<div align="center">

PHYS 514: General Relativity Lecture Notes
Shereen Elaidi

</div>

Contents

1 Lecture 1: 7 January 2020

What is GR? GR is the modern theory of classical gravity that incorporates the effects of special relativity. The basic idea is the following: gravity is a force. Most forces are described by fields. *What are fields?* We have the following examples of fields:

(1) The Newtonian gravitational Potential, Φ.
(2) For electromagnetism, we have the electromagnetic fields E and B.

We call such a theory a **field theory**. A classical field theory has two ingredients that make up a theory of forces described by fields:

(1) **A field equation**: an equation of motion which tells us exactly how the field is determined by some set of forces. This is typically a 2nd order ODE. For example:

   (1) The field equation for Newtonian gravity is $\nabla^2 \Phi = 4\pi G\rho$.

(2) The field equations for E&M are Maxwell's equations.

(2) **Force Law:** an equation which determines how objects move in the presence of a field. For example:

   (1) For Newtonian gravity, this is $F = ma = m\nabla\Phi$.
   (2) For E&M, this is the Lorentz force law $F = q(E + v \times B)$.

The fields are functions of points in space-time. We write them as $\Phi(t, x)$ for the gravitational field or $E(t, x)$ for the electric field. The force law tells us how the objects move, i.e., how they deviate from a straight line ($a = 0$). The modern perspective of GR is the following: it is not a field theory; it is something totally different– gravity is not due to a field Φ that is a function of space-time, but is instead a feature of space-time itself. Hence, we replace the Φ with a "metric tensor" which describes the curvature / geometry of space-time.

The field equation of GR determines how space-time curves in the presence of matter or energy. This field equation is a differential equation for the metric of space-time that helps us determine how space-time curves in the presence of matter.

The force law is the geodesic equation, which tells us how objects move through curved space-time. In short, this means that objects travel on **geodesics**, which are as straight as possible given the curvature of space-time. For example, on a plane, the shortest way to travel between two points is simply a line. However, for a sphere, the curve that minimises distance is the arc of a great circle.

2 Lecture 2: 12 January 2021

Space time is a set of points at which events could take place, which can be parameterized in a smooth manner by a coordinate system, such as cartesian coordinates (t, x). We call points in space-time **events**. Physics should be independent of the coordinate system used.

In Newtonian physics, for two points (t_1, x_1) and (t_2, x_2), we can define the time-separation and spatial separations:

$$\Delta t = t_2 - t_1,$$
$$\Delta x = \sqrt{(x_2 - x_1)^2}.$$

These two quantities have independent meaning and make sense in Newtonian physics (this is called the principle of general covariance). In SR, there is no separate notion of time and space (time separation/distance) between two events. In other words, Δt and Δx do not make sense independently. These depend on the frame of reference in which they are measured. But, there is a notion of an "invariant interval" between two events. Given two events,

$$\Delta s^2 = -c^2 \Delta t^2 + \Delta x^2$$
$$= -\Delta t^2 + \Delta x^2 \text{ (units where } c = 1 \text{)}$$

When we talk about space-time, it's useful to draw pictures. We'll plot pictures in space-time using a **space-time diagram**, with time running vertically. Curves of a fixed Δs trace out hyperbolae. We have that light-rays travel on 45 degree lines. We have three cases:

(1) $\Delta s^2 = 0$: these two points can be connected by a light ray. These pairs of events are called **null-separated** or **light-like separated**.
(2) $\Delta s^2 < 0$: these points/events can be connected with a trajectory with a velocity less than c. We call such events **time-like separated**.

(3) $\Delta s^2 > 0$: these are points which cannot be connected as above. We call such events **space-like separated**.

Claim: all of special relativity can be reduced to the following statement:

For two time-like separated events, the proper time $\Delta\tau$ measured by an observer moving at a constant velocity between the events is given by:

$$(\Delta\tau)^2 = -(\Delta s)^2.$$

This reproduces various phenomena. For example, consider **time-dilation**. Using the above, we can deduce that the relative age difference between the two twins would be $\frac{2\Delta\tau}{\Delta t} = \sqrt{1 - v^2}$. We see a similar result for **length contraction**: $\frac{\Delta\tau^2}{t^2} = 1 - v^2$. In our world, we have three spatial coordinates, and so the invariant interval takes the following form:

$$\Delta s^2 = -\Delta t^2 + \Delta x_1^2 + \Delta x_2^2 + \Delta x_3^2.$$

To write this more compactly, we'll introduce a new notation. Let $x^\mu = (t, x)$ where $\mu = 0, 1, 2, 3$. Then, we can write the invariant interval as:

$$(\Delta s)^2 = \sum_{\mu,\nu=0}^{3} g_{\mu\nu}\Delta x, \tag{1}$$

where,

$$g_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

That matrix is the "metric" of flat space-time in cartesian coordinates. We can write things even more compactly by introducing **Einstein summation notation**:

$$\Delta s^2 = g_{\mu\nu}\Delta x^\mu \Delta x^\nu.$$

We read this by knowing that repeated indices are summed. Since μ and ν range from 0 to 3, we will be summing the above from 0 to 3.

You are familiar with Euclidean space $\mathbb{R}^2$ and $\mathbb{R}^3$. Minkowski space is denoted by $\mathbb{R}^{3,1}$; this is just like $\mathbb{R}^4$ except we have a negative sign in the formula for the distance. This gives us a **Lorentzian space-time**.

3   Lecture 3: 14 January 2021

3.1   Special Relativity and the Equivalence Principle

We left off last time with a reformulation of special relativity as the geometry of space-time. In particular,

The proper time $\Delta\tau$ measured by an observer moving at a constant velocity between two points in space-time is given by the invariant interval Δs:

$$(\Delta\tau)^2 = -(\Delta s)^2,$$

where $\Delta s^2 = -\Delta t^2 + (\Delta x)^2$ in Cartesian coordinates.

Recalling that $\Delta x^\mu = (\Delta t, \Delta x)$, we can write the above using the metric as:

$$\Delta s^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu,$$

where the metric is given by the following matrix:

$$\eta_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now, we want to build up to the equivalence principle. This requires us to think about observers who are accelerating (basic idea of the equivalence principle). For us, accelerating means that observers are moving through a path in space-time which is not in a straight line.

A general path through space-time is known as a **world-line**. It is parameterized by four functions. Let λ be our parameter which labels different points on the curve. Then:

$$x^\mu(\lambda) = (t(\lambda), x^1(\lambda), x^2(\lambda), x^3(\lambda)).$$

Q: How would you measure the proper time of an observer moving on this trajectory: $\Delta\tau$?
The proper time $\Delta\tau$ measured on an observer can be obtained by taking the following limiting process: break the curve down into small segments, approximate each segment by straight lines, then apply the formula from the previous lecture to proper time along a straight line. Then, take the limit as the length of each interval goes to zero. This gives us that the proper time along a path is $\sqrt{-(\Delta s)^2}$, where

$$\Delta s = \int \sqrt{-\left(\frac{dt}{d\lambda}\right)^2 + \left(\frac{dx^1}{d\lambda}\right)^2 + \left(\frac{dx^2}{d\lambda}\right)^2 + \left(\frac{dx^3}{d\lambda}\right)^2}\, d\lambda. \tag{2}$$

Observe that the above is similar to the arc-length of a curve from calculus. Using the Einstein summation notation, we can write this more compactly as:

$$\Delta s = \int \sqrt{\eta_{\mu\nu}\frac{\partial x^\mu}{\partial \lambda}\frac{\partial x^\nu}{\partial \lambda}}\, d\lambda. \tag{3}$$

We will write this as $\Delta s = \int ds$, where ds is an infinitesimal interval of space-time. We call it the **line-element of ST in GR**:

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

which we got by cancelling the $d\lambda$ in the above. Now, special relativity is the statement that "line element" is given by

$$ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu \text{ with } \eta_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

GR is a very simple modification of the above statement; GR is the statement that all the effects of gravity are packages in terms of the replacement of $\eta_{\mu\nu}$ by $g_{\mu\nu}(x)$; the components depend on the coordinates themselves. We call this the "metric of curved space-time." $\eta_{\mu\nu}$ is parameterized by 16 functions of x. Hence, to summarise, every statement of ST goes to GR with the proviso that

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu. \tag{4}$$

4

In GR, light rays still travel along null trajectories, but the null trajectory is defined as $ds = 0$ rather than $\Delta s = 0$ as previously. Similarly, time-like trajectories (trajectories which massive objects travel along) has $(ds)^2 < 0$ for every point along the trajectory. The proper time measured by an observer who travels along a wordline $x^\mu(\lambda)$ is again given by:

$$\Delta\tau^2 = -\Delta s^2, \tag{5}$$

where $\Delta s = \int ds$ where ds is given by:

$$\Delta s = \int \sqrt{g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}d\lambda. \tag{6}$$

With this geometric formulation, general relativity is a nearly trivial modification of special relativity. Just replace $\eta_{\mu\nu}$ with $g_{\mu\nu}$. We still need some dynamical principle (this will be Einstein's field equations) and we need some mathematical tools. For example, the above gives us time dilation, which then give us phenomena such as red shift in cosmology and event horizons for black holes.

