

MATH635  
Riemannian Geometry

Pingbang Hu

March 27, 2024

## Abstract

This is the advanced graduate-level differential geometry course focused on Riemannian geometry taught by [Lydia Bieri](#) at University of Michigan. Topics include local and global aspects of differential geometry and the relation with the underlying topology. We'll use do Carmo's *Riemannian Geometry* [FC13] as our reference. Apart from this, I also found [Sch15] very useful.

A noticeable different is that we introduce [geodesics](#) differently from do Carmo [FC13], where we set the solution of the variations of [energy](#) to define a [geodesic](#) first, and then draw connection to the “[curve](#) with zero acceleration” after introduce the [covariant derivative](#); however, do Carmo [FC13] first introduce [covariant derivative](#) and then return the variation view point much later.



# Contents

<b>1</b>	<b>Smooth Manifolds</b>	<b>3</b>
1.1	Topological Manifolds . . . . .	3
1.2	Differentiable Manifolds . . . . .	4
1.3	Tangent and Cotangent Spaces . . . . .	10
1.4	Vector Fields and Brackets . . . . .	13
1.5	Submanifolds, Immersions, and Embeddings . . . . .	15
<b>2</b>	<b>Riemannian Manifolds</b>	<b>17</b>
2.1	Riemannian Metrics . . . . .	17
2.2	Geodesics . . . . .	19
2.3	Hopf-Rinow Theorem . . . . .	24
2.4	Vector Bundles and Tensor Fields . . . . .	28
2.5	Other Metrics . . . . .	31
<b>3</b>	<b>Connections and Curvatures</b>	<b>33</b>
3.1	Levi-Civita Connections . . . . .	33
3.2	Riemannian Curvatures . . . . .	36
3.3	Flows of Vector Fields . . . . .	38
3.4	Covariant Derivatives and Parallelism . . . . .	40
3.5	More on Tangent and Cotangent Bundles . . . . .	43
3.6	Sectional Curvatures . . . . .	47
3.7	More on Covariant Derivatives . . . . .	48
<b>4</b>	<b>Isometric Immersions</b>	<b>50</b>
4.1	Riemannian Covering Maps . . . . .	50
4.2	The Second Fundamental Form . . . . .	52
4.3	The Fundamental Equations . . . . .	55
<b>5</b>	<b>Jacobi Fields</b>	<b>57</b>
5.1	Jacobi Fields . . . . .	57
5.2	Variations of Length and Energy . . . . .	58
5.3	Index Form . . . . .	59
5.4	Jacobi Fields and Geodesics . . . . .	63
5.5	Conjugate Points . . . . .	65
5.6	The Cut Locus . . . . .	69
<b>6</b>	<b>Morse Index, Rauch Comparison, Sphere Theorems, and More</b>	<b>72</b>
6.1	Morse Index Theorem . . . . .	72
6.2	Morse Theory and Flow Homology . . . . .	75
6.3	The Rauch Comparison Theorem . . . . .	80
6.4	The Sphere Theorem . . . . .	83
<b>7</b>	<b>Epilogue</b>	<b>89</b>
7.1	Uniformization Theorem . . . . .	89
7.2	Lorentzian Manifolds and General Relativity . . . . .	93
7.3	Ricci Flow . . . . .	94

---

<b>A</b>	<b>Additional Notes</b>	<b>96</b>
A.1	The $C^\infty(\mathcal{M})$ -Module Viewpoint of Tensor Fields . . . . .	96
A.2	Lie Groups and Lie Algebra . . . . .	97

# Chapter 1

## Smooth Manifolds

### Lecture 1: A Foray to Smooth Manifolds

#### 1.1 Topological Manifolds

5 Jan. 13:00

Let's start with a common definition.

**Definition 1.1.1 (Topological manifold).** A *topological manifold*  $\mathcal{M}$  of dimension  $n$  is a Hausdorff and second-countable (topological) space such that each point  $p \in \mathcal{M}$  has a neighborhood  $U$  homeomorphic via  $\varphi: U \rightarrow U'$  to an open subset  $U' \subseteq \mathbb{R}^n$ .

**Definition 1.1.2 (Local coordinate map).** For every  $p \in \mathcal{M}$ , the corresponding homeomorphism  $\varphi$  is called the *local coordinate map*.

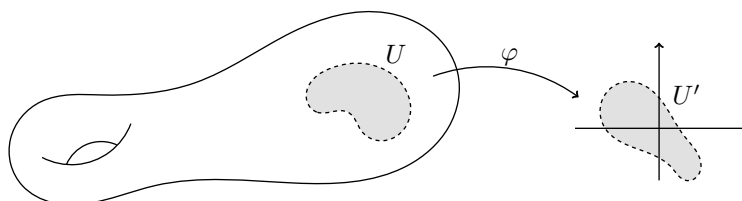
**Definition 1.1.3 (Local coordinate).** The pull-back  $(x^1, \dots, x^n)$  of the *local coordinate map*  $\varphi$  from  $\mathbb{R}^n$  is called the *local coordinates* on  $U$ , given by

$$\varphi(p) = (x^1(p), \dots, x^n(p)).$$

**Definition 1.1.4 (Coordinate chart).** The pair  $(U, \varphi)$  is called a *(coordinate) chart* on  $\mathcal{M}$ .

**Remark.** The reason why we want the space to be Hausdorff and second-countable is because we can then have *partition of unity*.

In other words, a *topological manifold* can be thought of as a space such that it looks like  $\mathbb{R}^n$  locally.



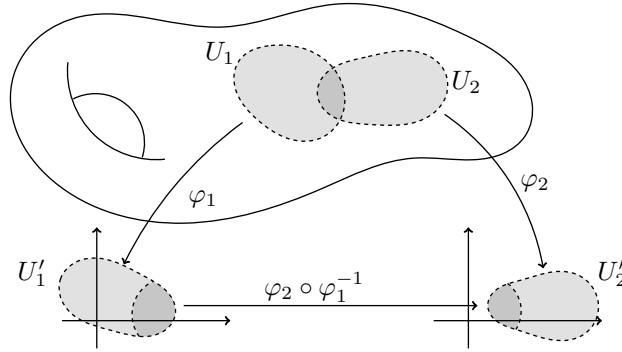
**Definition 1.1.5 (Atlas).** An *atlas*  $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}_\alpha$  for a *manifold*  $\mathcal{M}$  is a collection of *charts* such that  $\{U_\alpha \subseteq \mathcal{M} \mid U_\alpha \text{ open}\}_\alpha$  are an open covering of  $\mathcal{M}$ , i.e.,  $\mathcal{M} = \bigcup_\alpha U_\alpha$ .

In other words, for all  $p \in \mathcal{M}$ , there exists a neighborhood  $U \subseteq \mathcal{M}$  and homeomorphism  $h: U \rightarrow U' \subseteq \mathbb{R}^n$  open.

**Definition 1.1.6** (Locally finite). An **atlas** is said to be *locally finite* if each point  $p \in \mathcal{M}$  is contained in only a finite collection of its open sets.

Clearly, without any help of ambient space such as  $\mathbb{R}^n$ , there's no clear way to make sense of differentiability of a **manifold**. But thankfully, we now have an explicit relation to the ambient space  $\mathbb{R}^n$  via  $\varphi_\alpha$ . To formalize, let  $\mathcal{A}$  be an **atlas** for a **manifold**  $\mathcal{M}$ , and assume that  $(U_1, \varphi_1), (U_2, \varphi_2)$  are 2 elements of  $\mathcal{A}$ . Then clearly, the map  $\varphi_2 \circ \varphi_1^{-1}: \varphi_1(U_1 \cap U_2) \rightarrow \varphi_2(U_1 \cap U_2)$  is a homeomorphism between 2 open sets of Euclidean spaces since both  $\varphi_1$  and  $\varphi_2$  are homeomorphisms. Due to this map's importance, it has its own name.

**Definition 1.1.7** (Coordinate transition). The map  $\varphi_2 \circ \varphi_1^{-1}$  is called the *coordinate transition* of  $\mathcal{A}$  for the pair of **charts**  $(U_1, \varphi_1), (U_2, \varphi_2)$ .



## 1.2 Differentiable Manifolds

Notice that the **coordinate transitions** are from  $\mathbb{R}^n$  to  $\mathbb{R}^n$ ; hence differentiability makes sense now, which induces the following.

**Definition 1.2.1** (Differentiable atlas). The **atlas**  $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}$  is *differentiable* if all **transitions** are differentiable.

**Remark.** Here, the differentiability depends on the content. Sometimes, we may want it to be  $C^\infty$ , and sometimes may be  $C^k$  for some finite  $k$ . On the other hand, smooth always refers to  $C^\infty$ . We'll use them interchangeably if it's clear which case we're referring to.

**Definition 1.2.2** (Equivalence atlas). Two **atlases**  $\mathcal{U}, \mathcal{V}$  of a **manifold** are equivalent if for every  $(U, \varphi) \in \mathcal{U}, (V, \psi) \in \mathcal{V}$ ,

$$\varphi \circ \psi^{-1}: \psi(U \cap V) \rightarrow \varphi(U \cap V)$$

and

$$\psi \circ \varphi^{-1}: \varphi(U \cap V) \rightarrow \psi(U \cap V)$$

are diffeomorphisms between subsets of Euclidean spaces.

Notably, we have the following notation.

**Notation** (Smoothly compatible). Two **charts**  $(U, \varphi)$  and  $(V, \psi)$  are *smoothly compatible* if either  $U \cap V = \emptyset$  or  $\psi \circ \varphi^{-1}$  is a diffeomorphism.

This suggests the following.

**Definition 1.2.3** (Smooth structure). A *smooth structure* on  $\mathcal{M}$  is an equivalence class  $\mathcal{U}$  of **coordinate atlas** with the property that all **transition functions** are diffeomorphisms.

**Remark.** We can also use the *maximal differentiable atlas* to be our differentiable structure.

**Definition 1.2.4** (Smooth manifold). A *smooth manifold* is a manifold  $\mathcal{M}$  with a *smooth structure*.

In this way, we can do calculus on *smooth manifolds*! Furthermore, it now makes sense to say that a function  $f: \mathcal{M} \rightarrow \mathbb{R}$  is differentiable (or  $C^\infty$ ) by considering differentiability of  $f \circ \varphi^{-1}$  around  $p$ . This leads to the following.

**Notation.** The collection of *smooth functions* on *smooth manifold*  $\mathcal{M}$  is denoted by  $C^\infty(\mathcal{M}, \mathbb{R})$ , or  $C^k(\mathcal{M}, \mathbb{R})$ .<sup>a</sup>

<sup>a</sup>We will formalize this later in [Definition 1.2.8](#).

**Remark.** The class  $C^\infty(\mathcal{M}, \mathbb{R})$  consists of functions with property is well-defined.

**Proof.** Let  $\mathcal{A}$  be any given *atlas* from *equivalence class* that defines the *smooth structure*, and as we have shown, if  $(U, \varphi) \in \mathcal{A}$ , then  $f \circ \varphi^{-1}$  is smooth on  $\mathbb{R}^n$ . This requirement defines the same set of *smooth functions* no matter the choice of representative *atlas* by the nature of [Definition 1.2.2](#) requirement that defines the equivalent *manifolds*.  $\circledast$

### 1.2.1 Orientation

Another essential property of a *manifold* is its orientability.

**Definition.** Consider an *atlas*  $\mathcal{A}$  for a *differentiable manifold*  $\mathcal{M}$ .

**Definition 1.2.5** (Oriented).  $\mathcal{A}$  is *oriented* if all *transitions* have positive functional determinant.

**Definition 1.2.6** (Orientable).  $\mathcal{M}$  is *orientable* if  $\mathcal{A}$  is an *oriented atlas*.

Motivated by the above definitions, we see that we can actually use an *atlas* to define an *orientation*.

**Definition 1.2.7** (Orientation). Let  $\mathcal{M}$  be an *orientable manifold*. Then a *oriented differentiable structure* is called an *orientation* of  $\mathcal{M}$ .

If  $\mathcal{M}$  possesses an *orientation*, we can also say that it's *oriented*. But we don't bother to make a new definition to confuse ourselves with [Definition 1.2.5](#).

**Remark.** Two *differentiable structures* obeying [Definition 1.2.5](#) determine the same *orientation* if the union again satisfying [Definition 1.2.5](#).

**Remark.** If  $\mathcal{M}$  is *orientable* and connected, then there exists exactly 2 distinct *orientations* on  $\mathcal{M}$ .

Now, we can see some examples of *smooth manifolds*.

**Example** (Sphere). The sphere  $S^n \subseteq \mathbb{R}^{n+1}$  given by

$$S^n = \{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \dots + x_{n+1}^2 = 1\}.$$

Consider  $U_i^+ = \{x \in S^n \mid x_i > 0\}$ ,  $U_i^- = \{x \in S^n \mid x_i < 0\}$  for  $i = 1, \dots, n+1$ , and  $h_i^\pm: U_i^\pm \rightarrow \mathbb{R}^n$  such that

$$h_i^\pm(x_1, \dots, x_{n+1}) = (x_1, \dots, \hat{x}_i, \dots, x_{n+1}).$$

Note that the minimum *charts* needed to cover  $S^n$  is 2.

**Example.** Let  $\mathcal{M} = U \subseteq \mathbb{R}^n$ , then  $\{(U, \varphi)\}$  is a **smooth structure** with  $\varphi = \text{id}$ .

**Example.** Open sets of  $C^\infty$ -**manifolds** are  $C^\infty$ -**manifolds**.

**Example (General linear group).**  $\text{GL}(n) = \{A \in M_n(\mathbb{R}) \mid \det A \neq 0\} \subseteq M_n(\mathbb{R}) = \mathbb{R}^{n^2}$ , open.

**Example (Real projective space).**  $\mathbb{RP}^n = S^n / \sim$  where  $x \sim -x$  with  $\pi: S^n \rightarrow \mathbb{RP}^n$ ,  $x \mapsto [x]$ .

**Proof.**  $\pi$  is a homeomorphism on each  $U_i^+$  for  $i = 1, \dots, n+1$ , with

$$\{(\pi(U_i^+), \varphi_i^+ \circ \pi^{-1}), i = 1, \dots, n+1\}$$

is a  $C^\infty$ -**atlas** for  $\mathbb{RP}^n$ . \*

**Note.** Observe that  $\mathbb{RP}^n = \mathbb{R}^{n+1} \setminus \{0\} / \sim$ .

## Lecture 2: Maps Between Smooth Manifolds

### 1.2.2 Smooth Maps

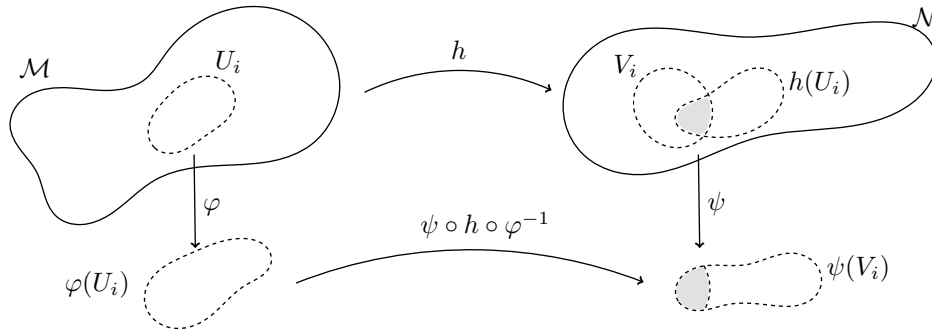
10 Jan. 13:00

We can now consider the maps between **manifolds**, specifically, the **smooth manifolds**.

**Definition 1.2.8 (Smooth function).** Let  $\mathcal{M}, \mathcal{N}$  be two **smooth manifolds**, and let  $\mathcal{U}$  be **locally finite atlas** from the **equivalence class** that gives the **smooth structure** on  $\mathcal{M}$ , and let  $\mathcal{V}$  be the corresponding for  $\mathcal{N}$ . A map  $h: \mathcal{M} \rightarrow \mathcal{N}$  is said to be *smooth* if each map in the collection

$$\{\psi \circ h \circ \varphi^{-1} : h(U) \cap V \neq \emptyset\}$$

where  $(U, \varphi) \in \mathcal{U}$ ,  $(V, \psi) \in \mathcal{V}$  is  $C^\infty$ -differentiable as a map from one Euclidean space to another.



**Remark.** **Equivalence relation** guarantees that **Definition 1.2.8** depends only on the **smooth structure** of  $\mathcal{M}, \mathcal{N}$ , but not on the chosen representative **coordinate atlas**.

**Definition.** Consider two **smooth manifolds**  $\mathcal{M}, \mathcal{N}$  and a **smooth homeomorphism**  $h: \mathcal{M} \rightarrow \mathcal{N}$  with **smooth** inverse.

**Definition 1.2.9 (Diffeomorphic).** The two **manifolds**  $\mathcal{M}, \mathcal{N}$  are said to be *diffeomorphic*.

**Definition 1.2.10 (Diffeomorphism).** The map  $h$  is said to be a *diffeomorphism*.



Let  $\mathcal{M}_1, \mathcal{M}_2$  be two smooth manifolds, and let  $\varphi: \mathcal{M}_1 \rightarrow \mathcal{M}_2$  be a diffeomorphism. Then

- (a)  $\mathcal{M}_1$  is orientable if and only if  $\mathcal{M}_2$  is orientable.
- (b) If in addition,  $\mathcal{M}_1$  and  $\mathcal{M}_2$  are both connected and oriented, then  $\varphi$  induces an orientation on  $\mathcal{M}_2$  that may or may not coincide with the initial orientation of  $\mathcal{M}_2$ .

If the induced orientation coincides, then we say  $\varphi$  preserves the orientation, otherwise  $\varphi$  reverses the orientation.

### 1.2.3 Grassmannian Manifold

Before proceeding, let's consider an interesting smooth manifold.

**Definition 1.2.11 (Grassmannian manifold).** Given  $m, n \in \mathbb{N}$ , the so-called *Grassmannian manifold*  $G(n, m)$  is the set of all  $n$ -dimensional subspaces of  $\mathbb{R}^{n+m}$ .

**Note.**  $G(1, m)$  is just  $\mathbb{R}P^m$ , and  $G(0, m), G(n, 0)$  are one-point sets.

As we will soon see,  $G(n, m)$  has the smooth structure of an  $mn$ -dimensional manifold.

**Intuition.** We obtain the structure by exhibiting an atlas whose transitions are diffeomorphisms.

Firstly, we give  $G(n, m)$  a suitable topology, i.e., the metric topology. Let  $\Pi \in G(n, m)$ , and let  $\mathcal{L}(\Pi, \Pi^\perp)$  denote the  $mn$ -dimensional space of linear maps from  $\Pi$  to  $\Pi^\perp$ . Define the map

$$\varphi_\Pi: \mathcal{L}(\Pi, \Pi^\perp) \rightarrow G(n, m), \quad \varphi_\Pi(\alpha) = (\mathbb{1}_\Pi \oplus \alpha)(\Pi)$$

where  $\mathbb{1}_\Pi \oplus \alpha$  is regarded as a map  $\Pi \rightarrow \Pi \oplus \Pi^\perp = \mathbb{R}^{n+m}$ .<sup>1</sup> Clearly,  $\varphi_\Pi$  is injective, and thus,  $(\mathcal{L}(\Pi, \Pi^\perp), \varphi_\Pi)$  is an  $mn$ -dimensional chart of  $G(n, m)$ .

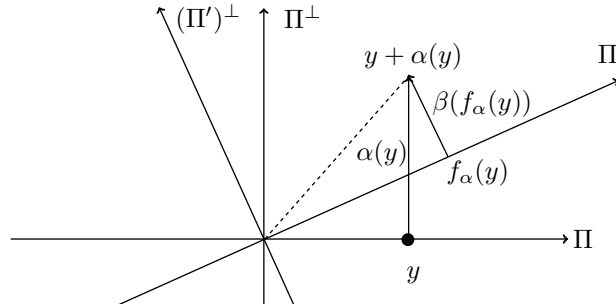
**Remark.** The images  $\varphi_\Pi(\mathcal{L}(\Pi, \Pi^\perp))$  cover  $G(n, m)$ .

**Example.**  $\Pi = \varphi_\Pi(0) \in \varphi_\Pi(\mathcal{L}(\Pi, \Pi^\perp))$ .

We can now prove that these charts are mutually compatible. Let  $\Pi, \Pi' \in G(n, m)$ , and let  $P, P'$  be orthogonal projections from  $\mathbb{R}^{n+m}$  onto  $\Pi, \Pi'$  respectively. Firstly,

$$F = \varphi_{\Pi'}^{-1} \varphi_\Pi: \varphi_\Pi^{-1}(\varphi_{\Pi'}(\mathcal{L}(\Pi', (\Pi')^\perp))) \rightarrow \varphi_{\Pi'}^{-1}(\varphi_\Pi(\mathcal{L}(\Pi, \Pi^\perp)))$$

is smooth.



Consider  $\alpha \in \mathcal{L}(\Pi, \Pi^\perp)$ , and  $\beta \in \mathcal{L}(\Pi', (\Pi')^\perp)$ , then for  $\alpha, \beta$ , the equality  $F(\alpha) = \beta$  means that  $\varphi_\Pi(\alpha) = \varphi_{\Pi'}(\beta)$ . Let  $f_\alpha: \Pi \rightarrow \Pi'$  be defined by

$$f_\alpha = P' \circ (\mathbb{1}_\Pi \oplus \alpha).$$

We need to check

<sup>1</sup>In other words,  $\varphi_\Pi(\alpha)$  is the graph of  $\alpha$  in  $\Pi \oplus \Pi^\perp = \mathbb{R}^{n+m}$ .

- (a)  $f_\alpha$  is invertible, and  
 (b)  $\forall y \in \Pi, y + \alpha(y) = f_\alpha(y) + \beta(f_\alpha(y))$ .

**Note.** The condition  $\det f_\alpha \neq 0$  gives an exact description of the subset  $\varphi_{\Pi^{-1}}(\varphi_{\Pi'}(\mathcal{L}(\Pi', (\Pi')^\perp)))$  of  $\mathcal{L}(\Pi, \Pi^\perp)$ , which is therefore open.

For  $\beta$ , it is  $(\mathbb{1}_{\Pi'} \oplus \beta) \circ f_\alpha = \mathbb{1}_\Pi \oplus \alpha$ , and hence

$$\beta = F(\alpha) = (\mathbb{1}_\Pi \oplus \alpha) \circ f_\alpha^{-1} - \mathbb{1}_{\Pi'}.$$

It follows by the construction that the image of  $\beta$  is contained in  $(\Pi')^\perp$ .

**Remark.** We obtain an infinite atlas for  $G(n, m)$  with charts labeled by  $\Pi \in G(n, m)$ . But it suffices to consider only  $\binom{n+m}{n}$  charts corresponding to subspaces  $\Pi$  spanned with  $n$  coordinate axes.

## 1.2.4 Other Manifold Properties

We now introduce two notions.

**Definition 1.2.12** (Closed manifold). A manifold is *closed* if it is compact and without boundary.

**Definition 1.2.13** (Open manifold). A manifold is *open* if it has only non-compact components without boundary.

**Lemma 1.2.1.** If  $\mathcal{M}$  can be covered by two coordinate neighborhoods  $V_1, V_2$  such that  $V_1 \cap V_2$  is connected, then  $\mathcal{M}$  is *orientable*.

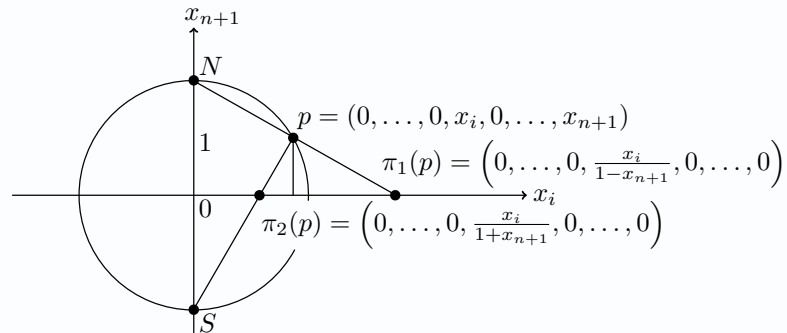
**Proof.** The determinant of the differential of the coordinate change  $\neq 0$ , so it does not change sign in  $V_1 \cap V_2$ . If it's negative at a single point, it's enough to change the sign of one of the coordinates to make it positive at that point, hence on  $V_1 \cap V_2$ . ■

**Example.** Let  $S^n = \{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} x_i^2 = 1\} \subseteq \mathbb{R}^{n+1}$  is *orientable*.

**Proof.** Let  $N = (0, \dots, 0, 1)$  and  $S = (0, \dots, 0, -1)$ , consider given  $p = (0, \dots, 0, x_i, 0, \dots, x_{n+1})$ , then  $\pi_1: S^n \setminus \{N\} \rightarrow \mathbb{R}^n$  given by

$$\pi_1(p) = \left(0, \dots, 0, \frac{x_i}{1 - x_{n+1}}, 0, \dots, 0\right)$$

to be the stereographic projection from the North Pole  $N$ .



More generally, it takes  $p(x_1, \dots, x_{n+1}) \in S^n - \{N\}$  into the intersection at the hyperplane

$x_{n+1} = 0$  with the line passing through  $p$  and  $N$ . In this way, we have

$$\pi_1(x_1, \dots, x_n) = \left( \frac{x_1}{1 - x_{n+1}}, \frac{x_2}{1 - x_{n+1}}, \dots, \frac{x_n}{1 - x_{n+1}} \right),$$

hence  $\pi_1: S^n \setminus \{N\} \rightarrow \mathbb{R}^n$  is differentiable, and is injective. Similarly,  $\pi_2: S^n \setminus \{S\} \rightarrow \mathbb{R}^n$  for  $S$  can also be defined and everything holds similarly. We see that these two parametrizations  $(\mathbb{R}^n, \pi_1^{-1}), (\mathbb{R}^n, \pi_2^{-1})$  cover  $S^n$ . The [change of coordinate](#) is given by

$$y_j = \frac{x_j}{1 - x_{n+1}} \leftrightarrow y'_j = \frac{x_j}{1 + x_{n+1}}, \quad (y_1, \dots, y_n) \in \mathbb{R}^n, \quad j = 1, \dots, n,$$

where

$$y'_j = \frac{y_j}{\sum_{i=1}^n y_i^2}.$$

This implies that  $\{(\mathbb{R}^n, \pi_1^{-1}), (\mathbb{R}^n, \pi_2^{-1})\}$  is a [differentiable structure](#) for  $S^n$ . Now, consider

$$\pi_1^{-1}(\mathbb{R}^n) \cap \pi_2^{-1}(\mathbb{R}^n) = S^n \setminus \{N, S\},$$

which is connected, hence  $S^n$  is [orientable](#), and the above [structure](#) gives an [orientation](#) of  $S^n$ .  $\circledast$

## Lecture 3: Tangent Spaces and Bundles

Let's look at two more examples about [orientation](#).

12 Jan. 13:00

**Example.** Let  $A: S^n \rightarrow S^n$  be the antipodal map given by  $A(p) = -p$  for  $p \in \mathbb{R}^{n+1}$ . It's easy to see that  $A$  is differentiable with  $A^2 = \mathbb{1}$ . Furthermore,  $A$  is a [diffeomorphism](#) of  $S^n \subseteq \mathbb{R}^{n+1}$ . We see that

- if  $n$  is even,  $A$  reverses the [orientation](#);
- if  $n$  is odd,  $A$  preserves the [orientation](#).

**Example.**  $G(k, n)$  is [orientable](#) if and only if  $n$  is even or  $n = 1$ .

Finally, we introduce the notion of [complex manifolds](#).

**Definition 1.2.14 (Complex manifold).** A *complex manifold*  $\mathcal{M}$  of complex dimension  $d$  ( $\dim_{\mathbb{C}} \mathcal{M} = d$ ) is a [differentiable manifold](#) of (real) dimension  $2d$  ( $\dim_{\mathbb{R}} \mathcal{M} = 2d$ ) whose [charts](#) take values in open subsets of  $\mathbb{C}^d$  with holomorphic [chart transitions](#).

**As previously seen.** The [chart transitions](#)  $z_\beta \circ z_\alpha^{-1}: z_\alpha(U_\alpha \cap U_\beta) \rightarrow z_\beta(U_\alpha \cap U_\beta)$  is holomorphic if  $\partial z_\beta^j / \partial \bar{z}_\alpha^k = 0$  for all  $j, k$  where

$$\frac{\partial}{\partial \bar{z}^k} = \frac{1}{2} \left( \frac{\partial}{\partial x^k} + i \frac{\partial}{\partial y^k} \right).$$

We're not going to spend more time on [complex manifolds](#) in this course, however, one important thing to note is the following.

**Remark.** [Complex Grassmannians](#)  $G_{\mathbb{C}}(k, n)$  are all [orientable](#). More generally, [complex manifolds](#) are always [orientable](#) because holomorphic maps always have positive functional determinant.

### 1.2.5 Partition of Unity

So far, we have defined [functions](#) between [manifolds](#) in a local way. However, to get a global definition from the local definitions, it's not that straightforward. One way to do this is by the so-called [partition of unity](#).

**Definition 1.2.15** (Partition of unity). Let  $\mathcal{M}$  be a [differentiable manifold](#), and let  $(U_\alpha)_{\alpha \in \mathcal{A}}$  be an open covering of  $\mathcal{M}$ . Then a *partition of unity* is a [locally finite](#) refinement  $(V_\beta)_{\beta \in \mathcal{B}}$  of  $(U_\alpha)$  and  $C^\infty$ -functions  $\varphi_\beta: \mathcal{M} \rightarrow \mathbb{R}$  with

- (a)  $\text{supp}(\varphi_\beta) \subseteq V_\beta$  for all  $\beta \in \mathcal{B}$ ;
- (b)  $0 \leq \varphi_\beta(x) \leq 1$  for all  $x \in \mathcal{M}$ ,  $\beta \in \mathcal{B}$ ;
- (c)  $\sum_{\beta \in \mathcal{B}} \varphi_\beta = 1$  for all  $x \in \mathcal{M}$ .<sup>a</sup>

<sup>a</sup>There are only finitely many non-vanishing summands of each point, since only finitely many  $\varphi_\beta$  are non-zero of any given point as the covering  $(V_\beta)$  is [locally finite](#).

**Intuition.** Essentially, a [partition of unity](#) gives us a way to weight all the “local functions” with domains (i.e., the [local coordinates](#)) containing a particular point. By patching these weighted function values together, we get a final, consistent function value of that point.

We state, without proof, of an important lemma about the [partition of unity](#).

**Lemma 1.2.2** (Partition of unity). Let  $\mathcal{M}$  be a [differentiable manifold](#), and let  $(U_\alpha)_{\alpha \in \mathcal{A}}$  be an open covering of  $\mathcal{M}$ . Then there exists a [partition of unity](#) subordinate to  $(U_\alpha)$ ,

**Remark.** [Lemma 1.2.2](#) is only possible by requiring  $\mathcal{M}$  being Hausdorff and second-countable.

## 1.3 Tangent and Cotangent Spaces

### 1.3.1 Tangent Spaces in Euclidean Spaces

To discuss the concept of calculus between [manifolds](#) formally, we start with our discussion in Euclidean spaces, where we naturally have the coordinates for every point.

**Definition.** Let  $\mathcal{M}$  be a Euclidean [manifold](#) of dimension  $d$ ,  $x = (x^1, \dots, x^d)$  be Euclidean coordinates of  $\mathbb{R}^d$ , and  $x_0 \in \Omega \subseteq \mathbb{R}^d$  where  $\Omega$  is open.

**Definition 1.3.1** (Tangent space of Euclidean space). The *tangent space*  $T_{x_0}\Omega$  of  $\Omega$  at  $x_0$  is the vector space  $\{x_0\} \times E^a$  spanned by the basis  $(\partial/\partial x^1, \dots, \partial/\partial x^d)$ .

<sup>a</sup> $E$  is a  $d$ -dimensional Euclidean space.

**Definition 1.3.2** (Tangent vector of Euclidean space). The elements in the [tangent space of Euclidean spaces](#) is called *tangent vectors*.

Before proceeding, we introduce a shorthand notation.

**Notation** ([Einstein notation](#)). The *Einstein notation* abbreviates the summation  $\sum_i v^i x_i$  as  $v^i x_i$ , where we implicitly sum over the upper and lower index.

**Definition 1.3.3** (Differential of Euclidean space). If  $\Omega, \Omega' \subseteq \mathbb{R}^d$  are open, and  $f: \Omega \rightarrow \Omega'$  is differentiable, then the *differential*  $df(x_0)$  for  $x_0 \in \Omega$  is the induced linear map between [tangent spaces](#)

$$df(x_0): T_{x_0}\Omega \rightarrow T_{f(x_0)}\Omega', \quad v = v^i \frac{\partial}{\partial x^i} \mapsto v^i \frac{\partial f^j}{\partial x^i}(x_0) \frac{\partial}{\partial f^j}.$$

**Definition 1.3.4 (Tangent bundle of Euclidean space).** The *tangent bundle* is defined as  $T\Omega := \coprod_{x \in \Omega} T_x \Omega \cong \Omega \times E \cong \Omega \times \mathbb{R}^d$ , which is an open subset of  $\mathbb{R}^d \times \mathbb{R}^d$ .

**Note (Total space).**  $T\Omega$  is also called the *total space*.

**Remark.** Given a *tangent bundle*  $T\Omega$ , we define  $\pi$  to be the projection  $\pi: T\Omega \rightarrow \Omega$  given by  $\pi(x, v) = x$ . This makes  $T\Omega$  naturally a *differentiable manifold*.

With the notion of *tangent bundle*, given  $f: \Omega \rightarrow \Omega'$ , we can also define  $df: T\Omega \rightarrow T\Omega'$  as

$$\left(x, v^i \frac{\partial}{\partial x^i}\right) \mapsto \left(f(x), v^i \frac{\partial f^j}{\partial x^i}(x_0) \frac{\partial}{\partial f^j}\right).$$

**Notation.** We often write  $df(x)(v)$  instead of  $df(x, v)$  to coincide with the notation of *differential*.

In particular, for  $v = v^i \partial/\partial x^i$ , we have

$$df(x)(v) = v^i \frac{\partial f}{\partial x^i}(x) \in T_{f(x)} \mathbb{R} \cong \mathbb{R},$$

and we write  $v(f)(x)$  for  $df(x)(v)$ .

### 1.3.2 Tangent Spaces in Manifolds

We now try to formally define the *tangent space* on a *smooth manifold*. A natural idea is the following.

**Intuition.** Let  $\mathcal{M}^d$  be a *differentiable manifold* with a *chart*  $x: U \rightarrow \Omega \subseteq \mathbb{R}^d$  and  $p \in U \subseteq \mathcal{M}$  where  $U$  is open. The *tangent space*  $T_p \mathcal{M}$  of  $\mathcal{M}$  at  $p$  should be represented in the *chart*  $x$  by  $T_{x(p)}x(U)$ .

To see that the above are well-defined, i.e.,  $T_p \mathcal{M}$  are independent of the choice of *charts*, let  $x': U' \rightarrow \mathbb{R}^d$  to be another *chart* with  $p \in U' \subseteq \mathcal{M}$  where  $U'$  is also open. Denote  $\Omega := x(U)$ , and  $\Omega' := x'(U')$ , then the transition map

$$x' \circ x^{-1}: x(U \cap U') \rightarrow x'(U \cap U')$$

induces a vector space isomorphism

$$L := d(x' \circ x^{-1})(x(p)): T_{x(p)}\Omega \rightarrow T_{x'(p)}\Omega',$$

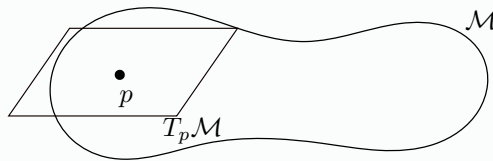
such that  $v \in T_{x(p)}\Omega$  and  $L(v) \in T_{x'(p)}\Omega'$  represent the same *tangent vector* in  $T_p \mathcal{M}$ .

**Remark.** A *tangent vector* in  $T_p \mathcal{M}$  is given by the family of the *coordinate representations*.

Now, we want to define the similar notion of *differential of Euclidean spaces*. Let consider a simple case first, where we let  $f: \mathcal{M} \rightarrow \mathbb{R}$  to be a differentiable function, and assume that the *tangent vector*  $w \in T_p \mathcal{M}$  is represented by  $v \in T_{x(p)}x(U)$ .

**Intuition.** We want to define  $df(p)$  as a linear map from  $T_p \mathcal{M} \rightarrow \mathbb{R}$ . In *chart*  $x$ , let  $w \in T_p \mathcal{M}$  be given as  $v = v^i \partial/\partial x^i \in T_{x(p)}x(U)$ . Say that  $df(p)(w)$  in this chart represented by

$$d(f \circ x^{-1})(x(p))(v).$$



**Remark.**  $T_p\mathcal{M}$  is a vector space of dimension  $d$  isomorphic to  $\mathbb{R}^d$ , where the isomorphism depends on choice of **chart**.

**Intuition.** Pull functions on  $\mathcal{M}$  back by a **chart** to an open subset of  $\mathbb{R}^d$ , differentiate there.

In order to obtain a **tangent space** which does not depend on **charts**, we need to have transformation behavior under change of **charts**. Let  $F: \mathcal{M}^d \rightarrow \mathcal{N}^c$  be a differentiable map where  $\mathcal{M}, \mathcal{N}$  are **smooth manifolds**. Then we want to represent  $dF$  in **local charts**  $x: U \subseteq \mathcal{M} \rightarrow \mathbb{R}^d, y: V \subseteq \mathcal{N} \rightarrow \mathbb{R}^c$  by  $d(y \circ F \circ x^{-1})$ . The **local coordinates** on  $U$  is given by  $(x^1, \dots, x^d)$ , and on  $V$  is  $(F^1, \dots, F^c)$  such that

$$F(x) = (F^1(x^1, \dots, x^d), \dots, F^c(x^1, \dots, x^d)).$$

Then,  $dF$  induces a linear map  $dF: T_p\mathcal{M} \rightarrow T_{F(x)}\mathcal{N}$  which in our **coordinate representation** is given by the matrix

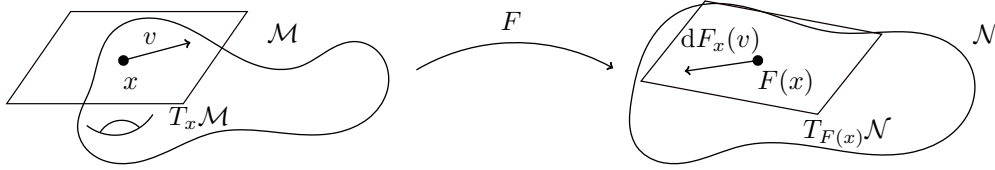
$$\left( \frac{\partial F^\alpha}{\partial x^i} \right)_{\substack{\alpha=1, \dots, c \\ i=1, \dots, d}},$$

and a change of **charts** is then just the base change at **tangent spaces**: if

$$\begin{aligned} (x^1, \dots, x^d) &\mapsto (\xi^1, \dots, \xi^d) \\ (F^1, \dots, F^c) &\mapsto (\phi^1, \dots, \phi^c) \end{aligned}$$

are **coordinate changes**, then  $dF$  represented in the new **coordinates** is given by

$$\left( \frac{\partial \phi^\beta}{\partial \xi^j} \right) = \left( \frac{\partial \phi^\beta}{\partial F^\alpha} \frac{\partial F^\alpha}{\partial x^i} \frac{\partial x^i}{\partial \xi^j} \right).$$



## Lecture 4: Tangent Bundles, Vector Fields, and Submanifolds

**Definition.** Let  $\mathcal{M}^d$  be a **differentiable manifold** with a **chart**  $x: U \rightarrow \Omega \subseteq \mathbb{R}^d$  and  $p \in U \subseteq \mathcal{M}$  where  $U$  is open. On  $\{(x, v) \mid v \in T_x(p)\Omega\}$ , we define an equivalence relation by  $(x, v) \sim (y, w)$  if and only if  $w = d(y \circ x^{-1})v$ .

**Definition 1.3.5 (Tangent space).** The space of equivalence classes is called the *tangent space*  $T_p\mathcal{M}$  at point  $p$  to  $\mathcal{M}$ .

**Definition 1.3.6 (Tangent vector).** The elements in the **tangent space** is called *tangent vectors*.

**Remark.**  $T_p\mathcal{M}$  naturally carries the structure of a vector space.

Now,  $T\mathcal{M}$  is defined as

$$T\mathcal{M} := \coprod_{p \in \mathcal{M}} T_p\mathcal{M}.$$

Recall the projection  $\pi: T\mathcal{M} \rightarrow \mathcal{M}$  with  $\pi(V) = p$  for  $V \in T_p\mathcal{M}$ . Then we can define the following.

**Definition 1.3.7 (Derivation).** If  $x: U \rightarrow \mathbb{R}^d$  be a **chart** for  $\mathcal{M}$ , and let  $TU = \coprod_{p \in U} T_pU$ . Then we define the *derivation*  $dx: TU \rightarrow T_x(U) := \coprod_{p \in x(U)} T_p\mathcal{M}$  by  $w \mapsto dx(\pi(w))(w) \in T_{x(\pi(w))}x(U)$ .

17 Jan. 13:00

The transition maps

$$dx' \circ (dx)^{-1} = d(x' \circ x^{-1})$$

are differentiable.  $\pi$  is local represented by  $x \circ \pi \circ dx^{-1}$  maps  $(x_0, v) \in Tx(U)$  to  $x_0$ .

**Definition 1.3.8 (Tangent bundle).** The triple  $(T\mathcal{M}, \pi, \mathcal{M})$  is called the *tangent bundle* of  $\mathcal{M}$ .

**Definition 1.3.9 (Total space).**  $T\mathcal{M}$  is called the *total space* of the *tangent bundle*.

We can choose the courses (the initial) topology for *total space*  $T\mathcal{M}$  such that  $\pi$  is continuous. Furthermore, we can construct a  *$C^\infty$ -atlas*  $\mathcal{A}_{T\mathcal{M}}$  on  $T\mathcal{M}$  from the  *$C^\infty$ -atlas*  $\mathcal{A}$  of  $\mathcal{M}$ . Specifically, consider  $\mathcal{A}_{T\mathcal{M}} := \{(TU, \xi_x) \mid (U, x) \in \mathcal{A}\}$  where  $\xi_x: TU \rightarrow \mathbb{R}^{2 \cdot d}$  such that

$$x \mapsto ((x^1 \circ \pi)(x), \dots, (x^d \circ \pi)(x), (dx^1)_{\pi(x)}(X), \dots, (dx^d)_{\pi(x)}(X)).$$

**Intuition.** We know that  $X = X_x^i (\partial/\partial x^i)_{\pi(x)}$ , and we might tempt to write  $X^i$  as the last  $d$  components. But we write it in the above way is because

$$(dx^j)_{\pi(x)}(X) = (dx^j)_{\pi(x)} \left( X_x^i \left( \frac{\partial}{\partial x^i} \right)_{\pi(x)} \right) = X_x^i \delta_i^j = X_x^j.$$

**Note.** We can check that  $\xi_x^{-1}$  exists, and it's also smooth, hence  $T\mathcal{M}$  has a natural topology and a  *$C^\infty$ -atlas* making it a  $2 \dim \mathcal{M}$ -dimensional *smooth manifold*.

### 1.3.3 Cotangent Spaces

Another important objects is the *cotangent spaces*.

**Definition.** Let  $\mathcal{M}^d$  be a *differentiable manifold*, and  $T_p\mathcal{M}$  be the *tangent space* at  $p$  to  $\mathcal{M}$ .

**Definition 1.3.10 (Cotangent space).** The *cotangent space*  $T_p^*\mathcal{M}$  to  $\mathcal{M}$  is the dual of  $T_p\mathcal{M}$ , i.e.,  $T_p^*\mathcal{M} = (T_p\mathcal{M})^*$ .

