SUMMATION CONVENTION

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 5.4 and Problems 5.3, 5.4.

We've met the summation convention briefly before, but we'll examine it again, since it is central to relativity calculations.

Many formulas involving coordinate systems require the summation over the various coordinates. For example, suppose we have a function of n independent variables x^i (where $i = 1, \cdots, n$):

$$f = f(x) \tag{1}$$

If this function defines an m-dimensional subsurface within the n-dimensional manifold, we need m parameters to describe the subsurface. That means that we can write the function in parametric form, where each of the x^i is a function of m parameters u^j (where $j = 1, \cdots, m$). That is

$$f = f(x(u)) \tag{2}$$

The derivative of this function with respect to one of the parameters can be found using the chain rule:

$$\frac{\partial f}{\partial u^a} = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} \frac{\partial x^i}{\partial u^a} \tag{3}$$

The index i is repeated in the summand, with one occurrence being 'on the top' and one 'on the bottom'. Einstein noticed that this pattern occurs regularly in relativity theory: whenever an index is repeated with one occurrence up and the other down, this index is summed over. As such, the summation sign is redundant and, since it's a pain to have to write it over and over, it can be dropped without any ambiguity. Thus the above formula can be written

$$\frac{\partial f}{\partial u^a} = \frac{\partial f}{\partial x^i} \frac{\partial x^i}{\partial u^a} \tag{4}$$

As we'll see when we study tensors properly, a tensor can have upper and lower indices (the upper ones are *contravariant* and the lower ones are *covariant*), and terms involving products of tensors use the same summation convention. Thus we have, as examples

$$A_b^a B^{bc} = \sum_{b=1}^{n} A_b^a B^{bc} = T^{ac} \tag{5}$$

$$C_l^{ijk} D_i^p E_j^{qr} = \sum_{i=1}^{n} \sum_{j=1}^{n} C_l^{ijk} D_i^p E_j^{qr} = S_l^{kpqr} \tag{6}$$

The last terms in these examples show that we can write the result of these implied sums as a single tensor that contains only the non-repeated indices, so that T^{ac} and S_l^{kpqr} are the results of the sums.

A repeated index is known as a *dummy index*, since it is just a summation label, and it can be changed to any symbol (that doesn't appear as another index) without affecting the formula. So we could write

$$A_b^a B^{bc} = A_i^a B^{ic} = A_t^a B^{tc} = \cdots \tag{7}$$

Note that we *cannot* replace a non-repeated index, since it doesn't get summed over and appears in the final result. Thus we couldn't replace a or c in the last example.

As an example of the interchangeability of indices, consider the expression

$$(Z_{abc} + Z_{cab} + Z_{bca}) X^a X^b X^c \tag{8}$$

Since all three indices are repeated, they are all dummies and can be changed. Further, the factor $X^a X^b X^c$ is symmetric with respect to the three indices, so it doesn't matter what order the Xs are written. Multiplying the terms out we get

$$(Z_{abc} + Z_{cab} + Z_{bca}) X^a X^b X^c = Z_{abc} X^a X^b X^c + Z_{cab} X^a X^b X^c + Z_{bca} X^a X^b X^c \tag{9}$$

In the second term we can relabel the indices as $a \to b$, $b \to c$ and $c \to a$, and in the third term we can relabel as $a \to c$, $b \to a$ and $c \to b$. As a result, all three terms are equal, and we get

$$(Z_{abc} + Z_{cab} + Z_{bca}) X^a X^b X^c = 3 Z_{abc} X^a X^b X^c \tag{10}$$

We've used the Kronecker delta many times in these posts. It's defined as

$$\delta_b^a = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \tag{11}$$

Many uses of the Kronecker delta are sloppy with the indices, in that they usually both appear at the bottom as in δ_{ab}. In relativity, it's important to write the indices in the correct location so that if the symbol appears in a summation, the correct summation is performed. For example

$$\delta_a^b X^a = \sum_{a=1}^{n} \delta_a^b X^a = X^b \tag{12}$$

$$\delta_a^b X_b = \sum_{b=1}^{n} \delta_a^b X_b = X_a \tag{13}$$

If an summation expression contains several deltas, only those terms where all the deltas are non-zero survive. Thus

$$\delta_a^b \delta_b^c \delta_c^d = \delta_a^d \tag{14}$$

This follows, since the indices b and c are summed, and in the first delta the only surviving term is when $a = b$, and in the last delta, we must have $c = d$. The middle delta requires $b = c$, so combining this with the other two results gives us $a = d$ as the only surviving term. Again, note that it's only the non-repeated indices that survive the summation process.

PINGBACKS

Pingback: Contravariant tensors