Riemannian Geometry and General Relativity

Peng Zhao

July 10, 2009

Contents

1 History and Overview 2
  1.1 What is straight? ............................................. 2
  1.2 What is curvature? ........................................... 3
  1.3 Foucault Pendulum and Gauss-Bonnet ...................... 3

2 Riemannian Geometry .................................. 4
  2.1 Metric ......................................................... 4
  2.2 Vectors and Tensors ....................................... 5
  2.3 Geodesics ..................................................... 7
  2.4 Connection and Curvature .................................. 7

3 General Relativity ....................................... 9
  3.1 Special relativity ........................................... 9
  3.2 Einstein equation .......................................... 10
  3.3 Variational principle ...................................... 10

4 Black Holes ............................................. 10
  4.1 The Schwarzschild black hole ............................ 10
  4.2 Geodesics in Schwarzschild spacetime .................. 11
  4.3 Projects ...................................................... 12
1 History and Overview

It all started in 1854. Bernhard Riemann was a young postdoc at the University of Göttingen. Carl Gauss, his supervisor, had asked him to propose three topics for his Habilitation lecture. Riemann assumed his supervisor will choose the first or the second topic, so he put geometry, a subject he was least sure about, as his third topic. A genius though he was, Riemann forgot the cardinal rule of exams: what you think may not show up will show up. Gauss was most interested in what Riemann had to say about geometry...

After only a few weeks of intense work, Riemann delivered his monumental lecture “Über die Hypothesen, welche der Geometrie zu Grunde liegen”. In this lecture, he described how to generalize Gauss’ idea of surfaces and their curvatures to higher dimensions and thus singlehandedly created the subject of Riemannian geometry. Gauss was extremely impressed by the depth of Riemann’s thought and praised his work without reservation. However, Riemann’s work was so ahead of his time that it did not receive its proper attention in the mathematical community until Einstein’s formulation of general relativity in 1915. In his theory of general relativity, Einstein imagined the spacetime as a geometric object whose curvature is determined by the distribution of energy and matter. Thus gravitational force is no longer a force in the Newtonian sense but a mere manifestation of the curvature of spacetime.

This was one of the great triumphs of mathematics and physics, where a

In general relativity, the spacetime is a manifold \( \mathcal{M} \) and a spacetime event is represented by a point \( p \in \mathcal{M} \). The worldline of the particle is traced by a curve in \( \mathcal{M} \).

There are several basic tenets of general relativity.

- Spacetime is a semi-Riemannian manifold.
- Free particles follow geodesics.
- Curvature tells matter how to move, and matter tells spacetime how to curve.

\[ R_{ij} - \frac{1}{2} R g_{ij} = 8 \pi T_{ij}. \]

Intrinsic properties - those that can be measured on the surface without regard to how it is embedded in an ambient space. Sure, we all know that a sphere is curved by looking at a soccer ball. But it took a while for human to realized that we lived on a sphere. We are still not quite sure of geometry of the universe. We believe that our spacetime is the entire universe and not embedded in some bigger space, it is important that geometric quantities are measured intrinsically.

Remarkably, almost every tool we study in Riemannian geometry will carry over to the study of the spacetime. It leads to exciting concepts such as the black holes, big bang.

1.1 What is straight?

To discuss curved objects, we first need a notion of straightness. In a flat space such as the plane, a straight line is often defined as the shortest path between two points. But this definition requires a notion of distance and to determine the shortest path we need to look at all possible paths. Suppose you are on an airplane. If you have just enough gasoline to get from point \( A \) to point \( B \) but cannot afford to take detours, how do you find out which path is the shortest? More mathematically, we are looking for a local definition of straightness. It is useful to examine how the velocity of a curve changes when . In other words, the change in its tangent vector is along its direction. For a curve with unit velocity, (e.g., parametrized by arc-length), its its acceleration is along the normal direction. Thus a straight line with unit speed cannot have any acceleration where as a curved line turns away from its direction.

We summarize the the two notions of straightness.

- Metric: A straight line is the shortest path between two points.

\(^1\)Remember that Gauss was not easily impressed. There are many stories of other mathematicians showing their latest discovery to Gauss, to which his reaction was “I discovered that twenty years ago”.

2
• **Affine**: A straight line is one whose tangent vector can be parallel transported along itself.

Later we shall generalize these ideas to define straight lines on a manifold using the Riemannian metric and the affine connection. A straight line on a manifold is called a **geodesic** because it is confusing to talk about straight curves. The affine definition of geodesics relies on the notion of angles, and does not require any notion of distance. Sometimes angles can also be defined using the metric, e.g., the inner product. In fact, we shall use the metric to define the **metric connection**.

