General Relativity

Installment VII.

September 28, 2003

1 Transformation of Vectors

The displacement is the vector *par excellence*. We express it in terms of the change of coordinates (dx^i) caused by the displacement $(d\mathbf{x})$. When we transform coordinates, we transform the components of $d\mathbf{x}$. To see how this works, recall that in a transformation of coordinates the new coordinates can be expressed as functions of the old coordinates, and vice versa. So we have

$$x^{i'} = x^{i'}(x^i). \tag{1}$$

Here the notation is such that the new coordinates get the primes on their indices so that $x^{i'}$ are the new coordinates and x^i are the old ones. In the new system, the components of the displacement vector are

$$dx^{i'} = \frac{\partial x^{i'}}{\partial x^i} dx^i \ . \tag{2}$$

So the transformation matrix $\partial x^{i'}/\partial x^i$ tell us what the components of $d\mathbf{x}$ are in the new coordinates (with the implied new basis vectors).

The infinitesimal displacement vector is such an exemplary vector that we shall demand that in our space of n dimensions any set of n quantities posing as the components of a vector has to establish its credentials by transforming like the components of $d\mathbf{x}$ as seen in (2). Suppose that we have a set of n numbers a^i given at every point in space. If these are supposed to be the components of a vector field $\mathbf{a}$ in the old system, then its components in the new coordinate system had better be

$$a^{i'} = \frac{\partial x^{i'}}{\partial x^i} a^i \ . \tag{3}$$

This then is our *d*efinition of what a vector field is — a set of n functions of position that transforms as in (3) when the coordinates transform as in (1).

We shall deal only with coordinate transformations that are invertible, so we can expect that

$$dx^i = \frac{\partial x^i}{\partial x^{i'}} dx^{i'} \tag{4}$$

where $\partial x^i/\partial x^{i'}$ is the inverse of the transformation matrix. That is, as in Installment VI,

$$\frac{\partial x^j}{\partial x^{i'}}\frac{\partial x^{i'}}{\partial x^k} = \delta^j_{\cdot k} \tag{5}$$

and

$$\frac{\partial x^{j'}}{\partial x^i}\frac{\partial x^i}{\partial x^{k'}} = \delta^{j'}_{\cdot k'} \tag{6}$$

If a point P has coordinates x^i in the old system and $x^{i'}$ in the new, a function $\phi(x^i)$ is called a *scalar* (or scalar field) if, under the transformation $x^i \to x^{i'}$, ϕ becomes

$$\phi'(x^{i'}) = \phi(x^i) \tag{7}$$

Now consider the derivative of a scalar with respect to the coordinates:

$$\frac{\partial \phi'}{\partial x^{i'}} = \frac{\partial \phi}{\partial x^j}\frac{\partial x^j}{\partial x^{i'}} = \frac{\partial x^j}{\partial x^{i'}}\frac{\partial \phi}{\partial x^j}\ . \tag{8}$$

If we let $b_i = \partial\phi/\partial x^i$, we see that this quantity transforms like

$$b_{i'} = \frac{\partial x^i}{\partial x^{i'}} b_i\ . \tag{9}$$

Components that transform in this way are called covariant while those (with upper indices) that transform as in (3) are called contravariant. The names have to do with the comparison to the transformation of coordinates with the transformation matrices

$$\Lambda^{i'}_{\cdot i} = \frac{\partial x^{i'}}{\partial x^i}; \qquad\qquad \Lambda^{i}_{\cdot i'} = \frac{\partial x^{i}}{\partial x^{i'}}\ . \tag{10}$$

The two sets of components actually represent geometrically different objects but, as they both form vector spaces of dimension n, we can match them up one by one using g_{ij}. The distinction is more important in situations where there is no g_{ij} provided, but that is a hardship we shall not face. (Mathematicians do seem to enjoy roughing it in that way, though.) It is nevertheless worth keeping the distinction in mind. It is exemplified by the tangent vector to a curve — a one dimensional object — being identified with vectors and the derivative of a function — with codimension one — being identified with differential forms. The former is often represented as an arrow and the latter by a pair of parallel line segments. (Burke's 'Spacetime, Geometry and Cosmology' (Univ. Science Books, 1980) has a long, chatty and bewildering discussion of all that at what is intended to be a very elementary level. Have a look at it if you can find it.)

