# Geometry of Manifolds I

*Based on lectures by Prof. Bill Minicozzi*

Robert Koirala

Last updated: January 7, 2023

# Contents

# Preface

These (evolving) notes grew out of a course (18.965) on the geometry of manifolds I took with Prof. Bill Minicozzi in the Fall of 2020 at MIT. They started more or less as transcripts of Bill speaking in the class. Over time, I have taken polishing the notes as an opportunity to review topics and take breaks from reading research papers. I have done nothing more than add footnotes and partly organize the presentation of the content.

**Words about the course:** We developed Riemannian geometry covering material in Chapters 1 to 7, 12, and 13 in [dC92]. Then we covered some comparison geometry: eigenvalue comparison and the Laplacian comparison. The course assumed some familiarity with the geometry of manifolds at the level of 18.950 at MIT or at the level of [dC16]. However, we will do a rapid review of the basics. Some extra references for the courses were: [SY94] and [CM11].

**Words about the content:** There aren't many philosophies to study Riemannian manifolds. Some of them are to understand:

1. the space of submanifolds (subspaces) in its own right: for instance, paths and geodesics,

2. the space of functions and more generally the space of geometric objects on manifolds along with the space of (elliptic) operators acting on these spaces: for instance, harmonic functions, and

3. the space of functions into manifolds from simpler spaces: for instance the space of functions from $n$-spheres and $n$-*simplices*.

In these notes, we will mostly focus on the first two since the third philosophy leads to algebraic topology which is out of the scope of these notes although we will scratch the fundamental groups of manifolds here and there.

In Chapters IV and VI-IX we will carry out our study based on the first philosophy, focusing mainly on the space of one-dimensional submanifolds. In the last two chapters, we will scratch the study of the Laplacian $\Delta$ acting on the space of functions, focusing on its first non-zero eigenvalue as well as the space of harmonic functions (kernel of $\Delta$).

The rest of the chapters can be viewed as setting up the language of Riemannian manifolds to achieve the aforementioned goals.

A more detailed description of the chapters is as follows:

**Chapter I**: We will review the concepts of manifolds vector fields, lie brackets, and flows.

**Chapter II**: We will define Riemannian metric and tensors. We will jot down tensor contraction, a way to get tensors or lower rank. With the help of a metric, one can change the rank of a tensor by raising and lowering indices.

**Chapter III**: We will make sense of taking derivatives of vector fields and define a notion of parallel vectors. Based on the notion of derivatives, we will define the Laplacian and the Hessian operator.

**Chapter IV**: We will define a notion of acceleration-free paths on manifolds. Then using geodesics, we show that a neighborhood of a point on the manifold is diffeomorphic to a neighborhood of the origin in tangent space. Then we will study metrics in polar coordinates. Finally, we will see that the Riemannian manifold can be viewed as a metric space where the distance between two points is the infimum of the distance of curves joining the points.

**Chapter V**: We will give an algebraic definition of the Riemannian curvature tensor as well as Ricci curvature and sectional curvature.

**Chapter VI**: We will focus on extrinsic quantities of the manifold by viewing it as a submanifold of larger space. It allows us to compute the (intrinsic) curvature. Then we will scratch the study of minimal submanifolds although we won't show that the definition is justified by the fact that they actually minimize the volume.

**Chapter VII**: In the next three chapters, we will discuss the space of paths on $M$. In this chapter, we will study the space of geodesics and make sense of a tangent at a geodesic. Further, we characterize spaces where geodesics can be extended forever. In any case the goal is to infer the (topological and geometric) properties of the underlying manifold by studying the space of paths.

**Chapter VIII**: We will formalize the idea that geodesics are length-minimizing by showing that they are critical points of the length functional. We will compute the first derivative and the second derivative of the length. As an application, we will show that complete manifolds with positive Ricci curvature are compact and have a finite fundamental group.

**Chapter IX**: The second derivative of the length gives us a notion of Hessian. To understand more about the critical points of the length, we will study the linear algebraic properties (eigenvalues, number of positive/negative eigenvalues, etc) of the Hessian.

**Chapter X**: We will compare quantities on manifolds with Euclidean space: the Laplacian of the distance function and eigenvalues of the Laplacian. Further, we will see that the smoothness of the distance function on $M$ does not come for free as in Euclidean space. In some sense, the points where the distance function fails to be smooth capture the topology of $M$.

**Chapter XI**: Without worrying too much about the existence of a solution to the Laplace equation, we will study the uniqueness (follows from the maximum principle) and estimates of the derivatives of the solutions.

**Apology:** Since these notes were not written from the perspective of reference material, I apologize for failing to provide extensive bibliography. Further, I have decided not to include any pictures due to my laziness, but I encourage the reader to draw them while reading the proofs. I also apologize for the mistakes (which are due to the writer and not the instructor). To warn you of a logical hole, possible mistakes, or the fact that some paragraphs are not well written, I have included a road sign:

Please feel free to email me at rkoirala(at)ucsd(dot)edu if you have any comments or want to report mistakes.

**Acknowledgements:** I am grateful to Bill for teaching a wonderful class. My fondness for Riemannian geometry is partly because of this class. Thanks, Bill for sharing your stories during the breaks you offered in our hybrid classes. The first few years of the pandemic were hard for all of us.

I would like to thank dad, mom, and sister for listening to me talk about math every time I have a call with them halfway across the globe. Dad, mom, and sister, as I have said in our calls, I don't study anything that I can't explain to you in minutes no matter how simple the concepts sound.

Robert
La Jolla, CA

# Notation

### Linear algebra

| | |
|---|---|
| $[\bullet, \bullet]$ | Lie bracket |
| $\dim V$ | Dimension of a vector space $V$ |
| $\operatorname{ind} A$ | Index of $A$ |
| $\ker A$ | Kernel of $A$ |
| $\oplus$ | Direct sum |
| $\otimes$ | Tensor product |
| $\operatorname{Tr} A$ | Trace of $A$ |
| $V^*$ | The dual space of a vector space $V$ |

### General

| | |
|---|---|
| $(x^1, \dots, x^n)$ | Coordinates at $p$ |
| $\langle \bullet, \bullet \rangle$ | Inner product induced by the metric |
| $\mathrm{Cut}(p)$ | Cut locus at $p$ |
| $\Delta$ | Laplacian–Beltrami operator |
| $\mathrm{diam}(M)$ | Diameter of $M$ |
| $\boldsymbol{\nabla}\cdot$ | Divergence |
| $\Gamma_{ij}^k$ | Christoffel symbols |
| $\boldsymbol{\nabla}$ | Gradient |
| $\mathrm{Hess}_u$ | Hessian of $u : M \to \mathbb{R}$ |
| $\kappa$ | Sectional curvature |
| $\mathfrak{X}(M)$ | The space of vector fields on $M$ |

| | |
|---|---|
| $\pi_1(M)$ | Fundamental group of $M$ |
| Ric | Ricci curvature |
| $\varkappa_p$ | Coordinate map at $p$ |
| $A$ | Second fundamental form |
| $g$ | Riemannian metric |
| $H$ | Mean curvature |
| $M$ | Riemannian manifold |
| $R$ | Riemannian curvature tensor |
| $S$ | Scalar curvature |
| $T_pM$ | Tangent space at $p$ |