3.2   Why GR is a simple modification of SR (Equivalence Principle)

The equivalence principle starts with a very simple but powerful observation about gravity. In Newtonian gravity, the way that you describe dynamics is by determining force. Let m_i be the inertial mass of an object and let m_g be the gravitational mass of an object. Let Φ be the gravitational potential. Then,

$$F = m_i a$$
$$= -m_g \nabla\Phi.$$

In principle, m_i may not be equal to m_g. But, it is an *experimental fact* that for every object that we know, $m_i = m_g$. You can cancel them out and get the first **equivalence principle**:

$$\boxed{a = -\Delta\Phi} \tag{7}$$

for all objects, independent of their mass (think of the Leaning tower of Pisa experiment or the feather dropping experiments from space). Note that it is possible that we could discover a new form of matter for which this is not true, in which case we would need to modify GR to deal with that. The above implies that in the presence of a gravitational potential, there is a preferred set of trajectories which any freely-falling object would prefer. These trajectories are the solutions of $a = -\Delta\Phi$. We call those trajectories **inertial trajectories**.

Consider a small region of space-time where $\nabla\Phi$ is approximately constant. Call that constant $-a_0$. The basic idea of the **Weak Equivalence Principle** is the following:

> In this very small region, the effects of gravity are *indistinguishable* from the effects of being in a constantly accelerating reference frame.

We can see this via a thought experiment. Consider two labs: one situated on Earth, and another attached to the end of a rocket in space which is accelerating with $a_0 = g$. The EOM of an object in the first lab is $\ddot{x} = -g$. The EOM of an object in the second lab would be $x' = x - \frac{1}{2}gt^2$, i.e., $\ddot{x}' = -g$. This means that if you were to do experiments in these labs, there is no way to tell which scenario that you are in, i.e., the laws are the same. We call forces which arise due to acceleration **fictitious forces**.

The equivalence principle is the statement that in a very small area, the effects of gravity are indistinguishable from a fictitious force due to acceleration. In this case, the metric of flat space-time (Minkowski space) will no longer be constant when written in terms of accelerating coordinates, which then gives us

that the effects of gravity can be captured by changes in the metric.

Once $\Delta\Phi$ is no longer approximately constant, i.e., when we are no longer in a region of space-time which is small, we cannot just think of gravity in terms of acceleration (fictitious forces). For example, in a very lab due to tidal forces, we can use experiments to tell if there is varying gravitational potential present; this cannot be mimicked by a constantly accelerating reference frame.

Let's conclude by reformulating the EP (Equivalence Principle) a bit:

> **Einstein Equivalence Principle**: in a sufficiently small region of space-time, the laws of physics reduce to those of SR. In other words, space-time *locally* looks like Minkowski space $\mathbb{R}^{3,1}$.

We'll spend the next few weeks unpacking the above statement. There's also a V2 of the EEP:

> **EEP'**: for any point x_0^μ, there is a coordinate system such that the line element takes the form of that in SR near x_0^μ. Mathematically:
>
> $$ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu + O(x - x_0^\mu),$$
>
> where $O(x - x_0^\mu)$ represents the extent to which gravity is not a fictitious force.

4 Lecture 4: 19 January 2020

4.1 Manifolds

Space-time locally "looks like" Minkowski space $\mathbb{R}^{3,1}$. On larger scales, gravity cannot just be regarded as a fictitious force. It's in the sense that a sphere locally looks like flat space. Precisely, this is the statement that ST is a four-dimensional pseudo-Riemannian **manifold**. We say that A D-dimensional **manifold** M is a set of points that can be labelled by coordinate systems in a smooth way.

Precisely, a coordinate system is a set of d functions x^i which label different points in some region $U \subseteq M$ in a unique way. Mathematically, this is expressed as:

$$x^i : \mathcal{U} \to \mathbb{R}^d, \qquad i = 1, ..., D.$$

The region $\mathcal{U}$ which these coordinates label uniquely is called a **patch** or **coordinate patch** of the manifold which is "covered" by the x^i. A **global coordinate system** is one which covers all of M. Generally, there are many coordinate systems which may cover some patch $\mathcal{U}$ of M.

E.g.: x^i and $x^{i'}$ are two different ways of labelling points in $\mathcal{U}$. Then, we can talk about the coordinate transformation or a change of coordinates between x^i and $x^{i'}$. These functions will be parameterized and C^∞.

$$x^i(x^{i'}) \text{ or } x^{i'}(x^i).$$

If the coordinates x^i and $x^{i'}$ uniquely, label points in some region $\mathcal{U}$, then the Jacobian matrix

$$\frac{\partial x^i}{\partial x^{i'}} = \frac{\partial x^i}{\partial x^{j'}},$$

is invertible (each index runs from $1, ...D$). This describes the components of a $D \times D$ matrix. If it's invertible, then we have that $\det\left(\frac{dx^i}{dx^j}\right) \neq 0$. If the determinant is zero, then one of the coordinate system

has become degenerate in some way in the sense that it no longer labels points in a unique way. Hence, the coordinate system becomes singular.

For example, the plane $\mathbb{R}^2$ describes a two-dimensional manifold. Two coordinate systems on this manifold are the cartesian coordinates (x, y) (global coordinates) and polar coordinates (r, θ). They are related by a chance of coordinates:

$$x^i = (x^1, x^2) = (x, y),$$
$$x^{i'} = (x^{1'}.x^{2'}) = (r, \theta).$$

Then, the coordinate transformations arse given by:

$$x^i(x^{i'}) : x = r\cos\theta, \ y = r\sin\theta$$

and the inverse transformation is given by:

$$x^{i'}(x^i) : r = \sqrt{x^2 + y^2}, \ \theta = \arctan(y/x).$$

Hence, the Jacobian matrix is given by:

$$\left(\frac{\partial x^i}{\partial x^{j'}}\right) = \begin{bmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{bmatrix}$$

We have that $\det(\mathrm{jac}) = r\cos^2\theta + r\sin^2\theta = r$, which is non-zero only if $r \neq 0$.

When describing space-time, we usually use a spatial notation for the coordinates:

$$x^\mu, \mu = 0, ..., 3.$$

Having greek indices mean that we are describing space-time; having Roman indices mean that we are describing a Euclidean manifold. For example:

$$x^\mu = (t, x) \text{ of } \mathbb{R}^{3,1} \text{ is an example of a coordinate system on a space-time.}$$

Most important sentence of this lecture: our goal is to write the laws of physics in such a way that any physical prediction made by these laws is independent of the choice of coordinate systems that we use to write the law. For example, we had the same laws in Newtonian physics for both polar and cartesian coordinates. We give this goal the following name: **Principle of General Covariance**.

E.g., in previous lectures we met the line element $ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu$. Under some change of coordinates $x^\mu \to x^{\mu'}(x^\mu)$, we have that the line element transforms by multiplication by the Jacobian matrix:

$$dx^\mu = \left(\frac{\partial x^\mu}{\partial x^\nu}\right)dx^{\nu'},$$

What this is saying is the following: if I displaced ν' by some infinitesimal amount, and asked how much μ changed, then the above equation (basically thanks to the chain rule), will answer that question for us. This becomes clear with an example:

E.g.: set $x = r\cos\theta, y = r\sin\theta$. Then,

$$dx = \cos\theta r - r\sin\theta d\theta,$$
$$dy = \sin\theta dr + r\cos\theta d\theta.$$

How does the line element transform? We apply the formula:

$$ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu$$
$$= \left(\frac{\partial x^\mu}{\partial x^{\mu'}}dx^{\mu'}\right)\left(\frac{\partial x^\nu}{\partial x^{\nu'}}dx^{\nu'}\right)\eta_{\mu\nu}$$
$$= g_{\mu'\nu'}dx^{\mu'}dx^{\nu'},$$

where

$$g_{\mu' nu'} = \eta_{\mu\nu}\frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^\nu}{\partial x^{\nu'}}.$$

In other words, the metric transforms as:

$$\eta_{\mu\nu} \to g_{\mu'\nu'} = \eta_{\mu\nu}\frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^\nu}{\partial x^{\nu'}}.$$

E.g.: the line element of flat space $\mathbb{R}^2$ is the thing that we use to measure distances:

$$ds^2 = dx^2 + dy^2.$$

In polar coordinates, this becomes:

$$ds^2 = dx^2 + dy^2$$
$$= (\cos\theta dr - r\sin\theta d\theta)^2 + (\sin\theta dr + r\cos\theta d\theta)^2$$
$$= dr^2 + r^2 d\theta$$

5   Lecture 5: 21 January 2021

5.1   Tensors

Our goal is to formulate the laws of physics such that things are independent of the coordinate systems in which we use to measure them. This goal is called the **Principle of Covariance**. This goal underlies everything that we will be discussing today. To tackle this goal, we will need to understand how quantities transform under coordinate transformations.

To that end, consider a D-dimensional manifold M. A patch M can be described by multiple different coordinate systems; say, x^μ and $x^{\mu'}$. The simples object that can transform is a **scalar/scalar function**.