**Remark 1.1.** A subtle, but absolutely crucial, issue is how to compare vectors at different points of a manifold. In Euclidean space we can just translate the vector at point $q$ to point $p$. This cannot be done for general manifolds. For example, consider vectors on a sphere. A vector at $q \in S^2$ cannot be translated to $p \in S^2$ without leaving the sphere. In the next section we shall see how to define a notion of an affine connection which captures the idea of parallel transport along the manifold.

### 1.2 What is curvature?

What does it mean for a curve or a surface to be curved? It is much easier to talk about the curvature of embedded surfaces than the curvature measured intrinsically. The ancient Greeks already knew that the earth is curved, by looking at ships disappearing over the horizon. But it took Einstein to find out that space and time themselves are curved.

Let us review the notion of curvature, starting from the easier, extrinsic definition. One can define the curvature of a curve at a point $p$ as the inverse of the radius of the osculating circle. Osculating (kissing) is a romantic way of saying that the circle is tangent to the curve at $p$. A point has a lot of curvature if its osculating circle is small and very little curvature if its osculating circle is large. Notice that the notion of an osculating circle is extrinsic to the curve itself. An equivalent, but intrinsic notion of curvature would be measure the acceleration of the curve parametrized by arc-length (unit speed).

**Example 1.1.** Show that the acceleration of a unit-speed curve is inverse to the radius of its osculating circle at every point.

Let $M$ be a two-dimensional surface. Then every normal to the surface at $p \in M$ is contained in some plane.

The Gaussian curvature $K$ at a point $p$ in a surface is defined as the product of the two principle curvatures $\kappa_1$ and $\kappa_2$ at $p$. Gauss’ Theorema Egregium (Totally Awesome Theorem) states that the Gaussian curvature is intrinsic, i.e., invariant under local isometries.

**Remark 1.1.** There is an interesting real-life application of Gauss’ Theorema Egregium. Many of us pizza-lovers have trouble preventing the toppings from falling off when we hold the crust with one hand. One can fight the force of gravity either eat quickly before the toppings fall, or can hold the other end with the other hand. Either ways are pretty clumsy. However, there is an easy way out if you know some geometry! Notice that, when you hold the pizza with one hand, the principle curvature along the crust is much smaller than along the direction of falling toppings. You can change the principle curvatures by creating a large principle curvature along the crust. Because the Gaussian curvature is invariant, this forces the principle curvature along the other direction to diminish, thus saving the toppings from falling off! You may wonder what caused the pizza to **curve** in the first place, hmm...

### 1.3 Foucault Pendulum and Gauss-Bonnet

There is one apparatus at the heart of all physics – the harmonic oscillator. In 1851, the French physicist Léon Foucault designed a large pendulum to test the rotation of the earth. It consisted of a large bob hang from the roof of the Panthéon in Paris. When you enter the Panthéon, you will see the pendulum swinging back and forth in one direction. If you take a look again after taken a tour of the rest of the place, you will see a small change in its plane of rotation. If you come back again the next day, you will see the plane of oscillation rotated by almost 270 degrees! Why did this happen? The rotation of the earth caused the pendulum to be parallel transported along a closed curve. It is purely a geometric effect!
Example 1.2. Show that a vector parallel transported along a spherical triangle obtains a phase shift of sum of interior angles $-\pi$. What is the phase shift of a vector parallelly transported along an arbitrary closed path?

We conclude that the phase shift of the vector transported along a closed path is proportional to the area enclosed by the path. At Paris ($\phi \approx 48^\circ$). The area enclosed by the latitude line is $2\pi \sin \phi \approx 267^\circ$, which is the phase shift measured by the Foucault pendulum!

$$\int K \, dA = \delta \theta.$$  

The Gauss-Bonnet formula for a compact surface with Euler character $\chi$ is

$$\int K \, dA = \chi.$$  

The Gauss-Bonnet theorem is a central theorem of differential geometry. It relates the local geometry of a manifold (the left-hand side) to its global topology (the right-hand side). In 1944, the Chinese mathematician Shiing-Shen Chern gave a intrinsic proof of the Gauss-Bonnet theorem. In project 1, you are invited to read the original paper and present his result.

## 2 Riemannian Geometry

Let us get a bit formal here and give precise definitions. We will rely on the intuition developed so far to understand the underlying geometry.

### 2.1 Metric

**Definition 2.1.** A Riemannian metric $g : V \times V \to \mathbb{R}$ is a symmetric bilinear form.
The metric one usually meets in an analysis course is also supposed to be positive definite, meaning \( g(v, w) \geq 0 \) for all \( v, w \in V \) unless \( v = w = 0 \). We relax this condition because soon we shall see that the metric one uses for spacetime does not satisfy this condition.