2 Tensors

What happens to two-indexed quantities like g_{ij} when we transform coordinates? To think about this, consider the formula for the distance between two points that are only slightly separated so that their coordinates differ by only dx^i. The distance between them, ds, is given by the formula

$$ds^2 = g_{ij} dx^i \, dx^j \tag{11}$$

where g_{ij} is specified for the space we are in.

The value of the separation between the two points, ds, should not change when we change the coordinates, so we may also write that

$$ds^2 = g_{i'j'} dx^{i'} \, dx^{j'} \tag{12}$$

where $x^{i'}$ is the new coordinate system. On equating the two expressions for the distance interval, we find

$$g_{i'j'} dx^{i'} \, dx^{j'} = g_{ij} dx^i \, dx^j = g_{ij} \left(\frac{\partial x^i}{\partial x^{i'}} dx^{i'} \right) \left(\frac{\partial x^j}{\partial x^{j'}} dx^{j'} \right) . \tag{13}$$

That is,

$$\left(g_{i'j'} - g_{ij} \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^j}{\partial x^{j'}} \right) dx^{i'} dx^{j'} = 0 \ . \tag{14}$$

Since the coordinate differentials are arbitrary, we may conclude that

$$g_{i'j'} = g_{ij} \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^j}{\partial x^{j'}} \ . \tag{15}$$

Or, in the new notation,

$$g_{i'j'} = \Lambda^{i}_{\cdot i'} \, \Lambda^{j}_{\cdot j'} \, g_{ij} \tag{16}$$

So g_{ij}, with two indices downstairs, has a transformation matrix operating on each of its indices in the way that the components of vectors have transformations that work on their single indices. If you go through the same exercise with g^{ij}, you will find a similar result but with the upper indices transforming like the index of contravariant vector components.

More generally, one says that any set of n^2 functions labeled by two indices, as T_{ij}, where n is the range of i and j, makes up the components of a covariant tensor of second rank if, under transformation of coordinates, these components transform like

$$T_{i'j'} = \Lambda^{i}_{\cdot i'} \, \Lambda^{j}_{\cdot j'} \, T_{ij} \tag{17}$$

Likewise, a set of quantities T^{ij} that transforms like

$$T^{i'j'} = \Lambda^{i'}_{\cdot i} \, \Lambda^{j'}_{\cdot j} \, T^{ij} \tag{18}$$

is a contravariant tensor of rank two. Again, we may similarly define a mixed tensor $T^i_{\cdot j}$ of rank two. What would be the appropriate transformation law for this object?

If a^i and b^j are the contravariant components of two vectors, and T_{ij} are the components of a covariant tensor of rank two, what are the natures of $T_{ij}a^i$ and of $T_{ij}a^i\,b^i$ as determined by their transformation properties?

In this language, the vector components a^i are said to constitute the components of a contravariant tensor of rank one. And, naturally, a scalar is a tensor of rank zero. Tensors of higher rank than two, contravariant, covariant and mixed may also be defined, as we shall see later. (Try writing down now what would be the sensible definition for such objects.) But just because some quantities have lots of indices, you cannot assume that they are the components of a tensor. For example, the Christoffel symbols $\Gamma^{\;i}_{jk}$ are not the components of a third rank tensor.

When Columbia physicist I.I. Rabi learned of the discovery of a new and unexpected particle, he famously said "Who ordered that?" Perhaps you are entertaining similar thoughts about tensors. However, you can certainly see the usefulness of g_{ij}. But you may still wonder whether there is a real need for a whole class of such objects. You would not wonder at this if you had studied elasticity in any depth. In fact, that is where the word tensor (as in tension) comes from. If you have a continuous medium and you cut it into two parts (in your mind) by putting a plane through it, you could ask what is the force at a point in the plane exerted by the medium on one side on the material on the other side. There you have two vectors, one representing the surface (the plane could have any orientation) and the other representing the force itself. If you wanted to codify the information contained in that result, you would do well to introduce a two-index quantity. Of course, the two-indexed quantity need not be a tensor; you could use something called a dyad, for example. But the dyad does not have the nice properties of a tensor. For example, the sum of two dyads is not necessarily a dyad, and this is no good at all. So here the subject begins. Before we are through, you will stop worrying and learn to love tensors.

3 Differentiation of Tensors

3.1 Vectors First

As with vectors, you can add tensors by simply adding the corresponding components: for example $T^{ij} + S^{ij}$ is a second rank contravariant tensor with components U^{ij} say. But you may not add components of tensors of differing types: that is, $T^{ij} + S_{ij}$, for instance, has no meaning.

