Massachusetts Institute of Technology
Department of Physics
8.033 Fall 2021

Lecture 5
Introduction to 4-vectors

## 5.1 A mathematical interlude

We are going to have occasional lectures in this class that are more heavy on formalism, notation, and mathematics than on the physics. This is one of those lectures. The goal of Lecture 5 is to introduce you to the way in which we represent an important class of geometric objects: 4-vectors, vectors with four components that point along the 3 spatial dimensions as well as time. The reward for setting everything up in this careful way will be a representation of many quantities which we use in physics that automatically builds into it *Lorentz covariance*, meaning that quantities are defined in such a way that it is straightforward for us to transform them between reference frames.

## 5.2 More spacetime geometry: The displacement 4-vector

We begin with the spacetime diagram and events $A$ and $B$ discussed previously (using $ct$ and $ct'$ for the "time" directions so that they have the same units as the "space" directions):

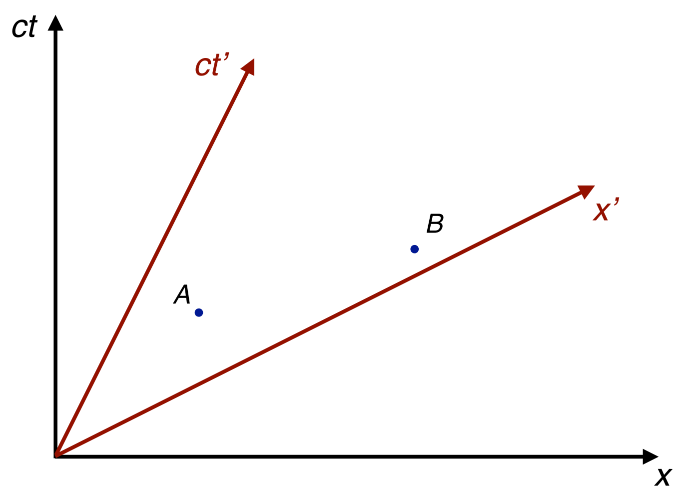
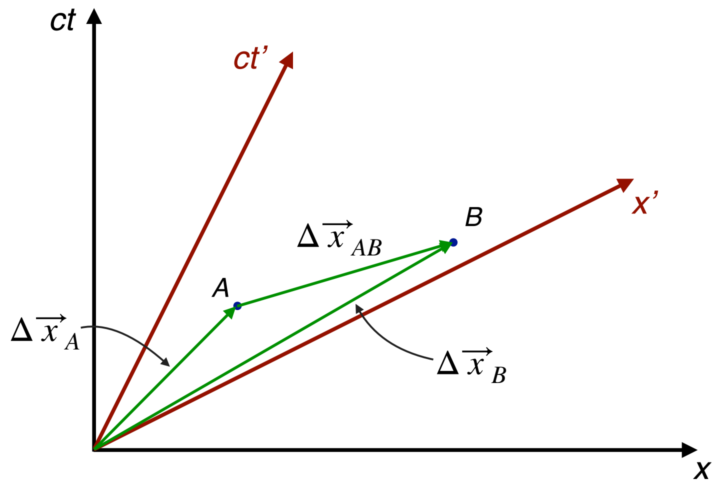



Figure 1: Two events in spacetime measured by inertial observers $\mathcal{O}$ and $\mathcal{O}'$. Observer $\mathcal{O}'$ travels with velocity $\mathbf{v} = (c/2)\mathbf{e}_x$ according to $\mathcal{O}$. We show the axes $(ct, x)$ of $\mathcal{O}$, and the axes $(ct', x')$ of $\mathcal{O}'$, both as seen by observer $\mathcal{O}$. In the coordinates of $\mathcal{O}$, the events $A$ and $B$ have spacetime coordinates $(ct_A, x_A) = (2\,\text{lightsec}, 2\,\text{lightsec})$, $(ct_B, x_B) = (3\,\text{lightsec}, 5\,\text{lightsec})$. Transforming to the frame of $\mathcal{O}'$, these become $(ct'_A, x'_A) = (2/\sqrt{3}\,\text{lightsec}, 2/\sqrt{3}\,\text{lightsec})$, $(ct'_B, x'_B) = (1/\sqrt{3}\,\text{lightsec}, 7/\sqrt{3}\,\text{lightsec})$.

