Tensors without Tears

1 General statements

I have been asked by several students if I knew a good book to read about tensors. I am sure that they exist, but I could not immediately think of any. So instead, I will try and give a quick review about the rules for tensors.

First of all, why do we worry about tensors? In other words, why do we spend a lot of time trying to put physical quantities into tensor form. The reason is that we know how tensors transform in going from one inertial frame to another inertial frame. In fact, we can even generalize this to transforming from any frame to any other frame, but we won't worry about that here. So if we can put physical quantities into tensor form, then we have a recipe for finding these quantities in any inertial frame.

Tensors can come with two types of indices. These are "upper" indices and "lower" indices. The distinction between the two is important because the two types of indices transform differently under Lorentz transformations.

Let us start with the simplest tensor, the displacement "4-vector" Δx^μ. Since this has an upper index, we say that this is a contravariant vector. The upper index μ refers to one of 4 possible values 0, 1, 2 or 3. The 0 component is the component along the time direction. The other three components are called *spatial* components. The 4 components of Δx^μ can now be written as

$$\Delta x^\mu : \quad (\Delta x^0, \Delta x^1, \Delta x^2, \Delta x^3) = (c\Delta t, \Delta x, \Delta y, \Delta z)\,. \tag{1}$$

Notice the first coefficient when written with Δt has a factor of c so that the Δx^0 has units of length, just like the spatial components. We also sometimes write Δx^μ : $(\Delta x^0, \Delta\vec{x})$. Lower case Greek letters are used for space-time indices, while lower case Roman letters (i, j, k etc.) are used for spatial indices only. The indices are associated with a particular inertial frame $\mathbf{S}$, and the variables x^μ are known as the *coordinates* of that frame.

Now that we have a tensor, in this case a contravariant vector, let us transform it to a new frame. Suppose we start with the 4-vector A^μ in frame $\mathbf{S}$. The goal is to find the 4-vector in the new frame $\mathbf{S}'$. This is done through a Lorentz transformation. There are 6 independent transformations. Three of these are called *boosts* and the 3 independent transformations are the boosts in the three spatial directions. The boost is completely determined by the relative velocity $\vec{v}$ that $\mathbf{S}'$ is moving *with respect to* (wrt) $\mathbf{S}$. The other 3 independent transformations are *spatial rotations*. A rotation takes place in a plane, so the independent transformations are the 3 ways of choosing 2 spatial directions, (x, y), (y, z) and (z, x). Along with the plane, we should also specify the angle through which we rotate. In any case, any Lorentz transformation can be made with some combination of these types of transformations.

A Lorentz transformation can be given by a matrix $\Lambda^{\mu'}{}_{\mu}$. The index μ' refers to the index in the frame $\mathbf{S}'$ and μ refers to the index in the frame $\mathbf{S}$. You should not assume that the two indices are the same (for instance, we could have μ be 0 while μ' is $1'$). We can use this matrix to transform our contravariant 4-vector. A typical example for $\Lambda^{\mu'}{}_{\mu}$ is a boost in the x direction

$$\Lambda^{\mu'}{}_{\mu} = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{2}$$

where v is the velocity in the x direction and γ is

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} . \tag{3}$$

Another example is the rotation in the $x - y$ plane

$$\Lambda^{\mu'}{}_{\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{4}$$

A general form for a rotation is

$$\Lambda^{\mu'}{}_{\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & R_{11} & R_{12} & R_{13} \\ 0 & R_{21} & R_{22} & R_{23} \\ 0 & R_{31} & R_{32} & R_{32} \end{pmatrix} , \tag{5}$$

The index μ' refers to the row of the matrix (with $0'$ corresponding to the first row) and μ refers to the column. So for example, the component $\Lambda^{0'}{}_{1}$ refers to the entry in the first row and second column of the matrix, which is $-\frac{v}{c}\gamma$ in (2) and 0 in (5). We can now express our 4-vector in $\mathbf{S}'$ as

$$A^{\mu'} = \Lambda^{\mu'}{}_{\mu}A^{\mu} \equiv \sum_{\mu=0}^{3} \Lambda^{\mu'}{}_{\mu}A^{\mu} . \tag{6}$$