A **scalar function** f is an assignment of a number to each point on a manifold M which is *independent* of the coordinates. For example:

$$f(x^\mu) \text{ in } x^\mu \text{ coordinates.}$$
$$f(x^{\mu'}) \text{ in } x^{\mu'} \text{ coordinates} = f(x^{\mu'}(x^\mu)) \text{ in } x^{\mu'} \text{ coordinates.}$$

We say that f is invariant under the change of coordinates $x^\mu \to x^{\mu'} = x^{\mu'}(x^\mu)$. A scalar function is called a **(0,0)-tensor**.

A **vector** is a set of D functions V^μ (in each coordinate system) which transforms under changes of coordinates in a particular way:

$$V^\mu(x^\mu) \text{ in } x^\mu \text{ coordinates.}$$
$$V^{\mu'}(x^{\mu'}) = \frac{\partial x^{\mu'}}{\partial x^\mu}V^\mu(x^{\mu'}(x^\mu)).$$

8

This is known as a **contravariant** vector or a $(1,0)$-tensor. This is a generalisation of the usual idea of a vector in multivariable calculus. For example, consider $\mathbb{R}^3$, and let v be a vector. We can represent v in a specific coordinate system with the following components: $v = (v^1, v^2, v^3)$. If we were to rotate the axes, then we can express the rotation with a rotation matrix R:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The components of v in this new coordinate system would accordingly change:

$$\begin{bmatrix} v^{1'} \\ v^{2'} \\ v^{3'} \end{bmatrix} = \begin{bmatrix} R_2^1 & R_2^1 & R_3^1 \\ \vdots & \ddots & \vdots \\ R_1^r & \cdots & R_3^3 \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \\ v^3 \end{bmatrix}$$

The *components* of the vector will transform by the multiplication of the transformation matrix. To see how this exactly works out, let v^i be the components of v in the unprimed coordinate frame, and let $v^{i'}$ be the components in the un-primed frame. Using index multiplication, we will discover that Einstein summation notation is basically repackaging matrix multiplication:

$$v^{i'} = R_j^{i'} v^j,$$

that is, R is simply the Jacobian:

$$R_j^{i'} = \frac{\partial x^{i'}}{\partial x^j}.$$

So, our definition of a vector is simply an object that obeys the following transformation rule: $v^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} v^\mu$; it's just a generalisation from vector calculus which accommodates more interesting and non-linear coordinate transformations. This is interesting to us if we think about Einstein's Equivalence Principle. Even tho the components of our vecttor,

$$v^\mu = \begin{bmatrix} v^0 \\ \vdots \\ v^3 \end{bmatrix}$$

will look different in different coordinate systems, a vector should be regarded as a **geometric object** with *coordinate-independent meaning*. A typical example is as follows: consider a world line $x^\mu(\lambda)$ which is moving through the manifold. The points on this curve are parameterised by some variable λ. We can then define the vector $v^\mu := \frac{\partial x^\mu}{\partial \lambda}$; this is the tangent vector to the world-line.

Why is this a vector in the sense of tensor calculus? This is a vector because of the chain rule; the chain rule says that in the new coordinates $x^{\mu'}(x^\mu)$:

$$v^{\mu'} = \frac{\partial x^{\mu'}}{\partial \lambda} = \left(\frac{\partial x^{\mu'}}{\partial x^\mu} \right) \frac{\partial x^\mu}{\partial \lambda},$$

where $\left(\frac{\partial x^{\mu'}}{\partial x^\mu} \right)$ is the Jacobian matrix. We can consider the set of all tangent vectors – the **tangent space**, denoted $T_p(M)$ – is the set of all tangent vectors attached to a point p of M. This is a vector space. A tangent vector is an example of a vector that is defined for each point on a curve. A vector attached to each point in M is called a **vector field**; in this class, if we talk about a vector, we mean a VECTOR FIELD.

We can also use our vectors to define differential operators. Given a scalar f and a vector $v^\mu(x^\mu)$, we can define the **differential operator** v by:

$$vf := v^\mu \frac{\partial}{\partial x^\mu} f = v^\mu \partial_\mu f = v^\mu f_{,\mu}.$$

This is precisely the directional derivative in the direction v. For this to be well-defined, vf must be a scalar. Indeed, we have that it is a scalar by the chain rule.

$$vf = v^{\mu'} \frac{\partial}{\partial x^{\mu'}} f = v^\mu \left(\frac{\partial x^{\mu'}}{\partial x^\mu} \right) \left(\frac{\partial x^\nu}{\partial x^{\mu'}} \right) \frac{\partial}{\partial x^\nu} f.$$

In particular, by the chain rule,

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

Since the product of the two Jacobians,

$$J = \begin{bmatrix} \frac{\partial x^{1'}}{\partial x^1} & \cdots & \frac{\partial x^{D'}}{\partial x^1} \\ \vdots & & \vdots \\ \frac{\partial x^{1'}}{\partial x^D} & \cdots & \frac{\partial x^{D'}}{\partial x^D} \end{bmatrix}$$

and

$$\hat{J} = \begin{bmatrix} \frac{\partial x^1}{\partial x^{1'}} & \cdots & \frac{\partial x^D}{\partial x^{1'}} \\ \vdots & & \vdots \\ \frac{\partial x^1}{\partial x^{D'}} & \cdots & \frac{\partial x^D}{\partial x^{D'}} \end{bmatrix}$$

gives us the identity, that means:

$$vf = v^\mu \delta^\nu_\mu \frac{\partial f}{\partial v} = v^\mu \frac{\partial}{\partial x^\mu} f.$$

This verifies that the differential operator $v = v^\mu \partial_\mu$ takes scalars to scalars. We call this the **directional derivative in the direction of** v. If we had some scalar function,

$$f(x^\mu(\lambda)) = f(\lambda),$$

then we can intepret the directional derivative as:

$$\frac{df}{d\lambda} = \frac{\partial x^\mu}{\partial \lambda} \frac{\partial f}{\partial x^\mu} = v^\mu \frac{\partial}{\partial x^\mu} f.$$

The geometric interpretation of this is as follows: this operator takes a derivative along the curve to which v is a tangent vector. Sometimes, we think of this ∂_μ as a "basis" for the set of vector in the same sense that the e_i form a standard basis of $\mathbb{R}^n$:

$$v = v^\mu \partial_\mu$$

6   Lecture 6: Tensors 26 January 2021

A tensor is a generalisation to the manifold the idea of a vector or a matrix. We already met other tensors before:

- A **scalar**, or a $(0,0)$-tensor, is a number for each point in a manifold M. It is of the form of $f(x^\mu)$.
- A **vector**, or a $(1,0)$-tensor, is a collection of D functions of the coordinates which transforms as:

$$v^\mu \to v^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} v^\mu.$$

We call the matrix $\frac{\partial x^{\mu'}}{\partial x^\mu}$ the Jacobian matrix. We can write this more familiarly as a Jacobian matrix:

$$J \cdot v,$$

where v is a column vector.

These are the two tensors which we have met so far. We can represent a vector v^μ in a *coordinate-independent way* as a differential operator which takes scalars to tensors:

$$v = v^\mu \frac{\partial}{\partial x^\mu},$$

where μ is summed over according to the Einstein summation notation.

(1) A **one-form** is a set of D functions of the coordinates which transform as:

$$\omega_\mu \to \omega_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \omega_\mu.$$

Note that $\frac{\partial x^\mu}{\partial x^{\mu'}}$ is the *inverse* of the Jacobian. A one form looks similar to a vector. It is known as a **covariant vector**, or a $(0,1)$-tensor. Note that a one-form has a LOWER index whereas a vector has an UPPER index.

This is useful because if we have a one form ω_μ and a vector v^μ, then we can combine them to make a scalar. To see why, let's **contract** them together:

$$
\begin{aligned}
v^\mu \omega_\mu \to v^{\mu'} \omega_{\mu'} &= \left(v^{\mu'} \frac{\partial x^{\mu'}}{\partial x^\mu} \right) \left(\frac{\partial x^\nu}{\partial x^{\nu'}} \omega_\nu \right) \\
&= v^\mu \left(\frac{\partial x^{\mu'}}{\partial x^\mu} \frac{\partial x^\nu}{\partial x^\mu} \right) \omega_\nu \\
&= v^\mu \delta^\nu_\mu \omega_\nu \text{ (by the chain rule)} \\
&= v^\mu \omega_\mu.
\end{aligned}
$$

This shows that a vector and a one-form get combined together to make a scalar. Formally, we can think of a one-form ω_μ as defining a map from vectors to numbers. We see one-forms referred to as **dual vectors** which live in a **dual tangent space**, $T^*_p(M)$. You've encountered one-forms before. In linear algebra, we usually represent vectors as column vectors and one-forms as row vectors.

$$\omega_\mu = \begin{bmatrix} \omega_1 & \cdots & \omega_d \end{bmatrix} \qquad v^\mu = \begin{bmatrix} v^1 \\ \vdots \\ v^d \end{bmatrix}.$$

The scalar you would have formed from these two objects is the dot product:

$$\omega_\mu v^\mu = \begin{bmatrix} \omega_1 & \cdots & \omega_d \end{bmatrix} \begin{bmatrix} v^1 \\ \vdots \\ v^d \end{bmatrix}.$$

The location of an index – namely, if it's an upper index or a lower index – tells us how an object transforms. When we set an upper and a lower index equal to each other and sum over them, we say that we are **contracting the indices**.