**Definition 1.3.11 (Cotangent vector).** The elements in the *cotangent space* is called *cotangent vectors*.

**Remark.**  $T_p^*\mathcal{M}$  is the space of 1-forms on  $T_p\mathcal{M}$ .

**Notation** (Covariant vector). The *cotangent vectors* are also called *covariant vectors*.

**Notation** (Contravariant vector). The *tangent vectors* are also called *contravariant vectors*.

Similarly, we can define the projection  $\pi: T^*\mathcal{M} \rightarrow \mathcal{M}$  with  $\pi(\omega) = p$  for  $\omega \in T_p^*\mathcal{M}$ , and we have the following.

**Definition 1.3.12 (Cotangent bundle).** The triple  $(T^*\mathcal{M}, \pi, \mathcal{M})$  is called the *cotangent bundle* of  $\mathcal{M}$ .

## 1.4 Vector Fields and Brackets

### 1.4.1 Vector Fields

We now introduce the notion of *vector field*.

**Definition 1.4.1 (Vector field).** A (tangent) vector field  $X$  on a differentiable manifold  $\mathcal{M}$  is a correspondence associating to each point  $p \in \mathcal{M}$  a vector  $X(p) \in T_p\mathcal{M}$ , i.e.,  $X: \mathcal{M} \rightarrow T\mathcal{M}$ .

**Remark.** Naturally, we say that the field  $X$  is differentiable if the map  $X$  is differentiable.

Considering a local chart  $x: U \subseteq \mathbb{R}^n \rightarrow \mathcal{M}$ , we can write

$$X(p) = \sum_{i=1}^n a_i(p) \frac{\partial}{\partial x_i},$$

where  $a_i: U \rightarrow \mathbb{R}$  are functions on  $U$  for  $i = 1, \dots, n$ , and  $\{\partial/\partial x_i\}_i$  is the basis associated to  $x$ .

**Remark.**  $X$  is differentiable if and only if  $a_i$  are differentiable for some (and, therefore, for any)  $x$ .

It's convenient to think of a vector field as a mapping  $X: \mathcal{D} \rightarrow \mathcal{F}$  from the set  $\mathcal{D}$  of differentiable functions on  $\mathcal{M}$  to the set  $\mathcal{F}$  of the functions on  $\mathcal{M}$ , defined by

$$(Xf)(p) = \sum_{i=1}^n a_i(p) \frac{\partial f}{\partial x_i}(p),$$

where  $f$  is implicitly denoting the expression of  $f$  in the chart  $x$ .

**Intuition.** This idea of a vector as a directional derivative is precisely what was used to define the notion of tangent vector.

**Remark.**  $Xf$  does not depend on the choice of  $x$ .

**Remark.**  $X$  is differentiable if and only if  $X: \mathcal{D} \rightarrow \mathcal{D}$ , i.e.,  $Xf \in \mathcal{D}$  for all  $f \in \mathcal{D}$ .

Observe that if  $\varphi: \mathcal{M} \rightarrow \mathcal{M}$  is a diffeomorphism,  $v \in T_p\mathcal{M}$  and  $f$  differentiable function in a neighborhood of  $\varphi(p)$ , we have

$$(d\varphi(v)f)\varphi(p) = v(f \circ \varphi)(p)$$

since by letting  $\alpha: (-\epsilon, \epsilon) \rightarrow \mathcal{M}$  be a differentiable curve with  $\alpha'(0) = v$ ,  $\alpha(0) = p$ ,<sup>2</sup> then

$$(d\varphi(v)f)\varphi(p) = \left. \frac{d}{dt}(f \circ \varphi \circ \alpha) \right|_{t=0} = v(f \circ \varphi)(p).$$

### 1.4.2 Brackets

By viewing  $X$  as an operator on  $\mathcal{D}$ , we can consider the iterates of  $X$ , i.e, given differentiable fields  $X$  and  $Y$  and  $f: \mathcal{M} \rightarrow \mathbb{R}$  being a differentiable function, consider  $X(Yf)$  and  $Y(Xf)$ .

**Note.** In general,  $X(Yf)$  (and hence  $Y(Xf)$ ) is not a field.

**Proof.** It involves derivatives of order higher than one. ⊛

But we have the following.

**Lemma 1.4.1.** Let  $X, Y$  be differentiable vector fields on a smooth manifold  $\mathcal{M}$ . Then there exists a unique vector field  $Z$  such that for all  $f \in \mathcal{D}$ ,  $Zf = (XY - YX)f$ .

**Proof.** See do Carmo [FC13, §0 Lemma 5.2]. ■

This  $Z$  is called the bracket.

<sup>2</sup>This is the way do Carmo [FC13] used to define tangent vectors.



**Definition 1.4.2 (Bracket).** Given two differentiable **vector fields**  $X, Y$  on a **smooth manifold**  $\mathcal{M}$ , the *bracket* of  $X$  and  $Y$  is defined by

$$[X, Y] := XY - YX.$$

Clearly,  $[X, Y]$  is differentiable.

**Proposition 1.4.1.** If  $X, Y$  and  $Z$  are differentiable **vector fields** on  $\mathcal{M}$ ,  $a, b \in \mathbb{R}$ ,  $f, g$  are differentiable functions, then we have the following.

- (a)  $[X, Y] = -[Y, X]$  (*anti-commutativity*),
- (b)  $[aX + bY, Z] = a[X, Z] + b[Y, Z]$  (*linearity*),
- (c)  $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$  (*Jacobi identity*),
- (d)  $[fX, gY] = fg[X, Y] + fX(g)Y - gY(f)X$ .

**Proof.** See do Cargo [FC13, §0 Proposition 5.3]. ■

**Example.**  $[\partial/\partial x^i, \partial/\partial x^j] = 0$  for  $i = j$ .

## 1.5 Submanifolds, Immersions, and Embeddings

We now study the relation between **manifolds**.

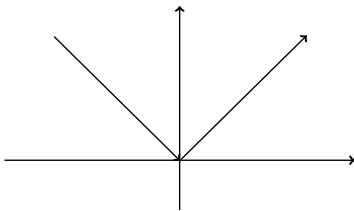
**Definition 1.5.1 (Immersion).** Let  $\mathcal{M}^m, \mathcal{N}^n$  be **smooth manifolds**. A differentiable mapping  $\varphi: \mathcal{M} \rightarrow \mathcal{N}$  is an *immersion* if

$$d\varphi_p: T_p\mathcal{M} \rightarrow T_{\varphi(p)}\mathcal{N}$$

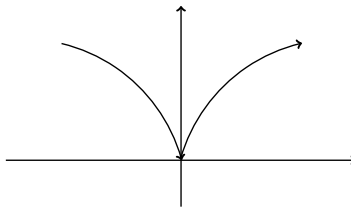
is injective for every  $p \in \mathcal{M}$ .

**Definition 1.5.2 (Embedding).** An **immersion**  $\varphi: \mathcal{M} \rightarrow \mathcal{N}$  is an *embedding* if it is also a homeomorphism onto  $\varphi(\mathcal{M}) \subseteq \mathcal{N}$ , with  $\varphi(\mathcal{M})$  having the subspace topology induced from  $\mathcal{N}$ .

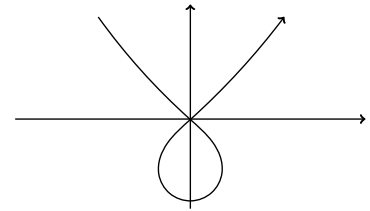
**Definition 1.5.3 (Submanifold).** If the inclusion  $\iota: \mathcal{M} \hookrightarrow \mathcal{N}$  between two **manifolds** is an **embedding**, then  $\mathcal{M}$  is a *submanifold* of  $\mathcal{N}$ .



(a) Non-differentiable curve.



(b) Non-immersion curve.



(c) Non-embedding curve.

Figure 1.1: Three simple examples

**Lemma 1.5.1.** Let  $f: \mathcal{M}^m \rightarrow \mathcal{N}^n$  to be an **immersion** and  $x \in \mathcal{M}$ .<sup>a</sup> Then there exists a neighborhood  $U$  of  $x$  and a **chart**  $(V, y)$  on  $\mathcal{N}$  with  $f(x) \in V$  such that  $f|_U$  is a differentiable **embedding** and  $y^{m+1}(p) = \dots = y^n(p) = 0$  for all  $p \in f(U \cap V)$ .

<sup>a</sup>Hence,  $n \geq m$ .

**Proof.** In the **local coordinates**  $(z^1, \dots, z^n)$  on  $\mathcal{N}$ , and  $(x^1, \dots, x^m)$  on  $\mathcal{M}$ , without loss of generality,<sup>a</sup> let

$$\left( \frac{\partial z^\alpha(f(x))}{\partial x^i} \right)_{i, \alpha=1, \dots, m}$$

be non-singular. Consider

$$F(z, x) := (z^1 - f^1(x), \dots, z^n - f^n(x)),$$

which has maximal rank in  $x^1, \dots, x^m, z^{m+1}, \dots, z^n$ . By the **implicit function theorem**, locally, there exists a map  $\varphi: U \rightarrow \mathbb{R}^n$  such that

$$(z^1, \dots, z^m) \mapsto (\varphi^1(z^1, \dots, z^m), \dots, \varphi^n(z^1, \dots, z^m)) = x$$

such that  $F(z, x) = 0$ , i.e.,

$$\varphi^i(z^1, \dots, z^m) = \begin{cases} x^i, & \text{if } i = 1, \dots, m; \\ z^i, & \text{if } i = m+1, \dots, n, \end{cases}$$

for which

$$\left( \frac{\partial \varphi^i}{\partial z^\alpha} \right)_{\alpha, i=1, \dots, m}$$

has maximal rank. Now, we choose a new coordinate

$$(y^1, \dots, y^n) = (\varphi^1(z^1, \dots, z^m), \dots, \varphi^m(z^1, \dots, z^m), \\ z^{m+1} - \varphi^{m+1}(z^1, \dots, z^m), \dots, z^n - \varphi^n(z^1, \dots, z^m)).$$

Then, we have  $z = f(x) \Leftrightarrow F(z, x) = 0$ , i.e.,  $(y^1, \dots, y^n) = (x^1, \dots, x^n, 0, \dots, 0)$ , proving the result. ■

<sup>a</sup>Since  $df(x)$  is injective.

**Lemma 1.5.2.** Let  $f: \mathcal{M}^m \rightarrow \mathcal{N}^n$  be a differentiable map such that  $m \geq n$  with  $p \in \mathcal{N}$ . Let  $df(x)$  has rank  $n$  for all  $x \in \mathcal{M}$  with  $f(x) = p$ . Then  $f^{-1}(p)$  is the union of differentiable **submanifolds** of  $\mathcal{M}$  of dimension  $m - n$ .

**Remark.** Let  $\mathcal{N}^n$  be a **smooth manifold**, and let  $1 \leq m \leq n$ . Then an arbitrary subset  $\mathcal{M} \subseteq \mathcal{N}$  has the structure of **differentiable submanifold** of  $\mathcal{N}$  of dimension  $m$  if and only if for all  $p \in \mathcal{M}$ , there exists a smooth **chart**  $(U, \varphi)$  of  $\mathcal{N}$  such that  $p \in U$ ,  $\varphi(p) = 0$ ,  $\varphi(U)$  is open, and

$$\varphi(U \cap \mathcal{M}) = (-\epsilon, +\epsilon)^n \times \{0\}^{n-m},$$

where  $(-\epsilon, +\epsilon)^n$  is the cube. Noticeably, the  **$C^\infty$ -manifold structure** of  $\mathcal{M}$  is uniquely determined.

**Remark.** Let  $\mathcal{M} \subseteq \mathcal{N}$  be a **differentiable submanifold** of  $\mathcal{N}$ , and let  $\iota: \mathcal{M} \hookrightarrow \mathcal{N}$  be the inclusion. Then, for  $p \in \mathcal{M}$ ,  $T_p\mathcal{M}$  can be considered as subspace of  $T_p\mathcal{N}$ , namely as the image of  $d\iota(T_p\mathcal{M})$ .

**Lemma 1.5.3.** Let  $f: \mathcal{M}^m \rightarrow \mathcal{N}^n$  be a differentiable map such that  $m \geq n$  with  $p \in \mathcal{N}$ . Let  $df(x)$  has rank  $n$  for all  $x \in \mathcal{M}$  with  $f(x) = p$ . For the **submanifold**  $X = f^{-1}(p)$  and for  $q \in X$ , it is true that

$$T_q X = \ker df(q) \subseteq T_q \mathcal{M}.$$

# Chapter 2

## Riemannian Manifolds

### Lecture 5: Riemannian Manifolds

In this chapter, we start our discussion on [Riemannian manifolds](#).

19 Jan. 13:00

#### 2.1 Riemannian Metrics

We start by defining the [Riemannian metric](#).

**Definition 2.1.1** (Riemannian metric). A *Riemannian metric*  $g$  on a [differentiable manifold](#)  $\mathcal{M}$  is given by a scalar product  $I$  on each  $T_p\mathcal{M}$  which depends smoothly on the base point  $p$ .

**Definition 2.1.2** (Riemannian manifold). A *Riemannian manifold*  $(\mathcal{M}, g)$  is a [smooth manifold](#)  $\mathcal{M}$  equipped with a [Riemannian metric](#)  $g$ .

Let  $x = (x^1, \dots, x^d)$  be the [local coordinates](#). In these, a [metric](#) is represented by a positive definite symmetric matrix  $(g_{ij}(x))_{i,j=1,\dots,d}$ , i.e.,  $g_{ij} = g_{ji}$ , and  $g_{ij}\xi^i\xi^j > 0$  for all  $\xi = (\xi^1, \dots, \xi^d) \neq 0$  with coefficients smoothly depending on  $x$ .

##### 2.1.1 Transformation Behavior

We now see that the smoothness does not depend on [coordinates](#), i.e., the smooth dependence on the base point (as required in [Definition 2.1.1](#)) can be represented in the [local coordinates](#). Given 2 [tangent vectors](#)  $v, w \in T_p\mathcal{M}$  with [coordinate representations](#)  $(v^1, \dots, v^d), (w^1, \dots, w^d)$  given by  $x$  such that  $v = v^i \frac{\partial}{\partial x^i}$  and  $w = w^j \frac{\partial}{\partial x^j}$ , their product is

$$\langle v, w \rangle := g_{ij}(x(p))v^i w^j.$$

In particular,

$$\left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle = g_{ij}.$$

**Remark.** The length of  $v$  is given as  $\|v\| := \langle v, v \rangle^{1/2}$ .

Let  $y = f(x)$  define different [local coordinates](#). In these,  $v, w$  are given as

$$(\tilde{v}^1, \dots, \tilde{v}^d), (\tilde{w}^1, \dots, \tilde{w}^d)$$

with  $\tilde{v}^j = v^i \frac{\partial f^j}{\partial x^i}$  and  $\tilde{w}^j = w^i \frac{\partial f^j}{\partial x^i}$ . Denote the [metric](#) in new [coordinates](#)  $y$  by  $h_{k\ell}(y)$ , then we have

$$h_{k\ell}(f(x))\tilde{v}^k \tilde{w}^\ell = \langle v, w \rangle = g_{ij}(x)v^i w^j.$$

Plug everything in, we have

$$h_{k\ell}(f(x)) \frac{\partial f^k}{\partial x^i} \frac{\partial f^\ell}{\partial x^j} v^i w^j = g_{ij}(x)v^i w^j.$$

We see that this holds for any **tangent vectors**  $v, w$ , therefore,

$$h_{k\ell}(f(x)) \frac{\partial f^k}{\partial x^i} \frac{\partial f^\ell}{\partial x^j} = g_{ij}(x),$$

which is the transformation behavior under **coordinates changes**.

**Remark.** This shows that the smoothness does not depend on the choice of coordinates!

**Example.** Consider the Euclidean space  $\Omega$ , then given  $v, w \in T_p\Omega$ , we have

$$\langle v, w \rangle = \delta_{ij} v^i w^j = v^i w_i.$$

**Theorem 2.1.1.** Every **differentiable manifold** can be equipped with a **Riemannian metric**.

**Proof.** From **Lemma 1.2.2**, there exists a differentiable **partition of unity**  $\{f_\alpha\}$  of  $\mathcal{M}$  subordinate to a covering  $\{V_\alpha\}$  of  $\mathcal{M}$ . Consider the induced **metric**  $\langle \cdot, \cdot \rangle^\alpha$  of the system of **local coordinates** on each  $V_\alpha$ . Then, for every  $p \in M$ , a **Riemannian metric**  $\langle \cdot, \cdot \rangle_p$  can be defined naturally as

$$\langle u, v \rangle_p = \sum_\alpha f_\alpha(p) \langle u, v \rangle_p^\alpha$$

for all  $u, v \in T_p M$ . Given the fact that  $\{f_\alpha\}$  is the **partition of unity**, we know that

- (a)  $f_\alpha \geq 0$ , and  $f_\alpha = 0$  on  $\overline{V_\alpha}^c$ ,
- (b)  $\sum_\alpha f_\alpha(p) = 1$  for all  $p$  on  $\mathcal{M}$ ,

it's then immediate that the defined is indeed a **Riemannian metric**. ■

## 2.1.2 Isometry

After introducing any type of mathematical structure, we must introduce a notion of when two objects are the same, hence we now characterize  $g$ .

**Definition 2.1.3 (Isometry).** A **diffeomorphism**  $h: \mathcal{M} \rightarrow \mathcal{N}$  is an *isometry* between two **Riemannian manifolds** if it preserves the **Riemannian metric**, i.e., for  $p \in \mathcal{M}$ ,  $v, w \in T_p \mathcal{M}$ ,

$$\langle v, w \rangle_{\mathcal{M}} = \langle dh(v), dh(w) \rangle_{\mathcal{N}}.$$

**Definition 2.1.4 (Local isometry).** A **diffeomorphism**  $h: \mathcal{M} \rightarrow \mathcal{N}$  is a *local isometry* between two **Riemannian manifolds** if for every  $p \in \mathcal{M}$ , there exists a neighborhood  $U$  such that  $h|_U: U \rightarrow h(U): \mathcal{M} \rightarrow \mathcal{N}$  is an **isometry** and  $h(U) \subseteq \mathcal{N}$  is open.

It's common to say that a **Riemannian manifold**  $\mathcal{M}$  is **locally isometric** to a **Riemannian manifold**  $\mathcal{N}$  if for every  $p \in \mathcal{M}$ , there exists a neighborhood  $U$  of  $p$  in  $\mathcal{M}$  and a **local isometry**  $f: U \rightarrow f(U) \subseteq \mathcal{N}$ .

**Example (Euclidean space).** The *Euclidean space of dimension  $n$*   $\mathcal{M} = \mathbb{R}^n$  with  $\partial/\partial x_i$  identified with  $e_i = (0, \dots, 1, \dots, 0)$  is with the metric

$$\langle e_i, e_j \rangle = \delta_{ij}.$$

The Riemannian geometry of this space is metric Euclidean geometry.

**Example (Lie group).** See **Appendix A.2** for reference.

## 2.2 Geodesics

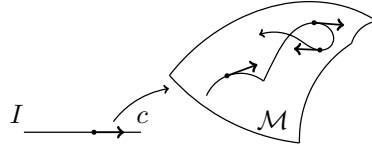
This is the first focus on the study of Riemannian geometry, i.e., the [geodesics](#). The up-shot is that a [geodesic](#) minimizes the [arc length](#) for points *sufficiently close* (in a sense to be made precise); in addition, if a [curve](#) minimizes [arc length](#) between any two of its points, it is a [geodesic](#).

### 2.2.1 Vector Fields along Curves

We are now going to show how a [Riemannian metric](#) can be used to calculate the [length](#) of a [curve](#).

**Definition 2.2.1 (Curve).** A (parametrized) *curve* is a differentiable mapping  $c: I \subseteq \mathbb{R} \rightarrow \mathcal{M}$  to a [smooth manifold](#)  $\mathcal{M}$ .

**Note.** A parametrized [curve](#) can admit self-intersections as well as corners.



**Definition 2.2.2 (Vector field along a curve).** A (smooth) *vector field*  $X$  along a curve  $c: I \subseteq \mathbb{R} \rightarrow \mathcal{M}$  on a [smooth manifold](#)  $\mathcal{M}$  is defined as  $X: I \rightarrow T\mathcal{M}$  such that  $X(t) \in T_{c(t)}\mathcal{M}$  for all  $t \in I$ .

**Notation.** The set of smooth [vector fields along](#)  $c$  is denoted as  $\chi_c(\mathcal{M})$ .

**Note.** To say  $V$  is differentiable means that for any differentiable function  $f$  on  $\mathcal{M}$ , the function  $t \mapsto V(t)f$  is a differentiable function on  $I$ .

**Example (Velocity field).** The [vector field along](#)  $c$   $dc/dt := dc(d/dt)$  is called the *velocity field* or *tangent vector field*.

**Remark.** A [vector field along](#)  $c$  can't necessarily be extended to a [vector field](#) on an open set of  $\mathcal{M}$ .

**Notation (Segment).** The restriction of a [curve](#)  $c$  to a closed interval  $[a, b] \subseteq I$  is called a *segment*.

### 2.2.2 Lengths and Energies

We're interested in the following two quantities.

**Definition.** Let  $\gamma: [a, b] \rightarrow \mathcal{M}$  be a [curve](#) on a [Riemannian manifold](#)  $(\mathcal{M}, g)$ .

**Definition 2.2.3 (Length).** The *length* of  $\gamma$  is defined as

$$L(\gamma) := \int_a^b \left\| \frac{d\gamma}{dt}(t) \right\| dt.$$

**Definition 2.2.4 (Energy).** The *energy* of  $\gamma$  is defined as

$$E(\gamma) := \frac{1}{2} \int_a^b \left\| \frac{d\gamma}{dt}(t) \right\|^2 dt.$$

We now want to compute  $L(\gamma)$ ,  $E(\gamma)$  in **local coordinates**. Let the **local coordinates** be

$$(x^1(\gamma(t)), \dots, x^d(\gamma(t))),$$

we write

$$\dot{x}^i(t) = \frac{d}{dt}(x^i(\gamma(t))).$$

Then, we have

$$L(\gamma) = \int_a^b \sqrt{g_{ij}(x(\gamma(t))) \dot{x}^i(t) \dot{x}^j(t)} dt, \quad E(\gamma) = \frac{1}{2} \int_a^b g_{ij}(x(\gamma(t))) \dot{x}^i(t) \dot{x}^j(t) dt.$$

**Definition 2.2.5 (Distance).** Given a **Riemannian manifold**  $(\mathcal{M}, g)$ , the *distance* between 2 points  $p, q \in \mathcal{M}$  is defined as

$$d(p, q) := \inf \{L(\gamma) \mid \gamma: [a, b] \rightarrow \mathcal{M} \text{ piecewise curve with } \gamma(a) = p, \gamma(b) = q\}.$$

**Note.** Any 2 points  $p, q \in \mathcal{M}$  can be connected by a piecewise **curve**, hence  $d(p, q)$  always exists.

**Corollary 2.2.1.** The topology of  $\mathcal{M}$  induced by the **distance function**  $d$  coincides with the original manifold topology of  $\mathcal{M}$ .

**Lemma 2.2.1.** If  $\gamma: [a, b] \rightarrow \mathcal{M}$  is a **curve**, and  $\psi: [\alpha, \beta] \rightarrow [a, b]$  is a reparametrization, then  $L(\gamma \circ \psi) = L(\gamma)$ .

**Proof.** This can be proved by computation, and the take-away is that the **length functional** is invariant under parameter changes. ■

### 2.2.3 Geodesic Equations as Euler-Lagrange Equations

We want to find a **curve** which minimizes the **length** between sufficiently close two points. It turns out that instead of working with **length** directly, we should work with **energy** instead.

**Notation.** Let's first fix some common notations.

(a)  $(g^{ij})_{i,j=1,\dots,d} = (g_{ij})_{i,j=1,\dots,d}^{-1}$ .<sup>a</sup>

(b)  $g_{j\ell,k} := \frac{\partial}{\partial x^k} g_{j\ell}$ .

<sup>a</sup>Technically,  $g^{-1}$  is not an inverse:  $g$  is a **(0, 2)-tensor field**, while  $g^{-1}$  is a **(2, 0)-tensor field**.

**Note.** In the above notations, we have  $g^{i\ell} g_{\ell j} = \delta_j^i$ .

And the following is particularly important.

**Notation (Christoffel symbol).** The *Christoffel symbol* is defined for all  $i$  as

$$\Gamma_{jk}^i := \frac{1}{2} g^{i\ell} (g_{j\ell,k} + g_{k\ell,j} - g_{j\ell,k}).$$

**Remark.** The notion of  $\Gamma$  is a bit cryptic at first, and we will come back to this after. Now, just treat it as a calculation tool.

The up-shot is that the **Euler-Lagrange equations** for the **energy**  $E$  has a nice form, and the solution of which has exactly the properties we want, hence we define it as **geodesics**.

**Proposition 2.2.1.** The **Euler-Lagrange equations** for the **energy**  $E$  are

$$\ddot{x}^i(t) + \Gamma_{jk}^i(x(t))\dot{x}^j(t)\dot{x}^k(t) = 0 \text{ for } i = 1, \dots, d. \quad (2.1)$$

**Proof.** The **Euler-Lagrange equations** of a functional<sup>a</sup>

$$I(x) = \int_a^b f(t, x(t), \dot{x}(t)) dt$$

are

$$\frac{d}{dt} \frac{\partial f}{\partial \dot{x}^i} - \frac{\partial f}{\partial x^i} = 0$$

for  $i = 1, \dots, d$ . Just by plugging in, we obtain for  $E$ , we have

$$\frac{d}{dt} (g_{ik}(x(t))\dot{x}^k(t) + g_{ji}(x(t))\dot{x}^j(t)) - g_{jk,i}(x(t))\dot{x}^j(t)\dot{x}^k(t) = 0$$

for  $i = 1, \dots, d$ . Hence,

$$g_{ik}\ddot{x}^k + g_{ji}\ddot{x}^j + g_{ik,\ell}\dot{x}^\ell\dot{x}^k + g_{ji,\ell}\dot{x}^\ell\dot{x}^j - g_{jk,i}\dot{x}^\ell\dot{x}^j = 0$$

Rename some indices and use  $g_{ij} = g_{ji}$ , we have that

$$2g_{\ell m}\ddot{x}^m + (g_{k\ell,j} + g_{j\ell,k} - g_{jk,\ell})\dot{x}^j\dot{x}^k = 0$$

for  $\ell = 1, \dots, d$ . Hence, we have

$$g^{i\ell}g_{\ell m}\ddot{x}^m + \frac{1}{2}g^{i\ell}(g_{\ell k,j} + g_{j\ell,k} - g_{jk,\ell})\dot{x}^j\dot{x}^k = 0$$

for  $i = 1, \dots, d$ . Finally, observe that  $g^{i\ell}g_{\ell m} = \delta_{im}$ , i.e.,  $g^{i\ell}g_{\ell m}\ddot{x}^m = \ddot{x}^i$ , hence the claim follows. ■

<sup>a</sup>The Lagrangian is  $\mathcal{L} = \frac{1}{2}g_{jk}\dot{x}^j\dot{x}^k$ .

Finally, we define the **geodesics** as the solution of **Equation 2.1**.

**Definition 2.2.6 (Geodesic).** A **curve**  $\gamma: [a, b] \rightarrow \mathcal{M}$  that obeys **Equation 2.1** is called a *geodesic*.

**Intuition.** **Geodesic** is the critical points of **energy**.<sup>a</sup>

<sup>a</sup>In fact, we can also start from **length** and get the same thing, which might be more natural.

## 2.2.4 Variation of Energies

We now discuss why **geodesic** is well-defined, i.e., we want to show that **Equation 2.1** has a unique solution. We solve this via the variational principal, and we first define the **action functional**.

**Definition 2.2.7 (Action functional).** Let  $\mathcal{L}$  be the Lagrangian, then the *action functional*

$$I[w(\cdot)] := \int_0^t \mathcal{L}(\dot{w}(s), w(s)) ds$$

is defined for functions  $w(\cdot) = (w^1(\cdot), \dots, w^n(\cdot))$  of the admissible class

$$\mathcal{A} = \{w(\cdot) \in C^2([0, t]; \mathbb{R}^n) \mid w(0) = y, w(t) = x\}.$$

**Example.** Both **length** and **energy** are **action functionals**.

From the calculus of variation, we can find a **curve**  $x(\cdot) \in \mathcal{A}$  such that  $I[x(\cdot)] = \min_{w(\cdot) \in \mathcal{A}} I[w(\cdot)]$ .

**Theorem 2.2.1** (Euler-Lagrangian equations). The solution  $x(\cdot)$  from  $I[x(\cdot)] = \min_{w(\cdot) \in \mathcal{A}} I[w(\cdot)]$  solves the system of **Euler-Lagrangian equations**

$$\frac{d}{ds} (D_{\dot{x}} \mathcal{L}(\dot{x}(s), x(s)) + D_x \mathcal{L}(\dot{x}(s), x(s))) = 0$$

for  $0 \leq s \leq t$ .

## Lecture 6: Geodesics and the Exponential Map

Now, we draw some relations between **length** and **energy** and see why starting from **energy** makes sense. 24 Jan. 13:00

**Proposition 2.2.2.** For all **curves**  $\gamma: [a, b] \rightarrow \mathcal{M}$ ,

$$\mathcal{L}(\gamma)^2 \leq 2(b-a)E(\gamma)$$

with equality if and only if  $\|d\gamma/dt\|$  is a constant.

**Proof.** From **Hölder's inequality**,

$$\int_a^b \left\| \frac{d\gamma}{dt} \right\| dt \leq (b-a)^{1/2} \left( \int_a^b \left\| \frac{d\gamma}{dt} \right\|^2 dt \right)^{1/2}$$

with equality if and only if  $\|d\gamma/dt\|$  is a constant. ■

**Example.** Let

$$\mathcal{L}(q, x) = \frac{1}{2}m|q|^2 - V(x)$$

with  $m > 0$ ,  $q = \dot{x}$ , the Euler-Lagrangian equations is given by  $m\ddot{x}(s) = F(x(s))$  for  $F := -DV$ .

Since regular curves can be parametrized by **arc length** with unit speed  $\|d\gamma/dt\| = \|\dot{\gamma}\| \equiv 1$ , the following is natural.

**Lemma 2.2.2.** Each **geodesic** is parametrized proportionally to the **arc length**, i.e.,  $\|\dot{\gamma}\|$  is a constant.

**Proof.** For a solution  $x(t)$  of  $\ddot{x}^i(t) + \Gamma_{jk}^i(x(t))\dot{x}^j(t)\dot{x}^k(t) = 0$  (i.e., the **geodesic**), we have

$$\frac{d}{dt} \langle \dot{x}, \dot{x} \rangle = \frac{d}{dt} (g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t)) = 0.$$

■

**Remark.** This is one of the advantages of working with the **energy** rather than the **length**.

Since the **length** and the **energy** functionals are invariants under parameter changes, it's enough to look at **curves** parametrized by **arc length**.

**Theorem 2.2.2.** Let  $\mathcal{M}$  be a **Riemannian manifold**,  $p \in \mathcal{M}$  and  $v \in T_p\mathcal{M}$ . Then there exists an  $\epsilon > 0$  and a unique **geodesic** such that  $c: [0, \epsilon] \rightarrow \mathcal{M}$  with  $c(0) = p$  and  $\dot{c}(0) = v$ . In addition,  $c$  smoothly depend on  $p, v$ .

**Proof.** Since **Equation 2.1** is a system of second order ODE, by **Picard-Lindelöf theorem**, we have local existence and uniqueness of the solution with prescribed initial values and derivative such that the solution depends smoothly on  $p, v$ . ■

If  $x(t)$  is the solution of **Equation 2.1**, then  $x(\lambda t)$  is also a solution for any constant  $\lambda \in \mathbb{R}$ . Denote **geodesic** from **Theorem 2.2.2** by  $c_v$ , then

$$c_v(t) = c_{\lambda v}(t/\lambda)$$



for  $\lambda > 0$ ,  $t \in [0, \epsilon]$ , and hence  $c_{\lambda v}$  defined on  $[0, \epsilon/\lambda]$ .

**Remark.** Since  $c_v$  depends smoothly on  $v$ , the set  $\{v \in T_p\mathcal{M} \mid \|v\| = 1\}$  is compact, hence there exists  $\epsilon_0 > 0$  such that for  $\|v\| = 1$ ,  $c_v$  defined at least on  $[0, \epsilon_0]$ , implying that for all  $w \in T_p\mathcal{M}$  with  $\|w\| \leq \epsilon_0$ ,  $c_w$  is defined at least on  $[0, 1]$ .

### 2.2.5 Exponential Maps and Normal Coordinates

The above discussion permits us to introduce the concept of the [exponential map](#) in the following manner.

**Definition 2.2.8 (Exponential map).** Let  $(\mathcal{M}, g)$  be a [Riemannian manifold](#),  $p \in \mathcal{M}$ , and  $V_p := \{v \in T_p\mathcal{M} \mid c_v \text{ defined on } [0, 1]\}$ . The *exponential map* of  $\mathcal{M}$  at  $p$ ,  $\exp_p: V_p \rightarrow \mathcal{M}$ , is defined as  $v \mapsto c_v(1)$ .

Clearly,  $\exp$  is differentiable, and we shall utilize the restriction of  $\exp$  to an open subset of the [tangent space](#)  $T_q\mathcal{M}$ , i.e., we define

$$\exp_p: B(0, \epsilon) \subseteq T_p\mathcal{M} \rightarrow \mathcal{M},$$

where  $B(0, \epsilon)$  is an open ball with center at the origin 0 of  $T_p\mathcal{M}$  of radius  $\epsilon$ . It's easy to see that  $\exp_p$  is differentiable and that  $\exp_p(0) = p$ .

**Intuition.** Geometrically,  $\exp_p(v)$  is a point of  $\mathcal{M}$  obtained by going out the [length](#) equal to  $|v|$ , starting from  $p$ , along a [geodesic](#) which passes through  $p$  with velocity equal to  $v/|v|$ .

**Proposition 2.2.3.** The [exponential map](#)  $\exp_p$  maps a neighborhood of  $0 \in T_p\mathcal{M}$  [diffeomorphically](#) onto a neighborhood of  $p \in \mathcal{M}$ .

**Proof.** We see that

$$d(\exp_p)_0(v) = \left. \frac{d}{dt} \exp_p(tv) \right|_{t=0} = \left. \frac{d}{dt} c_{tv}(1) \right|_{t=0} = \left. \frac{d}{dt} c_v(t) \right|_{t=0} = v,$$

i.e.,  $d(\exp_p)_0$  is the identity of  $T_q\mathcal{M}$ . By the inverse function theorem,  $\exp_p$  is a local [diffeomorphism](#) on a neighborhood of 0. ■

**Example.** Let  $\mathcal{M} = \mathbb{R}^n$ , then the [exponential map](#) is the identity.<sup>a</sup>

<sup>a</sup>With the usual identification of  $T_p\mathbb{R}^n$  at  $p$  with  $\mathbb{R}^n$ .

**Example.** Let  $\mathcal{M} = S^2$ .



Now we know that  $\exp_p: B(0, \epsilon) \subseteq T_p\mathcal{M} \rightarrow \mathcal{M}$  maps [diffeomorphically](#) onto its image, we then define the following.

**Definition 2.2.9 (Normal coordinate).** Given an [exponential map](#)  $\exp_p: B(0, \epsilon) \rightarrow \mathcal{M}$ , let  $(e_1, \dots, e_n)$  be the orthonormal basis of  $T_p\mathcal{M}$ . Then the associated [local coordinates](#) are the *normal coordinates*.

In this case, given  $p \in \mathcal{M}$ ,  $0 \in \mathbb{R}^n$ , for all  $i, j, k$ ,<sup>1</sup>

$$g_{ij}(p) = \delta_{ij}, \quad \Gamma_{ij}^k(p) = 0, \quad g_{ij,k} = 0.$$

**Intuition.** The first derivative vanishes, so locally, the manifold looks Euclidean.

**Note.** [FC13] introduces everything above using  $T\mathcal{M}$  instead of  $T_p\mathcal{M}$ .

## 2.3 Hopf-Rinow Theorem

With all the tools we have developed, we now want to characterize the minimizing property of geodesics.

### 2.3.1 Riemannian Polar Coordinates

A particular useful tool is the Riemannian polar coordinates, which is introduced as follows.

**Theorem 2.3.1.** For all  $p \in \mathcal{M}$ , there exists  $\rho > 0$  such that the Riemannian polar coordinates may be introduced on  $B(p, \rho) = \{q \in \mathcal{M} \mid d(p, q) \leq \rho\}$ . For any such  $\rho$  and  $q \in \partial B(p, \rho)$ , there exists a unique geodesic of shortest length ( $= \rho$ ) from  $p$  to  $q$ . In the polar coordinates, this geodesic is given by the straight line  $x(t) = (t, \varphi_0)$ ,  $0 \leq t \leq \rho$ , with  $q$  represented by coordinates  $(\rho, \varphi_0)$ ,  $\varphi_0 \in S^{d-1}$ .

**Proof.** Take an arbitrary curve from  $p$  to  $q$ , namely  $c(t) = (r(t), \varphi(t))$ ,  $0 \leq t \leq T$ , which does not have to be entirely contained in  $B(p, \rho)$ . Let  $t_0$  be defined as

$$t_0 := \inf \{t \leq T \mid d(x(t), p) \geq \rho\}.$$

Then  $t_0 \leq T$  such that  $c|_{[0, t_0]}$  lies entirely in  $B(p, \rho)$ . We want to show that

- (a)  $L(c|_{[0, t_0]}) \geq \rho$ , and
- (b)  $L(c|_{[0, t_0]}) = \rho$  only for a straight line in the polar coordinates,

where

$$L(c|_{[0, t_0]}) := \int_0^{t_0} \sqrt{g_{ij}(c(t)) \dot{c}^i \dot{c}^j} dt.$$

Observe that  $g_{r\varphi} = 0$ , with  $g_{\varphi\varphi}$  being positive definite and  $g_{rr} \equiv 1$ , we have

$$L(c|_{[0, t_0]}) \geq \int_0^{t_0} \sqrt{g_{rr}(c(t)) \dot{r}^2} dt = \int_0^{t_0} |\dot{r}| dt \geq \int_0^{t_0} \dot{r} dt = r(t_0) = \rho.$$

■

**Corollary 2.3.1.** Let  $\mathcal{M}$  be a compact Riemannian manifold. Then there exists  $\rho_0 > 0$  such that

- (a) for any  $p \in \mathcal{M}$ , Riemannian polar coordinates may be introduced on  $B(p, \rho_0)$ ;
- (b) for any  $p, q \in \mathcal{M}$  with  $d(p, q) \leq \rho_0$ , they can be connected by precisely one geodesic or shortest length which depends continuously on  $p$  and  $q$ .

## Lecture 7: Hopf-Rinow Theorem

### 2.3.2 Hopf-Rinow Theorem

26 Jan. 13:00

By using Corollary 2.3.1, we have shown the following in the homework.

<sup>1</sup>Note that this only holds at  $p$ . We will come back to this when we formally introduce the linear connection.

**Lemma 2.3.1.** Let  $(\mathcal{M}, g)$  be a compact Riemannian manifold. Then for all  $p \in \mathcal{M}$ , the exponential map  $\exp_p$  is defined on all of  $T_p\mathcal{M}$  and any geodesic may be extended indefinitely in each direction.

Then we use Lemma 2.3.1 to show the following.

**Theorem 2.3.2.** Let  $(\mathcal{M}, g)$  be a compact Riemannian manifold.

- (a) For any 2 points  $p, q \in \mathcal{M}$ , there exists a geodesic in every homotopy class of curves from  $p$  to  $q$ . Moreover, we can choose a shortest curve as the geodesic in the homotopy class.
- (b) Every homotopy class of closed curves in  $\mathcal{M}$  contains a curve that is shortest and geodesic.

We now want to generalize it. However, this is not true in the most general setting, and we need one more requirement.

**Definition 2.3.1** (Geodesically complete). A Riemannian manifold  $(\mathcal{M}, g)$  is *geodesically complete* if for all  $p \in \mathcal{M}$ ,  $\exp_p$  is defined on all of  $T_p\mathcal{M}$ .

In other words, a Riemannian manifold  $\mathcal{M}$  is *geodesically complete* if any geodesic  $c(t)$  with  $c(0) = p$  can be extended for all  $t \in \mathbb{R}$ . Then, we have the following.

**Theorem 2.3.3** (Hopf-Rinow theorem). Let  $(\mathcal{M}, g)$  be a Riemannian manifold, then the following statements are equivalent.

- (a)  $\mathcal{M}$  is complete as a metric space.<sup>a</sup>
- (b) The closed and bounded subsets of  $\mathcal{M}$  are compact.
- (c) There exists  $p \in \mathcal{M}$  such that  $\exp_p$  is defined on all  $T_p\mathcal{M}$ .
- (d)  $\mathcal{M}$  is geodesically complete.

Furthermore, (d) (and hence (a), (b), and (c)) implies

- (e) for two points  $p, q \in \mathcal{M}$  can be joined by a minimizing geodesic, i.e., geodesic of the shortest distance  $d(p, q)$ .

<sup>a</sup>Hence, equivalently, complete as a topological space w.r.t. the underlying topology.

**Proof.** We start by proving (d) implies (e). Let  $\mathcal{M}$  be geodesically complete, and let  $r := d(p, q)$ , and let  $\rho$  be as in Corollary 2.3.1. Let  $p_0 \in \partial B(p, \rho)$  be a point where the continuous functional  $d(q, \cdot)$  attains its minimum on the compact set  $\partial B(p, \rho)$ . Then, for some  $V \in T_{p_0}\mathcal{M}$ ,  $p_0 = \exp_{p_0} \rho V$ .

Consider the geodesic  $c(t) = \exp_{p_0} tV$ , by showing  $c(r) = q$ ,  $c|_{[0, r]}$  will be the shortest geodesic from  $p$  to  $q$ . We start by defining

$$I := \{t \in [0, r] \mid d(c(t), q) = r - t\},$$

and referring to the following diagram to guide us.



Now, we want to show that  $I = [0, r]$ , which will follow from showing that  $I$  is open.

**Note.**  $I$  is not empty since by definition it contains 0 and  $r$ . Further,  $I$  is closed by continuity.

Let  $t_0 \in I$ , and let  $\rho_1 > 0$  be the radius as in the corollary, without loss of generality,  $\rho_1 < r - t_0$ . Let  $p_1 \in \partial B(c(t_0), \rho_1)$  be the point where the continuous functional  $d(q, \cdot)$  attains its minimum on the compact set  $\partial B(c(t_0), \rho_1)$ . By the triangle inequality,

$$d(p, q) \leq d(p, p_1) + d(p_1, q).$$

For every curve  $\gamma$  from  $c(t_0)$  to  $q$ , there exists  $\gamma(t) \in \partial B(c(t_0), \rho_1)$ , hence

$$L(\gamma) \geq d(c(t_0), \gamma(t)) + d(\gamma(t), q) = \rho_1 + d(p_1, q) \Rightarrow d(q, c(t_0)) \geq \rho_1 + d(p_1, q)$$

where we use  $d(c(t_0), \gamma(t)) = \rho_1$ . But from the triangle inequality, we actually have

$$d(q, c(t_0)) = \rho_1 + d(p_1, q) \Leftrightarrow d(p_1, q) = \underbrace{d(q, c(t_0))}_{r-t_0} - \rho_1,$$

hence  $d(p_1, p) \geq r - (r - t_0 - \rho_1) = t_0 + \rho_1$ , i.e., this is a minimizing curve!

On the other hand, there exists a curve from  $p$  to  $p_1$  of length  $t_1 + \rho_1$  since it's composed by the portion from  $p$  to  $c(t_0)$  along  $c(t)$  and the portion being the **geodesic** from  $c(t_0)$  to  $p_1$  of length  $\rho_1$ . Then, by [Theorem 2.3.2](#), this curve is a **geodesic** curve. Finally, from the uniqueness of **geodesic** with the given extra data, this **geodesic** coincides with  $c$ . Hence,  $p_1 = c(t_0 + \rho_1)$ . With  $d(p_1, q) = r - t_0 - \rho_1$ ,

$$d(c(t_0 + \rho_1), q) = d(p_1, q) = r - t_0 - \rho_1 = r - (t_0 + \rho_1),$$

so  $t_0 + \rho_1 \in I$ , implying that  $I$  is open, i.e.,  $I = [0, r]$ , so  $c(r) = q$  follows.<sup>a</sup>

<sup>a</sup>For a detailed proof, see [\[FC13, Corollary 3.9\]](#).