We notice that the index μ in (6) is repeated, with it appearing once with the index down and once with the index up. When you see such a repeated index, you should assume that it is summed over the 4 components of space-time. A repeated index is also called a *dummy* index. The index μ' in (6) is not repeated. Furthermore, note that it appears on both the left and the right hand side of the equation, and in both cases, it is up. An index that is not repeated is called a *free index*. The free indices need to match exactly on the left and right hand sides of the equations, otherwise the equation is nonsense. This is not the case for dummy indices. We can think of the equation in (6) as a matrix multiplying a vector. In other words, we can write the 4-vector as

$$A^{\mu} = \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix} . \tag{7}$$

Then the transformation can be written as

$$A^{\mu'} = \begin{pmatrix} A^{0'} \\ A^{1'} \\ A^{2'} \\ A^{3'} \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix} . \tag{8}$$

Summing over the repeated index μ corresponds to the usual sum that appears in matrix multiplication. So for example, $A^{1'}$ is found by taking the second row of Λ and multiplying the entry in column μ with the entry in row μ of A^μ. More explicitly, using the transformation in (2),

$$\begin{aligned} A^{1'} &= \Lambda^{1'}{}_0 A^0 + \Lambda^{1'}{}_1 A^1 + \Lambda^{1'}{}_2 A^2 + \Lambda^{1'}{}_3 A^3 \\ &= -\frac{v}{c}\gamma A^0 + \gamma A^1 + 0 \cdot A^2 + 0 \cdot A^3 = \gamma(-\frac{v}{c}A^0 + A^1) \,. \end{aligned} \tag{9}$$

Now we should also be able to make a transformation from $\mathbf{S}'$ back to $\mathbf{S}$. Since $\mathbf{S}$ is moving with velocity $-v$ with respect to $\mathbf{S}'$, we should get the correct transformation by replacing v with $-v$ in our transformation. Hence we have

$$\Lambda^\mu{}_{\mu'} = \begin{pmatrix} \gamma & +\frac{v}{c}\gamma & 0 & 0 \\ +\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{10}$$

Then we can write A^μ as

$$A^\mu = \Lambda^\mu{}_{\mu'} A^{\mu'} \,. \tag{11}$$

Now if we first transform from $\mathbf{S}$ to $\mathbf{S}'$ and then transform back again, that is the same as doing nothing. If we put together our transformations we have

$$A^\mu = \Lambda^\mu{}_{\mu'} A^{\mu'} = \Lambda^\mu{}_{\mu'} \Lambda^{\mu'}{}_\nu A^\nu \,. \tag{12}$$

The right hand side should equal the left hand side, so we see that

$$\Lambda^\mu{}_{\mu'} \Lambda^{\mu'}{}_\nu = \delta^\mu{}_\nu , \tag{13}$$

where $\delta^\mu{}_\nu$ is the identity matrix, in other words, $\delta^\mu{}_\nu A^\nu = A^\mu$. One can also check that $\Lambda^\mu{}_{\mu'}$ is the inverse matrix of $\Lambda^{\mu'}{}_\mu$. Notice that the distinction is made between the matrix and its inverse depending on whether the primed index is up or down.

The second type of tensor we can consider is a *covariant* 4-vector, B_μ. As you can see, the index is down in this case. It is important to distinguish between the two types of indices, because they transform differently. The transformation for B_μ from $\mathbf{S}$ to $\mathbf{S}'$ is given by

$$B_{\mu'} = \Lambda^\mu{}_{\mu'} B_\mu \,, \tag{14}$$

Notice that in this transformation, the inverse transformation is used to go from $\mathbf{S}$ to $\mathbf{S}'$. Likewise, the transformation from $\mathbf{S}'$ back to $\mathbf{S}$ is

$$B_\mu = \Lambda^{\mu'}{}_\mu B_{\mu'} \tag{15}$$

We can also combine 4-vectors into more general types of tensors. For example, we can create an object $T^{\mu\nu} = A^\mu C^\nu$ out of the two contravariant 4-vectors A^μ and C^ν. Each 4-vector comes with its own index, so $T^{\mu\nu}$ naturally comes with two indices. Clearly,

$$T^{\mu'\nu'} = \Lambda^{\mu'}{}_{\mu} A^\mu \Lambda^{\nu'}{}_{\nu} C^\nu = \Lambda^{\mu'}{}_{\mu} \Lambda^{\nu'}{}_{\nu} T^{\mu\nu} \,. \tag{16}$$