$$\omega_\mu v^\nu \to \omega_\mu v^\mu.$$

Now consider a scalar f. Then, we can define a one-form whose components are just the derivatives of f in a given coordinate system:

$$\frac{\partial f}{\partial x^\mu} = \partial_\mu f.$$

We can obtain the transformation rule for this by using the chain rule:

$$\partial_\mu f \to \partial_{\mu'} f = \frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu f,$$

which is indeed the multiplication rule that defines a one-form. Just as we write vectors in a basis-independent way as $v = v^\mu \partial_\mu$, we can write one-forms in a basis-independent was as $\omega = \omega_\mu dx^\mu$. We can think of the dx^μ as infinitesimal displacements of the coordinate x^μ. They also form a basis of the cotangent space, and they transform as vectors:

$$\omega = \omega_\mu dx^\mu$$
$$= \left(\omega_{\mu'} \frac{\partial x^{\mu'}}{\partial x^\mu} \right) \left(\frac{\partial x^\mu}{\partial x^{\nu'}} dx^{\nu'} \right)$$
$$= \omega_{\mu'} dx^{\mu'},$$

where the final equality follows from the chain rule. Since the notation is consistent, this proves that $\omega_\mu dx^\mu$ is a basis-independent representation of a one-form. There is a special notation for the one-form $\partial_\mu f$:

$$df = \partial_\mu f dx^\mu.$$

If, for example, $f = x^1$, then this defines, by taking derivatives, the one-form whose components are the derivative of that function:

$$\partial_\mu(x^1) = (1, 0, ..., 0),$$

which is written as dx^1. Now we are ready to introduce the general definition of a tensor.

Definition 6.1 (Tensor). A (k, l)-**tensor**, $T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \ldots \nu_l}$, is a set of D^{k+l} functions which transform under coordinate transformations in a particular way:

$$\boxed{T^{\mu'_1 \cdots \mu'_k}{}_{\nu'_1 \ldots \nu'_l} = \left(\frac{\partial x^{\mu'_1}}{\partial x^{\mu_1}} \right) \cdots \left(\frac{\partial x^{\mu'_k}}{\partial x^{\mu_k}} \right) \left(\frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}} \right) \cdots \left(\frac{\partial x^{\nu_l}}{\partial x^{\nu'_l}} \right) T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \ldots \nu_l}}$$

An easy way to remember Definition 6.1 is to remember that this is the only way to "balance the indices" in a way that respects the Einstein summation convention: upper and lower indices are summed over. We call (k, l) the **rank of the tensor**, and k is the number of **contravariant indices** and l is the number of **covariant indices**. Crucial note: the order and placement of the indices in Definition 6.1 matter!

Sometimes you'll see people writing tensors in the coordinate-invariant way by contracting it with a bunch of partial derivatives and a bunch of one-forms that are multiplied together by using some cross product as a circle – this is called a **tensor product**, but we'll leave that formal stuff to the mathematicians to deal with.

$$T = T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} \partial_{\mu_1} \otimes \dots \otimes \partial_{\mu_k} \otimes dx^{\nu_1} \otimes \dots \otimes dx^{\nu_l},$$

e.g., we can write a $(0, 2)$-tensor with components $g_{\mu\nu}$ as:

$$g = g_{\mu\nu} dx^\mu \otimes dx^\nu$$
$$= g_{\mu\nu} dx^\mu dx^\nu.$$

Now, a few comments about tensors. There are three operations that we can do on tensors:

(1) **Add** together two tensors of the same rank to get another tensor of the same rank:

$$T^{\mu_1 \dots \mu_k}{}_{1 \ \nu_1 \dots \nu_l} + T^{\mu_1 \dots \mu_k}{}_{2 \ \nu_1 \dots \nu_l} = T^{\mu_1 \dots \mu_k}{}_{3 \ \nu_1 \dots \nu_l}.$$

(2) **Multiply** a tensor with rank (k, l) with a tensor of rank (k', l') to get a tensor of rank $(k + k', l + l')$:

$$T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} S^{\mu_{k+1} \dots \mu_{k+k'}}{}_{\nu_{l+1} \dots \nu_{l+l'}} \text{ is a tensor of rank } (k + k', l + k').$$

(3) Given a tensor of rank (k, l), we can **contract** an upper and a lower index together to get a tensor of rank $(k - 1, l - 1)$:

$$T^{\mu_1 \dots \alpha \dots \mu_{k-1}}{}_{\nu_1 \dots \alpha \dots \nu_{l-1}}.$$

e.g., given a vector and a one-form, you can define a $(1, 1)$-tensor $v^\mu \omega_\nu$ with μ and ν separate indices, and $v^\mu \omega_\mu$ as a scalar (a $(0, 0)$-tensor).

7 Lecture 7: 28 January 2021

7.1 Tensors and the Geometry of Space-Time

From the operations on tensors introduced last class, we can build other operations. But first, let's define a math term. A **permutation** σ is a re-ordering of the numbers $(1, ..., l) \to (1, ..., l)$. A result from algebra tells us that σ is the product of pairwise interchanges. Using this notation, we can define two more operations on tensors:

(1) The **symmetrization** of a tensor makes it symmetric:

$$T_{(\mu_1, \dots, \mu_l)} = \frac{1}{l!} \sum_{\text{perms of } \sigma} T_{\sigma(\mu_1) \dots \sigma(\mu_l)}.$$

(2) The **anti-symmetrization** of a tensor makes it anti-symmetric:

$$T_{[\mu_1 \dots \mu_l]} = \frac{1}{l!} \sum_{\text{perms of } \sigma} \text{sgn}(\sigma) T_{\sigma(\mu_1) \dots \sigma(\mu_l)}.$$

13

For example:

$$T_{(\mu_1,\mu_2)} = \frac{1}{2}(T_{\mu_1\mu_2} + T_{\mu_2\mu_1})$$

$$T_{[\mu_1,\nu_2]} = \frac{1}{2}(T_{\mu_1\mu_2} - T_{\mu_2\mu_1}).$$

We can also do this procedure with upper indices or with only some fraction of the indices. E.g., we can consider something like this:

$$T^{[\mu_1\mu_2]}{}_{\nu_1(\nu_2\nu_3)}.$$

Definition 7.1 ((Anti)-Symmetric Tensor)**.** A **(anti)-symmetric tensor** is one which equals its (anti)-symmetrization.

- A symmetric tensor is unchanged when we permute two of its indices.
- An anti-symmetric tensor will pick up a factor that is $\text{sgn}(\sigma)$ when we interchange two of the indices. Hence, if two of the indices are equal for an anti-symmetric tensor, then the tensor is zero.

An anti-symmetric tensor of rank $(0, n)$ is a **differential form of rank** n, e.g., an anti-symmetric $(0, 2)$-tensor is a 2-form. You've already met some tensors:

$$f = (0,0) - \text{tensor.}$$
$$\partial_\mu f = (0,1) - \text{tensor.}$$
$$\frac{\partial x^\mu}{\partial \lambda} = (1,0) - \text{tensor (tangent line of a world line).}$$
$$\delta^\mu{}_\nu = (1,1) - \text{tensor.}$$
$$g_{\mu\nu} \text{ a symmetric } (0,2)\text{-tensor.}$$

Why are we studying tensors? The reason why is the Principle of General Covariance: the laws of physics should make predictions which should be *independent* of the coordinate system that we use to write those laws. Hence, the laws of physics are TENSOR equations, i.e., every law of physics better take the form of

$$T^{\mu_1\dots\mu_k}{}_{\nu_1\dots\nu_l} = S^{\mu_1\dots\mu_k}{}_{\nu_1\dots\nu_l}$$

in one reference frame, and

$$T^{\mu'_1\dots\mu'_k}{}_{\nu'_1\dots\nu'_l} = S^{\mu'_1\dots\mu'_k}{}_{\nu'_1\dots\nu'_l},$$

in another reference frame, because of the tensor transformation rule (Jacobian matrices are invertible). For example, Maxwell's equations are tensor equations.

We need to take derivatives of tensors now. We don't know how to do this in a **covariant manner**. Consider the only notion of differentiation that we have for now:

$$v^\mu \to \partial_\nu v^\mu \neq (1,1) - \text{tensor,}$$

This specifies a direction, and is not what we need. There is more than one way of formulating calculus on tensors. We'll be studying differential geometry / Riemannian geometry to do so.

Recall the Einstein Equivalence Principle: the effect of gravity in a very small region of space-time is

indistinguishable from an accelerating reference frame. Let's try to interpret this in a slightly more geometric point of view. *Space-time is a manifold with a metric $g_{\mu\nu}$ that is used to define the line-element:*

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

such that for any point x_0^μ, there is a coordinate system such that

$$g_{\mu\nu} = \eta_{\mu\nu} + \text{ terms that vanish as } x_0^\mu \to x_0^\mu,$$

i.e., we've Taylor expanded the metric about a point! From this more geometric formulation of the Einstein Equivalence Principle, we find that $g_{\mu\nu}$ is a $(0,2)$-tensor with three properties:

(1) **Symmetric**: $g_{\mu\nu} = g_{\nu\mu}$.
(2) **Invertible**: $\det(g_{\mu\nu}) \neq 0$.
(3) **diagonalizablity**: g is diagonalizable with eigenvalues $\lambda = -1, 1, 1, 1$, i.e., there exists some basis where the metric can be expressed as $\text{diag}(-1, 1, 1, 1)$.