## Lecture 8: Injectivity Radius and Vector Bundles

Let's finish the proof of [Hopf-Rinow theorem](#).

31 Jan. 13:00

**Proof of Theorem 2.3.3 (Continued).** We see that (d) implies (e), hence we only need to show that (a), (b), (c), and (d) are equivalent.

- (d)  $\Rightarrow$  (c): It is trivial.
- (c)  $\Rightarrow$  (b): Let  $K \subseteq \mathcal{M}$  be closed and bounded. As  $K$  bounded,  $K \subseteq B(p, r)$  for some  $r > 0$ . Then any point in  $B(p, r)$  can be joined with  $p$  by **geodesic** of length  $\leq r$ , and  $B(p, r)$  is the image of the compact ball in  $T_p \mathcal{M}$  of radius  $r$  under continuous map  $\exp_p$ , hence  $B(p, r)$  is compact. As  $K$  closed and  $K \subseteq B(p, r)$ ,  $K$  is compact.
- (b)  $\Rightarrow$  (a): Let  $(p_n)_{n \in \mathbb{N}} \subseteq \mathcal{M}$  be a Cauchy sequence, so it's bounded, and by (b), its closure is compact. It contains a convergent subsequence, so it converges, i.e.,  $\mathcal{M}$  is **complete**.
- (a)  $\Rightarrow$  (d): Let  $c$  be a **geodesic** in  $\mathcal{M}$ , parametrized by arc length defined on a maximal interval  $I$ . Since  $I$  is non-empty, and we can show that  $I$  is both open and closed.

■

**Remark.** It's worth mentioning that we do have uniqueness after choosing  $p_0$ ,<sup>a</sup> so the non-uniqueness really comes from the initial choice of  $p_0$ .

<sup>a</sup>In other words, after choosing  $p_0$ , everything is fixed.

**Example.** Consider  $S^2$ , after fixing  $p_0$ ,  $c(t_0)$  is extended uniquely.



### 2.3.3 Injectivity Radius

One might wonder, though we have [Lemma 2.3.1](#), how far can  $\exp_p$  extend while maintaining injectivity?

**Definition 2.3.2** (Injectivity radius). Let  $\mathcal{M}$  be a [Riemannian manifold](#), and  $p \in \mathcal{M}$ . The *injectivity radius*  $i(p)$  of  $p$  is

$$i(p) := \sup \{ \rho > 0 \mid \exp_p \text{ defined on } B(0, \rho) \subseteq T_p \mathcal{M} \text{ and injective} \}.$$

Similarly, the *injectivity radius*  $i(\mathcal{M})$  of  $\mathcal{M}$  is defined as  $i(\mathcal{M}) := \inf_{p \in \mathcal{M}} i(p)$ .

**Example** (Sphere).  $i(S^n) = \pi$ .

**Example** (Torus).  $i(T^n) = 1/2$ .

**Remark.** Any [manifold](#) carries a [complete Riemannian metric](#). If  $(\mathcal{M}, g_1)$  is not [complete](#), we can find  $g_2$  such that  $(\mathcal{M}, g_2)$  is [complete](#).

**Example** (Hyperbolic half-plane). The half-plane  $P = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$  equipped with metric induced by the Euclidean metric on  $\mathbb{R}^2$ , which is not [complete](#).

However, it becomes [complete](#) when equipped with the following metric

$$\frac{1}{y^2}(dx^2 + dy^2).$$

In fact,  $P$  with the above metric is called the *hyperbolic half-plane*  $\mathbb{H}^2$ , and we can extend it to  $\mathbb{H}^n$ .

Another question we may ask is the following.

**Problem.** Is the converse of [Hopf-Rinow theorem](#) true? I.e., can we show that (e) implies (d)?

**Answer.** No! Any 2 points in the open half-sphere can be joined by a unique minimal [geodesic](#), but this manifold is not [geodesically complete](#). ⊗

**Example.** The [injectivity radius](#) of  $H^n$  is  $\infty$ .

**Remark.** Given a compact  $\mathcal{M}$ , the [injectivity radius](#) is always  $> 0$  by continuity argument.

Now, given a [complete](#) but not compact  $\mathcal{M}$ , the [injectivity radius](#) can be 0.

**Example.** Take the quotient of the Poincaré half-plane by the translations

$$(x, y) \mapsto (x + n, y), \quad n \in \mathbb{Z}.$$

We then obtain a [complete Riemannian manifold](#)  $\mathcal{M}$  with  $i(\mathcal{M}) = 0$ .

**Note.** Finding lower bounds for  $i(\mathcal{M})$  introduces curvature estimates.

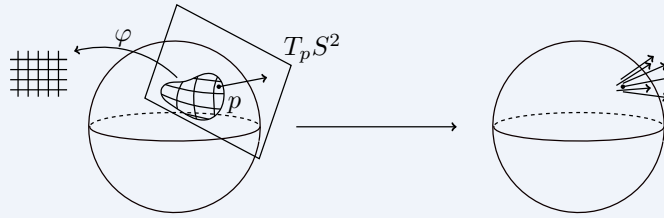
## 2.4 Vector Bundles and Tensor Fields

We now introduce the theory of **bundles**, which allows us to introduce **vector fields**, and hence **tensor fields**. Noticeably, nearly every structure we can put on a **Riemannian manifold** will be in the form of **tensor fields**, which is why we care about them.

As a motivating example, recall the **tangent bundle**<sup>2</sup>  $(T\mathcal{M}, \pi, \mathcal{M})$ , which captures the idea of “for every  $p \in \mathcal{M}$ , we associate a space  $T_p\mathcal{M}$ ”.

**Intuition.** This helps us construct **tangent vector fields** since a **tangent vector field**  $X$  of  $\mathcal{M}$  is defined by associating  $p$  to a **tangent vector**  $X(p)$  in the associated **tangent space**  $T_p\mathcal{M}$ .

**Example.** The **tangent vector field** assigns every  $p \in S^2$  a “point” in the associated **tangent space**.



### 2.4.1 Bundles

The above example of **tangent bundle** generalizes quite easily for defining a general **bundle**.

**Definition 2.4.1 (Bundle).** A *bundle* is a tuple  $(E, \pi, \mathcal{M})$  consists of the **total space**  $E$ , the **base space**  $\mathcal{M}$ , and the **bundle projection**  $\pi: E \rightarrow \mathcal{M}$ .

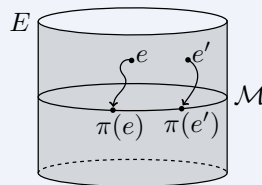
**Definition 2.4.2 (Total space).** The **differentiable manifold**  $E$  is called the *total space*.

**Definition 2.4.3 (Base space).** The **differentiable manifold**  $\mathcal{M}$  is called the *base space*.

**Definition 2.4.4 (Bundle projection).** The (differentiable) continuous surjection  $\pi: E \rightarrow \mathcal{M}$  is called the *bundle projection*.

**Example.** A **tangent bundle**  $(T\mathcal{M}, \pi, \mathcal{M})$  is a **bundle**.

**Example.** Let  $E$  be a cylinder,  $\mathcal{M}$  be a circle.



As we can see, the number of possible  $\pi$  is enormous, as long as it's surjective and smooth.

<sup>2</sup>Where just use the name “**bundle**” and don't know what it is. We'll see now!

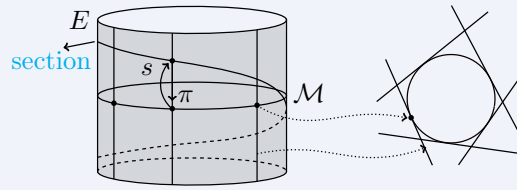
**Notation.** Sometimes, we will just denote a **bundle** as  $E \xrightarrow{\pi} \mathcal{M}$ , or even more compactly, just  $\pi$  since it captures all the data.

**Definition 2.4.5 (Fiber).** Given a **bundle**  $(E, \pi, \mathcal{M})$ , the *fiber* over  $p \in \mathcal{M}$  under  $\pi$  is the preimage of a  $\{p\}$ , i.e.,  $\pi^{-1}(\{p\})$ .

**Definition 2.4.6 (Section).** A *section* of a **bundle**  $(E, \pi, \mathcal{M})$  is a differentiable map  $s: \mathcal{M} \rightarrow E$  such that  $\pi \circ s = \text{id}_{\mathcal{M}}$ .

**Remark.** We see that a **section**  $s$  encodes lots of information of a **bundle**, since  $s$  includes  $E, \mathcal{M}$ , and the condition deal with  $\pi$ .

**Example.** Again let  $E$  be a cylinder,  $\mathcal{M}$  be a circle. This time, we choose  $\pi$  to be the trivial one.



We see that in this way, this **bundle** really captures all the **tangent spaces** structure of a circle!

## 2.4.2 Vector Bundles

Then, we're interested in the so-called **vector bundle**.

**Definition 2.4.7 (Vector bundle).** A (differentiable) *vector bundle* of rank  $n$  is a **bundle**  $(E, \pi, \mathcal{M})$  such that each **fiber**  $E_x := \pi^{-1}(x)$  of  $x \in \mathcal{M}$  carries a structure of an  $n$ -dimensional (real) vector space, and **local triviality** condition holds.

**Definition 2.4.8 (Local trivialization).** For all  $x \in \mathcal{M}$ , the *local trivialization*  $(U, \varphi)$  consists a neighborhood  $U$  and **diffeomorphism**  $\varphi: \pi^{-1}(U) \rightarrow U \times \mathbb{R}^n$  such that for all  $y \in U$ ,

$$\varphi_y := \varphi|_{E_y} : E_y \rightarrow \{y\} \times \mathbb{R}^n$$

is a vector space isomorphism.



Figure 2.1: An illustration of **vector bundle**  $(E, \pi, \mathcal{M})$ .

**Definition 2.4.9 (Trivial).** A **vector bundle** is *trivial* if it's isomorphic to  $\mathcal{M} \times \mathbb{R}^n$ .<sup>a</sup>

<sup>a</sup> $n$  is the rank of the **vector bundle**.

**Intuition.** Local trivialization shows that *locally*  $\pi$  looks like the **projection** of  $U \times \mathbb{R}^n$  on  $U$ .

**Definition 2.4.10** (Bundle chart). The pair  $(\varphi, U)$  is the *bundle chart* in **local trivialization**.

**Remark.** From **Definition 2.4.7**, **vector bundle** is locally, but not necessarily globally a product of **base space** and the **fiber**.

**Intuition.** We may look at a **vector bundle** as a family of vector spaces, all isomorphic to a fixed  $\mathbb{R}^n$ , “parametrized” (**locally trivially**) by a **manifold**.

### 2.4.3 Vector Fields

We can now introduce the notion of **vector field** in terms of **section**.

**Definition 2.4.11** (Vector field). A (smooth) *vector field*  $X$  is a smooth **section** of a **vector bundle**.

**Note.** We see that a smooth **tangent vector field** is indeed a smooth **vector field** with the **bundle** being the **tangent bundle**.

**Notation.** Since we will nearly always be talking about **tangent vector fields**, we will abuse the notation a bit and just simply call it **vector fields**. But always keep in mind that more broadly, a **vector field** should be a **section** of a **vector bundle**, not always  $T\mathcal{M}$ .

## Lecture 9: Tensors and Connections

### 2.4.4 Tensor Fields

2 Feb. 13:00

We can introduce the notion of “**tensor fields**” in a brute-force way<sup>3</sup> by following a similar path of how we define **vector field**, i.e., we first define some **bundle** called **tensor bundle**, and then the **tensor field** is just a smooth **section** of which. But first, it might be beneficial to see how does a **tensor** look like.

**Definition 2.4.12** (Tensor). Let  $V$  be a vector space of dimension  $m < \infty$ , and the dual space  $V^*$ . The vector space of the *r-times contravariant and s-times covariant tensors over  $V$* , denoted as  $T_s^r(V)$ , is defined as

$$T_s^r(V) = \{T: \underbrace{V^* \times \cdots \times V^*}_r \times \underbrace{V \times \cdots \times V}_s \rightarrow \mathbb{R}\} = (V^*)^{\otimes r} \otimes V^{\otimes s}.$$

**Intuition.** Just as **vector field**, we’re trying to assign a vector on every point  $p \in \mathcal{M}$ . Here, we’re trying to assign a **tensor** on every point  $p \in \mathcal{M}$ , which is just an element in  $(V_p^*)^{\otimes r} \otimes V_p^{\otimes s}$ .

As one might imagine, since a **tensor** is an element in  $(V_p^*)^{\otimes r} \otimes V_p^{\otimes s}$ , the corresponding **tensor bundle** is defined as follows.

**Definition 2.4.13** (Tensor bundle). A *(r, s)-tensor bundle*  $T_s^r\mathcal{M} = \bigcup_{p \in \mathcal{M}} T_s^r(T_p\mathcal{M}) = T_s^r(T\mathcal{M})$  on  $\mathcal{M}$  is a **fiber bundle** where the **fiber** is the tensor product of  $s$  **tangent spaces** and  $r$  **cotangent spaces**.

**Remark.** It’s clear that one can also define **tensor bundle** in terms of a general **vector bundle**  $V$  on  $\mathcal{M}$  instead of the **tangent bundle**  $T\mathcal{M}$  specifically.

<sup>3</sup>See **Appendix A.1** for another view point.



So in a **tensor bundle**, the **fiber** is a vector space and the **tensor bundle** is a special kind of **vector bundle**.<sup>4</sup> Finally, the **tensor field** is defined as follows.

**Definition 2.4.14** (Tensor field). A  $(r, s)$ -*tensor field* is a **section** of a  $(r, s)$ -**tensor bundle**.

A convenient notation is the following.

**Notation.** Let  $\mathcal{M}^n$  be a **smooth manifold** and  $\pi: E \rightarrow \mathcal{M}$  a **smooth vector bundle**, then the set of **sections** is denoted as

$$\Gamma(E) := \{s \in C^\infty(\mathcal{M}, E) \mid \pi \circ s = \text{id}_{\mathcal{M}}\}.$$

Then, we see the following.

**Example.** Consider the **vector bundle**  $(T\mathcal{M}, \pi, \mathcal{M})$ , then  $\Gamma(T\mathcal{M}) := \{\text{vector fields on } \mathcal{M}\}$ .

**Example.** A  $(r, s)$ -**tensor field** on  $\mathcal{M}$  is then equivalently defined as an element in  $\Gamma(T_s^r \mathcal{M})$ .

**Notation.** For  $s \in \mathbb{N}$ , let  $\Lambda^s(V^*) := \{A \in T_s^0(V) \mid A \text{ skew-symmetric}\}$ .

**Example.**  $\Gamma(\Lambda_s \mathcal{M}) := \{s\text{-forms on } \mathcal{M}\}$  with  $\Lambda_s \mathcal{M} = \Lambda^s \left( \bigcup_{p \in \mathcal{M}} T_p^* \mathcal{M} \right)$ .

**Example.** A **Riemannian metric**  $g$  on  $\mathcal{M}$  is a  $(0, 2)$ -**tensor field**, i.e.,  $g \in \Gamma(T_2^0(\mathcal{M}))$  for all  $p \in \mathcal{M}$ .

**Proof.** Since  $g_p: T_p \mathcal{M} \times T_p \mathcal{M} \rightarrow \mathbb{R}$ , so by regarding  $p$  as the argument of the map  $g$ ,  $g: \Gamma(T\mathcal{M}) \times \Gamma(T\mathcal{M}) \rightarrow C^\infty(\mathcal{M})$ . ⊗

**Note.** It's in fact unnecessary to have such a general **Definition 2.4.14** on a **Riemannian manifold**.

**Proof.** Since given a **Riemannian metric**  $g$ , it associates to each  $X \in \Gamma(T\mathcal{M})$  a unique  $\omega \in \Gamma(T^* \mathcal{M})$  given by  $\omega(Y) = g(X, Y)$  for all  $X, Y \in \Gamma(T\mathcal{M})$ . ⊗

## 2.5 Other Metrics

Finally, we conclude this chapter by introducing some other metrics a **manifold** can equip with.

**Definition 2.5.1** (Pseudo-Riemannian metric). A *pseudo-Riemannian metric* on a **differentiable manifold**  $\mathcal{M}$  is a  $(0, 2)$ -**tensor field**  $g \in \Gamma(T_2^0(\mathcal{M}))$  for all  $p \in \mathcal{M}$  with

- (a)  $g(X, Y) = g(Y, X)$  for all  $X, Y \in T\mathcal{M}$ ;
- (b) for all  $p \in \mathcal{M}$ ,  $g_p$  is non-degenerate bilinear form on  $T_p \mathcal{M}$ , i.e.,  $g_p(X, Y) = 0$  for all  $X, Y \in T_p \mathcal{M}$  if and only if  $Y = 0$ .

**Note.** A **pseudo Riemannian metric** is a **Riemannian metric** if it's positive definite at every  $p \in \mathcal{M}$ .

**Definition 2.5.2** (Lorentzian metric). A *Lorentzian metric*  $g$  is a continuous assignment of a non-degenerate<sup>a</sup> quadratic form  $g_p$  of index 1<sup>b</sup> in  $T_p \mathcal{M}$  for all  $p \in \mathcal{M}$ .

<sup>a</sup> $g_p(X, Y) = 0$  for all  $Y \in T_p \mathcal{M}$  implies  $X = 0$ .

<sup>b</sup>It means that the maximal dimension of a subspace of  $T_p \mathcal{M}$  on which  $g_p$  is negative definite is 1.

<sup>4</sup>There are **vector bundles** which are not **tensor bundles**.

An equivalent definition is the following.

**Definition 2.5.3 (Lorentzian).** A quadratic form  $g_p$  in  $T_p\mathcal{M}$  is *Lorentzian* if there exists a vector  $V \in T_p\mathcal{M}$  such that  $g_p(V, V) < 0$  while setting  $\Sigma_V = \{X \mid g_p(X, V) = 0\}$  such that  $g_p|_{\Sigma_V}$ <sup>a</sup> is positive definite.

<sup>a</sup>The  $g_p$ -orthogonal complement of  $V$ .

**Example (Minkowski space).** The Minkowski space on  $\mathbb{R}^4$  is the prototypical example from physics (flat spacetime). Namely, the metric is given by the quadratic form

$$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

with the coordinates being  $(t, x, y, z)$ .

## Chapter 3

# Connections and Curvatures

So far, we saw that a [vector field](#)  $X$  can be used to provide a directional derivative since it gives us a [tangent vector](#) at each point smoothly. Now, we will introduce a new symbol  $\nabla$  where for  $f \in C^\infty(\mathcal{M})$ ,

$$\nabla_X f := Xf.$$

**Problem.** Does this notation overkill? We already know that  $Xf = (df)(X)$ !

**Answer.** No! While  $\nabla, X: C^\infty(\mathcal{M}) \rightarrow C^\infty(\mathcal{M})$ , while  $df: \Gamma(TM) \rightarrow C^\infty(\mathcal{M})$ , we can generalize  $\nabla_X$  to act from [vector fields](#) to [vector fields](#)! The insight is that if  $X$  can be extended naturally (without providing any extra structures), then we certainly won't bother introducing a new symbol. However, as you might guess, to let  $\nabla$  doing this, we do need to provide extra structures, and  $\nabla$  stands exactly for these extra structures!  $\circledast$

In some sense, this new notions  $\nabla$  allows us to “connect” [tangent spaces](#), which allows us to make sense of “curvatures” and other geometric property of a [Riemannian manifold](#).

### 3.1 Levi-Civita Connections

We start by talking about [linear connections](#), and then realize that after specifying a [Riemannian metric](#)  $g$ , with an additional (technical) assumption, a unique [linear connection](#), defined as [Levi-Civita connections](#), exists for any [Riemannian manifold](#). In other words, specifying  $g$  is the same as specifying the “shape of the space.” We'll make sense of all these on the way.

#### 3.1.1 Affine Connections

We first formulate a *wish list* of properties which the  $\nabla_X$  should have. Any remaining freedom in choosing  $\nabla$  will need to be provided as additional structures beyond the structures on  $\mathcal{M}$  we already have.

**Definition 3.1.1** (Linear connection). A *linear connection* (or *affine connection*) on a [smooth manifold](#)  $\mathcal{M}$  is a bilinear map

$$\nabla: \Gamma(TM) \times \Gamma(TM) \rightarrow \Gamma(TM),$$

which is denoted by  $\nabla(X, Y) = \nabla_X Y$  and which satisfies

- (a)  $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$ ;
- (b)  $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z$ ;
- (c)  $\nabla_X fY = f\nabla_X Y + X(f)Y$ ;

for all [vector fields](#)  $X, Y, Z \in \Gamma(TM)$  and  $f, g \in C^\infty(\mathcal{M})$ .

**Remark.** Definition 3.1.1 (c) shows that this is actually a local notion as we will see.

**Note.** There's a similar notation called [covariant derivative](#), denoted by  $D$ , satisfies similar properties as a [linear connection](#). Hence, we often write  $D$  and  $\nabla$  interchangeably.<sup>a</sup>

<sup>a</sup> $\nabla$  is more general than  $D$ ; however, we treat them as the same as suggested by [Proposition 3.4.1](#).

Now, one might be wondering that, after fixing these rules we want, how much freedom is left? To see this, let's first do some calculations...

### 3.1.2 Connection Coefficients

Choose a [system of coordinates](#)  $(x_1, \dots, x_n)$  at  $p \in \mathcal{M}$ , we can write  $X = X^i \frac{\partial}{\partial x_i}$ ,  $Y = Y^j \frac{\partial}{\partial x_j}$ , then

$$\nabla_X Y = \nabla_{X^i \frac{\partial}{\partial x_i}} \left( Y^j \frac{\partial}{\partial x_j} \right) = X^i Y^j \nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} + X^i \frac{\partial}{\partial x_i} (Y^j) \frac{\partial}{\partial x_j}.$$

Now, we see that  $\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j}$  is another [vector field](#), hence can again write

$$\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} =: \sum_k \Gamma_{ij}^k \frac{\partial}{\partial x_k}$$

in terms of the basis with a new set of coefficients  $\Gamma$ .

**Notation** (Connection coefficient). The coefficients  $\Gamma_{ij}^k$  is called the *connection coefficients*.

**Intuition.**  $\Gamma$  are the *corrections* to an ordinary derivative on a “curved” [manifold](#) w.r.t.  $\nabla$ .

It's tempting to say that the [connection coefficients](#) are the same as [Christoffel symbols](#) since we're using the same symbols. Indeed, they are the same if  $\nabla$  is chosen to be the [Levi-Civita connection](#).<sup>1</sup>

**Note.** It's clear that  $\Gamma_{ij}^k$  are differentiable and [charts](#)-dependent and hence  $\nabla$  is local.

Finally, in a particular domain  $U$ , we have

$$\nabla_X Y = (X^i Y^j \Gamma_{ij}^k + X(Y^k)) \frac{\partial}{\partial x_k} \Rightarrow (\nabla_X Y)^k = X(Y^k) + \Gamma_{ij}^k X^i Y^j,$$

meaning that we have  $(\dim \mathcal{M})^3$  many  $\Gamma$ 's (freedom) when choosing  $\Gamma_{ij}^k$  with [Definition 3.1.1](#).

**Remark.** One might ask what about other [tensor fields](#)? Fortunately, the same set of  $\Gamma$ 's fix the action of  $\nabla$  on any [tensor fields](#).

**Proof.** The key observation is that if we define  $\nabla_{\frac{\partial}{\partial x^j}} (dx^i) =: \Sigma_{jk}^i dx^k$ , then

$$\nabla_{\frac{\partial}{\partial x^j}} \left( dx^i \left( \frac{\partial}{\partial x^k} \right) \right) = \begin{cases} \frac{\partial}{\partial x^j} (\delta_k^i) = 0; \\ \left( \nabla_{\frac{\partial}{\partial x^j}} dx^i \right) \frac{\partial}{\partial x^k} + dx^i \left( \underbrace{\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k}}_{\Gamma_{jk}^\ell \frac{\partial}{\partial x^\ell}} \right), \end{cases}$$

and since  $dx^i \frac{\partial}{\partial x^i} = \delta_\ell^i$ , the above leads to

$$\left( \nabla_{\frac{\partial}{\partial x^j}} dx^i \right) \frac{\partial}{\partial x^k} = -dx^i \left( \Gamma_{jk}^\ell \frac{\partial}{\partial x^\ell} \right) \Rightarrow \left( \nabla_{\frac{\partial}{\partial x^j}} dx^i \right)_k = -\Gamma_{jk}^i.$$

⊗

In summary, we have

$$\begin{cases} (\nabla_X Y)^k = X(Y^k) + \Gamma_{ij}^k X^i Y^j, & \text{if } Y \text{ is a vector field;} \\ (\nabla_X \omega)_k = X(\omega_k) - \Gamma_{ik}^j X^i \omega_j, & \text{if } \omega \text{ is a co-vector field.} \end{cases}$$

<sup>1</sup>As we will soon see, it means the [torsion free](#) and [Riemannian connection](#). See [this](#) for a more detailed explanation.

### 3.1.3 Levi-Civita Connections

The basic insight is that, after choosing a particular [connection](#),<sup>2</sup> the space is basically fixed: i.e., the *shape*, or “curvature”, of the space is determined by the choice of  $\nabla$ ! We now formalize this idea. A particularly natural notion related to “curvature” is the [torsion](#), defined as follows.

**Definition 3.1.2 (Torsion).** The *torsion*  $T$  of a [linear connection](#)  $\nabla$  is the [\(1,2\)-tensor field](#)

$$T(\omega, X, Y) := \omega(\nabla_X Y - \nabla_Y X - [X, Y]).$$

**Notation.** We usually write this as  $T(X, Y)$  by neglecting  $\omega$ .

**Remark.**  $T$  is actually  $C^\infty$ -linear in each entry,<sup>a</sup> hence a [tensor field](#).

<sup>a</sup>See [Appendix A.1](#).

**Proof.** Since  $T(f \cdot \omega, X, Y) = f \cdot \omega(\dots) = fT(\omega, X, Y)$  and  $T(\omega + \psi, X, Y) = \dots = T(\omega, X, Y) + T(\psi, X, Y)$ , and also

$$\begin{aligned} T(\omega, fX, Y) &= \omega(\nabla_{fX} Y - \nabla_Y(fX) - [fX, Y]) \\ &= \omega(f\nabla_X Y - (Yf)X - f\nabla_Y X - f[X, Y] + (Yf)X) = f \cdot T(\omega, X, Y) \end{aligned}$$

since

$$([fX, Y])g = f \cdot X(Yg) - Y(f \cdot Xg) = f \cdot X(Yg) - (Yf)(Xg) - f \cdot Y(Xg) = (f \cdot [X, Y] - (Yf)X)g.$$

Finally, we claim that the additivity at  $X$  holds, with  $T(\omega, X, Y) = -T(\omega, Y, X)$ , we’re done.  $\circledast$

**Intuition.** [Definition 3.1.2](#) makes sense (in such a form) since this will make  $T$  actually a [tensor field](#). For example, without the [Lie bracket](#) term, we don’t have the linearity at  $X$  (hence  $Y$ ).

**Definition 3.1.3 (Torsion-free).** A [linear connection](#)  $\nabla$  is *torsion-free* if  $T = 0$ .

**Notation** (symmetric). A [torsion-free](#)  $\nabla$  is sometimes said to be *symmetric*.

In a [chart](#),

$$T_{jk}^i := T\left(dx^i, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right) = \Gamma_{jk}^i - \Gamma_{kj}^i = 2\Gamma_{[jk]}^i,$$

hence if  $T = 0$ , we can interchange the lower two indexes of  $\Gamma_{ij}^k$ , i.e.,  $\Gamma_{ij}^k = \Gamma_{ji}^k$ .

**Definition 3.1.4 (Riemannian).** Let  $\nabla$  be a [linear connection](#) and  $g$  be a [Riemannian metric](#) on  $\mathcal{M}$ . Then  $\nabla$  is *Riemannian* (or *metric*) if for all  $X, Y, Z \in \Gamma(T\mathcal{M})$ ,<sup>a</sup>

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

<sup>a</sup>We view  $g(X, Y) \in C^\infty(\mathcal{M})$  as suggested by [Appendix A.1](#).

**Notation** (Compatible). A [Riemannian](#)  $\nabla$  is sometimes said to be *compatible*.

**Remark.** Equivalently, [Definition 3.1.4](#) can be formulated as  $\nabla g = 0$ .

We are now able to state the fundamental theorem of this section.

<sup>2</sup>Remember that we have freedom to choose  $\Gamma$ ’s.

**Theorem 3.1.1 (Levi-Civita).** On each Riemannian manifold  $(\mathcal{M}, g)$ , there exists a unique Riemannian, torsion-free connection  $\nabla$  on  $T\mathcal{M}$  determined by the Koszul formula

$$\langle \nabla_X Y, Z \rangle = \frac{1}{2} (X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle - \langle X, [Y, Z] \rangle + \langle Y, [Z, X] \rangle + \langle Z, [X, Y] \rangle). \quad (3.1)$$

**Proof sketch.** Firstly, we show that every Riemannian and torsion-free connection satisfies Koszul formula, which implies uniqueness. For existence, we verify that the unique map  $\nabla: \Gamma(T\mathcal{M}) \times \Gamma(T\mathcal{M}) \rightarrow \Gamma(T\mathcal{M})$  given by Koszul formula is Riemannian and torsion-free.<sup>a</sup> ■

<sup>a</sup>For a detail proof, see [FC13, §2 Theorem 3.6].

**Note.** I rearrange the Koszul formula to make it easier to memorize.

Finally, we define the following.

**Definition 3.1.5 (Levi-Civita connection).** The Levi-Civita connection is the unique linear connection  $\nabla$  defined by the Koszul formula.

**Remark.** This means, given a Riemannian metric  $g$ , with the condition of torsion-free, the shape of the space is also fixed since there's a unique linear connection  $\nabla$  such that  $T = \nabla g = 0$ .

## Lecture 10: Curvatures and Flow of Vector Fields

### 3.2 Riemannian Curvatures

7 Feb. 13:00

Given all these definitions, we can now introduce the notion of “curvatures.” Consider the following.<sup>3</sup>

**Definition 3.2.1 (Riemannian curvature).** The Riemannian curvature  $R$  of a Levi-Civita connection  $\nabla$  is the  $(1, 3)$ -tensor field<sup>a</sup>

$$R(\omega, Z, X, Y) := \omega (\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z).$$

<sup>a</sup> $R$  is indeed  $C^\infty$ -linear in each entry (see Appendix A.1) although we omit the proof here.

**Notation.** We usually write this as  $R(X, Y)Z$  by emphasizing  $Z$  and neglecting  $\omega$ .

**Example (Euclidean space).** If  $\mathcal{M} = \mathbb{R}^n$  (with the “flat”  $\nabla$ ),  $R(X, Y)Z = 0$  for all  $X, Y, Z \in \Gamma(T\mathbb{R}^n)$ .

**Proof.** Since given  $Z = (z_1, \dots, z_n)$  with the components from natural coordinates of  $\mathbb{R}^n$ ,  $\nabla_X Z = (X z_1, \dots, X z_n)$ , then  $\nabla_Y \nabla_X Z = (Y X z_1, \dots, Y X z_n)$ , hence  $R(X, Y)Z = 0$ . ⊛

Hence, we see the following.

**Intuition.**  $R(X, Y)Z$  is trying to measure how much  $\mathcal{M}$  deviates from being Euclidean.

Another way to look at this is the following.

**Intuition.** Consider a system of coordinates  $\{x_i\}$  around  $p \in \mathcal{M}$ . Since  $[\partial/\partial x_i, \partial/\partial x_j] = 0$ ,

$$R\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) \frac{\partial}{\partial x_k} = (\nabla_{\frac{\partial}{\partial x_i}} \nabla_{\frac{\partial}{\partial x_j}} - \nabla_{\frac{\partial}{\partial x_j}} \nabla_{\frac{\partial}{\partial x_i}}) \frac{\partial}{\partial x_k},$$

i.e.,  $R(X, Y)Z$  is trying to measure the non-commutativity of the covariant derivative.

<sup>3</sup>In do Carmo [FC13], the corresponding definition of Definition 3.2.1 differs by a sign.

### 3.2.1 Local Expressions

It's convenient to express things in a **local coordinates**. Consider a **chart**  $(U, x)$  at  $p \in \mathcal{M}$  and let  $\partial/\partial x_i = X_i$ . We define  $R_{ijk}^\ell$  as<sup>4</sup>

$$R_{ijk}^\ell X_\ell := R(X_i, X_j)X_k.$$

If  $X = u^i X_i, Y = v^j X_j, Z = w^k X_k$ , from the linearity of  $R$ ,

$$R(X, Y)Z = R_{ijk}^\ell u^i v^j w^k X_\ell.$$

Then the above **intuition** can be rewritten as follows.

**Remark** (Algebraic significant of Riemannian curvature). Since

$$(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z) = R(\cdot, Z, X, Y) + \nabla_{[X, Y]} Z,$$

by letting  $\nabla_i := \nabla_{\frac{\partial}{\partial x^i}}, \nabla_j := \nabla_{\frac{\partial}{\partial x^j}}$ , in a **chart**  $(U, x)$ , we have

$$(\nabla_i \nabla_j Z)^k - (\nabla_j \nabla_i Z)^k = R_{lij}^k Z^\ell + \underbrace{\nabla_{[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}]} Z}_{=0} = R_{lij}^k Z^\ell,$$

i.e., the components of  $R$  contains all the information of how  $\nabla_i$  and  $\nabla_j$  fail to commute.

We can also express  $R_{ijk}^\ell$  in terms of  $\Gamma_{ij}^k$  by observing

$$R(X_i, X_j)X_k = \nabla_{X_i} \nabla_{X_j} X_k - \nabla_{X_j} \nabla_{X_i} X_k = \nabla_{X_i} (\Gamma_{jk}^\ell X_\ell) - \nabla_{X_j} (\Gamma_{ik}^\ell X_\ell),$$

hence,

$$R_{ijk}^s = \Gamma_{jk}^\ell \Gamma_{i\ell}^s - \Gamma_{ik}^\ell \Gamma_{j\ell}^s + \Gamma_{jk, i}^s - \Gamma_{ik, j}^s.$$

Lastly, we write

$$\langle R(X_i, X_j)X_k, X_\ell \rangle = R_{ijk}^s g_{\ell s} =: R_{ijk\ell}.$$

### 3.2.2 Identities

There are many important identities related to  $R$ , and we should see some of them.

**Note.** Although the above interpretations and intuitions are more or less formal, we should first get used to the formal properties of  $R$  and postpone a more geometric interpretation of **curvature** later.

The following two are due to Bianchi (both are proved in homework 2).

**Proposition 3.2.1** (First Bianchi identity). Given the **Riemannian curvature tensor**  $R$ , for all **vector fields**  $X, Y, Z$ ,

$$R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0;$$

or equivalently,  $R_{k\ell ij} + R_{kij\ell} + R_{kj\ell i} = 0$ .

**Proof.** See also do Carmo [FC13, Proposition 2.4]. ■

**Proposition 3.2.2** (Second Bianchi identity). Given the **Riemannian curvature tensor**  $R$ ,

$$\frac{\partial}{\partial x^h} R_{k\ell ij} + \frac{\partial}{\partial x^k} R_{\ell hij} + \frac{\partial}{\partial x^\ell} R_{hki j} = 0;$$

or equivalently,  $\nabla_{[\alpha} R_{\beta\gamma]\delta\epsilon} := \nabla_\alpha R_{\beta\gamma\delta\epsilon} + \nabla_\beta R_{\gamma\alpha\delta\epsilon} + \nabla_\gamma R_{\alpha\beta\delta\epsilon} = 0$ .<sup>a</sup>

<sup>a</sup>This notation is a bit cryptic: see **Ricci calculus**.

Moreover, we can also talk about exchanging two indices.

<sup>4</sup>This is how we define **connection coefficients**, i.e.,  $R_{ijk}^\ell$  are components of  $R$  in  $(U, x)$ .

**Proposition 3.2.3.** Given the [Riemannian curvature tensor](#)  $R$ ,

- (a)  $R(X, Y)Z = -R(Y, X)Z$ , i.e.,  $R_{k\ell ij} = -R_{k\ell ji}$ ;
- (b)  $\langle R(X, Y)Z, W \rangle = -\langle R(X, Y)W, Z \rangle$ , i.e.,  $R_{k\ell ij} = -R_{\ell k ij}$ ;
- (c)  $\langle R(X, Y)Z, W \rangle = -\langle R(Y, X)Z, W \rangle$ , i.e.,  $R_{k\ell ij} = -R_{\ell k ji}$ ;
- (d)  $\langle R(X, Y)Z, W \rangle = -\langle R(Z, W)X, Y \rangle$ , i.e.,  $R_{k\ell ij} = R_{ij\ell k}$ .

**Proof.** See also do Carmo [FC13, Proposition 2.5]. ■

### 3.2.3 Other Curvatures

There are other notions of curvature, but they all depend on the [Riemannian curvature](#), and appearing to be some sorts of “average” of  $R$ . We have already seen the first one.

**Definition 3.2.2** ([Riemannian-Christoffel curvature](#)). The *Riemannian-Christoffel curvature* is defined by

$$R_{k\ell ij} := g_{km} R_{\ell ij}^m = \left\langle R \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) \frac{\partial}{\partial x^\ell}, \frac{\partial}{\partial x^k} \right\rangle.$$

**Definition 3.2.3** ([Ricci curvature](#)). The *Ricci curvature* is defined by  $R_{ab} = g^{cm} R_{camb} = R_{amb}^m$ .

**Definition 3.2.4** ([Ricci scalar curvature](#)). The *(Ricci) scalar curvature* is defined by  $R = g^{ab} R_{ab}$ .

**Note.** For a more formal treatment, see do Carmo [FC13, §4.4].<sup>a</sup>

<sup>a</sup>Notice that the order in do Carmo [FC13] is a bit different: it introduces [sectional curvature](#) first.

## 3.3 Flows of Vector Fields

Let  $\mathcal{M}$  be a [smooth manifold](#), and  $X$  a [vector field](#) on  $\mathcal{M}$ . Then  $X$  defines a first order differential equation<sup>5</sup>

$$\dot{c} = X(c).$$

And this ODE has a solution, as guaranteed by [Proposition 3.3.1](#).

**Proposition 3.3.1.** For all  $p \in \mathcal{M}^d$ , there exists an open interval  $I = I_p \subseteq \mathbb{R}$  with  $0 \in I_p$  such that a [smooth curve](#)  $c: I_p \rightarrow \mathcal{M}$  solves

$$\begin{cases} \frac{dc(t)}{dt} = X(c(t)), & t \in I; \\ c(0) = p. \end{cases}$$

Further, the solution depends smoothly on the initial data (i.e.,  $p$ ).<sup>a</sup>

<sup>a</sup>This directly follows from ODE theory.

**Proof.** For all  $p \in \mathcal{M}$ , we want to find an open interval  $I = I_p$  around  $0 \in \mathbb{R}$  and a solution of the following ODE for  $c: I \rightarrow \mathcal{M}$ :

$$\begin{cases} \frac{dc(t)}{dt} = X(c(t)), & t \in I; \\ c(0) = p. \end{cases}$$

<sup>5</sup>If  $\dim \mathcal{M} > 1$ , it is a system of first order differential equations.



We can check in [local coordinates](#) that this is a system of ODE. In such [coordinates](#), let  $c(t)$  be given by  $c(t) = (c^1(t), c^2(t), \dots, c^d(t))$ . Let  $X =: X^i \partial / \partial x^i$ , then the above system becomes

$$\frac{dc^i(t)}{dt} = X^i(c(t)), \quad i = 1, \dots, d.$$

From the [Picard-Lindelöf theorem](#), with the initial data  $c(0) = p$ , there is a unique solution. ■

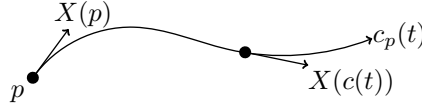
**Proposition 3.3.2.** For all  $p \in \mathcal{M}$ , there exists an open neighborhood  $U$  of  $p$  and an open interval  $I_p$  with  $0 \in I_p$  such that for all  $q \in U$ , the [curve](#)  $c_q$  with

$$\dot{c}_q(t) = X(c_q(t)), \quad c_q(0) = q$$

is defined on  $I$  and the map  $c: I \times U \rightarrow \mathcal{M}$ ,  $(t, q) \mapsto c_q(t)$  is smooth.

[Proposition 3.3.2](#) suggests the following definition.

**Definition 3.3.1** (Local flow). The map  $c_q(t): I \times U \rightarrow \mathcal{M}$ ,  $(t, q) \mapsto c_q(t)$  from [Proposition 3.3.2](#) is called the *local flow* of the [vector field](#)  $X$ .



**Definition 3.3.2** (Integral curve). The [local flow](#)  $c_q(t)$  is called the *integral curve* of  $X$  through  $q$ .

### 3.3.1 Local 1-Parameter Groups

Now, given a [local flow](#)  $c_q(t)$  of a [vector field](#)  $X$ , by fixing  $t$ , we can vary  $q$  and see the following.

**Theorem 3.3.1.** Let  $\varphi_t(q) := c_q(t)$  such that  $\varphi_t \circ \varphi_s(q) = \varphi_{t+s}(q)$  for  $s, t, (t+s) \in I_q$ . If  $\varphi_t$  is defined on  $U \subseteq \mathcal{M}$ , it maps  $U$  [diffeomorphically](#) onto its image.

We see that  $\varphi_t$  defines a family of [diffeomorphism](#) around  $p$ , which gives the following.

**Definition 3.3.3** (Local 1-parameter group). A family  $(\varphi_t)_{t \in I}$  of [diffeomorphism](#) from  $\mathcal{M}$  to  $\mathcal{M}$  satisfying [Theorem 3.3.1](#) is called a *local 1-parameter group* of [diffeomorphisms](#).

In general, a [local 1-parameter group](#) needs not be extendible to a group because the maximum interval  $I = I_q$  in [Definition 3.3.3](#) need not be all of  $\mathbb{R}$ .

**Example.** Let  $\mathcal{M} = \mathbb{R}$ ,  $X(t) = \tau^2 d/d\tau$ . Then the solution of  $\dot{c}(t) = c^2(t)$  is not defined over all  $\mathbb{R}$ .

To get the whole group structure, consider the following.

**Theorem 3.3.2.** Let  $X$  be a [vector field](#) on a [smooth manifold](#)  $\mathcal{M}$  with a compact support. Then the corresponding [local flow](#) is defined for every  $q \in \mathcal{M}$  and  $t \in \mathbb{R}$ , and the [local 1-parameter group](#) becomes a group of [diffeomorphisms](#).

**Proof.** By using  $\text{supp}(X) \subseteq K$ ,  $K$  compact, we can cover  $K$  by a finite covering, then using [Proposition 3.3.2](#), we're done. ■

This leads to the following.

**Corollary 3.3.1.** On a compact [differentiable manifold](#)  $\mathcal{M}$ , any [vector field](#) generates a [local 1-parameter group](#).