**Sets, Functions and Spaces**

| | |
|---|---|
| $\bar{A}$ | Closure of the set $A$ |
| $\delta_{ij}$ | The Kronecker $\delta$ |
| $\mathbb{H}^n$ | $n$-dimensional upper half plane |
| $\mathbb{S}^n$ | $n$-dimensional sphere |
| $\mathbb{R}^n$ | $n$-dimensional real vector space |
| $A \Subset B$ | $A$ is a compact subset of $B$ |
| $A^\circ$ | Interior of the set $A$ |
| $B_r(x)$ | Ball of radius $r$ centered at $x$ |
| $C(M;N)$ | Space of continuous functions from $M$ to $N$ |
| $C^\infty(M;N)$ | Space of smooth functions from $M$ to $N$ |
| $C^k(M;N)$ | Space of $k$-differentiable functions from $M$ to $N$ |

# Chapter 1

# Review of manifolds

The Euclidean space $\mathbb{R}^n$ with the Euclidean metric is a mathematical abstraction of space around us, but it fails to capture the essence of curved objects. However, if we look locally, the curved objects look flat, for instance, the earth.[1] A mathematical object that captures the essence of locally smooth curved space is a smooth manifold, which is a space locally modeled on $\mathbb{R}^n$. The earth looks like $\mathbb{R}^2$ locally. The goal of this chapter is to set up the basics of smooth manifolds so that we can study geometry on them. In particular, we will study length, area, volume, curvature, etc.

## 1.1 Smooth manifolds

**Definition 1.1.1.** An *n-dimensional manifold* $M$ is a second countable, Hausdorff topological space such that for each point $p \in M$ there is a homeomorphism $\varkappa_p : U_p \to V_p$ where $U_p \subset M$ is an open set containing $p$ and $V_p$ is an open set in $\mathbb{R}^n$. We say that $\varkappa_p : U_p \to V_p$ is a *chart* at $p$.[2] If $\varkappa_p(p) = 0$, we say that the chart is *centered at* $p$.

Suppose that $\varkappa_p : U_p \to V_p$ and $\varkappa'_p : U'_p \to V'_p$ are two charts at $p$. If the composition $\varkappa'_p \circ \varkappa_p^{-1}$ is a ($k$-differentiable) smooth map on $\mathbb{R}^n$, we say that the manifold is *($k$-differentiable) smooth*.[3]

**Example 1.1.2.** 1. The Euclidean space $\mathbb{R}^n$ with the identity map $\mathbb{R}^n \to \mathbb{R}^n$ is a smooth manifold.

[1]One could argue that, at a smaller scale, the earth has a rough surface. However, for the sake of simplicity, we will only study the spaces that are locally smooth. The study of spaces that are locally rough is beyond the scope of this book. Try looking for metric measure spaces, that encompass fractal spaces.

[2]We often abuse and say that $\varkappa_p$ is a chart at $p$.

[3]One can consider the collection of charts to form an *atlas*, giving a *differentiable structure* to a manifold. However, we will leave the topic for a class on differential topology. A major open problem in differential topology as of January 7, 2023 is the smooth Poincaré conjecture in four dimension.

2. The unit sphere $\mathbb{S}^n$ defined as the level set of a function $f : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ such that $x \mapsto |x|^2$ i.e. $\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid f(x) = 1\}$ is a smooth manifold.

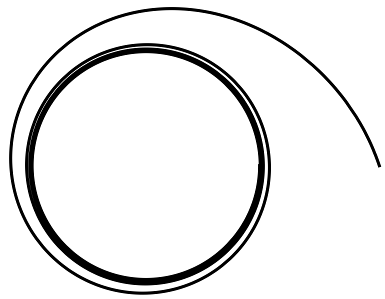

Figure 1.1: In polar coordinates $(3 - \arctan\theta, \theta)$

3. A convenient way to generate manifolds is to look at non-empty level sets at regular values of *proper*[4] functions $f : \mathbb{R}^n \to \mathbb{R}$.[5] Here $c$ is a regular value of $f$ is $\{\boldsymbol{\nabla} f = 0\} \cap \{f = c\} = \emptyset$, i.e. there are no critical points in $\{f = c\}$.

**Definition 1.1.3.** A map $f : M \to N$ between two manifold is *smooth* if for any chart $\varkappa_p$ at $p \in M$ and $\varkappa_{f(p)}$ at $f(p) \in N$ the map $\varkappa_{f(p)} \circ f \circ \varkappa_p^{-1}$ is smooth as a map between subsets of Euclidean spaces.

Going forward we will write $C^\infty(M; N)$ to mean the algebra of smooth functions $f : M \to N$. When $N = \mathbb{R}$, we write $C^\infty(M)$.

**Definition 1.1.4.** A map $f : M \to N$ is a *diffeomorphism* if $f$ is bijective and $f^{-1} \in C^\infty(N; M)$.

## 1.2 Tangent spaces

We rushed to define manifold to study curved spaces, but curved spaces are harder to study in practice. As is the general philosophy in other branches of math, we linearize (flatten) manifolds up to the first order via tangent spaces. We will later see in Chapter 4 precisely the amount of information the linear space encodes about the manifold.

Thinking of a smooth manifold $M$, via the Whitney embedding, as a subspace of $\mathbb{R}^n$, we can define the tangent space $T_pM$ at each point $p \in M$ by considering parametrized paths

[4] A map is proper if the preimages of compact sets are compact. We want properness so that we can avoid pathological extrinsic properties. Consider the level sets that look the one in Figure 1.1. Note that at a point in the limiting circle, there is no chart when viewed as a subset of $\mathbb{R}^2$.

[5] It might seem that the theory of smooth manifold fails to produce anything new than the subspaces of $\mathbb{R}^n$. In fact, the Whitney embedding theorem implies that any smooth manifold can be embedded in a Euclidean space. However, the embedding forces us to think about the ambient space of $M$. Meanwhile, it often helps to think of $M$ as an abstract object rather than a subspace of $\mathbb{R}^n$, especially, when it comes to defining intrinsic quantities, which are independent of the ambient space.

passing through $p$. In fact, say $\gamma : [0,1] \to M \subset \mathbb{R}^n$ is a curve such that $\gamma' \neq 0$ and $\gamma(0) = p$. Then the $n$ dimensional linear space spanned by $\gamma'(0)$ for all such $\gamma$ should be the tangent space of $p$. Although intuitive as it is, the construction of tangent space depends on the ambient space and embedding. Before giving an abstract definition of tangent spaces in terms of directional derivatives, which basically capture $\gamma'(0)$, let's make the intuition precise at least when $M$ can be realized as graphs of functions.

