Curvature.

Asaf Pe'er[1]

May 14, 2020

This part of the course is based on Refs. [1], [2] and [3]. All figures are taken from Sean Carroll's notes in "Level 5: A Knowledgebase for Extragalactic Astronomy and Cosmology" (http://ned.ipac.caltech.edu/level5/March01/Carroll3/Carroll3.html).

1. Introduction

Mastering (hopefully) the treatments of vectors and tensors in curved space time, we can now write covariant physical equations in such spaces. Recall our basic goal: **we want to know how do gravity curves space-time**.

Before answering this question, we first have to clarify what do we exactly mean by "curved space time"? **How do we know that a space-time is "curved"?**

What we look for is a mathematical description of "curvature" of space-time. Alternatively, we would like to be able to tell, being *inside* space - time, that it is "curved". (You may think of earth as being "curved", but this is because earth is embedded in 3-d universe; namely, we can look at it from the "outside", as astronauts do... and indeed, many years ago, it was not at all obvious that earth is round !. But what about the entire universe ? We can't look at the universe from the outside, so is there alternative way?)

The purpose of this chapter is to show that there is a way (first found by *Gauss*) of determining the curvature of space-time, using only intrinsic quantities.

2. Parallel transport

The first thing we need to discuss is **parallel transport** of vectors and tensors, which we touched upon in the last part of the last chapter. Recall that in flat space it was unnecessary to be very careful about the fact that vectors were elements of tangent spaces defined at individual points; it is actually very natural to compare vectors at different points (where by "compare" we mean add, subtract, take the dot product, etc.). The reason why it is natural

[1]Department of Physics, Bar-Ilan University

is because it makes sense, in flat space, to “move a vector from one point to another while keeping it constant” (see Figure 1). Then once we get the vector from one point to another we can do the usual operations allowed in a vector space.

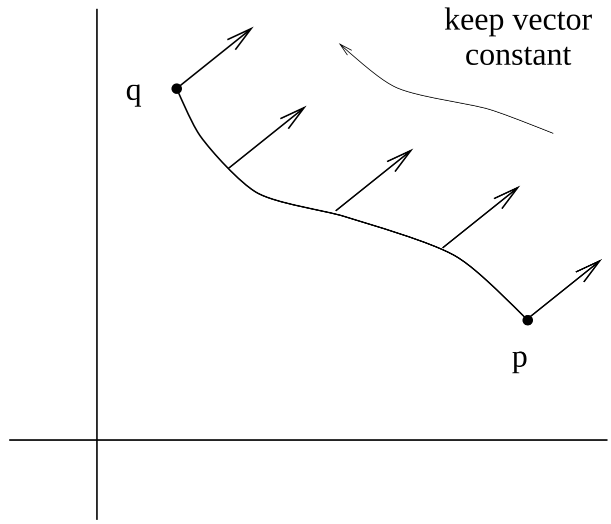

Fig. 1.— Parallel transport of a vector.

The concept of moving a vector along a path, keeping constant all the while, is known as parallel transport. As we shall see, parallel transport is defined whenever we have a connection; the intuitive manipulation of vectors in flat space makes implicit use of the Christoffel connection on this space. The crucial difference between flat and curved spaces is that, in a curved space, **the result of parallel transporting a vector from one point to another will depend on the path taken between the points**. This is most clearly demonstrated when parallel transporting a vector along a sphere, as is demonstrated in Figure 2. Clearly, the vector, parallel transported along two paths, arrived at the same destination with two different values (rotated by an angle θ).

It therefore appears as if **there is no natural way to uniquely move a vector from one tangent space to another**; we can always parallel transport it, but the result depends on the path, and there is no natural choice of which path to take. Unlike some of the problems we have encountered, **there is no solution to this one** — we simply must learn to live with the fact that two vectors can only be compared in a natural way if they are elements of the same tangent space !.

As an example, two particles passing by each other have a well-defined relative velocity (which cannot be greater than the speed of light). But two particles at different points on a curved manifold do not have any well-defined notion of relative velocity — the concept simply

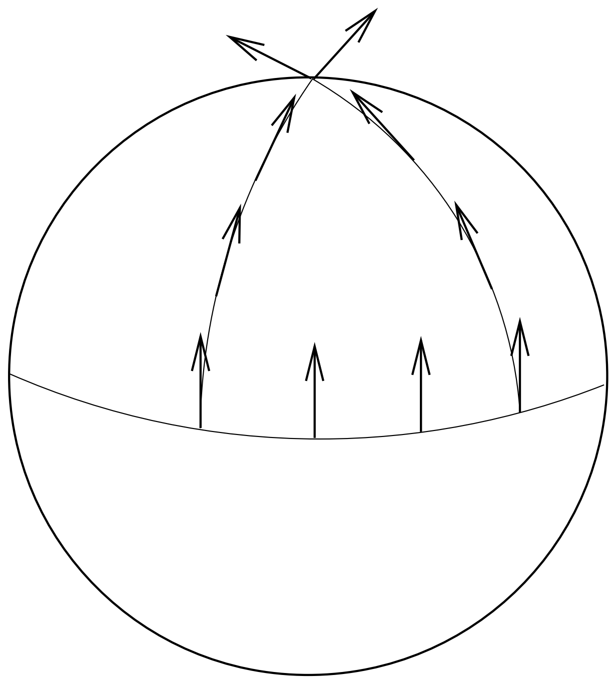

Fig. 2.— A vector, parallel-transported along different path on the surface of a sphere arrive to the same destination with two different values.

makes no sense. Of course, in certain special situations it is still useful to talk as if it did make sense, but it is necessary to understand that occasional usefulness is not a substitute for rigorous definition !. In cosmology, for example, the light from distant galaxies is redshifted with respect to the frequencies we would observe from a nearby stationary source. Since this phenomenon bears such a close resemblance to the conventional Doppler effect due to relative motion, it is very tempting to say that the galaxies are "receding away from us" at a speed defined by their redshift. At a rigorous level this is nonsense: the galaxies are not receding, since the notion of their velocity with respect to us is not well-defined. What is actually happening is that the metric of space-time between us and the galaxies has changed (the universe has expanded) along the path of the photon from here to there, leading to an increase in the wavelength of the light. As an example of how you can go wrong, naive application of the Doppler formula to the redshift of galaxies implies that some of them are receding faster than light, in apparent contradiction with relativity. The resolution of this apparent paradox is simply that the very notion of their recession should not be taken literally.