You may also multiply components pretty freely. So if a^i and b_j are components of vectors, then $a^i\,b_j$ is a mixed tensor of rank two as you may verify by checking the transformation properties. Also, you may multiply components and contract indices. For example, you should verify that if T^{ij} and S^k_i are tensors of rank two, then is $T^{ij}S^k_i$ is a tensor of rank three.

All those manipulations are worth practicing and they make good sense. But life becomes a little more complicated when we ask what tensor calculus may be like. You may want to look at some tensor books to really get into this subject — the one by Synge and Schild is particularly good, I find. Here, we need only some basics of the calculus for our purposes. Let us begin with differentiation of vectors.

Suppose we have the covariant components of a vector, a_i. They transform like

$$a_i = \frac{\partial x^{j'}}{\partial x^i} a_{j'} \tag{19}$$

If we differentiate with respect to x^k, using the notation $a_{i,k} = \partial a_i/\partial x^k$, we find from (19) on using the rule for differentiation of a product that

$$a_{i,k} = \frac{\partial x^{j'}}{\partial x^i} a_{j',\ell'} \frac{\partial x^{\ell'}}{\partial x^k} + \frac{\partial^2 x^{j'}}{\partial x^i\,\partial x^k} a_{j'} \tag{20}$$

In terms of the earlier notation for transformation matrices,

$$\Lambda^{j'}_{.i} = \frac{\partial x^{j'}}{\partial x^i} \tag{21}$$

we may write this as

$$a_{i,k} = \Lambda^{j'}_{.i}\Lambda^{\ell'}_{.k} a_{j',\ell'} + \frac{\partial^2 x^{j'}}{\partial x^i\,\partial x^k} a_{j'} \tag{22}$$

Because of the second term on the right, $a_{i.j}$ does not follow the rule for the transformation of tensor components: partial differentiation of a vector field destroys the tensor character of the field. This seems to be an undesirable feature and we may ask, what went wrong? (The problem does not arise when the transformation matrices are constant in space. For then $\partial^2 x^{j'}/(\partial x^i\,\partial x^k) = (\Lambda^{j'}_i)_{,k} = 0$.)

3.2 Another Difficulty

To see further into the problems with differentiation, let us consider the components of a vector, a^i, at point P with coordinates x^j. At a nearby point Q, the coordinates are $x^j + d\,x^j$. You might be tempted to suggest that the components at Q should be $a^i + d\,a^i$, where $d\,a^i = (\partial a^i/\partial x^j)dx^j$. But this would be wrong. Though dx^j is a vector, $a^i_{,j}$ is not, so the product is not a vector. We need to ask how to move $\mathbf{a}$ from P to Q in order to see what the difference in the components between the two places really is.

In a small enough region around P, we can go into Cartesian coordinates, designated as $x^{\hat{i}}$. This is possible because we are considering spaces that are Euclidean in small enough regions and it is standard to take advantage of this feature of Riemannian geometry. (In 'General Relativity' by H. Stephani (Cambridge, 1980) there is a nice discussion of locally flat coordinates.) The components of $\mathbf{a}$ in the local Cartesian system are

$$a^{\hat{i}} = \Lambda^{\hat{i}}_{\cdot i} a^i \tag{23}$$

In this situation, we can locally slide $\mathbf{a}$ from P to Q and we may expect expect the change $\delta\mathbf{a}$ in $\mathbf{a}$ from that operation to vanish. That is, $\delta a^i = 0$ since there is no change in the Cartesian components from a slight parallel displacement of $\mathbf{a}$. But of course, there will be a change in the components with respect to general coordinates. Hence, from (23), we obtain

$$\delta a^{\hat{i}} = 0 = x^{\hat{i}}_{,i} \delta a^i + x^{\hat{i}}_{\;,i,j} a^i d\,x^j \tag{24}$$

(The corresponding change in coordinates, δx^i, is the same as $d\,x^i$.) Thus, we find that

$$x^{\hat{i}}_{\;,i} \delta a^i = -x^{\hat{i}}_{\;,i,j} a^i d\,x^j \tag{25}$$

When we multiply this by $x^k_{\;,\hat{i}}$ and recall that

$$x^k_{\;,\hat{i}} x^{\hat{i}}_{\;,i} = \frac{\partial x^k}{\partial x^{\hat{i}}} \frac{\partial x^{\hat{i}}}{\partial x_i} = \delta^k_{\cdot i} \;, \tag{26}$$

we obtain

$$\delta a^k = -\, x^k_{\;,\hat{i}} x^{\hat{i}}_{\;,i,j}\, a^i\, d\,x^j \tag{27}$$