These three properties are proven by noting that there exists some basis where $g_{\mu\nu} = \eta_{\mu\nu}$):

$$\Rightarrow g_{\mu\nu} = \left(\frac{\partial x^{\mu'}}{\partial x^\mu}\right)\left(\frac{\partial x^{\nu'}}{\partial x^\nu}\right)\eta_{\mu'\nu'} + ...,$$

i.e., in matrix notation:

$$g_{\mu\nu} = J^t \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} J.$$

The above properties are coordinate-invariant statements in the sense that if they are true in one coordinate system then they are true in every coordinate system.

An invertible symmetric $(0,2)$-tensor is a **metric**. If the eigenvalues of the matrix are all positive, then we call it a **Riemannian metric**. Otherwise, we call it a **pseudo-Riemannian metric**. The number of positive and negative eigenvalues of $g_{\mu\nu}$ is called the **signature** of the metric. Using this new vocabulary, we can now say that in general relativity, space-time is a manifold with a pseudo-Riemannian metric of signature $(-, +, +, +)$, or signature $(1, 3)$.

What can we do with this metric?
We can invert the metric. We'll denote the inverse of $g_{\mu\nu}$ as $g^{\mu\nu}$. It is a symmetric $(2,0)$-tensor. Using the Einstein summation notation, we find:

$$g^{\mu\nu}g_{\nu\rho} = \delta^\mu{}_\rho, \text{ or}$$
$$g_{\mu\nu}g^{\nu\rho} = \delta^\rho{}_\mu$$

Conclusion: the inverse of a metric is a tensor. We will use $g_{\mu\nu}$ and $g^{\mu\nu}$ to raise and lower indices via Einstein summation notation. I.e., given a vector v^μ we define a one-form $v_\mu := g_{\mu\nu}v^\nu$. Given a one-form ω^μ, we define a vector $\omega^\mu := g^{\mu\nu}\omega_\nu$. We can "lower an index" as:

$$T^{\mu\nu}{}_\rho = g_{\rho\sigma}T^{\mu\nu\sigma}.$$

Note that the order of the indices still matter. For example:

$$T^\mu{}_\rho{}^\nu = g_{\rho\sigma}T^{\mu\sigma\nu}.$$

Given a vector v^μ or a one-form ω_μ, we can define the **norm**:

$$v^2 = v_\mu v^\mu \text{ and } \omega^2 = \omega_\mu \omega^\mu.$$

These norms are NOT positive-definite: a vector is space-like, time-like, or null depending on whether v^2 is $> 0, < 0, \text{ or} = 0$.

8 Lecture 8: 2 February 2021

8.1 Geometry

The metric $g_{\mu\nu}$ determines the geometry of space-time in the sense that it determines the distance (i.e., the invariant interval) between points in space-time. It also determines the length of paths through space-time. For example, consider a world line $x^\mu(\lambda)$ connecting two points, λ_0 and λ_1. Then, we get the distance, or the interval, connecting $x^\mu(\lambda_0)$ and $x^\mu(\lambda_1)$ as measured along $x^\mu(\lambda)$.

$$
\begin{aligned}
s &= \int_{\lambda_0}^{\lambda_1} ds \\
&= \int_{\lambda_0}^{\lambda_1} \sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}\, d\lambda,
\end{aligned}
$$

where $\dot{x}^\mu = \frac{dx^\mu}{d\lambda}$ is the tangent vector to the curve. The proper time as measured by an observer on this path is $T = is$.

Definition 8.1 (Geodesic). A path which extremizes $S[x^\mu(\lambda)]$ is called a **geodesic**. The length of such a path is the geodesic between $x^\mu(\lambda_0)$ an $x^\mu(\lambda_1)$.

- If S is real, then the end-points are space-like separated.
- If S is imaginary, then the points are time-like separated.
- If S is zero, then the points are null separated.

This means that given a metric, we now have a notion of distance.

Q: What do geodesics look like?
Let's derive it:

$$
S[x^\mu(\lambda)] = \int \sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}\, d\lambda.
$$

We have that $\frac{ds}{d\lambda} = \sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}$. The above quantity is extremized under $x^\mu \to x^\mu + \delta x^\mu$ if:

$$
0 = \frac{\delta S}{\delta x^\mu} = \frac{d}{d\lambda}\left(\frac{g_{\mu\alpha}\dot{x}^\alpha}{\frac{dS}{d\lambda}}\right) - \frac{1}{2}\frac{g_{\alpha\beta,\mu}.\dot{x}^\mu\dot{x}^\nu}{\frac{dS}{d\lambda}}. \tag{8}
$$

c.f. Euler-Lagrange Equations:

$$
S[x^i(t)] = \int dt \mathcal{L}(x^i, \dot{x}^i) \Rightarrow \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right) - \frac{\partial \mathcal{L}}{\partial x^i} = 0.
$$

Let's try to unpack (8) a bit:

$$
0 = -g_{\mu\alpha}\dot{x}^\alpha \frac{\frac{d^2 s}{d\lambda}}{\left(\frac{ds}{d\lambda}\right)^2} + \frac{1}{\frac{ds}{d\lambda}}\left(g_{\mu\alpha}\ddot{x}^\alpha + (g_{\mu\alpha,\beta} - \frac{1}{2}g_{\alpha\beta,\mu})\dot{x}^\alpha\dot{x}^\beta\right).
$$

Re-writing the terms in brackets more cleverly as $\frac{1}{2}(g_{\mu\alpha,\beta} + g_{\mu\beta,\alpha} - g_{\alpha\beta,\mu})$, we obtain:

$$
0 = -g_{\mu\alpha}\dot{x}^\alpha \frac{\frac{d^2 s}{d\lambda}}{\left(\frac{ds}{d\lambda}\right)^2} + g_{\mu\alpha}\ddot{x}^\alpha + \frac{1}{2}(g_{\mu\alpha,\beta} + g_{\mu\beta,\alpha} - g_{\alpha\beta,\mu})\dot{x}^\alpha\dot{x}^\beta.
$$

We want to re-write the above equation as the second derivative of x, which involves multiplying the above by the inverse of the metric, $g^{\mu\nu}$. This gives us the **geodesic equation**:

$$\ddot{x}^\nu + \Gamma^\nu{}_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = \dot{x}^\nu \frac{d^2 s}{d\lambda^2}\left(\frac{ds}{d\lambda}\right)^{-1} \tag{9}$$

where,

$$\Gamma^\nu{}_{\alpha\beta} := \frac{1}{2}g^{\mu\nu}(g_{\alpha\mu,\beta} + g_{\beta\mu,\alpha} - g_{\alpha\beta,\mu}), \tag{10}$$

are called the **christoffel symbols**, or the connection coefficients, associated to the metric $g_{\mu\nu}$. Sometimes we write $\ddot{x}^\mu + \Gamma^\mu{}_{\nu\rho}\dot{x}^\nu\dot{x}^\rho = f(\lambda)\dot{x}^\mu$. In particular, λ is some parameter that we would use to label points along the path. We could choose a new parameter, $f'(\lambda)$, which would change the right hand side of the geodesic equation.

Since $f(\lambda) = \frac{d^2 s}{d\lambda}\left(\frac{ds}{d\lambda}\right)^{-1}$, we could set $f = 0$ by choosing to label points using the parameter s itself, i.e., choose λ such that as we travel along the path, $s(\lambda_0, \lambda) = \lambda$. This would imply that

$$\frac{ds}{d\lambda} = 1 \text{ and } \frac{d^2 s}{d\lambda} = 0.$$

Such a parameter has a name: λ like the above are known as **affine parameters**. For an affine parameter, the geodesic equation looks like:

$$\ddot{x}^\mu + \Gamma^\mu{}_{\nu\rho}\dot{x}^\nu\dot{x}^\rho = 0 \tag{11}$$

In many cases, the above will be the geodesic equation that we want to study. For a null geodesic, it is impossible to choose an affine parameter. The Christoffel symbol,

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

has various different properties:

(1) $\Gamma^\mu{}_{\nu\rho} = \Gamma^\mu{}_{\rho\nu}$ (symmetric in its lower indices).
(2) $\Gamma^\mu{}_{\nu\rho}$ is not a tensor. Instead, it transforms as:

$$\Gamma^{\mu'}{}_{\nu'\rho'} = \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^{\nu'}}\frac{\partial x^\rho}{\partial x^{\rho'}}\Gamma^\mu{}_{\nu\rho} - \frac{\partial x^\rho}{\partial x^{\rho'}}\frac{\partial x^\nu}{\partial x^{\nu'}}\frac{\partial^2 x^{\mu'}}{\partial x^\rho \partial x^\nu}$$

The correction terms arise from some non-linear coordinate change. This reflects that going into an accelerating coordinate system gives rise to fictitious forces.

8.2   Derivatives of the Metric of Spacetime

The Christoffel symbol captures in some sense how geodesics deviate from being "straight lines." with $\ddot{x}^\mu = 0$. That is, the christoffel symbol tells us how much we deviate from being on straight line due to both being in an accelerating reference frame AND due to the curvature of space-time (by the Einstein Equivalence Principle) [this is the effect of gravity]. You should think of a geodesic as a curve which is as straight as possible given the curvature of space-time. For example, the geodesics on spheres are arcs of great circles. Curvature is the mathematical way in which we describe the non-fictitiousness of the force of gravity.