Let \( \{e^a\} \) be a basis of \( V \). We can write \( v, w \in V \) as \( v = v^a e_a, w = w^a e_a \). Here we use the Einstein summation convention and discard the summation sign. By linearity,

\[
g(v, w) = g(e_a, e_b)v^aw^b \equiv g_{ab}v^aw^b.
\]

Here \( g_{ab} \) are called the components of \( g \).

The metric allows us to define notions of length and angle. The length of a vector \( v \in V \) is \( |v| \equiv \sqrt{g(v, v)} = \sqrt{g_{ab}v^av^b} \). The angle between two vectors \( v, w \in V \) is \( \cos^{-1} \frac{g(v, w)}{|v||w|} \). Two vectors \( v, w \in V \) are orthogonal if \( g(v, w) = 0 \).

**Example 2.1.** Two metrics \( g, \bar{g} \) are conformally related if \( \bar{g} = \Omega^2 g \) for some \( \Omega \neq 0 \). Show that conformally related metrics preserve the angle between vectors but not the length.

Lorentzian metric.

### 2.2 Vectors and Tensors

Tensor is one of the things that really confused me when learning linear algebra. It is used ubiquitously from algebra to relativity to fluid dynamics. There are several ways of formulating what a tensor is, depending on who you ask. If you ask an algebraist, they will show you a commutative diagram and lots of abstract nonsense. Physicists will tell you “tensor is an object that transforms like a tensor under coordinate transformations”. Here I will take the viewpoint that a tensor is simply a multilinear map, which I believe is a natural way to generalizing the notion of linear maps between vector spaces.

Remember from linear algebra that every vector \( v \in V \) is associated with a covector \( v^* \in V^* \) such that \( v^*(v) = \delta^a_0 \). Formally, \( v^* \) is a linear map \( v^* : V \to \mathbb{R} \) sending each vector to a number. Alternatively, we can think of \( v \) as a linear map \( v : V^* \to \mathbb{R} \) sending each covector to a number. ² What about \( v \otimes w^* : V^* \times V \to \mathbb{R} \).

**Definition 2.2.** An \( \binom{r}{s} \) tensor is a multilinear map

\[
T : V^* \times \cdots \times V^* \times V \times \cdots V \to \mathbb{R}.
\]

**Example 2.2.** A vector can be thought of as a column vector, a covector as a row vector and a rank 2 tensor as a matrix. In this sense we see that tensor is a generalization of the notion of vectors and matrices.

Now the metric \( g \) is a \( \binom{0}{0} \) tensor. It is symmetric because \( g_{ab} = g_{ba} \). We can extend this to the notion of symmetric tensors. In components, a totally-symmetric tensor is invariant under permutation of indices.

**Example 2.3.** Show that the Bach brackets are “infectious” under contraction, i.e.,

\[
S_{(ab)}T^{ab} = S_{(ab)}T^{(ab)} = S_{ab}T^{(ab)}, \quad S_{[ab]}T^{ab} = S_{[ab]}T^{[ab]} = S_{ab}T^{[ab]}.
\]

In fact, many mathematical objects are tensors in disguise. For example, the ubiquitous Kronecker delta

\[
\delta_{ab} = \begin{cases} 1 & \text{ if } a = b \\ 0 & \text{ if } a \neq b \end{cases}
\]

are the components of a \( \binom{0}{0} \) tensor. What is it?³ The Minkowski metric, denote by \( \eta_{ab} \) differs from \( \delta_{ab} \) by a sign. It is defined by \( -\eta(\partial_t, \partial_t) = \eta(\partial_x, \partial_x) = \eta(\partial_y, \partial_y) = \eta(\partial_z, \partial_z) = 1 \). It can be written \( ds^2 = -dt \otimes dt + dx \otimes dx + dy \otimes dy + dz \otimes dz \), where \( ds^2 \) is physicists’ convention for \( g \). In shorthand, we write \( ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \).

**Example 2.4.** Show that the Minkowski metric in polar coordinates \( (t, r, \theta, \phi) \) is \( ds^2 = -dt^2 + dr^2 + r^2d\theta^2 + r^2\sin^2 \theta d\phi^2 \).

---

²This really should be a double dual \( v^{**} \). The space of all double duals \( V^{**} \) is isomorphic to \( V \).

³Hint: The answer is somewhere in this page.
When the manifold is equipped with a metric, there is a canonical way to relate a covector to a vector, and vice versa by the isomorphism \( g : v \mapsto g(v, \cdot) \). In index notation, this corresponds to lowering the vector components \( v_a = g_{ab} v^b \). In a calculation, we are free to raise or lower indices as long as they are consistent on both sides of an equation.

**Example 2.5.** Show that the Minkowski metric \( \eta_{ab} \) can be written in matrix form as

\[
\begin{pmatrix}
-1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{pmatrix}
\text{ in Cartesian coordinates}
\quad \text{and}
\begin{pmatrix}
-1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & r^2 & 0 \\
0 & 0 & 0 & r^2 \sin \theta \\
\end{pmatrix}
\text{ in polar coordinates.}
\]

What is the inverse metric \( \eta^{ab} \) in these coordinates?