Lets get to the math now. Parallel transport is supposed to be the curved-space generalization of the concept of "keeping the vector constant" as we move it along a path; similarly for a tensor of arbitrary rank. Given a curve $x^\mu(\lambda)$, the requirement of constancy of a tensor T along this curve in flat space is simply $\frac{dT}{d\lambda} = \frac{dx^\mu}{d\lambda}\frac{\partial T}{\partial x^\mu} = 0$. We therefore define the covariant derivative along the path to be given by an operator

$$\frac{D}{D\lambda} = \frac{dx^\mu}{d\lambda}\nabla_\mu \ . \tag{1}$$

We then **define parallel transport** of the vector V along the path $x^\mu(\lambda)$ to be the requirement that, along the path,

$$\left(\frac{DV}{D\lambda}\right)^\mu = \frac{dx^\sigma}{d\lambda}\nabla_\sigma V^\mu \equiv \frac{dx^\sigma}{d\lambda}V^\mu{}_{;\sigma} = \frac{dV^\mu}{d\lambda} + \Gamma^\mu_{\sigma\rho}\frac{dx^\sigma}{d\lambda}V^\rho = 0 \ . \tag{2}$$

Similarly, the parallel transport of a tensor T along the path $x^\mu(\lambda)$ is the requirement that along the path,

$$\left(\tfrac{D}{D\lambda}T\right)^{\mu_1\mu_2\cdots\mu_k}{}_{\nu_1\nu_2\cdots\nu_l} \equiv \tfrac{dx^\sigma}{d\lambda}\nabla_\sigma T^{\mu_1\mu_2\cdots\mu_k}{}_{\nu_1\nu_2\cdots\nu_l} = 0 \ . \tag{3}$$

For example,

$$\left(\frac{D}{D\lambda}T\right)^\mu{}_\nu = \frac{dT^\mu{}_\nu}{d\lambda} + \Gamma^\mu_{\sigma\rho}\frac{dx^\sigma}{d\lambda}T^\rho{}_\nu - \Gamma^\rho_{\sigma\nu}\frac{dx^\sigma}{d\lambda}T^\mu{}_\rho \tag{4}$$

Equations 2, 3 are known as the **equation of parallel transport**. We can look at the parallel transport equation as a first-order differential equation defining an initial-value problem: given a tensor at some point along the path, there will be a unique continuation of the tensor to other points along the path such that the continuation solves Equation 3. We say that such a tensor is parallel transported.

Not surprisingly, parallel transport equation is closely related to the geodesic equation. As we have already discussed, we can **define** a geodesic as a path which parallel transports its own tangent vector. The tangent vector to the path $x^\mu(\lambda)$ is $dx^\mu/d\lambda$, and when putting this vector in the parallel transport Equation (2), it becomes the geodesic equation.

3. Curvature

Finally, we are ready to discuss the curvature of space time. As we will shortly show, the curvature is quantified by the **Riemann tensor**, which is derived from the affine connection. The basic idea is that the entire information about the intrinsic curvature of a space is given in the **metric** from which we derive the **affine connection**. We "see" the curvature by parallel transporting a vector over an infinitesimal, closed loop, and comparing the initial and final values of the vector.

For example, in flat (Minkowski) space time, $g_{\mu\nu} = \eta_{\mu\nu}$, all the derivatives of $\eta_{\mu\nu}$, and all the components of the affine connection are 0, and when parallel transporting a vector in a closed loop it is not changed- this is what let us determine that the metric is "flat".

Let us now consider the more general curved space time. We have already argued, using the two-sphere as an example, that parallel transport of a vector around a closed loop in a curved space will lead to a transformation of the vector. The resulting transformation depends on the total curvature enclosed by the loop; from this reason, we choose to work with infinitesimally small loops. One conventional way to introduce the Riemann tensor, therefore, is to consider parallel transport around an infinitesimal loop.

While we will take a different (quicker) path, it is easy to demonstrate the idea. Imagine that we parallel transport a vector V^σ around a closed loop defined by two vectors A^μ and B^ν:

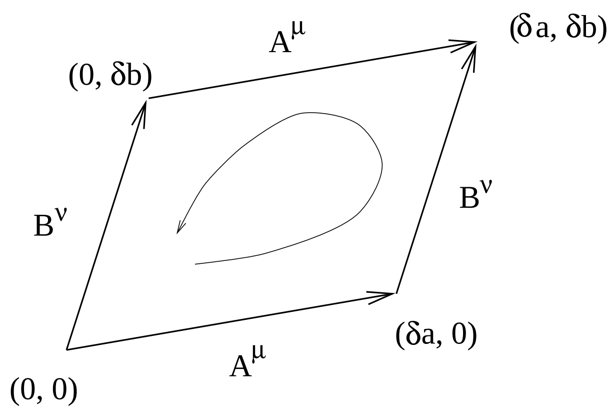

Fig. 3.— Parallel transport of a vector around an infinitesimal closed loop.

The (infinitesimal) lengths of the sides of the loop are δa and δb, respectively. Now, we use the fact that the action of parallel transport is independent of coordinates. Thus, there should be some tensor which tells us how the vector changes when it comes back to its starting point; it will be a linear transformation on a vector, and therefore involve one upper and one lower index. In addition, it will also depend on the two vectors A and B which define the loop; therefore there should be two additional lower indices to contract with A^μ and B^ν.

