Lecture Eight: Theory of Relativity

8.1 Covariant Differentiation

We have seen in our derivation of the geodesic equations, that the derivative of a tensor does not, in general, transform like a tensor.

For example, consider the derivative

$$\frac{\partial A^\mu}{\partial x^\nu}$$

of a contravariant vector A^μ. Since A^μ transforms as

$$A'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} A^\nu,$$

the derivative in a different coordinate system will transform as

$$\frac{\partial A'^\mu}{\partial x'^\lambda} = \frac{\partial A^\nu}{\partial x^\kappa}\frac{\partial x^\kappa}{\partial x'^\lambda}\frac{\partial x'^\mu}{\partial x^\nu} + \frac{\partial^2 x'^\mu}{\partial x^\nu \partial x^\kappa}\frac{\partial x^\kappa}{\partial x'^\lambda} A^\nu$$

unless the transformation is *linear* then this term is non-zero.

The derivative $\partial A^\mu/\partial x^\nu$ does not therefore transform as a tensor in general. However, it is straightfoward to prove that the combination

$$\frac{\partial A^\mu}{\partial x^\lambda} + \Gamma^\mu_{\lambda\kappa} A^\kappa, \tag{1}$$

does transform like a tensor. (In Section 6 we proved a similar relation when we looked at the derivative of a vector with respect to an affine parameter varying along a curve. In (1) we are taking the derivative with respect to a coordinate, but the proof that (1) transforms as a tensor is almost identical so I won't repeat the proof here).

Now, we want to be able to write differential equations in which derivatives are *tensorial* quantities so that the equations are invariant with respect to coordinate transformations (*i.e.* the equations are generally covariant). Derivatives of tensor quantities with respect to the coordinates do not produce tensor quantities, so we need to generalise the concept of a derivative. Equation (1) gives us a generalization of the concept of a derivative that *is* covariant. Instead of derivatives with respect to coordinates, we will therefore define the *covariant derivative* of a contravariant vector (which we will denote by a semi-colon),

$$A^\mu{}_{;\lambda} = \frac{\partial A^\mu}{\partial x^\lambda} + \Gamma^\mu_{\lambda\kappa} A^\kappa.$$

This is a mixed tensor since it transforms as

$$A'^\mu{}_{;\lambda} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\rho}{\partial x'^\lambda} A^\nu_{;\rho}.$$

Likewise, from the transformation laws for a covariant tensor, the covariant derivative of a covariant vector,

$$A_{\mu;\lambda} = \frac{\partial A_\mu}{\partial x^\lambda} - \Gamma^\kappa_{\mu\lambda} A_\kappa$$

transforms like a covariant tensor

$$A'_{\mu;\lambda} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\lambda} A_{\nu;\sigma}.$$

We can generalise this to contravariant, covariant and mixed tensors of arbitrary rank, for example,

$$B^{\mu\sigma}_{\ \ \ \lambda;\rho} = \frac{\partial B^{\mu\sigma}_\lambda}{\partial x^\rho} + \Gamma^\mu_{\rho\nu} B^{\nu\sigma}_\lambda + \Gamma^\sigma_{\rho\nu} B^{\mu\nu}_\lambda - \Gamma^\kappa_{\lambda\rho} B^{\mu\sigma}_\kappa.$$

The rules are straighforward: for each contravariant (upper) index (μ, σ, ...) include an affine connection term with a plus sign, ($\Gamma^\mu_{\rho\nu}$, $\Gamma^\sigma_{\rho\nu}$, ...) times the tensor; for each covariant (lower) index (λ, σ, ...) include an affine connection term with a negative sign ($\Gamma^\nu_{\lambda\rho}$, $\Gamma^\nu_{\sigma\rho}$, ...) times the tensor,

The process of covariant differentiation satisfies certain properties which I will state without detailed proof:

(I) Additivity:

$$(\alpha A^\mu_\nu + \beta B^\mu_\nu)_{;\lambda} = \alpha A^\mu_{\nu;\lambda} + \beta B^\mu_{\nu;\lambda},$$

where α and β are constants.

