General tensor analysis

Reading:

Mathews & Walker, *Mathematical Methods of Physics*, ch. 16.
S. Weinberg, *Gravitation and Cosmology*, ch. 3 & 4.
Ohanian & Ruffini, *Gravitation and Spacetime*, ch. 6.
McConnell, *Applications of Tensor Analysis*, ch. XI, XII
(Don't try to read all this in one afternoon!)

We begin with some obvious definitions. First, instead of linear coordinate transformations, we look at general ones:

$$\tilde{x}^\mu = \tilde{x}^\mu(x^0, x^1, x^2, x^3) \tag{16.1}$$

We consider a point P with coordinates x, a point Q with coordinates $x + dx$. In the tilde'd coordinates, P has coordinates $\tilde{x}$, Q has $\tilde{x} + d\tilde{x}$. The model of a *contravariant* vector is the transformation law

$$d\tilde{x}^\mu = \left(\frac{\partial \tilde{x}^\mu}{\partial x^\nu}\right) dx^\nu \tag{16.2a}$$

The model of a *covariant* vector is, as before,

$$\tilde{\varphi}_{,\mu} \equiv \frac{\partial \varphi}{\partial \tilde{x}_\mu} = \frac{\partial \varphi}{\partial x^\nu}\left(\frac{\partial x^\nu}{\partial \tilde{x}^\mu}\right) = \left(\frac{\partial x^\nu}{\partial \tilde{x}^\mu}\right)\varphi_{,\nu} . \tag{16.2b}$$

An object with the transformation law

$$\tilde{A}^{\mu\nu\ldots}_{\rho\sigma\ldots} = \left(\frac{\partial \tilde{x}^\mu}{\partial x^\kappa}\right)\left(\frac{\partial \tilde{x}^\nu}{\partial x^\lambda}\right)\ldots\left(\frac{\partial x^\alpha}{\partial \tilde{x}^\rho}\right)\left(\frac{\partial x^\beta}{\partial \tilde{x}^\sigma}\right) A^{\kappa\lambda\ldots}_{\alpha\beta\ldots} \left|\frac{\partial x}{\partial \tilde{x}}\right|^w \tag{16.3}$$

is called a *relative tensor* of weight w. Note

$$d^4\tilde{x} = d^4x \left|\frac{\partial \tilde{x}}{\partial x}\right| \equiv d^4x \left|\frac{\partial x}{\partial \tilde{x}}\right|^{-1} ; \tag{16.4}$$

thus d^4x is a relative scalar of weight –1.

A relative tensor $\mathsf{T}^{\mu\nu}$ of weight +1 is called a *tensor density*:

$$\int d^4x\, \mathsf{T}^{\mu\nu} = T^{\mu\nu} \tag{16.5}$$

is thus a tensor of weight 0.

A general space suitable for physical laws must have a fundamental notion of distance between points. This idea must be invariant under coordinate transformations or it is not fundamental. Thus we suppose $\exists$ a tensor $g_{\mu\nu}$, $\ni$

$$(ds)^2 = g_{\mu\nu}\, dx^{\mu}\, dx^{\nu} \,. \tag{16.6}$$

is invariant.

The derivative of a tensor

A covariant vector obeys the transformation law

$$\widetilde{v_\mu} = \left(\frac{\partial x^{\kappa}}{\partial \widetilde{x^{\mu}}}\right) v_\kappa \,. \tag{16.7}$$

Differentiating with respect to $\widetilde{x^{\nu}}$

$$\frac{\partial}{\partial \widetilde{x^{\nu}}}\, \widetilde{v_\mu} \overset{df}{=} \widetilde{v_{\mu,\nu}} = \left(\frac{\partial^2 x^{\kappa}}{\partial \widetilde{x^{\mu}}\, \partial \widetilde{x^{\nu}}}\right) v_\kappa + \left(\frac{\partial x^{\kappa}}{\partial \widetilde{x^{\mu}}}\right)\left(\frac{\partial x^{\lambda}}{\partial \widetilde{x^{\nu}}}\right) v_{\kappa\,,\,\lambda} \,. \tag{16.8}$$

Because of the extra term $\left(\frac{\partial^2 x^{\kappa}}{\partial \widetilde{x^{\mu}}\, \partial \widetilde{x^{\nu}}}\right) v_\kappa$, we see $v_{\kappa,\lambda}$ is *not* a tensor.