Furthermore, this tensor should be **antisymmetric** in these two indices, since interchanging the vectors corresponds to traversing the loop in the opposite direction, and should give the inverse of the original answer. (This is consistent with the fact that the transformation should vanish if A and B are the same vector.) We therefore expect that the expression for the change δV^ρ experienced by this vector when parallel transported around the loop should be of the form

$$\delta V^\rho = (\delta a)(\delta b) A^\mu B^\nu R^\rho{}_{\sigma\mu\nu} V^\sigma \ , \tag{5}$$

where $R^\rho{}_{\sigma\mu\nu}$ is a $(1,3)$ tensor known as the **Riemann tensor** (or simply "curvature tensor"). It is antisymmetric in the last two indices:

$$R^\rho{}_{\sigma\mu\nu} = -R^\rho{}_{\sigma\nu\mu} \ . \tag{6}$$

Generally, we can obtain the curvature tensor by performing parallel transport of the vector V^σ and look at the obtained result; we would get the curvature tensor as a function of the connection coefficients. Instead, we choose a quicker path: we look at the commutator of two covariant derivatives.

The relationship between the commutator of two covariant derivatives and parallel transport around a loop should be evident; the covariant derivative of a vector in a certain direction measures how much the vector changes relative to what it would have been if it had been parallel transported (since the covariant derivative of a vector in a direction along which it is parallel transported is zero; recall the discussion in the previous chapter, section 7 (Figure 4)). The commutator of two covariant derivatives, then, measures the difference between parallel transporting the tensor first one way and then the other, versus the opposite ordering.

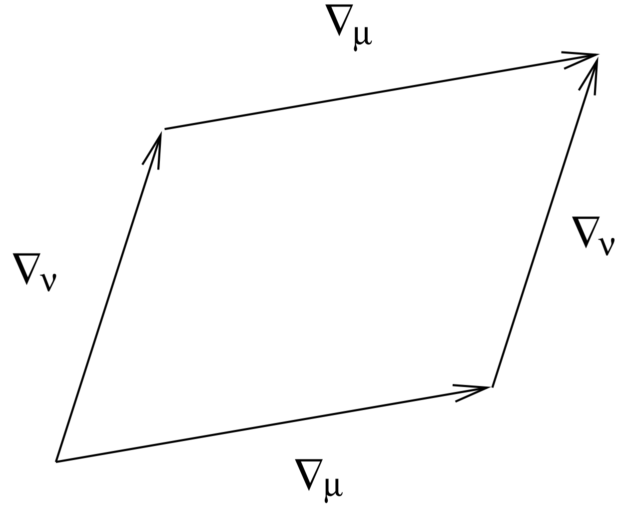

Fig. 4.— Commutator of a covariant derivative.

The calculation is really straightforward. Consider a vector field V^ρ. Then, (recalling that $V^\rho{}_{;\nu}$ is a (1,1) tensor)

$$\begin{aligned}
\nabla_\mu(\nabla_\nu V^\rho) &= (V^\rho{}_{;\nu})_{;\mu} \\
&= \nabla_\mu\left(\frac{\partial V^\rho}{\partial x^\nu} + \Gamma^\rho_{\nu\sigma}V^\sigma\right) \\
&= \frac{\partial}{\partial x^\mu}\left(\frac{\partial V^\rho}{\partial x^\nu} + \Gamma^\rho_{\nu\sigma}V^\sigma\right) + \Gamma^\rho_{\mu\sigma}\left(\frac{\partial V^\sigma}{\partial x^\nu} + \Gamma^\sigma_{\nu\lambda}V^\lambda\right) - \Gamma^\lambda_{\mu\nu}\left(\frac{\partial V^\rho}{\partial x^\lambda} + \Gamma^\rho_{\lambda\sigma}V^\sigma\right) \\
&= \partial_\mu\partial_\nu V^\rho + (\partial_\mu\Gamma^\rho_{\nu\sigma})V^\sigma + \Gamma^\rho_{\nu\sigma}\partial_\mu V^\sigma + \Gamma^\rho_{\mu\sigma}\partial_\nu V^\sigma + \Gamma^\rho_{\mu\sigma}\Gamma^\sigma_{\nu\lambda}V^\lambda - \Gamma^\lambda_{\mu\nu}\partial_\lambda V^\rho - \Gamma^\lambda_{\mu\nu}\Gamma^\rho_{\lambda\sigma}V^\sigma
\end{aligned} \tag{7}$$

We get the commutator by subtracting the same with the indices $\mu \leftrightarrow \nu$ changed. The first, sixth and seventh term immediately cancel ($\partial_\mu\partial_\nu V^\rho - \partial_\nu\partial_\mu V^\rho = 0$, $-\Gamma^\lambda_{\mu\nu}\partial_\lambda V^\rho + \Gamma^\lambda_{\nu\mu}\partial_\lambda V^\rho = 0$, etc.). So do the third and fourth term, and we are left with

$$[\nabla_\mu, \nabla_\nu]V^\rho = (\partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma})V^\sigma. \tag{8}$$

The expression in the left is clearly a tensor, and so the expression in parentheses must be a tensor itself. We thus write

$$[\nabla_\mu, \nabla_\nu]V^\rho = R^\rho{}_{\sigma\mu\nu}V^\sigma \tag{9}$$

where the **Riemann tensor** is identified as

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}\ . \tag{10}$$

A few points to note:

- We have not really demonstrated that the definition of the curvature tensor in Equation 10 is actually the same tensor that appeared in Equation 5, but in fact it's true (try to convince yourself, or see the book by Wald).
- You may find it surprising that the commutator $[\nabla_\mu, \nabla_\nu]$, which appears to be a differential operator, has an action on vector fields which is a simple multiplicative transformation. However, this is exactly what the Riemann tensor measures - the part of the commutator of covariant derivatives which is proportional to the vector field.
- The right hand side in Equation 10 is constructed from non-tensorial elements; you can check that the transformation laws all work out to make this particular combination a legitimate tensor.
- The antisymmetry of $R^\rho{}_{\sigma\mu\nu}$ in its last two indices is immediate from this formula and its derivation.
- The curvature tensor is constructed from the affine connection and its first derivatives. The affine connection itself is constructed from the metric tensor and its first derivatives. Thus, the Riemann tensor is constructed from the metric, its first and second derivatives.