We say that a geodesic is time-like, space-like, or null if S is imaginary, real, or zero. Similarly, the tangent vector to a geodesic is time-like/space-like/null if the tangent is.

9 Lecture 9: 4 February 2021

9.1 Calculus on Curved Spaces

We learned that space-time is a manifold with metric $g_{\mu\nu}$. From this metric, we can define a geometry and a distance along geodesics. The equation for a geodesic is:

$$\ddot{x}^\mu + \Gamma^\mu{}_{\nu\rho}\dot{x}^\rho\dot{x}^\nu = f(\lambda)\dot{x}^\mu.$$

The laws of physics involve derivatives. To write physics in a way that satisfies the principle of covariance, we need to take derivatives of tensors.

$$f : \text{scalar} \to \partial_\mu f = f_{,\mu}.$$

However, for a vector, $\partial_\mu v^\rho$ is NOT a $(1,1)$-tensor, even if the vector is a $(1,0)$-tensor. We want a new notion of derivatives – a **covariant derivative** – ∇_μ such that $\nabla_\mu v^\rho$ is a $(1,1)$-tensor. Such a covariant derivative should be the generalization of the derivative above. For example,

$$\nabla_\mu f := \partial_\mu f$$

Q: How will we define a covariant derivative on vectors?

(1) **Linearity:** $\nabla_\mu(v^\rho + \omega^\rho) = \nabla_\mu v^\rho + \nabla_\mu \omega^\rho$.
(2) **Leibnitz:** $\nabla_\mu(f v^\rho) = \nabla_\mu f v^\rho + f \nabla_\mu v^\rho = \partial_\mu f v^\rho + f \nabla_\mu v^\rho$.

You can convince yourself that this constrains the possible forms of the covariant derivative:

$$\nabla_\mu v^\rho = \partial_\mu v^\rho + G^\rho{}_{\mu\nu} v^\nu,$$

where G represents some linear transformation acting on a vector. For the $\nabla_\mu v^\rho$ to be a tensor, one must investigate how it transforms under coordinate transformations.

$$
\begin{aligned}
\nabla_{\mu'} v^{\rho'} &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\rho'}}{\partial x^\rho} \nabla^\rho_\mu\nu \\
&= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\rho'}}{\partial x^\rho} \left(\partial_\mu v^\rho + G^\rho{}_{\mu\nu} v^\nu \right) \\
&= \partial_{\mu'} \left(\frac{\partial x^{\rho'}}{\partial x^\rho} v^\rho \right) + G^{\rho'}{}_{\mu'\nu'} v^{\nu'} \\
&= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\rho'}}{\partial x^\rho} v^\rho{}_{,\mu} + G^{\rho'}{}_{\mu'\nu'} + \frac{\partial^2 x^{\rho'}}{\partial x^\rho \partial x^\mu} \frac{\partial x^\mu}{\partial x^{\mu'}} v^\rho.
\end{aligned}
$$

Re-arranging for $G^{\rho'}{}_{\mu'\nu'}$, we obtain:

$$
G^{\rho'}{}_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\rho'}}{\partial x^\rho} G^\rho{}_{\mu\nu} v^\nu - \frac{\partial^2 x^{\rho'}}{\partial x^\nu \partial x^\mu} \frac{\partial x^\mu}{\partial x^{\mu'}} v^\nu.
$$

Re-write the above so that everything is a tensor contracted with $v^{\nu'}$ (use that $v^\nu = \frac{\partial x^\nu}{\partial x^{\nu'}} v^{\nu'}$):

$$
G^{\rho'}{}_{\mu'\nu'} v^{\nu'} = \left(\frac{\partial x^{\rho'}}{\partial x^\rho} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial x^\mu}{\partial x^{\mu'}} G^\rho{}_{\mu\nu} - \frac{\partial^2 x^{\rho'}}{\partial x^\mu \partial x^\nu} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \right) v^{\nu'}.
$$

Since this is true for all $v^{\nu'}$, we know that G transforms under coordinate transformations as:

$$
G^{\rho'}{}_{\mu'\nu'} = \left(\frac{\partial x^{\rho'}}{\partial x^\rho} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial x^\mu}{\partial x^{\mu'}} G^\rho{}_{\mu\nu} - \frac{\partial^2 x^{\rho'}}{\partial x^\mu \partial x^\nu} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \right).
$$

Observe that this transformation rule for G has already been met before: this is exactly how the Christoffel symbol transforms. In fact, $\Gamma^\rho{}_{\mu\nu}$ is the unique object that we can form of the metric $g_{\mu\nu}$ an its first derivatives that has this transformation rule. So, we want to *define* the covariant derivative of a vector in terms of the christoffel symbols as:

$$\boxed{\nabla_\mu v^\rho := \partial_\mu v^\rho + \Gamma^\rho{}_{\mu\nu} v^\nu} \tag{12}$$

Now, let's go back and visit the geodesic equation. The geodesic equation for a path $x^\mu(\lambda)$ is:

$$v^\mu \nabla_\mu v^\nu = v^\mu \partial_\mu v^\nu + \Gamma^\rho{}_{\mu\nu} v^\mu v^\rho = f(\lambda) v^\nu$$

by definition, and so by the chain rule this gives us,

$$v^\mu \nabla_\mu v^\nu = \ddot{x}^\nu + \Gamma^\rho{}_{\mu\nu} \dot{x}^\mu \dot{x}^\rho = f(\lambda) \dot{x}^\nu.$$

This means that in terms of the covariant derivative, the geodesic equation is a particular example of the covariant derivative at work. Namely:

$$\boxed{v^\mu \nabla_\mu v^\nu = f(\lambda) v^\nu} \tag{13}$$

This gives us a geometric understanding of the geodesic equation. Namely, we can think of the quantity $v^\mu \nabla_\mu$ as the "directional derivative" along a path.

Q: What about derivatives of other tensors?

We *must* preserve the Leibniz rule as well as linearity. For example, suppose we want to differentiate a one-form. Then, using the fact that we know that $\omega_\mu v^\mu$ is a scalar, we can use its covariant derivative to find the covariant derivative of a one-form as follows:

$$\nabla_\alpha(\omega_\mu v^\nu) = \partial_\alpha(\omega_\mu v^\mu) = (\nabla_\alpha \omega_\mu) v^\mu + \omega_\mu(\nabla_\alpha v^\mu).$$

Solving for the covariant derivative of the one-form, we get the definition of a derivative of a one-form:

$$\boxed{\nabla_\mu \omega_\nu := \partial_\mu \omega - \Gamma^\rho{}_{\mu\nu} \omega_\mu.} \tag{14}$$

Noticing the pattern, we go on to define the covariant derivative in terms of Christoffel symbols, where upper indices come with a positive sign and lower indices come with a negative sign:

$$\boxed{\nabla_\alpha T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} := \partial_\alpha T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} + \Gamma^{\mu_1}{}_{\alpha\rho} T^{\rho, \mu_2, \dots \mu_k}{}_{\nu_1 \dots \nu_l} + \dots + \Gamma^\alpha{}_{\nu_1} T^{\mu_1 \dots \mu_k}{}_{\rho\nu_2 \dots \nu_l}} \tag{15}$$

This is the only possible expression that we can build out of the metric $g_{\mu\nu}$ that is consistent with linearity and the Leibniz rule. For example,

$$\nabla_\alpha(T^{\mu\nu} v_\rho) = (\nabla_\alpha T^{\mu\nu}) v_\rho + T^{\mu\nu}(\nabla_\alpha v_\rho).$$

Note. There is one important property that it does not have. Covariant derivatives DO NOT commute with each other, e.g.:

$$\nabla_\alpha(\nabla_\beta v^\rho) \neq \nabla_\beta(\nabla_\alpha v^\rho)$$

However, covariant derivatives DO commute in flat-space time. Hence, this is THE signature feature of curved space-time. To summarize: *space-time curved $\iff$ the covariant derivatives do NOT commute.*

Q: What properties does this covariant derivative have?

(1) **Metric compatible:** the covariant derivative of the metric is zero:

$$\nabla_\mu g_{\rho\sigma} = 0$$

Proof: Writing this derivative out,

$$\begin{aligned}
\nabla_\mu g_{\rho\sigma} &= \partial_\mu g_{\rho\sigma} - \Gamma^\lambda{}_{\mu\rho} g_{\lambda\sigma} - \Gamma^\lambda{}_{\mu\sigma} g_{\lambda\rho} \\
&= g_{\rho\sigma,\mu} - \frac{1}{2}\left(g_{\mu\sigma,\rho} + g_{\rho\sigma,\mu} - g_{\mu\rho,\sigma}\right) - \frac{1}{2}\left(g_{\mu\rho,\sigma} + g_{\rho\sigma,\mu} - g_{\mu\sigma,\rho}\right) \\
&= 0.
\end{aligned}$$

Similarly, you can show that $\nabla_\mu g^{\rho\sigma} = 0$ and that $\nabla_\mu \delta^\rho_\sigma = 0$ (the calculation is nearly identical). What does this mean for us? When we write down expressions with covariant derivatives, we can freely pass expressions including the metric and δ through ∇_μ; in other words, raising and lowering of indices commutes with the covariant derivative. For example:

$$v^\mu \nabla_\rho v_\mu = v_\mu \nabla_\rho v^\mu.$$

Note that in general the above is not true for partial derivatives.