We can also combine a contravariant and a covariant vector to form $T^\mu{}_\nu = A^\mu B_\nu$. This then transforms as

$$T^{\mu'}{}_{\nu'} = \Lambda^{\mu'}{}_{\mu} A^\mu \Lambda^{\nu}{}_{\nu'} B_\nu = \Lambda^{\mu'}{}_{\mu} \Lambda^{\nu}{}_{\nu'} T^\mu{}_\nu \,, \tag{17}$$

where the up index transforms with the Lorentz transformation matrix and the down index with inverse matrix. We can combine even more contravariant or covariant vectors to make a tensor with even more indices. In fact we can drop the condition that the tensor we have constructed is a product of vectors. In general we will define an $\begin{pmatrix} n \\ m \end{pmatrix}$ tensor as having n up indices and m down indices. Each up index transforms with a Lorentz transformation matrix and the down indices with the inverse. So we would have for a general tensor

$$T^{\mu_1'\mu_2'\ldots\mu_n'}_{\nu_1'\nu_2'\ldots\nu_m'} = \Lambda^{\mu_1'}{}_{\mu_1} \Lambda^{\mu_2'}{}_{\mu_2} \ldots \Lambda^{\mu_n'}{}_{\mu_n} \Lambda^{\nu_1}{}_{\nu_1'} \Lambda^{\nu_2}{}_{\nu_2'} \ldots \Lambda^{\nu_m}{}_{\nu_m'} T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m} \,, \tag{18}$$

where all repeated indices are summed over. Note that these multiple copies of Lorentz transformations should *not* be thought of as multiplying many matrices together using the usual rules of matrix multiplication. In matrix multiplication, say where we multiply one 4×4 matrix M_1 with another 4×4 matrix M_2 to get a third matrix $M_3 = M_1M_2$, the way the indices would be arranged are (for example)

$$M_3{}^{\mu'}{}_{\nu'} = M_1{}^{\mu'}{}_{\nu} M_2{}^{\nu}{}_{\nu'} \,. \tag{19}$$

There is a repeated index connecting M_1 and M_2 that corresponds to taking the column from M_1 and pairing it with the row of M_2. But in (18), no Lorentz matrix has a common index with another Lorentz matrix.

However, matrix multiplication is useful for thinking about multiple Lorentz transformations. So suppose that we first transform from $\mathbf{S}$ to $\mathbf{S}'$ and then transform to $\mathbf{S}''$. We should be able to write this as a single transformation from $\mathbf{S}$ to $\mathbf{S}''$. The way this works is through matrix multiplication, namely

$$\Lambda^{\mu''}{}_{\mu'} \Lambda^{\mu'}{}_{\mu} = \Lambda^{\mu''}{}_{\mu} \,, \tag{20}$$

The index μ' corresponding to the $\mathbf{S}'$ frame is repeated and the resulting matrix is the Lorentz transformation matrix that would be used in (18) if we wanted to transform $T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m}$ to $\mathbf{S}''$.

One thing I should stress is that the Lorentz transformation matrix is *not* a two index tensor. This is because each index is associated with a different frame, while a general two index tensor would have both indices associated with the same frame.