Just as we can define 3-dimensional displacement vectors from one point in space to another, we can define spacetime displacement 4-vectors from one event to another in spacetime. Let us define a 4-vector $\Delta\vec{x}_A$ that points from the origin to event $A$, a 4-vector $\Delta\vec{x}_B$ that points from the origin to event $B$, and a 4-vector $\Delta\vec{x}_{AB}$ that points from event $A$ to event $B$:




These displacement 4-vectors are each geometric objects, pointing from one geometric object (an event) to another (a different event). As geometric objects, 4-vectors have frame independent properties that transcend their representation in a particular frame. For example, all observers agree that $\Delta\vec{x}_{AB}$ points from event $A$ to event $B$. However, because different observers assign different coordinates to these events, they assign different values to the *components* of this 4-vector. In what follows, we will use an overarrow to denote a 4-vector. Any quantity with such an arrow can be assumed to be a frame-independent geometric object.

Let us first examine one of these 4-vectors in the frame of $\mathcal{O}$. Focusing on $\Delta\vec{x}_A$, we write

$$\begin{aligned}\Delta\vec{x}_A &= \Delta x^0_A\vec{e}_0 + \Delta x^1_A\vec{e}_1 + \Delta x^2_A\vec{e}_2 + \Delta x^3_A\vec{e}_3\\ &= \Delta x^t_A\vec{e}_t + \Delta x^x_A\vec{e}_x + \Delta x^y_A\vec{e}_y + \Delta x^z_A\vec{e}_z\ .\end{aligned} \tag{5.1}$$

The 4-vectors $\Delta\vec{x}_B$ and $\Delta\vec{x}_{AB}$ can be written similarly. In writing this out, we have introduced 4 new geometric objects, the *unit vectors* $\vec{e}_t$, $\vec{e}_x$, etc. These are dimensionless 4-vectors that point along a particular observer's coordinate axes. We will soon see that they have magnitude 1 (though we need to do a little more setup before we can define what "magnitude" means precisely). The 4-vector $\vec{e}_0 \equiv \vec{e}_t$ points along the $t$ or $ct$ axis; $\vec{e}_1 \equiv \vec{e}_x$ is a unit vector along the $x$ axis; etc. Because these unit vectors are geometric objects, all observers agree that (for instance) $\vec{e}_t$ points along the $ct$ axis of observer $\mathcal{O}$. As we will see shortly, other observers will use different unit vectors adapted to their own coordinates.

We have also used two systems to label the components and the unit vectors. We will sometimes find it useful to label the axes with a name like $t$ or $x$; other times, it is useful to label them with a number, like 0 or 1. Both are equivalent, and both are commonly used. The convention that is now[1] most commonly used has the time direction corresponding to 0, then the spatial directions numbered 1, 2, 3 in a right-handed system.

[1]You may find older sources that use variations on this scheme; labeling the timelike direction 4 was not uncommon especially in the early days of relativity.

In Eq. (5.1), we have also introduced the vector's four components:

$$\begin{aligned}
\Delta x^0_A &= \Delta x^t_A = c\Delta t_A \longrightarrow \text{Displacement along } ct \text{ axis from origin to } A \\
&= 2\,\text{lightseconds} \qquad (5.2) \\
\Delta x^1_A &= \Delta x^x_A = \Delta x_A \longrightarrow \text{Displacement along } x \text{ axis from origin to } A \\
&= 2\,\text{lightseconds} \qquad (5.3) \\
\Delta x^2_A &= \Delta x^y_A = \Delta y_A \\
&= 0 \qquad (5.4) \\
\Delta x^3_A &= \Delta x^z_A = \Delta z_A \\
&= 0\,. \qquad (5.5)
\end{aligned}$$

Likewise, we can write down values for the components $(\Delta x^0_B,\ \Delta x^1_B,\ \Delta x^2_B,\ \Delta x^3_B)$, $(\Delta x^t_{AB},\ \Delta x^x_{AB},\ \Delta x^y_{AB},\ \Delta x^z_{AB})$, etc.

