2008–11–10

The geometry is specified by $g_{\mu\nu}(x)$. But just looking at the metric does not really tell you what the geometry is, since we have an enormous freedom in choosing the coordinate system. Is it flat, is it curved? Finding a coordinate transformation that makes this obvious is not trivial.

We had the affine connection, constructed from the metric as

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}\, g^{\lambda\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu})$$

But this is not a tensor. It can be nonzero even in flat space, so it is not a good measure of whether space is curved or not. The main reason why this is important although it is not a tensor, is that we can use it to define a covariant derivative D_μ, since the ordinary derivative does not, in general, give us a tensor.

$$\mathrm{D}_\mu V^\nu = \partial_\mu V^\nu + \Gamma^{\mu}_{\nu\lambda} V^\lambda$$

This is the only sensible way of defining a derivative, as can be seen by using the equivalence principle. In the local inertial frame it is just the ordinary derivative, and the rest is just a coordinate transformation. For lower indices:

$$\mathrm{D}_\mu V_\nu = \partial_\mu V_\nu - \Gamma^\lambda_{\mu\nu} V_\lambda$$

Now we are ready to take the step to define curvature. We will not follow the beginning of Chapter 6 in Weinberg, but go directly to section 6.5. We look at how the covariant derivative works. Let's write $\mathrm{D}_\mu V^\nu$ in a slightly more suggestive way. Consider V^μ as a vector, and for each μ, we view $(\Gamma_\mu)^\nu_\lambda$ as a matrix:

$$\mathrm{D}_\mu V^\nu = [(\partial_\mu + \Gamma_\mu)V]^\nu$$

(The derivative ∂_μ is understood to be $I\,\partial_\mu$ where I is the identity matrix.) This is just ordinary matrix multiplication. The result is a tensor.

Now, what happens if we take two derivatives after each other? Say we are in some point P, and want to go to some other point Q. (Consider this to be infinitesimally small.) We take one step along one coordinate direction, and one step along another. Doing parallel transport of a vector. Γ will rotate the vector and we can get different vectors at Q depending on which order we took the displacements from P.

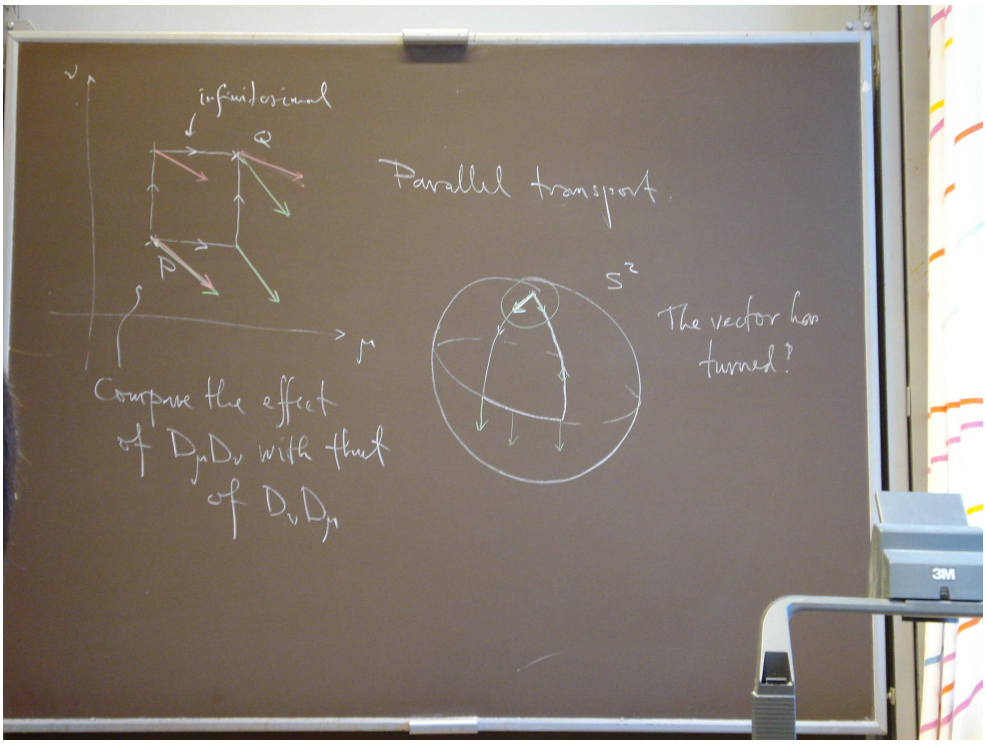

Figure 1.

Looking at parallel transport on a sphere: Say we start with a vector / a stick on the north pole, pointing southward. Then we take the vector and walk down to the equator — the vector still points southward. Then, without changing the direction of the stick, walk along the equator,

and then go back up to the north pole. The initial direction and the final direction of the vector will differ. The vector has turned, due to the geometry of the sphere.

