}$$

This definition contains an important point. The symbol $\overset{\mathcal{O}}{\rightarrow}$ is used instead of an equals sign, and is to be read 'the vector $\vec{A}$ in the coordinate system used by observer $\mathcal{O}$ has components $\left(A^0, A^1, A^2, A^3\right)$.

A curious point, if you haven't seen it before, is that the components of a vector are given with superscript indexes, rather than subscripts, as are more usual in linear algebra. The reason for this becomes more apparent when we get a bit deeper into the theory, but for now it should just be accepted. It is important to note that these superscripts are *not* exponents, but merely labels.

In Euclidean three-dimensional space, a three-vector (one with three components) has different components depending on which coordinate system we are using to describe it. The important point is that the vector exists independently of the coordinate system, and the components used to describe it depend on that coordinate system; the vector itself does not.

In two-dimensionsal space, for example, we might define a vector $\vec{x}$ in one coordinate system by the coordinates (1,0), which means that it extends one unit along (and is parallel to) the x axis. If we rotate the coordinate system by 90 degrees so that the x axis rotates into the y axis, the vector itself does not move and now extends one unit in the $-y$ direction, so its new coordinates are $(0,-1)$. It is not correct to say that $\vec{x}$ *equals* either of these coordinate descriptions; it is correct only to say it has these numerical coordinates in two particular systems. In yet other systems, it will have other coordinates.

In Euclidean space, a common vector is the *displacement* vector, which measures the distance and direction from one point to another. Again, the locations of the points and the vector connecting them do not depend on the coordinate system so that the vector will have different components in different systems. However, the *length* or *magnitude* of the distance between the points is invariant under a change of coordinates and is given by the standard Euclidean distance formula

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

assuming we are using a cartesian system.

1. Four-vectors in relativity

In relativity, the analog of the displacement vector is the interval between two events. Remember that an event is something that occurs at a specific time and place, although the actual numerical values of this time and place depend on the observer who measures them. We saw in an earlier post that the square of the interval between two events is an invariant, and is given by

$$\Delta s^2=-(\Delta t)^2+(\Delta x)^2+(\Delta y)^2+(\Delta z)^2 \tag{3}$$

We can therefore define the displacement vector in relativity as

$$\Delta\vec{x}=(\Delta t,\Delta x,\Delta y,\Delta z) \tag{4}$$

The magnitude of this vector is usually defined as the square of the interval, so we have

$$\Delta\vec{x}^2=-(\Delta t)^2+(\Delta x)^2+(\Delta y)^2+(\Delta z)^2 \tag{5}$$

As usual, the interval squared can be positive, zero or negative, depending on the separation of the events.

The individual components of a four-vector obey the usual rules for vectors under addition and scalar multiplication, so we get, for a scalar k

$$\vec{A}+\vec{B} \overset{\mathcal{O}}{\rightarrow} (A^0+B^0,A^1+B^1,A^2+B^2,A^3+B^3) \tag{6}$$

$$k\vec{A} \overset{\mathcal{O}}{\rightarrow} (kA^0,kA^1,kA^2,kA^3) \tag{7}$$

We've seen that the space and time coordinates of an event transform between observers by using the Lorentz transformations. In order to write this transformation in an efficient and compact way we need to introduce a bit more notation.

First, we have seen in 1 how a vector is represented in one coordinate system. In the system of another observer $\bar{\mathcal{O}}$ we can write for the same vector:

$$\vec{A} \xrightarrow{\bar{\mathcal{O}}} (A^{\bar{0}}, A^{\bar{1}}, A^{\bar{2}}, A^{\bar{3}}) \tag{8}$$

Notice that the bars go on the *indexes* of the vector components and not on the symbol A that we are using to represent the vector itself. This emphasizes the fact that the vector doesn't change when we change coordinate systems; only the *coordinates* used to describe the vector change. Thus any time we see a bar over the coordinate index, it is a coordinate measured by observer $\bar{\mathcal{O}}$.

Back to the Lorentz transformations. The form in which we've seen these transformations is that of four equations:

$$t_2 = \frac{1}{\sqrt{1-v^2}}(t_1 - vx_1) \tag{9}$$

$$x_2 = \frac{1}{\sqrt{1-v^2}}(x_1 - vt_1) \tag{10}$$

$$y_2 = y_1 \tag{11}$$

$$z_2 = z_1 \tag{12}$$

These transformations were derived for the case of displacement coordinates, but we can extend the definition to all four-vectors. That is we can say that the way to transform the unbarred coordinates into the barred coordinates is by using the Lorentz transformation on the vector's components.

$$A^{\bar{0}} = \gamma(A^0 - vA^1) \tag{13}$$

$$A^{\bar{1}} = \gamma(A^1 - vA^0) \tag{14}$$

$$A^{\bar{2}} = A^2 \tag{15}$$

$$A^{\bar{3}} = A^3 \tag{16}$$

where we've used the shorthand symbol $\gamma = 1/\sqrt{1-v^2}$. The index 0 corresponds to the time coordinate, and the indexes 1, 2 and 3 to the spatial coordinates.