The vector space that we will be mostly interested in is the tangent space. Intuitively, a tangent line to a curve and a tangent plane to a surface are those that just “touch” the curve or surface without crossing it more than once. A tangent vector allows us to define the directional derivative in the direction of the vector. In fact, we will take this as a definition of a tangent vector in Riemannian manifolds.

**Definition 2.3.** Let \( \gamma : [0, 1] \to M \) be a smooth curve and \( f : M \to \mathbb{R} \) be a smooth function. Then \( X_p : f \mapsto \frac{d}{ds}(f \circ \gamma)(s)|_p \) is called the tangent vector at \( p \). The tangent space \( T_p M \) is the collection of tangent vectors at \( p \).

Defining a tangent vector as a derivative that acts on functions (a linear derivation) seems a bit strange. For one thing, it is not clear if the \( T_p M \) really is a vector space. But as you can see that the definition is entirely intrinsic to the surface.

**Proposition 2.4.** \( T_p M \) is a vector space.

**Proof.** Exercise. (Hint: What does it mean to add two curves?)

The coordinates \((x^1, \ldots, x^n)\) of \( \mathbb{R}^n \) induces a natural basis of curves for \( T_p M \). Let \( x^i(t) = (0, \ldots, t, \ldots, 0) \) be the \( i \)th coordinate curve. Then the tangent vector along \( x^i(t) \) at \( p \) is given by \( X_p(f) = \left. \frac{d}{dt} f \right|_{t=p} \). We can conveniently write \( X_p(t) = \left. \frac{\partial f}{\partial x^i} \right|_p \), sometimes abbreviated by \( \partial_i \). The basis \( \{\partial_i\}_{i=1, \ldots, n} \) is called the coordinate-induced basis. Sometimes we consider non-coordinate basis for the tangent space, known as moving frames, which is very useful in calculations.

An arbitrary vector \( X_p \in T_p M \) can be written as \( X_p = X^i_p \partial_i \). Acting on \( f \) it gives \( X(f) = \left. X^i_p \frac{\partial f}{\partial x^i} \right|_p \). Thus we can think of a vector as the directional derivative along the direction defined by its components \( \{X^i_p\} \). How do vectors behave under a change of basis? Let \((y^1, \ldots, y^n)\) be another set of coordinates. It induces the basis vectors \( \{Y_p\} = \left\{ \left. \frac{\partial}{\partial y^j} \right|_p \right\} \). The basis vectors act on the coordinate curves \( x^i(t) \) as \( Y_p(x^i) = \left. \frac{\partial x^i}{\partial y^j} \right|_p \).

Then \( X_p \) can be written as \( X_p = \tilde{X}^i_p \partial_i \). Acting on \( f \) it gives \( X(f) = \left. \tilde{X}^i_p \frac{\partial f}{\partial y^i} \right|_p = \tilde{X}^j_p \frac{\partial x^i}{\partial y^j} \frac{\partial f}{\partial x^i} \). But \( \tilde{X}^i_p \frac{\partial}{\partial x^i} \). Thus \( \tilde{X}^j_p = \tilde{X}^i_p \frac{\partial x^i}{\partial y^j} \) is how components of \( X^i_p \) transform under a change of basis. In old-fashioned textbooks, this is taken as the definition of a vector. The concept of a manifold became widely used much later.

What about for covectors? A coordinate-induced basis of covectors \( \{dx^i\} \) dual to \( \left\{ \frac{\partial}{\partial x^i} \right\} \) is defined by

\[
dx^i \left( \frac{\partial}{\partial x^j} \right) = \delta^i_j.
\]

The vector space of covectors is denoted by \( T^*_p M \). An arbitrary covector \( \omega \in T^*_p M \) can be written as \( \omega = \omega_i dx^i \).
Example 2.6. Show that the components of $\omega$ transform as $\omega_i = \underaccent{\bar}{\omega}_j \frac{\partial y^j}{\partial x^i}$.

The components of a general $(^{i}_j)$ tensor transforms as

$$
\overset{\rightarrow}{T}^{a_1 \ldots a_r}_{b_1 \ldots b_r} = \frac{\partial x^{a_1}}{\partial \tilde{x}^{b_1}} \cdots \frac{\partial x^{a_r}}{\partial \tilde{x}^{b_r}} \cdot \overset{\rightarrow}{T}^{\tilde{a}_1 \ldots \tilde{a}_r}_{\tilde{b}_1 \ldots \tilde{b}_r}.
$$

(4)

Old-fashioned textbooks usually take this as the defining property of a tensor, saying something like “tensor is a quantity that transforms as 2.2 under a coordinate transformation. We understand now that it only refers to the way the components transform.