(II) Leibnitz's rule:
e.g.

$$\left(A^\mu_\nu B^\lambda\right)_{;\rho} = A^\mu_{\nu;\rho} B^\lambda + A^\mu_\nu {B^\lambda}_{;\rho}.$$

(III) The covariant derivative of a contracted tensor is equal to the contraction of the covariant derivative: *e.g.*

$$B^{\mu\sigma}_{\ \ \ \lambda;\rho} = \frac{\partial B^{\mu\sigma}_{\ \ \ \lambda}}{\partial x^\rho} + \Gamma^\mu_{\rho\nu} B^{\nu\sigma}_{\ \ \ \lambda} + \Gamma^\sigma_{\rho\nu} B^{\mu\nu}_{\ \ \ \lambda} - \Gamma^\kappa_{\lambda\rho} B^{\mu\sigma}_{\ \ \ \kappa},$$

put $\sigma = \lambda$

$$\begin{aligned} B^{\mu\lambda}_{\ \ \ \lambda;\rho} &= \frac{\partial B^{\mu\lambda}_{\ \ \ \lambda}}{\partial x^\rho} + \Gamma^\mu_{\rho\nu} B^{\nu\lambda}_{\ \ \ \lambda} + \Gamma^\lambda_{\rho\nu} B^{\mu\nu}_{\ \ \ \lambda} - \Gamma^\kappa_{\lambda\rho} B^{\mu\lambda}_{\ \ \ \kappa} \end{aligned}$$

these terms cancel

$$= \frac{\partial B^{\mu\lambda}_{\ \ \ \lambda}}{\partial x^\rho} + \Gamma^\nu_{\rho\nu} B^{\nu\lambda}_{\ \ \ \lambda}.$$

(IV) The covariant derivative of the metric tensor vanishes:

$$\begin{aligned} g_{\mu\nu;\lambda} &= 0, \\ g^{\mu\nu}{}_{;\lambda} &= 0. \end{aligned}$$

Here is a sketch of the proof: the covariant derivative of $g_{\mu\nu}$ is

$$g_{\mu\nu;\lambda} = \frac{\partial g_{\mu\nu}}{\partial x^\lambda} - \Gamma^\rho_{\lambda\mu} g_{\rho\nu} - \Gamma^\rho_{\lambda\nu} g_{\rho\mu}.$$

Now remember that

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} = \Gamma^\rho_{\lambda\mu} g_{\rho\nu} + \Gamma^\rho_{\lambda\nu} g_{\rho\mu}$$

which we used in deriving

$$\Gamma^\sigma_{\lambda\mu} = \frac{1}{2} g^{\nu\sigma} \left\{ \frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\nu} \right\}.$$

Hence

$$g_{\mu\nu;\lambda} = 0.$$

(V) The operations of covariant differentiation and raising/lowering indices commute: *e.g.*

$$(g^{\mu\nu} A_\nu)_{;\lambda} = g^{\mu\nu} A_{\nu;\lambda}$$

8.2 The Curvature Tensor

A space is defined to be *flat* if the metric tensor throughout space is

$$ds^2 = \epsilon_1 (dx^1)^2 + \epsilon_2 (dx^2)^2 + \epsilon_3 (dx^3)^2 + \ldots \tag{2}$$

where

$$\epsilon_i = \pm 1.$$

If we specify the metric in an arbitrary coordinate system, how can we test whether space is flat? Can we find a tensorial quantity depending just on the $g_{\mu\nu}$ and its derivates that measures the *curvature* of space?

Start with the covariant derivative of an arbitrary vector field,

$$A_{\mu;\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \Gamma^\rho_{\mu\nu} A_\rho$$

and take the covariant derivative of this equation,

$$A_{\mu;\nu;\lambda} = \frac{\partial A_{\mu;\nu}}{\partial x^\lambda} - \Gamma^\rho_{\mu\lambda} A_{\rho;\nu} - \Gamma^\rho_{\nu\lambda} A_{\mu;\rho}. \tag{3}$$

This equation follows because $A_{\mu;\nu}$ is a second rank tensor (write $A_{\mu;\nu} \equiv B_{\mu\nu}$ then apply the rules of covariant differentiation).

Now, permute the indices of the covariant derivatives in (3) and subtract to form

$$\begin{aligned}
A_{\mu;\nu;\lambda} - A_{\mu;\lambda;\nu} &= \frac{\partial}{\partial x^\lambda}\left[\frac{\partial A_\mu}{\partial x^\nu} - \Gamma^\beta_{\mu\nu}A_\beta\right] - \Gamma^\rho_{\mu\nu}\left[\frac{\partial A_\rho}{\partial x^\nu} - \Gamma^\beta_{\rho\nu}A_\beta\right] \\
&\quad -\Gamma^\rho_{\nu\lambda}\left[\frac{\partial A_\mu}{\partial x^\rho} - \Gamma^\gamma_{\mu\rho}A_\gamma\right] - \frac{\partial}{\partial x^\nu}\left[\frac{\partial A_\mu}{\partial x^\lambda} - \Gamma^\beta_{\mu\nu}A_\beta\right] \\
&\quad +\Gamma^\rho_{\mu\nu}\left[\frac{\partial A_\rho}{\partial x^\lambda} - \Gamma^\beta_{\rho\lambda}A_\beta\right] + \Gamma^\rho_{\lambda\nu}\left[\frac{\partial A_\mu}{\partial x^\rho} - \Gamma^\gamma_{\mu\rho}A_\gamma\right],
\end{aligned}$$