Now suppose that we combine a contravariant vector and a covariant vector as $A^\mu B_\mu$. The index μ is repeated, so it means that we should sum over it. Setting the indices equal like this is called *contracting* an index. Let us now see what this quantity is in $\mathbf{S}'$. When we have more than one index, each one is transformed with a Lorentz transformation matrix. Given these transformation properties, we find

$$A^{\mu'} B_{\mu'} = \Lambda^{\mu'}{}_\nu A^\nu \Lambda^\mu{}_{\mu'} B_\mu = \Lambda^\mu{}_{\mu'} \Lambda^{\mu'}{}_\nu A^\nu B_\mu \,. \tag{21}$$

In this last step we just rearranged the order of the terms in the equation. These are just numbers so we are allowed to do this. Now the $\Lambda^\mu{}_{\mu'}\Lambda^{\mu'}{}_\nu$ is precisely what appears in (13), so we can replace it by $\delta^\mu{}_\nu$, so we find

$$A^{\mu'} B_{\mu'} = \delta^\mu{}_\nu A^\nu B_\mu = A^\mu B_\mu \,. \tag{22}$$

Observe that $\delta^\mu{}_\nu$ is nonzero only when μ is the same as ν, and in this case it equals 1. Hence summing over ν will end up replacing the ν in A^ν with a μ. However, since both ν and μ are dummy indices, we could have equally replaced μ with ν. So you can see that you are free to relabel the dummy indices in which they appear. In other words

$$A^\mu B_\mu = A^\nu B_\nu \,. \tag{23}$$

We can see from the equation in (22) that the quantity $A^\mu B_\mu$ does not change when you transform to another frame. A quantity that does not change under Lorentz transformations is known as a *Lorentz invariant*, or a *Lorentz scalar* or sometimes just *invariant* or *scalar*. The multiplication of two 4-vectors to make a Lorentz scalar is called a *scalar product*, or also *inner product*. This expression has no free indices, so it is like having a tensor T with no indices. According to the rules in (18), we would have the equation $T' = T$ where the left hand side refers to the tensor in $\mathbf{S}'$. In other words, it is an invariant.

An important tensor is the two index tensor $\eta_{\mu\nu}$. This is defined as

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu\nu} \,. \tag{24}$$

Under the Lorentz transformation

$$\eta_{\mu'\nu'} = \Lambda^\mu{}_{\mu'} \Lambda^\nu{}_{\nu'} \eta_{\mu\nu} = \Lambda^{T\,\mu}_{\mu'} \eta_{\mu\nu} \Lambda^\nu{}_{\nu'} \,, \tag{25}$$

where we define $\Lambda^{T\,\mu}_{\mu'} \equiv \Lambda^\mu{}_{\mu'}$. In the last step in (25) we have arranged the indices so that it has the form of multiplying 3 matrices together. By rewriting the first Lorentz transformation this way we are transposing the rows and the columns, but are otherwise doing nothing else. If $\Lambda^{\mu'}{}_\mu$ is the general rotation in (5) then the matrix equation in (25) becomes

$$\Lambda^{T\,\mu}_{\mu'} \eta_{\mu\nu} \Lambda^\nu{}_{\nu'} = \left(\Lambda^T \eta \Lambda\right)_{\mu'\nu'} = \left(\eta \Lambda^T \Lambda\right)_{\mu'\nu'} \,, \tag{26}$$

where Λ^T is the transpose of Λ. Now for a rotation, it turns out that $\Lambda^T = \Lambda^{-1}$. Therefore, we have for rotations

$$\eta_{\mu'\nu'} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu'\nu'} . \tag{27}$$

If $\Lambda^{\mu'}{}_{\mu}$ corresponds to the boost in (2), then $\Lambda^T = \Lambda$ and we find

$$\begin{aligned}
\eta_{\mu'\nu'} &= \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ +\frac{v}{c}\gamma & -\gamma & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= \begin{pmatrix} \gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 & 0 \\ 0 & -\gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu'\nu'}
\end{aligned} \tag{28}$$

We would get the same result if we were to boost in any other direction because of the rotational symmetry of $\eta_{\mu\nu}$. Therefore, $\eta_{\mu'\nu'}$ has the same form as $\eta_{\mu\nu}$. But it is important to note that $\eta_{\mu\nu}$ is *not* a Lorentz invariant. This is because its indices changed under the transformation.