Now we define the **contravariant derivative** as follows:

$$\nabla^\rho := g^{\rho\mu} \nabla_\mu, \tag{16}$$

such that

$$\nabla_\mu v^\mu = v^\mu{}_\mu,$$

i.e., it's the divergence of a vector. By definition, this is a scalar. Similarly, we can define the **Laplacian** or **Laplace Beltrami** operator to be:

$$\nabla^2 := \nabla_\mu \nabla^\mu$$

10 Lecture 10: 9 February 2021

10.1 Curvature

ST is a manifold with a metric $g_{\mu\nu}$, from this metric, we define a notion of derivative: **covariant derivative**:

$$v_\mu v^\rho = \partial_\mu v^\rho + \Gamma^\mu{}_{\rho\sigma} v^\sigma.$$

where $\Gamma^\mu{}_{\rho\sigma} = \frac{1}{2} g^{\mu\sigma}(g_{\sigma\nu,\rho} + g_{\sigma\rho,\nu} - g_{\nu\rho,\sigma})$. Recall that the christoffel symbols are not a tensor in the exact opposite way that $\partial_\mu v^\rho$ is not a tensor. A metric defines a kind of geometry.

Q: What do we mean by a curved geometry?
We can informally think of this by transporting a vector around a curve such that it keeps the same angle with the tangent line. In flat space, the vector will return the same location. If we consider this exercise in a curved space, like a sphere, and transport the vector along geodesics (great circles), then the vector does not come back to it self. The whole question of if the vector comes back to itself is how we will capture the notion of curvature.

To make this precise, we need some way of understanding how vectors change when transported around loops. But the above gives us the intuitive notion of what happens.

Definition 10.1 (Parallel Transport)**.** A vector ω^μ is **parallely transported** along a curve $x^\mu(\lambda)$ if the directional derivative along the curve is zero:

$$v^\mu \nabla_\mu \omega^\nu = 0 \text{ where } v^\mu = \dot{x}^\mu = \frac{dx^\mu}{d\lambda}.$$

We now need a tensorial notion for what it means for something to be parallely transported.

E.g.: A geodesic is a curve which has the property where the tangent vector of the curve is parallely transported to itself along the curve. We can read this off of the geodesic equation:

$$v^\mu \nabla_\mu v^\rho = f(\lambda) v^\rho.$$

Q: How do vectors change as we parallel transport them around some small loop in space-time?
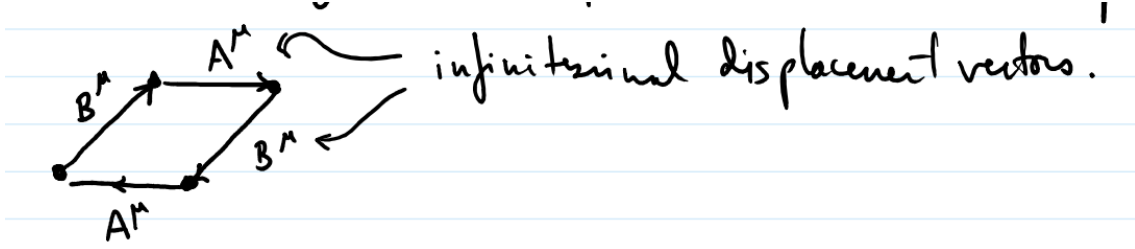
We can ask ourselves the following question: how does the same vector change under the above process, i.e., under the displacement $x^\mu \to x^\mu + B^\mu$. We would see that:

$$v \to v^\mu + B^\alpha v_\alpha v^\mu + \dots$$

If we add on the displacement, $x^\mu + B^\mu \to x^\mu + B^\mu + A^\mu$, we see that:

$$v^\mu + B^\alpha V_\alpha V^\mu \to V^\mu + B^\alpha V_\alpha V^\mu + A^\beta \nabla_\beta (V^\mu + B^\alpha \nabla_\alpha V^\mu),$$

and so on...

So if we transport around the entire loop, then V will change as:

$$\delta V^p = A^\mu A^\nu [\nabla_\mu, \nabla_\nu] V^\rho = A^\mu B^\nu (\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) V^\rho. \tag{17}$$

Note that covariant derivatives DO NOT commute. The above captures how vectors will change when you move around in small loops. This is a **characterization** of the local curvature at some point in space-time. We will define the **curvatue tensor/Riemann tensor** by:

$$\delta V^\rho = R^\rho{}_{\sigma\mu\nu} A^\mu B^\nu V^\sigma. \tag{18}$$

This is a $(1,3)$-tensor, which by definition is going to be anti-symmetric in its last two indices thanks to the commutator. A and B are arbitrary vectors that we use to measure curvature in different directions. Let's compute $R^\rho{}_{\sigma\mu\nu}$:

$$[\nabla_\mu, \nabla_\nu] v^\rho = \nabla_\mu (\nabla_\nu v^\rho) - \nabla_\nu (\nabla_\mu v^\rho)$$
$$= \partial_\mu (\nabla_\nu v^\rho) - \Gamma^\lambda{}_{\mu\nu} \nabla_\lambda v^\rho + \Gamma^\lambda{}_{\mu\nu} + \Gamma^\rho{}_{\mu\sigma} \nabla_\nu v^\sigma - (\mu \leftrightarrow \nu)$$

Observe that the Christoffel symbols by definition are symmetric in its lower indices. We then obtain:

$$= \partial_\mu (\partial_\nu v^\rho) - \partial_\mu (\Gamma^\mu{}_{\nu\sigma} v^\sigma) + \Gamma^\rho{}_{\mu\sigma} \partial_\nu v^\sigma + \Gamma^\rho{}_{\mu\sigma} \Gamma^\sigma{}_{\nu\lambda} v^\lambda - (\mu \leftrightarrow \nu)$$

By the equality of the mixed partials, the first two terms will cancel. We get, after re-labelling the indices $\sigma \leftrightarrow \lambda$:

$$= v^\sigma (\partial_\mu \Gamma^\partial_{\nu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\lambda} - (\mu \leftrightarrow \nu))$$
$$= R^\rho_{\sigma\mu\nu} v^\sigma.$$

So, to summarize:

$$\boxed{R^\rho_{\sigma\mu\nu} := \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}} \tag{19}$$

Now we need to unpack and understand this expression for curvature. First we'll state some facts about the curvature tensor:

(1) $R^\rho_{\sigma\mu\nu}$ is a tensor even though Γ is not.
(2) If $R^\rho_{\sigma\mu\nu} = 0$ then we say that the space-time is **flat.**
(3) **Theorem.** space-time is flat $\iff$ there is a coordinate system where the metric $g_{\mu\nu}$ is constant. The "proof": for the forward direction it follows from a complicated theorem that we will ignore; for the backward direction it follows easily since the Christoffel symbols vanish.

Now we'll define the following quantity by lowering the indices:

$$R_{\rho\sigma\mu\nu} := g_{\rho\alpha}R^\alpha_{\sigma\mu\nu}$$

We have the following properties of this:

(1) (Anti-symmetric in the second pair of indices): $R_{\rho\sigma\mu\nu} = -R_{\rho\sigma\nu\mu}$.
(2) (Anti-symmetric in its first two indices): $R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}$.
(3) $R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}$.
(4) (Total symmetrization vanishes): $R_{[\rho\sigma\mu\nu]} = 0$ (this is important).

The Riemann tensor involves the first and second derivatives of the metric, but the important feature is it packages them into a tensor, so if the Riemann tensor vanishes in one coordinate system then it much vanish in *every* coordinate system. However, we want something that's simpler than a four-index tensor. We ask ourselves the following question: **Q: What is the sort of thing we could construct from the Riemann tensor?** Let's contract the upper indices with a lower index to get the **Ricci tensor**:

$$R_{\mu\nu} := R^\rho_{\mu\rho\nu}. \tag{20}$$

By the symmetries of the Riemann tensor, this is the *only* possible contraction.

Definition 10.2 (Ricci Tensor). $R_{\mu\nu} := R^\rho_{\mu\rho\nu}$ is the **Ricci Tensor**, a $(0,2)$-tensor. This tensor is symmetric, and up to a sign is the only $(0,2)$-tensor formed by contracting the indices of the Riemann tensor.

We can use the metric to define something even simpler.

Definition 10.3 (Ricci Scalar). We define the **Ricci Scalar** as:

$$R := R_{\mu\nu}g^{\mu\nu}. \tag{21}$$

Let's take a step back. We are developing a theorem of differential geometry to define the curvature of space-time. The equations of motion obeyed by space-time will be an equation involving the curvature (so the Riemann tensor). One possible equation that could describe the curvature of space-time is:

$$R^\rho_{\sigma\mu\nu} = 0,$$

but this would describe flat space-time. This is too simple. Let's go to the next simplest one:

$$R_{\mu\nu} = 0.$$

This is Einstein's equation in the *absence* of matter!