This confirms what we have already seen (for example, in VI–(32)): that (27) can be written as

$$\delta a^k + \Gamma^{\,k}_{ij} a^i d\,x^j = 0 \tag{28}$$

where

$$\Gamma^{\,k}_{ij} = x^k_{\;,\hat{i}} x^{\hat{i}}_{\;,i,j} \tag{29}$$

Let us now, as promised earlier, see how this expression for Γ can be related to the other expression we gave for it in terms of g_{ij} and its derivatives. This calculation starts with the realization that, in local Cartesian coordinates, the metric is given by

$$g_{\hat{i}\hat{j}} = \delta_{\hat{i}\hat{j}} \tag{30}$$

When we transform to other, more general, coordinates we find the metric

$$g_{ij} = x^{\hat{i}}_{\;,i} x^{\hat{j}}_{\;,j}\, \delta_{\hat{i}\hat{j}} \tag{31}$$

When we differentiate (31) with respect to x^k, we find that

$$g_{ij,k} = \left(x^{\hat{i}}_{\;,i,k} x^{\hat{j}}_{\;,j} + x^{\hat{i}}_{\;,i} x^{\hat{j}}_{\;,j,k} \right)\, \delta_{\hat{i}\hat{j}} \tag{32}$$

We may interchange indices in this equation to obtain the additional relations

$$g_{ik,j} = \left(x^{\hat{i}}_{\ ,i,j} x^{\hat{j}}_{\ ,k} + x^{\hat{i}}_{\ ,i} x^{\hat{j}}_{\ ,k,j} \right) \delta_{\hat{i}\hat{j}} \tag{33}$$

$$g_{kj,i} = \left(x^{\hat{i}}_{\ ,k,i} x^{\hat{j}}_{\ ,j} + x^{\hat{i}}_{\ ,k} x^{\hat{j}}_{\ ,j,i} \right) \delta_{\hat{i}\hat{j}} \tag{34}$$

When we add (33) and (34) and subtract (32), allowing for the fact that the order of partial derivatives may be interchanged, we find that

$$g_{kj,i} + g_{ik,j} - g_{ij,k} = \left(x^{\hat{i}}_{\ ,i,j} x^{\hat{j}}_{\ ,k} + x^{\hat{i}}_{\ ,k} x^{\hat{j}}_{\ ,j,i} \right) \delta_{\hat{i}\hat{j}} = 2\, x^{\hat{i}}_{\ ,i,j} x^{\hat{j}}_{\ ,k}\, \delta_{\hat{i}\hat{j}} \tag{35}$$

An older notation is

$$[i\,j, k] = \frac{1}{2}(g_{kj,i} + g_{ik,j} - g_{ij,k}) \ , \tag{36}$$

and, when look back at V-(28), we see that the Christoffel symbol can be written as

$$\Gamma^{\ r}_{ij} = g^{rk}\,[i\,j, k] = \frac{1}{2} g^{rk}\,(g_{kj,i} + g_{ik,j} - g_{ij,k}) \ . \tag{37}$$

In analogy to (31), we may write

$$g^{rk} = x^{r}_{\ ,\hat{\ell}}\, x^{k}_{\ ,\hat{m}} \delta^{\hat{\ell}\hat{m}} \ . \tag{38}$$

We now multiply (**??**) by g^{rk} to find that

$$\Gamma^{\ r}_{ij} = x^{r}_{\ ,\hat{\ell}}\, x^{k}_{\ ,\hat{m}} \delta^{\hat{\ell}\hat{m}} x^{\hat{i}}_{\ ,i,j} x^{\hat{j}}_{\ ,k}\, \delta_{\hat{i}\hat{j}} = x^{r}_{\ ,\hat{\ell}}\, \delta^{\hat{j}}_{\hat{m}} \delta^{\hat{\ell}\hat{m}} x^{\hat{i}}_{\ ,i,j}\, \delta_{\hat{i}\hat{j}} = x^{r}_{\ ,\hat{\ell}}\, x^{\hat{\ell}}_{\ ,i,j} \ . \tag{39}$$

So we learn that, indeed,

$$\frac{1}{2} g^{rk}\,(g_{kj,i} + g_{ik,j} - g_{ij,k}) = x^{r}_{\ ,\hat{\ell}}\, x^{\hat{\ell}}_{\ ,i,j} \ , \tag{40}$$

as advertised. And the argument goes through even when the hatted coordinates are not Cartesian.

