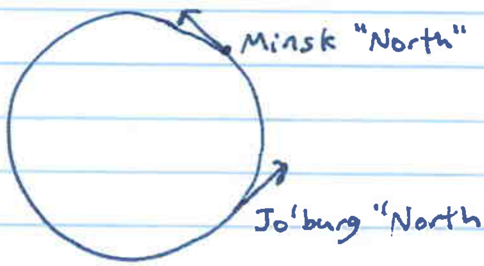
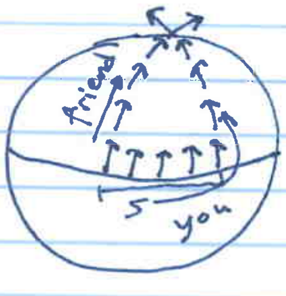
Curvature

You and a friend wanted to see the world. You decided to check out Johannesburg, South Africa, while your friend went up to Minsk, Belarus. During a skype call one day you say to your friend "Point North," just because. If you were away from Earth, you would see two people pointing like:



This is curious: though you are doing the exact same thing, the directions that you and your friend point are very different! (Flat earth is not assumed) It doesn't make sense to compare these two "North" vectors because they are different. However, they are different precisely because Minsk and Jo'burg are separated. On a non-linear manifold, like a sphere, we can't compare vectors at different points. It makes no sense because we know they are going to be different, because they have to follow the undulations of the manifold. On a non-linear manifold, we can only compare vectors at the same point. So, how do we do this and what information does that provide?

Let's imagine that you and your friend meet up again, this time on the equator. You decide to both travel to the North pole, and each of you will always point an arrow north along your walk. Your friend travels to the north pole directly, along a line of longitude. You, by contrast, travel a distance s along the equator, and then go North. Your two paths look like:



Once you both reach the North pole, you compare your arrows. If your friend's arrow can be represented by the vector:

$\vec{V}_1 = (1,0)$

your arrow is rotate with respect to it:

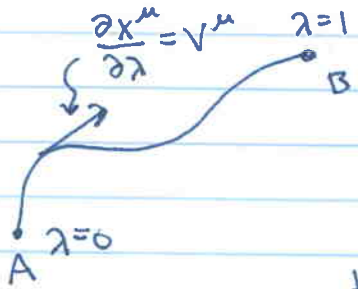
$\vec{V}_2 = \left(\cos\frac{s}{r}, \sin\frac{s}{r}\right)$, where r is the radius of Earth.

The length of the difference of these two vectors is

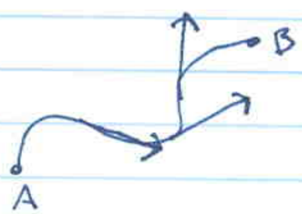
$|\vec{V}_1 - \vec{V}_2| = 2\sin\frac{s}{2r}$. So, knowing how much further you traveled than your friend, you can measure the radius of Earth! One might also say you can measure the curvature.

This procedure of moving a vector on a manifold is called "parallel transport" and is necessary for comparing vectors. You're likely familiar with parallel transport, at least on flat space. You know that, as long as orientation is maintained, you are free to move a vector around however you want, and it's the same vector. Parallel transport formalizes this action to an arbitrary manifold. As observed in the example above, however, the path along which the vector is parallel transported is important in determining the vector's final orientation. We'll discuss next lecture how we can formalize the observation that by comparing different paths of parallel transport, we can learn about the curvature of the manifold.

For now, let's use the notion of parallel transport to identify "shortest paths" on our manifold. On flat space the shortest path is of course a straight line between two points. A path between two points defines a special vector: its tangent vector. For a path P between points A and B, it's tangent vector is:



The path is parametrized by λ, which just is some map from [0,1] to every point on the curve. The tangent vector V^μ is just the vector formed from the derivatives of the coordinates wrt the parameter λ. This picture provides us with a beautiful definition of a shortest distance: the path of shortest distance parallel transports its tangent vector. Let's see how this works in flat space. Parallel transport maintains orientation, but the tangent vector on an arbitrary path flops around:



Not parallel transported! Only straight-line paths parallel transport their tangent vectors in flat space:

From this definition of shortest distance, the book derives the geodesic equation, which is a differential equation for the functional form of the path of shortest distance (a geodesic) on a general manifold. I will present another ~~der~~ derivation, directly from the "shortest path" requirement.