Let's look at how the observer $\mathcal{O}'$ represents $\Delta\vec{x}_A$: they write

$$\begin{aligned}
\Delta\vec{x}_A &= \Delta x^{0'}_A\vec{e}_{0'} + \Delta x^{1'}_A\vec{e}_{1'} + \Delta x^{2'}_A\vec{e}_{2'} + \Delta x^{3'}_A\vec{e}_{3'} \\
&= \Delta x^{t'}_A\vec{e}_{t'} + \Delta x^{x'}_A\vec{e}_{x'} + \Delta x^{y'}_A\vec{e}_{y'} + \Delta x^{z'}_A\vec{e}_{z'}\,. \qquad (5.6)
\end{aligned}$$

This observer uses different unit vectors — $\vec{e}_{t'}$ points along the $ct'$ axis, $\vec{e}_{x'}$ points along the $x'$ axis — and they break the displacement 4-vector into different components:

$$\begin{aligned}
\Delta x^{0'}_A &= \Delta x^{t'}_A = c\Delta t'_A \longrightarrow \text{Displacement along } ct' \text{ axis from origin to } A \\
&= 2/\sqrt{3}\,\text{lightseconds} \qquad (5.7) \\
\Delta x^{1'}_A &= \Delta x^{x'}_A = \Delta x'_A \longrightarrow \text{Displacement along } x' \text{ axis from origin to } A \\
&= 2/\sqrt{3}\,\text{lightseconds} \qquad (5.8)
\end{aligned}$$

etc. The key thing to bear in mind is that the vector $\Delta\vec{x}_A$ is *exactly the same object* in both frames. However, observers $\mathcal{O}$ and $\mathcal{O}'$ break the vector up into different components, and use a different set of unit vectors.

You hopefully are familiar with ideas like this from thinking about a 3-vector in space as represented in two different coordinate systems. One coordinate system may be oriented such that a vector $\mathbf{V} = V\,\mathbf{e}_z$; another may be oriented such that $\mathbf{V} = (V/\sqrt{3})(\mathbf{e}_{x'} + \mathbf{e}_{y'} + \mathbf{e}_{z'})$. This is simply telling us that the unprimed and primed coordinate systems differ by a rotation; $\mathbf{V}$ is the same object either way[2].

## 5.3 Einstein summation convention

In a moment, we will examine how to relate the components $(\Delta x^t_A,\ \Delta x^x_A,\ \Delta x^y_A,\ \Delta x^z_A)$ to $(\Delta x^{t'}_A,\ \Delta x^{x'}_A,\ \Delta x^{y'}_A,\ \Delta x^{z'}_A)$, and the unit vectors $(\vec{e}_t,\ \vec{e}_x,\ \vec{e}_y,\ \vec{e}_z)$ to $(\vec{e}_{t'},\ \vec{e}_{x'},\ \vec{e}_{y'},\ \vec{e}_{z'})$. Before doing so, it is worthwhile to pause in order to introduce conventions that are very useful, that are used throughout textbooks and literature on relativity, and that we will use extensively in this course.

Writing out

$$\Delta\vec{x}_A = \Delta x^0_A\vec{e}_0 + \Delta x^1_A\vec{e}_1 + \Delta x^2_A\vec{e}_2 + \Delta x^3_A\vec{e}_3 \qquad (5.9)$$

[2] This way of thinking about how the two representations are connected is often called a *passive coordinate transformation* in mathematical literature.

over and over again is cumbersome. Notice, though, that each term on the right-hand side is the same except for the index, which shifts in value. This suggests that we rewrite this using a variable for the indices:

$$\Delta\vec{x}_A = \sum_{\mu=0}^{3} \Delta x^\mu_A \vec{e}_\mu \ . \tag{5.10}$$

We can do the same thing using the components and unit vectors for observer $\mathcal{O}'$:

$$\Delta\vec{x}_A = \sum_{\mu'=0}^{3} \Delta x^{\mu'}_A \vec{e}_{\mu'} \ . \tag{5.11}$$

(It's worth noting that a very common convention, which we will use in this class, is that if the index is a Greek letter, it denotes a spacetime direction, and so ranges from 0 to 3. If the index is a Latin letter, it is a spatial direction, and ranges from 1 to 3.)