Consider a smooth function $z = u(x,y) : \mathbb{R}^2 \to \mathbb{R}$ for simplicity. Note that the graph $G \coloneqq \{(x,y,z) : z = u(x,y)\}$ defines a manifold with projection onto the $xy$-plane as charts. Note that a normal vector at each point on $G$ is given by

$$\frac{(-u_x, -u_y, 1)}{\sqrt{1 + |\boldsymbol{\nabla} u|^2}}. \tag{1.2.1}$$

Therefore, $T_pG$ is the subspace of $\mathbb{R}^3$ whose normal vector is given by (1.2.1). Concretely, consider the parametrized paths of the form $(t, y_0, u(t, y_0))$ where we fix $y_0$ and vary $t$. Differentiating with respect to $t$, we see that $(1, 0, u_x)$ gives tangent vectors. Similarly, $(0, 1, u_y)$ gives tangent vectors. Note that these vectors are linearly independent, so the tangent space is two-dimensional. Check that note that the vectors are perpendicular to the normal (1.2.1).

**Definition 1.2.1.** A *tangent* $v \in T_pM$ is a map $C^\infty(M) \to \mathbb{R}$ such that for any $f, g \in C^\infty(M)$ it satisfies:

- $\mathbb{R}$-linearity: $v(af + bg) = av(f) + bv(g)$ for any $a, b \in \mathbb{R}$,
- Leibniz rule: $v(fg) = f(p)v(g) + g(p)v(f)$.

**Remark 1.2.2.**

1. Note that the Leibniz rule implies that $v(1) = 0$ which combined with $\mathbb{R}$-linearity implies that $v(c) = 0$ for any constant function $c$.
2. $T_pM$ is non-empty. In fact, the coordinate charts generate tangent spaces. Consider a coordinate chart $\varkappa_p = (x^1, \ldots, x^n)$. Define the action of $\partial_{x^i} \coloneqq \frac{\partial}{\partial x^i}$ on $C^\infty(M)$ as

$$\partial_{x^i}(f) \coloneqq \partial_{x^i}(f \circ \varkappa_p^{-1})\Big|_{\varkappa_p}. \tag{1.2.2}$$

where the right-hand side is taken as a partial derivative of a function $f \circ \varkappa_p^{-1} : \mathbb{R}^n \to \mathbb{R}$ evaluated at $\varkappa_p$. Check that the action in (1.2.2) is $\mathbb{R}$-linear and satisfies the Leibniz rule.

**Proposition 1.2.3.** *$T_pM$ is an $n$-dimensional $\mathbb{R}$-vector space.*

*Proof.* Write $\varkappa_p = (x^1, \ldots, x^n)$. Define a bump function $\eta$ by

$$\eta = \begin{cases} 1 \text{ in } B_r(p), \\ 0 \text{ outside } B_{2r}(p) \end{cases}$$

where $B_r(p)$ is a ball of radius $r$ centered at $p$. Then $\eta x^i \in C^\infty(M)$. Since $\partial_{x^i}$ is a local operator, we see that $\partial_{x^j}(\eta x^i) = \delta^i_j$. In particular, $\dim T_pM \geq n$ since there are $n$ linearly independent coordinate vectors $\partial_{x^i}$.

To prove the other direction, we have to prove that any element $v \in T_pM$ is a linear combination of $\partial_{x^i}$. We prove that by looking at the action of $v$ on any $f \in C^\infty(M)$. Note that, by multiplying a bump function, we can consider a chart $\varkappa_p$ centered at $p$ to be an element in $C^\infty(M;\mathbb{R}^n)$. Then for any $f \in C^\infty(M)$, we have the first order Taylor expansion:

$$f \circ \varkappa_p^{-1}(x) = f(0) + \sum_{i=1}^n x^i \int_0^1 \frac{\partial(f \circ \varkappa_p^{-1})}{\partial x^i}(tx)dt =: f(0) + \sum_{i=1}^n x^i F_i \tag{1.2.3}$$

where $F_i \coloneqq \int_0^1 \frac{\partial(f\circ\varkappa_p^{-1})}{\partial x^i}(tx)dt$. To see (1.2.3), integrate the following expression with respect to $t$ from 0 to 1 :

$$\frac{\mathrm{d}}{\mathrm{d}t} f \circ \varkappa_p^{-1}(tx) = \sum_{i=1}^n x^i \frac{\partial(f \circ \varkappa_p^{-1})}{\partial x^i}(tx).$$

To this end, apply $v$ on both sides of (1.2.3). Using $v(c) = 0$ for constants, $\mathbb{R}$-linearity and Leibniz, we get

$$\begin{aligned} v(f) = \sum_{i=1}^n v(x^i F_i) &= \sum_{i=1}^n x^i(\varkappa_p(p))v(F_i) + v(x^i)F_i(\varkappa_p(p)) \\ &= \sum_{i=1}^n v(x^i) \frac{\partial(f \circ \varkappa_p^{-1})}{\partial x^i}(0) \\ &= \sum_{i=1}^n v(x^i)\partial_{x^i}(f). \end{aligned}$$

Therefore, $v = \sum_{i=1}^n v(x^i)\partial_{x^i}$. $\square$

**Proposition 1.2.4 (Change of basis).** [6] *Suppose $\varkappa_p \coloneqq (x^1, \ldots, x^n)$ and $\varkappa'_p \coloneqq (y^1, \ldots, y^n)$ are two coordinate charts centered at $p$. Then*

$$\partial_{y^j} = \sum_{i=1}^n \partial_{y^j}(x^i)\partial_{x^i}. \tag{1.2.4}$$

**Remark 1.2.5.** The equation (1.2.4) should be interpreted with caution. Suppose $F : U_p \to U'_p$ be a diffeomorphism that gives rise to the change of coordinates between $\varkappa_p$ and $\varkappa'_p$. Then $x^i$ in (1.2.4) should have been written as $x^i \circ F^{-1}$ to think of it as a function on $U'_p$ so that $\partial_{y^j}$ can act on it.

[6] Physicists define tangent vectors equivalently in terms of coordinates and require them to satisfy the change of basis rule (1.2.4) under a coordinate transformation. See [DFN84] for this approach.

Oftentimes, we are interested in how tangent spaces interact with smooth maps between manifolds.

**Definition 1.2.6.** Suppose $\phi \in C^\infty(M;N)$. Then the *differential* of $\phi$ is the induced map $d\phi_p : T_pM \to T_{\phi(p)}N$ such that for every $f \in C^\infty(N)$ and $v \in T_p(M)$, we have

$$d\phi_p(v)(f) := v(f \circ \phi). \tag{1.2.5}$$

Check that (1.2.5) actually defines a tangent vector in $T_{\phi(p)}(N)$ so that the map $d\phi_p$ is well defined.

## 1.3 Vector fields

Once we have set up a space, we might want to move around. A simple local way to capture our movement is by assigning tangent vectors at each point and saying that our velocity is given by the tangent vector. Vector fields give us a way of smoothly assigning tangent vectors. We will later see that each vector field gives rise to a flow.

**Definition 1.3.1.** A *smooth vector field* on $M$ is a collection of tangent vectors at each point $p \in M$ that varies smoothly with respect to $p$ i.e, in a coordinate chart $\varkappa_p := (x^1, \dots, x^i)$

$$v(x) = \sum_{i=1}^{n} v^i(x)\partial_{x^i} \tag{1.3.1}$$

such that $v^i(x) \in C^\infty_{loc}(M)$. Here *loc* in subscript emphasizes that $v^i(x)$ are smooth functions on charts.