This leads us to the notion of *covariant* derivative†: first,

$$A_\mu = \left(\frac{\partial \widetilde{x^{\sigma}}}{\partial x^{\mu}}\right) \widetilde{A}_\sigma \,. \tag{16.9}$$

Now defining

$$\begin{aligned} dA_\mu &= A_\mu(x+dx) - A_\mu(x) \\ &= \left(\frac{\partial \widetilde{x^{\sigma}}}{\partial x^{\mu}}\right)\!(x+dx)\; \widetilde{A}_\sigma(x+dx) - \left(\frac{\partial \widetilde{x^{\sigma}}}{\partial x^{\mu}}\right)\!(x)\; \widetilde{A}_\sigma(x) \end{aligned} \tag{16.10}$$

we see that

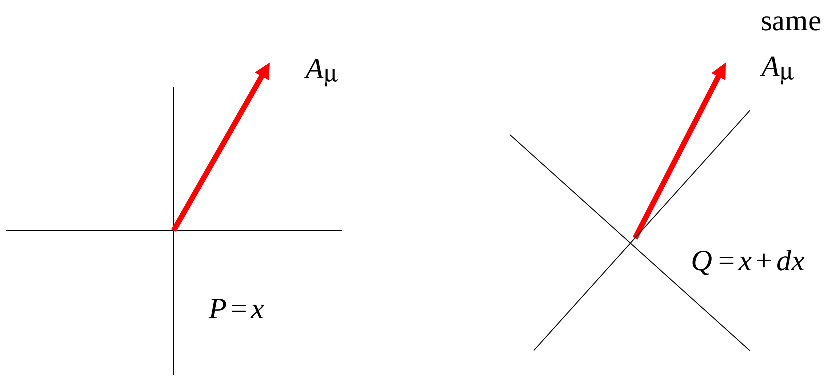
$$dA_\mu = \left(\frac{\partial \widetilde{x^{\sigma}}}{\partial x^{\mu}}\right) d\widetilde{A}_\sigma + \left(\frac{\partial^2 \widetilde{x^{\sigma}}}{\partial x^{\mu}\, \partial x^{\nu}}\right)\left(\frac{\partial x^{\nu}}{\partial \widetilde{x^{\kappa}}}\right) \widetilde{A}_\sigma\, \widetilde{dx}^{\kappa} \tag{16.11}$$

That is, as expected, dA_μ is not a vector.

This failure suggests that we try to find an object $\mathrm{D}\, A_\mu$ that *will* transform like a covariant vector. To guess what form we should try, we need to know what went wrong.

As the figure above shows, when we move from point x to point $x+dx$, the coordinate system x^{μ} might change its orientation. For example, imagine a coordinate system erected on a surface in three dimensional space---say the surface of a sphere. We have at each point two orthogonal directions,

† …in the sense of transforming like a tensor under general coordinate transformations.



$\hat{e}_\theta$ and $\hat{e}_\varphi$; but these vectors change orientation relative to a three-dimensional Cartesian system, as we move from one point to another on the sphere.

Now, suppose we have a vector that points in a constant direction in 3-space; it will have the same direction at x and at $x+dx$. But its components with respect to the local coordinate system will appear to have changed because the local system has rotated! We want a definition of a differential $\mathsf{D}\, A_\mu$ that gives zero for a vector that is constant with respect to a fixed coordinate system in which the curvilinear system is embedded.

One way to do this is to calculate the total change in A_μ going from x to $x+dx$, and subtract from it the part of the change due solely to the change in the coordinate system. We shall do just this later on.