The next convention we introduce works as follows: if two symbols with the same index appear in an expression, one symbol has the index in the "upstairs" position and the other is "downstairs," then the summation can be assumed:

$$\Delta\vec{x}_A = \Delta x^\mu_A \vec{e}_\mu \ . \tag{5.12}$$

This is known as the *Einstein summation convention*; it appears to have been introduced in a 1916 paper[3] by Einstein describing the foundations of general relativity. We will use it extensively, and you will have plenty of chances to practice using it.

## 5.4 Transformations I: Displacement vector components

We now have two ways of writing a displacement vector $\Delta\vec{x}$, depending on whether we expand using the components and unit vectors of $\mathcal{O}$, or those of $\mathcal{O}'$. Putting the summation back in momentarily, we have

$$\Delta\vec{x} = \sum_{\mu=0}^{3} \Delta x^\mu \vec{e}_\mu \tag{5.13}$$

$$= \sum_{\alpha'=0}^{3} \Delta x^{\alpha'} \vec{e}_{\alpha'} \ . \tag{5.14}$$

(Comment: note that I've used changed which Greek letter I sum over in these two expressions. Because this variable gets summed over, it is called a "dummy index" — it is necessary to have *some* index in place as we expand the sum, but any letter will serve. Once the sum is performed, the index is no longer needed, and its name becomes irrelevant. We are going to start transforming quantities between reference frames, and it is a good idea to use names that do not get confused as we go between frames.)

As we have emphasized, the vector $\Delta\vec{x}$ is the same frame-independent geometric object in both of these equations. The components and unit vectors are not. How do we relate the components in one frame to the components in the other, and how do we relate the unit vectors in one frame to the unit vectors in the other?

[3] A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*, Ann. der Physik **49**, 769 (1916).

Transforming the components is easy: This is a displacement vector, and the components are simply the difference between the coordinates we use to label events in the two frames. Since these coordinates transform with the Lorentz transformation, it follows that the components of the displacement vector also transform with the Lorentz transformation:

$$\begin{pmatrix} \Delta x^{0'} \\ \Delta x^{1'} \\ \Delta x^{2'} \\ \Delta x^{3'} \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x^{0} \\ \Delta x^{1} \\ \Delta x^{2} \\ \Delta x^{3} \end{pmatrix} . \tag{5.15}$$

Writing out the matrix every time is cumbersome, so we use index notation to simplify this:

$$\begin{aligned} \Delta x^{\alpha'} &= \sum_{\mu=0}^{3} \Lambda^{\alpha'}{}_{\mu} \Delta x^{\mu} \\ &= \Lambda^{\alpha'}{}_{\mu} \Delta x^{\mu} . \end{aligned} \tag{5.16}$$

The second line of Eq. (5.16) uses the Einstein summation convention. Notice that we sum over the index $\mu$, but we do *not* sum over the index $\alpha'$. We call $\alpha'$ a *free index* — it appears on both sides of the equation, and we are free to set $\alpha'$ to any of its allowed values. Equation (5.16) is thus shorthand for 4 different equations, one for each value that $\alpha'$ is free to take.

In Eq. (5.16), we've introduced the following notation:

$$\Lambda^{\alpha'}{}_{\mu} \equiv \text{row } \alpha', \text{ column } \mu \text{ of the Lorentz transformation matrix from } \mathcal{O} \text{ to } \mathcal{O}' . \tag{5.17}$$

Notice in writing this quantity, one index connects to quantities defined in the frame of $\mathcal{O}$, the other index connects to quantities defined in the frame of $\mathcal{O}'$. The convention we use is that this quantity is an element of the Lorentz transformation matrix that takes coordinates from the frame of the *lower* index to the frame of the *upper* index. This convention implies that the elements of the inverse transformation matrix can be written by swapping the position of the indices:

$$\Lambda^{\mu}{}_{\alpha'} \equiv \text{row } \mu, \text{ column } \alpha' \text{ of the Lorentz transformation matrix from } \mathcal{O}' \text{ to } \mathcal{O} . \tag{5.18}$$