Four of the terms on the rhs of this equation cancel, leaving

$$\begin{aligned}
A_{\mu;\nu;\lambda} - A_{\mu;\lambda;\nu} &= -\frac{\partial\left(\Gamma^\beta_{\mu\nu}A_\beta\right)}{\partial x^\lambda} - \Gamma^\rho_{\mu\lambda}\frac{\partial A_\rho}{\partial x^\nu} + \Gamma^\rho_{\mu\lambda}\Gamma^\beta_{\rho\nu}A_\beta \\
&\quad +\frac{\partial\left(\Gamma^\beta_{\mu\lambda}A_\beta\right)}{\partial x^\nu} + \Gamma^\rho_{\mu\nu}\frac{\partial A_\rho}{\partial x^\lambda} - \Gamma^\rho_{\mu\nu}\Gamma^\beta_{\rho\lambda}A_\beta \\
&= -\frac{\partial\Gamma^\beta_{\mu\nu}}{\partial x^\lambda}A_\beta - \Gamma^\beta_{\mu\nu}\frac{\partial A_\beta}{\partial x^\lambda} - \Gamma^\rho_{\mu\lambda}\frac{\partial A_\rho}{\partial x^\nu} + \Gamma^\rho_{\mu\lambda}\Gamma^\beta_{\rho\nu}A_\beta \\
&\quad +\frac{\partial\Gamma^\beta_{\mu\lambda}}{\partial x^\nu}A_\beta + \Gamma^\beta_{\mu\lambda}\frac{\partial A_\beta}{\partial x^\nu} + \Gamma^\rho_{\mu\nu}\frac{\partial A_\rho}{\partial x^\lambda} - \Gamma^\rho_{\mu\lambda}\Gamma^\beta_{\rho\lambda}A_\beta \\
&= \left[\frac{\partial\Gamma^\beta_{\mu\lambda}}{\partial x^\nu} - \frac{\partial\Gamma^\beta_{\mu\nu}}{\partial x^\lambda} + \Gamma^\rho_{\mu\lambda}\Gamma^\beta_{\rho\nu} - \Gamma^\rho_{\mu\nu}\Gamma^\beta_{\rho\lambda}\right]A_\beta.
\end{aligned}$$

Look at the term in square brackets. This involves just the affine connections and its derivatives, so we have created a *mixed tensor* $R^\beta_{\mu\lambda\nu}$ from the *metric tensor* $g_{\mu\nu}$ and its first and second derivatives:

$$R^\beta_{\mu\lambda\nu} = \frac{\partial\Gamma^\beta_{\mu\lambda}}{\partial x^\nu} - \frac{\partial\Gamma^\beta_{\mu\nu}}{\partial x^\lambda} + \Gamma^\rho_{\mu\lambda}\Gamma^\beta_{\rho\nu} - \Gamma^\rho_{\mu\nu}\Gamma^\beta_{\rho\lambda}. \tag{4}$$

This tensor is called the *curvature tensor*[1].

From the curvature tensor, we can construct a rank 2 tensor by contracting indices

$$R_{\mu\nu} \;=\; R^\beta_{\mu\beta\nu} \quad \leftarrow \text{this is called the Ricci tensor,}$$

and from this we can form a scalar

$$R \;=\; g^{\mu\nu}R_{\mu\nu} \quad \leftarrow \text{this is called the curvature scalar.}$$

8.3 Physical significance of curvature

Why are the tensorial quantities $R^\beta_{\mu\lambda\nu}$, $R_{\mu\nu}$ and R important? Recall our original problem – given the metric

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu,$$

[1]Note that different books use different sign conventions for the curvature tensor $R^\mu_{\alpha\beta\gamma}$. I am using the same sign convection as in Weinberg's book, which is the oppositive of that used by Rindler and d'Inverno.

we wanted to form a quantity which told us whether the $g_{\mu\nu}$ represent flat space independently of the choice of coordinates x^μ. For example, consider the following metric for a three-dimensional space

