

PARALLEL TRANSPORT OF TENSORS

Link to: [physicspages home page](#).

To leave a comment or report an error, please use the auxiliary blog.

Reference: d'Inverno, Ray, *Introducing Einstein's Relativity* (1992), Oxford Uni Press. - Section 6.4; Problem 6.7.

From basic linear algebra, we're familiar with the idea of moving a vector parallel to itself. To add two vectors **A** and **B** in 3-d Euclidean space, for example, we're usually told to move the vector **B** parallel to itself so that its tail coincides with the head of **A** and then draw the sum as the vector from the tail of **A** to the head of **B**.

Moving a vector in this way always works in 'flat' space such as 3-d space spanned by the three rectangular basis vectors because this space is uniform everywhere. To see that this isn't always the case for different types of space, consider the example of a sphere.

Here, the notion of a vector has to be defined a bit more carefully. In 3-d flat space, a vector belongs to the same space as the space used to define the points between which we can move. Said another way, the tangent space to a point in 3-d flat space is just that very 3-d flat space itself. On a sphere, though, the tangent space to a point on the sphere is a plane which intersects the sphere at that single point. The tangent space (the plane) is no longer the same space as the surface of the sphere. The vector tangent to the sphere at a point exists in a different space from the sphere itself.

To make matters worse, every tangent vector on the sphere exists in a *different* tangent space, since the plane tangent to the sphere varies as we move around the sphere. Thus if we try to move the tangent vector and, in some sense, keep it parallel to itself, we are faced with the problem of comparing vectors from different tangent spaces.

A common but illustrative example is to start with a vector tangent to the Earth (assuming the Earth is a sphere, which isn't quite right, but it will do) at the intersection of the Greenwich meridian (0° longitude) and the equator (0° latitude), with the vector pointing north. If we now walk due north along this line until we reach the north pole and keep the vector parallel to itself (in the sense of keeping it pointing due north and tangent to the sphere at each point), then when we reach the north pole, we'll have a vector that is horizontal and which points down the 180° meridian.

However, if we return to our starting point on the equator and now walk due east until we reach a longitude of 90° E, still keeping our vector pointing due north, then walk north along this meridian until we again reach the north

pole, we will now end up with a vector that is horizontal and pointing down the 90°W meridian. The angle between the two north pole vectors is thus 90° , even though we took great care to keep the vector 'constant' as we moved it through the two paths.

This example serves to show that if we are in a curved space, the notion of parallel transport of a vector (and by extension, a tensor) depends on the path we take. This is in fact a general problem and there isn't a clever way of defining some new term that eliminates this effect.

How can we describe this situation mathematically? Returning to our 3-d flat-space situation for a moment, suppose we define some curve by means of a parameter u , so that

$$x^a = x^a(u) \quad (1)$$

Now suppose we have a vector \mathbf{v} and we want to transport this vector along this curve in such a way that the vector doesn't change. Mathematically, what we're saying is that

$$\frac{dv^a}{du} = 0 \quad (2)$$

for each vector component v^a . Introducing the coordinates, we can expand this using the chain rule to get

$$\frac{dv^a}{du} = \frac{\partial v^a}{\partial x^b} \frac{dx^b}{du} = 0 \quad (3)$$

Since the tangent vector to the curve has components dx^b/du , this equation is taking the contraction of the tangent vector with the derivatives of v^a .

Now we've seen that in general, the partial derivative of a vector is not a tensor, and that we got round this problem by introducing the covariant derivative, which is a tensor. Since the tangent vector dx^b/du is a tensor we don't need to modify it, but to make the derivative of v^a along the curve a tensor, we need to generalize the ordinary derivative above to a covariant derivative. That is, we define the condition for parallel transport of a vector in a general curved coordinate system to be

$$\frac{dv^a}{du} = \frac{dx^b}{du} v^a_{;b} = 0 \quad (4)$$

Expanding the covariant derivative, we get

$$\frac{dv^a}{du} = \frac{dx^b}{du} \left(\frac{\partial v^a}{\partial x^b} + v^c \Gamma^a_{cb} \right) \quad (5)$$