We nail this down by requiring these matrix elements to have the following property:

$$\sum_{\alpha'=0}^{3} \Lambda^{\mu}{}_{\alpha'} \Lambda^{\alpha'}{}_{\nu} = \delta^{\mu}{}_{\nu} \tag{5.19}$$

or

$$\Lambda^{\mu}{}_{\alpha'} \Lambda^{\alpha'}{}_{\nu} = \delta^{\mu}{}_{\nu} . \tag{5.20}$$

The quantity $\delta^{\mu}{}_{\nu}$ is called the Kronecker delta:

$$\delta^{\mu}{}_{\nu} = 1 \qquad \text{if } \mu = \nu \tag{5.21}$$

$$= 0 \qquad \text{otherwise} . \tag{5.22}$$

The Kronecker delta is thus an element of the identity matrix, and so Eqs. (5.19) and (5.20) do exactly what we expect if $\Lambda^{\mu}{}_{\alpha'}$ and $\Lambda^{\alpha'}{}_{\mu}$ are elements of matrices which are in inverse relation to each other.

For completeness, it should be noted that Eqs. (5.19) and (5.20) can be written in a form in which the index associated with the frame of $\mathcal{O}$ is the summed-over dummy, and the index associated with the frame of $\mathcal{O}'$ is free:

$$\sum_{\mu=0}^{3} \Lambda^{\alpha'}{}_{\mu} \Lambda^{\mu}{}_{\beta'} = \delta^{\alpha'}{}_{\beta'}$$
$$\Lambda^{\alpha'}{}_{\mu} \Lambda^{\mu}{}_{\beta'} = \delta^{\alpha'}{}_{\beta'} \ . \tag{5.23}$$

## 5.5 Transformations II: Unit vectors

With all this in mind, let's deduce what the transformation rule must be for the unit vectors. Working in index notation and putting sums explicitly in for clarity, we know that the components transform as

$$\Delta x^{\alpha'} = \sum_{\mu=0}^{3} \Lambda^{\alpha'}{}_{\mu} \Delta x^{\mu} \ . \tag{5.24}$$

Let's assume there exists *some* matrix whose elements are used to relate the unit vectors $\vec{e}_\mu$ and $\vec{e}_{\alpha'}$:

$$\vec{e}_{\alpha'} = \sum_{\mu=0}^{3} M^{\mu}{}_{\alpha'} \vec{e}_{\mu} \ . \tag{5.25}$$

We also know that the displacement 4-vector $\Delta\vec{x}$ is the same geometric object no matter how we represent it:

$$\Delta\vec{x} = \sum_{\mu=0}^{3} \Delta x^{\mu} \vec{e}_{\mu} \tag{5.26}$$
$$= \sum_{\alpha'=0}^{3} \Delta x^{\alpha'} \vec{e}_{\alpha'} \ . \tag{5.27}$$

Let us plug in the transformation rules to this final line:

$$\Delta\vec{x} = \sum_{\alpha'=0}^{3} \left( \sum_{\mu=0}^{3} \Lambda^{\alpha'}{}_{\mu} \Delta x^{\mu} \right) \left( \sum_{\nu=0}^{3} M^{\nu}{}_{\alpha'} \vec{e}_{\nu} \right) \ . \tag{5.28}$$

Notice that in writing Eq. (5.28) I was very careful to make sure that all the dummy indices we sum over are distinct from one another. A common error[4] is to get indices "crossed" and accidentally connect the wrong elements in an expression to one another.

[4]Two weeks of my life when I was a postdoctoral researcher were spent debugging a computer code in which the root issue was exactly this — bad notation that led to indices not being properly distinguished.