- Using what are by now our usual methods, the action of $[\nabla_\rho, \nabla_\sigma]$ can be computed on a tensor of arbitrary rank. The answer is

$$\begin{aligned}[\nabla_\rho, \nabla_\sigma] X^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} &= R^{\mu_1}{}_{\lambda\rho\sigma} X^{\lambda\mu_2 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} + R^{\mu_2}{}_{\lambda\rho\sigma} X^{\mu_1 \lambda \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} + \cdots \\ & \quad - R^{\lambda}{}_{\nu_1\rho\sigma} X^{\mu_1 \cdots \mu_k}{}_{\lambda\nu_2 \cdots \nu_l} - R^{\lambda}{}_{\nu_2\rho\sigma} X^{\mu_1 \cdots \mu_k}{}_{\nu_1 \lambda \cdots \nu_l} - \cdots \ .\end{aligned} \tag{11}$$

  In particular, the commutation of a covariant vector is

$$[\nabla_\rho, \nabla_\sigma] V_\mu = -R^{\lambda}{}_{\mu\rho\sigma} V_\lambda \tag{12}$$

  Equations 9, 11 and 12 imply that if the curvature tensor vanishes, the covariant derivatives **commute** - as we would expect for a coordinate system that can be transformed into a Minkowski coordinate system.

4. Riemann Tensor and the curvature of space

As we have seen, both the affine connection and the curvature tensor are derived from the metric. This fact allows us to finally properly define what we mean by "flat" spaces, or spaces for which the metric looks Euclidean or Minkowskian. For this we have the following theorem: **The necessary and sufficient conditions for the metric $g_{\mu\nu}(x)$ to be equivalent to the Minkowski metric $\eta_{\mu\nu}$ (in the sense that we can find a transformation $x \to \xi$) are (1) that the components of the curvature tensor vanish everywhere; and (2) that the metric $g_{\mu\nu}$ has three positive and one negative eigenvalues.**

Showing it is a *necessary* condition is easy. If we are in some coordinate system such that $\partial_\sigma g_{\mu\nu} = 0$ (everywhere, not just at a point), then $\Gamma^\rho_{\mu\nu} = 0$ and $\partial_\sigma \Gamma^\rho_{\mu\nu} = 0$; thus $R^\rho{}_{\sigma\mu\nu} = 0$ by Equation 10. But this is a tensor equation, and if it is true in one coordinate system it must be true in any coordinate system. Therefore, the statement that the Riemann tensor vanishes is a necessary condition for it to be possible to find coordinates in which the components of $g_{\mu\nu}$ are constant everywhere.

Showing it is *sufficient* condition requires some work. We start by looking at a point p, for which we can choose (locally) coordinate system such that $g_{\mu\nu} = \eta_{\mu\nu}$ at p. Denote the basis vectors at p by $\hat{e}_{(\mu)}$, with components $\hat{e}^\sigma_{(\mu)}$. Then by construction we have

$$g_{\sigma\rho} \hat{e}^\sigma_{(\mu)} \hat{e}^\rho_{(\nu)}(p) = \eta_{\mu\nu} \ . \tag{13}$$

Next we parallel transport the entire set of basis vectors from p to another point q; the vanishing of the Riemann tensor ensures that the result will be independent of the path taken between p and q. Since parallel transport with respect to a metric compatible connection preserves inner products, we must have

$$g_{\sigma\rho}\hat{e}^{\sigma}_{(\mu)}\hat{e}^{\rho}_{(\nu)}(q) = \eta_{\mu\nu} \ . \tag{14}$$

We therefore have specified a set of vector fields which everywhere define a basis in which the metric components are constant. What we need to show is that this is a coordinate basis - which can be done only if the curvature vanishes.

While we do know that if the $\hat{e}_{(\mu)}$'s are a coordinate basis, their commutator will vanish:

$$[\hat{e}_{(\mu)}, \hat{e}_{(\nu)}] = 0 \tag{15}$$

what we need is the converse: that if the commutator vanishes we can find coordinates y^μ such that $\hat{e}_{(\mu)} = \frac{\partial}{\partial y^\mu}$. This is a true result, known as **Frobenius's Theorem** in differential topology. We will not prove it here (see Schutz's *Geometrical Methods* book for proof). The commutator of the vector fields $\hat{e}_{(\mu)} = \frac{\partial}{\partial y^\mu}$ is

$$[\hat{e}_{(\mu)}, \hat{e}_{(\nu)}] = \nabla_{\hat{e}_{(\mu)}}\hat{e}_{(\nu)} - \nabla_{\hat{e}_{(\nu)}}\hat{e}_{(\mu)}. \tag{16}$$

The covariant derivatives vanish, given the method by which we constructed our vector fields; they were made by parallel transporting along arbitrary paths. If the fields are parallel transported along arbitrary paths, they are certainly parallel transported along the vectors $\hat{e}_{(\mu)}$, and therefore their covariant derivatives in the direction of these vectors will vanish. The commutator in Equation 16 thus vanishes, implying that we can find a coordinate system y^μ for which these vector fields are the partial derivatives. In this coordinate system the metric will have components $\eta_{\mu\nu}$, as desired.

5. Algebraic properties of the Riemann tensor

In n-dimensional space, one could naively claim that the Riemann tensor, having four indices, have n^4 independent components. The anti-symmetry properties of the last two indices (Equation 6) implies that these last two indices can have only $n(n-1)/2$ independent values.

There are a number of other symmetries that reduce the independent components further. Let's consider these now.