We can use the η-tensor to change contravariant indices to covariant indices. So for example consider the transformation of $A^\mu \eta_{\mu\nu}$

$$\begin{aligned}
A^{\mu'}\eta_{\mu'\nu'} &= \Lambda^{\mu'}{}_{\mu} A^\mu \Lambda^{\lambda}{}_{\mu'} \Lambda^{\nu}{}_{\nu'} \eta_{\lambda\nu} = \Lambda^{\nu}{}_{\nu'} \Lambda^{\lambda}{}_{\mu'} \Lambda^{\mu'}{}_{\mu} A^\mu \eta_{\lambda\nu} \\
&= \Lambda^{\nu}{}_{\nu'} \delta^{\lambda}{}_{\mu} A^\mu \eta_{\lambda\nu} = \Lambda^{\nu}{}_{\nu'} A^\mu \eta_{\mu\nu} .
\end{aligned} \tag{29}$$

Hence $A^\mu \eta_{\mu\nu}$ transforms as if there is only the down index ν. Therefore, we can define a covariant 4-vector from a contravariant vector by $A_\nu \equiv A^\mu \eta_{\mu\nu}$. This process is known as *lowering* an index. We can also consider the inverse of $\eta_{\mu\nu}$ with two raised indices, such that $\eta^{\mu\lambda}\eta_{\lambda\nu} = \delta^{\mu}{}_{\nu}$. But we can also think of this as $\eta^{\mu}{}_{\nu}$ where the η-tensor with the down indices lowered one of the indices of $\eta^{\mu\nu}$. We can then use $\eta^{\mu\nu}$ to raise indices; $A^\mu = \eta^{\mu\nu} A_\nu$. Obviously, this is called *raising* an index. We can raise or lower several indices by using more than one

η-tensor. If we go back to our invariant $A^\mu B_\mu$, we now see that we can write this as

$$A^\mu B_\mu = A^\mu B^\nu \eta_{\mu\nu} = A_\nu B^\nu = A_\mu B^\mu \tag{30}$$

In other words, if we have a repeated index with one raised and the other lowered, we can switch the lowered and raised indices. Notice that the last step is just a relabeling of the dummy index. We could have used any greek letter we like, although not one that is already being used as a free index. Also, it is not wise to use the same dummy index within a product of tensors, say like, ${T^\mu}_\mu {R^\mu}_\mu$ since it is not clear which index is being contracted with which. In general ${T^\mu}_\mu {R^\nu}_\nu \neq {T^\mu}_\nu {R^\nu}_\mu$.

Just to be clear about which combinations of tensors are equal to each other, the following set of equalities holds

$$T^{\mu\nu} {R_\mu}^\lambda = T^{\rho\nu} {R_\rho}^\lambda = {T_\mu}^\nu R^{\mu\lambda} \,. \tag{31}$$

Notice that the free indices ν and λ stay the same.

Sometimes a tensor has some extra symmetry. For example, we can have the symmetric tensor $T^{\mu\nu} = T^{\nu\mu}$. It is easy to show that this is symmetric in all inertial frames, to wit

$$T^{\mu'\nu'} = {\Lambda^{\mu'}}_\mu {\Lambda^{\nu'}}_\nu T^{\mu\nu} = {\Lambda^{\mu'}}_\mu {\Lambda^{\nu'}}_\nu T^{\nu\mu} = {\Lambda^{\nu'}}_\nu {\Lambda^{\mu'}}_\mu T^{\nu\mu} = T^{\nu'\mu'} \,. \tag{32}$$

Likewise for an antisymmetric tensor $G^{\mu\nu} = -G^{\nu\mu}$, this too is antisymmetric in all frames:

$$A^{\mu'\nu'} = {\Lambda^{\mu'}}_\mu {\Lambda^{\nu'}}_\nu A^{\mu\nu} = -{\Lambda^{\mu'}}_\mu {\Lambda^{\nu'}}_\nu A^{\nu\mu} = -{\Lambda^{\nu'}}_\nu {\Lambda^{\mu'}}_\mu A^{\nu\mu} = -A^{\nu'\mu'} \,. \tag{33}$$