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2, \tag{5}$$

for which the metric tensor is diagonal with components

$$g_{11} = 1, \qquad g_{22} = r^2, \qquad g_{33} = r^2 \sin^2\theta.$$

Of course, we can recognise the metric (5) as the metric of ordinary three-dimensional Euclidean space written in spherical polar coordinates. In other words, the transformation

$$\left.\begin{array}{rcl} x & = & r\sin\theta\cos\phi \\ y & = & r\sin\theta\sin\phi \\ z & = & r\cos\theta \end{array}\right\} \begin{array}{l}\text{will turn the}\\ \text{metric into}\end{array}\Big\} \longrightarrow \qquad ds^2 = dx^2 + dy^2 + dz^2.$$

But, what about other metrics, *e.g.*, recall our three-space described by the metric

$$ds^2 = \frac{a^2}{(a^2 - r^2)} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2.$$

How can we tell whether this metric, or more complicated metrics, correspond to flat space but merely look complicated because of a weird choice of coordinates. It would be immensely tedious to try and prove for each metric whether there does, or does not, exist a coordinate transformation which reduces the metric to the form (2). We would like some sort of coordinate free machine that will tell us whether or not the metric is flat directly from the $g_{\mu\nu}$, *independently of the coordinate system*. This is the *mathematical* statement of the problem, and it is central to the development of General Relativity.

The *physical* significance of the problem is as follows. If we can reduce a metric

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu$$

to Minkowski form, *over all space*, then there can be no gravitational field. The equivalence of a general metric to flat space[2] therefore guarantees that the gravitational field will vanish. The solution to our mathematical problem of finding a coordinate free way of defining the curvature of space will lead us to the field equations of gravity.

8.4 Mathematical measure of curvature

We can find a solution to the problem of measuring curvature in the following simple way. If the space is flat, the *covariant derivative must commute*

$$A_{\mu;\nu;\lambda} - A_{\mu;\lambda;\nu} = 0. \tag{6}$$

This is a tensor equation and so is independent of the coordinates. It obviously applies in flat space with metric (2), since all of the affine connection terms vanish at every point in the space.

[2] The identification of flat space with Minkowski space requires further that $g_{\mu\nu}$ has three negative and one positive eigenvalue.

But the left hand side of (6) is exactly how we constructed the curvature tensor,

$$A_{\mu;\nu;\lambda} - A_{\mu;\lambda;\nu} = R^{\beta}_{\mu\lambda\nu} A_{\beta}.$$

Hence the covariant derivative will only commute at every point in the space if the curvature tensor

$$R^{\beta}_{\mu\lambda\nu} = 0$$

i.e. if the curvature tensor vanishes at every point in the space. The commutation condition applies for *any* tensor of *arbitrary* rank, *e.g.* for the mixed tensor

$$B^{\lambda}_{\mu;\nu;\kappa} - B^{\lambda}_{\mu;\kappa;\nu} = B^{\lambda}_{\mu} R^{\mu}_{\sigma\nu\kappa} - B^{\lambda}_{\sigma} R^{\sigma}_{\mu\nu\kappa}.$$

In summary, we have done some mathematics and some physics in this section:

Mathematics: If the tensor $R^{\beta}_{\mu\lambda\nu}$ vanishes everywhere, then the space must be flat. (We have created a 'machine' – the curvature tensor – that tells us about deviations from flatness whatever the choice of coordinates.)

Physics: If the tensor $R^{\beta}_{\mu\lambda\nu}$ vanishes everywhere, then the space must be flat and hence *there can be no gravitational field.*

8.5 Properties of the curvature tensor

I will end this section with a summary of some of the properties of the curvature tensor, which we will use frequently in later lectures.

The mixed form of the curvature tensor

$$R^{\lambda}_{\mu\nu\kappa}$$

has the following properties:

- Antisymmetry with respect to the indices ν and κ:

$$R^{\lambda}_{\mu\nu\kappa} = -R^{\lambda}_{\mu\kappa\nu}$$

- Cyclicity:

$$R^{\lambda}_{\mu\nu\kappa} + R^{\lambda}_{\nu\kappa\mu} + R^{\lambda}_{\kappa\mu\nu} = 0$$

The covariant form of the curvature tensor,

$$R_{\lambda\mu\nu\kappa} = g_{\lambda\sigma} R^{\sigma}_{\mu\nu\kappa}$$

satisfies:

- Symmetry with respect to swaps of the *pairs* of indices λ, μ and ν, κ:

$$R_{\underbrace{\lambda\mu}\,\underbrace{\nu\kappa}} = R_{\underbrace{\nu\kappa}\,\underbrace{\lambda\mu}}$$

