Lecture III: Vector algebra in special relativity

(Dated: August 28, 2019)

I. OVERVIEW

This lecture introduces 4-dimensional vectors (*4-vectors*) in special relativity. Some of you may have seen these before. In any case, my purpose here is to both introduce vectors themselves, and also the index notation that we will use throughout the course.

This material corresponds to approximately §§2.1, 2.2, and 2.5 of Schutz. (The physics applications, which are the rest of Chapter 2, will be in my Lecture IV.)

II. DEFINITIONS AND NOTATION

In undergraduate physics, you learned that a *vector* had a magnitude and direction, but was not fixed to any position in space. It could be described by N components, where N is the number of dimensions. This remains true in special relativity. When we consider curved spacetime, we will have to re-define the concept of a vector, but we won't do that yet.

The book denotes the components of a vector $\boldsymbol{x}$ in a given frame $\mathcal{O}$ by the notation:

$$\boldsymbol{x} \underset{\mathcal{O}}{\longrightarrow} (x^0, x^1, x^2, x^3) = \{x^\alpha\}, \tag{1}$$

If we consider another frame $\bar{\mathcal{O}}$, from here forward we will write

$$\boldsymbol{x} \underset{\bar{\mathcal{O}}}{\longrightarrow} (x^{\bar{0}}, x^{\bar{1}}, x^{\bar{2}}, x^{\bar{3}}) = \{x^{\bar{\alpha}}\}, \tag{2}$$

where we place the bar over the coordinate index to indicate that it is the same vector $\boldsymbol{x}$ but written in the barred coordinates.

If we transform from one frame to another, we have already written the matrix notation $\bar{\boldsymbol{x}} = \boldsymbol{\Lambda}\boldsymbol{x}$. If we write this in component notation, it is

$$x^{\bar{\alpha}} = \sum_{\beta=0}^{3} \Lambda^{\bar{\alpha}}{}_{\beta} x^\beta. \tag{3}$$

Here $\Lambda^{\bar{\alpha}}{}_{\beta}$ is an entry in the 4×4 transformation matrix $\boldsymbol{\Lambda}$. A coordinate transformation matrix (here $\boldsymbol{\Lambda}$; in GR, it will be a Jacobian matrix $\partial x^{\bar{\alpha}}/\partial x^\beta$) will be written with one index up and one index down.

A. Einstein summation convention

It is cumbersome to write explicitly the sums over indices, especially later when we do tensors. Therefore, we introduce the *Einstein summation convention* (EΣC): if an index is repeated and appears once up and once down, then unless otherwise indicated this implies it is summed over all of its legal values. Thus instead of Eq. (3), we may write

$$x^{\bar{\alpha}} = \Lambda^{\bar{\alpha}}{}_{\beta} x^\beta, \tag{4}$$

with the sum over β implied. Note that Eq. (4) is actually 4 equations, one for each value of $\bar{\alpha}$; in EΣC notation it is presumed that all 4 of these equations are valid.

Because β is a dummy index in Eq. (4), we can replace it with another value, just as $\int_0^1 f(\nu)\,\mathrm{d}\nu$ could also be written $\int_0^1 f(\psi)\,\mathrm{d}\psi$:

$$x^{\bar{\alpha}} = \Lambda^{\bar{\alpha}}{}_{\epsilon} x^\epsilon. \tag{5}$$

We can also replace a non-summed index such as $\bar{\alpha}$, which indicates any of 0, 1, 2, or 3, with any other index that indicates 0, 1, 2, or 3 – so long as we do so on both sides:

$$x^{\bar{\delta}} = \Lambda^{\bar{\delta}}{}_{\beta} x^\beta. \tag{6}$$

B. Multiple transformations and inverse transformations

It is possible to have a sequence of two Lorentz transformations, $\mathcal{O} \to \bar{\mathcal{O}} \to \mathcal{O}'$. In 4×4 matrix notation, we would write

$$\boldsymbol{x}' = \Lambda_{\bar{\mathcal{O}} \to \mathcal{O}'} \Lambda_{\mathcal{O} \to \bar{\mathcal{O}}} \boldsymbol{x}. \tag{7}$$

In index notation, we write

$$x^{\alpha'} = [\Lambda_{\bar{\mathcal{O}} \to \mathcal{O}'}]^{\alpha'}{}_{\bar{\beta}} x^{\bar{\beta}} = [\Lambda_{\bar{\mathcal{O}} \to \mathcal{O}'}]^{\alpha'}{}_{\bar{\beta}} [\Lambda_{\mathcal{O} \to \bar{\mathcal{O}}}]^{\bar{\beta}}{}_{\gamma} x^{\gamma}. \tag{8}$$