This set of four equations can be written as a matrix equation, if we define the matrix Λ as

$$\Lambda = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{17}$$

We can now write the transformations as

$$A^{\bar{\alpha}}=\sum_{\beta=0}^{3}\Lambda^{\bar{\alpha}}{}_{\beta}A^{\beta} \tag{18}$$

Here $\Lambda^{\bar{\alpha}}{}_{\beta}$ is the entry from row $\bar{\alpha}$ and column β of the matrix. Again, the convention of writing the row index as a superscript comes in handy in what follows, so don't make the mistake of taking this superscript as an exponent.

2. Summation convention

A notational simplification that is used throughout special and general relativity is the *summation convention*. Sums over indexes such as that over the index β above are very common, and having to write out a summation sign over and over again gets very tedious. As a result, Einstein introduced a summation convention into relativity. Whenever a product of two or more terms contains a pair of identical indexes with one of the pair a superscript and the other a subscript, a summation is *automatically* performed on that index. If the index is a Greek letter, the summation extends over all four components (from 0 to 3); if it is a Latin letter, the summation extends over only the three spatial coordinates (1 to 3). Thus we can write the Lorentz transformation for a general four-vector in the simplified form

$$A^{\bar{\alpha}}=\Lambda^{\bar{\alpha}}{}_{\beta}A^{\beta} \tag{19}$$

Whenever the same index appears on both sides of an equation (such as $\bar{\alpha}$ here), it may be taken to be any of the coordinates within its range (again, Greek = 0, 1, 2, 3; Latin = 1, 2, 3), so this simplified equation actually contains four equations depending on the value assigned to $\bar{\alpha}$.

In both these cases (the index pair used in the summation, and the index pair on opposite sides of the equation), the index concerned is merely a dummy index, and can be replaced by any other index (as long as Greek is replaced by Greek, and Latin by Latin) without changing the meaning of the equation. Thus we could just as well write

$$A^{\bar{\xi}}=\Lambda^{\bar{\xi}}{}_{\tau}A^{\tau} \tag{20}$$

Here, the summation is over $\tau=0..3$ and the index $\bar{\xi}$ can stand for any of 0, 1, 2 or 3.

However, what you *cannot* do is replace a barred index by an unbarred one (or vice versa), since each type of index refers to a particular coordinate system, and the set of coordinates in one system will usually be totally different to the set in the other.

3. Scalar product

Just as a scalar product (often called the 'dot product') exists in 3-dim Euclidean space, there is a scalar product defined in the 4-dim spacetime used in relativity. We've already seen one example of this in the definition 5 of the invariant interval between two events. Here 'invariant' means that the interval has the same value in any inertial frame. Another way of putting this is that the interval is invariant under Lorentz transformations. The scalar product of two four-vectors $\vec{A}$ and $\vec{B}$ is defined as

$$\vec{A}\cdot\vec{B} \equiv -A^0B^0 + A^1B^1 + A^2B^2 + A^3B^3 \tag{21}$$

We can write this more compactly if we define the *Minkowski metric* by the matrix

$$\eta_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{22}$$

Using the summation convention, the scalar product is

$$\vec{A}\cdot\vec{B} = \eta_{\mu\nu}A^\mu B^\nu \tag{23}$$

The scalar product of any two four-vectors is Lorentz invariant. The proof of this is the same as that used for the interval between two events. Consider, as usual, one inertial frame at rest in the lab and another inertial frame moving at speed v relative to the lab frame. The Lorentz transformations for a four-vector are

$$\begin{aligned} A'^0 &= \gamma\left(A^0 - vA^1\right) \\ A'^1 &= \gamma\left(A^1 - vA^0\right) \\ A'^2 &= A^2 \\ A'^3 &= A^3 \end{aligned} \tag{24}$$

and similarly for $\vec{B}$. In the primed frame, the scalar product is (I'll ignore the 2 and 3 components since they remain the same in both frames):

$$\vec{A}'\cdot\vec{B}' = \gamma^2\left[-\left(A^0-vA^1\right)\left(B^0-vB^1\right)+\left(A^1-vA^0\right)\left(B^1-vB^0\right)\right] \tag{25}$$

$$= \gamma^2\left[-A^0B^0+A^1B^1+\right.$$
$$\left.\gamma^2\left[v\left(A^1B^0+A^0B^1-A^0B^1-A^1B^0\right)-v^2\left(A^1B^1-A^0B^0\right)\right]\right. \tag{26}$$

$$= \gamma^2\left[\left(1-v^2\right)\left(-A^0B^0+A^1B^1\right)\right] \tag{27}$$

$$= -A^0B^0+A^1B^1 \tag{28}$$

$$= \vec{A}\cdot\vec{B} \tag{29}$$

PINGBACKS

Pingback: Four-velocity
Pingback: Tensors and one-forms
Pingback: Tensor index notation