But for now, let the x be embedded in a Cartesian space of 1 more dimension, so the coordinates are functions of $\xi^a : x^\mu = x^\mu(\xi)$. Then

$$x^\mu + dx^\mu = x^\mu(\xi) + \frac{\partial x^\mu}{\partial \xi^a} d\xi^a \tag{16.12}$$

and

$$\frac{\partial (x^\mu + dx^\mu)}{\partial x^\lambda} = \delta^\mu{}_\lambda + \frac{\partial^2 x^\mu}{\partial \xi^a \partial \xi^b} \frac{\partial \xi^a}{\partial x^\lambda} \frac{\partial \xi^b}{\partial x^\kappa} dx^\kappa \quad . \tag{16.13}$$

Now, from our general transformation law, we see that *by construction*,

$$\mathsf{D}\, A_\mu = A_\mu(x+dx) \frac{\partial (x^\mu + dx^\mu)}{\partial x^\lambda} - A_\mu(x) \tag{16.14}$$

is a vector (the reason is that we have here treated the point Q as a change of coordinates).

Hence, applying Eq. 16.13 we find

$$\mathsf{D}\, A_\mu = \left[A_{\mu,\nu} + \frac{\partial^2 x^\lambda}{\partial \xi^a \partial \xi^b} \frac{\partial \xi^a}{\partial x^\mu} \frac{\partial \xi^b}{\partial x^\nu} A_\lambda \right] dx^\nu \tag{16.15}$$

is a vector.

To proceed, we now eliminate reference to the special embedding coordinates ξ^a, in favor of the intrinsic coordinates x^μ. That is, we want to express

$$B^\lambda_{\mu\nu} \stackrel{df}{=} \frac{\partial^2 x^\lambda}{\partial\xi^a\,\partial\xi^b}\,\frac{\partial\xi^a}{\partial x^\mu}\,\frac{\partial\xi^b}{\partial x^\kappa} \tag{16.16}$$

(which we notice is homogeneous in the ξ but not in x) in terms of the intrinsic coordinates.

The only thing we have to play with is the metric tensor $g_{\mu\nu}(x)$. We have the invariant interval

$$g_{\mu\nu}(x)\,dx^\mu\,dx^\nu \equiv \eta_{ab}\,d\xi^a\,d\xi^b \tag{16.17}$$

where η_{ab} is *constant*. Note also that

$$\frac{\partial x^\lambda}{\partial\xi^a}\,\frac{\partial\xi^a}{\partial x^\mu} \equiv \delta^\lambda{}_\mu \tag{16.18}$$

so, since (obviously!)

$$\partial_\nu\,\delta^\lambda{}_\mu = 0\,,$$

we have

$$\frac{\partial^2 x^\lambda}{\partial\xi^a\,\partial\xi_b}\,\frac{\partial\xi^a}{\partial x^\mu}\,\frac{\partial\xi^b}{\partial x^\nu} + \frac{\partial x^\lambda}{\partial\xi^a}\,\frac{\partial^2\xi^a}{\partial x^\mu\,\partial x^\nu} = 0\,.$$

Thus, the vector $\mathsf{D}\,A_\mu$ may be re-expressed as

$$\mathsf{D}\,A_\mu = \left[A_{\mu,\nu} - \frac{\partial x^\lambda}{\partial\xi^a}\,\frac{\partial^2\xi^a}{\partial x^\mu\,\partial x^\nu}\,A_\lambda\right] dx^\nu\,. \tag{16.19}$$

How do we eliminate

$$\begin{Bmatrix}\lambda\\ \mu\,\nu\end{Bmatrix} \stackrel{df}{=} \frac{\partial x^\lambda}{\partial\xi^a}\,\frac{\partial^2\xi^a}{\partial x^\mu\,\partial x^\nu}\,?$$

(The $\{\}$ is called a *Christoffel symbol*.) Consider something related to $\{\}$:

$$\left[\mu\nu,\sigma\right] \stackrel{df}{=} g_{\lambda\sigma}\begin{Bmatrix}\lambda\\ \mu\,\nu\end{Bmatrix} \tag{16.20}$$

$$g_{\lambda\sigma} = \eta_{ab}\,\frac{\partial\xi^a}{\partial x^\lambda}\,\frac{\partial\xi^b}{\partial x^\sigma} \tag{16.21}$$

$$\left[\mu\nu,\sigma\right] = \eta_{ab}\,\frac{\partial^2\xi^a}{\partial x^\mu\,\partial x^\nu}\,\frac{\partial\xi^b}{\partial x^\sigma}\,. \tag{16.22}$$