We see that if the connections Γ_{cb}^a are all zero, this reduces to the flat-space case. Further, the condition for parallel transport becomes

$$\frac{dx^b}{du} \left(\frac{\partial v^a}{\partial x^b} + v^c \Gamma_{cb}^a \right) = 0 \quad (6)$$

For a general tensor, the covariant derivative is

$$T_{cd\dots;e}^{ab\dots} = \partial_e T_{cd\dots}^{ab\dots} + T_{cd\dots}^{fb\dots} \Gamma_{fe}^a + T_{cd\dots}^{af\dots} \Gamma_{fe}^b + \dots - T_{fd\dots}^{ab\dots} \Gamma_{ce}^f - T_{cf\dots}^{ab\dots} \Gamma_{de}^f - \dots \quad (7)$$

so we'd need to plug this into the equation above to get

$$\begin{aligned} \frac{dT_{cd\dots}^{ab\dots}}{du} &= \frac{dx^e}{du} T_{cd\dots;e}^{ab\dots} \\ &= \frac{dx^e}{du} \left[\partial_e T_{cd\dots}^{ab\dots} + T_{cd\dots}^{fb\dots} \Gamma_{fe}^a + T_{cd\dots}^{af\dots} \Gamma_{fe}^b + \dots - T_{fd\dots}^{ab\dots} \Gamma_{ce}^f - T_{cf\dots}^{ab\dots} \Gamma_{de}^f - \dots \right] \end{aligned} \quad (8)$$

The tangent vector can be written as

$$\frac{dx^a}{du} \equiv X^a \quad (10)$$

and we get what is called the *absolute derivative* of a tensor:

$$\frac{dT_{cd\dots}^{ab\dots}}{du} = \frac{dx^e}{du} T_{cd\dots;e}^{ab\dots} \quad (11)$$

We can generalize this idea to get the contraction of any vector X with the covariant derivative of a tensor T . The notation used for this is:

$$\nabla_X T_{cd\dots}^{ab\dots} \equiv X^e T_{cd\dots;e}^{ab\dots} \quad (12)$$

This quantity satisfies some standard properties. If we restrict ourselves to vector fields, then if we have three vector fields W , Y and Z , two scalar functions f and g , and two constants λ and μ , then we can derive some of these properties. First:

$$\nabla_W (\lambda Y + \mu Z)^a = W^b \left(\lambda \frac{\partial Y^a}{\partial x^b} + \lambda Y^c \Gamma_{cb}^a + \mu \frac{\partial Z^a}{\partial x^b} + \mu Z^c \Gamma_{cb}^a \right) \quad (13)$$

$$= \lambda W^b \left(\frac{\partial Y^a}{\partial x^b} + Y^c \Gamma_{cb}^a \right) + \mu W^b \left(\frac{\partial Z^a}{\partial x^b} + Z^c \Gamma_{cb}^a \right) \quad (14)$$

$$= \lambda \nabla_W Y^a + \mu \nabla_W Z^a \quad (15)$$

Second:

$$\nabla_{fW+gY}Z^a = (fW^b + gY^b) \left(\frac{\partial Z^a}{\partial x^b} + Z^c \Gamma_{cb}^a \right) \quad (16)$$

$$= fW^b \left(\frac{\partial Z^a}{\partial x^b} + Z^c \Gamma_{cb}^a \right) + gY^b \left(\frac{\partial Z^a}{\partial x^b} + Z^c \Gamma_{cb}^a \right) \quad (17)$$

$$= f\nabla_w Z^a + g\nabla_Y Z^a \quad (18)$$

Finally:

$$\nabla_W (fY)^a = W^b \left(\frac{\partial (fY)^a}{\partial x^b} + (fY)^c \Gamma_{cb}^a \right) \quad (19)$$

$$= W^b (Y^a \partial_b f + f \partial_b Y^a + f Y^c \Gamma_{cb}^a) \quad (20)$$

$$= Y^a W^b \partial_b f + f \nabla_W Y^a \quad (21)$$

$$= Y^a (W \cdot \nabla f) + f \nabla_W Y^a \quad (22)$$

where the ∇f is the ordinary gradient of f . (I believe the answer given in d'Inverno's problem 6.7 is wrong here.)

PINGBACKS

Pingback: [Affine geodesic and affine parameter](#)

Pingback: [Riemann tensor and covariant contraction](#)