## Lecture 11: Geodesic & Cogeodesic Flows and Parallel Transport

### 3.3.2 Geodesic and Cogeodesic Flows

A particularly interesting [flow](#) is the [cogeodesic flow](#): let's first transform [Equation 2.1](#) (which is a second order ODE) into a first order system on the [cotangent bundle](#)  $T^*\mathcal{M}$ , and [locally trivialize](#)  $T^*\mathcal{M}$  by [chart](#)  $T^*\mathcal{M}|_U \cong U \times \mathbb{R}^d$  with coordinates  $(x^1, \dots, x^d, p_1, \dots, p_d)$ . Now, set

$$H(x, p) = \frac{1}{2} g^{ij}(x) p_i p_j, \quad (3.2)$$

**Theorem 3.3.3.** [Equation 2.1](#) is equivalent to the system on  $T^*\mathcal{M}$ :

$$\begin{cases} \dot{x}^i = \frac{\partial H}{\partial p_i} g^{ij}(x) p_j; \\ \dot{p}_i = -\frac{\partial H}{\partial x^i} = -\frac{1}{2} g^{jk}_{,i}(x) p_j p_k. \end{cases} \quad (3.3)$$

**Proof.** This is just computation (recall that  $g^{ik} g_{kj} = \delta_j^i$ ). ■

**Definition 3.3.4** (Cogeodesic flow). The *cogeodesic flow* is the [local flow](#) determined by [Equation 3.3](#).

**Definition 3.3.5** (Geodesic flow). The [geodesic flow](#) on  $T\mathcal{M}$  is obtained from the [cogeodesic flow](#) by the first equation in [Equation 3.3](#).

**Intuition.** The [geodesic](#) is the projection of the [integral curve](#) of the [geodesic flow](#) onto  $\mathcal{M}$ .

**Note** (Hamiltonian flow). The [cogeodesic flow](#) is a *Hamiltonian flow* for the Hamiltonian  $H$ .

**Proof.** By [Equation 3.3](#), along the [integral curves](#),

$$\frac{dH}{dt} = H_{x^i} \dot{x}^i + H_{p_i} \dot{p}_i = -\dot{p}_i \dot{x}^i + \dot{x}^i \dot{p}_i = 0.$$

⊛

Observe that the [cogeodesic flow](#) maps  $E_\lambda := \{(x, p) \in T^*\mathcal{M} \mid H(x, p) = \lambda\}$ <sup>6</sup> onto itself for all  $\lambda \geq 0$ , and if  $\mathcal{M}$  is compact, then all  $E_\lambda$  are compact, then all [geodesic flows](#) are defined on all  $E_\lambda$  for all  $\lambda$ .

## 3.4 Covariant Derivatives and Parallelism

An important concept related to [curvatures](#) is “parallelism,” which needs a formal introduction of [covariant derivatives](#). As a motivating example, the following is an equivalent definition of [geodesic](#).

**Example** (Autoparallel). The [geodesic](#)  $c$  satisfies  $\nabla_{\dot{c}} \dot{c} = 0$ . This is called [autoparallel](#).

**Proof.** In the [local coordinates](#), we have  $\dot{c} = \dot{c}^i \partial / \partial x^i$ , and note that

$$\nabla_{\dot{c}} \dot{c} = \dot{c}^i \nabla_{\frac{\partial}{\partial x^i}} \dot{c}^j \frac{\partial}{\partial x^j} = \dot{c}^i \dot{c}^j \Gamma_{ij}^k \frac{\partial}{\partial x^k} + \dot{c}^k \frac{\partial}{\partial x^k} = (\dot{c}^k + \Gamma_{ij}^k \dot{c}^i \dot{c}^j) \frac{\partial}{\partial x^k} = 0 \quad (3.4)$$

since a [geodesic](#) is the solution of [Equation 2.1](#). ⊛

**Intuition.** A [geodesic](#) is a [curve](#) with “zero acceleration”.

To understand what  $\nabla_{\dot{c}} \dot{c}$  is doing beyond just calculation, we need to understand [parallel transports](#).

<sup>6</sup>  $\mathcal{M} = \bigcup_{\lambda \geq 0} P E_\lambda$  for  $P$  being the projection.

9 Feb. 13:00

This section is weird...  
Need fix

### 3.4.1 Covariant Derivatives

We can now finally define **covariant derivative** formally.

**As previously seen.** Let  $X = X^i \frac{\partial}{\partial x_i}$ ,  $V = V^k \frac{\partial}{\partial x_k}$ , and let  $D$  be the **Levi-Civita connection**. Then

$$D_V X = D_V \left( X^i \frac{\partial}{\partial x_i} \right) = V(X^i) \frac{\partial}{\partial x_i} + X^i D_V \frac{\partial}{\partial x_i} = V(X^i) \frac{\partial}{\partial x_i} + V^k X^i \Gamma_{ki}^j \frac{\partial}{\partial x_j}.$$

**Proposition 3.4.1 (Covariant derivative).** Let  $(\mathcal{M}, g)$  be a **Riemannian manifold**,  $D$  the **Levi-Civita connection**, and  $c$  a **smooth curve** in  $\mathcal{M}$  with the set of smooth **vector fields along  $c$**   $\mathcal{X}_c(\mathcal{M})$ . Then there exists a unique operator  $D/dt$  defined as the vector space of **vector fields along  $c$**  satisfying

- (i) (a)  $\frac{D}{dt}(fY)(t) = f'(t)Y(t) + f(t)\frac{D}{dt}Y(t)$  for all  $f \in C^\infty(I)$  and  $Y \in \mathcal{X}_c(\mathcal{M})$ ;
- (b)  $\frac{D}{dt}(V+W) = \frac{DV}{dt} + \frac{DW}{dt}$  for all  $V, W \in \mathcal{X}_c(\mathcal{M})$ ;
- (ii) if there exists a neighborhood of in  $I$  such that  $Y$  is the restriction to  $c$  of a **vector field  $X$**  defined on a neighborhood of  $c(t_0)$  in  $\mathcal{M}$ , then  $\frac{D}{dt}Y(t_0) = (D_{c(t_0)}X)_{c(t_0)}$ .

**Proof.** Consider defining such an operator  $D/dt$  as

$$\frac{D}{dt} \left( Y^i(t) \frac{\partial}{\partial x_i} \right) = \frac{dY^i}{dt} \frac{\partial}{\partial x_i} + \dot{c} Y^i \Gamma_{ji}^k(c(t)) \frac{\partial}{\partial x_k},$$

where  $\dot{c} = \dot{c}^k \frac{\partial}{\partial x_k}$ . This shows (i) (a) and (b) hold. Next, to show (ii), let  $x$  be a smooth **vector field** in  $\mathcal{M}$ . Then the induced **vector field along  $c$**  is given by  $Y(t) = X_{c(t)}$ , i.e., in terms of the coordinate basis, we have  $Y(t) = Y^i(t) \frac{\partial}{\partial x_i}$ ,  $X_x = X^i(x) \frac{\partial}{\partial x_i}$ , and  $Y^i(t) = X^i(c(t))$ . Then,

$$\begin{aligned} D_i X &= D_i \left( X^i \frac{\partial}{\partial x_i} \right) = \dot{c}(X^i) \frac{\partial}{\partial x_i} + X^i D_i \frac{\partial}{\partial x_i} = X^i \dot{c}^k \underbrace{D_{\frac{\partial}{\partial x_k}} \frac{\partial}{\partial x_i}}_{\Gamma_{ki}^\ell \frac{\partial}{\partial x_\ell}} \\ &= \partial_t(X^i \circ c) \frac{\partial}{\partial x_i} + \dot{c}^k X^i \Gamma_{ki}^\ell \frac{\partial}{\partial x_\ell} = \partial_t(X^i \circ c) \frac{\partial}{\partial x_i} + \dot{c}^k Y^i \Gamma_{ki}^\ell \frac{\partial}{\partial x_\ell} = \frac{D}{dt} Y. \end{aligned}$$

■

**Note.** From **Proposition 3.4.1**,  $D/dt$  is what we want, and note how it depends on  $c$ .

**Definition 3.4.1 (Covariant derivative).** The *covariant derivative* of  $V$  along  $c$  is the **vector field**  $DV/dt$ .

**Problem 3.4.1.** Why not just define  $DY/dt$  by (ii)?

**Answer.** A **vector field  $Y$  along  $c$**  may not always be extended to a neighborhood of  $c$  in  $\mathcal{M}$ . But, in **local coordinates**,  $Y$  is always a linear combination of **vector fields along  $c$**  since

$$Y(t) = \sum_{i=1}^n Y^i(t) \left( \frac{\partial}{\partial x^i} \right)_{c(t)},$$

i.e., it can be extended. \*

**Proposition 3.4.1** shows that the choice of a **linear connection** on  $\mathcal{M}$  leads to a bona fide (satisfying (a) and (b)) derivative of **vector fields along curves**.

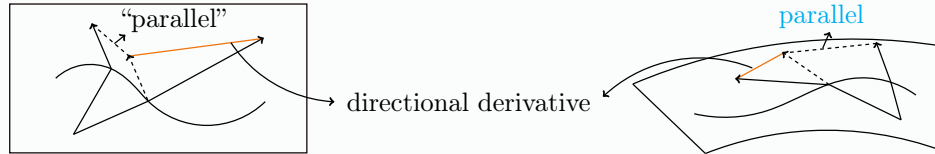
**Remark.** The notion of **connection** furnishes a manner of differentiating vectors along **curves**.

### 3.4.2 Parallel Transports

Finally, we introduce the notion of [parallel](#).

**Definition 3.4.2 (Parallel).** A [vector field](#)  $X$  on  $\mathcal{M}$  along a [curve](#)  $c$  is *parallel* (or *parallelly transported*) along  $c$  if  $DX/dt = 0$  for all  $t \in I$ .

**Intuition.** In the (flat) Euclidean space, we know what is “parallel,” and hence we can define the directional derivative. But now the logic is reversed: we first define what is [parallel](#) in a curved space, and then we can make sense of directional derivative in a curved space!

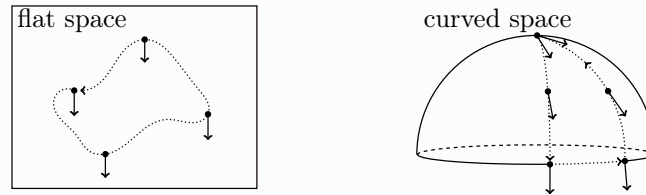


Given the definition of a [parallel vector fields along curves](#), we can talk about [parallel transport](#).

**Definition 3.4.3 (Parallel transport).** The *parallel transport* from  $c(0)$  to  $c(t)$  along the [curve](#)  $c$  in a [Riemannian manifold](#)  $(\mathcal{M}, g)$  is the linear map  $P_t: T_{c(0)}\mathcal{M} \rightarrow T_{c(t)}\mathcal{M}$  associating  $v \in T_{c(0)}\mathcal{M}$  with  $X_v(t) \in T_{c(t)}\mathcal{M}$  with  $X_v$  being the [parallel vector field along](#)  $c$  such that  $X_v(0) = v$ .

It's clear that how we can extend [Definition 3.4.3](#) for a piece-wise smooth [curve](#).

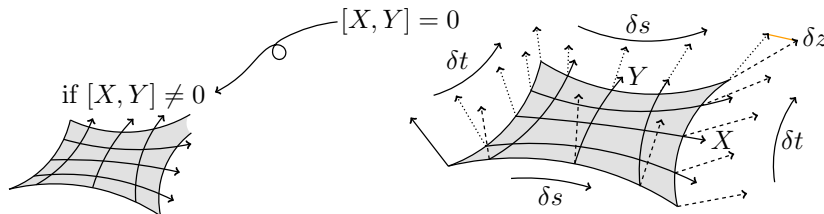
**Intuition.** When the space is flat, keeping the “arrow” (which defines a [vector field](#)) in one direction and moving around won't produce any changes, while when the space is curved, it will.



We make a surprising remark on the relation between [Riemannian curvature](#) and [parallel transport](#).

**Remark (Geometric significant of Riemannian curvature).** The idea is that for a [manifold](#) with [torsion free](#)  $\nabla$ , if we [parallel transporting](#) along two paths on an infinitesimal patch (which induces  $X, Y$ ) such that  $[X, Y] = 0$ , we can detect [curvature](#) in terms of  $\delta z$ , where<sup>a</sup>

$$(\delta z)^i = \dots = R^i_{jkl} X^k Y^l Z^j \cdot \delta s \delta t + O(\delta s^2 \delta t, \delta s \delta t^2).$$



We will come back to this later.

<sup>a</sup>This is a deep theorem! In the ..., we use  $T \equiv 0$ .

**Proposition 3.4.2.** The [parallel transport](#) exists, uniquely.

**Proof.** do Carmo [FC13, Proposition 2.6] ■

**Proposition 3.4.3.** Let  $(\mathcal{M}, g)$  be a Riemannian manifold. The parallel transport defines for all  $t$  an isometry from  $T_{c(0)}\mathcal{M}$  onto  $T_{c(t)}\mathcal{M}$ ; more generally, if  $X, Y$  are vector fields along  $c$ , then

$$\frac{d}{dt}g(x(t), y(t)) = g\left(\frac{DX(t)}{dt}, Y(t)\right) + g\left(X(t), \frac{DY(t)}{dt}\right).$$

**Proof.** See do Carmo [FC13, Proposition 3.2] ■

### 3.4.3 Autoparallel Curves

Now we can formally introduce the notion of autoparallel.

**Definition 3.4.4 (Autoparallel).** Let  $\nabla$  be a connection on  $T\mathcal{M}$  of a differentiable manifold  $\mathcal{M}$ . A curve  $c: I \rightarrow \mathcal{M}$  is called *autoparallel* (or *geodesic*) w.r.t.  $\nabla$  if

$$\nabla_{\dot{c}}\dot{c} = 0.$$

**Intuition.** An autoparallel curve is the *straightest line* (hence *geodesic*) in the space w.r.t.  $\nabla$ !

**Remark (Physical interpretation).** One can start from introducing  $\nabla$ , considering  $\nabla_{\dot{c}}\dot{c} := 0$  (which is just Equation 2.1), and realize that we don't need to consider gravity as a force, rather a "curvature of spacetime," in order to make sense of Newton's first law, i.e., mass without forces will undergo a autoparallel curve.

**Example (Euclidean plane).** Let  $U = \mathbb{R}^2$ ,  $x = \text{id}_{\mathbb{R}^2}$ ,  $\Gamma_{jk}^i = 0$ , then  $\ddot{c}^k = 0$  in Equation 3.4. Hence,

$$c^k(t) = a^k t + b^k \text{ for } a^k, b^k \in \mathbb{R}^d.$$

**Example (Round sphere).** The geodesics on a "round sphere" are the great circles.

**Proof.** Consider a "unit round sphere"  $\mathcal{M} = S^2$  with spherical coordinates  $x(p) = (r, \theta, \varphi)$  such that  $r = 1$ ,  $\theta \in (0, \pi)$ , and  $\varphi \in [0, 2\pi)$ . The "roundness" is given by  $\nabla_{\text{round}}$  where we specify (at one point)

$$\Gamma_{22}^1 := -\sin \theta \cos \theta, \quad \Gamma_{21}^2 = \Gamma_{12}^2 := \cot \theta,$$

where we let  $x^1(p) = \theta(p)$ ,  $x^2(p) = \varphi(p)$ . The autoparallel equation tells us

$$\begin{cases} \ddot{\theta} + \Gamma_{22}^1 \dot{\varphi}^2 = 0; \\ \ddot{\varphi} + 2\Gamma_{12}^2 \dot{\theta} \dot{\varphi} = 0; \end{cases} \Leftrightarrow \begin{cases} \ddot{\theta} - \sin(\theta) \cos(\theta) \dot{\varphi}^2 = 0; \\ \ddot{\varphi} + 2 \cot(\theta) \dot{\theta} \dot{\varphi} = 0. \end{cases}$$

Then, we see that  $\theta(t) = \pi/2$ ,  $\varphi(t) = \omega t + \varphi_0$  is a solution.<sup>a</sup> Hence, we conclude that if we run at a constant speed around the great circle of  $S^2$ , it'll be autoparallel, hence a geodesic. ⊛

<sup>a</sup>Note that  $\theta(t) = \pi/2$ ,  $\varphi(t) = \omega t^2 + \varphi_0$  is not a solution.

Similarly, given any  $\nabla$  on a space, we can find the straightest curve on which.

## Lecture 12: Tangent and Cotangent Bundles

### 3.5 More on Tangent and Cotangent Bundles

14 Feb. 13:00

Let  $f: \mathcal{M} \rightarrow \mathcal{N}$  be a differentiable map between two differentiable manifolds, until now, we have only talked about how to transform tangent vectors or 1-form via  $f$ . Implicitly, these are just pullback ( $f^*$ ) and pushforward ( $f_*$ ), as we now define formally.

**Definition.** Let  $f: \mathcal{M} \rightarrow \mathcal{N}$  be a smooth map between two smooth manifolds and  $p \in \mathcal{M}$ .

**Definition 3.5.1 (Pushforward).** The *pushforward* is the linear map  $f_* := df_p: T_p\mathcal{M} \rightarrow T_{f(p)}\mathcal{N}$ .

**Definition 3.5.2 (Pullback).** The *pullback* is the linear map  $f^*: T_{f(p)}^*\mathcal{N} \rightarrow T_p^*\mathcal{M}$  where

$$(f^*\omega)(X) = \omega(f_*X)$$

for  $\omega \in T_{f(p)}^*\mathcal{N}$  and  $X \in T_p\mathcal{M}$ .

In all, the following diagram commutes:

$$\begin{array}{ccc} T_p^*\mathcal{M} & \xleftarrow{f^*} & T_p^*\mathcal{N} \\ \pi \downarrow & & \downarrow \pi \\ \mathcal{M} & \xrightarrow{f} & \mathcal{N} \end{array} \quad \begin{array}{ccc} T_p\mathcal{M} & \xrightarrow{f_*} & T_{f(p)}\mathcal{N} \\ \pi \downarrow & & \downarrow \pi \\ \mathcal{M} & \xrightarrow{f} & \mathcal{N} \end{array}$$

### 3.5.1 Pullbacks and Pushforwards on Bundles

Now, consider a vector bundle  $(E, \pi, \mathcal{N})$  over  $\mathcal{N}$ , we want to use  $f$  to “pull back” the vector bundle, i.e., construct a vector bundle, denote as  $f^*E$ , for which the fiber over  $x \in \mathcal{M}$  is  $E_{f(x)}$ .

**Definition 3.5.3 (Pullback bundle).** The *pullback bundle*  $f^*E$  is the vector bundle over  $\mathcal{M}$  with the bundle charts  $(\varphi \circ f, f^{-1}(U))$  if  $(\varphi, U)$  is the bundle charts of  $E$ .

Similarly, we can “push forward” a vector bundle  $(E, \pi, \mathcal{M})$  over  $\mathcal{M}$  via  $f$  in the same fashion.

**Definition 3.5.4 (Pushforward bundle).** The *pushforward bundle*  $f_*E$  is the vector bundle over  $\mathcal{N}$  with the bundle charts  $(\varphi \circ f^{-1}, f(U))$  if  $(\varphi, U)$  is the bundle charts of  $E$ .

**Note.** In Definition 3.5.4, it only makes sense if  $\mathcal{M} \hookrightarrow \mathcal{N}$ .

**Definition 3.5.5 (Bundle homomorphism).** Consider 2 vector bundles  $(E_1, \pi_1, \mathcal{M}), (E_2, \pi_2, \mathcal{M})$  over  $\mathcal{M}$ , and let the differentiable map  $f: E_1 \rightarrow E_2$  be fiber preserving, i.e.,  $\pi_2 \circ f = \pi_1$ . If the fiber maps  $f_x: E_{1,x} \rightarrow E_{2,x}$  is linear,<sup>a</sup> then  $f$  is called a *bundle homomorphism*.

<sup>a</sup>I.e., vector homomorphisms.

**Definition 3.5.6 (Subbundle).** Let  $(E, \pi, \mathcal{M})$  of rank  $n$  be a vector bundle. Let  $E^1 \subseteq E$ , and assume that for all  $x \in \mathcal{M}$ , there exists a bundle chart  $(\varphi, U)$  for  $x \in U$  and

$$\varphi(\pi^{-1}(U) \cap E^1) = U \times \mathbb{R}^m \subseteq U \times \mathbb{R}^n$$

for  $m \leq n$ . Then the *subbundle* of  $E$  of rank  $m$  is the vector bundle  $(E^1, \pi|_{E^1}, \mathcal{M})$ .

**Example.** Consider  $f: \mathcal{M} \hookrightarrow \mathcal{N}$  where  $g_{\mathcal{N}}$  is a metric on  $\mathcal{N}$ . Then,  $g_{\mathcal{N}}$  induces a metric  $g_{\mathcal{M}}$  on  $\mathcal{M}$  by  $f$  since we can define

$$g_{\mathcal{M}}(X, Y) := g_{\mathcal{N}}(f_*(X), f_*(Y)).$$

### 3.5.2 Pullbacks and Pushforwards of Vector Fields

Now, we consider to “pull back” or “push forward” a vector field, i.e., a section of a bundle.

**Definition 3.5.7 (Pushforward).** Let  $\psi: \mathcal{M} \rightarrow \mathcal{N}$  be a diffeomorphism between smooth manifolds, and let  $X$  be a vector field on  $\mathcal{M}$ . Then the pushforward vector field  $Y = \psi_*X = d\psi X$  on  $\mathcal{N}$  is

$$Y(p) = d\psi(X(\psi^{-1}(p))).$$

**Definition 3.5.8 (Pullback).** Let  $\psi: \mathcal{M} \rightarrow \mathcal{N}$  be a diffeomorphism between smooth manifolds, and let  $Y$  be a vector field on  $\mathcal{N}$ . Then the pullback vector field  $X = \psi^*Y$  on  $\mathcal{M}$  is just  $X(p) = Y_{\psi(p)}$ .

**Note.** We let  $\psi$  be a diffeomorphism just for convenient.

**Lemma 3.5.1.** For every differentiable function  $f: \mathcal{N} \rightarrow \mathbb{R}$ ,  $(\psi_*X)(f)(p) = X(f \circ \psi)(\psi^{-1}p)$ .

**Lemma 3.5.2.** Let  $X$  be a vector field on  $\mathcal{M}$  and  $\psi: \mathcal{M} \rightarrow \mathcal{N}$  be a diffeomorphism. If the local 1-parameter group  $(\varphi_t)_{t \in I}$  generated by  $X$ , then the local 1-parameter group generated by  $\psi_*X$  is  $\psi \circ \varphi_t \circ \psi^{-1}$ .

### 3.5.3 Induced Bundle Metrics

Let  $(\mathcal{M}, g)$  be a Riemannian manifold, then  $g$  induces the bundle metrics on all vector bundles over  $\mathcal{M}$ : for  $T^*\mathcal{M}$ , it is given by

$$g(\omega, \eta) := g^{ij}\omega_i\eta_j$$

for  $\omega = \omega_i dx^i, \eta = \eta_i dx^i$ . Hence, we can talk about the identification between  $T\mathcal{M}$  and  $T^*\mathcal{M}$  through  $g$ :

$$\begin{array}{c} V = V^i \frac{\partial}{\partial x^i} \in T\mathcal{M} \\ \updownarrow \\ \omega = \omega_j dx^j \in T^*\mathcal{M} \end{array}$$

with  $\omega_j = g_{ij}V^i$  (or  $V^i = g^{ij}\omega_j$ ) such that

- (a)  $g(X, Y) = g_{ij}X^iY^j$  for  $X, Y \in T\mathcal{M}$ ;
- (b)  $g(\omega, \eta) = g^{ij}\omega_i\eta_j$  for  $\omega, \eta \in T^*\mathcal{M}$ .

Thus, for  $V \in T_x\mathcal{M}$ , there corresponds a 1-form  $\omega \in T_x^*\mathcal{M}$  via the metric  $\omega(Y) := g(V, Y)$  for all  $Y$ , and we further have  $\|\omega\| = \|V\|$ .

We can also consider the coordinate transformation behavior. Let  $(e_i)_{i=1,\dots,d}$  be a basis of  $T_x\mathcal{M}$  and  $(\omega^j)_{j=1,\dots,d}$  the dual basis of  $T_x^*\mathcal{M}$ , i.e.,  $\omega^j(e_i) = \delta_i^j$ . Given  $V = V^i e_i \in T_x\mathcal{M}$ ,  $\eta = \eta_j \omega^j \in T_x^*\mathcal{M}$ , we then have  $\eta(V) = \eta_i V^i$ . Now, consider bases  $(e_i), (\omega^j)$  in the local coordinates, i.e.,  $e_i = \partial/\partial x^i$  and  $\omega^j = dx^j$ . Let  $f$  be a local coordinates change, then  $V$  and  $\eta$  transformed as

$$f_*(V) := V^i \frac{\partial f^\alpha}{\partial x^i} \frac{\partial}{\partial f^\alpha}, \quad f^*(\eta) := \eta_j \frac{\partial x^j}{\partial f^\beta} df^\beta$$

correspondingly, and we see that

$$f^*(\eta)(f_*(V)) = \eta_j \frac{\partial x^j}{\partial f^\alpha} V^i \frac{\partial f^\alpha}{\partial x^i} = \eta_i V^i = \eta(V).$$

**Intuition.** The above means that

- the tangent vectors transform with the functional matrix of coordinates change;
- the cotangent vectors transform with the transposed inverse of the above matrix.

To compute the **coordinates** change  $y \mapsto x(y)$  for  $\omega = \omega_i dx^i$ ,  $\eta = \eta_i dx^i$  with  $\langle \omega, \eta \rangle = g^{ij} \omega_i \eta_j$ , we have

$$\omega_i dx^i = \omega_i \frac{\partial x^i}{\partial y^\alpha} dy^\alpha =: \tilde{\omega}_\alpha dy^\alpha.$$

**As previously seen.**  $g^{ij}$  is transformed as

$$h^{\alpha\beta} = g^{ij} \frac{\partial y^\alpha}{\partial x^i} \frac{\partial y^\beta}{\partial x^j}.$$

Then, we see that  $h^{\alpha\beta} \tilde{\omega}_\alpha \tilde{\eta}_\beta = g^{ij} \omega_i \eta_j$  and  $\|\omega(x)\| = \sup \{\omega(x)(V) \mid V \in T_x \mathcal{M}, \|v\| = 1\}$ .

**Remark.** If we consider  $T\mathcal{M} \otimes T\mathcal{M}$ , then metric is

$$\langle V \otimes Y, \xi \otimes \eta \rangle = g_{ij} V^i Y^j g_{kl} \xi^k \eta^l.$$

**As previously seen** (Lie derivative). Consider a **vector field**  $X$  with a **local 1-parameter group**  $(\psi_t)_{t \in I}$  and a **tensor field**  $S$  on  $\mathcal{M}$ . The **Lie derivative** of  $S$  in the direction of  $X$  is defined as

$$\mathcal{L}_X S := \left. \frac{d}{dt} (\psi_t^* S) \right|_{t=0}.$$

## Lecture 13: Sectional Curvatures and Space Forms

Let  $X = X^i \partial / \partial x^i$  be a **vector field**. Then consider  $(\psi_t)_* X(\psi_t(x))$  to get a **curve**  $X_t$  in  $T_x \mathcal{M}$  for  $t \in I$ . 16 Feb. 13:00  
By differentiate that curve, i.e.,

$$(\psi_t)_* \frac{\partial}{\partial x^i} (\psi_t(x)) = \frac{\partial \psi_t^k}{\partial x^i} \frac{\partial}{\partial x^k}.$$

**Note.** For  $\varphi: \mathcal{M} \rightarrow \mathcal{N} := \mathcal{M}$  and  $X$  and  $\varphi(x)$  are in the same **coordinate neighborhood**,

$$\varphi^* \frac{\partial}{\partial x^i} = \frac{\partial \varphi^k}{\partial x^i} \frac{\partial}{\partial \varphi^k}$$

since  $\frac{\partial}{\partial \varphi^k} = \frac{\partial}{\partial x^k}$ .

On the other hand, let  $\omega = \omega_i dx^i$  be a 1-form, then we have

$$(\psi_t^*)(\omega)(x) = \omega_i(\psi_t(x)) \frac{\partial \psi_t^i}{\partial x^k} dx^k,$$

which is a **curve** in  $T_x^* \mathcal{M}$ .

**Note.** For  $\varphi: \mathcal{M} \rightarrow \mathcal{N}$  ( $\varphi$  need not be a **diffeomorphism**) with for the 1-form  $\omega = \omega_i dx^i$  on  $\mathcal{N}$ ,

$$\varphi^* \omega = \omega_i(\varphi(x)) \frac{\partial z^i}{\partial x^k} dx^k.$$

Let  $\varphi: \mathcal{M} \rightarrow \mathcal{N}$  be a **diffeomorphism**,  $Y$  be a **vector field** on  $\mathcal{N}$ . Then, set

$$\varphi^* Y := (\varphi^{-1})_* Y,$$

and for other **contravariant tensors**,  $\varphi^*$  can be defined in an analogous way.

**Example.** For a **vector field**  $X$  and a **local 1-parameter group**  $(\psi_t)_{t \in I}$ , it is  $(\psi_t^* X) = (\psi_t)_* X$ .



### 3.6 Sectional Curvatures

Beyond [Riemannian curvature](#) and other “averaging” variations of which, the following one is in particular interesting and is the one considered by Riemann.

**Definition 3.6.1** (Sectional curvature). The *sectional curvature* of the plane  $\Sigma$  spanned by the (linearly independent) [tangent vectors](#)  $X = X^i \frac{\partial}{\partial x^i}, Y = Y^i \frac{\partial}{\partial x^i} \in T_x \mathcal{M}$  of a [Riemannian manifold](#)  $(\mathcal{M}, g)$  is

$$K(\sigma) := K(X \wedge Y) = \frac{g(R(X, Y)Y, X)}{|X \wedge Y|^2}$$

where  $|X \wedge Y|^2 = g(X, X)g(Y, Y) - g(X, Y)^2$ .<sup>a</sup>

<sup>a</sup>Given a vector space  $V$  and  $x, y \in V$ ,  $|x \wedge y| := \sqrt{|x|^2|y|^2 - \langle x, y \rangle^2}$  represents the area of the two-dimensional parallelogram spanned by  $x, y$ .

**Note.** [Definition 3.6.1](#) is well-defined since  $K(\sigma)$  is invariant under different bases of  $\sigma$ .

**Remark.** [Sectional curvature](#) determines the whole [Riemannian curvature](#).

**Proof.** Given  $g(R(X, Y)Z, W)$ , we can express this entirely by  $K$  [[FC13](#), Lemma 3.3]. ⊛

**Remark** ([Gauss curvature](#)). For  $\dim \mathcal{M} = 2$ ,  $R_{ijkl} = K(g_{ik}g_{jl} - g_{ij}g_{kl})$  since  $T_x \mathcal{M}$  contains only one plane, i.e.,  $T_x \mathcal{M}$  itself. In this case,  $K$  is called the *Gauss curvature*.

In particular, the [space form](#) considers the space with constant [sectional curvature](#).

**Definition 3.6.2** (Space form). A [Riemannian manifold](#)  $(\mathcal{M}, g)$  is a *space form* if  $K(X \wedge Y)$  is a constant for all linearly independent [tangent vectors](#)  $X, Y \in T_p \mathcal{M}$  for all  $p \in \mathcal{M}$ .

**Definition 3.6.3** (Spherical). A [space form](#) is called *spherical* if  $K > 0$ .

**Definition 3.6.4** (Flat). A [space form](#) is called *flat* if  $K = 0$ .

**Definition 3.6.5** (Hyperbolic). A [space form](#) is called *hyperbolic* if  $K < 0$ .

Generalize [Definition 3.6.2](#) a bit, we have the so-called [Einstein manifolds](#).

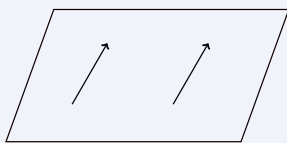
**Definition 3.6.6** (Einstein manifold). A [Riemannian manifold](#)  $(\mathcal{M}, g)$  is called an *Einstein manifold* if  $R_{ik} = cg_{ik}$  for a constant  $c$ .<sup>a</sup>

<sup>a</sup>Which does not depend on the choice of [local coordinates](#).

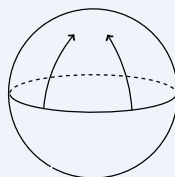
**Remark.** Every [space form](#) is an [Einstein manifold](#).

**Example.**  $\mathbb{R}^n$  is [flat](#),  $S^n$  is [spherical](#), and  $\mathbb{H}^n$  is [hyperbolic](#). And all are [Einstein manifolds](#).

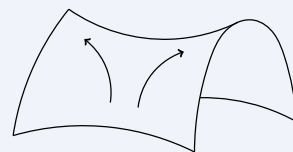
$K = 0$



$K > 0$



$K < 0$



**Definition 3.6.7 (Flat).** A connection  $\nabla$  on  $T\mathcal{M}$  is *flat* if each point in  $\mathcal{M}$  has a neighborhood  $U$  with *local coordinates* for which all the coordinate *vector fields*  $\partial/\partial x^i$  are *parallel*, i.e.,  $\nabla \partial/\partial x^i = 0$ .

**Theorem 3.6.1.** A connection  $\nabla$  on  $T\mathcal{M}$  is *flat* if and only if its *curvature* and *torsion* vanish identically.

**Proof.** *Flat connection* implies  $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = 0$ , hence all  $\Gamma_{ij}^k = 0$ , so  $T, R$  vanish. Conversely, find the *local coordinates* such that  $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = 0$  for all  $i, j$  and use *Frobenius theorem*. ■

**Example.** The following are *flat manifolds* with their usual shape, i.e., *connections*.

- $\mathbb{R}^n$ .
- Products of *flat manifolds*.
- Torus  $T^2$ .
- Every 1-dimensional *Riemannian manifold*.
- Tori.

**Theorem 3.6.2 (Schur theorem).** Let  $(\mathcal{M}, g)$  be a *Riemannian manifold* with  $\dim \mathcal{M} \geq 3$ .

- (a) If the *sectional curvature* of  $\mathcal{M}$  is constant at each point, i.e.,  $K(X \wedge Y) = f(x)$  for  $X, Y \in T_x \mathcal{M}$ , then  $f(x)$  is a constant on  $\mathcal{M}$ , hence  $\mathcal{M}$  is a *space form*.
- (b) If the *Ricci curvature* is a constant at each point, i.e.,  $R_{ik} = c(x)g_{ik}$ , then  $c(x)$  is a constant, hence  $\mathcal{M}$  is an *Einstein manifold*.

**Remark.** *Schur theorem* says that the isotropy<sup>a</sup> of a *Riemannian manifold* implies the homogeneity.<sup>b</sup> Hence, a point-wise property implies a global one!

<sup>a</sup>I.e., the property that at each point, all directions are geometrically indistinguishable.

<sup>b</sup>I.e., all points are geometrically indistinguishable.

### 3.7 More on Covariant Derivatives

To end this chapter, we revisit *covariant derivative*. But this time, we generalize it from *vector field* to *tensor field*, i.e., we will show that it's also possible to covariantly differentiate *tensors*. The motivation is that given a 1-form  $\omega$ , and *vector fields*  $X, Y$ , we have

$$X(\omega(Y)) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y),$$

and for arbitrary *tensors*  $S, T$ , we similarly have

$$\nabla_X(S \otimes T) = \nabla_X S \otimes T + S \otimes \nabla_X T.$$

Consider the following.<sup>7</sup>

**Definition 3.7.1 (Covariant differential).** Let  $T$  be a  $(0, s)$ -*tensor*. The *covariant differential*  $\nabla T$  of  $T$  is a  $(0, s+1)$ -*tensor* given by

$$\nabla T(Y_1, \dots, Y_s, Z) = Z(T(Y_1, \dots, Y_s)) - T(\nabla_Z Y_1, \dots, Y_s) - \dots - T(Y_1, \dots, Y_{s-1}, \nabla_Z Y_s).$$

**Definition 3.7.2 (Covariant derivative).** For each  $Z \in \Gamma(T\mathcal{M})$ , the *covariant derivative*  $\nabla_Z T$  of  $T$  relative to  $Z$  is a  $(0, s)$ -*tensor* given by

$$\nabla_Z T(Y_1, \dots, Y_s) = \nabla T(Y_1, \dots, Y_s, Z).$$

We primarily focus on covariant *tensor*, however, we also have the following.

<sup>7</sup>Definition 3.7.2 is natural by considering a certain *frame* [FC13, §4.5].

**Remark.** For  $T$  a  $(p, q)$ -tensor,

$$\begin{aligned} (\nabla_Y T)(\alpha_1, \dots, \alpha_q, X_1, \dots, X_p) &= Y(T(\alpha_1, \dots, \alpha_q, X_1, \dots, X_p)) \\ &\quad - \sum_{i=1}^q T(\alpha_1, \dots, \nabla_Y \alpha_i, \dots, \alpha_q, X_1, \dots, X_p) \\ &\quad - \sum_{i=1}^p T(\alpha_1, \dots, \alpha_q, X_1, \dots, \nabla_Y X_i, \dots, X_p). \end{aligned}$$

**Example.** Consider the metric tensor  $g = g_{ij} dx^i \otimes dx^j$ , then  $\nabla_X g = 0$  for all vector fields  $X$ .

**Proof.** For all  $X, Y, Z \in \Gamma(T\mathcal{M})$ ,

$$\nabla g(X, Y, Z) = Z\langle X, Y \rangle - \langle \nabla_Z X, Y \rangle - \langle X, \nabla_Z Y \rangle = 0$$

since  $\nabla$  is Riemannian. \*

It's convenient to use the following identification.

**Notation.** Let  $X \in \Gamma(T\mathcal{M})$  and identify  $X$  with the tensor that associates to  $Y \in \Gamma(T\mathcal{M})$  the function  $\langle X, Y \rangle$ .

**Intuition.** Consider the covariant derivative of the tensor  $X$  relative to  $Z \in \Gamma(T\mathcal{M})$ , which is such that for all  $Y \in \Gamma(T\mathcal{M})$ ,

$$\nabla_Z X(Y) = \nabla X(Y, Z) = Z(X(Y)) - X(\nabla_Z Y) = Z\langle X, Y \rangle - \langle X, \nabla_Z Y \rangle = \langle \nabla_Z X, Y \rangle.$$

This shows that the tensor  $\nabla_Z X$  can be identified with the vector field  $\nabla_Z X$  as well by our new notation!

**Remark.** This justifies the notation adopted, and shows that the Definition 3.7.2 is a generalization of Definition 3.4.1.

# Chapter 4

## Isometric Immersions

Consider  $f: \mathcal{M} \rightarrow \widetilde{\mathcal{M}}$  be a differentiable immersion of a manifold  $\mathcal{M}^n$  into a Riemannian manifold  $\widetilde{\mathcal{M}}^k$  for  $k = n + m$ . The Riemannian metric of  $\widetilde{\mathcal{M}}$  induces, naturally, a Riemannian metric on  $\mathcal{M}$ : if  $v_1, v_2 \in T_p\mathcal{M}$ , we let

$$\langle v_1, v_2 \rangle := \langle df_p(v_1), df_p(v_2) \rangle.$$

This makes  $f$  an isometric immersion of  $\mathcal{M}$  into  $\widetilde{\mathcal{M}}$ , and we want to study the relationship between the geometry of  $\mathcal{M}$  and that of  $\widetilde{\mathcal{M}}$ .

While do Carmo [FC13] directly discusses the second fundamental form, we start by introducing the Riemannian covering map, which has a strong connection to the second fundamental form and furnishes a broader view of the theory of isometric immersions.

### 4.1 Riemannian Covering Maps

Let's first review the basic notion in algebraic topology.

**Definition 4.1.1** (Covering map). Let  $\mathcal{M}, \widetilde{\mathcal{M}}$  be two manifolds. A map  $p: \widetilde{\mathcal{M}} \rightarrow \mathcal{M}$  is a *covering map* if

- (a)  $p$  is smooth and surjective;
- (b) for all  $m \in \mathcal{M}$ , there exists a neighborhood  $U$  at  $m$  in  $\mathcal{M}$  with  $p^{-1}(U) = \coprod_{i \in I} U_i$  with  $p: U_i \rightarrow U$  being a diffeomorphism and  $U_i$  are disjoint open subsets of  $\widetilde{\mathcal{M}}$ .

**Notation** (Covering space).  $\widetilde{\mathcal{M}}$  in Definition 4.1.1 is called the *covering space*.

**Notation** (Universal covering space). A covering space is *universal* if it's simply connected.

By introducing local isometry, we have the so-called Riemannian covering map.

**Definition 4.1.2** (Riemannian covering map). Let  $(\mathcal{M}, g), (\mathcal{N}, h)$  be Riemannian manifolds. A map  $p: \mathcal{N} \rightarrow \mathcal{M}$  is a *Riemannian covering map* if  $p$  is a smooth covering map and is a local isometry.

#### 4.1.1 Induced Riemannian Covering Maps

Given a covering map, from a Riemannian metric  $g$  on the covering space, we obtain an induced Riemannian metric on the base space and a Riemannian covering map.

**Proposition 4.1.1.** Let  $p: \mathcal{N} \rightarrow \mathcal{M}$  be a smooth covering map. For every Riemannian metric  $g$  on  $\mathcal{M}$ , there exists a unique Riemannian metric  $h$  on  $\mathcal{N}$  such that  $p$  is a Riemannian covering map.

**Note.** The converse of [Proposition 4.1.1](#) is generally not true.

Let's first see some examples.

**Example.** Every space [covers](#) itself trivially.

**Example.**  $\mathbb{R}$  is the [universal covering space](#) of  $S^1$ .

**Example.**  $U(n)$  has [universal covers](#)  $U(n) \times \mathbb{R}$ .

**Example.**  $S^n$  is a double [cover](#) for  $\mathbb{R}P^n$  and is [universal](#) for  $n > 1$ .

## Lecture 14: The Second Fundamental Form

21 Feb. 13:00

**Proposition 4.1.2.** Let  $(\mathcal{N}, h)$  be a [Riemannian manifold](#) and  $G$  be a [free](#) and [proper](#) group of [isometries](#) of  $(\mathcal{N}, h)$ . Then, there exists a unique [Riemannian metric](#)  $g$  on the quotient manifold  $\mathcal{M} = \mathcal{N} / G$  such that the connected projection  $p: \mathcal{N} \rightarrow \mathcal{M}$  is a [Riemannian covering map](#).

**Proof.** Let  $n, n' \in \mathcal{N}$  such that  $n, n' \in p^{-1}(m)$  for  $m \in \mathcal{M}$ . Hence, there exists an [isometry](#)  $f \in G$  such that  $f(n) = n'$ . Also,  $p \circ f = p$ , and  $p$  is a local [diffeomorphism](#), so we can define a scalar product  $g_m$  on  $T_m\mathcal{M}$ : for all  $u, v \in T_m\mathcal{M}$ ,

$$g_m(u, v) = h_n((T_n p)^{-1}u, (T_n p)^{-1}v)$$

for  $n \in p^{-1}(m)$ . This does not depend on the choice of  $n \in p^{-1}(m)$  since  $(T_n p)^{-1} = T_n f \circ (T_{n'} p)^{-1}$  and  $T_n f$  is an [isometry](#) of the Euclidean vector spaces  $T_n\mathcal{N}$  and  $T_{n'}\mathcal{N}$ . It can be shown that  $g$  is smooth. Thus, we have constructed a [metric](#)  $g$  on  $\mathcal{M}$  such that  $p$  is a [Riemannian covering map](#), which is unique. ■

### 4.1.2 Totally Geodesic

A particular interesting condition is the following.

**Definition 4.1.3** (Totally geodesic). A [submanifold](#)  $\mathcal{M}$  of  $(\widetilde{\mathcal{M}}, \widetilde{g})$  is called *totally geodesic* if for all  $m \in \mathcal{M}$  and  $v \in T_m\mathcal{M}$ , the [geodesic](#)  $c$  of  $(\widetilde{\mathcal{M}}, \widetilde{g})$  with  $c(0) = m$  and  $c'(0) = v$  is contained fully in  $\mathcal{M}$ .

**Proposition 4.1.3.** Let  $p: (\mathcal{N}, h) \rightarrow (\mathcal{M}, g)$  be a [Riemannian covering map](#). The [geodesic](#) of  $(\mathcal{M}, g)$  are the projections of the [geodesic](#) in  $(\mathcal{N}, h)$ ; and the [geodesic](#) of  $(\mathcal{N}, h)$  are the liftings of those in  $(\mathcal{M}, g)$ .