Example 2.7. Let $f : \mathcal{M} \to \mathbb{R}$ be a smooth function. Consider $T_{ij} = \frac{\partial f}{\partial x_i \partial x_j}$, the Hessian of $f$. Show that $T_{ij}$ transforms as the component of a $(^{0}_2)$ tensor at a point where $df = 0$. Can you give a coordinate-independent expression of this tensor?

Example 2.8. Define $g \equiv \sqrt{g_{ab}}$, where $g_{ab}$ is seen as a matrix. How does $g$ transform under a coordinate transformation? Show that $\sqrt{|g|} d^nx$ is invariant under a coordinate transformation.

Definition 2.5. A vector field $X$ on $\mathcal{M}$ is the collection of tangent vectors at each point in $\mathcal{M}$. Covector and tensor fields are defined similarly.

Given two vector fields $X, Y$, we can form another vector field $X(Y)$ defined by $X(Y)f = X(Y(f))$.

Definition 2.6. The vector field commutator is defined by $[X, Y] = X(Y) - Y(X)$.

Example 2.9. Prove the Jacobi identity $[X, [Y, Z]] + [[Y, Z], X] + [Z, [X, Y]] = 0$.

2.3 Geodesics

Let $\gamma : (a, b) \to \mathcal{M}$ be a smooth curve and $\partial_t$ be the tangent vector. Then the length of the curve is defined as $L(\gamma) = \int_a^b ds \sqrt{g(\partial_t(\gamma), \partial_t(\gamma))}$. Note that $L$ is reparametrization invariant: $ds \sqrt{g(\partial_t(\gamma), \partial_t(\gamma))}$, as it should be. Often we parametrize

Example 2.10. If we define $d(x, y) = \inf \{L(\gamma) : \gamma$ is a curve joining $x, y \in \mathcal{M}\}$, then $\mathcal{M}$ is a metric space.

2.4 Connection and Curvature

collection to “connect” vectors at different points

Definition 2.7. A connection is a map $\nabla : (X, Y) \mapsto \nabla_X Y$ mapping a pair of vector fields to another vector field, such that for all $f \in C^\infty(\mathcal{M})$

- $\nabla_{(f_1 + f_2)}X = f_1 \nabla_X Y + f_2 \nabla_Y X$ ($\nabla_Y : X \mapsto \nabla_X Y$ is linear),
- $\nabla_X (f_1 + f_2) = \nabla_X f_1 + \nabla_X f_2$, $\nabla_X (fY) = (\nabla_X f)Y + f \nabla_X Y$ ($\nabla_X : Y \mapsto \nabla_X Y$ is a derivation).

Definition 2.8. A geodesic $\gamma$ along a vector field $X$ is an integral curve satisfying $\nabla_X X = aX$ for some $a \in \mathbb{R}$.

This captures the notion of straightness in terms of parallel transport along itself. For calculation, it is useful to find a coordinate expression for the geodesic condition. Let $X = \dot{x}^a e_a$ be the tangent vector to $\gamma$. Define $\nabla_{e_a} e_b = \Gamma^c_{ab} e_c$, where $\Gamma^c_{ab}$ are called the Christoffel symbols. By Leibniz rule and linearity,

$$
\nabla_X X = X^a \nabla_a (X^b e_b) = X^a \partial_a X^b e_b + X^a X^b \Gamma_{ab}^c e_c = X^a (\partial_a X^c) + X^b \Gamma_{ab}^c e_c = (\dot{x}^c + \Gamma^c_{ab} \dot{x}^a \dot{x}^b) e_c = 0.
$$

(5)

The last expression is known as the geodesic equation. It is a second order ODE. From ODE theory, we know that there exists a unique solution locally for any initial conditions $x^a(0)$ and $\dot{x}^a(0)$. 

7
Example 2.11. For a 2-sphere whose metric is $ds^2 = r^2 d\theta d\phi + r^2 \sin \theta d\phi^2$. Show that the nonzero Christoffel symbols are $\Gamma^\theta_{\phi\phi} = -\sin \theta \cos \theta$, $\Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot \theta$. Using the geodesic equation, show that the geodesics are great circles.

Note that this captures the key idea that a curve is “straight” if it is parallel transported. Remember that a straight line between two points is also one that minimizes the distance in between. Is it true for a curved space?

Example 2.12. Show that a geodesic minimizes the length.

Definition 2.9. Let $\gamma(s)$ be a timelike curve parametrized by its proper time. Then the vector $v = v^i \partial_i$ tangent to $\gamma(s)$ is called the four-velocity.