Let us now describe some particular tensors. The displacement vector Δx^μ has already been mentioned as an example of a contravariant 4-vector. From this and the η-tensor we can construct a Lorentz invariant, Δs^2,

$$\begin{aligned} \Delta s^2 &= \Delta x^\mu \Delta x^\nu \eta_{\mu\nu} = \Delta x^\mu \Delta x_\mu = \Delta x_\mu \Delta x^\mu \\ &= \Delta x^0 \Delta x^0 - \Delta x^1 \Delta x^1 - \Delta x^2 \Delta x^2 - \Delta x^3 \Delta x^3 \\ &= c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = c^2 \Delta t^2 - \Delta\vec{x} \cdot \Delta\vec{x} \\ &= c^2 \Delta\tau^2 \,. \end{aligned} \tag{34}$$

As you can see, I am trying to make the point that all of the expressions in the above equation are the same. This invariant is loosely speaking a length squared, although not exactly because of the minus signs. Instead one can think of it as the square of the *proper time* τ between two events, multiplied by a factor of c^2 to get the dimensions right.

Instead of displacements, we can also consider differentials dx^μ and construct an invariant out of these

$$ds^2 = c^2 d\tau^2 = dx^\mu dx^\nu \eta_{\mu\nu} = dx^\mu dx_\mu \tag{35}$$

The differentials dx^μ are themselves contravariant 4-vectors and so transform as

$$dx^{\mu'} = \Lambda^{\mu'}{}_{\mu} dx^\mu \tag{36}$$

But we also know from elementary calculus that when one changes variables, the differentials of the new variables (in this case $x^{\mu'}$) are related to the differentials of the old variables (x^μ) by

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

Comparing these two equations, we see that

$$\frac{\partial x^{\mu'}}{\partial x^\mu} = \Lambda^{\mu'}{}_{\mu} \tag{38}$$

It also follows that by transforming back that

$$\frac{\partial x^{\mu}}{\partial x^{\mu'}} = \Lambda^{\mu}{}_{\mu'} \, . \tag{39}$$

We can now see that there is a natural covariant vector that we can build. Consider a scalar function $f(x^\nu)$ which depends on the coordinates. Since this function is a scalar, if it is transformed to a new frame, the function becomes

$$f'(x^{\mu'}) = f(x^\mu) \tag{40}$$

Notice that $f'(x^\mu) \neq f(x^\mu)$. The function is different, but it is different so that the *new* function as a function of the *new* variables is equal to the *old* function as a function of the *old* variables. We can then consider a derivative of $f(x^\mu)$ with respect to one of the coordinates x^ν

$$\frac{\partial}{\partial x^\nu} f(x^\mu) \, . \tag{41}$$

Now change variables, in other words, compute the derivative of the new variables acting on the new function. We then find

$$\frac{\partial}{\partial x^{\nu'}} f'(x^{\mu'}) = \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial}{\partial x^\nu} f(x^\mu) = \Lambda^{\nu}{}_{\nu'} \frac{\partial}{\partial x^\nu} f(x^\mu) \, . \tag{42}$$

In other words $\frac{\partial}{\partial x^\nu} f(x^\mu)$ transforms as a covariant 4-vector. In fact, it does not matter what scalar function the derivative is acting on, it transforms as a covariant 4-vector. We can also generalize this to a derivative acting on a $\begin{pmatrix} n \\ m \end{pmatrix}$ tensor function, $\frac{\partial}{\partial x^\lambda} T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m}$ where we see that the new object is now a $\begin{pmatrix} n \\ m+1 \end{pmatrix}$ tensor with an extra down index λ. Since the derivative adds a covariant index, it is common to write this as

$$\partial_\nu \equiv \frac{\partial}{\partial x^\nu} \, . \tag{43}$$

Another common terminology is to write the derivative of a tensor as

$$T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m,\lambda} = \frac{\partial}{\partial x^\lambda} T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m} \,, \tag{44}$$

where the comma (,) indicates that this index is coming from a derivative. If you go on to study general relativity, you will see that these ideas have to be modified somewhat, but in a very interesting way.