**Proof.** Since  $p$  is a [local isometry](#), if  $\gamma$  is a [geodesic](#) of  $\mathcal{N}$ , then  $c = p \circ \gamma$  is also a [geodesic](#) of  $\mathcal{M}$ . From the [uniqueness theorem](#) for [geodesics](#) shows that these are indeed the only [geodesics](#) on  $\mathcal{M}$ . Conversely, if  $p \circ \gamma$  is a [geodesic](#) in  $\mathcal{M}$ , then  $\gamma$  is a [geodesic](#) in  $\mathcal{N}$ . ■

**Example.** In Euclidean spaces, the [totally geodesic submanifold](#) are affine linear subspaces and their open subsets.

**Example.** Each closed [geodesic](#) in [Riemannian manifolds](#) defines a 1-dimensional compact [totally geodesic submanifold](#).

**Example.** The **totally geodesic submanifolds** of  $S^n \subseteq \mathbb{R}^{n+1}$  are the intersections of  $S^n$  with linear subspaces of  $\mathbb{R}^{n+1}$ .

**Example.** In general, **Riemannian manifolds** do not have any **totally geodesic submanifolds** of dimension  $> 1$ .

**Remark.** We will see that  $\mathcal{M}$  is **totally geodesic** in  $\widetilde{\mathcal{M}}$  if and only if all the **second fundamental forms** vanish identically.

## 4.2 The Second Fundamental Form

Let  $f: \mathcal{M}^m \rightarrow \mathcal{N}^n$  be an **immersion** between two **Riemannian manifolds**. We already know that a **metric** on  $\mathcal{N}$  induces a **metric** on  $\mathcal{M}$  naturally by the **immersion** (inclusion). We now ask: given the **Levi-Civita connection**  $\nabla^{\mathcal{N}}$  of  $\mathcal{N}$ , how to get  $\nabla^{\mathcal{M}}$  of  $\mathcal{M}$ ?

**Note.** In the following discussion, we consider  $\mathcal{M}^m \subseteq \mathcal{N}^n$ , i.e., we simply consider the case of inclusion. However, everything works out nicely by identifying  $\mathcal{M}$  with the image of  $f(\mathcal{M})$  in  $\mathcal{N}$ .

### 4.2.1 The Immersion-Induced Levi-Civita Connection

This **immersion-induced Levi-Civita connection** is given by the central object  $(\nabla_X^{\mathcal{N}} Y)^\top$ , where  $\top: T_x \mathcal{N} \rightarrow T_x \mathcal{M}$  for  $x \in \mathcal{M}$  is the orthogonal projection.<sup>1</sup> The formal guarantee is given by **Theorem 4.2.1**.

**Theorem 4.2.1.** For  $X, Y \in \Gamma(T\mathcal{M})$ ,  $\nabla_X^{\mathcal{M}} Y = (\nabla_X^{\mathcal{N}} Y)^\top$ .

**Proof.** Firstly, we have to make sure that the right-hand side is defined. This can be done by extending **vector fields**  $X, Y$  locally to a neighborhood of  $\mathcal{M}$  in  $\mathcal{N}$ . We do this in the **local coordinates** around  $x \in \mathcal{M}$  locally mapping  $\mathcal{M}$  to  $\mathbb{R}^m \subseteq \mathbb{R}^n$ . Specifically, the extension of  $X = \xi^i(x) \partial / \partial x^i$  is

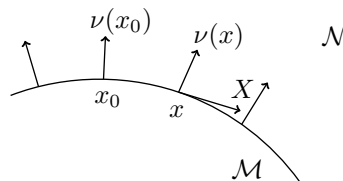
$$\widetilde{X}(x^1, \dots, x^n) = \sum_{i=1}^m \xi^i(x^1, \dots, x^n) \frac{\partial}{\partial x^i}.$$

Then  $\langle \widetilde{X}, \widetilde{Y} \rangle(x) = \langle X, Y \rangle(x)$  and  $[\widetilde{X}, \widetilde{Y}](x) = [X, Y](x)$ . From **Levi-Civita theorem**, the **Koszul formula** holds for both  $\mathcal{N}$  and  $\mathcal{M}$ , hence

- $(\nabla_X^{\mathcal{N}} Y)^\top$  does not depend on the choice of extensions: follows from the fact that the representation of  $\nabla^{\mathcal{N}}$  is done by  $\Gamma$ ;
- $(\nabla_X^{\mathcal{N}} Y)^\top$  defines a **torsion-free connection** on  $\mathcal{M}$ : as  $\nabla_X^{\mathcal{N}} Y - \nabla_Y^{\mathcal{N}} X - [X, Y]$  vanishes, also the tangential (to  $\mathcal{M}$ ) part has to vanish.

■

Let  $\nu(x)$  be a **vector field** in a neighborhood of  $x_0 \in \mathcal{M} \subseteq \mathcal{N}$  that is orthogonal to  $\mathcal{M}$ , i.e.,  $\langle \nu(x), X \rangle = 0$  for all  $X \in T_x \mathcal{M}$ .



<sup>1</sup>We note this one last time: this makes sense since we can identify  $T_x \mathcal{M} \subseteq T_x \mathcal{N}$  by the **immersion**  $f: \mathcal{M} \rightarrow \mathcal{N}$ .

**Notation.** Let  $T_x\mathcal{M}^\perp$  be the orthogonal complement of  $T_x\mathcal{M}$  in  $T_x\mathcal{N}$ .

With this notation, we see that  $\langle \nu(x), X \rangle = 0$  for all  $X \in T_x\mathcal{M}$  means  $\nu(x) \in T_x\mathcal{M}^\perp$ .

**Notation** (Normal bundle). The *normal bundle*  $T\mathcal{M}^\perp$  of  $\mathcal{M}$  in  $\mathcal{N}$  is the **bundle** with **fiber**  $T_x\mathcal{M}^\perp$  of  $x \in \mathcal{M}$ .

**Lemma 4.2.1.**  $(\nabla_X^\mathcal{N}\nu)^\top(x)$  only depends on  $\nu(x)$ .

**Proof.** For a real-valued function  $f$  on a neighborhood of  $x$ , we have

$$(\nabla_X^\mathcal{N}f\nu)^\top(x) = (X(f)(x)\nu(x))^\top + f(x)(\nabla_X^\mathcal{N}\nu)^\top(x) = f(x)(\nabla_X^\mathcal{N}\nu)^\top(x)$$

as  $(X(f)(x)\nu(x))^\top = 0$  with  $\nu(x) \in T_x\mathcal{M}^\perp$ . ■

## 4.2.2 The Second Fundamental Form

With the notations we have developed, we define the following.

**Definition 4.2.1** (Second fundamental tensor). The *second fundamental tensor*  $S: T_x\mathcal{M} \times T_x\mathcal{M}^\perp \rightarrow T_x\mathcal{M}$  of  $\mathcal{M}$  at point  $x \in \mathcal{M}$  is defined by

$$S(X, \nu) = (\nabla_X^\mathcal{N}\nu)^\top.$$

**Lemma 4.2.2.** For  $X, Y \in T_x\mathcal{M}$ ,  $\ell_\nu(X, Y) := \langle S(X, \nu), Y \rangle$  is symmetric in  $X, Y$ .

**Proof.** We see that

$$\begin{aligned} \ell_\nu(X, Y) &= \langle (\nabla_X^\mathcal{N}\nu)^\top, Y \rangle \\ &= \langle \nabla_X^\mathcal{N}\nu, Y \rangle && (Y \in T_x\mathcal{M}) \\ &= -\langle \nu, \nabla_X^\mathcal{N}Y \rangle && (\nabla^\mathcal{N} \text{ is metric and } \langle \nu, Y \rangle = 0) \\ &= -\langle \nu, \nabla_Y^\mathcal{N}X + [X, Y] \rangle && (\nabla^\mathcal{N} \text{ is torsion-free}) \\ &= -\langle \nu, \nabla_Y^\mathcal{N}X \rangle - \langle \nu, [X, Y] \rangle \\ &= -\langle \nu, \nabla_Y^\mathcal{N}X \rangle && (\nu \in T_x\mathcal{M}^\perp, [X, Y] \in T_x\mathcal{M}) \\ &= \langle \nabla_Y^\mathcal{N}\nu, X \rangle && (\nabla^\mathcal{N} \text{ is metric}) \\ &= \langle (\nabla_Y^\mathcal{N}\nu)^\top, X \rangle && (X \in T_x\mathcal{M}) \\ &= \ell_\nu(Y, X). \end{aligned}$$
■

**Definition 4.2.2** (Second fundamental form). The *second fundamental form*  $\ell_\nu(\cdot, \cdot)$  of  $\mathcal{M}$  w.r.t.  $\mathcal{N}$  is defined as  $\ell_\nu(X, Y) := \langle S(X, \nu), Y \rangle$ .

**Note.** do Carmo [FC13] defines the **second fundamental form** as  $\ell_\nu(X, X)$ .

**Note** (First fundamental form). The *first fundamental form* is the **metric** applied to  $X, Y \in T_x\mathcal{M}$ , i.e.,  $\langle X, Y \rangle$ .

Now, fix a **normal field**  $\nu$ , and let  $S_\nu(X) := S(X, \nu)$ , then  $S_\nu: T_x\mathcal{M} \rightarrow T_x\mathcal{M}$  is self-adjoint w.r.t. the **metric**  $\langle \cdot, \cdot \rangle$  by **Lemma 4.2.2**.

### 4.2.3 Curvatures and Second Fundamental Forms

Due to the time, we can only talk about the definition of the following. For a detailed discussion, see [FC13, §6 Example 2.4 – Example 2.8].

**Definition.** Assume that  $\langle \nu, \nu \rangle \equiv 1$ , i.e.,  $\nu$  is the unit **normal field**, then  $S_\nu$  has  $m$  real eigenvalues.

**Definition 4.2.3 (Principal curvature).** The eigenvalues are called *principal curvatures* of  $\mathcal{M}$  in direction  $\nu$ .

**Definition 4.2.4 (Principal curvature vector).** The corresponding eigenvectors are called *principal curvature vectors* of  $\mathcal{M}$  in direction  $\nu$ .

**Definition 4.2.5 (Mean curvature).** The *mean curvature* of  $\mathcal{M}$  in direction  $\nu$  is defined by

$$H_\nu := \frac{1}{m} \operatorname{Tr} S_\nu.$$

**Definition 4.2.6 (Gauss-Kronecker curvature).** The *Gauss-Kronecker curvature* of  $\mathcal{M}$  in direction  $\nu$  is defined by

$$K_\nu := \det S_\nu.$$

## Lecture 15: The Second Fundamental Form

### 4.2.4 Totally Geodesic and Second Fundamental Form

23 Feb. 13:00

Let  $\dim \mathcal{N} = m + 1$ ,  $\dim \mathcal{M} = m$ , then for all  $x \in \mathcal{M}$ , there are exactly 2 normal vectors  $\nu \in T_x \mathcal{M}^\perp$  with  $\langle \nu, \nu \rangle \equiv 1$ , i.e.,  $\nabla_X^\mathcal{N} \nu$  always tangential to  $\mathcal{M}$ . Now, we fix locally such a **normal field** and drop the subscript  $\nu$  in the following discussion.

**Note.** If we choose an opposite **normal field**, then  $\ell$ ,  $S$ , and **mean curvature** will change their sign. However, for even  $m$ , the **Gauss-Kronecker curvature** does not depend on the choice of the direction of  $\nu$ .

**Intuition.**  $\nabla_X^\mathcal{N} \nu$  measures the “tilting velocity” with which  $\nu$  is tilted relative to a fixed **parallel vector field** in  $\mathcal{N}$ , when on  $\mathcal{M}$  in direction  $X$ .

**Theorem 4.2.2.** Given  $\mathcal{M} \subseteq \widetilde{\mathcal{M}}$ , then  $\mathcal{M}$  is **totally geodesic** in  $\widetilde{\mathcal{M}}$  if and only if all **second fundamental form** of  $\mathcal{M}$  vanish identically.

**Proof.** Let  $c: I \rightarrow \mathcal{M}$  be a **geodesic** in  $\mathcal{M}$ , i.e.,  $\nabla_{\dot{c}}^\mathcal{M} \dot{c} = 0$ . By **Theorem 4.2.1**,  $\nabla_{\dot{c}}^\mathcal{M} \dot{c} = (\nabla_{\dot{c}}^{\widetilde{\mathcal{M}}} \dot{c})^\top = 0$ , implying  $c$  is a **geodesic** in  $\widetilde{\mathcal{M}}$  if and only if  $(\nabla_{\dot{c}}^{\widetilde{\mathcal{M}}} \dot{c})^\top = 0$ , i.e.,  $\langle \nabla_{\dot{c}}^{\widetilde{\mathcal{M}}} \dot{c}, \nu \rangle = 0$  for all  $\nu \in T\mathcal{M}^\perp$ . Notice that

- $\langle \dot{c}, \nu \rangle = 0$ , and hence
- $\dot{c} \langle \dot{c}, \nu \rangle = \langle \nabla_{\dot{c}}^{\widetilde{\mathcal{M}}} \dot{c}, \nu \rangle + \langle \dot{c}, \nabla_{\dot{c}}^{\widetilde{\mathcal{M}}} \nu \rangle = 0$ .

In all, we have  $0 = \langle \nabla_{\dot{c}}^{\widetilde{\mathcal{M}}} \dot{c}, \nu \rangle = -\langle \dot{c}, \nabla_{\dot{c}}^{\widetilde{\mathcal{M}}} \nu \rangle = -\ell_\nu(\dot{c}, \dot{c})$ , proving the theorem. ■

**Note.** **Theorem 4.2.2** also holds for **Lorentzian manifolds**  $(\widetilde{\mathcal{M}}, \widetilde{g})$ .



**Example** (Initial value problem for Einstein equations). Given a  $(\widetilde{\mathcal{M}}^4, \widetilde{g})$  a Lorentzian manifold satisfying Einstein equations, and a  $(\mathcal{M}^3, g)$  non-degenerate Riemannian manifold. If the second fundamental form of  $\mathcal{M}^3$  in  $\widetilde{\mathcal{M}}^4$  vanishes identically, then  $\mathcal{M}^3$  is totally geodesic.<sup>a</sup>

<sup>a</sup>This is just a special case of Theorem 4.2.2; in general, it does not vanish.

Theorem 4.2.2 allows us to get what is probably the best geometric interpretation of sectional curvature. Let  $\mathcal{M}$  be a Riemannian manifold and let  $p \in \mathcal{M}$ . Let  $B \subseteq T_p\mathcal{M}$  be an open ball in  $T_p\mathcal{M}$  on which  $\exp_p$  is a diffeomorphism, and let  $\sigma \subseteq T_p\mathcal{M}$  be a subspace of dimension 2. Then,  $\exp_p(\sigma \cap B) = S$  is a submanifold of dimension 2 of  $\mathcal{M}$  passing through  $p$ .

**Intuition.**  $S$  is the surface formed by “small” geodesics that start from  $p$  and are tangent to  $\sigma$  at  $p$ .

**Note.** By Theorem 4.2.2,  $S$  is geodesic at  $p$ , hence the second fundamental forms of the inclusion  $\iota: S \subseteq \mathcal{M}$  vanish at  $p$ .

As a submanifold of  $\mathcal{M}$ ,  $S$  has an induced Riemannian metric whose Gauss curvature at  $p$  will be denoted by  $K_S$ . It follows from the Gauss formula [FC13, §6 Theorem 2.5]<sup>2</sup> that

$$K_S(p) = K(p, \sigma),$$

i.e., the sectional curvature  $K(p, \sigma)$  is the Gauss curvature, at  $p$ , of a small surface formed by geodesics of  $\mathcal{M}$  that start from  $p$  and are tangent to  $\sigma$ .

**Remark.** This was exactly the way in which Riemann defined sectional curvature.

### 4.3 The Fundamental Equations

Given an isometric immersion  $f: \mathcal{M}^m \rightarrow \mathcal{N}^n$  with  $n = m + k$ , at each  $p \in \mathcal{M}$ , we have

$$T_p\mathcal{N} = T_p\mathcal{M} \oplus (T_p\mathcal{M})^\perp,$$

which varies differentiably with  $p$ .

**Intuition.** Locally, the portion of the tangent bundle  $T\mathcal{N}$  which sits over  $\mathcal{M}$  can be decomposed into the direct sum of the tangent bundle  $T\mathcal{M}$  and the normal bundle  $T\mathcal{M}^\perp$ .

Everything about immersions occurs as if the geometry decomposes into two geometries: the geometry of the tangent bundle and the geometry of the normal bundle, and these geometries are related by the second fundamental form of the immersions.

**Notation.** Greek indices  $(\alpha, \beta, \dots)$  occurring twice are summed from 1 to  $k$  for  $X, Y, Z, W \in T_x\mathcal{M}$ .

**Theorem 4.3.1** (Gauss' equations). Let  $\mathcal{N}$  be a Riemannian manifold with  $\dim \mathcal{N} = n$ , and let  $\mathcal{M} \subseteq \mathcal{N}$  be a submanifold with  $\dim \mathcal{M} = m$ . Let  $k = n - m$ , and  $x \in \mathcal{M}$ ,  $\nu_1, \dots, \nu_k$  be an orthonormal basis of  $(T_x\mathcal{M})^\perp$ ,  $S_\alpha := S_{\nu_\alpha}$ ,  $\ell_\alpha := \ell_{\nu_\alpha}$ ,  $\alpha = 1, \dots, k$ . Then,

$$R^\mathcal{M}(X, Y)Z - (R^\mathcal{N}(X, Y)Z)^\top = \ell_\alpha(Y, Z)S_\alpha(X) - \ell_\alpha(X, Z)S_\alpha(Y).$$

Thus, we also have

$$\langle R^\mathcal{M}(X, Y)Z, W \rangle - \langle R^\mathcal{N}(X, Y)Z, W \rangle = \ell_\alpha(Y, Z)\ell_\alpha(X, W) - \ell_\alpha(X, Z)\ell_\alpha(Y, W).$$

**Proof.** We can extend  $X, Y, Z, W$ , and  $\nu_1, \dots, \nu_k$  to vector fields in  $T\mathcal{M}$  and  $T\mathcal{M}^\perp$ , respectively. Let

<sup>2</sup>Which is just a special case of Gauss' equations.

$\nu_\alpha$  be orthonormal, then

$$\nabla_Y^{\mathcal{N}} Z = (\nabla_Y^{\mathcal{N}} Z)^\top = (\nabla_X^{\mathcal{N}} Z)^\perp = \nabla_Y^{\mathcal{M}} Z + \langle \nu_\alpha, \nabla_Y^{\mathcal{N}} Z \rangle \nu_\alpha$$

as  $\nu_\alpha$  form orthonormal basis of  $T\mathcal{M}^\perp$ . Hence,

$$\nabla_X^{\mathcal{N}} \nabla_Y^{\mathcal{N}} Z = \nabla_X^{\mathcal{N}} \nabla_Y^{\mathcal{M}} Z + X(\langle \nu_\alpha, \nabla_Y^{\mathcal{N}} Z \rangle) \nu_\alpha + \langle \nu_\alpha, \nabla_Y^{\mathcal{N}} Z \rangle \nabla_X^{\mathcal{N}} \nu_\alpha.$$

Then,

$$(\nabla_X^{\mathcal{N}} \nabla_Y^{\mathcal{N}} Z)^\top = \nabla_X^{\mathcal{M}} \nabla_Y^{\mathcal{M}} Z + \underbrace{\langle \nu_\alpha, \nabla_Y^{\mathcal{N}} Z \rangle}_{-\ell_\alpha(Y, Z)} \underbrace{(\nabla_X^{\mathcal{N}} \nu_\alpha)^\top}_{S_\alpha(X)} = \nabla_X^{\mathcal{M}} \nabla_Y^{\mathcal{M}} Z - \ell_\alpha(Y, Z) S_\alpha(X).$$

Analogously, we have

$$(\nabla_Y^{\mathcal{N}} \nabla_X^{\mathcal{N}} Z)^\top = \nabla_Y^{\mathcal{M}} \nabla_X^{\mathcal{M}} Z - \ell_\alpha(X, Z) S_\alpha(Y),$$

and also, we have

$$(\nabla_{[X, Y]}^{\mathcal{N}} Z)^\top = \nabla_{[X, Y]}^{\mathcal{M}} Z.$$

By collecting terms, we have

$$\begin{aligned} & (\nabla_X^{\mathcal{N}} \nabla_Y^{\mathcal{N}} Z)^\top - (\nabla_Y^{\mathcal{N}} \nabla_X^{\mathcal{N}} Z)^\top - (\nabla_{[X, Y]}^{\mathcal{N}} Z)^\top \\ &= \nabla_X^{\mathcal{M}} \nabla_Y^{\mathcal{M}} Z - \nabla_Y^{\mathcal{M}} \nabla_X^{\mathcal{M}} Z - \nabla_{[X, Y]}^{\mathcal{M}} Z - \ell_\alpha(Y, Z) S_\alpha(X) + \ell_\alpha(X, Z) S_\alpha(Y), \end{aligned}$$

equivalently,

$$R^{\mathcal{M}}(X, Y)Z - (R^{\mathcal{N}}(X, Y)Z)^\top = \ell_\alpha(Y, Z) S_\alpha(X) - \ell_\alpha(X, Z) S_\alpha(Y).$$

■

**Theorem 4.3.1** tells us that for a surface  $\mathcal{M}$  in  $\mathbb{R}^3$ , the **Gauss-Kronecker curvature** coincides with the **Riemannian curvature** of  $\mathcal{M}$ , which is independent of the **embedding**. Therefore, **Gauss-Kronecker curvature** does not depend on **embeddings** of  $\mathcal{M}$  into  $\mathbb{R}^3$ .

**Remark (Codazzi equations).** Let  $\mathcal{M}^m \subseteq \mathcal{N}^{m+1}$  where  $N$  is unit normal on  $\mathcal{M}$ . Then, the *Codazzi equations* is defined as

$$\langle R(X, Y)e_j, N \rangle = (\nabla_X^{\mathcal{M}} \ell)(Y, e_j) - (\nabla_Y^{\mathcal{M}} \ell)(X, e_j) = X^k Y^i \nabla_k^{\mathcal{M}} \ell_{ij} - Y^k X^i \nabla_k^{\mathcal{M}} \ell_{ij}, \quad (4.1)$$

i.e.,  $\langle R(X, Y)Z, N \rangle = (\nabla_X^{\mathcal{M}} \ell)(Y, Z) - (\nabla_Y^{\mathcal{M}} \ell)(X, Z)$ .

The **Codazzi equations**, together with **Gauss' equations**, form the fundamental equations of the local theory of **isometric immersions**.

# Chapter 5

## Jacobi Fields

### Lecture 16: Jacobi Field

In this chapter, we derive a first relation between the two basic concepts introduced, i.e., [geodesics](#) and [curvatures](#). This is done by introducing [Jacobi field](#): [vector fields along geodesics](#), defined by means of differential equations naturally from [exponential map](#). Moreover, [Jacobi fields](#) allow us to obtain a simple characterization of the singularities of the [exponential map](#).

7 Mar. 13:00

**Intuition.** The upshot is, as we will see, the [curvature](#)  $K(p, \sigma)$ ,  $\sigma \subseteq T_p \mathcal{M}$ , determines how fast the [geodesics](#), that start from  $p$  and are tangent to  $\sigma$ , spread apart. This is described by [Jacobi field](#).

### 5.1 Jacobi Fields

As mentioned, we want to consider neighboring [geodesics](#) under a [vector field along which](#), and study how do they move. Their behaviors are essentially governed by [curvature](#).

**Definition 5.1.1** (Jacobi field). Let  $\mathcal{M}$  be a  $d$ -dimensional [Riemannian manifold](#). Let  $c: I \rightarrow \mathcal{M}$  be a [geodesic](#). A [vector field](#)  $X$  along  $c$  is called a *Jacobi field* if it satisfies the *Jacobi equation*

$$\nabla_{\frac{d}{dt}} \nabla_{\frac{d}{dt}} X + R(X, \dot{c})\dot{c} = 0. \quad (5.1)$$

**Notation.** We write  $\dot{X} := \nabla_{\frac{d}{dt}} X$  and  $\ddot{X} := \nabla_{\frac{d}{dt}} \nabla_{\frac{d}{dt}} X$ .

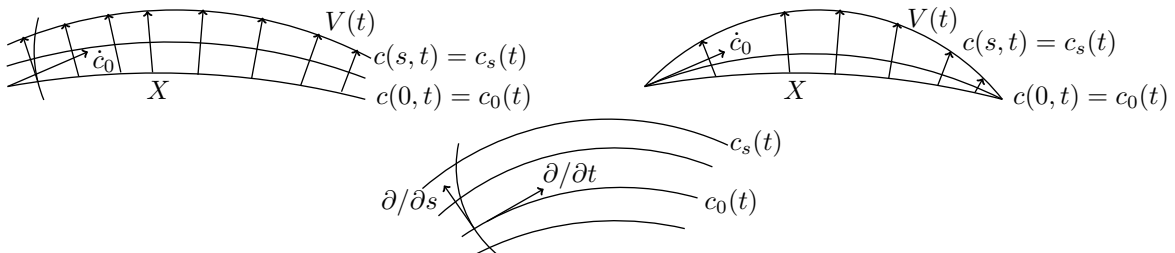
Using new notations, the [Jacobi equation](#) is rewritten as

$$\ddot{X} + R(X, \dot{c})\dot{c} = 0.$$

To understand [Jacobi field](#), we first recall the [variation](#).

**As previously seen** (Variation). For some  $\epsilon > 0$ , the *variation* of a [smooth curve](#)  $c: [a, b] \rightarrow \mathcal{M}$  is a differentiable map  $F: [a, b] \times (-\epsilon, \epsilon) \rightarrow \mathcal{M}$  such that  $F(t, 0) = c(t)$  for  $t \in [a, b]$  with  $s \in (-\epsilon, \epsilon)$ .

Essentially, a [Jacobi field](#) studies the [variation of geodesics](#): we can label [geodesics](#)  $c$  as



**Notation** (Proper variation). A *proper variation* is a [variation](#) where the endpoints are fixed, i.e.,  $F(a, s) = c(a)$  and  $F(b, s) = c(b)$  for all  $s \in (-\epsilon, \epsilon)$ .

**Note.** We might either fix the endpoints or left them open, i.e., we can consider both [proper](#) and [non-proper](#) cases.

**Intuition.** The [Jacobi equation](#) can be viewed as the linearization of the [geodesic equation](#).

Formally, we define the following.

**Definition 5.1.2** (Geodesic variation). Let  $\mathcal{M}$  be a [\(semi-\)Riemannian manifold](#). A [variation of curves](#)  $c: I \times (-\epsilon, \epsilon) \rightarrow \mathcal{M}$  is called a *geodesic variation* if for all  $s \in (-\epsilon, \epsilon)$ , the [curve](#)  $t \mapsto c_s(t) := c(t, s)$  is a [geodesic](#).

**Notation.** We set  $c_s(t) = c(t, s) = F(t, s)$ , and

- $\dot{c}(t, s) := \frac{\partial}{\partial t} c(t, s)$ , i.e.,  $dF(\partial/\partial t)c(t, s)$ ;
- $c'(t, s) := \frac{\partial}{\partial s} c(t, s)$ , i.e.,  $dF(\partial/\partial s)c(t, s)$ .

## 5.2 Variations of Length and Energy

Recall the following.

**As previously seen.** Given a [variation](#) of a [geodesic](#)  $c_s(t)$ , The [energy](#) for  $c_s$  is defined as

$$E(s) := \frac{1}{2} \int_a^b \left\langle \frac{\partial c(t, s)}{\partial t}, \frac{\partial c(t, s)}{\partial t} \right\rangle dt,$$

and the [length](#) for  $c_s$  is defined as

$$L(s) := \int_a^b \left\langle \frac{\partial c(t, s)}{\partial t}, \frac{\partial c(t, s)}{\partial t} \right\rangle^{1/2} dt.$$

And we want to compute

- the first [variations](#)  $E'(0)$  and  $L'(0)$ , i.e., the first derivatives;
- for  $c = c_0$  [geodesic](#), compute the second [variations](#)  $E''(0)$  and  $L''(0)$ , i.e., the second derivatives.

### 5.2.1 First Variations

Let's consider the first [variations](#), i.e.,  $E'(0)$  and  $L'(0)$ .

**Lemma 5.2.1.** If  $L(s)$ ,  $E(s)$  are differentiable w.r.t.  $s$ , then

$$L'(0) = \int_a^b \left( \frac{\frac{\partial}{\partial t} \langle c', \dot{c} \rangle}{\langle \dot{c}, \dot{c} \rangle^{1/2}} - \frac{\langle c', \nabla_{\frac{\partial}{\partial t}} \dot{c} \rangle}{\langle \dot{c}, \dot{c} \rangle^{1/2}} \right) dt,$$

and

$$E'(0) = \langle c'(b, 0), \dot{c}(b, 0) \rangle - \langle c'(a, 0), \dot{c}(a, 0) \rangle - \int_a^b \left\langle \frac{\partial c}{\partial s}, \nabla_{\frac{\partial}{\partial t}} \frac{\partial c}{\partial t}(t, s) \right\rangle dt.$$

**Proof.** We have already proved this in different notations. ■

**Note.** If  $c = c_0$  is parametrized proportionally to the arc-length, i.e.,  $\|\dot{c}(t, 0)\|$  is a constant. Then  $L'(0)$  becomes

$$L'(0) = \frac{1}{\langle \dot{c}, \dot{c} \rangle^{1/2}} \left( \langle c', \dot{c} \rangle \Big|_{t=a, s=0}^{t=b, s=0} - \int_a^b \langle c', \nabla_{\frac{\partial}{\partial t}} \dot{c} \rangle dt \right).$$

If we consider the fixed endpoints case (i.e., **proper variation**), we observe that  $E$  and  $L$  are stationary if and only if

$$\nabla_{\frac{\partial}{\partial t}} \dot{c}(t, 0) = 0,$$

i.e., when  $c$  is a **geodesic**.

### 5.2.2 Second Variations

Now, let  $c = c_0$  be a **geodesic**. Then we compute the second derivatives w.r.t.  $s$  of  $E$  and  $L$  at  $s = 0$ .

**Theorem 5.2.1.** Let  $c: [a, b] \rightarrow \mathcal{M}$  be a **geodesic**. Then

$$E''(0) = \int_a^b \left\langle \nabla_{\frac{\partial}{\partial t}} c'(t, 0), \nabla_{\frac{\partial}{\partial t}} c'(t, 0) \right\rangle dt - \int_a^b \left\langle R(\dot{c}, c') c', \dot{c} \right\rangle dt \Big|_{s=0} + \left\langle \nabla_{\frac{\partial}{\partial s}} c', \dot{c} \right\rangle \Big|_{t=a, s=0}^{t=b, s=0}.$$

By letting  $c'^\perp := c' - \left\langle \frac{\dot{c}}{\|\dot{c}\|}, c' \right\rangle \frac{\dot{c}}{\|\dot{c}\|}$ ,<sup>a</sup> we have

$$L''(0) = \frac{1}{\|\dot{c}\|} \left( \int_a^b \left\langle \nabla_{\frac{\partial}{\partial t}} c'^\perp, \nabla_{\frac{\partial}{\partial t}} c'^\perp \right\rangle dt - \int_a^b \left\langle R(\dot{c}, c'^\perp) c'^\perp, \dot{c} \right\rangle dt + \left\langle \nabla_{\frac{\partial}{\partial s}} c', \dot{c} \right\rangle \Big|_{t=a}^{t=b} \right) \Big|_{s=0}.$$

<sup>a</sup>I.e., the component of  $c'$  orthogonal to  $\dot{c}$ .

**Remark.** By keeping the endpoints fixed, if the **sectional curvature** of  $\mathcal{M}$  is non-positive, then the **Riemannian curvature** in  $E''(0)$  and  $L''(0)$  are non-negative. This implies  $E''(0) > 0$ , then  $E(c_s) > E(c_0)$  for small  $|s|$ .

**Corollary 5.2.1.** On a manifold with non-positive **sectional curvature**, the **geodesics** with fixed endpoints are always locally minimizing.

## 5.3 Index Form

### 5.3.1 Pullback Connections

Let  $\mathcal{M}$  be a **Riemannian manifold** of dimension  $d$ , and  $\mathcal{H}$  be a **differentiable manifold**.<sup>1</sup> Let  $f: \mathcal{H} \rightarrow \mathcal{M}$ , smooth, and  $f$  may not be injective. We ask the following question.

**Problem 5.3.1.** What is the **tangent space** of  $f(\mathcal{H})$  of point  $p \in f(\mathcal{H})$ ?

We see that even if  $f$  is an **immersion**, since it can be non-injective, there maybe issues.

**Example.** Let  $p = f(x) = f(y)$  for  $x \neq y$ . For  $f$  being an **immersion**, we may restrict  $f$  to a sufficiently small neighborhood  $U, V$  at  $x, y$ , respectively, such that  $f(U), f(V)$  have well-defined **tangent spaces** at  $p$ . Then, in a double point (e.g.,  $p$ ) of  $f(\mathcal{H})$ , the **tangent space** can be specified by specifying the preimage ( $x$  or  $y$ ).

Formally, consider  $f^*(T\mathcal{M})$ , the **tangent bundle**  $T\mathcal{M}$  **pullback** by  $f$ .

<sup>1</sup>Often times,  $\mathcal{H}$  is an interval  $I \subseteq \mathbb{R}$  or a square  $I \times I \subseteq \mathbb{R}^2$ .

**Note.** The [fiber](#) over  $x \in \mathcal{H}$  is  $T_{f(x)}\mathcal{M}$ .

Then, we can introduce a [connection](#)  $f^*(\nabla)$  on  $f^*(T\mathcal{M})$ : let  $X \in T_x\mathcal{H}$ ,  $Y$  a [section](#) of  $f^*(T\mathcal{M})$ . Set

$$(f^*\nabla)_X Y := \nabla_{df(X)} Y,$$

where  $f^*(T\mathcal{M})_x$  is identified with  $T_{f(x)}\mathcal{M}$  with  $\nabla$  for  $f^*\nabla$ .

**Note.** For  $\nabla_{df(X)} Y$  to be well-defined, we need to extend  $Y$  to a neighborhood of  $f(\mathcal{H})$ .<sup>a</sup>

<sup>a</sup>Hence, it does not depend on the choice of extension.

**Notation.** Write  $\nabla$  for  $f^*\nabla$  in what follows.

### 5.3.2 Index Form

Now, let  $f = c: I \rightarrow \mathcal{M}$  (often,  $c$  is a [geodesic](#)), i.e., we consider [vector field along  \$c\$](#) .<sup>2</sup> Specifically, let  $X$  be a [vector field along  \$c\$](#)  where  $c$  is a [geodesic](#). Then, there exists a [geodesic variation](#)

$$c: [a, b] \times (-\epsilon, \epsilon) \rightarrow \mathcal{M}$$

of  $c(t)$  with  $\frac{\partial c}{\partial s}|_{s=0} = X$ . Consider the second [variation](#) of [energy](#): inspired from [Theorem 5.2.1](#), we write

$$I(X, X) := \int_a^b \left( \langle \nabla_{\frac{\partial}{\partial t}} X, \nabla_{\frac{\partial}{\partial t}} X \rangle - \langle R(\dot{c}, X)X, \dot{c} \rangle \right) dt,$$

i.e.,  $I(X, X) = \frac{d^2}{ds^2} E(0)$  if  $X(a) = X(b) = 0$ . Moreover, instead of considering a 1-parameter [variation](#), we can also consider a 2-parameter [variation](#) on  $X$  and  $Y := \frac{\partial c}{\partial t}$ . In this case, we propose the following.

**Definition 5.3.1** (Index form). The *index form* of a [geodesic](#)  $c$  on  $X = \frac{\partial c}{\partial s}|_{s=0}$  and  $Y = \frac{\partial c}{\partial t}$  is

$$I(X, Y) := \int_a^b \left( \langle \nabla_{\frac{\partial}{\partial t}} X, \nabla_{\frac{\partial}{\partial t}} Y \rangle - \langle R(\dot{c}, X)Y, \dot{c} \rangle \right) dt.$$

**Note.** We see that  $I(X, Y)$  is a bilinear and symmetric in  $X, Y$ .

**As previously seen.** Recall the [Jacobi equation](#), i.e.,  $\nabla_{\frac{d}{dt}} \nabla_{\frac{d}{dt}} X + R(X, \dot{c})\dot{c} = 0$ .

**Proposition 5.3.1** (Jacobi field). A [vector field  \$X\$  along a geodesic  \$c: \[a, b\] \rightarrow \mathcal{M}\$](#)  is a [Jacobi-field](#) if and only if the [index form](#) of  $c$  satisfies  $I(X, Y) = 0$  for all [vector fields  \$Y\$  along  \$c\$](#)  with  $Y(a) = Y(b) = 0$ .

**Proof.** Observe that

$$\begin{aligned} I(X, Y) &= \int_a^b \left( \langle \nabla_{\frac{\partial}{\partial t}} X, \nabla_{\frac{\partial}{\partial t}} Y \rangle - \langle R(\dot{c}, X)Y, \dot{c} \rangle \right) dt \\ &= \int_a^b \left( \langle \nabla_{\frac{\partial}{\partial t}} X, \nabla_{\frac{\partial}{\partial t}} Y \rangle - \langle R(X, \dot{c})\dot{c}, Y \rangle \right) dt = \int_a^b \left( \langle -\nabla_{\frac{\partial}{\partial t}} \nabla_{\frac{\partial}{\partial t}} X, Y \rangle - \langle R(X, \dot{c})\dot{c}, Y \rangle \right) dt, \end{aligned} \quad (5.2)$$

where the second inequality follows from the fact that  $\nabla$  is [Riemannian](#), hence

$$\nabla_{\frac{d}{dt}} \langle \nabla_{\frac{d}{dt}} X, Y \rangle = \langle \nabla_{\frac{d}{dt}} \nabla_{\frac{d}{dt}} X, Y \rangle + \langle \nabla_{\frac{d}{dt}} X, \nabla_{\frac{d}{dt}} Y \rangle,$$

<sup>2</sup>In deed, a [section](#) of  $f^*(T\mathcal{M})$  is a *vector field along  $f$*  in general even for  $f: I^2 \rightarrow \mathcal{M}$ .

with  $Y(a) = 0 = Y(b)$ ,

$$\int_a^b \nabla_{\frac{d}{dt}} \langle \nabla_{\frac{d}{dt}} X, Y \rangle dt = \langle \nabla_{\frac{d}{dt}} X, Y \rangle \Big|_a^b = 0,$$

so

$$\int_a^b \langle \nabla_{\frac{d}{dt}} \nabla_{\frac{d}{dt}} X, Y \rangle dt = - \int_a^b \langle \nabla_{\frac{d}{dt}} X, \nabla_{\frac{d}{dt}} Y \rangle dt.$$

We see that the right-hand side of Equation 5.2 vanishes for every  $Y$  if and only if

$$\nabla_{\frac{d}{dt}} \nabla_{\frac{d}{dt}} X + R(X, \dot{c})\dot{c} = 0,$$

which is just the [Jacobi equation](#), so the result follows.  $\blacksquare$

**Intuition.** Proposition 5.3.1 is really where the [Jacobi equation](#) comes from.

**Remark.** do Carmo [FC13] introduce [Jacobi equation](#) slightly differently, but it's basically the same.

## Lecture 17: Jacobi Fields and General Relativity

**Lemma 5.3.1.** A [vector field along a geodesic](#)  $c: [a, b] \rightarrow \mathcal{M}$  is a [Jacobi field](#) if and only if it is a critical point of  $I(X, X)$  w.r.t. all [variations](#) with fixed endpoints, i.e.,

$$\left. \frac{d}{ds} I(X + sY, X + sY) \right|_{s=0} = 0$$

for every [vector field along  \$c\$](#)  with  $Y(a) = Y(b) = 0$ .

**Proof.** We just use the proof of Proposition 5.3.1 with the fact that

$$\left. \frac{d}{ds} I(X + sY, X + sY) \right|_{s=0} = 2 \int_a^b \left( -\langle \nabla_{\frac{d}{dt}} \nabla_{\frac{d}{dt}} X, Y \rangle - \langle R(X, \dot{c})\dot{c}, Y \rangle \right) dt.$$

**Remark.** Lemma 5.3.1 tells us that the [Jacobi equation](#) is the [Euler-Lagrange equations](#) for  $I(X) := I(X, X)$ .

### 5.3.3 Existence and Uniqueness of Jacobi Fields

Given the initial data, how can we characterize the [Jacobi equation](#) on a [Riemannian manifold](#)  $(\mathcal{M}, g)$  with  $\dim \mathcal{M} = d$ ? Firstly, we know that the [Jacobi equation](#) is a system of  $d$  linear second order ODE.

**Theorem 5.3.1.** Let  $c: [a, b] \rightarrow \mathcal{M}$  be a [geodesic](#). For all  $v, w \in T_{c(a)}\mathcal{M}$ , there exists a unique [Jacobi field](#)  $X$  along  $c$  with  $X(a) = v$ ,  $\dot{X}(a) = w$ .

**Proof.** Let  $\{v_i\}_{i=1}^d$  be an orthonormal basis of  $T_{c(a)}\mathcal{M}$ . Let  $\{X_i\}_{i=1}^d$  be [parallel vector field along  \$c\$](#)  with  $X_i(a) = v_i$  for  $i = 1, \dots, d$ . Then for all  $t \in [a, b]$ ,  $X_1(t), \dots, X_d(t)$  is an orthonormal basis of  $T_{c(t)}\mathcal{M}$ . Choose arbitrary [vector field  \$X\$  along  \$c\$](#)  as  $X = \xi^i X_i$ , i.e.,  $\xi^i(t) = \langle X(t), X_i(t) \rangle$ . As [vector fields](#)  $X_i$ 's are [parallel](#), we have

$$\nabla_{\frac{d}{dt}} X = \frac{d\xi^i}{dt} X_i + \underbrace{\xi_i \nabla_{\frac{d}{dt}} X_i}_0 = \frac{d\xi^i}{dt} X_i \Rightarrow \nabla_{\frac{d}{dt}} \nabla_{\frac{d}{dt}} X = \frac{d^2 \xi^i}{dt^2} X_i.$$

To write the **Jacobi equation** in these coordinates, we first write the **curvature** as

$$R(X, \dot{c})\dot{c} = \xi^i \rho_i^k X_k,$$

where we let  $\rho_i^k := \langle R(X_i, \dot{c})\dot{c}, X_k \rangle$ ,<sup>a</sup> i.e.,  $R(X_i, \dot{c})\dot{c} = \rho_i^k X_k$ . Then, the **Jacobi equation** becomes

$$\left( \frac{d^2 \xi^k}{dt^2} + \xi^i \rho_i^k \right) X_k = 0 \Rightarrow \frac{d^2 \xi^k(t)}{dt^2} + \xi^i(t) \rho_i^k(t) = 0, \quad k = 1, \dots, d$$

since  $\{X_i\}$  is a orthonormal basis. Then, by the linear algebra and ODE theory, we have existence and uniqueness. ■

<sup>a</sup> $\rho_i^k$  is sometimes referred to *rotation*.

Let's see some examples of **Jacobi fields**.