**Remark 1.3.2.** The Definition 1.3.1 is independent of the choice of coordinate charts.

We will write $\mathfrak{X}(M)$ to denote the vector space of all vector fields. Note that $\mathfrak{X}(M)$ is also module over $C^\infty(M)$.

**Proposition 1.3.3 (Change of basis).** *Suppose $\varkappa_p := (x^1, \dots, x^n)$ and $\varkappa'_p := (y^1, \dots, y^n)$ are two charts centered at $p \in M$. Then for any $v = \sum_{i=1}^n v^i(x)\partial_{x^i} \in \mathfrak{X}(M)$ we have*

$$v^i(y) = \sum_{j=1}^{n} v^j(x)\partial_{x^j}(y^i). \tag{1.3.2}$$

**Remark 1.3.4.**

1. We should interpret (1.3.2) as in Remark 1.2.5.
2. If $v \in \mathfrak{X}(M)$ vanishes in one coordinate system, it vanishes in every coordinate system.

3. Vector fields, regardless of coordinates, are the same abstract object. However, they could look different in different coordinates. In practice, there is difficulty in recognizing when two representations represent the same vector field. Let us consider $x \in \mathbb{R}$ and a tangent vector $\partial_x : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ that maps $f \mapsto f'(x)$. At $x$, consider the chart $y := 2x + 1$. Then

$$\begin{aligned} \partial_y f &= \partial_y (f \circ y^{-1}) \\ &= \partial_y f\left(\frac{1}{2}y - \frac{1}{2}\right) \\ &= \frac{1}{2} f'\left(\frac{1}{2}y - \frac{1}{2}\right). \end{aligned}$$

**Definition 1.3.5.** A *derivation* is a map $X : C^\infty(M) \to C^\infty(M)$ such that for any $f, g \in C^\infty(M)$, we have

- $\mathbb{R}$-linearity: for any $a, b \in \mathbb{R}$, we have $X(af + bg) = aX(f) + bX(g)$, and
- Leibniz rule: $X(fg) = fX(g) + gX(f)$.

Denote $\mathfrak{D}(M)$ to be the vector space of all derivations on $M$. Note that $\mathfrak{D}(M)$ is a module over $C^\infty(M)$.

**Proposition 1.3.6.** *The spaces $\mathfrak{X}(M)$ and $\mathfrak{D}(M)$ are isomorphic as $\mathbb{R}$-vector spaces as well as $C^\infty(M)$-modules.*

*Proof.* Look at the proof of Proposition 1.2.3. □

**Remark 1.3.7.** A slightly different but equivalent perspective of a smooth vector field is to consider the *tangent bundle* $TM = \bigcup_{p \in M} T_p M$. It turns out that $TM$ is a manifold with charts given by $(x^1, \ldots, x^n, \partial_{x^1}, \ldots, \partial_{x^n})$. Then a vector field is a *smooth section* of this bundle. Although it is useful, we won't think about vector bundles. See [Tau11] for a bundle theory perspective.

## 1.4 Lie brackets

We can put a multiplicative structure in $\mathfrak{X}(M)$ defined by composition to turn it into an algebra. However, $\mathfrak{X}(M)$ is non-commutative. We can measure the non-commutativity using a commutator, Lie Bracket.

**Definition 1.4.1.** The *Lie bracket* $[v, w]$ of $v, w \in \mathfrak{X}(M)$ is an element of $\mathfrak{X}(M)$ such that for any $f \in C^\infty(M)$ such that for $f \in$

$$[v, w]f = v(w(f)) - w(v(f)). \tag{1.4.1}$$

**Remark 1.4.2.** Note that (1.4.1) forces $[v,w]\in\mathfrak{X}(M)$, so the definition makes sense. In fact, we have the Leibniz rule

$$\begin{aligned}[v,w](fg)=v(w(fg))-w(v(fg)) =&v(w(f)g+fw(g))-w(v(f)g+fv(g))\\ =&v(w(f))g+w(f)v(g)+v(f)w(g)+fv(w(g))\\ &-w(v(f))g-v(f)w(g)-w(f)v(g)-fw(v(g))\\ =&[v,w](f)g+f[v,w](g).\end{aligned}$$

**Example 1.4.3.** 1. $[\partial_{x^i},\partial_{x^j}]=0$.

2. On $\mathbb{R}$, we have $[\partial_x, x\partial_x]=\partial_x$.

In a local coordinate, say $v=\sum_{i=1}^n v^i(x)\partial_{x^i}$ and $w=\sum_{i=1}^n w^i(x)\partial_{x^i}$. Then

$$[v,w]=\sum_{i,j=1}^n\big(v^i\partial_{x^i}w^j-w^i\partial_{x^i}v^j\big)\partial_{x^j}.$$

We list some properties of the Lie bracket below:

- $\mathbb{R}$ **bilinearity:** $[\bullet,\bullet]:\mathfrak{X}(M)\times\mathfrak{X}(M)\to\mathfrak{X}(M)$ is a $\mathbb{R}$-linear map.
- **Skew symmetric:** For any $u,v\in\mathfrak{X}(M)$, $[u,v]=-[u,v]$.
- **Jacobi identity:** For any $u,v,w\in\mathfrak{X}(M)$, we have

$$[u,[v,w]]+[w,[u,v]]+[v,[w,u]]=0. \tag{1.4.2}$$

**Remark 1.4.4.** Any vector space $\mathfrak{g}$ with a skew-symmetric $\mathbb{R}$-bilinear map $[\bullet,\bullet]:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ that satisfies the Jacobi identity (1.4.2) is called a *Lie algebra.* In particular, $\mathfrak{X}(M)$ is a Lie algebra with the Lie bracket defined above.[7]

## 1.5 Flows

Vector fields give a local picture of how an object moves, however, it is often helpful to look at the global trajectories. This brings us to the discussion of flows.

Consider the vector field $v=x\partial_x$ in $\mathbb{R}$. Then we are interested in finding a parametrized curve $\gamma:[0,1]\to\mathbb{R}$ such that

$$\gamma'(t)=v(t) \tag{1.5.1a}$$
$$\gamma(0)=x_0 \tag{1.5.1b}$$

[7]Lie algebras naturally arise as tangent spaces of Lie groups (a manifold that has a group structure such that the group operators are smooth maps). For a detailed study of Lie groups and Lie algebra see [Eti22]. In our setting, $\mathfrak{X}(M)$ is the "Lie algebra" for the diffeomorphism group $\mathrm{Diff}(M)$ that consists of diffeomorphisms of $M$. There is a technical adjustment since $\mathrm{Diff}(M)$ turns out to be infinite-dimensional.

Writing $\gamma(t) = x(t)$ and $x(0) = x_0$, the data in (1.5.1a) and (1.5.1b) become $x'(t) = x$ with $x_0 = x_0$ for which the solution is $x(t) = e^t x_0$. In other words, the vector field $x\partial_x$ integrates to give a family of diffeomorphisms given by $\phi(x,t) = e^t x$.

