Physics 110C Relativity Notes

Michael Johnson

April 11, 2008

1 Motivation

The goal of relativity is to reformulate physical principles so that they are independent of our choice of coordinates. However, it is observed that the speed of light is constant in all inertial frames, so we have to be a little careful about what we mean by the sought coordinate independence. Fully understanding the answer requires some differential geometry, but the result is that we seek physical laws which are invariant under a particular class of coordinate transformations - Lorentz transformations. To understand these, we must first introduce the notion of a 4-vector.

2 4-Vectors

Vectors are usually defined as quantities that transform in the same way as coordinates (x_1, x_2, x_3). In relativity, we make time a coordinate as well, and multiply by c in order to obtain consistency between units. Thus, an event is space-time is described by the coordinates $x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$. We also need to introduce a metric, which tells us how lengths and distances behave in our space.

The metric is just a generalization of the ordinary dot product, only it is written as a tensor. In the class, we will be primarily interested in flat spacetime (the domain of special relativity). In this case, the metric tensor is universally denoted η and has the following matrix representation:

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The last thing we need to define is our transformation group, which is the set of all Lorentz transformations. These are defined as transformations which obey $\Lambda^T \eta \Lambda = \eta$. Solutions can be classified as ordinary spatial rotations, which aren't very interesting, and boosts (we also need to include reflections for full generality).

Let's put these in a form that you've seen. Suppose you want to transform a stationary frame into one where all coordinates are identical except that the modified frame is moving at velocity $v = \beta c$

in the positive x-direction with respect to the first. The corresponding Lorentz transformation tensor is:

$$\Lambda^{\mu}_{\nu} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Where $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. This allows us to write the ordinary Lorentz transformations between coordinate systems in the following way:

$$x^{\mu'} = \Lambda^{\mu'}_{\nu} x^{\nu}$$

Here I've used the Einstein summation convention of implicitly summing over all repeated indices. This concisely tells us how position vectors change under Lorentz transformations. We extend this concept by saying that a general vector $v^{\nu}(x^{\mu})$ is any quantity that transforms as:

$$v^{\nu'}(x^{\mu'}) = \Lambda^{\nu'}_{\nu} v^{\nu}(x^{\mu})$$

Such a v (one with an upper index) is called **contravariant**. There is another type of vector that transforms in a slightly different way:

$$w_{\nu'}(x^{\mu'}) = \Lambda^{\nu}_{\nu'} w_{\nu}(x^{\mu})$$

In this case, w_{ν} is said to be a **covariant** vector.

Example: The 4-velocity of a particle is a vector, but for it to transform properly it picks up factors of the Lorentz factor γ. If $\vec{v}$ is the ordinary velocity then we have:

$$v^{\mu} = (\gamma c, \gamma\vec{v})$$

3 Tensors

A tensor is just a generalization of the familiar concepts of scalars, vectors, and matrices. A scalar is a rank-0 tensor, a vector is a rank-1 tensor, a matrix is a rank-2 tensor, and so forth. A scalar is just a quantity which is invariant under a particular type of transformation - in special relativity, it is a quantity which is invariant under Lorentz transformations. As we saw above, vectors can either be covariant or contravariant and are defined by how they transform. A rank-n tensor has n indices, each of which can either be up (contravariant) or down (covariant).

In general, a tensor of rank (or type) (p,q) is an object with p contravariantly transforming indices (superscripts) and q covariantly transforming indices (subscripts).

Example: The Electromagnetic Stress Energy Tensor $T^{\alpha\beta}$ is a tensor of type (2,0). Suppose we want to transform $T^{\alpha\beta}$ using a particular Lorentz transformation Λ^{ν}_{μ}. We must transform each index so we get the following equation:

$$T^{\alpha'\beta'} = \Lambda^{\alpha'}_{\alpha} \Lambda^{\beta'}_{\beta} T^{\alpha\beta}$$

Note: The above equation involves summation over α and β which is why they don't appear on the left. Also, all these letters are just dummy indices - the letters themselves have no importance. Finally, it is **crucial** that we keep track of whether indices are up or down, since changing them gives different equations (as we shall see later).

4 Index Gymnastics

Although it initially appears confusing, unneccessary, and cumbersome, the index notation is actually used to make things simpler. By following a set of simple rules, it is possible to do tedious calculations in a way that requires no deep understanding of the mathematics. Here are the basic ones to remember:

Index Rules

1. Repeated indices **must** come in up-down pairs and are summed.
2. The same index cannot appear more than twice in an equation (the summation is ambiguous).
3. Indices appear in the same position (up or down) on both sides of an equation.
4. Indices may be raised or lowered using the metric tensor η.