**Example ( $\mathbb{R}^n$ ).** Since the **geodesics** are “straight lines”, consider the **Jacobi field**  $X$  along straight line  $c$  with  $X(a) = v, \dot{X}(a) = w$ . Let  $V(t), W(t)$  be **parallel vector fields along  $c$**  with  $V(a) = v, W(a) = w$ , by linearizing, we have

$$X(t) = V(t) + (t - a)W(t).$$

**Example ( $S^n \subseteq \mathbb{R}^{n+1}$ ).** Let  $c: [0, T] \rightarrow S^n$  be a **geodesic** with  $\|\dot{c}\| = 1$ , and  $v, w \in T_{c(0)}S^n$ ,  $V, W$  **parallel vector fields along  $c$**  with  $V(0) = v, W(0) = w$ . Also, assume that  $\langle v, \dot{c}(0) \rangle = 0 = \langle w, \dot{c}(0) \rangle$ , then the **Jacobi field**  $X$  is

$$X(t) = V(t) \cos t + W(t) \sin t.$$

**Proof.** We see that

$$\dot{X}(t) = -V(t) \sin t + W(t) \cos t,$$

and

$$\ddot{X}(t) = -V(t) \cos t - W(t) \sin t.$$

By using the **Riemannian curvature** on  $S^n$ , we have

$$R(X, \dot{c})\dot{c} = \underbrace{\langle \dot{c}, \dot{c} \rangle}_1 X - \underbrace{\langle X, \dot{c} \rangle}_0 \dot{c} = X.$$

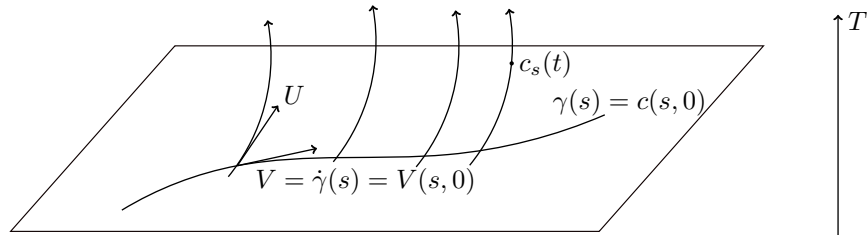
Then  $\ddot{X} + R(X, \dot{c})\dot{c} = 0$ . ⊛

**Remark.** We can also consider  $S_\rho^n \subseteq \mathbb{R}^{n+1}$  with  $\|\dot{c}\| = 1$  and play the above game, i.e., by letting

$$X(t) = V(t) \cos \frac{t}{\rho} + W(t) \sin \frac{t}{\rho}.$$

### 5.3.4 Application of General Relativity

We take a quick detour to see a huge break through in general relativity related to **Jacobi field**. Consider the universe as a  $(\mathcal{M}^4, g)$  a **Lorentzian manifold**,



Here, we have  $[\partial/\partial s, \partial/\partial t] = [U, V] = 0$ . Hence, the **Jacobi equation** is now

$$\nabla_U^2 V + R(V, U, U) = 0.$$



For given  $U$ , the right-hand side defines of each  $p \in \mathcal{M}$  a linear map

$$N \mapsto R(N, U)U$$

for  $N$  unit normal of subspace of  $T_p\mathcal{M}$  perpendicular to  $U$ .<sup>3</sup> Hence, locally,

- the gravitational field  $g$ , the “fields strengths”  $\Gamma$  can be transformed away;
- **variation** of gravitational fields strengths can be described by **Riemannian curvature tensor**, hence cannot be transformed away.

All these imply that the **Jacobi equation** with **Riemannian curvature tensor** can describe the relative accelerations (or field forces) of nearby **geodesics**.

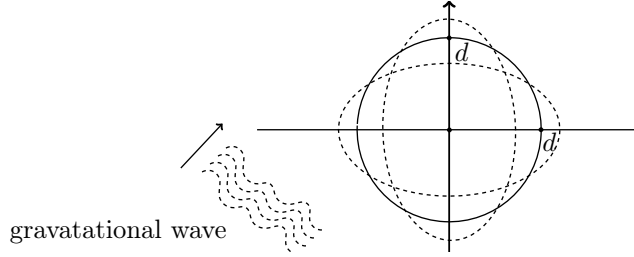


Figure 5.1: LIGO [Abb+16],  $\frac{\Delta\lambda}{\lambda} \approx 10^{-21}$ .

## 5.4 Jacobi Fields and Geodesics

Consider a **Jacobi field** transversal along  $c$ , then we can split the **Jacobi field** into

- tangential component: do not depend on geometry of  $\mathcal{M}$ , hence no information about  $\mathcal{M}$ ;
- normal component: very useful!

Specifically, consider  $X = X^\top + X^\perp$ , we have the following.

**Lemma 5.4.1.** Let  $c: [a, b] \rightarrow \mathcal{M}$  be a **geodesic**, and  $\lambda, \mu \in \mathbb{R}$ . Then, the **Jacobi field**  $X$  along  $c$  with  $X(a) = \lambda\dot{c}(a)$ ,  $\dot{X}(a) = \mu\dot{c}(a)$  is given by  $X(t) = (\lambda + (t - a)\mu)\dot{c}(t)$ .

## Lecture 18: Jacobi Fields and Geodesics

### 5.4.1 Jacobi Fields and the Linearization of Geodesic Equations

14 Mar. 13:00

**As previously seen.** Recall that **Equation 5.2** is linear, hence the sum of solutions is a solution.

**Theorem 5.4.1.** Consider a **geodesic**  $c: [0, 1] \rightarrow \mathcal{M}$ ,  $t \mapsto c(t)$ , and the **geodesic variation**  $c: [0, 1] \times (-\epsilon, \epsilon) \rightarrow \mathcal{M}$  of  $c$ .<sup>a</sup> Then  $X(t) := \frac{\partial}{\partial s} c(t, s)|_{s=0}$  is a **Jacobi field** along  $c(t) = c_0(t)$ . Conversely, every **Jacobi field** along  $c(t)$  can be obtained in this way, i.e., by **variation** of **geodesics**.

<sup>a</sup>I.e., for all **curves**  $c(\cdot, s) =: c_s(\cdot)$  are **geodesics**.

**Proof.** The forward direction is straightforward: since  $c(t, s)$  for a fixed  $s$  is a **geodesic**, hence

<sup>3</sup>This is often called the *field force operator*.

$\nabla_{\frac{\partial}{\partial t}} \frac{\partial}{\partial t} c(t, s) = 0$  for all  $s$ , implying  $\nabla_{\frac{\partial}{\partial s}} \nabla_{\frac{\partial}{\partial t}} \frac{\partial}{\partial t} c(t, s) = 0$ . Then,

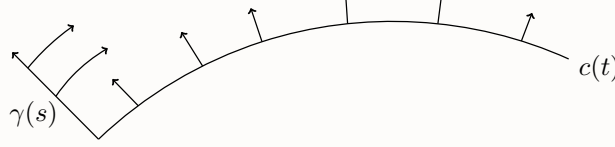
$$\begin{aligned} 0 &= \nabla_{\frac{\partial}{\partial s}} \nabla_{\frac{\partial}{\partial t}} \frac{\partial}{\partial t} c(t, s) \\ &= \nabla_{\frac{\partial}{\partial t}} \nabla_{\frac{\partial}{\partial s}} \frac{\partial}{\partial t} c(t, s) + \left( -\nabla_{\frac{\partial}{\partial t}} \nabla_{\frac{\partial}{\partial s}} + \nabla_{\frac{\partial}{\partial s}} \nabla_{\frac{\partial}{\partial t}} \right) \frac{\partial}{\partial t} c(t, s) \\ &= \nabla_{\frac{\partial}{\partial t}} \nabla_{\frac{\partial}{\partial t}} \frac{\partial}{\partial s} c(t, s) + R \left( \frac{\partial c}{\partial s}, \frac{\partial c}{\partial t} \right) \frac{\partial c}{\partial t}, \end{aligned}$$

since  $[\frac{\partial}{\partial s}, \frac{\partial}{\partial t}] = 0$ , and hence also  $\nabla_{\frac{\partial}{\partial s}} \frac{\partial}{\partial t} = \nabla_{\frac{\partial}{\partial t}} \frac{\partial}{\partial s}$ . Plugging in the definition of  $X$ , we have

$$\nabla_{\frac{\partial}{\partial t}} \nabla_{\frac{\partial}{\partial t}} X + R \left( X, \frac{\partial c}{\partial t} \right) \frac{\partial c}{\partial t} = 0,$$

i.e.,  $X$  is a **Jacobi field**.

The converse direction is left as a homework. As a hint, consider the following:



Then, let

$$c(t, s) = \exp_{\gamma(s)}(t(\dot{c}(0) + s \cdot V))$$

for some  $V$ . Once we have this, we just let  $X(t) = \frac{\partial}{\partial s} c(t, s)|_{s=0}$ . ■

**Remark.** This confirms the intuition: **Jacobi equation** is the linearization of the **geodesic equation**!

**Theorem 5.4.1** gives us a clear picture of how **Jacobi field** arise from the **geodesic variation**.

## 5.4.2 Killing Fields

To proceed, we need a new definition called **killing field**.

**As previously seen.** Recall that

$$\begin{aligned} (\mathcal{L}_X S)(Y_1, \dots, Y_p) &= X(S(Y_1, \dots, Y_p)) - \sum_{i=1}^p S(Y_1, \dots, [X, Y_i], \dots, Y_p) \\ &= (\nabla_X S)(Y_1, \dots, Y_p) + \sum_{i=1}^p S(Y_i, \dots, \nabla_{Y_i} X, \dots, Y_p) \end{aligned}$$

since  $\nabla$  is **torsion-free**, we have  $\nabla_X Y_i - \nabla_{Y_i} X = [X, Y_i]$ .

**Definition 5.4.1 (Killing field).** Consider a **Riemannian manifold**  $(\mathcal{M}, g)$ , and  $g = g_{ij} dx^i \otimes dx^j$ . Then a **vector field**  $X$  such that

$$\mathcal{L}_X g = 0$$

is called a **killing field** (or *infinitesimal isometry*).

Here are two basic facts about **killing fields**.

**Lemma 5.4.2.** A **vector field**  $X$  on  $(\mathcal{M}, g)$  is a **killing field** if and only if the **local 1-parameter group** generated by  $X$  consisted of **local isometries**.

**Lemma 5.4.3.** The **killing fields** of a **Riemannian manifold** constitute a **Lie algebra**.

**Theorem 5.4.1** implies the following.

**Corollary 5.4.1.** Every **killing field**  $X$  on  $\mathcal{M}$  is a **Jacobi field** along any **geodesic** in  $\mathcal{M}$ .

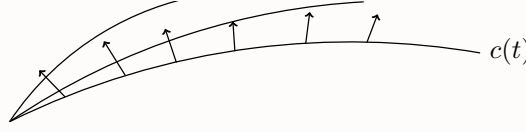
**Proof idea.** Since we have a **killing field**  $X$ , we use it to construct  $\Phi_s: \mathcal{M} \rightarrow \mathcal{M}$ , which is an **isometry** since  $X$  is a **killing field**.

The idea is to consider  $c(t, s) = \Phi_s \circ c(t)$ , and let  $X = \frac{\partial}{\partial s} c(t, s)$ . By **Theorem 5.4.1**, we're done. ■

**Corollary 5.4.2.** Let  $c: [0, T] \rightarrow \mathcal{M}$  be a **geodesic** with  $p = c(0)$ , i.e.,  $c(t) = \exp_p(t\dot{c}(0))$ . For  $W \in T_p\mathcal{M}$ , the **Jacobi field**  $x$  along  $c$  with  $X(0) = 0$ ,  $\dot{X}(0) = W$ , is given as

$$X(t) = D(\exp_p)|_{(t\dot{c}(0))} (tW).$$

**Proof.** This is a direct consequence of **Theorem 5.4.1**, since now  $X(0) = 0$ , we don't need to worry about constructing  $\gamma(s)$ , i.e., we have the following:



Now, we consider  $c(t, s) = \exp_p(t(\dot{c}(0) + s \cdot W))$ , hence

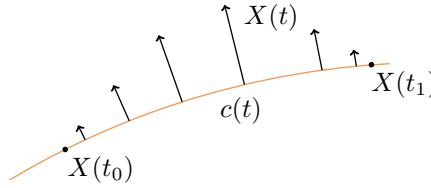
$$\frac{\partial}{\partial s} c(t, s) = \frac{\partial}{\partial s} \exp_p(t\dot{c} + s \cdot W) \Big|_{s=0}.$$

To have  $D(\exp_p)|_V (W)$ , we construct a **Jacobi field**  $W$  such that  $X(0) = 0$ ,  $\dot{X}(0) = W$ . ■

**Remark.** Thus, derivative of exp can be computed from **Jacobi field** along radial **geodesics**.

## 5.5 Conjugate Points

Consider the following, where  $c: I \rightarrow \mathcal{M}$  is a **geodesic**, and  $X(t)$  is a **Jacobi field** with  $X(t_0) = X(t_1) = 0$  such that  $t_0 \neq t_1 \in I$ .



**Note.** Notice that  $X$  is always normal to  $c$ .

To characterize this scenario, we define the following.

**Definition 5.5.1 (Conjugate point).** Let  $c: I \rightarrow \mathcal{M}$  be a **geodesic**. For  $t_0, t_1 \in I$  with  $t_0 \neq t_1$ ,  $c(t_0)$  and  $c(t_1)$  are called *conjugate* along  $c$  if there exists a **Jacobi field**  $X(t)$  along  $c$  which does not vanish identically but satisfies  $X(t_0) = 0 = X(t_1)$ .

**Note.** We see that  $\langle X(t), \dot{c}(t) \rangle = 0$  for all  $t$ .

**Proof.** Since  $\nabla_{\partial t} \langle X(t), \dot{c}(t) \rangle = \langle \dot{X}, \dot{c} \rangle$ , so

$$\nabla_{\partial t} \nabla_{\partial t} \langle X(t), \dot{c}(t) \rangle = \langle \ddot{X}, \dot{c} \rangle = -\langle R(X, \dot{c})\dot{c}, \dot{c} \rangle = 0.$$

This is a linear function, and if two endpoints are both 0, everything is 0.  $\circledast$

**Note.** If  $t_0, t_1 \in I$ ,  $t_0 \neq t_1$  are not **conjugate** along  $c$ , then for  $V \in T_{c(t_0)}\mathcal{M}$ ,  $W \in T_{c(t_1)}\mathcal{M}$ , there exists a unique **Jacobi field**  $Y(t)$  along  $c$  such that  $Y(t_0) = V$ ,  $Y(t_1) = W$ .

**Proof.** Let  $\mathcal{J}_c$  be the vector space of **Jacobi fields** along  $c$ . Construct the linear map

$$A: \mathcal{J}_c \rightarrow T_{c(t_0)}\mathcal{M} \times T_{c(t_1)}\mathcal{M}, \quad Y \mapsto (Y(t_0), Y(t_1)).$$

Since  $\mathcal{J}_c$  is a vector space with  $\dim \mathcal{J}_c = 2n$ , and the target space is also with dimension  $2n$ , and because  $t_0 \neq t_1$  are not **conjugate**,  $\ker A = \{0\}$ , i.e.,  $A$  is injective, hence  $A$  is bijective as the domain and the range of  $A$  have the same dimension.  $\circledast$

**Example.** Any antipodal points of  $S^n$  are **conjugate points**.

**Example.**  $\mathbb{R}^n$  with flat **metric** doesn't have **conjugate points**.

**Example.** **Riemannian manifolds** with non-positive **sectional curvature** has no **conjugate points**.

### 5.5.1 Length-Minimizing Geodesics

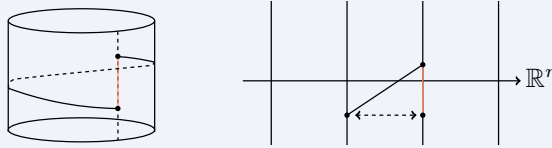
We can formalize the above examples by the following.

**Theorem 5.5.1.** Let  $c: [a, b] \rightarrow \mathcal{M}$  be a **geodesic**.

- (a) If there does not exist a point **conjugate** to  $c(a)$  along  $c(t)$ , then there exists  $\epsilon > 0$  such that for all piecewise **smooth curve**  $g: [a, b] \rightarrow \mathcal{M}$  with  $g(a) = c(a)$ ,  $g(b) = c(b)$  and  $d(g(t), c(t)) < \epsilon$  for all  $t \in [a, b]$ , we have  $L(c) \leq L(g)$ , and the equality holds when if and only if  $g$  is a reparametrization of  $c$ .
- (b) If there is  $\tau \in (a, b)$  such that  $c(a)$  and  $c(\tau)$  are **conjugate points** along  $c$ , then there exists a **proper variation**  $c(t, s): [a, b] \times (-\epsilon, \epsilon) \rightarrow \mathcal{M}$  such that  $L(c_s) < L(c)$  for  $s \in (-\epsilon, \epsilon) \setminus \{0\}$ .

**Theorem 5.5.1 (a)** implies that if there are no **conjugate points**, a **geodesic** is length-minimizing w.r.t. *sufficiently close curves*. As we have seen multiple times, this is not global.

**Example (Cylinder).** Consider the cylinder, where we identify every (integer-multiple) line of  $\mathbb{R}^n$  below.



There are two **geodesics**, but one is strictly longer.

**Example (Torus).** Consider **geodesics** on a “flat” torus which winds around more than once on the torus. Then, even without **conjugate points**, it's not length-minimizing globally.

To prove **Theorem 5.5.1**, we need the following.

**Corollary 5.5.1.** Let  $p \in \mathcal{M}$  and  $V \in T_p\mathcal{M}$  is contained in the domain of definition of  $\exp_p$ . Let  $c(t) = \exp_p(tV)$ , and  $\gamma: [0, 1] \rightarrow T_p\mathcal{M}$  be a piecewise **smooth curve** contained in the domain of  $\exp_p$ .

with  $\gamma(0) = 0, \gamma(1) = V$ . Then

$$\|v\| = L \left( \exp_p(tV) \Big|_{t \in [0,1]} \right) \leq L(\exp_p \circ \gamma(t))$$

and the equality holds if and only if  $\gamma$  differs from the curve  $tV$ ,  $t \in [0, 1]$  only by reparametrization.

**Proof hint.** We directly estimate

$$L(\exp \circ \gamma) = \int_0^1 \left| \frac{d}{dt} \exp \circ \gamma \right| dt = \int_0^1 |D \exp \circ \gamma| dt.$$

■

## Lecture 19: Length-Minimizing Geodesics and Conjugacy

Now, let's prove [Theorem 5.5.1](#).

14 Mar. 13:00

**Proof of Theorem 5.5.1.** We prove them one by one.

- (a) We want to show that if there's no [conjugate point](#), then for all [curves](#) as in (a), there exists a [curve](#)  $\gamma$  as in [Corollary 5.5.1](#). Without loss of generality, let  $a = 0$ ,  $b = 1$ , and we set  $V := \dot{c}(0)$ . Then, we know that since there are no [conjugate points](#) along  $c$ ,  $\exp_p$  of maximal rank along any radial [curve](#)  $tV$ ,  $0 \leq t \leq 1$ . By the inverse function theorem, for all  $t$ ,  $\exp_p$  is a [diffeomorphism](#) in a neighborhood of  $tV$ .

Now, cover  $\{tV \mid 0 \leq t \leq 1\}$  by finitely many such neighborhoods  $\{\Omega_i\}_{i=1}^k$ , and let  $U_i = \exp_p \Omega_i$ . Assume that  $tV \in \Omega_i$ , for  $t_{i-1} \leq t \leq t_i$  (with  $t_0 = 0, t_k = 1$ ). Let  $\epsilon > 0$  sufficiently small. Then, for all [curve](#)  $g: [0, 1] \rightarrow \mathcal{M}$  satisfying the assumption,  $g([t_{i-1}, t_i]) \subseteq U_i$ .

**Claim.** For all  $g$  satisfying  $g([t_{i-1}, t_i]) \subseteq U_i$ , there exists a [curve](#)  $\gamma \subseteq T_p \mathcal{M}$  such that  $\exp_p \gamma = g$  with  $\gamma(0) = 0, \gamma(1)gV$ .

**Proof.** Put  $\gamma(t) = (\exp_p|_{\Omega_i})^{-1}(g(t))$  for  $t_{i-1} \leq t \leq t_i$ , so  $\gamma$  satisfies [Corollary 5.5.1](#).  $\otimes$

- (b) Without loss of generality, let  $a = 0, b = 1$ . Let  $X$  be a non-trivial [Jacobi field](#) along  $c$  with  $X(0) = 0 = X(\tau)$ . We have  $\dot{X}(\tau) \neq 0$ , as otherwise  $X \equiv 0$  by the uniqueness. Let  $Z(t)$  be an arbitrary [vector field](#)  $X$  along  $c$  with  $Z(0) = 0 = Z(1)$ ,  $Z(\tau) = -\dot{X}(\tau)$ . Let  $\eta > 0$ , set

$$Y_\eta(t) = \begin{cases} Y_\eta^1(t) = X(t) + \eta Z(t), & \text{if } 0 \leq t \leq \tau; \\ Y_\eta^2(t) = \eta Z(t), & \text{if } \tau \leq t \leq 1, \end{cases}$$

and we let  $Z^1 := Z|_{[0, \tau]}, Z^2 := Z|_{[\tau, 1]}$ . Now, since

$$I(Y_\eta^1, Y_\eta^1) = \langle \dot{X}(\tau), 2\eta Z(\tau) \rangle + \eta^2 I(Z^1, Z^1) - 2\eta \|\dot{X}(\tau)\|^2 + \eta^2 I(Z^1, Z^1),$$

with

$$I(Y_\eta^2, Y_\eta^2) = \eta^2 I(Z^2, Z^2),$$

with

$$I(Y_\eta, Y_\eta) = I(Y_\eta^1, Y_\eta^1) + I(Y_\eta^2, Y_\eta^2) = -2\eta \|\dot{X}(\tau)\|^2 + \eta^2 I(Z, Z)$$

for sufficiently small  $\eta > 0$ . Now, consider the variation  $c(t, s) := \exp_{c(t)} s Y_\eta(t)$ , we have  $L'(0) = 0^a$  and

$$L''(0) = I(Y_\eta, Y_\eta) < 0.$$

By the Taylor theorem, this is a minimum, i.e.,  $L(c_s) < L(c)$ .

■

<sup>a</sup>Note that  $L(s) := L(c_s)$ ,  $L(0) = L(c)$ .

**Remark.** Given a [geodesic](#)  $\gamma$  from  $q$  to  $p$ ,  $q$  is [conjugate](#) to  $p$  along  $\gamma$  if there exists a non-trivial [Jacobi field](#) along  $\gamma$  vanishing at  $p$  and  $q$ .

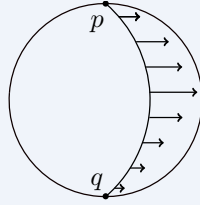
We finally note that there's one concept that's related to our discussion.

**Definition 5.5.2 (Order).** The *order* (or multiplicity) of [conjugacy](#) is the dimension of the space of [Jacobi fields](#) vanishing at two [conjugate points](#).

Given  $\dim \mathcal{M} = n$ , by the existence and uniqueness theorem for [Jacobi fields](#), we see that

- there is an  $n$ -dimensional space of [Jacobi fields](#) vanishing at  $p \in \mathcal{M}$ ;
- there is an at most  $(n - 1)$ -dimensional space of [Jacobi fields](#) vanishing at  $p, q \in \mathcal{M}$ , as tangential [Jacobi fields](#) vanishes at most at one point.

**Example ( $S_r^n$ ).** On  $S_r^n$  and  $p, q$  antipodal points on  $S_r^n$ , there is a [Jacobi field](#) vanishing at  $p$  and  $q$  for all [parallel normal vector field](#) along  $\gamma$ , thus  $p, q$  [conjugate](#) to *order* exactly  $(n - 1)$ .



### 5.5.2 Characterization via Exponential Maps and Index Forms

We now characterize the [conjugate points](#) by [exponential map](#) and the [index form](#): firstly, they are precisely the images of singularities of the [exponential map](#).

**Proposition 5.5.1.** Let  $p \in \mathcal{M}$ ,  $V \in T_p \mathcal{M}$ ,  $q = \exp V$ . Then  $\exp_p$  is a local [diffeomorphism](#) in a neighborhood of  $V$  if and only if  $q$  does not [conjugate](#) to  $p$  along [geodesic](#)  $\gamma(t) = \exp_p tV$ ,  $t \in [0, 1]$ .

For simplicity, let's develop some shorthand notations.

**Notation.** Let  $c: [a, b] \rightarrow \mathcal{M}$  be a [curve](#). Denote

- $\nu_c$ : the space of [vector field](#)  $X$  along  $c$ , i.e.,  $\nu_c = \Gamma(c^*(T\mathcal{M}))$ ;
- $\dot{\nu}_c$ : the space of [vector field](#)  $X$  along  $c$  with  $V(a) = V(b) = 0$ .

Another characterization is related to [index form](#).

**Lemma 5.5.1.** Let  $c: [a, b] \rightarrow \mathcal{M}$  be a [geodesic](#). Then there is no pair of [conjugate points](#) along  $c$  if and only if the [index form](#)  $I$  of  $c$  is strictly positive definite on  $\dot{\nu}_c$ .

**Proof.** Assume that  $c$  has no [conjugate points](#), then [Theorem 5.5.1 \(a\)](#) implies that  $I(X, X) \geq 0$  for all  $X \in \dot{\nu}_c$  because otherwise  $c(t, s) := \exp_{c(t)} sX(t)$  would be locally length-decreasing.

If  $I(Y, Y) = 0$  for some  $Y \in \dot{\nu}_c$ , then by  $I(X, X) \geq 0$ , for all  $Z \in \dot{\nu}_c$  and  $\lambda \in \mathbb{R}$ ,

$$0 \leq I(Y - \lambda Z, Y - \lambda Z) = 0 - 2\lambda I(Y, Z) + \lambda^2 I(Z, Z).$$

This inequality holds only if  $I(Y, Z) = 0$  for all  $Z \in \dot{\nu}_c$ , implying  $Y$  is a [Jacobi field](#) from [Proposition 5.3.1](#). As there are no [conjugate points](#) along  $c$ ,  $Y = 0$ , i.e.,  $I$  is strictly positive definite.

## Lecture 20: Sobolev Spaces and Cut Locus

**Proof of Lemma 5.5.1 (Continue).** For the backward direction, assume that for  $t_0, t_1 \in [a, b]$  (without loss of generality, let  $t_0 < t_1$ ) such that  $c(t_0), c(t_1)$  are two **conjugate points** along  $c$ . Then, there exists a non-trivial **Jacobi field**  $X$  along  $c$  such that  $X(t_0) = 0 = X(t_1)$ . Now, consider

$$Y(t) = \begin{cases} 0, & \text{if } a \leq t \leq t_0; \\ X(t), & \text{if } t_0 \leq t \leq t_1; \\ 0, & \text{if } t_1 \leq t \leq b; \end{cases} \Rightarrow J(Y, Y) = 0,$$

hence  $I$  is not positive definite, a contradiction.  $\blacksquare$

## 5.6 The Cut Locus

We end this chapter by developing one more notion for later. We first take a detour to **Sobolev spaces**.

### 5.6.1 Sobolev Spaces

On  $\dot{\nu}_c$ , we introduce the norm

$$\|X\| := \left( \int_a^b (\langle \dot{X}, \dot{X} \rangle + \langle X, X \rangle) dt \right)^{1/2},$$

and denote  $\dot{H}_c$  the completion of  $\dot{\nu}_c$  w.r.t.  $\|\cdot\|$ .

**Definition 5.6.1 (Schwartz space).** A *Schwartz space*  $\mathcal{S}(\mathbb{R}^d)$  is defined as

$$\mathcal{S}(\mathbb{R}^d) := \left\{ u \in C^\infty(\mathbb{R}^d) \mid \forall \alpha, \beta \in \mathbb{N}^d \sup_{x \in \mathbb{R}^d} |x^\alpha \partial^\beta u(x)| < \infty \right\}.$$

**Definition 5.6.2 (Tempered distribution).** A *tempered distribution* is a continuous linear functional  $f$  on  $\mathcal{S}(\mathbb{R}^d)$ , i.e.,  $f: \mathcal{S}(\mathbb{R}^d) \rightarrow \mathbb{C}$ .

**Notation.** The space of **tempered distributions** is denoted as  $\mathcal{S}^1(\mathbb{R}^d)$ .

**Definition 5.6.3 (Locally integrable).** Let  $\Omega \subseteq \mathbb{R}^n$  be open, and let  $f: \Omega \rightarrow \mathbb{C}$  be Lebesgue measurable. Then the *locally integrable* (or locally summable) space is defined as

$$L_{\text{loc}}^1(\Omega) := \{ f: \Omega \rightarrow \mathbb{C} \text{ measurable} \mid f|_K \in L^1(K) \forall \text{ compact } K \subseteq \Omega \}.$$

**Definition 5.6.4 (Weak derivative).** Let  $U \subseteq \mathbb{R}^n$  be open, and  $u, v \in L_{\text{loc}}^1(U)$ . Let  $\alpha$  be a multi-index. Then  $v$  is the  $\alpha^{\text{th}}$ -weak derivative of  $u$ , denoted as  $D^\alpha u = v$  provided

$$\int_U u \cdot D^\alpha \varphi dx = (-1)^{|\alpha|} \int_U v \varphi dx$$

for all test functions  $\varphi \in C_c^\infty(U)$ .

**Notation.**  $C_c^\infty(U)$  is the space of smooth functions with compact support defined on  $U$ .

**Notation.** Here,  $D^\alpha \varphi$  means

$$D^\alpha \varphi = \frac{\partial^{|\alpha|} \varphi}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

**Note.** We can write  $D^\alpha u$  since it's unique (up to measure 0).

**Remark.** If the [weak derivative](#) exists, then it's unique up to a set of measure zero.

**Definition 5.6.5 (Sobolev space).** Fix  $1 \leq p \leq \infty$ , and let  $k$  be a non-negative integer. The *Sobolev space*  $W^{k,p}(U)$  consists of all [locally integrable](#) functions  $u: U \rightarrow \mathbb{R}$  for all  $\alpha$  with  $|\alpha| \leq k$  such that  $D^\alpha u$  exists in the [weak](#) sense and belongs to  $L^p(U)$ .

**Remark.** If  $p = 2$ ,  $H^k(U) := W^{k,2}(U)$  for  $k = 0, 1, \dots$  is a Hilbert space.

**Example.**  $H^0(U) = L^2(U)$ .

On  $W^{k,p}(U)$ , we introduce the norm

$$\|u\|_{W^{k,p}(U)} = \begin{cases} \left( \sum_{|\alpha| \leq k} \int_U |D^\alpha u|^p dx \right)^{1/p}, & \text{if } 1 \leq p < \infty; \\ \sum_{|\alpha| \leq k} \operatorname{ess\,sup}_U |D^\alpha u|, & \text{if } p = \infty. \end{cases}$$

**Notation.** Denote the closure of  $C_c^\infty(U)$  in  $W^{k,p}(U)$  as  $W_0^{k,p}(U)$ .

Thus,  $u \in W_0^{k,p}(U)$  if and only if there exists functions  $u_n \in C_c^\infty(U)$  such that  $u_n \rightarrow u$  in  $W^{k,p}(U)$ .

**Remark.** Lastly, the upshot is that  $u \in W_0^{k,p}(U)$  if  $u \in W^{k,p}(U)$  such that “ $D^\alpha u = 0$  on  $\partial U$ ” for all  $|\alpha| \leq k - 1$ , more precisely, use traces.

## 5.6.2 The Index

Let  $\{\dot{V}_i\}_{i=1}^d$  for  $d = \dim \mathcal{M}$  be an orthonormal basis of [parallel vector fields](#). Now, write  $X = \xi^i \dot{V}_i$ , so  $\dot{X}_i = \dot{\xi}^i \dot{V}_i$ , hence

$$\|X\| = \left( \int_a^b \left( \dot{\xi}^i \dot{\xi}^j + \xi^i \xi^j \right) dt \right)^{1/2}.$$

Then,  $\dot{H}_c^1$  can be identified with [Sobolev space](#)  $\dot{H}^{1,2}(I, \mathbb{R}^d)$ . Next, consider  $I$  (the [index form](#)) of  $c$  as quadratic form on  $\dot{H}_c^1$ , i.e.,  $I: \dot{H}_c^1 \times \dot{H}_c^1 \rightarrow \mathbb{R}$  with

$$I(X, Y) = \int_a^b \left( \langle \dot{X}, \dot{Y} \rangle - \langle R(\dot{c}, X)Y, \dot{c} \rangle \right) dt,$$

and we define the following.

**Definition 5.6.6 (Index).** The *index* of  $c$ ,  $\operatorname{Ind}(c)$ , is the dimension of the largest subspace of  $\dot{H}_c^1$ , on which  $I$  is negative definite.

**Definition 5.6.7 (Extended index).** The *extended index* of  $c$ ,  $\operatorname{Ind}_0(c)$ , is the dimension of the largest subspace of  $\dot{H}_c^1$ , on which  $I$  is negative semi-definite.

**Definition 5.6.8 (Nullity).** The *nullity* is defined as  $N(c) := \operatorname{Ind}_0 - \operatorname{Ind}(c)$ .



**Notation.** For  $t \in (a, b]$ , let  $\mathcal{J}_c^t$  be the space of **Jacobi fields**  $x$  along  $c$  with  $X(a) = 0 = X(t)$ .

**Lemma 5.6.1.**  $\text{Ind}(c)$  and  $N(c)$  are always finite.

**Proof.** See [FC13] (by contradiction). ■

**Lemma 5.6.2.**  $\dim \mathcal{J}_c^b = N(c)$ .

### 5.6.3 The Cut Locus

Let  $(\mathcal{M}^n, g)$  be a **complete Riemannian manifold**. Let  $p \in \mathcal{M}$ , and denote  $d(\cdot) := d(p, \cdot)$ . Then we have seen that there is a normal neighborhood where **geodesics** are minimizing and  $d$  is smooth away from  $p$ . For all  $v \in S^{n-1}$ , we can find a **geodesic**  $c_v(t) = \exp_t(tv)$ . Let  $R(v) := \sup \left\{ T \mid c_v|_{[0, T]} \text{ minimizing} \right\}$ .

**Note.** If  $t < R(v)$ , then  $d(p, c_v(t)) = t$ ; moreover, if  $R(v) = \infty$ ,  $c_v$  is minimizing.

We can then define the following.

**Definition 5.6.9 (Cut locus).** The *cut locus* of  $p$  is defined as

$$C(p) := \{c_v(R(v)) \mid v \in S^{n-1} \text{ such that } R(v) < \infty\}.$$

**Definition 5.6.10 (Cut point).** Consider a **geodesic**  $c$  with  $d(c(0), c(t))d(p, c(t)) = t$  on  $t \in [0, t_0]$  for  $t_0$  being the last point this holds. Then we say  $c(t_0)$  is a *cut point* of  $p$  along  $c$ .

**Intuition.** The **cut locus**  $C(p)$  of  $p$  is the union of the **cut points** of  $p$  along all **geodesics** starting from  $p$ .

## Lecture 21: Morse Index Theorem

**As previously seen.** Fix  $p \in \mathcal{M}$ , let  $q \in C(p)$  a **cut point**. Then there exists a **geodesic**  $c$  such that

- (a)  $c$  minimizing up to and including  $q$ , and
- (b)  $c$  uniquely minimizing up to but not including  $q$ .

Thus,  $c$  is not minimizing after that point.

Let's first see the following under the above setting.

**Proposition 5.6.1.** At each **cut point**  $q \in C(p)$ ,  $q$  is either a **conjugate point** or there exists two minimizing **geodesics** connecting  $p, q$ .

**Proof.** See [FC13]. ■

23 Mar. 13:00

## Chapter 6

# Morse Index, Rauch Comparison, Sphere Theorems, and More

Now, we have everything to prove three important theorems: the [Morse index theorem](#), [Rauch comparison theorem](#), and the [sphere theorem](#).

**Intuition.** In short,

- [Morse index theorem](#) relates the number (with multiplicities) of [conjugate points](#) on a [geodesic](#) segment to the [index](#).
- [Rauch comparison theorem](#) is one of the basic facts in Riemannian geometry. Intuitively, it expresses the plausible fact that as the [curvature](#) grows, lengths shorten.
- [Sphere theorem](#) is one of the most beautiful theorems of global differential geometry, which says that under some mild [curvature bounds](#), the space is homeomorphic to a sphere.

In what follows, we prove each theorem one by one.

**Note.** After proving the [Morse index theorem](#), we detour to study the [Morse function](#) and [Morse homology](#) before going to the [Rauch comparison theorem](#).

**Note.** After prove the [sphere theorem](#), we prove another important [uniformization theorem](#), which is worth noting here.

Let's start by proving the [Morse index theorem](#).

### 6.1 Morse Index Theorem

In this section, we study the [Morse index theorem](#), which gives information about [conjugate points](#) via [index form](#).

**As previously seen.** The [index form](#)  $I(X, Y)$ , [index](#)  $\text{Ind}(c)$ , and also  $\text{Ind}_0$ , and [Lemma 5.6.1](#).

Let  $c: [0, T] \rightarrow \mathcal{M}$  be a [geodesic](#). Then the [index](#)  $\text{Ind}(c)$  on the space  $\mathcal{V}_c$  is finite and equals the number of points  $c(t)$  [conjugate](#) to  $c(0)$  for  $t \in (0, T)$ , counted with multiplicities.

#### 6.1.1 The Conjugate Locus

Before proving the [Morse index theorem](#), let's see one last definition.

**Definition 6.1.1** (Conjugate locus). Let  $(\mathcal{M}, g)$  be a [Riemannian manifold](#). The set of (first) [conjugate points](#) of point  $p \in \mathcal{M}$  for all [geodesics](#) starting at  $p$  is called the *conjugate locus* of  $p$ .

**Proposition 6.1.1.** Let  $(\mathcal{M}, g)$  be a complete Riemannian manifold. Let  $c: [0, \infty) \rightarrow \mathcal{M}$  be a normalized geodesic with  $c(0) = p$ . Assume that  $c(t_0)$  is the cut point at  $p = c(0)$  along  $c$ . Then, either  $c(t_0)$  is the first conjugate point of  $c(0)$  along  $c$  or there exists another geodesic  $\sigma \neq c$  from  $p$  to  $c(t_0)$  such that  $\ell(\sigma) = \ell(c)$ . Conversely, if either the above are true, then there exists  $t_0 \in (0, t_0]$  such that  $c(t_1)$  is the cut point of  $p$  along  $c$ .

**Proof.** See [FC13]. ■

## 6.1.2 Morse Index Theorem

Consider the following.

**Theorem 6.1.1 (Morse index theorem).** Let  $c: [a, b] \rightarrow \mathcal{M}$  be a geodesic. Then, there are at most finitely many points conjugate to  $c(a)$  along  $c$ , and

$$\text{Ind}(c) = \sum_{t \in (a, b)} \dim \mathcal{J}_c^t, \quad \text{Ind}_0(c) = \sum_{t \in (a, b]} \dim \mathcal{J}_c^t.$$

**Proof.** For all  $t_i \in (a, b]$ , for which  $c(t_i)$  conjugate to  $c(a)$ , there exists a Jacobi field  $X_i$  along with  $X_i(a) = 0 = X_i(t_i)$ . Set

$$Y_i(t) := \begin{cases} X_i(t), & \text{if } a \leq t \leq t_i; \\ 0, & \text{otherwise,} \end{cases}$$

we have that  $Y_i(t)$  are linearly independent such that  $I(Y_i, Y_i) = 0$  for all  $i$ . This implies that the number of conjugate points is at most  $\text{Ind}_0(c)$ , which is finite from Lemma 5.6.1.

For  $\tau \in (a, b]$ , set

$$\varphi(\tau) := \text{Ind} \left( c|_{[a, \tau]} \right), \quad \varphi_0(\tau) := \text{Ind}_0 \left( c|_{[a, \tau]} \right).$$

**Claim.**  $\varphi(\tau)$  is left-continuous.

**Proof.** For  $\tau \in (a, b]$ , let  $I_\tau$  be the index form of  $c|_{[a, \tau]}$ , and let  $X$  be a vector field along  $c|_{[a, \tau]}$  satisfy  $I_\tau(X, X) < 0$  and  $\|X\| = 1$ .

Let  $\tilde{X}$  be vector field defined by  $\tilde{X}(t) := X(\tau t / \sigma)$  on  $[a, \sigma]$ . Then,

$$\int_0^\sigma \langle \dot{\tilde{X}}(t), \dot{\tilde{X}}(t) \rangle dt = \int_0^\sigma \left( \frac{\tau}{\sigma} \right)^2 \langle \dot{X}(\tau t / \sigma), \dot{X}(\tau t / \sigma) \rangle dt = \frac{\tau}{\sigma} \int_0^\tau \langle \dot{X}(s), \dot{X}(s) \rangle ds,$$

implying

$$\int_0^\sigma \langle \dot{\tilde{X}}(t), \dot{\tilde{X}}(t) \rangle dt \rightarrow \int_0^\tau \langle \dot{X}(t), \dot{X}(t) \rangle dt$$

for  $\sigma \rightarrow \tau$ . Also, we have  $\|X\| = 1$ , and  $X$  is continuous,<sup>a</sup> we see that  $\tilde{X}$  converges point-wise to  $X$  as  $\sigma \rightarrow \tau$ , hence

$$\int_0^\sigma \langle R(\dot{c}, \tilde{X})\tilde{X}, \dot{c} \rangle dt \rightarrow \int_0^\tau \langle R(\dot{c}, X)X, \dot{c} \rangle dt$$

as  $\sigma \rightarrow \tau$ , hence  $I_\sigma(\tilde{X}, \tilde{X}) \rightarrow I_\tau(X, X)$  as  $\sigma \rightarrow \tau$ . Notice that the above also implies  $I_\sigma(\tilde{X}, \tilde{X}) < 0$  if  $\sigma$  is sufficiently close to  $\tau$ .

Finally, for all orthonormal basis of a space on which  $I_\tau$  is negative definite, we may also find a basis of some space on which  $I_\sigma$  is negative definite if  $\sigma$  is sufficiently close to  $\tau$ . As  $\varphi$  is monotonically increasing, we have left-continuity. ⊗

<sup>a</sup>This is from something called Sobolev theorem.

**Claim.**  $\varphi_0(\tau)$  is right-continuous.

**Proof.** Let  $(\tau_n)_{n \in \mathbb{N}} \subseteq (a, b]$  converge to  $\tau \in (a, b]$  for all  $n \in \mathbb{N}$ , let  $X_n$  be a **vector field along**  $c|_{[0, \tau_n]}$  with  $\|X\| = 1$  and  $I_{\tau_n}(X_n, X_n) \leq 0$ . After selecting a subsequence,  $X_n$  converges weakly in Sobolev space  $H^{1,2}$  topology to some **vector field  $X$  along**  $c|_{[a, \tau]}$ . Then, we just check every ingredient of **index form** (see [FC13]).  $\otimes$

Finally, let  $a < t_1 < t_2 < \dots < t_k \leq b$  be the points  $c(t_i)$  **conjugate** to  $c(a)$ . Then,  $\varphi_0(t) - \varphi(t) = 0$  for  $t_i \in (a, b]$ . Then,

$$\sum_{t \in (a, b]} \dim \mathcal{J}_c^t = \sum_{t \in (a, b]} (\varphi_0(t) - \varphi(t)) = \sum_{i=1}^k (\varphi_0(t_i) - \varphi(t_i)).$$

Since  $\varphi$  is left-continuous, and  $\varphi_0$  is right-continuous, hence we have

$$\varphi_0(t_i) = \varphi(t_{i+1})$$

for  $i = 1, \dots, k-1$ , we finally have

$$\sum_{i=1}^k (\varphi_0(t_i) - \varphi(t_i)) = \varphi_0(t_k) - \varphi(t_1).$$

From  $\varphi$  being left-continuous again,  $\varphi(t_1) = 0$ . Finally, again, from the continuity properties of  $\varphi, \varphi_0$ , they can “jump” only at the points  $\tau$  where  $\varphi_0(\tau) \neq \varphi(\tau)$ , i.e., at the **conjugate points**. In particular,  $\varphi_0$  is constant on  $[t_k, b]$  hence,  $\varphi_0(t_k) = \varphi_0(b)$ , i.e.,

$$\varphi_0(b) = \sum_{t \in (a, b]} \dim \mathcal{J}_c^t.$$

■

**Intuition.** The “jump” only happens at **conjugate points**.

## Lecture 22: Bonnet-Mayers Theorem and Morse Functions

### 6.1.3 Bonnet-Mayers Theorem

28 Mar. 13:00

**Definition 6.1.2 (Diameter).** The *diameter* of a **manifold**  $\mathcal{M}$  is defined as

$$\text{diam}(\mathcal{M}) := \sup_{p, q \in \mathcal{M}} d(p, q).$$

**Theorem 6.1.2 (Bonnet-Mayers theorem).** Let  $(\mathcal{M}^n, g)$  be a **complete<sup>a</sup> Riemannian manifold** and **Ricci curvature**  $\geq \lambda > 0$ , i.e.,

$$\text{Ric}(X, X) \geq \lambda \langle X, X \rangle$$

for all  $X \in T\mathcal{M}$ . Then the **diameter** of  $\mathcal{M}$  is less than  $\pi \sqrt{(n-1)/\lambda}$ . In particular,  $\mathcal{M}$  is compact and has finite fundamental group  $\pi_1(\mathcal{M})$ .