In EΣC, the summation over $\bar{\beta}$ actually **is** matrix multiplication:

$$[\Lambda_{\bar{\mathcal{O}} \to \mathcal{O}'} \Lambda_{\mathcal{O} \to \bar{\mathcal{O}}}]^{\alpha'}{}_{\gamma} = [\Lambda_{\bar{\mathcal{O}} \to \mathcal{O}'}]^{\alpha'}{}_{\bar{\beta}} [\Lambda_{\mathcal{O} \to \bar{\mathcal{O}}}]^{\bar{\beta}}{}_{\gamma}. \tag{9}$$

For example, the 13 entry of the product matrix is obtained by summing over 4 products constructed from row 1 of the first matrix and row 3 of the second matrix. In this sense, EΣC can be thought of as a generalization of matrix multiplication. Its advantage over matrix notation is that it will apply to higher-dimensional structures than matrices; for example, we will introduce the curvature tensor later, which is a $4 \times 4 \times 4 \times 4$ hypercube of numbers, and the index notation will help us keep track of all the different axes. It can also keep track of the order of multiplication: matrix multiplication notation tells you which rows to multiply by which columns using the order of matrices, whereas in index notation the repeated indices tell you what to multiply by what, and the factors can be written in any order you want (they are components, and therefore are numbers rather than matrices).

A special case of the above is when $\mathcal{O}' = \mathcal{O}$, i.e., where the second transformation is the inverse of the first. In this case, the left-hand side of Eq. (9) is the identity:

$$[\mathbb{I}]^{\alpha}{}_{\gamma} = [\boldsymbol{\Lambda}_{\bar{\mathcal{O}} \to \mathcal{O}}]^{\alpha}{}_{\bar{\beta}} [\boldsymbol{\Lambda}_{\mathcal{O} \to \bar{\mathcal{O}}}]^{\bar{\beta}}{}_{\gamma}, \tag{10}$$

and we see that $\boldsymbol{\Lambda}_{\bar{\mathcal{O}} \to \mathcal{O}}$ is the matrix inverse of $\boldsymbol{\Lambda}_{\mathcal{O} \to \bar{\mathcal{O}}}$. We can denote this matrix inverse by $[\boldsymbol{\Lambda}^{-1}_{\mathcal{O} \to \bar{\mathcal{O}}}]^{\alpha}{}_{\bar{\beta}}$. Note that inverting a matrix means that the barred and unbarred indices switch positions.

Finally, we note that the identity matrix has entries that are the Kronecker delta: $[\mathbb{I}]^{\alpha}{}_{\gamma}$ is 1 if $\alpha = \gamma$ and 0 otherwise. For this reason, we will generally just write $\delta^{\alpha}{}_{\gamma}$.

III. BASIS VECTORS

You know that in N dimensions, a vector can be written as a linear combination of a set of N *basis vectors.* In special relativity, these are:

$$\boldsymbol{e}_0 \underset{\mathcal{O}}{\longrightarrow} (1,0,0,0), \quad \boldsymbol{e}_1 \underset{\mathcal{O}}{\longrightarrow} (0,1,0,0), \quad \boldsymbol{e}_2 \underset{\mathcal{O}}{\longrightarrow} (0,0,1,0), \quad \text{and} \quad \boldsymbol{e}_3 \underset{\mathcal{O}}{\longrightarrow} (0,0,0,1). \tag{11}$$

We write the basis vectors with a down index because a general vector $\boldsymbol{v}$ can be decomposed as:

$$\boldsymbol{v} = v^0 \boldsymbol{e}_0 + v^1 \boldsymbol{e}_1 + v^2 \boldsymbol{e}_2 + v^3 \boldsymbol{e}_3 = v^{\alpha} \boldsymbol{e}_{\alpha}. \tag{12}$$

Note, however, that there is a big difference between the use of the index label "1" in v^1 and $\boldsymbol{e}_1$: the 1 in v^1 indicates a component of $\boldsymbol{v}$, and the 1 in $\boldsymbol{e}_1$ tells us which basis vector we are considering.

We may write the components of a basis vector are simply the Kronecker delta:

$$(\boldsymbol{e}_{\alpha})^{\mu} = \delta^{\mu}_{\alpha}. \tag{13}$$

The basis vectors in a different frame can be written as follows:

$$(\boldsymbol{e}_{\alpha})^{\bar{\mu}} = \Lambda^{\bar{\mu}}{}_{\nu} (\boldsymbol{e}_{\alpha})^{\nu} = \Lambda^{\bar{\mu}}{}_{\nu} \delta^{\nu}_{\alpha} = \Lambda^{\bar{\mu}}{}_{\alpha}, \tag{14}$$

and so $\mathcal{O}$'s basis vectors can be written in terms of $\bar{\mathcal{O}}$'s vectors as:

$$\boldsymbol{e}_{\alpha} = \Lambda^{\bar{\mu}}{}_{\alpha} \boldsymbol{e}_{\bar{\mu}}. \tag{15}$$

The description of $\bar{\mathcal{O}}$'s vectors in terms of $\mathcal{O}$'s vectors is the same, but with an inverse transformation:

$$\boldsymbol{e}_{\bar{\mu}} = [\boldsymbol{\Lambda}^{-1}]^{\alpha}{}_{\bar{\mu}} \boldsymbol{e}_{\alpha}. \tag{16}$$

In general in this class, the lower index will be used for indices whose basis transformation rule contains a $\boldsymbol{\Lambda}^{-1}$, and the upper indices when the basis transformation rule contains a $\boldsymbol{\Lambda}$.