We next exchange the order of sums in Eq. (5.28) and reorganize the terms slightly:

$$\Delta\vec{x} = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3}\left(\sum_{\alpha'=0}^{3}\Lambda^{\alpha'}{}_{\mu}M^{\nu}{}_{\alpha'}\right)\Delta x^{\mu}\vec{e}_{\nu}\ . \tag{5.29}$$

Examining this expression, we see that it will work out perfectly *if*

$$\sum_{\alpha'=0}^{3}\Lambda^{\alpha'}{}_{\mu}M^{\nu}{}_{\alpha'} = \delta^{\nu}{}_{\mu}\ . \tag{5.30}$$

If this is the case, then

$$\begin{aligned}\Delta\vec{x} &= \sum_{\mu=0}^{3}\sum_{\nu=0}^{3}\delta^{\nu}{}_{\mu}\Delta x^{\mu}\vec{e}_{\nu} \\ &= \sum_{\mu=0}^{3}\Delta x^{\mu}\vec{e}_{\mu}\ .\end{aligned} \tag{5.31}$$

To go from the first line of (5.31) to the second, we use the fact that the Kronecker delta is zero unless $\nu = \mu$.

Thus, everything works as long as Eq. (5.30) holds. But this equation tells us that

$$M^{\nu}{}_{\alpha'} = \Lambda^{\nu}{}_{\alpha'}\ . \tag{5.32}$$

In other words, *the matrix that takes the unit vectors from the frame of $\mathcal{O}$ to the frame of $\mathcal{O}'$ is the* **inverse of the matrix** *that does this for the vector components.*

## 5.6 Summary: A glossary of transformation rules

With the results of the previous section in hand, we now have a complete set of rules describing how to Lorentz transform both vector components and unit vectors between two different inertial reference frames:

$$\Delta x^{\alpha'} = \sum_{\mu=0}^{3}\Lambda^{\alpha'}{}_{\mu}\Delta x^{\mu} = \Lambda^{\alpha'}{}_{\mu}\Delta x^{\mu}\ ; \tag{5.33}$$

$$\vec{e}_{\alpha'} = \sum_{\mu=0}^{3}\Lambda^{\mu}{}_{\alpha'}\vec{e}_{\mu} = \Lambda^{\mu}{}_{\alpha'}\vec{e}_{\mu}\ . \tag{5.34}$$

To go in the other direction (from quantities in $\mathcal{O}'$ to $\mathcal{O}$), we have

$$\Delta x^{\mu} = \sum_{\alpha'=0}^{3}\Lambda^{\mu}{}_{\alpha'}\Delta x^{\alpha'} = \Lambda^{\mu}{}_{\alpha'}\Delta x^{\alpha'}\ ; \tag{5.35}$$

$$\vec{e}_{\mu} = \sum_{\alpha'=0}^{3}\Lambda^{\alpha'}{}_{\mu}\vec{e}_{\alpha'} = \Lambda^{\alpha'}{}_{\mu}\vec{e}_{\alpha'}\ . \tag{5.36}$$

You may be somewhat in despair at this moment, worried that your 8.033 life is going to be filled with tons of algebra and fiddling with indices and matrices. There will be some of that, but I ask you to carefully study the expressions (5.33)–(5.36). Notice that the results are actually quite simple in form provided we remember some simple rules:

- Whether we are transforming the components or the unit vectors, we have one object in the frame of $\mathcal{O}$ on one side of the equation, and one in the frame of $\mathcal{O}'$ on the other.
- We connect them with an element of the "Lambda matrix" that carries out the Lorentz transformation.
- We "line up the indices" so that we have one free index on the left-hand side (corresponding to the object in the frame that we are transforming to), and we sum over dummy indices on the other side to "use up" all the indices on the object in the frame that is being transformed from.

## 5.7 Matrix multiplication versus index notation

We all hope that you soon become fluent in using the index notation, and that equations like those in the previous few sections become intuitive and simple to manipulate. However, experience has shown that many 8.033 students at least start by wanting to write out quantities as column vectors and matrices and then combine things using techniques learned in linear algebra courses.