<sup>a</sup>I.e., closed and any two points can be joined by a minimizing **geodesic**.

**Proof.** For all  $\rho < \text{diam}(\mathcal{M})$ , there exists  $p, q \in \mathcal{M}$  with  $d(p, q) = \rho$ . As  $\mathcal{M}$  **complete**, there exists a shortest **geodesic** arc  $c: [0, \rho] \rightarrow \mathcal{M}$  with  $c(0) = p$  and  $c(\rho) = q$ . Now, let  $\{e_i\}_{i=1}^n$  be an orthonormal basis of  $T_p\mathcal{M}$ , such that  $e_1 = \dot{c}(0)$ . Now, consider a **parallel** orthonormal basis  $\{\dot{c}(t), X_1(t), \dots, X_n(t)\}$  along  $c$ . Furthermore, consider  $Y_i(t) := (\sin \pi t / \rho) X_i(t)$  with  $i = 2, \dots, n$ .

Then,

$$I(Y_i, Y_i) = \int_0^\rho -\langle \ddot{Y}_i, Y_i \rangle - \langle R(Y_i, \dot{c})\dot{c}, Y_i \rangle dt = \int_0^\rho \sin^2 \frac{\pi t}{\rho} \left( \frac{\pi^2}{\rho^2} - \langle R(X_i, \dot{c})\dot{c}, X_i \rangle \right) dt.$$

Since  $c$  is the shortest [curve](#) connecting  $p, q$ , it follows that there are no [conjugate points](#) between  $p, q$ . Hence,  $I(Y_i, Y_i) \geq 0$  for all  $i$ , so

$$0 \leq \sum_{i=2}^n I(Y_i, Y_i) = \int_0^\rho \sin^2 \frac{\pi t}{\rho} \left( \frac{\pi^2}{\rho^2} (n-1) - R(\dot{c}, \dot{c}) \right) dt \leq \left( \frac{\pi^2}{\rho^2} (n-1) - \lambda \right) \int_0^\rho \sin^2 \frac{\pi t}{\rho} dx,$$

implying

$$0 \leq \frac{1}{2}\rho \left( \frac{\pi^2(n-1)}{\rho^2} - \lambda \right) \Rightarrow \rho^2 \leq \frac{\pi^2(n-1)}{\lambda} \Rightarrow \rho \leq \pi \sqrt{\frac{n-1}{\lambda}}.$$

Since this is true for all  $\rho < \text{diam}(\mathcal{M})$ , hence we see that  $\text{diam}(\mathcal{M}) \leq \pi \sqrt{(n-1)/\lambda}$ .

Furthermore, the universal cover of  $\mathcal{M}$  satisfies the same assumptions as [Ricci curvature](#), by computation, we have finite  $\pi_1(\mathcal{M})$ . ■

**Remark.** We choose  $Y_i(t) = \sin(\pi t/\rho)X_i(t)$  is just because it satisfies the needed condition, and makes the computation works out nicely.

**Intuition.** [Bonnet-Mayers theorem](#) says that if  $\mathcal{M}$  has [Ricci curvature](#) not less than the one of  $S_r^n$ , then  $\text{diam}(\mathcal{M})$  is at most the one of  $S_r^n$ .

Consider the hyperbolic space  $\mathbb{H}^n$  in  $\mathbb{R}^{n+1}$  where we define

$$\langle x, x \rangle := -(x^0)^2 + (x^1)^2 + \cdots + (x^n)^2$$

for  $x = (x^0, \dots, x^1)$ . Then

$$\mathbb{H}^n := \{x \in \mathbb{R}^{n+1} \mid \langle x, x \rangle = -1, x^0 > 0\}.$$

Also, consider the half-space of  $\mathbb{R}^n$  such that

$$H^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_n > 0\}$$

with metric on  $\mathbb{H}^n$ , we have

$$g_{ij}(x_1, \dots, x_n) = \frac{\delta_{ij}}{x_n^2}.$$

Then, we see that we have a constant sectional curvature of  $-1$ .

## 6.2 Morse Theory and Flow Homology

We detour to study [Morse functions](#), and some related topics. In what follows, we focus on critical points of functions.

### 6.2.1 Morse Functions

Let  $(\mathcal{M}, g)$  be a [complete Riemannian manifold](#). Let  $f: \mathcal{M} \rightarrow \mathbb{R}$  be a smooth<sup>1</sup> function. Then

$$df(x) = 0$$

means that  $x$  is a critical point of  $f$ .

**Definition 6.2.1 (Non-degenerate).** A critical point  $a$  of  $f$  is *non-degenerate* if the Hessian of  $f$  is non-singular at  $a$ .

<sup>1</sup>It's typically enough to ask for  $f \in C^3(\mathcal{M}, \mathbb{R})$ .

**Definition 6.2.2 (Morse index).** The *index*  $\mu(p)$  of **non-degenerate** critical point  $p$  of  $f$  is the dimension of the largest subspace of  $T_p\mathcal{M}$  on which the Hessian is negative definite.

**Intuition.** That is, the number of directions in which  $f$  decreases.

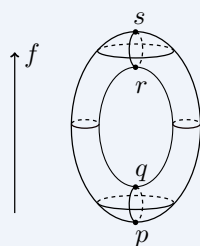
**Note.** The **degeneracy** and **index** are independent of coordinate choice.

Now, we define the critical set of  $f$  as

$$C(f) := \{x \in \mathcal{M} \mid df(x) = 0\}.$$

**Definition 6.2.3 (Morse function).** A *Morse function*  $f$  is a function as introduced such that all critical points are **non-degenerate**.

**Example.** Consider  $f$  is the height function, which is a **Morse function** such that the **index** of  $s$  is 2,  $r$  is 1,  $q$  is 1, and  $p$  is 0 by looking at the decreasing directions.



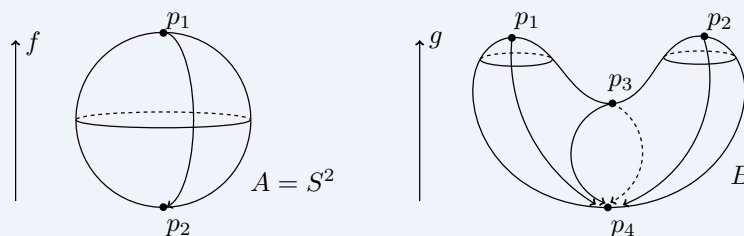
Now, define  $M^a = f^{-1}(-\infty, a]$ , then we see that

- (a) Pass  $p$ :  $M^a$  for  $0 < a < f(q)$  is a disk, which is homotopy equivalent to a point, i.e., 0-cell.
- (b) Pass  $q$ :  $M^a$  for  $f(q) < a < f(r)$  is a cylinder, where we attach a 1-cell.
- (c) Pass  $r$ :  $M^a$  for  $f(r) < a < f(s)$  is a torus with disk removed.
- (d) Pass  $s$ :  $M^a$  for  $a > f(s)$  is a torus.

## Lecture 23: Morse Theory and Flow Homology

Throughout this lecture, let  $\mathcal{M}$  be a compact **Riemannian manifold**, also, we'll keep mentioning the following example. 30 Mar. 13:00

**Example.** Consider the following.



Then,

- $f$ :  $\mu_f(p_1) = 2$ ,  $\mu_f(p_2) = 0$ .
- $g$ :  $\mu_g(p_1) = \mu_g(p_2) = 2$ ,  $\mu_g(p_3) = 1$ , and  $\mu_g(p_4) = 0$ .

Additionally, we can study some “invariants” about spaces, such as the Euler characteristic<sup>2</sup> of  $\mathcal{M}$ , which is “defined” as

$$\chi(\mathcal{M}) = \sum_{p: \text{critical point of } f} (-1)^{\mu(p)} \mu(p).$$

**Example.**  $\chi(A) = \chi(S^2) = 2$ , and  $\chi(B) = 4 - 1 = 3$ .

To make  $\chi$  formal, we need to consider the homology as we did in algebraic topology, i.e., we need to define the “boundary map” and “complexes”.

**Intuition.** Intuitively, our complexes should be a vector space built on top of critical points; on the other hand, for two critical points  $p, q$  such that  $\mu(p) - \mu(q) = 1$ , we want to count the trajectories from  $p$  to  $q$  modulo 2, i.e., the boundary map  $\partial$  should be somehow defined as

$$\partial p := \sum_{\substack{p \text{ critical point of } f \\ \mu(q) = \mu(p) - 1}} (\#\{\text{flow lines from } p \text{ to } q\} \bmod 2) \cdot q.$$

### 6.2.2 Morse Complex

We study so-called **Morse complex**, and define the so-called *flow homology*. We start by defining the so-called **negative gradient flow**.

**Definition 6.2.4 (Negative gradient flow).** The *negative gradient flow* of  $f$  on  $\mathcal{M}$  is defined as the solution  $\phi: \mathcal{M} \times \mathbb{R} \rightarrow \mathcal{M}$  of

$$\begin{cases} \frac{\partial}{\partial t} \phi(x, t) = -\text{grad}(f(\phi(x, t))); \\ \phi(x, 0) = x \end{cases} \quad \text{for } x \in \mathcal{M}.$$

**Note.** We will simply call **negative gradient flow** a *flow* for simplicity.

**Remark.** From the **Picard-Lindelöf theorem**, local existence of the **flow** is guaranteed. Moreover, if we have “very good” conditions, such a **flow** may even exist globally.

More generally, the Euler characteristic is an example of a **flow**  $\phi: \mathcal{M} \times \mathbb{R} \rightarrow \mathcal{M}$  such that

$$\begin{cases} \frac{\partial}{\partial t} \phi(x, t) = -V(f(\phi(x, t))); \\ \phi(x, 0) = x \end{cases}$$

with some **vector field**  $V$  on  $\mathcal{M}$ .

**Note (Autonomous).** We have  $V(\phi(x, t))$ , not  $V(\phi(x, t), t)$ , i.e., it doesn’t explicitly depend on  $t$ .

**Remark.** The **flow** satisfies group property, i.e.,  $\phi(x, t_1 + t_2) = \phi(\phi(x, t_1), t_2)$  for all  $t_1, t_2 \in \mathbb{R}$ .

- Moreover, for all  $x \in \mathcal{M}$ , the **flow line** or orbit  $\gamma_x := \{\phi(x, t) \mid t \in \mathbb{R}\}$  through point  $x$  is flow-invariant, i.e., for all  $y \in \gamma_x$ ,  $t \in \mathbb{R}$ , we have  $\phi(y, t) \in \gamma_x$ .
- Finally, for all  $t \in \mathbb{R}$ ,  $\phi(\cdot, t): \mathcal{M} \rightarrow \mathcal{M}$  is a **diffeomorphism** of  $\mathcal{M}$  onto its image.

Naturally, we can consider the following two kinds of points of  $\mathcal{M}$ .

<sup>2</sup>We will make it formal.

**Definition 6.2.5** (Stable manifold). The *stable manifold* at  $x_0$  of the flow  $\phi$  are defined as

$$W^s(x_0) := \left\{ y \in \mathcal{M} \mid \lim_{t \rightarrow \infty} \phi(y, t) = x_0 \right\}.$$

**Definition 6.2.6** (Unstable manifold). The *unstable manifold* at  $x_0$  of the flow  $\phi$  are defined as

$$W^u(x_0) := \left\{ y \in \mathcal{M} \mid \lim_{t \rightarrow -\infty} \phi(y, t) = x_0 \right\}.$$

That is to say, we should focus on the dimension of  $W^u(p)$ .

**Intuition.** For  $t \rightarrow \pm\infty$ , each flow line  $x(t)$  defined as  $x: \mathbb{R} \rightarrow \mathcal{M}$  with  $\dot{x}(t) = -\text{grad } f(x(t))$  for all  $t \in \mathbb{R}$  converges to critical point, i.e.,

$$p = x(-\infty), \quad p = x(+\infty) \Rightarrow W^u(p) \text{ the all flow lines } x(t) \text{ with } x(-\infty) = p.$$

**As previously seen.** For a Morse function  $f$ , the set of critical points  $C(f) := \{x \in \mathcal{M} \mid df(x) = 0\}$

**Notation.** Denote the set of critical points of  $f$  of index  $k$  as  $\text{Crit}_k(f)$ .

**Definition 6.2.7** (Morse complex). Define the vector space over  $\mathbb{Z}/2\mathbb{Z}$  as

$$C_k(f, \mathbb{Z}_2) = C_k(f) := \left\{ \sum_{a \in \text{Crit}_k(f)} m_a a \mid m_a \in \mathbb{Z}/2\mathbb{Z} \right\}.$$

**Definition 6.2.8** (Boundary operator). The *boundary operator*  $\partial_k: C_k(f) \rightarrow C_{k-1}(f)$  by specifying its behavior on the basis elements. Given a critical point  $a \in C_k(f)$ ,  $\partial_k$  sends  $a$  to a linear combination of points in  $\text{Crit}_{k-1}(f)$  defined as

$$\partial_k(a) = \sum_{b \in \text{Crit}_{k-1}(f)} m(a, b) b$$

with  $m(a, b) \in \mathbb{Z}/2\mathbb{Z}$  being the number mod 2 of trajectories from  $a$  to  $b$ .<sup>a</sup>

<sup>a</sup>We can check that  $\partial \circ \partial = 0$ .

### 6.2.3 Morse Homology

With all the notions we have established, we have the following naturally.

**Definition 6.2.9** (Morse homology group). The *Morse homology group* is defined as

$$H_k(\mathcal{M}, f, \mathbb{Z}_2) := \ker \partial \text{ on } C_k(f) / \text{Im } \partial \text{ from } C_{k+1}(f).$$

**Remark.** The image of  $\partial$  from  $C_{k+1}(f, \mathbb{Z}_2)$  is always contained in the kernel of  $\partial$  on  $C_k(f, \mathbb{Z}_2)$ .

**Definition 6.2.10** (Betti number). The *Betti number* is defined as  $b_k := \dim_{\mathbb{Z}_2} H_k(\mathcal{M}, f, \mathbb{Z}_2)$ .



**Definition 6.2.11 (Euler characteristic).** The *Euler characteristic* of  $\mathcal{M}$  is defined as

$$\chi(\mathcal{M}) = \sum_i (-1)^i b^i.$$

Let's now calculate all these on our examples  $A = S^2$  and  $B$ .

**Example.** Revisit the example for  $f$ , we have

- $C_2(f) = \mathbb{Z} / 2\mathbb{Z}[p_1];$
- $C_0(f) = \mathbb{Z} / 2\mathbb{Z}[p_2];$
- $C_1(f) = 0.$

Also, we see that the chain complexes are

- $C_2 = \mathbb{Z}_2[p_1];$
- $C_0 = \mathbb{Z}_2[p_2];$
- $C_1 = 0;$

For kernels,

- $\ker \partial_2 = \{p_1\};$
- $\ker \partial_0 = \{p_2\};$
- and since  $\partial_1$  is trivial, so all images are trivial.

Finally, we calculate the homology groups as

- $H_2(A, f, \mathbb{Z}_2) = \mathbb{Z}_2;$
- $H_1(A, f, \mathbb{Z}_2) = 0;$
- $H_0(A, f, \mathbb{Z}_2) = \mathbb{Z}_2.$

**Example.** Revisit the example for  $g$ , we have

- $C_2(g) = \mathbb{Z} / 2\mathbb{Z}[p_1];$
- $C_1(g) = \mathbb{Z} / 2\mathbb{Z}[p_3];$
- $C_0(g) = \mathbb{Z} / 2\mathbb{Z}[p_4];$
- $C_k(g) = 0$  for  $k \geq 3.$

Also, we see that the chain complexes are

- $\partial_2(p_1) = p_3 = \partial_2(p_2),$  and  $\partial_2(p_1 + p_2) = 2p_3 = 0;$
- $\partial_1(p_3) = 2p_4 = 0.$

So the kernels are

- $\ker \partial_1 = \text{Im } \partial_2 = \mathbb{Z}_2[p_3];$
- $\ker \partial_2 = \mathbb{Z}_2[p_1 + p_2];$
- $\ker \partial_0 = \mathbb{Z}_2[p_4],$

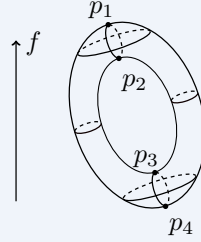
and all other images and kernels are trivial. Finally, we calculate the homology groups as

- $H_2(B, g, \mathbb{Z}_2) = \ker \partial_2 = \mathbb{Z}_2$ ;
- $H_1(B, g, \mathbb{Z}_2) = \ker \partial_1 / \text{Im } \partial_2 = 0$ ;
- $H_0(B, g, \mathbb{Z}_2) = \mathbb{Z}_2$ .

## Lecture 24: Introduction to the Rauch Comparison Theorem

**Example (Tilted torus).** Consider the tilted torus

4 Apr. 13:00



We see that

- $C_2 = \mathbb{Z}_2[p_1]$ ;
- $C_1 = \mathbb{Z}_2[p_2] \oplus \mathbb{Z}_2[p_3]$ ;
- $C_0 = \mathbb{Z}_2[p_1]$ .

Moreover, the chain complex is

$$0 \xrightarrow{\partial_3} \mathbb{Z}[p_1] \xrightarrow{\partial_2} \mathbb{Z}_2[p_2] + \mathbb{Z}_2[p_3] \xrightarrow{\partial_1} \mathbb{Z}_2[p_4] \xrightarrow{\partial_0} 0$$

## 6.3 The Rauch Comparison Theorem

In this section, our goal is to compare [Riemannian manifolds](#)  $(\mathcal{M}, g)$  with other [Riemannian manifolds](#) of constant [curvatures](#) model spaces, e.g.,  $S^n$ ,  $\mathbb{R}^n$ , and  $\mathbb{H}^n$ .

**Notation (Model space).** The set of *model spaces* is denoted as  $\mathcal{M}_m \in \{S^n, \mathbb{R}^n, \mathbb{H}^n\}$ .

### 6.3.1 Preliminary Estimations

Let  $c(t)$  be a [geodesic](#) with  $\|\dot{c}\| = 1$ ,  $v \in T_{c(0)}\mathcal{M}$ . Furthermore, let  $\mathcal{J}(t)$  be the [Jacobi field](#) along  $c(t)$  with  $\mathcal{J}(0) = 0$  and  $\dot{\mathcal{J}}(0) = v$  given by

$$\begin{cases} (\sin t)v, & \text{for } S^n; \\ tv, & \text{for } \mathbb{R}^n; \\ (\sinh t)v, & \text{for } \mathbb{H}^n. \end{cases}$$

Now, consider  $(\mathcal{M}, g)$  such that  $\lambda \leq \kappa \leq \mu$  with  $\lambda \leq 0$  and  $\mu \geq 0$ .

**Notation.** For  $\rho \in \mathbb{R}$ ,

$$c_\rho(t) = \begin{cases} \cos(\sqrt{\rho}t), & \text{if } \rho > 0; \\ 1, & \text{if } \rho = 0; \\ \cosh(\sqrt{-\rho}t), & \text{if } \rho < 0, \end{cases}$$

and also,

$$s_\rho(t) = \begin{cases} \frac{1}{\sqrt{\rho}} \sin(\sqrt{\rho}t), & \text{if } \rho > 0; \\ t, & \text{if } \rho = 0; \\ \frac{1}{\sqrt{-\rho}} \sinh(\sqrt{-\rho}t), & \text{if } \rho < 0, \end{cases}$$

These are solutions of **Jacobi equations** for constant **sectional curvature**  $\rho$ , i.e.,

$$\ddot{f}(t) + \rho f(t) = 0$$

with corresponding initial values  $f(0) = 0$ ,  $\dot{f}(0) = 1$ , respectively,  $f(0) = 1$ ,  $\dot{f}(0) = 0$ .

**Theorem 6.3.1.** Assume  $\kappa \leq \mu$  and  $\|\dot{c}\| \equiv 1$ , and assume either  $\mu \geq 0$  or  $\mathcal{J}^{\tan} \equiv 0$ . Let  $f_\mu := |\mathcal{J}(0)|c_\mu + |\mathcal{J}'(0)s_\mu$  solve

$$\ddot{f} + \mu f = 0$$

with  $f(0) = |\mathcal{J}(0)|$  and  $\dot{f}(0) = |\mathcal{J}'(0)|$ . If  $f_\mu(t) > 0$  for  $0 < t < \tau$ , then the following holds.

- (a)  $\langle \mathcal{J}, \dot{\mathcal{J}} \rangle f_\mu \geq \langle \mathcal{J}, \mathcal{J} \rangle \dot{f}_\mu$  on  $[0, \tau]$ .
- (b)  $1 \leq \frac{|\mathcal{J}(t_0)|}{f_\mu(t_1)} \leq \frac{|\mathcal{J}(t_2)|}{f_\mu(t_2)}$  if  $0 < t_1 \leq t_2 < \tau$ .
- (c)  $|\mathcal{J}(0)|c_\mu(t) + |\mathcal{J}'(0)s_\mu(t) \leq |\mathcal{J}(t)|$  for  $0 \leq t \leq \tau$ .

**Proof.** Firstly, we have that

$$|\mathcal{J}'| = \frac{\langle \mathcal{J}, \dot{\mathcal{J}} \rangle}{|\mathcal{J}|}, \quad |\mathcal{J}''| = \frac{\langle \dot{\mathcal{J}}, \dot{\mathcal{J}} \rangle}{|\mathcal{J}|} + \frac{\langle \mathcal{J}, \ddot{\mathcal{J}} \rangle}{|\mathcal{J}|} - \frac{\langle \mathcal{J}, \dot{\mathcal{J}} \rangle^2}{|\mathcal{J}|^3},$$

so

$$|\mathcal{J}''| + \mu|\mathcal{J}| = \frac{1}{|\mathcal{J}|} (-\langle R(\mathcal{J}, \dot{c})\dot{c}, \mathcal{J} \rangle + \mu\langle \mathcal{J}, \mathcal{J} \rangle) + \frac{1}{|\mathcal{J}|^3} (|\dot{\mathcal{J}}|^2|\mathcal{J}|^2 - \langle \mathcal{J}, \dot{\mathcal{J}} \rangle^2) \geq 0$$

since  $\kappa \leq \mu$  for  $0 < t < \tau$ , provided  $\mathcal{J}$  has no zeros on  $(0, \tau)$ . Moreover,

$$\left( |\mathcal{J}'|f_\mu - |\mathcal{J}|\dot{f}_\mu \right)' = |\mathcal{J}''|f_\mu - |\mathcal{J}|\ddot{f}_\mu \geq 0$$

since  $\ddot{f}_\mu + \mu f_\mu = 0$  for  $f_\mu(t) \geq 0$ . Also, we have  $|\mathcal{J}|(0) = f_\mu(0)$ ,  $|\mathcal{J}'|(0) = \dot{f}_\mu(0)$ , implying

$$|\mathcal{J}'|f_\mu - |\mathcal{J}|\dot{f}_\mu \geq 0,$$

which proves the first claim.

Furthermore,

$$\left( \frac{|\mathcal{J}|}{f_\mu} \right)' = \frac{1}{f_\mu^2} (|\mathcal{J}'|f_\mu - |\mathcal{J}|\dot{f}_\mu) \geq 0,$$

then since first zero of  $\mathcal{J}$  cannot occur before the first zero of  $f_\mu$ , the second claim is proved. The last claim follows directly from this.  $\blacksquare$

**Remark.**  $f_\mu(t) > 0$  for  $0 < t < \tau$  is necessary.

**Proof.** Take  $S^{n(\mu-\epsilon)}$  with  $\mathcal{J}(0) = 0$ . We see that  $f_\mu(t)$  has a zero at  $t = \pi/\sqrt{\mu}$  and  $\mathcal{J}(t)$  has one at  $t = \pi/\sqrt{\mu-\epsilon}$ . For small  $\epsilon > 0$  and any  $t$ , only slightly longer than  $\pi/\sqrt{\mu-\epsilon}$ , we have  $\frac{|\mathcal{J}(t)|}{f_\mu(t)} < 1$ .  $\circledast$

**Corollary 6.3.1.** Suppose that  $\kappa \leq \mu$ ,  $c_\mu \geq 0$  on  $(0, \tau)$ , and  $\mu \geq 0$  or  $\mathcal{J}^{\tan} \equiv 0$ . Let  $\|\dot{c}\| \equiv 1$ ,

$\mathcal{J}(0) = 0$ ,  $|R| < \Lambda$  with  $R$  being the **curvature tensor**. Then,

$$|\mathcal{J}(t) - t\dot{\mathcal{J}}(t)| \leq |\mathcal{J}(\tau)| \frac{1}{2} \Lambda t^2.$$

**Theorem 6.3.2.** Assume that  $\lambda \leq \kappa \leq \mu$ , and either  $\lambda \leq 0$  or  $\mathcal{J}^{\text{tan}} \equiv 0$ ,  $\|\dot{c}\| \equiv 1$ , and  $\mathcal{J}(0)$ ,  $\dot{\mathcal{J}}(0)$  be linearly dependent. Finally, assume  $s_{(\lambda+\mu)/2} > 0$  on  $(0, \tau)$ . Then, for  $0 \leq t \leq \tau$ ,

$$|\mathcal{J}(t)| \leq |\mathcal{J}(0)|c_\lambda(t) + |\mathcal{J}'(0)|s_\lambda(t).$$

**Proof idea.** Let  $\rho \in \mathbb{R}$ ,  $\eta := \max(\mu - \rho, \rho - \lambda)$ . Let  $A$  be a **vector field along  $c$**  with  $\ddot{A} + \rho A = 0$ ,  $A(0) = \mathcal{J}(0)$ , and  $\dot{A}(0) = \dot{\mathcal{J}}(0)$ .

Let  $a: I \rightarrow \mathbb{R}$  being a solution of  $\ddot{a} + (\rho - \eta)a = \eta|A|$ ,  $a(0) = \dot{a}(0) = 0$ , and  $b: I \rightarrow \mathbb{R}$  solving  $\ddot{b} + \rho b = \eta|\mathcal{J}|$ ,  $b(0) = \dot{b}(0) = 0$ . ■

## Lecture 25: Rauch Comparison Theorems and Sphere Theorem

### 6.3.2 Rauch Comparison Theorem

6 Apr. 13:00

We're now ready to provide the general statement of Rauch.

**Theorem 6.3.3 (Rauch comparison theorem).** Let  $(\mathcal{M}^m, g)$ ,  $(\overline{\mathcal{M}}^m, \overline{g})$  be **Riemannian manifolds** and  $\gamma: [0, a] \rightarrow \mathcal{M}$ ,  $\overline{\gamma}: [0, a] \rightarrow \overline{\mathcal{M}}$  be normalized **geodesics** with  $\gamma(0) = p$ ,  $\overline{\gamma}(0) = \overline{p}$ . Let  $X, \overline{X}$  be **Jacobi fields** along  $\gamma, \overline{\gamma}$ , respectively such that  $X(0) = \overline{X}(0) = 0$ ,  $|\nabla_{\dot{\gamma}(0)} X| = |\nabla_{\dot{\overline{\gamma}}(0)} \overline{X}|$ , and  $\langle \dot{\gamma}(0), \nabla_{\dot{\gamma}(0)} X \rangle = \langle \dot{\overline{\gamma}}(0), \nabla_{\dot{\overline{\gamma}}(0)} \overline{X} \rangle$ . Furthermore, assume that

- (a)  $\gamma$  has no **conjugate points** on  $[0, a]$ ;
- (b) **sectional curvatures**  $K, \overline{K}$  of  $\mathcal{M}, \overline{\mathcal{M}}$  satisfy  $\overline{K} \leq K$  for all 2-planes containing  $\dot{\gamma}, \dot{\overline{\gamma}}$ .

Then,  $\overline{\gamma}$  has no **conjugate points** on  $[0, a]$ , and for all  $t \in [0, a]$ ,

$$|X(t)| \leq |\overline{X}(t)|.$$

**Proof idea.** To prove this, we first see a lemma.

**Lemma 6.3.1.** Let  $(\mathcal{M}^m, g)$ ,  $(\overline{\mathcal{M}}^m, \overline{g})$  be **Riemannian manifolds** and  $\gamma: [0, a] \rightarrow \mathcal{M}$ ,  $\overline{\gamma}: [0, a] \rightarrow \overline{\mathcal{M}}$  be normalized **geodesics** with  $\gamma(0) = p$ ,  $\overline{\gamma}(0) = \overline{p}$ . Let  $X, \overline{X}$  be **Jacobi fields** along  $\gamma, \overline{\gamma}$ , respectively such that  $X(0) = \overline{X}(0) = 0$ . Furthermore, assume that

- (a)  $\gamma$  has no **conjugate points** on  $[0, a]$ ;
- (b) **sectional curvatures**  $K, \overline{K}$  of  $\mathcal{M}, \overline{\mathcal{M}}$  satisfy  $\overline{K} \leq K$  for all 2-planes containing  $\dot{\gamma}, \dot{\overline{\gamma}}$ .

Finally, assume that  $|X(a)| = |\overline{X}(a)|$ . Then,  $I(X, X) \leq I(\overline{X}, \overline{X})$ .

**Proof idea.** We first choose an orthonormal frame in  $(\mathcal{M}, g)$  and  $(\overline{\mathcal{M}}, \overline{g})$  with  $e_1 = \dot{\gamma}$  and  $\overline{e}_1 = \dot{\overline{\gamma}}$ , and  $e_2 = X(a)/|X(a)| \neq 0$ , etc. Consider  $X(t) = X^i(t)e_i(t)$  and the same for  $\overline{X}$ . Then, the second variation of the **energy** shows  $I(X, X) \leq I(\overline{X}, \overline{X})$ . ■

Now, consider normal components of  $X, \overline{X}$  only, and we can show that

$$\lim_{t \rightarrow 0} \frac{|X(t)|^2}{|\overline{X}(t)|^2} =: \lim_{t \rightarrow 0} \frac{\overline{u}(t)}{u(t)} = 1,$$

thus to prove  $|X| \leq |\overline{X}|$ , it's enough to show that

$$\frac{d}{dt} \frac{|X(t)|^2}{|\overline{X}(t)|^2} \geq 0,$$

equivalently,  $\dot{\bar{u}} - \bar{u}\dot{u} \geq 0$ . Then, since  $\gamma$  has no [conjugate points](#), we have  $u(t) > 0$ . Let  $c \in [0, a]$  be the greatest number such that  $\bar{u}(t) > 0$  on  $(0, c)$ . Then, for all  $b \in (0, c)$ , define

$$X_b(t) = \frac{X(t)}{|X(b)|}, \quad \bar{X}_b(t) = \frac{\bar{X}(t)}{|\bar{X}(b)|}.$$

From [Lemma 6.3.1](#) to  $I(X_b, X_b), I(\bar{X}_b, \bar{X}_b)$ , then we're done. ■

**Corollary 6.3.2.** Let  $(\mathcal{M}, g)$  be a [complete](#) and simply-connected [Riemannian manifold](#) with non-positive [sectional curvature](#), and  $\triangle ABC$  is a [geodesic](#) triangle in  $\mathcal{M}$ , then

- (a)  $|AB|^2 + |AC|^2 - 2|AB||AC|\cos\angle A \leq |BC|^2$ ;
- (b)  $\angle A + \angle B + \angle C \leq \pi$ .

**Corollary 6.3.3.** Suppose that [sectional curvature](#) of  $(\mathcal{M}, g)$  satisfies

$$0 < C_1 \leq K \leq C_2$$

for some constants  $C_1, C_2$ . Let  $\gamma$  be any [geodesic](#) in  $\mathcal{M}$ . Then, the distance  $d$  between any two [conjugate points](#) of  $\gamma$  satisfies

$$\frac{\pi}{\sqrt{C_2}} \leq d \leq \frac{\pi}{\sqrt{C_1}}.$$

**Corollary 6.3.4.** Let  $(\mathcal{M}, g)$  be compact [Riemannian manifold](#) where the [sectional curvature](#)  $K$  satisfies  $K \leq C$  for some constant  $C$ . Then, either the [injectivity radius](#)

$$i(\mathcal{M}, g) \geq \pi/\sqrt{C},$$

or there exists a closed [geodesic](#)  $\gamma$  in  $\mathcal{M}$  whose [length](#) is minimal among all closed [geodesics](#) such that

$$i(\mathcal{M}, g) \geq \frac{1}{2}L(\gamma).$$

## 6.4 The Sphere Theorem

In this section, we want to prove the following.

**Theorem 6.4.1** (Sphere theorem). Let  $\mathcal{M}^n$  be a compact and simply-connected [Riemannian manifold](#) with [sectional curvature](#)  $K$  such that

$$0 < hK_{\max} < K \leq K_{\max}.$$

Then if  $h = 1/4$ , then  $\mathcal{M}$  is homeomorphic to a sphere  $S^n$ .

**Notation** (Pinching number).  $h$  in the [sphere theorem](#) is called the *pinching number* of  $\mathcal{M}$ .

**Remark.** Another version of the [sphere theorem](#) is to assume  $0 < h < K \leq 1$  by scaling.

To prove this, Borger [Ber60], Klingenberg [Kli61] used [Rauch comparison theorem](#) with [Morse index theorem](#) in the 1960s.

### 6.4.1 Gauss-Bonnet Theorem and Theorem by Hamilton

To understand the [sphere theorem](#), we should consider  $n = 2, 3$ . In this case, it suffices to assume  $h \geq 0$ , i.e., for a compact and simply-connected [Riemannian manifold](#)  $\mathcal{M}^n$  with  $n = 2, 3$  such that it has positive

sectional curvature, then  $\mathcal{M}^n$  is homeomorphic to  $S^n$ .

**Note.** For  $n = 2$ , it follows from the Gauss-Bonnet theorem, and for  $n = 3$ , it follows from a theorem by R. Hamilton.

**Theorem 6.4.2** (Gauss-Bonnet theorem). Let  $\mathcal{M}$  be a compact connected 2-dimensional Riemannian manifold  $\mathcal{M}$  with Gauss curvature  $K$ . Then, its characteristic is given by

$$\chi(\mathcal{M}) = \frac{1}{2\pi} \int_{\mathcal{M}} K \, d\mu_{\mathcal{M}}.$$

The Gauss-Bonnet theorem generalizes the so-called Gauss-Bonnet formula.

**Note** (Gauss-Bonnet formula). Let  $\gamma$  be a curved polygon on an oriented Riemannian 2-manifold  $(\mathcal{M}, g)$  such that  $\gamma$  is positive oriented as the boundary of an open set  $\Omega$  with compact closure. Then,

$$\int_{\Omega} K \, dA + \int_{\gamma} k_N \, ds + \sum_i \epsilon_i = 2\pi,$$

where  $k_N(t) = \langle D_t \dot{\gamma}(t), N(t) \rangle$ ,<sup>a</sup> and  $\epsilon_i$  are the exterior angles.

<sup>a</sup> $N(t)$  is the normal vector field.

To understand all these, we need the following concept.

**Definition 6.4.1** (Smooth triangulation). For  $\mathcal{M}$  smooth, compact 2-manifold, a smooth triangulation of  $\mathcal{M}$  is a finite collection of curved triangles such that

- the union of the closed regions  $\overline{\Omega}_i$  bounded by the triangles is actually  $\mathcal{M}$ ;
- the intersection of any pair (if not empty) is either a single vertex of each or a single edge of each.

**Theorem 6.4.3** (Radó [Rad25]). Every compact topological 2-manifold has a triangulation.

**Note.** Let  $\mathcal{M}$  be a triangulated 2-manifold. Then, the Euler characteristic is

$$\chi(\mathcal{M}) = N_v - N_e + N_f.$$

This implies that

$$\int_{\mathcal{M}} K \, dA = 2\pi\chi(\mathcal{M}).$$

Then, we can start proving Gauss-Bonnet theorem.

**Proof of Theorem 6.4.2.** Let  $\{\Omega_i\}_{i=1}^N$  denote the faces of triangulation, and for all  $i$ , let  $\{\gamma_{ij} \mid j = 1, 2, 3\}$  be the edges of  $\Omega_i$  and  $\{\theta_{ij} \mid j = 1, 2, 3\}$  be its interior angles.

As each exterior angle is  $\pi$  minus the interior angle, by applying the Gauss-Bonnet formula to each triangle and sum over  $i$ , we have

$$\begin{aligned} \sum_{i=1}^{N_f} \int_{\Omega_i} K \, dA + \sum_{i=1}^{N_f} \sum_{j=1}^3 \int_{\gamma_{ij}} k_N \, ds + \sum_{i=1}^{N_f} \sum_{j=1}^3 (\pi - \theta_{ij}) &= \sum_{i=1}^{N_f} 2\pi \\ \Leftrightarrow \int_{\mathcal{M}} K \, d\mu_{\mathcal{M}} + 0 + 3\pi N_f - \sum_{i=1}^{N_f} \sum_{j=1}^3 \theta_{ij} &= 2\pi N_f \end{aligned}$$

where the second term vanishes since each edge appears twice but with opposite sign. Since degrees

at each vertex adds up to  $2\pi$ , we have

$$\int_{\mathcal{M}} K \, dA = 2\pi N_v - \pi N_f.$$

As each edge is in exactly 2 triangles and each triangle has 3 edges, we see that  $2N_e = 3N_f$ ,

$$\int_{\mathcal{M}} K \, dA = 2\pi N_v - 2\pi N_e + 2\pi N_f = 2\pi\chi(\mathcal{M}).$$

■

**Theorem 6.4.4 (Hamilton).** Let  $\mathcal{M}$  be a compact and simply-connected 3-dimensional Riemannian manifold  $\mathcal{M}$  with strictly positive Ricci curvature. Then,  $\mathcal{M}$  is diffeomorphic to  $S^3$ .

## Lecture 26: Toward Proving the Sphere Theorem

11 Apr. 13:00

**Corollary 6.4.1.** Let  $\mathcal{M}$  be a compact 2-dimensional Riemannian manifold, and  $K$  be the Gauss curvature.

- (a) If  $\mathcal{M}$  is homeomorphic to the sphere or the projective plane, then  $K > 0$  somewhere.
- (b) If  $\mathcal{M}$  is homeomorphic to torus or Klein bottle, then either  $K = 0$  or  $K$  takes on both positive and negative values.
- (c) If  $\mathcal{M}$  is any other compact surfaces, then  $K < 0$  somewhere.

**Corollary 6.4.2.** Let  $\mathcal{M}$  be a compact 2-dimensional Riemannian manifold, and  $K$  be the Gauss curvature.

- (a) If  $K > 0$ , then  $\mathcal{M}$  is homeomorphic to sphere or projective plane, and  $\pi_i(\mathcal{M})$  is finite.
- (b) If  $K \leq 0$ , then  $\pi_1(\mathcal{M})$  is infinite and  $\mathcal{M}$  has genus at least 1.

### 6.4.2 Beyond 2-Dimension

To go beyond 2-dimension, some consider the so-called Pfaffian.

**Definition 6.4.2 (Pfaffian).** Let  $\mathcal{P}$  be the map from  $(0,4)$ -tensors to  $\mathbb{R}$  with the domain carries symmetries as Riemannian curvature.

**Theorem 6.4.5.** On any oriented vector space, there exists a basis independent functions  $\mathcal{P}$  such that for all compact, even-dimension Riemannian manifold  $\mathcal{M}$ ,

$$\int_{\mathcal{M}} \mathcal{P}(R) \, dV = \frac{1}{2} \text{vol}(S^n) \chi(\mathcal{M}).$$

**Note.** This is too much information swallowed...

Another approach is to consider the following.

**Notation.** Let  $p \in \mathcal{M}$ , then  $d_p: \mathcal{M} \rightarrow \mathbb{R}$  such that  $d_p(q) = \text{dist}(p, q)$ .

We see that  $d_p$  is Lipschitz continuous and smooth on  $\mathcal{M} \setminus (\{p\} \cup \text{Cut}(p)) =: \mathcal{M}_p$ . At any  $q \in \mathcal{M}_p$ , the gradient  $\nabla d_p$  is the tangent vector at  $q$  of the unique normal minimizing geodesic from  $p$  to  $q$ . In particular,  $|\nabla d_p| = 1$  at most points of  $\mathcal{M}$ . Now, we want to compare distance functions on different manifolds. This requires comparing the Hessian.

As previously seen (Hessian). For all smooth function  $f$  on  $\mathcal{M}$ , its Hessian  $\nabla^2 f$  is defined as

$$\nabla^2 f(X, Y) = \langle \nabla_X \nabla f, Y \rangle.$$

The Hessian of  $f$  is symmetric, and we can write  $\Delta f = \text{Tr}(\nabla^2 f)$ .

**Notation.** Let  $K^+ := \max_{\sigma \subseteq T_p \mathcal{M}} K(\sigma)$  and  $K^- := \min_{\sigma \subseteq T_p \mathcal{M}} K(\sigma)$ .

**Theorem 6.4.6** (Hessian comparison theorem). Let  $(\mathcal{M}, g)$ ,  $(\widetilde{\mathcal{M}}, \widetilde{g})$  be complete Riemannian manifolds, and  $\gamma: [0, b] \rightarrow \mathcal{M}$  and  $\widetilde{\gamma}: [0, b] \rightarrow \widetilde{\mathcal{M}}$  be minimizing normal geodesics in  $\mathcal{M}$  and  $\widetilde{\mathcal{M}}$ , respectively, such that

$$\widetilde{K}^+(t) \leq K^-(t)$$

for all  $t \in [0, b]$ . Denote  $q = \gamma(a)$ ,  $\widetilde{q} = \widetilde{\gamma}(a)$  for  $a \leq b$ . Suppose  $X_q \in T_q \mathcal{M}$ ,  $\widetilde{X} \in T_{\widetilde{q}} \widetilde{\mathcal{M}}$  satisfy

$$\langle X_q, \dot{\gamma}(a) \rangle = \langle \widetilde{X}_{\widetilde{q}}, \dot{\widetilde{\gamma}}(a) \rangle$$

and  $|X_q| = |\widetilde{X}_{\widetilde{q}}|$ . Then,

$$\nabla^2 d_p(X_q, X_q) \leq \widetilde{\nabla}^2 \widetilde{d}_{\widetilde{p}}(\widetilde{X}_{\widetilde{q}}, \widetilde{X}_{\widetilde{q}}).$$

### 6.4.3 Toponogor Theorem

We now state the main tools we need in order to prove the sphere theorem.

**Definition.** Let  $(\mathcal{M}, g)$  be a complete Riemannian manifold.

**Definition 6.4.3** (Geodesic triangle). A geodesic triangle  $\triangle ABC$  consists of 3 points  $A, B, C \in \mathcal{M}$  and 3 minimizing geodesics (sides)  $\gamma_{AB}, \gamma_{BC}, \gamma_{CA}$  joining each 2 of them.

**Definition 6.4.4** (Generalized geodesic triangle). A generalized geodesic triangle  $\triangle ABC$  consists of 3 points  $A, B, C \in \mathcal{M}$  and 2 minimizing geodesics  $\gamma_{AB}, \gamma_{AC}$  and 1 geodesic  $\gamma_{BC}$  of length  $L(\gamma_{BC}) \leq L(\gamma_{AB}) + L(\gamma_{AC})$ , joining each 2 of them.

**Definition 6.4.5** (Geodesic hinge). A geodesic hinge  $\angle BAC$  consists of a point  $A \in \mathcal{M}$  and 2 minimizing geodesics  $\gamma_{AB}, \gamma_{AC}$  emanating from  $A$  with endpoints  $B, C$ .

**Definition 6.4.6** (Generalized geodesic hinge). A generalized geodesic hinge  $\angle BAC$  consists of a point  $A \in \mathcal{M}$  and 2 geodesics  $\gamma_{AB}, \gamma_{AC}$  emanating from  $A$  with endpoints  $B, C$ , with only one is minimizing.

