Theory of Relativity
Michaelmas Term 2009: M. Haehnelt

8 Curvature

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

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

$\uparrow$

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 define the *covariant derivative* of a contravariant vector (denoted by a semi-colon) as,

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

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 $\sigma = \lambda$

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

these terms cancel

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

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

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

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

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

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

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

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

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

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

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

and from this we can form a scalar

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

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

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

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

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

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

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

[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.

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.

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

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

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

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

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

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

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

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

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.

8.6 Parallel Transport

How do the components of a vector change as we move along a curve? Imagine a curve specified by a parameter τ that varies along the curve $\mathcal{C}$, as in the digram below:

Now imagine a vector field $A^{\mu}(\tau)$ defined along the curve. Normally we think of vector fields defined as functions of the coordinates, but I am asking you to think about a vector field defined along a curve (an example would be the spin of a single particle $S_{\mu}(\tau)$ along the particle's trajectory). In Euclidean space, the condition

$$\frac{dA^{\mu}}{d\tau} = 0 \tag{11}$$

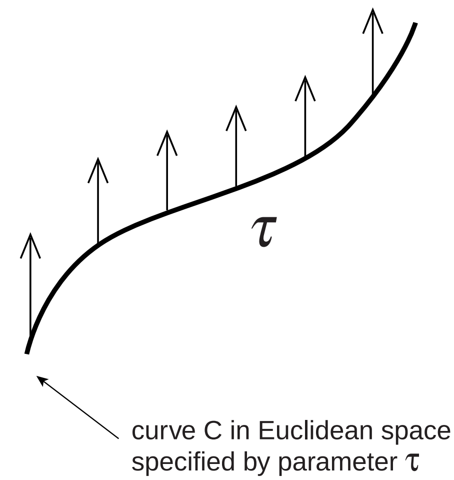

Figure 1: Parallel transport of a vector along a curve C.

tells us that the components of the vector field do not change as we move along the curve. In other words if the condition (11) is satisfied, then the vector is 'parallel transported' along the curve $\mathcal{C}$ (the components of A^μ are fixed as we move along the curve, as shown in the diagram).

Equation (11) can't express parallel transport in non-Euclidean spaces because the quantity $dA^\mu/d\tau$ is not a tensor and so depends on the choice of coordinates.

The obvious generalization is to use the covariant derivative with respect to the parameter τ that specifies the curve (not the covariant derivative with respect to the coordinates, because the vector field is defined only along the curve $\mathcal{C}$).

$$\frac{DA^\mu}{D\tau} = \frac{dA^\mu}{d\tau} + \Gamma^\mu_{\nu\lambda} A^\nu \frac{dx^\lambda}{d\tau} = 0,$$

or for a covariant vector field

$$\frac{DA_\mu}{D\tau} = \frac{dA_\mu}{d\tau} + \Gamma^\lambda_{\mu\nu} A_\lambda \frac{dx^\nu}{d\tau} = 0.$$

If these equations are satisfied, then they will be true in all coordinate systems and so they generalise the concept of parallel transport to non-Euclidean spaces. We can see this explicitly, since the 'parallel transport' condition tells us that

$$\frac{dA^\mu}{d\tau} = -\Gamma^\mu_{\nu\lambda} A^\nu \frac{dx^\lambda}{d\tau}. \tag{12}$$

In other words, if we specify the components A^ν at *some arbitrary point* τ along the curve, equation (12) fixes the components of A^μ *along the entire length of the curve.* Are you worried about whether the transportation is really parallel? Think about an infinitessimal displacement of the vector field. For a small displacement we can choose *locally* Euclidean coordinates in which the Γ's vanish and so setting the covariant derivative equal to zero describes an infinitessimal displacement which keeps the vector parallel ($dA^\mu = 0$).

Now what happens if we parallel transport a vector field around a closed curve $\mathcal{C}$? Define an arbitrary surface A bounding the curve $\mathcal{C}$ and break the area A up into lots of little closed curves $\mathcal{C}_N$, as in the diagram below.