The length ℓ of a path from A to B on a manifold is just the integral of the line element:

$$\ell = \int_A^B \sqrt{ds^2} = \int_A^B \sqrt{g_{\mu\nu} dx^\mu dx^\nu},$$

where x are coordinates on the manifold and $g_{\mu\nu}$ is its metric. We want to minimize l wrt the path, so this means function differentiation! It's useful to use the λ-parametrization from before, where

$$l = \int_0^1 \sqrt{g_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda}} \, d\lambda .$$

Now we want to vary the path coordinates x^μ and demand that it is stationary. That is, we consider modifying the coordinate as:

$$x^\mu \to x^\mu + \varepsilon^\mu, \quad \text{for some } \varepsilon^\mu \text{ that is infinitesimal.}$$

We also have to vary the metric, $g_{\mu\nu}$. It is some function of the coordinates, so we can Taylor expand:

$$g_{\mu\nu}(x^\sigma + \varepsilon^\sigma) = g_{\mu\nu}(x^\sigma) + \varepsilon^\sigma \partial_\sigma g_{\mu\nu} + \ldots$$

Then, the length of our modified curve is:

$$l[x^\mu + \varepsilon^\mu] = \int \sqrt{(g_{\mu\nu} + \varepsilon^\sigma \partial_\sigma g_{\mu\nu}) \frac{\partial}{\partial \lambda}(x^\mu + \varepsilon^\mu) \frac{\partial}{\partial \lambda}(x^\nu + \varepsilon^\nu)} \, d\lambda$$

$$= \int \sqrt{g_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda} + \varepsilon^\sigma \partial_\sigma g_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda} + g_{\mu\nu} \frac{\partial \varepsilon^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda} + g_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial \varepsilon^\nu}{\partial \lambda}} \, d\lambda$$

$$= \int \left\{ \sqrt{g_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda}} + \frac{1}{2\sqrt{g_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda}}} \left[\varepsilon^\sigma \partial_\sigma g_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda} + g_{\mu\nu} \frac{\partial \varepsilon^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda} + g_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial \varepsilon^\nu}{\partial \lambda} \right] \right\} d\lambda$$

Okay, we need to do the usual thing and move derivatives off of ε. Let's consider the first term with a derivative:

$$g_{\mu\nu}\frac{\partial \varepsilon^\mu}{\partial\lambda}\frac{\partial x^\nu}{\partial\lambda} = \frac{\partial}{\partial\lambda}\left(g_{\mu\nu}\varepsilon^\mu\frac{\partial x^\nu}{\partial\lambda}\right) - \varepsilon^\mu\frac{\partial}{\partial\lambda}\left(g_{\mu\nu}\frac{\partial x^\nu}{\partial\lambda}\right)$$

As always, we ignore the total derivative as it just evaluates at the fixed endpoints. The other derivative evaluates via the product rule:

$$\frac{\partial}{\partial\lambda}\left(g_{\mu\nu}\frac{\partial x^\nu}{\partial\lambda}\right) = \frac{\partial g_{\mu\nu}}{\partial\lambda}\frac{\partial x^\nu}{\partial\lambda} + g_{\mu\nu}\frac{\partial^2 x^\nu}{\partial\lambda^2}$$

$$= \partial_\sigma g_{\mu\nu}\frac{\partial x^\sigma}{\partial\lambda}\frac{\partial x^\nu}{\partial\lambda} + g_{\mu\nu}\frac{\partial^2 x^\nu}{\partial\lambda^2}$$

On the second line we used our old friend the chain rule. Then, we have, for this term

$$g_{\mu\nu}\frac{\partial\varepsilon^\mu}{\partial\lambda}\frac{\partial x^\nu}{\partial\lambda} = -\varepsilon^\mu(\partial_\sigma g_{\mu\nu})\frac{\partial x^\sigma}{\partial\lambda}\frac{\partial x^\nu}{\partial\lambda} - \varepsilon^\mu g_{\mu\nu}\frac{\partial^2 x^\nu}{\partial\lambda^2}$$