The simplest way to derive these additional symmetries is to examine the Riemann tensor with all lower indices,

$$R_{\rho\sigma\mu\nu} = g_{\rho\lambda}R^{\lambda}{}_{\sigma\mu\nu} \ . \tag{17}$$

Using the definitions of the Riemann tensor and the affine connection, we have

$$\begin{aligned} R_{\rho\sigma\mu\nu} &= g_{\rho\lambda}(\partial_\mu \Gamma^\lambda_{\nu\sigma} - \partial_\nu \Gamma^\lambda_{\mu\sigma}) + g_{\rho\lambda}(\Gamma^\lambda_{\mu\alpha}\Gamma^\alpha_{\sigma\nu} - \Gamma^\lambda_{\nu\alpha}\Gamma^\alpha_{\sigma\mu}) \\ &= \tfrac{1}{2} g_{\rho\lambda} g^{\lambda\tau}(\partial_\mu\partial_\nu g_{\sigma\tau} + \partial_\mu\partial_\sigma g_{\tau\nu} - \partial_\mu\partial_\tau g_{\nu\sigma} - \partial_\nu\partial_\mu g_{\sigma\tau} - \partial_\nu\partial_\sigma g_{\tau\mu} + \partial_\nu\partial_\tau g_{\mu\sigma}) \\ &\quad + \tfrac{1}{2} g_{\rho\lambda}\partial_\mu g^{\lambda\tau}(\partial_\nu g_{\sigma\tau} + \partial_\sigma g_{\tau\nu} - \partial_\tau g_{\nu\sigma}) - \tfrac{1}{2} g_{\rho\lambda}\partial_\nu g^{\lambda\tau}(\partial_\mu g_{\sigma\tau} + \partial_\sigma g_{\tau\mu} - \partial_\tau g_{\mu\sigma}) \\ &\quad + g_{\rho\lambda}(\Gamma^\lambda_{\mu\alpha}\Gamma^\alpha_{\sigma\nu} - \Gamma^\lambda_{\nu\alpha}\Gamma^\alpha_{\sigma\mu}) \end{aligned} \tag{18}$$

where in the 2^{nd} and 3^{rd} lines we expanded the affine connection: in the 2^{nd} line we took the derivative of $\partial_\mu g_{\alpha\beta}$... and in the 3^{rd} line the derivative of the inverse metric.

Using the fact that partial derivatives commute, the first line on the right hand side of equation 18 becomes

$$\frac{1}{2}(\partial_\mu\partial_\sigma g_{\rho\nu} - \partial_\mu\partial_\rho g_{\nu\sigma} - \partial_\nu\partial_\sigma g_{\rho\mu} + \partial_\nu\partial_\rho g_{\mu\sigma}).$$

We can use the relation

$$g_{\rho\lambda}\partial_\mu g^{\lambda\tau} = -g^{\lambda\tau}\partial_\mu g_{\rho\lambda} = -g^{\lambda\tau}(\Gamma^\eta_{\mu\rho} g_{\eta\lambda} + \Gamma^\eta_{\mu\lambda} g_{\eta\rho}) \tag{19}$$

to write the 2^{nd} and 3^{rd} lines in equation 18 as

$$-(\Gamma^\eta_{\mu\rho} g_{\eta\lambda} + \Gamma^\eta_{\mu\lambda} g_{\eta\rho})\Gamma^\lambda_{\nu\sigma} + (\Gamma^\eta_{\nu\rho} g_{\eta\lambda} + \Gamma^\eta_{\nu\lambda} g_{\eta\rho})\Gamma^\lambda_{\mu\sigma} + g_{\rho\lambda}(\Gamma^\lambda_{\mu\alpha}\Gamma^\alpha_{\sigma\nu} - \Gamma^\lambda_{\nu\alpha}\Gamma^\alpha_{\sigma\mu}).$$

Most of the $\Gamma\Gamma$ terms cancel (2^{nd} and 5^{th}, 4^{th} and 6^{th}), leaving us with

$$R_{\rho\sigma\mu\nu} = \frac{1}{2}(\partial_\mu\partial_\sigma g_{\rho\nu} - \partial_\mu\partial_\rho g_{\nu\sigma} - \partial_\nu\partial_\sigma g_{\rho\mu} + \partial_\nu\partial_\rho g_{\mu\sigma}) + g_{\eta\lambda}(-\Gamma^\eta_{\mu\rho}\Gamma^\lambda_{\sigma\nu} + \Gamma^\eta_{\nu\rho}\Gamma^\lambda_{\mu\sigma}) \tag{20}$$

(hopefully not too many typos on the way).

From Equation 20, we find the following algebraic properties of $R_{\rho\sigma\mu\nu}$:

1. Symmetry under interchange of the first pair of indices with the second pair:

$$R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}\ . \tag{21}$$

2. Anti-symmetric

$$R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu} = -R_{\rho\sigma\nu\mu} = R_{\sigma\rho\nu\mu} \tag{22}$$

3. Cyclicity: the sum of cyclic permutations of the last three indices vanishes:

$$R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} = 0\ . \tag{23}$$

It is frequently useful to consider contractions of the Riemann tensor. Even without the metric, we can form a contraction known as the **Ricci tensor**:

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu} \ . \tag{24}$$

In fact, the Ricci tensor is the **only** independent contraction (modulo conventions for the sign, which of course change from place to place) of the curvature tensor formed from the affine connection[2]. The Ricci tensor is symmetric,

$$R_{\mu\nu} = R_{\nu\mu} \ , \tag{25}$$

as a consequence of the various symmetries of the Riemann tensor. Using the metric, we can take a further contraction to form the **Ricci scalar**:

$$R = R^{\mu}{}_{\mu} = g^{\mu\nu} R_{\mu\nu} \ . \tag{26}$$

6. Riemann Tensor in n-dimensions

The symmetries considered above allow us to calculate the number of independent components of the Riemann tensor. Let's begin with the facts that $R_{\rho\sigma\mu\nu}$ is antisymmetric in the first two indices, antisymmetric in the last two indices, and symmetric under interchange of these two pairs. This means that we can think of it as a symmetric matrix $R_{[\rho\sigma][\mu\nu]}$, where the pairs $\rho\sigma$ and $\mu\nu$ are thought of as individual indices. An $m \times m$ symmetric matrix has $m(m+1)/2$ independent components, while an $n \times n$ antisymmetric matrix has $n(n-1)/2$ independent components. We therefore have

$$\frac{1}{2}\left[\frac{1}{2}n(n-1)\right]\left[\frac{1}{2}n(n-1)+1\right] = \frac{1}{8}(n^4 - 2n^3 + 3n^2 - 2n) \tag{27}$$

independent components. We still have the cyclicity (Equation 23), which implies that the totally antisymmetric part of the Riemann tensor vanishes,

$$R_{[\rho\sigma\mu\nu]} = 0 \ . \tag{28}$$

This equation thus adds $n(n-1)(n-2)(n-3)/4!$ further constraints[3] leaving us with

$$\frac{1}{8}(n^4 - 2n^3 + 3n^2 - 2n) - \frac{1}{24}n(n-1)(n-2)(n-3) = \frac{1}{12}n^2(n^2-1) \tag{29}$$

---

[2]Generally, we could form the Riemann tensor not from the metric, but we will not be interested in that here.

[3]The first index can have n options; the second $(n-1)$ options, since for anti-symmetric tensor, if the two indices are the smae $\nu = \mu$, it must be 0, and so on. However, since the order of which we select the indices is unimportant, since exchange of any 2 indices gives the same result up to a sign, hence the division by 4! to avoid double counting.

independent components of the Riemann tensor.

In four dimensions, therefore, the Riemann tensor has 20 independent components. (In one dimension it has none.) These twenty functions are precisely the 20 degrees of freedom in the second derivatives of the metric which we could not set to zero by a clever choice of coordinates. This should reinforce your confidence that the Riemann tensor is an appropriate measure of curvature.

7. The Bianchi identities

In addition to the algebraic identities discussed above, the curvature tensor obeys an important differential identity. Consider the covariant derivative of the Riemann tensor, evaluated in locally inertial coordinate system in which $\Gamma^\lambda_{\mu\nu}$ vanishes (but not its derivatives !)

$$\begin{aligned} \nabla_\lambda R_{\rho\sigma\mu\nu} &= \partial_\lambda R_{\rho\sigma\mu\nu} \\ &= \tfrac{1}{2}\partial_\lambda(\partial_\mu\partial_\sigma g_{\rho\nu} - \partial_\mu\partial_\rho g_{\nu\sigma} - \partial_\nu\partial_\sigma g_{\rho\mu} + \partial_\nu\partial_\rho g_{\mu\sigma}) \ . \end{aligned} \tag{30}$$

The sum of cyclic permutations of the first three indices vanish:

$$\begin{aligned} &\nabla_\lambda R_{\rho\sigma\mu\nu} + \nabla_\rho R_{\sigma\lambda\mu\nu} + \nabla_\sigma R_{\lambda\rho\mu\nu} \\ &= \tfrac{1}{2}(\partial_\lambda\partial_\mu\partial_\sigma g_{\rho\nu} - \partial_\lambda\partial_\mu\partial_\rho g_{\nu\sigma} - \partial_\lambda\partial_\nu\partial_\sigma g_{\rho\mu} + \partial_\lambda\partial_\nu\partial_\rho g_{\mu\sigma} \\ &\quad +\partial_\rho\partial_\mu\partial_\lambda g_{\sigma\nu} - \partial_\rho\partial_\mu\partial_\sigma g_{\nu\lambda} - \partial_\rho\partial_\nu\partial_\lambda g_{\sigma\mu} + \partial_\rho\partial_\nu\partial_\sigma g_{\mu\lambda} \\ &\quad +\partial_\sigma\partial_\mu\partial_\rho g_{\lambda\nu} - \partial_\sigma\partial_\mu\partial_\lambda g_{\nu\rho} - \partial_\sigma\partial_\nu\partial_\rho g_{\lambda\mu} + \partial_\sigma\partial_\nu\partial_\lambda g_{\mu\rho}) \\ &= 0 \ . \end{aligned} \tag{31}$$

Since this is an equation between tensors it is true in any coordinate system, even though we derived it in a particular one. We recognize by now that the antisymmetry $R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}$ allows us to write this result as

$$\nabla_{[\lambda} R_{\rho\sigma]\mu\nu} = 0 \ . \tag{32}$$

This is known as the **Bianchi identity**.

An especially useful form of the Bianchi identity comes from contracting twice on Equation 31:

$$\begin{aligned} 0 &= g^{\nu\sigma} g^{\mu\lambda}(\nabla_\lambda R_{\rho\sigma\mu\nu} + \nabla_\rho R_{\sigma\lambda\mu\nu} + \nabla_\sigma R_{\lambda\rho\mu\nu}) \\ &= \nabla^\mu R_{\rho\mu} - \nabla_\rho R + \nabla^\nu R_{\rho\nu} \ , \end{aligned} \tag{33}$$

or

$$\nabla^\mu R_{\rho\mu} = \frac{1}{2}\nabla_\rho R \ . \tag{34}$$

(Notice that, unlike the partial derivative, it makes sense to raise an index on the covariant derivative, due to metric compatibility, $\nabla_\rho g_{\mu\nu} = 0$.) If we define the **Einstein tensor** as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \ , \tag{35}$$

then we see that the twice-contracted Bianchi identity (Equation 34) is equivalent to

$$\nabla^\mu G_{\mu\nu} = 0 \ . \tag{36}$$

You will not find it hard to imagine that the Einstein tensor, which is symmetric due to the symmetry of the Ricci tensor and the metric, is of great importance in general relativity.