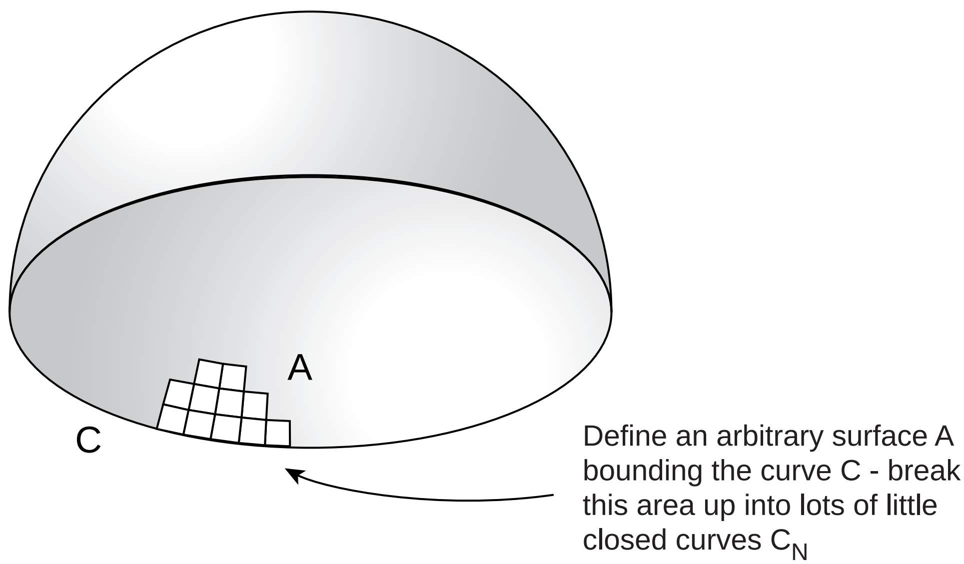

Figure 2: Surface broken up into a contiguous set of closed curves C_N.

The change in the components of A_μ around the closed curve $\mathcal{C}$ is

$$\Delta A_\mu = \sum_N (\Delta A_\mu)_N$$

where $(\Delta A_\mu)_N$ is the change in A_μ around a little closed curve. This follows because the changes in ΔA_μ around any of the interior closed curves cancel, leaving just the contributions around the outer edges that bound the curve $\mathcal{C}$.

Now let's calculate the $(\Delta A_\mu)_N$. Expand $\Gamma^\lambda_{\mu\nu}(x)$ in a Taylor series about an arbitrary point $X = x(\tau_0)$

$$\Gamma^\lambda_{\mu\nu}(x) = \Gamma^\lambda_{\mu\nu}(X) + (x^\rho - X^\rho)\frac{\partial\Gamma^\lambda_{\mu\nu}(X)}{\partial X^\rho} + \ldots$$

and also expand $A_\mu(\tau)$

$$A_\mu(\tau) = A_\mu(\tau_0) + \Gamma^\lambda_{\mu\nu}(X)(x^\nu - X^\nu)A_\lambda(\tau_0) + \ldots$$

Now substitute these expansions into the equation for parallel transport

$$\frac{dA_\mu}{d\tau} = -\Gamma^\lambda_{\mu\nu} A^\lambda \frac{dx^\nu}{d\tau},$$

and integrate

$$\begin{aligned} A_\mu(\tau) = A_\mu(\tau_0) \quad + & \quad \int_{\tau_0}^{\tau} \left[\Gamma^\lambda_{\mu\nu}(X) + (x^\rho - X^\rho) \frac{\partial \Gamma^\lambda_{\mu\nu}(X)}{\partial X^\rho} \right] \\ & \quad \left[A_\lambda(\tau_0) + A_\sigma(\tau_0) \Gamma^\sigma_{\lambda\rho}(X)(x^\rho - X^\rho) \right] \frac{dx^\nu}{d\tau} d\tau. \end{aligned}$$