For all  $k \in \mathbb{R}$ , denote  $\mathcal{M}_k^n$  the  $n$ -dimensional space form of constant curvature  $k$ , i.e.,

$$\mathcal{M}_k^n = S^n(k) \text{ or } \mathbb{R}^n \text{ or } \mathbb{H}^n(k).$$

**Lemma 6.4.1.** Let  $(\mathcal{M}^n, g)$  be a complete Riemannian manifold with sectional curvature  $K \geq k$ .

- (a) For all generalized geodesic hinge  $\angle BAC$  in  $\mathcal{M}$ , there exists a geodesic hinge  $\angle \widetilde{B}\widetilde{A}\widetilde{C}$  in  $\mathcal{M}_k^n$  with the same angle and corresponding sides are with the same length as  $\angle BAC$ .
- (b) For all generalized geodesic triangle  $\triangle ABC$  in  $\mathcal{M}$ , there exists a geodesic triangle  $\triangle \widetilde{A}\widetilde{B}\widetilde{C}$  in  $\mathcal{M}_k^n$  whose corresponding sides have the same length as  $\triangle ABC$ .



**Theorem 6.4.7** (Toponogor theorem). Let  $(\mathcal{M}, g)$  be a complete Riemannian manifold with sectional curvature  $K \geq k$ .

- (a) Let  $\angle BAC$  be a geodesic hinge in  $\mathcal{M}$  and  $\angle \tilde{B}\tilde{A}\tilde{C}$  in  $\mathcal{M}_k^n$ . Then,  $\text{dist}(B, C) = \text{dist}(\tilde{B}, \tilde{C})$ .
- (b) Let  $\triangle ABC$  be a geodesic triangle in  $\mathcal{M}$ ,  $\triangle \tilde{A}\tilde{B}\tilde{C}$  in  $\mathcal{M}_k^n$ . Then, the 3 angles in  $\triangle ABC$  are greater than the corresponding angles in  $\triangle \tilde{A}\tilde{B}\tilde{C}$ .

**Theorem 6.4.8** (Klingenberg). Let  $(\mathcal{M}, g)$  be a complete, simply-connected Riemannian manifold with sectional curvature  $1/4 < K \leq 1$ . Then,

$$\text{Inj}(\mathcal{M}, g) \geq \pi.$$

#### 6.4.4 Proof of the Sphere Theorem

Now, we can prove the sphere theorem. Let's first restate it (after scaling) for our reference.

**Theorem 6.4.9** ((Scaled) Sphere theorem). Let  $\mathcal{M}^n$  be a compact and simply-connected Riemannian manifold with sectional curvature  $K$  such that

$$\frac{1}{4} < K \leq 1.$$

Then  $\mathcal{M}$  is homeomorphic to a sphere  $S^n$ .

**Proof.** By Bonnet-Mayers theorem, we know that  $\mathcal{M}$  is compact, hence there exists  $k > 1/4$  such that  $k \leq K \leq 1$ . By the Klingenberg theorem,

$$\ell = \text{diam}(\mathcal{M}, g) \geq \text{Inj}(\mathcal{M}, g) \geq \pi > \frac{\pi}{2\sqrt{k}}.$$

Take  $p, q \in \mathcal{M}$  such that  $\text{dist}(p, q) = \text{diam}(\mathcal{M}, g)$ . Let  $q_0 \in \mathcal{M}$  such that  $\ell_1 = \text{dist}(p, q_0) > \pi/2\sqrt{k}$ , and  $\gamma_1$  be a minimizing normal geodesic connecting  $p = \gamma_1(0)$  and  $q_0 = \gamma_1(\ell_1)$ . Then, consider the following.

**Lemma 6.4.2.** Let  $(\mathcal{M}, g)$  be a compact Riemannian manifold where there exists  $p, q \in \mathcal{M}$  such that  $\text{dist}(p, q) = \text{diam}(\mathcal{M}, g)$ . Then, for all  $X_p \in T_p\mathcal{M}$ , there exists a minimizing geodesic  $\gamma$  connecting  $p = \gamma(0)$  to  $q$  such that

$$\langle \dot{\gamma}(0), X_p \rangle \geq 0.$$

From this, there exists a minimizing normal geodesic  $\gamma_2$  connecting  $p = \gamma_2(0)$  to  $q = \gamma_2(\ell)$  such that  $\langle \dot{\gamma}_1(0), \dot{\gamma}_2(0) \rangle \geq 0$ , i.e., the angle  $\alpha$  between  $\dot{\gamma}_1(0)$  and  $\dot{\gamma}_2(0)$  is no more than  $\pi/2$ . According to Toponogor theorem, by looking at the geodesic hinge  $\angle q_1 p q$ ,  $\text{dist}(q_1, q) \leq \text{dist}(\tilde{q}_1, \tilde{q})$  for a comparison geodesic hinge  $\angle \tilde{q}_1 \tilde{p} \tilde{q}$  in  $\mathcal{M}_k^n = S^n(1/\sqrt{k})$ . Now, due to the cosine law for  $S^n(1/\sqrt{k})$ ,

$$\cos(\sqrt{k} \cdot \text{dist}(q_1, q)) \geq \cos(\sqrt{k} \cdot \text{dist}(\tilde{q}_1, \tilde{q})) = \cos \sqrt{k} \ell \cdot \cos \sqrt{k} \ell_1 + \sin \sqrt{k} \ell \cdot \sin \sqrt{k} \ell_1 \cos \alpha \geq \dots > 0.$$

## Lecture 27: Uniformization Theorem

Before continue the proof of Theorem 6.4.9, we need one more tool.

13 Apr. 13:00

**Theorem 6.4.10** (Brown's theorem). Let  $\mathcal{M}$  be a smooth and compact manifold. If  $\mathcal{M} = U_1 \cup U_2$  with  $U_1, U_2$  open subsets in  $\mathcal{M}$  homeomorphic to  $\mathbb{R}^n$ , then  $\mathcal{M}$  is homeomorphic to  $S^n$ .

We can then finish the proof.

**Proof of Theorem 6.4.9 (Continue).** We have shown that

$$\overline{B_{\frac{\pi}{2\sqrt{k}}}(p)} \cup \overline{B_{\frac{\pi}{2\sqrt{k}}}(q)} = \mathcal{M}.$$

Denote  $r$  to be

$$r := \frac{1}{2} \left( \text{Inj}(\mathcal{M}, g) + \frac{\pi}{2\sqrt{k}} \right) > \frac{\pi}{2\sqrt{k}},$$

then  $\mathcal{M} = B_r(p) \cup B_r(q)$ . Moreover, since  $r < \text{Inj}(\mathcal{M}, g)$ , both  $B_r(p)$  and  $B_r(q)$  are homeomorphic to  $\mathbb{R}^n$ . By [Brown's theorem](#), the result follows. ■

### 6.4.5 The Family of the Sphere Theorem

The [sphere theorem](#) doesn't hold for  $1/4 \leq K \leq 1$ .<sup>3</sup>

**Example.** The complex projective spaces  $\mathbb{CP}^n$  are also compact, simply connected [Riemannian manifold](#) such that  $1/4 \leq K \leq 1$ . But they are not homeomorphic to  $S^{2m}$ .

However, this is “almost” the only counterexample in the following sense.

**Theorem 6.4.11.** Let  $\mathcal{M}^n$  be a compact and simply-connected [Riemannian manifold](#).

- (a) If  $m$  is even, then there exists  $\epsilon(m) > 0$  such that  $1/4 - \epsilon(m) \leq K \leq 1$ , then  $\mathcal{M}$  is either homeomorphic to  $S^n$  or [diffeomorphic](#) to either  $\mathbb{CP}^{m/2}$ ,  $\mathbb{HP}^{m/4}$ , or  $\mathbb{C}_a\mathbb{P}^2$  [[Ber83](#)].
- (b) If  $m$  is odd, then there exists  $\epsilon > 0$  such that if  $1/4 - \epsilon \leq K \leq 1$ , then  $\mathcal{M}$  homeomorphic to  $S^n$  [[AM96](#)].

Looking back,

- Rauch in 1951 proved the [sphere theorem](#) for  $3/4 < K \leq 1$  [[Rau51](#)];
- Klingenberg in 1959 proved the [sphere theorem](#) for  $0.55 < K \leq 1$  [[Kli59](#)];
- Berger in 1960 proved the [sphere theorem](#) for  $1/4 < K \leq 1$  when  $m$  is even [[Ber60](#)];
- Klingenberg in 1961 proved the [sphere theorem](#) [[Kli61](#)].

Now, what if we want [diffeomorphism](#)? “Exotic spheres” exists: [manifolds](#) that are homeomorphic to sphere but not [diffeomorphic](#).

**Example (J. Milnor).** If  $n = 7$ , we can construct as  $S^3$ -[bundles](#) over  $S^4$ .

Consider  $m = 2$ , from the [Gauss-Bonnet theorem](#),  $\mathcal{M}$  is [diffeomorphic](#) to  $S^2$  since

$$0 < \int_{\mathcal{M}} K \, dA = 2\pi\chi(\mathcal{M}),$$

with the fact that  $S^2$  is the only such object. As for  $m = 3$ , by the [Hamilton's proof](#), if  $(\mathcal{M}, g)$  is a 3-dimension compact [Riemannian manifold](#) with [Ricci curvature](#)  $> 0$  then  $(\mathcal{M}, g)$  is [diffeomorphic](#) to  $S^3$  using Ricci flow, e.g., [[BS08](#)].

<sup>3</sup>Where we originally have  $1/4 < K \leq 1$ .

# Chapter 7

## Epilogue

In the end of this long journey, we wrap up this course by showing some results and directions that can be further explored based on what we have learned.

### 7.1 Uniformization Theorem

Recall the [Gauss-Bonnet theorem](#), where we let  $(\mathcal{M}, g)$  be a compact, 2-dimension, [oriented Riemannian manifold](#) without boundary. And let  $K$  be the [Gauss curvature](#),  $\gamma$  be the genus of  $\mathcal{M}$ , then,

$$\int_{\mathcal{M}} K \, d\mu_g = 2\pi\chi.$$

In particular, there are three cases:

- (a)  $\gamma = 1$ , then  $\chi = 0$ ;
- (b)  $\gamma \geq 2$ , then  $\chi$  is negative integer;
- (c)  $\gamma = 0$ , then  $\chi = 2$ .

Now, we want to generalize it. To do this, we need the following notion.

**Definition 7.1.1** (Conformal). A [metric](#)  $g$  is *conformal* to another [metric](#)  $\tilde{g}$  if there exists a positive [smooth function](#)  $\Omega$  on  $\mathcal{M}$  such that

$$\tilde{g} = \Omega^2 \circ g$$

has constant [Gauss curvature](#).

**Note** (Constant rescaling). If we take  $\Omega = c$  for some constant, then  $\tilde{K} = c^{-2}K$ .

Now we state the [uniformization theorem](#).

**Theorem 7.1.1** (Uniformization theorem). If  $g$  is [conformal](#) to a [metric](#) of constant [Gauss curvature](#), then

$$\tilde{K} = \begin{cases} 0, & \text{if } \gamma = 1; \\ -1, & \text{if } \gamma \geq 2; \\ 1, & \text{if } \gamma = 0. \end{cases}$$

First, we check how [Gauss curvature](#) transforms under [conformal](#) transformations. Consider an  $n$ -dimensional [Riemannian manifold](#)  $(\mathcal{M}, g)$  and  $\tilde{g}_{ij} = \Omega^2 g_{ij}$ , then

$$\tilde{\Gamma}_{ij}^k = \Gamma_{ij}^k + \Omega^{-1}(\delta_i^k \partial_j \Omega + \delta_j^k \partial_i \Omega - g^{k\ell} g_{ij} \partial_\ell \Omega).$$

Moreover, we can also compute how do the [Riemannian curvature](#), [Ricci curvature](#), and also [sectional curvature](#) transform.

In particular, for  $n = 2$ , set  $\Omega = e^u$  for some smooth function  $u$ , then the Gauss curvature  $K$  is transformed as

$$\tilde{K} = e^{-2u}(K - \Delta_g u).$$

Now, we want to find some smooth solutions  $u$  of

$$\Delta_g u + \tilde{K} e^{2u} = K$$

when  $\gamma = 0$ ,  $\chi = 2$ , we have

$$\Delta_g u + e^{2u} = K.$$

**Note.** The maximum principle does not work.

### 7.1.1 Proof of the Uniformization Theorem

We now start to prove the uniformization theorem. We follow 5 steps to prove the theorem.

1. Find points  $p, o \in \mathcal{M}$  such that  $\text{dist}(p, o) = \text{diam}(\mathcal{M})$ .
2. Find a function  $w$  on  $\mathcal{M} \setminus \{p\}$  such that  $\tilde{g} = e^{2w}g$  on  $\mathcal{M} \setminus \{p\}$  makes  $(\mathcal{M} \setminus \{p\}, \tilde{g})$  isometric to a plane, i.e.,  $\tilde{K} = 0$ . From here, we get  $\Delta_g w - K = 0$  on  $\mathcal{M} \setminus \{p\}$ . So naturally, we define  $w$  to be a solution of

$$\Delta_g w - K = -4\pi\delta_p$$

on  $\mathcal{M}$ , where  $\delta_p$  is the delta Dirac distribution at  $p$ .

3. Let  $\tilde{d}_o$  be the distance function on  $(\mathcal{M} \setminus \{p\}, \tilde{g})$  from  $o$ . Set

$$e^v = \frac{1}{1 + \tilde{d}_o^2/4},$$

then metric  $\tilde{\tilde{g}} = 2^{2v}\tilde{g}$  has Gauss curvature  $\tilde{\tilde{K}} = 1$ . Namely,  $(\mathcal{M} \setminus \{p\}, \tilde{\tilde{g}})$  will be isometric to a standard sphere minus the North Pole  $N$ .

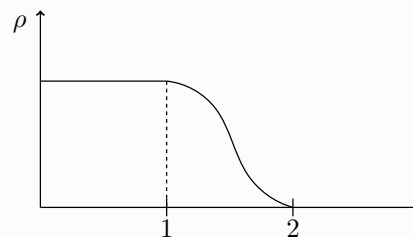
4. Set  $u = w + v$ , then we show that this function extends continuously to the point  $p$ , i.e.,  $\tilde{\tilde{g}} = e^{2u}g$  is a metric on  $\mathcal{M}$  such that  $(\mathcal{M}, \tilde{\tilde{g}})$  is isometric to a standard sphere.
5. Show that  $u$  is smooth on  $\mathcal{M}$ .<sup>1</sup>

## Lecture 28: Epilogue

18 Apr. 13:00

**Proof of Theorem 7.1.1.** We now start our proof following the 5 steps.

1. This step is trivial.
2. Use the exponential map. Let  $r_p$  be the injectivity radius (relative to  $(\mathcal{M}, g)$ ) of  $\exp_p$ . Choose  $\epsilon > 0$  such that  $2\epsilon < r_p$ . Let  $\rho$  be a  $C^\infty$  non-increasing function on  $[0, \infty)$  such that  $\rho = 1$  on  $[0, 1]$  and  $\rho = 0$  on  $[2, \infty)$ .



<sup>1</sup> $w \rightarrow \infty$  while  $v \rightarrow -\infty$  as we approach  $p$ . To show  $u = w + v$ , we need to analyze the blow-behavior for  $w$  and  $v$ .

Define the cut-off function for  $q \in \mathcal{M}$  such that

$$\eta(q) = \begin{cases} \rho(d_p/\epsilon), & \text{if } q \in B_{2\epsilon}(p); \\ 0, & \text{otherwise.} \end{cases}$$

Namely,  $\eta = 1$  on  $B_\epsilon(p)$  and  $\eta = 0$  on  $\mathcal{M} \setminus B_{2\epsilon}(p)$ . Define on  $\mathcal{M}$  the function

$$w_o = \begin{cases} -2\eta \log d_p, & \text{in } B_{2\epsilon}(p); \\ 0, & \text{on } \mathcal{M} \setminus B_{2\epsilon}(p). \end{cases}$$

In  $B_\epsilon(p)$ ,  $w_o = -2 \log d_p$ . As  $2\epsilon < r_p$ ,  $\exp_p$  is a **diffeomorphism** of the ball of radius  $2\epsilon$  with center 0 in  $T_p\mathcal{M}$  onto  $B_{2\epsilon}(p)$  in  $\mathcal{M}$ . Consider choosing a polar **normal coordinate**  $(r, \theta)$  in  $B_{2\epsilon}(p)$  such that  $d\rho = r$  and  $g = dr^2 + R^2(r, \theta)d\theta^2$  such that

$$\int_0^\pi R(r, \theta) d\theta = L(r)$$

be the perimeter of the **geodesic** circles. Then  $L(r)/r \rightarrow 2\pi$  as  $r \rightarrow 0$  by local euclidicity at  $p$ .

**Notation** (Geodesic curvature). The *geodesic curvature* of circles  $\kappa$  is defined as  $\kappa = \frac{1}{R} \frac{\partial R}{\partial r}$ .

We have

$$\frac{\partial \kappa}{\partial r} = -\kappa^2 - K.$$

Now, express the Laplace operator  $\Delta$  in polar coordinates.

**Remark.** In arbitrary coordinates, we have

$$\Delta_g = \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^a} \cdot \sqrt{\det g} (g^{-1})^{ab} \frac{\partial}{\partial x^b}.$$

Then in polar coordinates,

$$\Delta_g = \frac{1}{R} \frac{\partial}{\partial r} \cdot R \frac{\partial}{\partial r} + \frac{1}{r} \frac{\partial}{\partial \theta} \cdot \frac{1}{R} \frac{\partial}{\partial \theta},$$

and in  $B_\epsilon(p)$ , we have  $\omega_0 = -2 \log r$  and

$$\Delta_g \omega_0 = -\frac{2}{R} \frac{\partial R}{\partial r} \left( \frac{R}{r} \right) = -\frac{2\lambda}{r}$$

with  $\lambda = \frac{1}{R} \frac{\partial R}{\partial r} - \frac{1}{r} = \kappa - \frac{1}{r}$ . Hence,

$$\underbrace{\frac{\partial \kappa}{\partial r} + \kappa^2}_{-K} = \frac{\partial \lambda}{\partial r} + \frac{2\lambda}{r} + \lambda^2.$$

Set  $\mu = r^2 \lambda$ , it becomes

$$\frac{\partial \mu}{\partial r} + \frac{\mu^2}{r^2} = -r^2 K.$$

Since as  $r \rightarrow 0$ ,  $\mu \rightarrow 0$ , hence along each ray, the integral equation  $\mu(r, \theta)$  is

$$\mu(r, \theta) = - \int_0^r \frac{\mu(r', \theta)^2}{r'^2} + r'^2 K(r', \theta) dr',$$

and we have  $\lambda(r, \theta) = O(r)$ . Moreover,  $\lambda/r \rightarrow -K_p/3$  as  $r \rightarrow 0$ . It follows that  $\Delta_g \omega_0$  is bounded and  $\Delta_g \omega_0 \rightarrow 2K_p/3$  as approaching  $p$ .

Now, set  $\omega = \omega_0 + \omega_1$ , then  $\omega_1$  has to satisfy

$$\Delta_g \omega_1 = \Delta_g \omega - \Delta_g \omega_0 = K - \Delta_g \omega_0$$

on  $\mathcal{M} \setminus p$ . Let  $f = K - \Delta_g \omega_0$  be a function on  $\mathcal{M}$ . We can show that  $f$  extends to a continuous function on  $\mathcal{M}$ .

**Claim.** There is a solution  $w_1$  of  $\Delta_g w_1 = f$  unique up to an additive constant, provided that

$$\int_{\mathcal{M}} f \, d\mu_g = 0.$$

**Proof.** To prove this, we integrate  $f$  on  $\mathcal{M} \setminus B_\delta(p)$  with  $0 < \delta \leq \epsilon$ , i.e.,

$$-\int_{\mathcal{M} \setminus B_\delta(p)} \Delta_g w_0 \, d\mu_g = \int_{\partial B_\delta(p)} \nabla_N w_0 \, ds,$$

where  $ds$  is the element of arc length of  $\partial B_\delta(p)$ . In  $\overline{B}_\delta(p)$  we have, in polar coordinates,  $w_0 = -2 \log r$  and  $\nabla_N = \partial/\partial r$ , so  $\nabla_N w_0 = -2/r$ . Moreover, it is  $ds = R \, d\theta$ . So, we have

$$\int_{\partial B_\delta(p)} \nabla_N w_0 \, ds = -\frac{2}{\delta} \int_0^{2\pi} R(\delta, \theta) \, d\theta \rightarrow -4\pi \text{ as } \delta \rightarrow 0.$$

On the other hand,

$$\lim_{\delta \rightarrow 0} \int_{\mathcal{M} \setminus B_\delta(p)} K \, d\mu_g = \int_{\mathcal{M}} K \, d\mu_g = 4\pi$$

by [Gauss-Bonnet](#). We conclude that indeed  $\int_{\mathcal{M}} f \, d\mu_g = 0$ . ⊗

So, the equation is solvable for  $w_1$ . In fact, we can show that  $w_1$  is bounded on  $\mathcal{M}$  and is in fact continuous.<sup>a</sup>

3. Step 3 is also trivial.
4. Now, it is  $u = w + v$  where  $w = w_0 + w_1$ , and  $w_1$  is bounded on  $\mathcal{M}$ , while  $w_0 = -2\eta \log d_p$  and  $d_p$  is the  $g$ -distance from  $p$ . On the other hand,

$$e^v = \frac{1}{1 + \tilde{d}_o^2/4},$$

where  $\tilde{d}_o$  is the  $\tilde{g}$ -distance from  $o$ . Hence,

$$v = -2 \log \tilde{d}_o + O(1).$$

5. It follows that  $u$  is bounded on  $\mathcal{M}$  if and only if in  $B_{\epsilon(p)}$  (relative to  $g$ )  $d_p \cdot \tilde{d}_o$  is bounded above and below by positive constants.<sup>b</sup>

■

<sup>a</sup>For,  $f = \Delta_g w_1$  being bounded, in particular  $f \in L^2(\mathcal{M})$  implies  $w_1 \in H_2(\mathcal{M})$ , hence  $w_1$  is bounded.

<sup>b</sup>For detail, see the note.

## 7.1.2 Yamabe Problem

From the proof of [uniformization theorem](#), the following problem arises.

**Problem 7.1.1 (Yamabe problem).** Given a compact [Riemannian manifold](#)  $\mathcal{M}$ ,  $g$  of dimension  $n \geq 3$ . Find a [metric](#)  $\tilde{g}$  [conformal](#) to  $g$  such that the [scalar curvature](#) of  $\tilde{g}$  is constant.

If  $\mathcal{M}$  has no boundary, then Aubin [[Aub76b](#); [Aub76a](#)], Schoen [[Sch84](#)], Trudinger [[Tru68](#)] solves it.

With boundary, Escobar [Esc92]. If we write  $g = u^{\frac{4}{n-2}} g_0$ , the scalar curvature  $R_g$  is

$$R_g = u^{-\frac{n+2}{n-2}} \left( -\frac{4(n-1)}{n-2} \Delta_{g_0} u + R_{g_0} u \right). \quad (7.1)$$

$g$  has constant scalar curvature  $c$  if and only if  $u$  is a solution of the Yamabe equation

$$\frac{4(n-1)}{n-2} \Delta_{g_0} u - R_{g_0} u + c u^{\frac{n+2}{n-2}} = 0. \quad (7.2)$$

To solve this, consider the variational approach, where we have the following.

**Definition 7.1.2** (Einstein-Hilbert action). The *Einstein-Hilbert action*  $\mathcal{E}(g)$  is defined as

$$\mathcal{E}(g) = \frac{\int_{\mathcal{M}} R_g \, d\text{vol}_g}{\text{vol}(\mathcal{M}, g)^{\frac{n-2}{n}}}.$$

**Definition 7.1.3** (Einstein metric). A metric  $g$  is called an *Einstein metric* if  $\text{Ric}(g) = c \cdot g$  for some constant  $c$ .

**Remark.** A metric  $g$  is a critical point of  $\mathcal{E}$  if and only if  $g$  is an Einstein metric.

Given any positive function  $u$ , consider the Yamabe functional

$$\mathcal{E}_{g_0}(u) = \mathcal{E}(u^{\frac{4}{n-2}} g_0).$$

Equation 7.1 implies

$$\mathcal{E}_{g_0}(u) = \frac{\int_{\mathcal{M}} \left( \frac{4(n-1)}{n-2} |du_{g_0}|^2 + R_{g_0} u^2 \right) \, d\text{vol}_{g_0}}{\left( \int_{\mathcal{M}} u^{\frac{2n}{n-2}} \, d\text{vol}_{g_0} \right)^{\frac{n-2}{n}}}.$$

$u$  is a critical point if and only if  $u$  satisfies the Yamabe equation.

## 7.2 Lorentzian Manifolds and General Relativity

Consider

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}.$$

**Remark.** We have  $g(X, X) = -(X^0)^2 + \sum_{i=1}^3 (X_i)^2$ .

**Definition 7.2.1** (Arc length). The *arc length* of causal curve  $\gamma$  between 2 points corresponding to parameter values  $\lambda = a$  and  $\lambda = b$  is

$$L[\gamma](a, b) = \int_a^b \sqrt{-g(\dot{\gamma}(\lambda), \dot{\gamma}(\lambda))} \, d\lambda.$$

If  $q \in \mathcal{J}^+(p)$ , define temporal distance  $q$  from  $p$  as  $\tau(q, p) = \sup L[\gamma]$  over all future-directed casual curves  $(p, q)$ .

**As previously seen.** In the Riemannian manifold case, we have Hopf-Rinow theorem.

The analogous for Lorentzian manifolds is for maximization, i.e., it holds if spacetime admits Cauchy hypersurfaces when the supremum is achieved, and metric  $C'$ , maximizing curve is a causal geodesic.

## 7.3 Ricci Flow

The basic idea of Ricci flow is to consider the [metric](#)  $g$  is changing over time, i.e., the shape of the [manifold](#) changes w.r.t.  $g(t)$ , described by an O.D.E. related to the [Ricci curvature](#) as

$$\frac{\partial g(t)}{\partial t} = -2 \operatorname{Ric}(g).$$



# Appendix

# Appendix A

## Additional Notes

### A.1 The $C^\infty(\mathcal{M})$ -Module Viewpoint of Tensor Fields

To start this section, we need some primary tools.

**Definition A.1.1 (Left module).** Suppose  $R$  is a ring with 1. A *left  $R$ -module*  $M$  consists of an Abelian group  $(M, +)$  and an operation  $\cdot : R \times M \rightarrow M$  such that for all  $r, s \in R$  and  $x, y \in M$ ,

- (a)  $r \cdot (x + y) = r \cdot x + r \cdot y$ ;
- (b)  $(r + s) \cdot x = r \cdot x + s \cdot x$ ;
- (c)  $(rs) \cdot x = r \cdot (s \cdot x)$ ;
- (d)  $1 \cdot x = x$ .

**Note.** A *right  $R$ -module*  $M$  can also be defined similarly by consider  $\cdot : M \times R \rightarrow M$ .

**Definition A.1.2 (Module).** If  $R$  is commutative, then the *left and right  $R$ -module*  $M$  are the same, and we call  $M$  a *module*.

**Intuition.** We're basically relaxing the notion of  $\mathbb{F}$ -vector space, but this time, the field  $\mathbb{F}$  is replaced by a ring  $R$ .

**Remark.** The most noticeable difference between a *module* and a vector space is that a *module* usually don't have a basis.

The reason why we introduce the notion of *module* is because of the following: we can understand *tensor field* better in the following way. Observe that  $\Gamma(T\mathcal{M}) = \{X : \text{vector fields on } \mathcal{M}\}$  is actually a  $C^\infty(\mathcal{M})$ -*module*:

**Claim.**  $\Gamma(T\mathcal{M})$  carries a natural  $C^\infty(\mathcal{M})$ -*module* structure.

**Proof.** Firstly, observe that  $C^\infty(\mathcal{M}) = ((C^\infty(\mathcal{M}), +, \cdot))$  is not a field but a ring.<sup>a</sup> Then, naturally, the  $C^\infty(\mathcal{M})$ -*module*  $(\Gamma(T\mathcal{M}), \oplus, \odot)$  where

- $\oplus : (X \oplus \tilde{X})(f) := (Xf) + \tilde{X}(f)$ ;
- $\odot : (g \odot X)(f) := g \cdot X(f)$ ,

for  $X, \tilde{X} \in \Gamma(T\mathcal{M})$ ,  $g, f \in C^\infty(\mathcal{M})$ . \*

<sup>a</sup>Since given  $f \in C^\infty(\mathcal{M})$ , we might not have  $f^{-1}$ .

**Notation.** Notice that given a **vector field**  $X: \mathcal{M} \rightarrow T\mathcal{M}$  with  $p \mapsto X(p)$ , we let

$$Xf: \mathcal{M} \rightarrow \mathbb{R}, \quad p \mapsto X(p)f.$$

This makes sense since we can't always do things globally, e.g., **Hairy ball theorem**. Specifically, we can't choose a basis  $X_1, \dots, X_d \in \Gamma(T\mathcal{M})$  for our **vector field** globally as we already know. Similarly, we can define  $\Gamma(T^*\mathcal{M})$ , i.e., the set of “convector field”<sup>1</sup> is again a  $C^\infty(\mathcal{M})$ -**module**.

**Example.** Given  $\omega \in \Gamma(T^*\mathcal{M})$  and  $X \in \Gamma(T\mathcal{M})$ ,  $\omega$  acts on  $X$  to yield smooth functions by point-wise evaluation, i.e., we define

$$(\omega(X))(p) := \omega(p)(X(p)).$$

Then, the action of  $\omega$  on  $X$  is a  $C^\infty(\mathcal{M})$ -linear map since

$$(\omega(fX))(p) = f(p)\omega(p)(X(p)) = (f\omega)(p)(X(p)) = (f\omega(X))(p)$$

for  $f \in C^\infty(\mathcal{M})$ . This suggests that we should not regard  $\omega$  just as a **section** of  $T^*\mathcal{M}$ , but also a linear mapping of  $X \in \Gamma(T\mathcal{M})$  into  $C^\infty(\mathcal{M})$ .

Then, in this view point, we have the following.

**Definition A.1.3 (Tensor field\*).** A  $(r, s)$ -*tensor field*  $T$  on a **smooth manifold**  $\mathcal{M}$  is a  $C^\infty(\mathcal{M})$  multilinear map

$$T: \underbrace{\Gamma(T^*\mathcal{M}) \times \dots \times \Gamma(T^*\mathcal{M})}_r \times \underbrace{\Gamma(T\mathcal{M}) \times \dots \times \Gamma(T\mathcal{M})}_s \rightarrow C^\infty(\mathcal{M}).$$

Comparing to **Definition 2.4.14**, this definition is more general!

**Example.** The **linear connection**  $\nabla(X, Y) \mapsto \nabla_X Y$  does not define a **tensor field**.

**Proof.** Since  $\nabla$  is only  $\mathbb{R}$ -linear in  $Y$ . ⊗

## A.2 Lie Groups and Lie Algebra

### A.2.1 Lie Groups

**Lie groups** are an important topic to study for Riemannian geometry, hence we now introduce it.

**Definition A.2.1 (Lie group).** A *Lie group* is a group  $G$  with a **differentiable structure** such that the mapping  $G \times G \rightarrow G$  given by  $(x, y) \rightarrow xy^{-1}$ ,  $x, y \in G$ , is differentiable.

**Definition (Transformation).** Let  $G$  be a **Lie group**.

**Definition A.2.2 (Left transformation).** The *translations from the left*  $L_x: G \rightarrow G$  is defined as  $L_x(y) = xy$ .

**Definition A.2.3 (Right transformation).** The *translations from the right*  $R_x: G \rightarrow G$  is defined as  $R_x(y) = yx$ .

**Remark.** Both  $L_x$  and  $R_x$  are **diffeomorphisms**.

In the following discussion, let  $G$  be a **Lie group**. Turns out that  $G$  admits some nice properties on **left invariant vector fields**.

<sup>1</sup>We won't define it formally, but it's defined similarly.

**Definition** (Invariant of Riemannian metric). Let  $g$  be a [Riemannian metric](#) on  $G$ .

**Definition A.2.4** (Left invariant).  $g$  is *left invariant* if

$$\langle u, v \rangle_y = \langle d(L_x)_y u, d(L_x)_y v \rangle_{L_x(y)}$$

for all  $x, y \in G$ ,  $u, v \in T_y G$ , i.e.,  $L_x$  is an [isometry](#).

**Definition A.2.5** (Right invariant).  $g$  is *right invariant* if

$$\langle u, v \rangle_y = \langle d(R_x)_y u, d(R_x)_y v \rangle_{R_x(y)}$$

for all  $x, y \in G$ ,  $u, v \in T_y G$ , i.e.,  $R_x$  is an [isometry](#).

**Definition A.2.6** (Bi-invariant).  $g$  is *bi-invariant* if it's both [right](#) and [left invariant](#).

**Definition** (Invariant of vector field). Let  $X$  be a [vector field](#) on  $G$ .

**Definition A.2.7** (Left invariant).  $X$  is *left invariant* if  $dL_x X = X$  for all  $x \in G$ .

**Definition A.2.8** (Right invariant).  $X$  is *right invariant* if  $dR_x X = X$  for all  $x \in G$ .

**Definition A.2.9** (Bi-invariant).  $X$  is *bi-invariant* if it's both [right](#) and [left invariant](#).

As we mentioned, the [left invariant vector fields](#) are completely determined by their values at a single point of  $G$ , which allows us to introduce an additional structure on the [tangent space](#) to the neutral element  $e \in G$  in the following manner.

To each [vector](#)  $X_e \in T_e G$ , we associate the [left invariant](#)  $X$  defined by

$$X_a := dL_a X_e, \quad a \in G.$$

## A.2.2 Lie Algebras

Let  $X, Y$  be [left invariant vector fields](#) on  $G$ . Since for each  $x \in G$  and for any differentiable function  $f$  on  $G$ ,

$$dL_x[X, Y]f = [X, Y](f \circ L_x) = X(dL_x Y)f - Y(dL_x X)f = (XY - YX)f = [X, Y]f,$$

i.e.,  $[X, Y]$  is again a [left invariant vector field](#) if  $X, Y$  are. Now, if  $X_e, Y_e \in T_e G$ , we put  $[X_e, Y_e] = [X, Y]_e$ .

**Definition A.2.10** (Lie algebra). Given a [Lie group](#)  $G$ , the *Lie algebra*  $\mathfrak{g}$  is the vector space  $T_e G$  with the [bracket](#)  $[\cdot, \cdot]$ .

**Note.** The elements in the [Lie algebra](#)  $\mathfrak{g}$  will be thought of either as [vectors](#) in  $T_e G$  or as [left invariant vector fields](#) on  $G$ .

To introduce a [left invariant metric](#) on  $g$ , take any arbitrary inner product  $\langle \cdot, \cdot \rangle_e$  on  $\mathfrak{g}$  and define

$$\langle u, v \rangle_x := \langle (dL_{x^{-1}})_x(u), (dL_{x^{-1}})_x(v) \rangle_e \quad (\text{A.1})$$

for  $x \in G$ ,  $u, v \in T_x G$ . Since  $L_x$  depends differentiably on  $x$ , this is actually a [Riemannian metric](#), which is clearly [left invariant](#).

**Remark.** We can also construct a [right invariant metric](#) on  $G$ , and if  $G$  is compact,  $G$  possesses a [bi-invariant metric](#).

One important characterization for  $G$  having a [bi-invariant metric](#) is that the inner product that the [metric](#) determines on  $\mathfrak{g}$  satisfies the following relation.

**Proposition A.2.1.** If  $G$  has a [bi-invariant metric](#), then for any  $U, V, X \in \mathfrak{g}$ , the inner product that the [metric](#) determines on  $\mathfrak{g}$  satisfies

$$\langle [U, X], V \rangle = -\langle U, [V, X] \rangle.$$

**Proof.** See do Carmo [FC13, Page 40, 41]. ■

The important point about this relation is that it characterizes the [bi-invariant metrics](#) of  $G$  in the following sense.

**Remark.** If a positive bilinear form  $\langle \cdot, \cdot \rangle_e$  defined on  $\mathfrak{g}$  satisfies this relation, then the [Riemannian metrics](#) defined on  $G$  by [Equation A.1](#) is [bi-invariant](#).

### A.2.3 Lie Subalgebra

Consider  $(h_t^X)$  be a [local 1-parameter group](#) for a [vector field](#)  $X$ , and let  $\Gamma(TM)$  still denotes the set of all [vector fields](#), but now view it as just an  $\mathbb{R}$ -vector space. Then, we revise [Definition A.2.10](#) as follows.

**Definition A.2.11** (Lie algebra\*). Let  $\mathcal{M}$  be a [smooth manifold](#), the  $(\Gamma(TM), [\cdot, \cdot])$  is the *Lie algebra*.

This induces the following.

**Definition A.2.12** (Lie subalgebra). Let  $X_1, \dots, X_n$  be  $n$  [vector fields](#) on  $\mathcal{M}$  such that for all  $i, j$ ,

$$[X_i, X_j] = C_{ij}^k X_k$$

for  $C_{ij}^k \in \mathbb{R}$ . Then,  $L := (\text{span}_{\mathbb{R}}(\{X_1, \dots, X_n\}), [\cdot, \cdot])$  is called a *Lie subalgebra*.

**Notation** (Structure constant).  $C_{ij}^k$  in [Definition A.2.12](#) are called *structure constants*.

**Example.** On  $S^2$ , given  $[X_1, X_2] = X_3$ ,  $[X_2, X_3] = X_1$ ,  $[X_3, X_1] = X_2$ , we have

$$(\text{span}_{\mathbb{R}}(\{X_1, X_2, X_3\}), [\cdot, \cdot]) = \mathfrak{so}(3).$$

**Definition A.2.13** (Symmetry). A finite-dimensional [Lie subalgebra](#)  $(L, [\cdot, \cdot])$  is said to be a *symmetry* of a [metric tensor field](#)  $g$  if for every  $X \in L$  and  $t \in \mathbb{R}$ ,

$$g((h_t^X)_*(A), (h_t^X)_*(B)) = g(A, B).$$

This means that  $(h_t^X)_*$  defines an [isometry](#).

**Note.** Or equivalently,  $(h_t^X)^*g = g$  where for  $\varphi: \mathcal{M} \rightarrow \mathcal{M}$ ,

$$(\varphi^*g)(X, Y) := g(\varphi_*(X), \varphi_*(Y)).$$

### A.2.4 Lie Derivatives

Observe that for all  $X \in L$  with the corresponding [local 1-parameter group](#)  $(h_t^X)$ , if

$$\mathcal{L}_X := \lim_{t \rightarrow \infty} \frac{(h_t^X)^*g - g}{t} = 0,$$

then  $L$  is a [symmetry](#) of  $g$ .

**Definition A.2.14** (Lie derivative). The *Lie derivative*  $\mathcal{L}$  on a [smooth manifold](#)  $\mathcal{M}$  sends a pair of a [vector field](#)  $X$  and a [\( \$p, q\$ \)-tensor field](#) to a [\( \$p, q\$ \)-tensor field](#) such that

- (a)  $\mathcal{L}_X f = Xf$ ;
- (b)  $\mathcal{L}_X Y = [X, Y]$ ;
- (c)  $\mathcal{L}_X (T + S) = \mathcal{L}_X T + \mathcal{L}_X S$ ;
- (d)  $\mathcal{L}_X (T(\omega, Y)) = (\mathcal{L}_X T)(\omega, Y) + T(\mathcal{L}_X \omega, Y) + T(\omega, \mathcal{L}_X Y)$ , similarly for any other valence of  $T$ ;
- (e)  $\mathcal{L}_{X+Y} T = \mathcal{L}_X T + \mathcal{L}_Y T$ .

**Remark.**  $\nabla_X$  is  $C^\infty(\mathcal{M})$ -linear in the lower slot, while  $\mathcal{L}_X$  is not.

**Intuition.** Study neighboring [fibers](#) using a [local 1-parameter group](#) of [diffeomorphisms](#)  $(\psi_t)_{t \in I}$ .

# Bibliography

- [Abb+16] B. P. Abbott et al. “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: *Phys. Rev. Lett.* 116 (6 Feb. 2016), p. 061102. DOI: [10.1103/PhysRevLett.116.061102](https://doi.org/10.1103/PhysRevLett.116.061102). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.116.061102>.
- [AM96] Uwe Abresch and Wolfgang Meyer. “A sphere theorem with a pinching constant below  $1/4$ ”. In: *Journal of Differential Geometry* 44 (1996), pp. 214–261.
- [Aub76a] Thierry Aubin. “Equations différentielles non lineaires et probleme de Yamabe concernant la courbure scalaire”. In: *Journal de Mathématiques Pures et Appliquées* 55 (1976), pp. 269–296.
- [Aub76b] Thierry Aubin. “Problèmes isopérimétriques et espaces de Sobolev”. In: *Journal of Differential Geometry* 11 (1976), pp. 573–598.
- [Ber60] Marcel Berger. “Les variétés Riemanniennes  $(1/4)$ -pincées”. In: *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze* 14.2 (1960), pp. 161–170.
- [Ber83] Marcel Berger. “Sur les variétés riemanniennes pincées juste au-dessous de  $1/4$ ”. fr. In: *Annales de l’Institut Fourier* 33.2 (1983), pp. 135–150. DOI: [10.5802/aif.920](https://doi.org/10.5802/aif.920). URL: <http://www.numdam.org/articles/10.5802/aif.920/>.
- [BS08] S. Brendle and R. M. Schoen. *Manifolds with  $1/4$ -pinched Curvature are Space Forms*. 2008. arXiv: [0705.0766](https://arxiv.org/abs/0705.0766) [math.DG].
- [Esc92] José F Escobar. “The Yamabe problem on manifolds with boundary”. In: *Journal of Differential Geometry* 35.1 (1992), pp. 21–84.
- [FC13] F. Flaherty and M.P. do Carmo. *Riemannian Geometry*. Mathematics: Theory & Applications. Birkhäuser Boston, 2013. ISBN: 9780817634902. URL: <https://books.google.com/books?id=ct91XCWkWEUC>.
- [Kli59] W. Klingenberg. “Contributions to Riemannian Geometry in the Large”. In: *Annals of Mathematics* 69.3 (1959), pp. 654–666. ISSN: 0003486X. URL: <http://www.jstor.org/stable/1970029> (visited on 05/09/2023).
- [Kli61] Wilhelm Klingenberg. “Über Riemannsche Mannigfaltigkeiten mit positiver Krümmung.” In: *Commentarii mathematici Helvetici* 35 (1961), pp. 47–54. URL: <http://eudml.org/doc/139206>.
- [Rad25] T Radó. “Ober den Begriff der Riemannschen Fldche”. In: *Acta Univ. Szeged. (II) vol 2* (1925).
- [Rau51] H. E. Rauch. “A Contribution to Differential Geometry in the Large”. In: *Annals of Mathematics* 54.1 (1951), pp. 38–55. ISSN: 0003486X. URL: <http://www.jstor.org/stable/1969309> (visited on 05/09/2023).
- [Sch15] Frederic P Schuller. *International Winter School on Gravity and Light 2015*. Youtube. 2015. URL: [https://www.youtube.com/playlist?list=PLFeEvEPtX\\_OS6vxxiiNPrJbLu9aK1UVC\\_](https://www.youtube.com/playlist?list=PLFeEvEPtX_OS6vxxiiNPrJbLu9aK1UVC_).
- [Sch84] Richard M. Schoen. “Conformal deformation of a Riemannian metric to constant scalar curvature”. In: *Journal of Differential Geometry* 20 (1984), pp. 479–495.
- [Tru68] Neil S. Trudinger. “Remarks concerning the conformal deformation of riemannian structures on compact manifolds”. en. In: *Annali della Scuola Normale Superiore di Pisa - Scienze Fisiche e Matematiche Ser. 3*, 22.2 (1968), pp. 265–274. URL: [http://www.numdam.org/item/ASNSP\\_1968\\_3\\_22\\_2\\_265\\_0/](http://www.numdam.org/item/ASNSP_1968_3_22_2_265_0/).