IV. INNER PRODUCTS

We found earlier that there was an invariant inner product of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, given by

$$\boldsymbol{a} \cdot \boldsymbol{b} = -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3. \tag{17}$$

You will note that the inner product cannot simply be written with the EΣC without some new ingredient. This is with good reason – the EΣC and index notation is designed to describe vector spaces, and vector spaces don't necessarily have inner products (they have addition and scalar multiplication). To describe an inner product, we need to define a new object:

$$\eta_{\mu\nu} \underset{\mathcal{O}}{\longrightarrow} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{18}$$

which can be written as a 4×4 matrix. We then have

$$\boldsymbol{a} \cdot \boldsymbol{b} = \eta_{\mu\nu} a^\mu b^\nu. \tag{19}$$

The right-hand side has two repeated indices, and so it should really be interpreted as the sum of 16 terms:

$$\boldsymbol{a} \cdot \boldsymbol{b} = \eta_{\mu\nu} a^\mu b^\nu = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} a^\mu b^\nu. \tag{20}$$

Because only 4 of the components of $\eta_{\mu\nu}$ are non-zero, the sum turns out to collapse to Eq. (17).

When we transform to a different reference frame, we know that the components a^μ and b^ν have to transform. In order for $\eta_{\mu\nu} a^\mu b^\nu$ to be the same in another frame, you might wonder if the components $\eta_{\mu\nu}$ also have to transform. In principle, they do: we may write

$$\eta_{\mu\nu} a^\mu b^\nu = \eta_{\mu\nu} ([\boldsymbol{\Lambda}^{-1}]^\mu{}_{\bar{\alpha}} a^{\bar{\alpha}})([\boldsymbol{\Lambda}^{-1}]^\nu{}_{\bar{\beta}} b^{\bar{\beta}}) = ([\boldsymbol{\Lambda}^{-1}]^\mu{}_{\bar{\alpha}} [\boldsymbol{\Lambda}^{-1}]^\nu{}_{\bar{\beta}} \eta_{\mu\nu}) a^{\bar{\alpha}} b^{\bar{\beta}}. \tag{21}$$

This is equal to $\eta_{\bar{\alpha}\bar{\beta}} a^{\bar{\alpha}} b^{\bar{\beta}}$ if we impose the transformation law:

$$\eta_{\bar{\alpha}\bar{\beta}} = [\boldsymbol{\Lambda}^{-1}]^\mu{}_{\bar{\alpha}} [\boldsymbol{\Lambda}^{-1}]^\nu{}_{\bar{\beta}} \eta_{\mu\nu}. \tag{22}$$

Once again, we see that a quantity with a lower index has to have a transformation that contains $\boldsymbol{\Lambda}^{-1}$; in this case, since there are two indices, there are two factors of $\boldsymbol{\Lambda}^{-1}$. However, something very special happens for the quantity $\boldsymbol{\eta}$ in the case of Lorentz transformations. We can write Eq. (22) in matrix notation with the rearrangement:

$$\eta_{\bar{\alpha}\bar{\beta}} = [\boldsymbol{\Lambda}^{-1\,\mathrm{T}}]_{\bar{\alpha}}{}^\mu \eta_{\mu\nu} [\boldsymbol{\Lambda}^{-1}]^\nu{}_{\bar{\beta}}. \tag{23}$$

(note that the transpose superscript "T" switches the indices), and thus:

$$\bar{\boldsymbol{\eta}} = \boldsymbol{\Lambda}^{-1\,\mathrm{T}} \boldsymbol{\eta} \boldsymbol{\Lambda}^{-1}. \tag{24}$$

For a $\boldsymbol{\Lambda}$ that is specifically a Lorentz transformation matrix (and thus $\boldsymbol{\Lambda}^{-1}$ is a Lorentz transformation matrix), we saw in Lecture I that $\boldsymbol{\Lambda}^{-1\,\mathrm{T}} \boldsymbol{\eta} \boldsymbol{\Lambda}^{-1} = \boldsymbol{\eta}$. The components of $\boldsymbol{\eta}$ are therefore the same in all inertial frames.

We will call any basis in which $\boldsymbol{\eta}$ takes the form of Eq. (18) an *orthonormal basis*, because the inner product of two basis vectors is

$$\boldsymbol{e}_\rho \cdot \boldsymbol{e}_\sigma = \eta_{\rho\sigma}. \tag{25}$$

A Lorentz transformation, then, can be thought of as a transformation that changes from one orthonormal basis to another orthonormal basis.