Expanding the terms in brackets and retaining terms to first order in $x^\rho - X^\rho$,

$$\begin{aligned} A_\mu(\tau) &= A_\mu(\tau_0) + \Gamma^\lambda_{\mu\nu}(X) A_\lambda(\tau_0) \int_{\tau_0}^{\tau} \frac{dx^\nu}{d\tau} d\tau \\ &+ \left[\frac{\partial \Gamma^\sigma_{\mu\nu}(X)}{\partial X^\rho} + \Gamma^\sigma_{\lambda\rho}(X) \Gamma^\lambda_{\mu\nu}(X) \right] A_\sigma(\tau_0) \int_{\tau_0}^{\tau} (x^\rho - X^\rho) \frac{dx^\nu}{d\tau} d\tau. \end{aligned} \qquad (13)$$

If we integrate the coordinate differentials around a closed loop,

$$\oint dx^\nu = 0$$

and so the first integral in (13) vanishes giving,

$$\Delta A_\mu = \left[\frac{\partial \Gamma^\sigma_{\mu\nu}(X)}{\partial X^\rho} + \Gamma^\sigma_{\lambda\rho} \Gamma^\lambda_{\mu\nu} \right] A_\sigma(\tau_0) \oint x^\rho dx^\nu,$$

and interchanging the indices ν and ρ

$$\Delta A_\mu = \left[\frac{\partial \Gamma^\sigma_{\mu\rho}(X)}{\partial X^\nu} + \Gamma^\sigma_{\lambda\nu} \Gamma^\lambda_{\mu\rho} \right] A_\sigma(\tau_0) \oint x^\nu dx^\rho.$$

Using the relation

$$\oint x^\rho dx^\nu = -\oint x^\nu dx^\rho$$

we see that

$$\Delta A_\mu = \frac{1}{2} R^\sigma_{\mu\nu\rho} A_\sigma \oint x^\rho dx^\nu, \qquad (14)$$

where $R^\sigma_{\mu\nu\rho}$ is the *curvature tensor*,

$$R^\lambda_{\mu\nu\kappa} = \frac{\partial \Gamma^\lambda_{\mu\nu}}{\partial x^\kappa} - \frac{\partial \Gamma^\lambda_{\mu\kappa}}{\partial x^\nu} + \Gamma^\eta_{\mu\nu} \Gamma^\lambda_{\kappa\eta} - \Gamma^\eta_{\mu\kappa} \Gamma^\lambda_{\nu\eta}.$$

Equation (14) tells us that the components of A_μ will stay constant after parallel transportation around a small closed curve centred on X *if and only if* the curvature tensor vanishes at

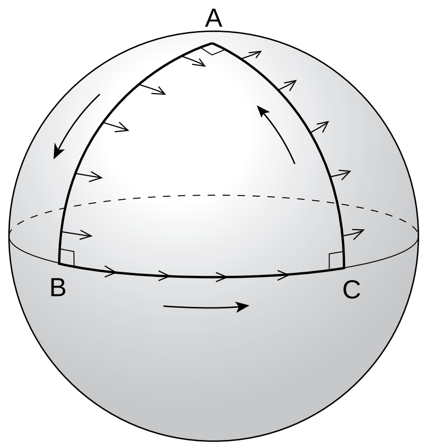

Figure 3: Parallel transport of a vector along the spherical triangle ABC.

the point X. So, returning to our construction of $(\Delta A_\mu)_N$, the vector A_μ will not change on parallel transportation around the entire closed curve $\mathcal{C}$ if the curvature tensor $R^\sigma_{\mu\nu\rho}$ vanishes over the entire area $\mathcal{A}$ bounding the curve.

As an example, consider the parallel transportation of a vector around the closed triangle ABC on the surface of a sphere:

It is self evident that the vector components change direction after parallel transportation around the triangle.