Example 2.13. In Newtonian gravity freely falling particles move under the force law $\vec{F} = m \ddot{x}$, where the force is the divergence of the gravitational potential: $\vec{F} = -m \nabla \phi$. Let $x^a = (t, \vec{x})$ be the four position and $\vec{x} = (1, \vec{x})$ be the four-velocity of the particle. We can express this in geodesic form as $\dot{x}^a + \partial^a \phi \dot{x}^0 \dot{x}^0 = 0$. We discover that there is a nonzero Christoffel symbol $\Gamma^a_{00} = \partial^a \phi$ and a Riemann tensor $R^a_{000} = \partial^a \partial^0 \phi$. Thus Newtonian spacetime is curved!

We can carry out the calculation further and compute the Einstein tensor $G_{00} = R_{00} - \frac{1}{2} R g_{00} = 2 \Delta \Phi$. By the Einstein equation, $\Delta \Phi = 4 \pi T_{00} = 4 \pi \rho$. Thus the Einstein equation is equivalent to the Poisson equation in the Newtonian limit.

Definition 2.10. The torsion tensor is a $(1, 1)$ tensor defined by $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$.

Example 2.14. Check that the torsion tensor really is a tensor. Show that a vanishing torsion is equivalent to symmetric connection: $\Gamma^a_{bc} = \Gamma^a_{cb}$ for coordinate-induced basis.

The torsion tensor measures how $X$ rotates as it is parallely transported along $Y$. Note that $T[X,Y] = -T[Y,X]$. In Einstein’s theory of general relativity, torsion is assumed to vanish. Alternative theories exists where the torsion is nonzero, e.g., Einstein-Cartan theory.

Definition 2.11. A connection is metric-compatible if $\nabla g = 0$.

Proposition 2.12. There exists a unique torsion-free, metric-compatible connection, called the Levi-Civita connection.

Proof. By the Leibniz rule and metric-compatibility,

$$X (g(Y, Z)) = \nabla_X g(Y, Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

$$= g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

$$Y (g(Z, X)) = g(\nabla_Y Z, X) + g(Z, \nabla_Y X)$$

$$Z (g(X, Y)) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y)$$

(6) (7) (8)

Take (1) - (2) + (3) and assuming zero torsion $\nabla_X Y - \nabla_Y X = [X, Y]$, we get

$$g(\nabla_X Y, Z) = \frac{1}{2} (-Z (g(X, Y)) + Y (g(Z, X)) + X (g(Y, Z)) + g(Z, [X, Y]) + g(Y, [Z, X]) - g(X, [Y, Z]).$$

(9)

Definition 2.13. The Riemann tensor is a $(1, 3)$ tensor defined by $R(X, Y)Z = (\nabla_X \nabla_Y - \nabla_Y \nabla_X)Z - \nabla_{[X,Y]}Z$.

Example 2.15. Check that the Riemann tensor really is a tensor. Show that, in coordinate-induced basis,

$$R^a_{bcd} = \partial_d \Gamma^a_{bc} - \partial_c \Gamma^a_{bd} + \Gamma^e_{cd} \Gamma^a_{be} - \Gamma^e_{ce} \Gamma^a_{bd}.$$  

(10)

Prove the Ricci identity $(\nabla_a \nabla_b - \nabla_b \nabla_a)X^a = R^a_{bce} X^b$. 

8
Clearly, \( R(X, Y)Z = -R(Y, X)Z \). Other symmetries of the Riemann tensor include \( g(R(X, Y)Z, W) = -g(R(X, Y)W, Z) \) and \( g(R(X, Y)Z, W) = g(R(Z, W)X, Y) \). In components, \( R_{abcd} = g(R(e_a, e_b)v^c, e_c) \). The symmetries are \( R_{abcd} = R_{cdab} = -R_{bacd} = -R_{abcd} \).

**Example 2.16.** In two dimensions, there is only one independent component of the Riemann tensor, namely \( R_{1212} \). Using the symmetries, we can write \( R_{abcd} = \frac{R}{2}(g_{ac}g_{bd} - g_{ad}g_{bc}) = K \det(g) \). \( K \) is called the Gaussian curvature and it is the product of the two principle curvatures. For a 2-sphere of radius \( r \), show that the Gaussian curvature is \( 1/r^2 \).

By the symmetries of the Riemann tensor, the only nonzero contraction is the Ricci tensor defined by \( Ric(X, Y) = g(R(Z, X)Y, Z) \). In components, \( Ric_{ab} = R_{ab} = R_{aeb} \). Note that the Ricci tensor is symmetric.

We can contract again and form the Ricci scalar \( R = R^a_a \). The Einstein tensor \( G_{ab} = R_{ab} - \frac{1}{2}Rg_{ab} \).