8. A few simple examples

The results of Equation 29 imply that in 1, 2, 3 and 4 dimensions there are 0, 1, 6 and 20 independent components of the curvature tensor, respectively. This means, e.g., that one-dimensional manifolds (such as S^1) are never curved; the intuition you have that tells you that a circle is curved comes from thinking of it embedded in a certain flat two-dimensional plane.

In two dimensions, the curvature has one independent component. Thus, in fact, all of the information about the curvature is contained in the single component of the Ricci scalar. Consider, for example a cylinder, $\mathbf{R} \times S^1$. We can put a metric on the cylinder whose components are constant in an appropriate coordinate system — simply unroll it and use the induced metric from the plane (see Figure 5). In this metric, the cylinder is flat. (we could choose a metric in which the cylinder is not flat, but the point is that it *can* be made flat by choosing the appropriate metric). In fact, a similar conclusion can be drawn for the torus (see Figure 6).

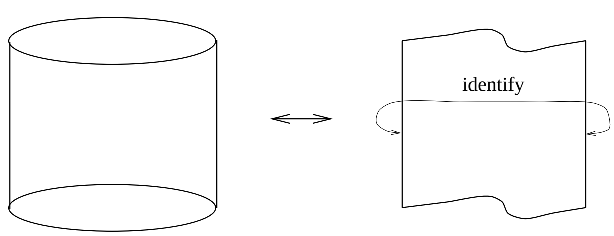

Fig. 5.— A cylinder can be made flat by “unrolling” it.

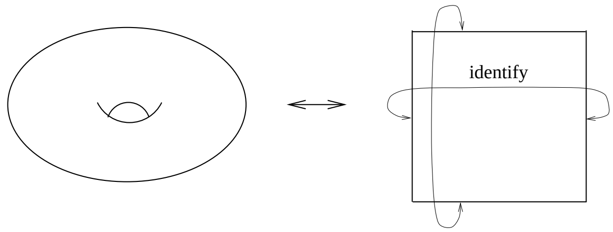

Fig. 6.— A torus can also be made flat by “unrolling” it.

We can think of the torus as a square region of the plane with opposite sides identified (in other words, $S^1 \times S^1$), from which it is clear that it can have a flat metric even though it looks curved from the embedded point of view.

Our favorite example is of course the two-sphere, with metric

$$ds^2 = a^2(\mathrm{d}\theta^2 + \sin^2\theta\ \mathrm{d}\phi^2)\ , \tag{37}$$

where a is the radius of the sphere (thought of as embedded in $\mathbf{R}^3$). Without going through the details (you were supposed to work it at home), the nonzero connection coefficients are

$$\begin{aligned} \Gamma^\theta_{\phi\phi} &= -\sin\theta\cos\theta \\ \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} &= \cot\theta\ . \end{aligned} \tag{38}$$

One component of the Riemann tensor will be:

$$\begin{aligned} R^\theta{}_{\phi\theta\phi} &= \partial_\theta\Gamma^\theta_{\phi\phi} - \partial_\phi\Gamma^\theta_{\theta\phi} + \Gamma^\theta_{\theta\lambda}\Gamma^\lambda_{\phi\phi} - \Gamma^\theta_{\phi\lambda}\Gamma^\lambda_{\theta\phi} \\ &= (\sin^2\theta - \cos^2\theta) - (0) + (0) - (-\sin\theta\cos\theta)(\cot\theta) \\ &= \sin^2\theta\ . \end{aligned} \tag{39}$$

(The notation is obviously imperfect, since the Greek letter λ is a dummy index which is summed over, while the Greek letters θ and ϕ represent specific coordinates.) Lowering an index, we have

$$\begin{aligned} R_{\theta\phi\theta\phi} &= g_{\theta\lambda}R^\lambda{}_{\phi\theta\phi} \\ &= g_{\theta\theta}R^\theta{}_{\phi\theta\phi} \\ &= a^2\sin^2\theta\ . \end{aligned} \tag{40}$$

It is easy to check that all of the components of the Riemann tensor either vanish or are related to this one by symmetry. We can go on to compute the Ricci tensor via $R_{\mu\nu} = g^{\alpha\beta}R_{\alpha\mu\beta\nu}$. We obtain

$$\begin{aligned} R_{\theta\theta} &= g^{\phi\phi}R_{\phi\theta\phi\theta} = 1 \\ R_{\theta\phi} &= R_{\phi\theta} = 0 \\ R_{\phi\phi} &= g^{\theta\theta}R_{\theta\phi\theta\phi} = \sin^2\theta\ . \end{aligned} \tag{41}$$

The Ricci scalar is similarly straightforward:

$$R = g^{\theta\theta} R_{\theta\theta} + g^{\phi\phi} R_{\phi\phi} = \frac{2}{a^2} \ . \tag{42}$$

Therefore the Ricci scalar, which for a two-dimensional manifold completely characterizes the curvature, is a constant over this two-sphere. This is a reflection of the fact that the manifold is "maximally symmetric," a concept we may define more precisely later (time permits; although it means what you think it should).

Notice that the Ricci scalar is not only constant for the two-sphere, it is manifestly positive. We say that the sphere is "positively curved". From the point of view of someone living on a manifold which is embedded in a higher-dimensional Euclidean space, if they are sitting at a point of positive curvature the space curves away from them in the same way in any direction, while in a negatively curved space it curves away in opposite directions. Negatively curved spaces are therefore saddle-like (see Figure 7).

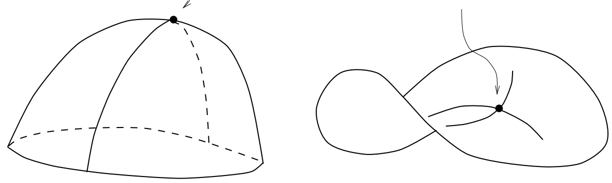

Fig. 7.— From a point of view of someone living on a manifold embedded in a higher dimension Euclidean space, positive curvature (left) implies that the space curves away in the same way in any direction, while negative curved space (right) the space curves in opposite directions.