8.7 Curvature of a spherical surface: a worked example

The metric of the surface of a sphere in spherical polar coordinates is

$$ds^2 = a^2 d\theta^2 + a^2 \sin^2\theta d\phi^2. \tag{15}$$

To get you used to handling problems involving curved spaces I would like you to calculate the components of the affine connection starting from the metric (15). The definition of the affine connection is

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

and in two dimensions there are six independent connection components:

$$\Gamma^1_{11}, \quad \Gamma^1_{12}, \quad \Gamma^1_{22}, \quad \Gamma^2_{11}, \quad \Gamma^2_{12}, \quad \Gamma^2_{22}.$$

I have worked out these components in full so that you can check your calculations:

$$\Gamma^1_{11} \;=\; \frac{1}{2} g^{11} \left\{ \frac{\partial g_{11}}{\partial x^1} + \frac{\partial g_{11}}{\partial x^1} - \frac{\partial g_{11}}{\partial x^1} \right\} = 0,$$

$$\begin{aligned}
\Gamma^1_{12} &= \frac{1}{2}g^{11}\left\{\frac{\partial g_{11}}{\partial x^2}+\frac{\partial g_{21}}{\partial x^2}-\frac{\partial g_{12}}{\partial x^1}\right\}=0,\\
\Gamma^1_{22} &= \frac{1}{2}g^{11}\left\{\frac{\partial g_{21}}{\partial x^2}+\frac{\partial g_{21}}{\partial x^2}-\frac{\partial g_{22}}{\partial x^1}\right\}=-\frac{1}{2}g^{11}\frac{\partial g_{22}}{\partial x^1},\\
\Gamma^2_{11} &= \frac{1}{2}g^{22}\left\{\frac{\partial g_{12}}{\partial x^2}+\frac{\partial g_{12}}{\partial x^2}-\frac{\partial g_{11}}{\partial x^2}\right\}=0,\\
\Gamma^2_{12} &= \frac{1}{2}g^{22}\left\{\frac{\partial g_{22}}{\partial x^1}+\frac{\partial g_{12}}{\partial x^2}-\frac{\partial g_{11}}{\partial x^2}\right\}=\frac{1}{2}g^{22}\frac{\partial g_{22}}{\partial x^1},\\
\Gamma^2_{22} &= \frac{1}{2}g^{22}\left\{\frac{\partial g_{22}}{\partial x^2}+\frac{\partial g_{22}}{\partial x^2}-\frac{\partial g_{22}}{\partial x^2}\right\}=0.
\end{aligned}$$

So, you should end up with only two non-zero components

$$\begin{aligned}
\Gamma^1_{22} &= -\frac{1}{2a^2}a^2 2\sin\theta\cos\theta=-\sin\theta\cos\theta,\\
\Gamma^2_{12} &= \frac{1}{2a^2\sin^2\theta}2a^2\sin\theta\cos\theta=\frac{\cos\theta}{\sin\theta}.
\end{aligned}$$

The curvature tensor is

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

and in two dimensions the symmetry properties of this tensor means that there is only *one* independent component. We can take this to be R_{1212}, so fortunately we only have to calculate this single component:

$$\begin{aligned}
R_{1212} &= \frac{1}{2}\left[\frac{\partial^2 g_{11}}{\partial (x^2)^2}-\frac{\partial^2 g_{21}}{\partial x^2\partial x^1}-\frac{\partial^2 g_{12}}{\partial x^1\partial x^2}+\frac{\partial^2 g_{22}}{\partial (x^1)^2}\right]+g_{\eta\sigma}\left[\Gamma^\eta_{11}\Gamma^\sigma_{22}-\Gamma^\eta_{21}\Gamma^\sigma_{21}\right]\\
&= \frac{1}{2}\left[\frac{\partial^2 g_{22}}{\partial (x^1)^2}\right]+g_{11}\left[\Gamma^1_{11}\Gamma^1_{22}-\Gamma^1_{12}\Gamma^1_{12}\right]+g_{22}\left[\Gamma^2_{11}\Gamma^2_{22}-\Gamma^2_{21}\Gamma^2_{21}\right]\\
&= -a^2\sin^2\theta.
\end{aligned}$$

So, you have now understood the origin of Gauss's curvature parameter that I introduced in Section 5. Gauss's curvature parameter is just

$$K=-\frac{R_{1212}}{g}=\frac{a^2\sin^2\theta}{a^4\sin^2\theta}=\frac{1}{a^2},$$

and if you write out R_{1212} in full in terms of derivatives of the $g_{\mu\nu}$ you will recover the formidably complicated equation for the Gaussian curvature K of Section 5.