**Example 2.17.** Show that \( G_{ab} \) is divergence free.

Ricci flow:

\[
\frac{\partial g_{ij}}{\partial t} = -2R_{ij}.
\] (11)

The Ricci flow can be thought as a heat equation on a manifold. The LHS depends on the first time derivative of the metric and the RHS depends on the second spacial derivatives of the metric. Recall that the heat equation “smooths out” an initial temperature profile as it evolves in time. Similarly, given an initial manifold, the Ricci flow smooths out its geometry. It was a key idea in proving the celebrated Poincaré conjecture, which states that any 3-manifold homotopic to a 3-sphere is homeomorphic to it.

**Example 2.18.** The Ricci curvature \( R_{ab} = R_{aeb} \) is often called the trace part of the Riemann tensor. The trace-free part of the Riemann tensor is known as the Weyl tensor. It is defined as

\[
C_{abcd} = R_{abcd} - \frac{2}{n-2}(g_{a[c}R_{d]b} - g_{b[c}R_{d]a}) + \frac{2}{(n-1)(n-2)}Rg_{a[c}g_{d]b}.
\] (12)

Show that \( C_{abcd} \) has the symmetries as \( R_{abcd} \) and that it is trace-free. The Einstein equation determines the Ricci tensor by the energy momentum tensor. The Weyl curvature accounts for the spacetime curvature in the absence of matter sources, i.e., it represents gravitational waves in vacuum.

**Example 2.19.** Show that the Einstein equation in two spacetime dimensions is trivial.

### 3 General Relativity

#### 3.1 Special relativity

**Definition 3.1.** Let \( v \) be a vector. Then according to the sign of its norm

- If \( |v|^2 < 0 \), it is called a timelike vector.
- If \( |v|^2 = 0 \), it is called a null vector.
- If \( |v|^2 > 0 \), it is called a spacelike vector.

A timelike curve is one whose tangent vector is timelike. Null and spacelike curves are defined similarly. A non-spacelike curve is called a causal curve.

**Example 3.1.** Future-directed.

Let \( v \) be the four-velocity of a timelike geodesic. The energy momentum four vector is \( P^a = mv^a = (E, \vec{p}) \). Let \( u \) be the four velocity a static observer. Then we can decompose \( v^a \) into the components parallel to \( u \) and components orthogonal to \( u \) via the projection tensor \( h^a_b = \delta^a_b + v^a u_b \). The components orthogonal to \( u \) are \( h^a_b v^b = v^a + u^a u^b v_b = v^a - \gamma^a u^a \). Thus we can write \( v^a = (\gamma^a - \gamma u^a) + \gamma u^a \). The energy momentum four vector can be decomposed as \( P^a = E u^a + P^a_\perp \), where \( P^a_\perp \) is orthogonal to \( u^a \). We define \( E = -P^a u_a \) as the energy measured by the static observer.
3.2 Einstein equation

\[ G_{ab} = 8\pi T_{ab}. \] (13)

3.3 Variational principle

In Newtonian theory, a particle with position \( \vec{x} \) in a gravitational field is associated with a kinetic energy \( \frac{1}{2}m\dot{x}^2 \) and a potential energy \( -mg\vec{x} \). The difference between the kinetic energy and the potential energy is called the Lagrangian \( L(\vec{x}, \dot{\vec{x}}) \). We can integrate the Lagrangian along the trajectory of the particle \( \gamma \) to form the action

\[
S = \int_{\gamma} dt \, L = \int_{\gamma} dt \left( \frac{1}{2}m\dot{x}^2 + mg\vec{x} \right).
\] (14)

The equation of motion is obtained by finding the stationary point of the action \( \delta S = 0 \). This gives rise to the Euler-Lagrange equation

\[
\frac{\partial L}{\partial \vec{x}} = \frac{d}{dt} \frac{\partial L}{\partial \dot{\vec{x}}}.
\] (15)

We find the equation of motion \( \ddot{x} = g \), as expected.

There is an elegant way to derive the Einstein equation from purely aesthetic principles. The Einstein-Hilbert action

\[
I_{EH} = \int d^4x \sqrt{-g} R.
\] (16)

4 Black Holes

4.1 The Schwarzschild black hole

The first exact solution to the vacuum Einstein’s equation \( R_{ij} = 0 \) is the Schwarzschild metric

\[
ds^2 = -\left(1 - \frac{2M}{r} \right) dt^2 + \left(1 - \frac{2M}{r} \right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2.
\] (17)

The metric takes a particularly simple form, independent of time \( t \) and spherical coordinates \( (\phi, \theta) \). In fact, the Schwarzschild black hole is the unique static, spherically symmetric solution of the vacuum Einstein equation. But there is a lot hidden in this metric then meets the eye. First of all, it predicts the existence of a region of spacetime causally disconnected from the rest of the universe, known as the black hole. It is one of the most mysterious and fascinating predictions of general relativity. Second, inside the black hole there is a singular point at which physics breaks down. Each of these surprising predictions continue to pose serious challenge for understanding.