We find a similar expression (with appropriate index switches) for the final term. Then, the term in square brackets is:

$$\varepsilon^\sigma\Big[\partial_\sigma g_{\mu\nu}\frac{\partial x^\mu}{\partial\lambda}\frac{\partial x^\nu}{\partial\lambda} - (\partial_\alpha g_{\sigma\nu})\frac{\partial x^\alpha}{\partial\lambda}\frac{\partial x^\nu}{\partial\lambda} - g_{\sigma\nu}\frac{\partial^2 x^\nu}{\partial\lambda^2}$$
$$- (\partial_\alpha g_{\mu\sigma})\frac{\partial x^\mu}{\partial\lambda}\frac{\partial x^\alpha}{\partial\lambda} - g_{\mu\sigma}\frac{\partial^2 x^\mu}{\partial\lambda^2}\Big]$$

The whole thing in square brackets must vanish for the shortest path. Relabeling indices and combining terms, this is:

$$2g_{\mu\sigma}\frac{\partial^2 x^\mu}{\partial\lambda^2} + \left(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right)\frac{dx^\mu}{\partial\lambda}\frac{\partial x^\nu}{\partial\lambda} = 0$$

Or, using the identity $g_{\mu\sigma} g^{\sigma\rho} = \delta^{\rho}_{\mu}$, we have

$$\boxed{\frac{d^2 x^{\rho}}{d\lambda^2} + \frac{1}{2} g^{\sigma\rho}\left(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}\right) \frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = 0}$$

This is called the <u>geodesic</u> <u>equation</u>. In flat space, the derivatives of the metric vanish $\partial g = 0$, so this reduces to

$$\frac{d^2 x^\rho}{d\lambda^2} = 0,$$ which is the equation

for a straight line, as expected. The combination of metrics is called the Levi-Civita connection or the Christoffel connection:

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

Don't let the indices fool you, though. The connection is <u>not</u> a tensor; i.e., it doesn't transform with Jacobian matrices. As a connection, these symbols encode curvature of a manifold. Straight lines on an arbitrary manifold aren't straight: the connection ensures that curvature is accounted for appropriately in measuring distances of extended curves.

In particular, the connection allows you to define a derivative on the manifold that satisfies all that we know and love: linearity and the Leibniz product rule. This derivative is denoted by ∇_μ and is called a <u>covariant</u> <u>derivative</u>. When acting on a vector V^ν, it transforms like a tensor:

$$\nabla_\mu V^\nu \to \frac{\partial x_\mu}{\partial x'_\alpha} \frac{\partial x^\nu}{\partial x'^\beta} \nabla_\alpha V^\beta$$

where x and x' are the new and old coordinates, respectively.

Recall that the regular partial derivative ∂_μ when acting on a vector does <u>not</u> transform like a tensor:

$$\partial_\mu V^\nu \to \frac{\partial x_\mu}{\partial x'^\alpha}\partial_\alpha\left(\frac{\partial x^\nu}{\partial x'^\beta}V^\beta\right) = \frac{\partial x_\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta}\partial_\alpha V^\beta + \frac{\partial x_\mu}{\partial x'^\alpha}V^\beta\frac{\partial^2 x^\nu}{\partial x^\alpha\partial x'^\beta}$$

The covariant derivative can then be expressed as:

$$\boxed{\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda}V^\lambda}$$

where $\Gamma^\nu_{\mu\lambda}$ is the connection. This linear form of the covariant derivative follows necessarily from linearity of a derivative. Additionally, at this stage, the connection is undefined; all that is required is that it transforms in such a way to ensure that the covariant derivative is a tensor. However, the Christoffel connection can be uniquely defined from the two requirements:

Torsion-free: $\Gamma^\rho_{\mu\nu} = \Gamma^\rho_{(\mu\nu)}$

Metric-Compatibility: $\nabla_\mu g_{\rho\sigma} = 0$

Everything we study in this class (and GR more broadly) is based on a torsion-free connection. Metric compatibility is a very nice property to have; among other things, it ensures that the raising and lowering of indices commutes with the covariant derivative. The book demonstrates how the Christoffel connection follows from these properties. Apparently, it is pretty important as we've seen two instances where it comes up. More next lecture...