8.8 Surface of a cylinder

Instead of a spherical surface, consider the surface of a cylinder of radius a. The metric of the surface in cylindrical polar coordinates is

$$ds^2 = a^2 d\theta^2 + dz^2,$$

and it is obvious that this two-dimensional space is spatially flat because we can transform the metric into the form

$$ds^2 = dx^2 + dz^2,$$

by the coordinate transformation $x = a\theta$. It therefore follows that the curvature of the cylindrical surface vanishes. Furthermore, since the curvature tensor vanishes everywhere, the components of a vector will remain unchanged if the vector is parallel transport around any closed curve on the cylinder. For example, this is evident on parallel transportation of a vector around the closed rectangle ABCD in the figure below:

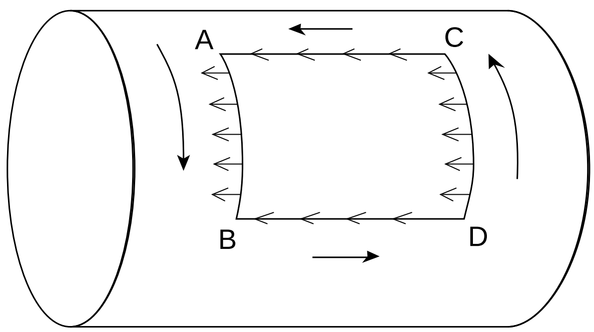

Figure 4: Parallel transport of a vector along the rectangle $ABCD$ on the surface of a cylinder

8.9 Geodesic Deviation

Consider the motion of two freely falling particles on nearby trajectories,

$$x^\mu(\tau),$$
$$x^\mu(\tau) + \delta x^\mu(\tau).$$

Each of these trajectories must obey the geodesic equations of motion:

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda}(x)\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0, \tag{16}$$

$$\frac{d^2\left[x^\mu + \delta x^\mu\right]}{d\tau^2} + \Gamma^\mu_{\nu\lambda}\left(x^\kappa + \delta x^\kappa\right)\frac{d}{d\tau}\left(x^\nu + \delta x^\nu\right)\frac{d}{d\tau}\left(x^\lambda + \delta x^\lambda\right) = 0. \tag{17}$$

Expand (17) in a Taylor series to first order in δx^λ,

$$\frac{d^2x^\mu}{d\tau^2}+\frac{d^2\delta x^\mu}{d\tau^2}+\Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}+\frac{\partial\Gamma^\mu_{\nu\lambda}}{\partial x^\rho}\delta x^\rho\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}+\Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{d\delta x^\lambda}{d\tau}+\Gamma^\mu_{\nu\lambda}\frac{dx^\lambda}{d\tau}\frac{d\delta x^\nu}{d\tau}=0,$$

and subtract (16)

$$\frac{d^2\delta x^\mu}{d\tau^2}+\frac{\partial\Gamma^\mu_{\nu\lambda}}{\partial x^\rho}\delta x^\rho\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}+2\Gamma^\mu_{\mu\lambda}\frac{dx^\nu}{d\tau}\frac{d\delta x^\lambda}{d\tau}=0.$$

This is quite a complicated equation to work with, but we can simplify it considerably by defining a set of *local* Euclidean coordinates ξ^μ. In such a coordinate system, we know that the Γ's vanish. The derivatives of Γ with respect to the coordinates will not necessarily vanish (as we will see below). In the locally Euclidean coordinate system (17) simplifies to

$$\frac{d^2\delta\xi^\mu}{d\tau^2}+\frac{\partial\Gamma^\mu_{\nu\lambda}}{\partial\xi^\rho}\delta\xi^\rho\frac{d\xi^\nu}{d\tau}\frac{d\xi^\lambda}{d\tau}=0. \qquad (18)$$

The covariant derivative of $\delta\xi^\mu$ with respect to τ is

$$\frac{D\delta\xi^\mu}{D\tau}=\frac{d\delta\xi^\mu}{d\tau}+\Gamma^\mu_{\nu\lambda}\frac{d\xi^\lambda}{d\tau}\delta\xi^\nu,$$

now don't set the Γ's equal to zero at this stage because we want to differentiate again and this will introduce a derivative of the Γ's.