In general, we can integrate $v \in \mathfrak{X}(M)$ to get diffeomorphism $\phi(x,t)$ such that

$$\begin{aligned} \frac{\mathrm{d}\phi}{\mathrm{d}t}(x,t) &= v(\phi(x,t)) \\ \phi(x,0) &= x \end{aligned}$$

where $\frac{\mathrm{d}\phi}{\mathrm{d}t} \coloneqq d\phi(\partial_t)$. This amounts to solving an ordinary differential equation just like in $\mathbb{R}$.

# Chapter 2

# Tensor calculus

After setting up the notion of paths that arise from vector fields, the next natural thing to study is the length of paths. A local way to compute the length of a parametrized path $\gamma : [0,t] \to M$ is to compute $\int_0^t |\gamma'(s)| ds$ where $|\gamma'(s)|$ is the speed at $\gamma(s)$. We will later see that a *metric* allows us to define a notion of speed (magnitude of tangent vectors). Once we generalize the metric, we get tensors.

## 2.1 Metric and tensors

**Definition 2.1.1.** A *metric* $g$ is a bilinear map $\mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$ that is

- symmetric: $g(v,w) = g(w,v)$,
- positive definite:[1] $g(v,v) \geq 0$ and
- non–degenerate:[2] $g(v,v) = 0$ if and only if $v = 0$.

**Remark 2.1.2.**

1. At each point $p \in M$, $g(\bullet, \bullet)$ defines an inner product on $T_p(M)$. Then we can set $\sqrt{g(v,v)}$ to be the norm $|v|$ of $v$.
2. We say that $(M, g)$ is a *Riemannian manifold.*

**Example 2.1.3.** On $\mathbb{R}^n$ with standard coordinates, $g$ is just the standard inner product.

**Example 2.1.4.** Any smooth manifold viewed as a subspace of $\mathbb{R}^n$ can be endowed with a metric induced from the Euclidean metric via the embedding.

[1] In local coordinates, $g$ looks like a matrix, and positive definiteness means that all of the eigenvalues are non-negative. One can relax positive definiteness in the definition of the metric. In special relativity, the Lorentz metric in the standard coordinate has one negative eigenvalue.

[2] In local coordinates, this means that $g$ viewed as matrix has positive eigenvalues.

When there is a vector space $V$, we can talk about its dual space $V^*$. Previously, we defined $\mathfrak{X}(M)$ as a vector space of vector fields. Then $\alpha \in \mathfrak{X}^*(M)$ is a smooth $C^\infty(M)$-linear map

$$\alpha : \mathfrak{X}(M) \to C^\infty(M).$$

**Definition 2.1.5.** An element $\alpha \in \mathfrak{X}^*(M)$ is called 1-*form.*

**Definition 2.1.6.** An $(r, s)$ *tensor* $A$ is a smooth $C^\infty(M)$-multi-linear map

$$A : \overbrace{\mathfrak{X}^*(M) \times \cdots \times \mathfrak{X}^*(M)}^{r} \times \overbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}^{s} \to C^\infty(M). \tag{2.1.1}$$

We say that the *rank of* $A$ is $(r, s)$.

**Example 2.1.7.**
1. $C^\infty(M)$ consists of $(0, 0)$ tensors.
2. Vector fields are $(1, 0)$ tensors.
3. A $(0, r)$ tensor which is *antisymmetric* is called $r$*-form.*
4. A $(0, 2)$ tensor $\mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$ is a bilinear form at each point on $M$. A metric is a special kind of $(0, 2)$ tensor.

**Remark 2.1.8.** We will write space of $(r, s)$ tensors as $\bigotimes_{i=1}^r \mathfrak{X}(M) \bigotimes_{j=1}^s \mathfrak{X}^*(M)$ where $\otimes_{i=1}^r A_i \otimes_{j=1}^s B^j \in \bigotimes_{i=1}^r \mathfrak{X}(M) \bigotimes_{j=1}^s \mathfrak{X}^*(M)$ is defined such that for any $w_i \in \mathfrak{X}(M)$ and $v^i \in \mathfrak{X}^*(M)$ we have

$$\otimes_{i=1}^r A_i \otimes_{j=1}^s B^k(v^1, \ldots, v^r, w_1, \ldots w_s) \coloneqq \prod_{i=1}^r v^i(A_i) \prod_{j=1}^s B^j(w_j). \tag{2.1.2}$$

Check that (2.1.2) actually defines a $C^\infty(M)$-multi-linear map as in (2.1.1).[3]

## 2.2 Tensors in local coordinates

In practice, we carry out computations in local coordinates, so we need expressions for metrics in charts. Fix a coordinate chart $\varkappa_p = (x^1, \ldots, x^n)$ centered at $p$. Write $dx^j$ to mean the dual of $\partial_{x^j}$, i.e. $dx^j(\partial_{x^i}) = \delta_i^j$. Check that $dx^j$ form a basis of $\mathfrak{X}^*(M)$.

**Proposition 2.2.1 (Change of basis).** *Suppose* $\varkappa_p \coloneqq (x^1, \ldots, x^n)$ *and* $\varkappa'_p \coloneqq (y^1, \ldots, y^n)$ *are two coordinate charts centered at* $p$. *Then*

$$dy^j = \sum_{i=1}^n \partial_{x^i}(y^i) dx^i. \tag{2.2.1}$$

[3] In literature, one defines the tensor product $\bigotimes_{i=1}^r \mathfrak{X}(M) \bigotimes_{j=1}^s \mathfrak{X}^*(M)$ using some equivalence relation $\sim$ on the free $C^\infty(M)$-module generated by $\prod_{i=1}^r \mathfrak{X}(M) \prod_{j=1}^s \mathfrak{X}^*(M)$. However, the universal property of tensor product guarantees that we can view tensors as $C^\infty(M)$-multi-linear maps, so there is no harm in using (2.1.2) as the definition of tensors. See Appendix A for more details.

In local coordinates, any 1-form $\alpha$ can be written as

$$\alpha = \sum_{j=1}^{n} \alpha(\partial_{x^j}) dx^j.$$

**Proposition 2.2.2 (Change of basis).** *Suppose $\varkappa_p := (x^1, \ldots, x^n)$ and $\varkappa'_p := (y^1, \ldots, y^n)$ are two coordinate charts centered at $p$. Suppose $\alpha$ is a $(0,1)$-form. Then*

$$\alpha(\partial_{y^j}) = \sum_{i=1}^{n} \alpha(\partial_{x^i}) \partial_{y^j}(x^i). \tag{2.2.2}$$

Suppose $v = \sum_{i=1}^{n} v^i \partial_{x^i}$ and $w = \sum_{j=1}^{n} w^j \partial_{x^j}$. Then by using the linearity of $g$, we see that

$$g(v, w) = g\left(\sum_i v^i \partial_{x^i}, \sum_j w^j \partial_{x^j}\right) = \sum_{ij} g(v^i \partial_{x^i}, w^j \partial_{x^j}) = \sum_{ij} v^i w^j g(\partial_{x^i}, \partial_{x^j}).$$

Writing $g(\partial_{x^i}, \partial_{x^j}) = g_{ij}$, we see that

$$g = \sum_{i,j=1}^{n} g_{ij} dx^i \otimes dx^j \tag{2.2.3}$$

where $dx^i \otimes dx^j(u, v) := dx^i(u) dx^j(v)$. We can view, $dx^i \otimes dx^j(u, v)$ as a $(0, 2)$ tensor such that $dx^i \otimes dx^j(\partial_{x^k}, \partial_{x^l}) = \delta^i_k \delta^j_l$.