This can be done — but it requires some care. If you want to do analyses in this way, here are a few tricks to bear in mind:

- Think of quantities with a single "upstairs" index as a column vector, e.g.

$$\Delta x^\mu \doteq \begin{pmatrix} \Delta x^0 \\ \Delta x^1 \\ \Delta x^2 \\ \Delta x^3 \end{pmatrix} . \tag{5.37}$$

- Think of quantities with a single "downstairs" index as a row vector, e.g.

$$\vec{e}_\mu \doteq \begin{pmatrix} \vec{e}_0 , & \vec{e}_1 , & \vec{e}_2 , & \vec{e}_3 \end{pmatrix} . \tag{5.38}$$

- Bearing in mind that in a quantity like $\Lambda^{\alpha'}{}_\mu$, the first index refers to the row and the second to the column, think carefully about how quantities are being combined. For example, in the relationship

$$\Delta x^{\alpha'} = \Lambda^{\alpha'}{}_\mu \Delta x^\mu , \tag{5.39}$$

  we see that we are going through the matrix elements row by row, combining the element in column $\mu$ of the matrix with row $\mu$ of the column vector $\Delta x^\mu$.

  This is just right multiplication of a column vector onto a matrix. Translating (5.39) into matrix form, we have

$$\begin{pmatrix} \Delta x^{0'} \\ \Delta x^{1'} \\ \Delta x^{2'} \\ \Delta x^{3'} \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x^0 \\ \Delta x^1 \\ \Delta x^2 \\ \Delta x^3 \end{pmatrix} = \begin{pmatrix} \gamma\,(\Delta x^0 - \beta\Delta x^1) \\ \gamma\,(-\beta\Delta x^0 + \Delta x^1) \\ \Delta x^2 \\ \Delta x^3 \end{pmatrix} . \tag{5.40}$$

Consider next the relationship

$$\vec{e}_{\alpha'} = \Lambda^{\mu}{}_{\alpha'}\vec{e}_{\mu} \ . \tag{5.41}$$

In this case we are going through the matrix column by column, combining the element in row $\mu$ with the $\mu$th column of the row vector $\vec{e}_{\mu}$. This translates into linear algebra as left multiplication of the row vector onto the matrix:

$$\begin{aligned}\left(\vec{e}_{0'}\ ,\quad \vec{e}_{1'}\ ,\quad \vec{e}_{2'}\ ,\quad \vec{e}_{3'}\right) &= \left(\vec{e}_{0}\ ,\quad \vec{e}_{1}\ ,\quad \vec{e}_{2}\ ,\quad \vec{e}_{3}\right)\begin{pmatrix}\gamma & \gamma\beta & 0 & 0\\ \gamma\beta & \gamma & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\\ &= \left(\gamma\left(\vec{e}_0 + \beta\vec{e}_1\right)\ ,\quad \gamma\left(\beta\vec{e}_0 + \vec{e}_1\right)\ ,\quad \vec{e}_2\ ,\quad \vec{e}_3\ .\right)\end{aligned} \tag{5.42}$$

Many, many mistakes that 8.033 students make when approaching problems using linear algebra arise because they translate quantities to column vectors that should be row vectors, and do not recognize that an analysis requires "left multiplication" rather than "right multiplication."

To illustrate how the index notation cleans things up, let us step through both of these calculations without introducing matrices. We just use the fact that the only non-zero elements of the $\Lambda$ matrices are $\Lambda^{0'}{}_0 = \Lambda^{1'}{}_1 = \gamma$, $\Lambda^{1'}{}_0 = \Lambda^{0'}{}_1 = -\gamma\beta$, $\Lambda^{2'}{}_2 = \Lambda^{3'}{}_3 = 1$; and $\Lambda^{0}{}_{0'} = \Lambda^{1}{}_{1'} = \gamma$, $\Lambda^{1}{}_{0'} = \Lambda^{0}{}_{1'} = \gamma\beta$, $\Lambda^{2}{}_{2'} = \Lambda^{3}{}_{3'} = 1$:

$$\Delta x^{\alpha'} = \Lambda^{\alpha'}{}_{\mu}\Delta x^{\mu} \longrightarrow \qquad \begin{aligned}\Delta x^{0'} &= \Lambda^{0'}{}_0\Delta x^0 + \Lambda^{0'}{}_1\Delta x^1 = \gamma\left(\Delta x^0 - \beta\Delta x^1\right)\\ \Delta x^{1'} &= \Lambda^{1'}{}_0\Delta x^0 + \Lambda^{1'}{}_1\Delta x^1 = \gamma\left(-\beta\Delta x^0 + \Delta x^1\right)\\ \Delta x^{2'} &= \Lambda^{2'}{}_2\Delta x^2 = \Delta x^2\\ \Delta x^{3'} &= \Lambda^{3'}{}_3\Delta x^3 = \Delta x^3\ .\end{aligned} \tag{5.43}$$

$$\vec{e}_{\alpha'} = \Lambda^{\mu}{}_{\alpha'}\vec{e}_{\mu} \longrightarrow \qquad \begin{aligned}\vec{e}_{0'} &= \Lambda^{0}{}_{0'}\vec{e}_0 + \Lambda^{1}{}_{0'}\vec{e}_1 = \gamma\left(\vec{e}_0 + \beta\vec{e}_1\right)\\ \vec{e}_{1'} &= \Lambda^{0}{}_{1'}\vec{e}_0 + \Lambda^{1}{}_{1'}\vec{e}_1 = \gamma\left(\beta\vec{e}_0 + \vec{e}_1\right)\\ \vec{e}_{2'} &= \Lambda^{2}{}_{2'}\vec{e}_2 = \vec{e}_2\\ \vec{e}_{3'} &= \Lambda^{3}{}_{3'}\vec{e}_3 = \vec{e}_3\ .\end{aligned} \tag{5.44}$$

Notice that the final results are exactly the same both ways, but the setup in index notation is simpler. (In this case, the simplicity is in part because we were able to leave out elements whose values we know to be zero.) It is worth developing "fluency" with this notation. Until you are comfortable with this form of things, linear algebra and matrix format will work. If you approach problems this way, be very careful how you translate from index format to linear algebraic equations.

## 5.8 An aside: Upstairs, downstairs; contravariant, covariant

The use of index notation abounds for representing vectorial quantities (and, more generally, *tensorial* quantities — but hold that thought until we start discussing and defining tensors very soon). We will soon encounter quite a few other geometric objects whose components have indices in the "upstairs" position, like the displacement vector components discussed above. Objects with indices up, like $\Delta x^{\mu}$, are often called *contravariant* vector components.

This name comes from the fact that the magnitude of such components "contravaries" with the scale of the reference axes to which they are attached. For instance, if we change units from meters to centimeters, decreasing the scale of our reference axes, then the numerical value of the displacement vector's components are *increased* by a factor of 100 — they *contra*vary with the scale.

There are other vector-like quantities we will define soon which are more naturally expressed with indices in the "downstairs" position. An example is the *gradient*. In electricity and magnetism, you presumably learned about electrostatic fields being the gradient of a scalar potential. In the index notation that we are beginning to use, such a relationship is most naturally expressed by writing $E_i = -\partial\phi/\partial x^i$. Objects with indices down are often called *covariant* vector components. This is because as we adjust the scale of reference axes, the magnitude of these components "covaries" with the scale. Applying the example above to this situation, if we change units from meters to centimeters, the components of a gradient all *decrease* by a factor of 100.

I mention this now because we will soon be encountering quite a few quantities with indices in both positions, and many of you are likely to encounter the terms "covariant" and "contravariant" in math classes or other physics classes. Complicating all this is that we will soon learn how to move an index from the up position to the down position and vice versa, and why this is useful and important for certain problems. Lowering the index of a vector produces what is known as its *dual vector*. We will get into these details very soon. Take these paragraphs as giving you a preview, as well as a heads up about how these quantities may be discussed elsewhere.

For what it's worth, I personally tend to just say "upstairs" and "downstairs" (a habit I picked up long ago from my Ph.D. supervisor). Some of the details of what's going on with contravariant and covariant components are worth knowing about, but (especially for 8.033) we often will not need to get into these details.