$$\frac{D^2\delta\xi^\mu}{D\tau^2}=\frac{d^2\delta\xi^\mu}{d\tau^2}+\frac{\partial\Gamma^\mu_{\nu\lambda}}{\partial\xi^\sigma}\frac{d\xi^\sigma}{d\tau}\frac{d\xi^\lambda}{d\tau}\delta\xi^\nu.$$

I have now set the Γ's equal to zero in this equation but retained the first derivative of Γ. Now substitute into equation (18),

$$\frac{D^2\delta\xi^\mu}{D\tau^2}+\frac{\partial\Gamma^\mu_{\nu\lambda}}{\partial\xi^\rho}\delta\xi^\rho\frac{d\xi^\nu}{d\tau}\frac{d\xi^\lambda}{d\tau}-\frac{\partial\Gamma^\mu_{\nu\lambda}}{\partial\xi^\sigma}\frac{d\xi^\sigma}{d\tau}\frac{d\xi^\lambda}{d\tau}\delta\xi^\nu=0,$$

i.e.,

$$\frac{D^2\delta\xi^\mu}{D\tau^2}+\left(\frac{\partial\Gamma^\mu_{\sigma\lambda}}{\partial\xi^\rho}-\frac{\partial\Gamma^\mu_{\rho\lambda}}{\partial\xi^\sigma}\right)\frac{d\xi^\sigma}{d\tau}\frac{d\xi^\lambda}{d\tau}\delta\xi^\rho=0.$$

But, in the locally Euclidean coordinate system in which the Γ's are zero, the curvature tensor is simply

$$R^\mu_{\lambda\sigma\rho}=\frac{\partial\Gamma^\mu_{\sigma\lambda}}{\partial\xi^\rho}-\frac{\partial\Gamma^\mu_{\lambda\rho}}{\partial\xi^\sigma}.$$

So, we can write a generally covariant equation describing *geodesic deviation*,

$$\frac{D^2\delta x^\mu}{D\tau^2}+R^\mu_{\lambda\sigma\rho}\frac{dx^\sigma}{d\tau}\frac{dx^\lambda}{d\tau}\delta x^\rho=0. \qquad (19)$$

Notice that I have replaced the locally Euclidean coordinates by arbitrary coordinates x^μ. I can do this because equation (19) is a tensorial equation and so is true in any coordinate system. The locally Euclidean coordinate system that I introduced is just a device to simplify the algebra in deriving the tensorial equation (19).

The physical meaning of equation (19) is straightforward. If space is curved ($R^\mu_{\lambda\sigma\rho} \neq 0$) two neighbouring geodesics (or freely falling particles) will either converge or diverge, depending on the local curvature. For example, the two neighbouring geodesics AB and AC on the surface of the sphere converge as we approach the point A at the pole because the surface is positively curved.

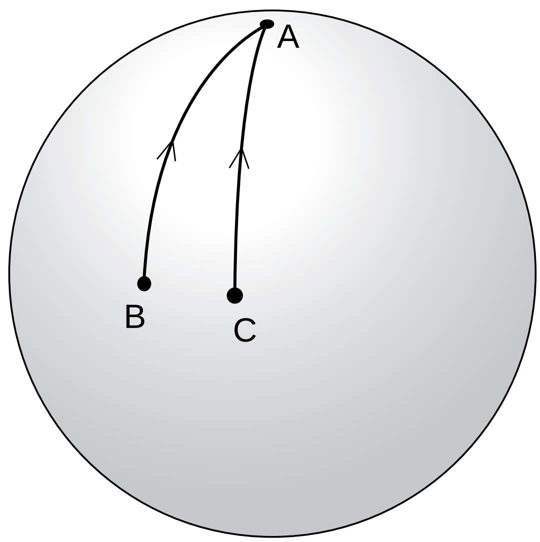

Figure 5: Convergence of two geodesics towards the point A on the surface of a sphere.

Equation (19) allows us to compute the rates of convergence or divergence of neighbouring geodesics for Riemannian spaces of arbitrary complexity. All you have to do is to compute the curvature tensor at each point from the metric.