Rather than dwell on the meaning of these, let's just do a few quick examples.

Example: Let's lower the indices of the electromagnetic stress-energy tensor $T^{\alpha\beta}$. To do this, we just use the equation $T_{\alpha\beta} = \eta_{\alpha\mu}\eta_{\beta\nu}T^{\mu\nu}$. Note how this obeys all the above rules. One nice thing about index notation is that we don't have to worry about non-commutativity since everything is implicitly written as a sum of terms which do commute. So we can start by summing over the μ index which just looks like matrix multiplication between $\eta_{\alpha\mu}$ and $T^{\mu\nu}$. Recall that:

$$\eta_{\alpha\mu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad T^{\mu\nu} = \begin{pmatrix} \frac{1}{2}(\epsilon E^2 + \frac{1}{\mu_0}B^2) & S_x & S_y & S_z \\ S_x & -T_{xx} & -T_{xy} & -T_{xz} \\ S_y & -T_{yx} & -T_{yy} & -T_{yz} \\ S_z & -T_{zx} & -T_{zy} & -T_{zz} \end{pmatrix}$$

$$\Rightarrow T_\alpha{}^\nu = \eta_{\alpha\mu}T^{\mu\nu} = \begin{pmatrix} -\frac{1}{2}(\epsilon E^2 + \frac{1}{\mu_0}B^2) & -S_x & -S_y & -S_z \\ S_x & -T_{xx} & -T_{xy} & -T_{xz} \\ S_y & -T_{yx} & -T_{yy} & -T_{yz} \\ S_z & -T_{zx} & -T_{zy} & -T_{zz} \end{pmatrix}$$

Now, to lower the other index, we must multiply by $\eta_{\beta\nu}$, but our summed index is the second one on T. One way to get around this is just move our equation to make it look like matrix multiplication again:

$$T_{\alpha\beta} = \eta_{\alpha\mu}\eta_{\beta\nu}T^{\mu\nu} = \eta_{\alpha\mu}T^{\mu\nu}\eta_{\beta\nu} = \eta_{\alpha\mu}T^{\mu\nu}\eta_{\nu\beta} = T_\alpha{}^\nu\eta_{\nu\beta}$$

Here I've used the fact that η is symmetric to swap the indices. Now we just multiply matrices:

$$T_{\alpha\beta} = T_\alpha{}^\nu\eta_{\nu\beta} = \begin{pmatrix} -\frac{1}{2}(\epsilon E^2 + \frac{1}{\mu_0}B^2) & -S_x & -S_y & -S_z \\ S_x & -T_{xx} & -T_{xy} & -T_{xz} \\ S_y & -T_{yx} & -T_{yy} & -T_{yz} \\ S_z & -T_{zx} & -T_{zy} & -T_{zz} \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2}(\epsilon E^2 + \frac{1}{\mu_0}B^2) & -S_x & -S_y & -S_z \\ -S_x & -T_{xx} & -T_{xy} & -T_{xz} \\ -S_y & -T_{yx} & -T_{yy} & -T_{yz} \\ -S_z & -T_{zx} & -T_{zy} & -T_{zz} \end{pmatrix} \neq T^{\alpha\beta}$$

So you see that we must be careful with index position.

Although it is sometimes difficult to tell whether or not an object is actually a tensor, it turns out that you can make tensors out of other tensors without having to check that properties are verified. Here are some guidelines:

Tensor Operations

1. *Linearity:* Given two tensors **of the same type**, linear combinations of them form another tensor.
   Ex. $T^{\mu}_{\ \nu} = aR^{\mu}_{\ \nu} + bS^{\mu}_{\ \nu}$

2. *Tensor Product:* Tensors may be multiplied to get another tensor. In this case, the ranks add.
   Ex. $U^{\rho\ \ \mu\nu}_{\ \lambda\ \ \ \ \sigma} = R^{\rho}_{\ \lambda} T^{\mu\nu}_{\ \ \ \sigma}$

3. *Contraction:* You may contract an upper index with a lower index resulting in a tensor with lower rank.
   Ex. This is a nice way of forming Lorentz invariance scalars. The simplest example is to find the norm squared of a vector - the generalization of length known as the space-time interval:

$$|x|^2 = x_\mu x^\mu = \eta_{\mu\nu} x^\nu x^\mu = -(ct)^2 + x^2 + y^2 + z^2$$