One can use a similar procedure to find a coordinate representation of any $(r, s)$-tensor as a linear combination of $\bigotimes_{k=1}^{r} \partial_{x^{i_k}} \bigotimes_{l=1}^{s} dx^{k_l}$.

## 2.3 Contraction of a tensor

Assume $r, s \geq 1$. Given an $(r, s)$-tensor $A \in \bigotimes_{i=1}^{r} \mathfrak{X}(M) \bigotimes_{j=1}^{s} \mathfrak{X}^*(M)$ there is a natural way to get a lower rank tensor $A'$ by pairing some coordinates in $\bigotimes_{j=1}^{s} \mathfrak{X}^*(M)$ with some coordinates in $\bigotimes_{i=1}^{r} \mathfrak{X}(M)$. In practice, it is easier to deal with $A'$ as an $C^\infty(M)$-multi-linear map since its domain has a smaller dimension than that for $A$. The caveat is we lose some information about $A$.

**Definition 2.3.1.** A *contraction* of $(1, 1)$-tensor is a *tensor map*

$$c_1^1 : \mathfrak{X}(M) \otimes \mathfrak{X}^*(M) \to C^\infty(M)$$

given by evaluating the second entry at the first one.

In a coordinated chart $\varkappa_p = (x_1, \ldots, x_n)$, any $A \in \mathfrak{X}(M) \otimes \mathfrak{X}^*(M)$ can be written as

$$A = \sum_{i,j=1}^{n} A^i_j \partial_{x^i} \otimes dx^j \tag{2.3.1}$$

where $A^i_j = A(dx^i, \partial_{x^j})$. Therefore,

$$c^1_1(A) = \sum_{i,j=1}^{n} A^i_j dx^j(\partial_{x^i}) = \sum_{i,j=1}^{n} A^i_j \delta^j_i = \sum_{i=1}^{n} A^i_i. \tag{2.3.2}$$

**Remark 2.3.2.** When we think of $(1,1)$-tensors as matrices, then (2.3.2) implies that $c^1_1$ is just the trace map.

Let's make sure that the computation in (2.3.2) is coordinate-independent. Say that $\varkappa'_p = (y_1, \ldots, y_n)$ is another coordinate centered at $p$. Then the change of basis formulae imply that

$$\begin{aligned}
\sum_{i=1}^{n} A(dx^i, \partial_{x^i}) &= \sum_{i=1}^{n} A\left(\sum_{j=1}^{n} \partial_{y^j}(x^i) dy^j, \sum_{k=1}^{n} \partial_{x^i}(y^k)\partial_{y^k}\right) \\
&= \sum_{i,j,k=1}^{n} \partial_{y^j}(x^i)\partial_{x^i}(y^k) A(dy^j, \partial_{y^k}) \\
&= \sum_{j,k=1}^{n} \delta^k_j A(dy^j, \partial_{y^k}) \\
&= \sum_{k=1}^{n} A(dy^k, \partial_{y^k})
\end{aligned}$$

where we used the fact that $\sum_{i=1}^n \partial_{y^j}(x^i)\partial_{x^i}(y^k)$ is a matrix multiplication of matrices inverse to each other.

In general, we have the following definition of contraction:

**Definition 2.3.3.** For any $1 \leq k \leq r$ and $1 \leq l \leq s$, an $(k,l)$-*contraction* is map of tensor product spaces $c^k_l : \bigotimes_{i=1}^r \mathfrak{X}(M) \bigotimes_{j=1}^s \mathfrak{X}^*(M) \to \bigotimes_{i=1}^{r-1} \mathfrak{X}(M) \bigotimes_{j=1}^{s-1} \mathfrak{X}^*(M)$ defined by evaluating $l^{th}$ coordinate of $\bigotimes_{j=1}^s \mathfrak{X}^*(M)$ on $k^{th}$ coordinate of $\bigotimes_{i=1}^r \mathfrak{X}(M)$. In particular, for any $v_1 \otimes \cdots \otimes v_r \otimes w^1 \otimes \cdots \otimes w^s \in \bigotimes_{i=1}^r \mathfrak{X}(M) \bigotimes_{j=1}^s \mathfrak{X}^*(M)$, we have

$$c^k_l(v_1 \otimes \cdots \otimes v_r \otimes w^1 \otimes \cdots \otimes w^s) \coloneqq w^l(v_k) v_1 \otimes \cdots \otimes \hat{v}_k \otimes \cdots \otimes v_r \otimes w^1 \otimes \cdots \otimes \hat{w}^l \otimes \cdots \otimes w^s,$$

where the $\hat{}$ represents that the entry is omitted.

We can even take multi-indices $K \coloneqq (k_1, \ldots, k_t)$ and $L = (l_1, \ldots, l_t)$ with $1 \leq t \leq \min(r,s)$ and define $c^K_L$ by pairing $l^{th}_m$ coordinate of $\bigotimes_{j=1}^s \mathfrak{X}^*(M)$ with $k^{th}_m$ coordinate of $\bigotimes_{i=1}^r \mathfrak{X}$ for all $1 \leq m \leq t$.

## 2.4 Raising and lowering indices

When we fix a metric $g$, there is an identification $\mathfrak{X}(M) \cong \mathfrak{X}^*(M)$ induced by the inner product structure $\langle \bullet, \bullet \rangle$ of $g$. In particular, we can think of $X \in \mathfrak{X}(M)$ as an element of $\langle X, \bullet \rangle \in \mathfrak{X}^*(M)$. Similarly, for any $X^* \in \mathfrak{X}^*(M)$ we can find a unique $X \in \mathfrak{X}(M)$ such that $X^*(Y) = \langle X, Y \rangle$ for any $Y \in \mathfrak{X}(M)$ i.e $X^* = \langle X, \bullet \rangle$. The identification allows us to think of $(r, s)$-tensors as $(r', s')$-tensors with $r + s = r' + s'$ and $r, r', s, s' \geq 0$. Indeed, we can make sense of contraction of $(r, 0)$ or $(0, s)$ tensors.

**Notation:** Going forward, we will follow Einstein notation.

1. If there is an index, we mean to sum over all possible values of the index. For instance, $v^i a_i := \sum_i v^i a_i$ and $v^i a_j := \sum_{ij} v^i a_j$.
2. As is evident in the notation so far, we will write local coordinates in upper indices i.e. $\varkappa_p = (x^1, \ldots, x^n)$.
3. Basis elements of $\mathfrak{X}(M)$ have lower indices i.e $\partial_{x^i}$, while their coefficients have upper indices. In particular, $v \in \mathfrak{X}(M)$ can be written as $v = v^i \partial_{x^i} := \sum_{i=1}^n v^i \partial_{x^i}$.
4. Basis elements of $\mathfrak{X}^*(M)$ have upper indices i.e. $dx^i$ while their coefficients have lower indices. In particular, any $w \in \mathfrak{X}^*(M)$ can be written as $w = w_i dx^i := \sum_{i=1}^n w_i dx^i$.
5. In general, we write an $(r, s)$-tensor $R$ as $R^{i_1 \ldots i_r}_{j_1 \ldots j_s} \partial_{x^{i_1}} \otimes \cdots \otimes \partial_{x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}$ which should be understood that we are summing over all $i_k$'s and $j_l$'s. Going forward, we will just write $R = R^{i_1 \ldots i_r}_{j_1 \ldots j_s}$ by making $\partial_{x^{i_1}} \otimes \cdots \otimes \partial_{x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}$ implicit. For instance, $g = g_{ij}$.