Looking at the metric, the values \( r = 2M \) and \( r = 0 \) should cause alarm. At \( r = 2M \), the \((t, r)\) components of the metric vanish and the metric is not invertible. The problem is more serious at \( r = 0 \), where the \((t, r)\) components blow up and the metric is not even defined there. As we shall see, the singularity at \( r = 2M \) can be removed by performing a suitable coordinate transformation, and is known as a coordinate singularity.

At each point in time, \( r = 2M \) defines a 2-sphere, called an event horizon. To get a glimpse of what happens here, consider an observer approaching the event horizon from the region \( r > 2M \). As he approaches the event horizon, his proper time \( s = g(\partial_t, \partial_t) = g_{00} = \left(1 - \frac{2M}{r} \right) \) tends to infinity. This means that as measured by an observer at infinity, it takes forever for an object to reach the event horizon.
To better see what happens near the event horizon, we adopt the tortoise coordinate \( r^* = r - 2M \log \frac{r}{2M} \).

Note that \( dr^* = \left(1 - \frac{2M}{r}\right) dr \). Let \( v = t + r^* \) be the new time coordinate. In ingoing Eddington-Finkelstein coordinates \((v, r, \theta, \phi)\), the Schwarzschild metric becomes

\[
ds^2 = -\left(1 - \frac{2M}{r}\right)^2 dv^2 + 2dvdr + r^2d\theta + r^2 \sin \theta d\phi.
\]

(18)

Ingoing null geodesics follow \( v = \text{const} \). Even though the \( v \) component of the metric vanishes at \( r = 2M \), the metric is invertible there and can be analytically continued to \( r < 2M \).

**Example 4.1.** Let \( u = t - r^* \). Show that in outgoing Eddington-Finkelstein coordinates \((u, r, \theta, \phi)\), the Schwarzschild metric becomes

\[
ds^2 = -\left(1 - \frac{2M}{r}\right)^2 du^2 - 2dudr + r^2d\theta + r^2 \sin \theta d\phi.
\]

(19)

Draw the spacetime diagram and show that there is a white hole region that no causal curve can enter from the outside.

### 4.2 Geodesics in Schwarzschild spacetime

Consider a particle outside of a spherically symmetric star or black hole. Its four velocity \( v = (\dot{t}, \dot{r}, \dot{\theta}, \dot{\phi}) \) satisfies

\[-l = -\left(1 - \frac{2M}{r}\right)^2 \dot{t}^2 + \left(1 - \frac{2M}{r}\right)^{-1} \dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin \theta \dot{\phi}^2.
\]

(20)

By spherical symmetry, it suffices to set \( \theta = \frac{\pi}{2} \) and (20) becomes

\[-l = -\left(1 - \frac{2M}{r}\right)^2 \dot{t}^2 + \left(1 - \frac{2M}{r}\right)^{-1} \dot{r}^2 + r^2 \dot{\phi}^2.
\]

(21)

Note that \( k = \partial_t \) is a Killing vector, so that \( E = g^{ab}k_a\dot{v}_b = -\left(1 - \frac{2M}{r}\right)\dot{t} \) is conserved along its worldline. \( m = \partial_\phi \) is also Killing, and is associated with the conserved quantity \( h = g_{ab}k^a\dot{v}^b = r^2\dot{\phi} \). Then we may write (21) as

\[
\frac{1}{2} \dot{r}^2 + V(r) = \frac{1}{2} E^2, \quad V(r) = l \left(1 - \frac{2M}{r}\right) + \frac{\hbar^2}{2r^2} - \frac{Mrh^2}{r^3}.
\]

(22)

Now compare this with the gravitational potential from Newtonian theory, \( V(r) = -\frac{M}{r} + \frac{\hbar^2}{2r^2} \). There are qualitative differences. In Newtonian theory, there is a infinity potential wall as \( r \to 0 \), preventing the particle from falling into the center. In GR, the \( -\frac{Mrh^2}{r^3} \) term dominate as \( r \to 0 \) and particles can fall into the center.

**Example 4.2.** Show that there is an unstable circular orbit at \( r = 3M \) for light \((l = 0)\).
4.3 Projects

**Project 1.** Chern’s proof of Gauss-Bonnet theorem.

**Project 2.** Cartan’s method for calculating curvature.

Introduce differential forms, moving frames, spin connection, Cartan’s structural equations. Compute the curvature outside of a Schwarzschild black hole.

**Project 3.** Classical Tests of GR

Gravitational redshift, procession of the perihelion of Mercury, bending of light.

**Project 4.** Penrose diagrams.

Conformal compactification of the Minkowski space. Kruskal extension and the Penrose diagram of Schwarzschild spacetime.