9. Geodesic deviation

The last topic we cover before getting to Einstein's Equation is geodesic deviation. While we motivated the curvature tensor by the need to construct a field equation to the gravitational field, it turns out to be useful in expressing the effects of gravitation on physical systems.

Consider a pair of freely-falling particles that travel on trajectories $x^\mu(\tau)$ and $x^\mu(\tau) + \delta x^\mu(\tau)$. The equations of motion are the geodesic equations,

$$\begin{aligned} &\tfrac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}(x)\tfrac{dx^\rho}{d\tau}\tfrac{dx^\sigma}{d\tau} = 0, \\ &\tfrac{d^2(x^\mu+\delta x^\mu)}{d\tau^2} + \Gamma^\mu_{\rho\sigma}(x+\delta x)\tfrac{d(x^\rho+\delta x^\rho)}{d\tau}\tfrac{d(x^\sigma+\delta x^\sigma)}{d\tau} = 0. \end{aligned} \tag{43}$$

Taking the difference between these two equations, to first order in δx^μ one gets

$$\frac{d^2(\delta x^\mu)}{d\tau^2} + \frac{\partial \Gamma^\mu_{\rho\sigma}}{\partial x^\alpha}\delta x^\alpha \frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} + 2\Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{d\delta x^\sigma}{d\tau} = 0 \tag{44}$$

We can use this result to write the relative acceleration between the two nearby geodesics as follows. Using Equation 2, we write

$$\begin{aligned} A^\mu &= \frac{D^2(\delta x^\mu)}{D\tau^2} \\ &= \frac{D}{D\tau}\left(\frac{d(\delta x^\mu)}{d\tau} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}(\delta x)^\sigma\right) \\ &= \frac{d^2(\delta x^\mu)}{d\tau^2} + \frac{d}{d\tau}\left(\Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}(\delta x)^\sigma\right) + \Gamma^\mu_{\alpha\beta}\frac{dx^\beta}{d\tau}\left(\frac{d(\delta x^\alpha)}{d\tau} + \Gamma^\alpha_{\rho\sigma}\frac{dx^\rho}{d\tau}(\delta x)^\sigma\right) \\ &= \frac{d^2(\delta x^\mu)}{d\tau^2} + 2\Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{d(\delta x)^\sigma}{d\tau} + \left(\frac{d\Gamma^\mu_{\rho\sigma}}{dx^\delta}\frac{dx^\delta}{d\tau}\right)\frac{dx^\rho}{d\tau}(\delta x)^\sigma + \Gamma^\mu_{\rho\sigma}\frac{d^2x^\rho}{d\tau^2}(\delta x)^\sigma + \Gamma^\mu_{\alpha\beta}\Gamma^\alpha_{\rho\sigma}\frac{dx^\beta}{d\tau}\frac{dx^\rho}{d\tau}(\delta x)^\sigma \end{aligned} \tag{45}$$

Using now Equation 44, we get

$$A^\mu = \frac{d\Gamma^\mu_{\rho\sigma}}{dx^\delta}\frac{dx^\delta}{d\tau}\frac{dx^\rho}{d\tau}(\delta x)^\sigma - \frac{\partial \Gamma^\mu_{\rho\sigma}}{\partial x^\alpha}\delta x^\alpha\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} + \Gamma^\mu_{\rho\sigma}\frac{d^2x^\rho}{d\tau^2}(\delta x)^\sigma + \Gamma^\mu_{\alpha\beta}\Gamma^\alpha_{\rho\sigma}\frac{dx^\beta}{d\tau}\frac{dx^\rho}{d\tau}(\delta x)^\sigma \tag{46}$$

We now use the geodesic equation,

$$\frac{d^2x^\rho}{d\tau^2} = -\Gamma^\rho_{\sigma\nu}\frac{dx^\sigma}{d\tau}\frac{dx^\nu}{d\tau}$$

and make a lot of relabeling of dummy indices (first term: $\sigma \to \alpha$, $\delta \to \sigma$; third term $\rho \to \lambda$, $\sigma \to \alpha$ and $\nu \to \rho$; and 4^{th} term, $\alpha \to \lambda$, $\sigma \to \alpha$, and $\beta \to \sigma$) to write Equation 46 as

$$A^\mu = \left(\frac{d\Gamma^\mu_{\rho\alpha}}{dx^\sigma} - \frac{d\Gamma^\mu_{\rho\sigma}}{dx^\alpha} + \Gamma^\mu_{\sigma\lambda}\Gamma^\lambda_{\rho\alpha} - \Gamma^\mu_{\alpha\lambda}\Gamma^\lambda_{\rho\sigma}\right)\frac{dx^\sigma}{d\tau}\frac{dx^\rho}{d\tau}(\delta x)^\alpha \tag{47}$$

As the term in the parenthesis is simply the Riemann tensor (see Equation 10), we finally write

$$\frac{D^2(\delta x^\mu)}{D\tau^2} = R^\mu{}_{\rho\sigma\alpha}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau}\delta x^\alpha \tag{48}$$

Equation 48 is known as the **geodesic deviation equation**. It expresses something that we might have expected: the relative acceleration between two neighboring geodesics is proportional to the curvature.

Thus, although a freely falling particle appears to be at rest in a coordinate system falling with the particle, a pair of nearby freely falling particles exhibit a relative motion that can reveal the presence of a gravitational field to an observer that falls with them. This is interpreted as a manifestation of gravitational tidal forces.

REFERENCES

[1] S. Carroll, *Lecture Notes on General Relativity*, part 3. Curvature (http://preposterousuniverse.com/grnotes/).

[2] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley & Sons), chapter 6.

[3] J. Hartle, *Gravity: An Introduction to Einstein's General Relativity* (Addison-Wesley), chapter 21.