Now, back to the infinitesimal case. Compare the effect of $\mathrm{D}_\mu\mathrm{D}_\nu$ with that of $\mathrm{D}_\nu\mathrm{D}_\mu$. To avoid drowning in indices, we will use the matrix notation, V denoting the vector with indices V^λ.

$$\mathrm{D}_\mu\mathrm{D}_\nu V = (\partial_\mu + \Gamma_\mu)(\partial_\nu + \Gamma_\nu)V = \partial_\mu\partial_\nu V + \partial_\mu(\Gamma_\nu V) + \Gamma_\mu\partial_\nu V + \Gamma_\mu\Gamma_\nu V$$

(Note: these are matrices. The order is important.)

$$= \widetilde{\partial_\mu\partial_\nu V} + (\partial_\mu\Gamma_\nu)V + \widetilde{\Gamma_\nu\partial_\mu V + \Gamma_\mu\partial_\nu V} + \Gamma_\mu\Gamma_\nu V$$

Let us now study the difference, let us study $(\mathrm{D}_\mu\mathrm{D}_\nu - \mathrm{D}_\nu\mathrm{D}_\mu)V$. I have marked the terms that are symmetric in μ and ν with a wide tilde above: these we do not have to care about when calculating $(\mathrm{D}_\mu\mathrm{D}_\nu - \mathrm{D}_\nu\mathrm{D}_\mu)V$.

$$(\mathrm{D}_\mu\mathrm{D}_\nu - \mathrm{D}_\nu\mathrm{D}_\mu)V = \underbrace{(\partial_\mu\Gamma_\nu - \partial_\nu\Gamma_\mu + \Gamma_\mu\Gamma_\nu - \Gamma_\nu\Gamma_\mu)}_{\text{Matrix! Two indices are suppressed.}}V$$

All derivatives of V have dropped out. $[\mathrm{D}_\mu, \mathrm{D}_\nu] \equiv \mathrm{D}_\mu\mathrm{D}_\nu - \mathrm{D}_\nu\mathrm{D}_\mu$ is not a differential operator. It is just multiplicative. It is, actually, what we call the curvature. With indices written out:

$$([\mathrm{D}_\mu, \mathrm{D}_\nu]V)^\lambda = -R_{\mu\nu}{}^\lambda{}_\kappa V^\kappa$$

(or matrix-wise, we can think of $R_{\mu\nu}{}^\lambda{}_\kappa$ as $(R_{\mu\nu})^\lambda{}_\kappa$) where

$$R_{\mu\nu}{}^\lambda{}_\kappa = -(\partial_\mu\Gamma_\nu - \partial_\nu\Gamma_\mu + \Gamma_\mu\Gamma_\nu - \Gamma_\nu\Gamma_\mu)^\lambda{}_\kappa =$$

$$= -\left(\partial_\mu\Gamma^\lambda_{\nu\kappa} - \partial_\nu\Gamma^\lambda_{\nu\kappa} + \Gamma^\lambda_{\mu\rho}\Gamma^\rho_{\nu\kappa} - \Gamma^\lambda_{\nu\rho}\Gamma^\rho_{\mu\kappa}\right)$$

$R_{\mu\nu}{}^\lambda{}_\kappa$ is the matrix that encodes the difference between vectors that are parallel transported along ν, μ and μ, ν. This is *the* tensor that tells us how the space is curved, and it tells us everything about how the space is curved. Why the minus sign? That is Weinberg's fault. It is more common to have a plus sign in the definition. It is a convention, and Martin Cederwall would have preferred the plus sign.

Let us take a couple of examples. Last week we did polar coordinates in flat space — let's calculate the curvature of flat space using polar coordinates. (If someone asked you to calculate the curvature of flat space, it would be more expedient to use Cartesian coordinates, making $\Gamma^\lambda_{\mu\nu}$ vanish and $R_{\mu\nu}{}^\lambda{}_\kappa = 0$.)

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\,\mathrm{d}\varphi^2$$

$$\Gamma^r_{\varphi\varphi} = -r, \quad \Gamma^\varphi_{r\varphi} = \Gamma^\varphi_{\varphi r} = \frac{1}{r}, \quad \text{all other } \Gamma^\mu_{\nu\lambda} \text{ vanish.}$$

$$-R_{r\varphi}{}^r{}_\varphi = \partial_r\Gamma^r_{\varphi\varphi} - \Gamma^r_{\varphi\varphi}\Gamma^\varphi_{r\varphi} = -1 - (-r)\frac{1}{r} = 0$$

[and the other one]

This vanishes. Since $R_{\mu\nu}{}^\lambda{}_\kappa$ is a tensor, we know that it vanishing in one coordinate system means that it's vanishes in all coordinate systems.