**Definition 2.4.1.** The identification of $(r, s)$-tensor as $(r + 1, s - 1)$-tensor is called *raising of indices.*

**Definition 2.4.2.** The identification of $(r, s)$-tensor as $(r - 1, s + 1)$-tensor is called *lowering of indices.*

Let's justify why the identifications have anything to do with raising and lowering indices. To this end, consider a $(0, 2)$ tensor $B$. Then for any $X, Y \in \mathfrak{X}(M)$, $B(X, Y) \in C^\infty(M)$. Note that $B(X, \bullet) : \mathfrak{X}(M) \to C^\infty(M)$ is a linear map. Therefore, the natural identification $\mathfrak{X}(M) \cong \mathfrak{X}^*(M)$ guarantees the existence of a unique vector field, also written as $B(X, \bullet)$, such that

$$B(X, Y) = \langle B(X, \bullet), Y \rangle. \tag{2.4.1}$$

Fix a coordinate. Then $B = B_{ij} dx^i \otimes dx^j$. Further, write $B(\partial_{x^i}, \bullet) = B_i^k \partial_{x^k}$. Then (2.4.1) means that $B_{ij} = B_i^k g_{kj}$. In particular, $B_i^k = B_{ij} g^{jk}$. Note that $B_i^k$ is a $(1, 1)$-tensor which is obtained by raising an index.

**Remark 2.4.3.** 1. Lowering the index amounts to multiplying by $g_{ij}$. For instance, if $v = v^i \in \mathfrak{X}(M)$ then we get a 1-form $v^i g_{ij} \in \mathfrak{X}^*(M)$ by lowering the index.

2. Tracing out a $(0, 2)$-tensor does not make any sense. However, we can raise the index and get $(1, 1)$ tensor and then take a trace.

# Chapter 3

# Differentiation

After the points in $M$, the first class of subspace of $M$ to study are curves and their lengths. If there is a parametrization of a curve, we can think of it as the trajectory of a particle. To study the trajectory, physicists often use acceleration, the derivative of velocity. Rephrasing in geometric terms, we have to define derivatives of vector fields. Then we can make sense of the constant velocities and trajectory of a "free" particle.

A naive way to think about differentiating a vector field $X \in \mathfrak{X}(M)$ at least along a curve $c(t)$ is to look at the limit of the difference quotients

$$\left.\frac{\mathrm{d}X(c(t))}{\mathrm{d}t}\right|_{t=r} \text{“} = \text{”} \lim_{s \to r} \frac{X(c(r)) - X(c(s))}{r - s}. \tag{3.0.1}$$

The problem with (3.0.1) is that $X(c(r))$ and $X(c(s))$ live in $T_{c(r)}M$ and $T_{c(s)}M$, respectively. Therefore, comparing elements from two vector spaces has to be justified.

There are different ways to compare vectors from two different tangent spaces. The first one which we will use not emphasize is the Lie derivative. When differentiating $Y \in \mathfrak{X}(M)$ in the direction of $X \in \mathfrak{X}(M)$, we use the pushforward of flow generated by $X$ to compare vectors. The second one is to fix a *connection*, which allows us to differentiate $Y$ in the direction of $X$.

## 3.1 Covariant derivative

**Definition 3.1.1.** An *affine connection* is a map $\boldsymbol{\nabla} : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ that sends $(X, Y)$ to $\boldsymbol{\nabla}_X Y$ such that the map satisfies

- $C^\infty(M)$-linearity in the first entry i.e. for any $f \in C^\infty(M)$ and $X, Y \in \mathfrak{X}(M)$ we have

$$\boldsymbol{\nabla}_{fX} Y = f \boldsymbol{\nabla}_X Y, \qquad \boldsymbol{\nabla}_{X_1 + X_2} Y = \boldsymbol{\nabla}_{X_1} Y + \boldsymbol{\nabla}_{X_2} Y.$$

- $\mathbb{R}$-linearity in the second entry i.e. for all $a_i \in \mathbb{R}$ and $X, Y_i \in \mathfrak{X}(M)$ we have
$$\nabla_X(a_1Y_1 + a_2Y_2) = a_1\nabla_X Y_1 + a_2\nabla_X Y_2.$$
- Leibniz rule in the second entry ie. for any $X, Y \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$
$$\nabla_X(fY) = f\nabla_X Y + X(f)Y.$$

**Remark 3.1.2.** There are too many affine connections (see [Tau11]). However, there is a natural one if we require the connection to satisfy:

- **Symmetry:** $\nabla_X Y - \nabla_Y X = [X, Y]$.[1]
- **Metric compatibility:** $X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$. By Leibniz rule, we know that $X[g(Y, Z)] = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) + (\nabla_X g)(Y, Z)$. Metric compatibility means that $(\nabla_X g)$ is zero. In particular, $g$ is constant in any direction.

**Theorem 3.1.3.** *There exists a unique affine connection that is symmetric and compatible with the metric. We call it the Levi-Civita connection.*

*Proof.* Let $X, Y \in \mathfrak{X}(M)$. Since the metric is non-degenerate, we can characterize the map $(X, Y) \mapsto \nabla_X Y$ by studying $\langle \nabla_X Y, Z\rangle$ for all $Z \in \mathfrak{X}(M)$. Note that metric compatibility implies that $X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$. Then cyclically permuting $X, Y, Z$. Then using the symmetry of connection i.e. $\nabla_X Y - \nabla_Y X = [X, Y]$ in the sum $X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle$, we get
$$2\langle \nabla_X Y, Z\rangle = X\langle Y, Z\rangle + Y\langle X, Z\rangle - Z\langle X, Y\rangle - \langle Y, [X, Z]\rangle - \langle X, [Y, Z]\rangle - \langle Z, [Y, X]\rangle. \tag{3.1.1}$$

Therefore, (3.1.1) implies the uniqueness of connection.

To prove the existence, define $\nabla_X Y$ by requiring it to satisfy (3.1.1) for all $Z \in \mathfrak{X}(M)$. Then it is an exercise to prove that the right-hand side of (3.1.1) satisfies the definition of an affine connection as well as is symmetric and compatible with the metric. □

Unless otherwise stated, we will mean Levi-Civita connection whenever we say connection.[2]

## 3.2 Christoffel symbols

In this section, we will study the covariant derivatives in local coordinates, say $\varkappa_p := (x^1, \ldots, x^n)$. Since $\nabla_{\partial_{x^i}}\partial_{x^j} \in \mathfrak{X}(M)$, we can write
$$\nabla_{\partial_{x^i}}\partial_{x^j} := \sum_k \Gamma^k_{ij}\partial_{x^k}.$$

[1] This condition is called *torsion free* in literature.