- Antisymmetry:

$$R_{\lambda\mu\nu\kappa} = -R_{\underbrace{\mu\lambda}\,\nu\kappa} = -R_{\lambda\mu\,\underbrace{\kappa\nu}} = R_{\underbrace{\mu\lambda}\,\underbrace{\kappa\nu}}$$

- Cyclicity:

$$R_{\lambda\mu\nu\kappa} + R_{\lambda\kappa\mu\nu} + R_{\lambda\nu\kappa\mu} = 0$$

These conditions *significantly* reduce the number of *independent* components of $R_{\lambda\mu\kappa\nu}$. For example: with no constraints, in 4 dimensions, the tensor $R_{\lambda\mu\kappa\nu}$ would have

$$4^4 = 256$$

independent components! In fact, in N-dimensional space, the various constraints reduce the number of independent components of $R_{\lambda\mu\kappa\nu}$ to

$$\frac{1}{12}N^2(N^2-1).$$

Thus

No. of space dimension	2	3	4
No. of independent components of $R_{\lambda\mu\nu\kappa}$	1	6	20

So you can see from this table that in four dimensions, the number of independent components is reduced from a possible 256 to 20. You will also see that in one dimension the curvature tensor is always equal to zero, $R_{1111} = 0$. How can this be? Can a line not be curved? Think about this – the curvature measures the "inner" properties of the space. When we say that a line is curved, we refer to a particular embedding in a higher dimensional space, but this doesn't tell us about the inner properties of the space. In one dimension, it is evident that we can *always* find a coordinate transformation that will reduce an arbitrary metric to the form (2).

The curvature tensor also satisfies some differential relations. From the definition of the curvature tensor:

$$R_{\lambda\mu\nu\kappa} = \frac{1}{2}\left[\frac{\partial^2 g_{\lambda\nu}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^\nu \partial x^\mu} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^\nu \partial x^\lambda}\right] + g_{\eta\sigma}\left[\Gamma^\eta_{\nu\lambda}\Gamma^\sigma_{\mu\kappa} - \Gamma^\eta_{\kappa\lambda}\Gamma^\sigma_{\mu\nu}\right],$$

the covariant derivative is

$$R_{\lambda\mu\nu\kappa;\eta} = \frac{1}{2}\frac{\partial}{\partial x^\eta}\left[\frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\nu} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} - \frac{\partial^2 g_{\lambda\kappa}}{\partial x^\mu \partial x^\nu} + \frac{\partial^2 g_{\mu\kappa}}{\partial x^\nu \partial x^\lambda}\right]$$

and, by permuting the indices ν, κ and η cyclically

$$R_{\lambda\mu\nu\kappa;\eta} + R_{\lambda\mu\eta\nu;\kappa} + R_{\lambda\mu\kappa\eta;\nu} = 0. \quad \leftarrow \quad \text{These are known as the Bianchi identies.}$$

Now contract the indices λ and ν in the Bianchi identities

$$g^{\lambda\nu}R_{\lambda\mu\nu\kappa;\eta} + g^{\lambda\nu}R_{\lambda\mu\eta\nu;\kappa} + g^{\lambda\nu}R_{\lambda\mu\kappa\eta;\nu} = 0, \tag{7}$$

and recall that the Ricci tensor is

$$R_{\mu\kappa} = g^{\lambda\nu} R_{\lambda\mu\nu\kappa}.$$

So,

$$R_{\mu\kappa;\eta} - R_{\mu\eta;\kappa} + R^{\nu}_{\mu\kappa\eta;\nu} = 0, \tag{8}$$

where the negative sign on the second term comes from the antisymmetric swap of indices ν and η in the second term of (7). Now contract equation (8)

$$g^{\mu\kappa} R_{\mu\kappa;\eta} - g^{\mu\kappa} R_{\mu\eta;\kappa} + g^{\mu\kappa} R^{\nu}_{\mu\kappa\eta;\nu} = 0$$

giving

$$R_{;\eta} - R^{\kappa}_{\eta;k} - R^{\nu}_{\eta;\nu} = 0,$$

which we can write compactly as

$$\left(R\delta^{\mu}_{\eta} - 2R^{\mu}_{\eta}\right)_{;\mu} = 0. \tag{9}$$

Now, raising the index η in (9)

$$R^{\mu\nu} = g^{\eta\nu} R^{\mu}_{\eta}$$

we get

$$\left(g^{\eta\nu}\delta^{\mu}_{\eta} R - 2R^{\mu\nu}\right)_{;\mu} = 0,$$

i.e.

$$\left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R\right)_{;\mu} = 0. \tag{10}$$

Equations (9) and (10) are equivalent and, as we will see, are important relations that will guide us to the field equations of gravity.