2 Particular tensors

In this section we will construct particular tensors corresponding to physical quantities.

2.1 4-velocity

We can see from (35) that $d\tau$, the differential of the proper time, is an invariant[1]. Since dx^μ is a 4-vector, then so is

$$u^\mu \equiv \frac{dx^\mu}{d\tau} \,, \tag{45}$$

which we call the *velocity 4-vector*, or simply *4-velocity*. In the rest frame of the particle, $d\tau = dt$, while $dx^i = 0$. Hence, in this frame $u^\mu = (c, 0, 0, 0)$. We can then exploit the fact that $u^\mu u_\mu$ is an invariant, and so $u^\mu u_\mu = c^2$ in all frames. In a frame where the particle is moving with velocity $\vec{v}$, we have that $dt = \gamma(\vec{v})d\tau$ by time dilation. Therefore

$$\frac{dx^0}{d\tau} = \frac{d(ct)}{d\tau} = c\gamma \qquad \frac{d\vec{x}}{d\tau} = \gamma\frac{d\vec{x}}{dt} = \gamma\vec{v}, \tag{46}$$

and so

$$u^\mu = \big(\gamma(v)c, \gamma(v)\vec{v}\,\big) \,. \tag{47}$$

Note that since $u^\mu u_\mu > 0$, the 4-velocity is time-like. Notice further that the 4-velocity is a tangent vector for the particle's space-time trajectory (also known as a world-line). Actually, this construction depends on the existence of a rest frame. For photons, this is clearly not the case, since the speed of light has to be the same in all inertial frames. Instead, for photons (or any other massless particles), we can define the 4-velocity as

$$u^\mu = \frac{dx^\mu}{d\lambda} \,, \tag{48}$$

where λ parameterizes the photon trajectory curve. In this case, $u^\mu u_\mu = 0$, that is the 4-velocity is light-like, and is a tangent vector for the photon's world-line.

[1] Why should the differential of the proper time be an invariant? The answer is that independent of what frame we start in, we always find the proper time by going to one particular frame, the particle's rest frame.

2.2 4-acceleration

Now that we have the 4-velocity, it is not hard to see how to find the acceleration 4-vector, or more simply called the *4-acceleration*. We just take the derivative of u^μ with respect to the proper time, namely

$$a^\mu = \frac{du^\mu}{d\tau}\,. \tag{49}$$

In the momentary rest frame of the particle we have that $u^\mu = (c, 0, 0, 0)$ at a time t, where t is the time coordinate in the momentary rest-frame. After an infinitesimal time $dt = d\tau$ elapses, the new 4-velocity in this frame is $du^\mu = (c\gamma(d\vec{v}) - c, \gamma(d\vec{v})d\vec{v})$. But $\gamma(d\vec{v}) = 1 + \mathrm{O}((d\vec{v})^2)$. Thus, to lowest order we have that $a^\mu = (0, \vec{\alpha})$ in the momentary rest frame, where

$$\vec{\alpha} = \frac{d\vec{v}}{dt} \tag{50}$$

and where $\vec{\alpha}$ is the *proper acceleration*. Again, constructing the invariant, we see that $a^\mu a_\mu = -(\vec{\alpha})^2 < 0$. Hence, the 4-acceleration is space-like (again assuming there is a rest-frame). We can also see that by going to the rest-frame that $a^\mu u_\mu = 0$, hence the 4-acceleration is a normal vector to the particle world-line.

2.3 Wave 4-vector

Consider a wave-train (the waves could be water waves or sound waves or even light waves) given by an amplitude $A(\Phi(t, \vec{x}))$, where $\Phi(t, \vec{x})$ is the phase of the wave which satisfies the periodicity condition

$$A(\Phi(t, \vec{x}) + 2\pi) = A(\Phi(t, \vec{x}))\,. \tag{51}$$

Let us write the space-time coordinates in the 4-vector form, $\Phi(t, \vec{x}) = \Phi(x^\mu)$. The phase is a Lorentz scalar, in other words if we go to a new frame, the new phase is related to the old phase by

$$\Phi'(x^{\mu'}) = \Phi(x^\mu)\,. \tag{52}$$

Notice that the function Φ' is different from Φ, that is, $\Phi'(x^\mu) \neq \Phi(x^\mu)$ (observe the difference between this statement and the one in (52)).