Let us take a close look at curvature and see what it is. We have already mentioned that there is antisymmetry in the $\mu\nu$ indices of $R_{\mu\nu}{}^{\lambda}{}_{\kappa}$.

$$R_{\mu\nu\rho\sigma} = g_{\rho\lambda} R_{\mu\nu}{}^{\lambda}{}_{\sigma} = \ldots =$$

$$= \frac{1}{2}\left(\partial_\mu\partial_\rho g_{\nu\sigma} - \partial_\nu\partial_\rho g_{\mu\sigma} - \partial_\mu\partial_\sigma g_{\nu\rho} + \partial_\nu\partial_\sigma g_{\mu\rho}\right) + g_{\kappa\lambda}\left(\Gamma^{\kappa}_{\mu\rho}\Gamma^{\lambda}_{\nu\sigma} - \Gamma^{\kappa}_{\mu\sigma}\Gamma^{\lambda}_{\nu\rho}\right)$$

The reason for writing it like this is that the symmetries are more obvious. We will not actually use this when calculating the curvature tensor. We see here that we have antisymmetry in $[\mu\nu]$ and in $[\rho\sigma]$. This is not all. $\mu\nu \leftrightarrow \rho\sigma$ symmetric.

$$R_{\nu\mu\rho\sigma} = -R_{\mu\nu\rho\sigma}, \quad R_{\mu\nu\sigma\rho} = -R_{\mu\nu\rho\sigma}, \quad R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma} \tag{1}$$

Sum the cyclic permutations of the last three indices:

$$R_{\mu\nu\rho\sigma} + R_{\mu\rho\sigma\nu} + R_{\mu\sigma\nu\rho} \equiv 3R_{\mu[\nu\rho\sigma]} = 0 \tag{2}$$

(Example: $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \equiv 2\,\partial_{[\mu}A_{\nu]}$).

Now we want to count the number of independent components in $R_{\mu\nu\rho\sigma}$, in n dimensional space.

$[\mu\nu]$ antisymmetric $D \times D$-matrix: $D(D-1)/2$. (Symmetric: $D\,(D+1)/2$)

Using (1): "R is symmetric $\frac{D(D-1)}{2} \times \frac{D(D-1)}{2}$-matrix"

$$\rightarrow \quad \frac{\frac{D(D-1)}{2}\left(\frac{D(D-1)}{2}+1\right)}{2} = \frac{1}{8}D(D-1)\left(D^2 - D + 2\right)$$

But this is not all, we have to take (2) into account. We subtract the number of components "in $R_{\mu[\nu\rho\sigma]}$". (This step is very easy to get wrong. Given the first properties (1), for any tensor having these properties: $R_{\mu[\nu\rho\sigma]} = R_{[\mu\nu\rho\sigma]}$ — nice homework.) $D(D-1)(D-2)(D-3)$ – the first index can be chosen in D ways, the next index must be different, and so on. And we have to divide by the number of permutations:

$$\frac{D(D-1)(D-2)(D-3)}{4!}$$

So we subtract this number. The final result for the number of components in $R_{\mu\nu\rho\sigma}$ is

$$\frac{1}{8}D(D-1)\left(D^2 - D + 2\right) - \frac{1}{24}D(D-1)(D-2)(D-3) = \frac{1}{12}D^2\left(D^2 - 1\right)$$

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

What can we do with this? We can form smaller tensors by contracting indices.

$$R_{\mu\nu} \equiv R_{\lambda\mu}{}^{\lambda}{}_{\nu}$$

We call $R_{\mu\nu}{}^{\kappa}{}_{\lambda}$ the curvature tensor (Riemann tensor). We call $R_{\mu\nu}$ the Ricci tensor. $R_{\mu\nu}$ is symmetric in $\mu\nu$. We can also form

$$R = g^{\mu\nu} R_{\mu\nu}$$

which is called the curvature scalar. Now Einstein's equations are not too far away.

Reminder: "current" for gravity is $T_{\mu\nu}$. $R_{\mu\nu} + ... \propto T_{\mu\nu}$? This could be field equations for gravity. Why is it nice? In the Ricci tensor we have two derivatives on $g_{\mu\nu}$, and we need to reproduce Newton's gravity: $\nabla^2\Phi \sim \rho$.

Riemann:

Riemann:	$\frac{1}{12} D^2\big(D^2-1\big)$
Ricci:	$\frac{1}{2} D(D+1)$
Curvature scalar:	1

D	Riemann	Ricci	R
1	0	0^*	0^*
2	1	1^*	1^*
3	6	6	1
4	20	10	1
5	...		

* We cannot use the formula here, due to the low dimensionality of space.

In two dimensions we only have to give R_{1212} to know everything. That is why, above, we only needed to calculate $R_{r\varphi r\varphi}$.

In three dimensions, the Ricci tensor contains all the information of the Riemann tensor. That means that the geometry is completely fixed by the matter distribution. No gravitational waves.

In four dimensions, we have more information in the Riemann tensor than in the Ricci tensor. This is very interesting. We can have gravitational waves.