So here is the dilemma: neither of the expressions (28), that is,

$$\delta a^k = -\Gamma^{\ k}_{ij} a^i d\, x^j, \tag{41}$$

nor

$$d\, a^i = a^i_{\ ,j}\, dx^j \tag{42}$$

(see (22)) is a vector.

3.3 The Covariant Derivative

Although the sum of tensor and a non-tensor is a non-tensor, the difference of two non-tensors might just turn out to be a tensor if the bad parts cancel. (There is no word that I know of for quantities that are not tensors, so I have improvised.) Thus we might try looking at

$$D\,a^i = d\,a^i - \delta\,a^i = \left(a^i{}_{,j} + \Gamma^{\,i}_{jk}\,a^k\right) d\,x^j \ . \tag{43}$$

This suggests defining what is called the covariant derivative,

$$a^i{}_{;j} = a^i{}_{,j} + \Gamma^{\,i}_{jk}\,a^k \tag{44}$$

This is in fact a mixed tensor of rank two. The argument for that is not given yet so that you will have a chance to try finding it for yourself.

A similar argument will lead you to the formula for the covariant derivative of covariant components, but there is a more pleasant way to find a suitable expression. Suppose you have the scalar

$$\phi = a^i b_i \tag{45}$$

where a^i and b_i are components of vectors. We have already seen that the derivative of a scalar is a covariant derivative (or differential form) and we would like the outcome of the derivative of this scalar product to behave in a like manner. We would also like differentiation of this product to satisfy the rule for differentiating product (Leibniz' rule), so we look at

$$\phi_{,j} = a^i{}_{,j}\,b_i + a^i\,b_{i,j} \tag{46}$$

This can be written as

$$\phi_{,j} = \left(a^i{}_{,j} + \Gamma^{\,i}_{jk}\,a^k\right)\,b_i + a^i\,\left(b_{i,j} - \Gamma^{\,k}_{ji}\,b_k\right) \tag{47}$$

This suggests the definition for the covariant derivative of a covariant vector:

$$b_{i;j} = b_{i,j} - \Gamma^{\,k}_{ji}\,b_k \tag{48}$$

for which the standard mnemonic is "co below and minus." Now we have the agreeable form

$$(a^i\,b_i)_{,j} = (a^i\,b_i)_{;j} = +b_i a^i{}_{;j} + a^i b_{i\,;j}, \tag{49}$$

which is a covariant vector.

For tensors, a like procedure will work. For the example of a second rank covariant tensor, look at $(T^{ij}a_ib_j)_{,j}$ and find the appropriate expression for $T^{ij}{}_{;k}$.

Before we go on to other (related) things we might recall that the real reason for all this complication is that the basis vectors change from place to place in general spaces or in flat space with non-Cartesian coordinates. If we have a vector

$$\mathbf{a} = a^i \mathbf{e}_i \tag{50}$$

where $\mathbf{e}_i$ are basis vectors, the partial derivative of $\mathbf{a}$ may be written out as

$$\mathbf{a}_{,j} = a^i{}_{,j}\mathbf{e}_i + a^i\mathbf{e}_{i,j} \tag{51}$$

The issue is then that we need to say what $\mathbf{e}_{i,j}$ is. If we suppose that this quantity (for a given fixed i) is a linear combination of the $\mathbf{e}_i$, as in

$$\mathbf{e}_{i,j} = \Gamma^{\,k}_{ij}\,\mathbf{e}_k \tag{52}$$

The coefficients Γ have to be figured out for this purpose, but the form is right. We have

$$\mathbf{a}_{,j} = \left(a^k{}_{,j} + \Gamma^{\,k}_{ij}\,a^i\right)\mathbf{e}_i \tag{53}$$

and, if we can (later) show that this is the same Γ as before, we have that

$$\mathbf{a}_{,j} = a^i{}_{;k}\mathbf{e}_i \tag{54}$$

where the covariant differentiation is seen to allow for the real changes with position as well as the changes in components because of changes in the basis vectors.