Differentiating Eq. 16.21 we find

$$\partial_\sigma g_{\mu\nu} = \eta_{ab}\, \partial_\sigma \left[\frac{\partial \xi^a}{\partial x^\mu}\, \frac{\partial \xi^b}{\partial x^\nu}\right] = \eta_{ab}\left[\frac{\partial^2 \xi^a}{\partial x^\mu\, \partial x^\sigma}\, \frac{\partial \xi^b}{\partial x^\nu} + \frac{\partial \xi^a}{\partial x^\mu}\, \frac{\partial^2 \xi^b}{\partial x^\nu\, \partial x^\sigma}\right]$$

$$= \left[\mu\sigma, \nu\right] + \left[\nu\sigma, \mu\right] , \tag{16.23}$$

and by permutation of indices,

$$\partial_\nu\, g_{\mu\sigma} = \left[\mu\nu, \sigma\right] + \left[\nu\sigma, \mu\right]$$

$$\partial_\mu\, g_{\sigma\nu} = \left[\mu\sigma, \nu\right] + \left[\mu\nu, \sigma\right] .$$

We seek a linear relation that eliminates the unwanted terms and keeps the one we are looking for:

$$A\, \partial_\sigma\, g_{\mu\nu} + B\, \partial_\nu\, g_{\mu\sigma} + C\, \partial_\mu\, g_{\sigma\nu} = \left[\mu\nu, \sigma\right] ,$$

or

$$A\Big(\left[\mu\sigma, \nu\right] + \left[\nu\sigma, \mu\right]\Big) + B\Big(\left[\mu\nu, \sigma\right] + \left[\nu\sigma, \mu\right]\Big) + C\Big(\left[\mu\nu, \sigma\right] + \left[\mu\sigma, \nu\right]\Big) = \left[\mu\sigma, \nu\right].$$

Therefore,

$$A + C = 0$$
$$A + B = 0$$
$$B + C = 1 \; .$$

The solution is thus $A = -\frac{1}{2}$, $B = C = \frac{1}{2}$, or

$$\left[\mu\nu, \sigma\right] = \frac{1}{2}\Big[\, \partial_\nu\, g_{\mu\sigma} + \partial_\mu\, g_{\sigma\nu} - \partial_\sigma\, g_{\mu\nu}\Big] .$$

Thus

$$\begin{Bmatrix} \lambda \\ \mu\ \nu \end{Bmatrix} = g^{\lambda\sigma}\left[\mu\nu, \sigma\right] = \frac{1}{2} g^{\lambda\sigma}\Big[\, \partial_\nu\, g_{\mu\sigma} + \partial_\mu\, g_{\sigma\nu} - \partial_\sigma\, g_{\mu\nu}\Big] . \tag{16.24}$$

We may therefore write

$$\mathsf{D}\, A_\mu = \left[A_{\mu,\nu} - \begin{Bmatrix} \lambda \\ \mu\ \nu \end{Bmatrix} A_\lambda\right] dx^\nu . \tag{16.25}$$

Since dx^ν is clearly a tensor, the bracketed quantity must be one also. That is, the covariant derivative of a vector field is defined to be

$$A_{\mu;\,\nu} \overset{df}{=} A_{\mu,\,\nu} - \begin{Bmatrix} \lambda \\ \mu\ \nu \end{Bmatrix} A_\lambda . \tag{16.26}$$

The Christoffel symbol $\begin{Bmatrix} \lambda \\ \mu\,\nu \end{Bmatrix}$ has now been defined entirely in terms of intrinsic quantities, namely the metric tensor and its derivatives with respect to the coordinates. We see that in Cartesian (Minkowski) space, the Christoffel symbol vanishes and $A_{\mu;\,\nu} = A_{\mu,\,\nu}$.

Exercises for the bold:

1. Show that $A_{\mu;\,\nu}$ transforms properly.
2. Derive the Christoffel symbol *without* making use of an embedding space.