11 Lecture 11: 11 Feb 2021

11.1 Curvature

We want to unpack the idea that $R^{\rho}{}_{\sigma\mu\nu}$ captures the extent to which space-time is not flat, i..e, the extent to which it is impossible to find a coordinate system where the metric is constant. We recall the Equivalence principle: for any point x_0^{μ}, we can find coordinates, say $\hat{x}^{\mu}$ such that $g_{\hat{\mu}\hat{\nu}} = \eta_{\hat{\mu}\hat{\nu}}$ (the metric is the usual flat metric at this point). Indeed, we can always make this choice where this point x_0 is the origin in the hatted coordinates:

$$\hat{x}^{\mu} = 0.$$

Let's unpack the equivalence in more detail: the Riemann tensor captures the extent to which you can't do the above *everywhere*.

Q: How would we implement this idea? How would we do this?

Imagine starting with a coordinate system with a metric $g_{\mu\nu}$ in some coordinates x^{μ}. Let's go to some new coordinate system, $\hat{x}^{\mu}$, where:

$$g_{\hat{\mu}\hat{\nu}} = \frac{\partial x^{\mu}}{\partial x^{\hat{\mu}}} \frac{\partial x^{\nu}}{\partial x^{\hat{\nu}}} g_{\mu\nu}.$$

We want to Taylor expand both sides of this equation in powers of $x^{\hat{\mu}}$ around the point that we're interested in – $x^{\hat{\mu}} = 0$. At $x^{\hat{\mu}} = 0$, the metric has 10 degrees of freedom and the Jacobian, $\frac{\partial x}{\partial \hat{x}}$, has 16 degrees of freedom. Since $16 > 10$, we can just choose the set of derivatives $\frac{\partial x}{\partial \hat{x}}$ to be such that

$$g_{\hat{\mu}\hat{\nu}} = \eta_{\hat{\mu}\hat{\nu}}.$$

In this implementation, we want to think of the Jacobian as a set of free parameters to find the coordinate system we used to get the Einstein Equivalence principle.

Q: What about *near* $\hat{x} = 0$ instead of just $\hat{x} = 0$?

Imagine expanding the Jacobian order by order in powers of $\hat{x}$. Then:

$$\frac{\partial x}{\partial \hat{x}} = \left[\frac{\partial x}{\partial \hat{x}}\right]_{\hat{x}=0} + \left[\frac{\partial^2 x}{\partial \hat{x}\partial \hat{x}}\right]_{\hat{x}=0} \hat{x} + \left[\frac{\partial^3 x}{\partial \hat{x}\partial \hat{x}\partial \hat{x}}\right]_{\hat{x}=0} \hat{x}^2 + \ldots$$

Let's count the degrees of freedom appearing in this expression:

- For the $\left[\frac{\partial x}{\partial \hat{x}}\right]_{\hat{x}=0}$ term, we have one upper index ranging from 0 to 3, and one lower index ranging from 0 to 3. Hence, we have $4 \times 4 = 16$ degrees of freedom.
- For the $\left[\frac{\partial^2 x}{\partial \hat{x}\partial \hat{x}}\right]_{\hat{x}=0}$ term, we have one upper index and two lower indices. By the equality of partial derivatives, this means we have $4 \times 10 = 40$ degrees of freedom, where 10 is the degrees of freedom in a symmetric 2-index tensor.
- For the $\left[\frac{\partial^3 x}{\partial \hat{x}\partial \hat{x}\partial \hat{x}}\right]_{\hat{x}=0}$ term, there is one upper index and three lower indices. Hence, we will have $4\times$ (number of degrees of freedom in a symmetric three-index tensor) $= 4 \times (4 + 4 \times 3 + \frac{4\times3\times2}{6}) = 80$ DOF.

We want to use the freedom to choose coordinates to make the metric as flat as possible $\Rightarrow$ we want to compare to the degrees of freedom of our metric $g_{\hat{\mu}\hat{\nu}}$ expanded about $\hat{x} = 0$:

$$\hat{g} = [\hat{g}]_{\hat{x}=0} + [\partial\hat{g}]_{\hat{x}=0}\hat{x} + [\partial^2\hat{g}]_{\hat{x}=0}\hat{x}^2 + \ldots$$

- $[\hat{g}]$: there are 10 degrees of freedom in a symmetric 4×4 matrix.
- $[\partial\hat{g}]$: there are $4 \times 10 = 40$ degrees of freedom.
- $[\partial^2\hat{g}]$: there are 10 degrees of freedom in the second partial derivative (by equality of the mixed partial derivatives) and there are 10 degrees of freedom in the metric, which gives us a total of 100 degrees of freedom for this term.

So, in particular,

(1) At $O(\hat{x}^0)$, we can use the 16 degrees of freedom in the $\frac{\partial x}{\partial \hat{x}}$ to set $\hat{g} = \hat{\eta}$.
(2) At $O(\hat{x}^1)$, we can get another 40 degrees of freedom in the Jacobian to set $\partial\hat{g} = 0$.
(3) At $O(\hat{x}^2)$, we have 80 degrees of freedom in the Jacobian but 100 degrees of freedom in the metric; this leaves us with 20 degrees of freedom remaining in the second derivative of the metric which cannot be removed by some choice of a coordinate system.

To summarize: the result is that we can always choose coordinates where, at $\hat{x}^\mu = 0$:

$$g_{\hat{\mu}\hat{\nu}} = \eta_{\hat{\mu}\hat{\nu}}$$
$$\partial_{\hat{\rho}} g_{\hat{\mu}\hat{\nu}} = 0,$$

at $\hat{x}^\mu = 0$. These coordinates are called the **Riemann Normal Coordinates** or **locally-inertial coordinates** or **freely-falling coordinates**. We say **locally-inertial coordinates** since this means that the Christoffel symbols vanish if the first derivatives vanish, which means that there are no fictitious forces at that point. This is why the Riemann tensor needs *second* derivatives, since if a tensor vanishes in one coordinate system then it vanishes in all coordinate systems.

Q: What is the Riemann Tensor in these freely-falling coordinates?

Let's compute, at $\hat{x} = 0$:

$$\begin{aligned}
R_{\rho\sigma\mu\nu} &= g_{\rho\alpha} R^\alpha{}_{\sigma\mu\nu} \\
&= g_{\rho\alpha}(\partial_\mu \Gamma^\alpha{}_{\sigma\nu} - \partial_\nu \Gamma^\alpha{}_{\sigma\mu}) \\
&= \frac{1}{2}(\partial_\mu(\partial_\nu g_{\sigma\rho} + \partial_\sigma g_{\nu\rho} - \partial_\rho g_{\sigma\nu}) - \partial_\nu(\partial_\mu g_{\sigma\rho} + \partial_\sigma g_{\mu\rho} - \partial_\rho g_{\sigma\mu})).
\end{aligned}$$

Using the equality of the mixed partial derivatives, we obtain:

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

Using this, we can read off the following properties:

(1) (Anti-symmetry when you switch the last two indices): $R_{\rho\sigma\mu\nu} = -R_{\rho\sigma\nu\mu}$.
(2) (Anti-symmetry under the interchange of the first index with the second): $R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}$.
(3) $R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}$.
(4) $R_{[\rho\sigma\mu\nu]} = R_{\rho[\sigma\mu\nu]} = 0$.

The above are tensor equations, which means that they are true in any coordinate system. Hence, we derived symmetry arguments for the Riemann tensor by thinking of it in inertial coordinates.

Q: How many degrees of freedom are there in the Riemann tensor?

- Conditions (1) and (2) force each pair to be like an anti-symmetric matrix. There are 6 degrees of freedom in a 4×4 anti-symmetric matrix.
- From condition (3), we get $\frac{6 \times 7}{2} = 21$ degrees of freedom.

- Condition (4) removes a degree of freedom from an anti-symmetric 5 index tensor (all four need to be different).

Hence, the number of degrees of freedom in the Riemann tensor is 20. But, we've encountered this number earlier in class. $R^{\rho}{}_{\sigma\mu\nu}$ is just a re-packaging of those 20 degrees of freedom in the second derivative of the metric in a tensorial way. Hence, you should think of the Riemann tensor as all of the coordinate-invariant information in $\partial^2 g$.

We have one final identity on the first derivatives of the Riemann tensor: **Bianchi Identity**:

$$\boxed{\nabla_{[\lambda} R_{\rho\sigma]\mu\nu} = 0} \tag{22}$$

Recall that $R_{\mu\nu} = R^{\rho}{}_{\mu\rho\nu}$ and $R = g^{\mu\nu} R_{\mu\nu}$. Then, the Bianchi identity can be used to show that $\nabla^{\mu} R_{\mu\nu} = \frac{1}{2}\nabla_{\nu} R$.

Now define $G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2}g_{\mu\nu} R$. This is called the **Einstein tensor**. It has the property that:

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

This follows from the Bianchi identity.

Up to a unique sign, $G_{\mu\nu}$ is the unique tensor formed from second derivatives of the metric $g_{\mu\nu}$ with this property. The changes of coordinates which leave the metric $\eta_{\mu\nu}$ unchanged give us the **Symmetries of flat Minkowski space**. They are:

(1) The three rotations
(2) The three boosts

These are called the Lorentz transformations of flat Minkowski space, and they generate the Lorentz group.