The angular frequency ω is found from the phase by taking its time derivative, giving

$$\omega = \frac{\partial \Phi}{\partial t} = c\,\frac{\partial \Phi}{\partial x^0} \tag{53}$$

while the wave-vector $\vec{k}$ is

$$\vec{k} = -\vec{\nabla}\Phi\,. \tag{54}$$

Note that the wave-vector is related to the wavelength λ by $|\vec{k}| = 2\pi/\lambda$. We also note that the *dispersion relation* for the wave is

$$\omega = v_w|\vec{k}| \tag{55}$$

where v_w is the speed of the wave. Since Φ is a scalar, if we compare (53) and (54) with (41) we see that we can express everything in terms of a covariant 4-vector k_μ whose components are

$$k_\mu = \left(\frac{\omega}{c}, -\vec{k}\right) . \tag{56}$$

It is very important that we write the index for k_μ as down. Otherwise we would find the wrong transformations when going to a different inertial frame.

We can also see that k_μ is space-like or light-like. The scalar product of k_μ with itself gives

$$k_\mu k^\mu = k_\mu k_\nu \eta^{\mu\nu} = \frac{\omega^2}{c^2} - \vec{k}\cdot\vec{k} = \left(\frac{v_w^2}{c^2} - 1\right) \vec{k}\cdot\vec{k} \leq 0 , \tag{57}$$

since $v_w \leq c$ and $\vec{k}\cdot\vec{k}$ is positive definite. Obviously for light-waves k_μ is light-like. Recall from quantum mechanics that there is a particle wave duality, so light waves can also be thought of as a collection of photons. Under this duality, it turns out that the wave 4-vector of the light wave is parallel to the 4-velocity of the photon.

2.4 4-momentum

Recall that for nonrelativistic physics, the momentum of a particle is related to its velocity by

$$\vec{p} = m\vec{v} \tag{58}$$

where m is the particle's mass. So it is clear that we can make a *4-momentum*, which we exress as p^μ, by taking the 4-velocity and multiplying by a scalar quantity which has the units of mass,

$$p^\mu = m_0 u^\mu . \tag{59}$$

m_0 is called the *rest mass* of the particle, the mass of the particle when it is at rest. Let us suppose that the particle itself has velocity $\vec{v}$. In this case we have

$$p^\mu = (m_0\gamma(v)c, m_0\gamma(v)\vec{v}) . \tag{60}$$

If we look at the spatial components of p^μ for very small velocities, $v << c$, then we see that $p^i \approx m_0 v^i$. The spatial components of the 4-momentum reduce to the ordinary Newtonian momentum in the nonrelativistic limit.

Now how should one interpret p^0? Again let us consider the nonrelativistic limit and let us consider the Taylor expansion of $\gamma(v)$,

$$\gamma(v) = \frac{1}{\sqrt{1 - v^2/c^2}} \approx 1 + \frac{1}{2}\frac{v^2}{c^2} . \tag{61}$$

If we then consider cp^0 in this limit, then we find

$$cp^0 \approx m_0 c^2 + \frac{1}{2} m_0 v^2 . \tag{62}$$

The second term in (62) is simply the ordinary Newtonian kinetic energy of a particle with mass m_0. So p^0/c is interpreted as an energy. This means that when the particle is at rest, it has energy

$$E = m_0 c^2 \,, \tag{63}$$

which is known as its rest energy.

Sometimes one uses the *relativistic mass*, m, where

$$m = \gamma(v) m_0 \,. \tag{64}$$

In this case, the spatial components of the 4-vector have the Newtonian form $p^i = mv^i$. Note that while the velocity of a particle is limited by the speed of light, its momentum is not because the relativistic mass approaches ∞ as v approaches c.