[2] There is a general group theoretic approach to talk about connections on principal bundles, see for instance [KN63].

for some $\Gamma^k_{ij}$, which are called *Christoffel symbols*.

**Remark 3.2.1.** 1. Note that $\Gamma^k_{ij}$ are not tensors since they are $C^\infty(M)$-linear only in one slot.

2. The Christoffel symbols are symmetric in $i$ and $j$ indices i.e. $\Gamma^k_{ij} = \gamma^k_{ji}$.

Suppose $v = v^j \partial_{x^j}$ and $w = w^i \partial_{x^i}$. Then using the property of covariant derivate, we see that

$$\boldsymbol{\nabla}_w v = (v^j \Gamma^k_{ij} + \partial_{x^i} v^k) w^i \partial_{x^k}. \tag{3.2.1}$$

The definition of Christoffel symbols merely does not allow us to do any computation. Let's compute them in terms of metric $g$. Recall that $g_{ij} = \langle \partial_{x^i}, \partial_{x^j} \rangle$. Therefore, $\langle \boldsymbol{\nabla}_{\partial_{x^i}} \partial_{x^j}, \partial_{x^l} \rangle = \Gamma^k_{ij} g_{kl}$. Further, using (3.1.1), we see that $2\langle \boldsymbol{\nabla}_{\partial_{x^i}} \partial_{x^j}, \partial_{x^l} \rangle = \partial_{x^i} g_{jl} + \partial_{x^j} g_{li} - \partial_{x^l} g_{ij}$. Therefore,

$$2\Gamma^k_{ij} g_{kl} = \partial_{x^i} g_{jl} + \partial_{x^j} g_{li} - \partial_{x^l} g_{ij}.$$

Write $g^{lk}$ to mean the inverse of $g_{kl}$. Then multiplying $g^{lk}$ on both sides, we get

$$\Gamma^k_{ij} = \frac{1}{2} g^{lk} (\partial_{x^i} g_{jl} + \partial_{x^j} g_{li} - \partial_{x^l} g_{ij}). \tag{3.2.2}$$

**Remark 3.2.2.** 1. If $\partial_X g = 0$ (for instance in $\mathbb{R}^n$ with the standard metric) then $\Gamma^k_{ij} = 0$. In this case the covariant derivative $\boldsymbol{\nabla}_w v$ in (3.2.1) is just the directional derivative.

2. Note that $\Gamma^k_{ij}$ depends on the metric and the first derivative and is not a tensor. In fact, there is no tensor that depends just on the metric and its first derivative. The proof of the statement requires representation theory which is out of the scope of these notes. However, there is a tensor called *curvature tensor* that depends on the metric, and its first and second derivatives. We won't prove but any such tensor can be derived from the curvature tensor.

## 3.3 Parallel transport

In this section, we will make sense of when a vector field is constant (parallel). In $\mathbb{R}^n$, a vector field $X$ is constant means that $X(p)$ has the same magnitude and direction for all $p \in \mathbb{R}^n$. On a Riemannian manifold, a metric gives a notion of constant magnitude. However, we need to compare tangent vectors at different points to define direction canonically. We will make a comparison by picking a curve in a manifold and "transporting" vectors along the curve preserving the "direction".

Suppose $\gamma : [0,1] \to M$ is a parametrized curve. Say that $v$ is a vector field along $\gamma$ i.e. $v(t) \in T_{\gamma(t)} M$. Then $\boldsymbol{\nabla}_{\gamma'} \mathrm{v}$ is a way to differentiate $v$ along $\gamma$. If $v$ is a restriction of an element in $\mathfrak{X}(M)$ then $\boldsymbol{\nabla}_{\gamma'} v$ is well defined. To see that it is a well-defined object in general, let's fix a local coordinate $\varkappa_p := (x^1, \ldots, x^n)$. Then $v(t) = v^j(t)\partial_{x^j}$. By Leibniz rule and linearity of $\boldsymbol{\nabla}$, we get

$$\boldsymbol{\nabla}_{\gamma'} v = \boldsymbol{\nabla}_{\gamma'}(v^j(t)\partial_{x^j}) = v^j \boldsymbol{\nabla}_{\gamma'}\partial_{x^j} + \frac{\mathrm{d}v^j}{\mathrm{d}t}\partial_{x^j}.$$

Since the right-hand side does not depend on how we define $v^j(t)$ off of $\gamma$, the left-hand side is well defined. Further, the definition satisfies a Leibniz rule on the entry $v$, tensorial in $\gamma'$, symmetric and compatible with metric.

Note that $\boldsymbol{\nabla}_{\gamma} v = 0$ in one coordinate implies it is zero in all coordinates. This motivates a definition of transporting $v$ along $\gamma$ without changing direction.

**Definition 3.3.1.** A *parallel transport* along a curve $\gamma : [0,1] \to M$ such that $\gamma(0) = p$ is a map $P_t : T_pM \to T_{\gamma(t)}M$, where $v \in T_pM$ gets sent to $v(t) \in T_{\gamma(t)}M$, such that

$$\boldsymbol{\nabla}_{\gamma'} v = 0, \qquad v(0) = v. \tag{3.3.1}$$

**Remark 3.3.2.**

1. Fix a chart $\varkappa_p = (x^1 \ldots, x^n)$ and write $\gamma = (\gamma^1, \ldots, \gamma^n)$ and $\gamma_t := \partial_t \gamma$. Then finding a parallel transport amounts to solving the following system of equation

$$0 = v^j \gamma_t^i \boldsymbol{\nabla}_{\partial_{x^i}} \partial_{x^j} + \frac{\mathrm{d}v^j}{\mathrm{d}t}\partial_{x^j} = \big(v^j \gamma_t^i \Gamma_{ij}^k + \partial_t v^k\big)\partial_{x^k}. \tag{3.3.2}$$

   Note that (3.3.2) is a system of $n$ linear ODE on $n$ variables $v^k$. Therefore, if we specify $v^k(0)$ then we get a unique solution for $v(t)$. Therefore, we can say that $v(t)$ *the* parallel transport of $v$ along $\gamma(t)$.

2. Note that $P_t : T_pM \to T_{\gamma(t)}M$ is a linear isometry. In fact, say $v(t)$ and $w(t)$ are the parallel transport of $v$ and $w$ along $\gamma$. Then

$$\partial_t \langle v(t), w(t)\rangle = \langle \boldsymbol{\nabla}_{\gamma'} v, w\rangle + \langle v, \boldsymbol{\nabla}_{\gamma'} w\rangle = 0.$$

   In particular, the "relative direction" of two vectors under parallel transport is preserved.

3. In $\mathbb{R}^n$, the Christoffel symbols vanish. Therefore, (3.3.2) implies that $\partial_t v^k = 0$ which implies that $v$ is constant in magnitude and direction under parallel transport.

4. Although we use local coordinates to find the parallel transport of $v \in T_pM$, the solution is independent of the coordinates since we are really solving an abstract equation (3.3.1). However, the representation of the solution could depend on the coordinates.

5. To define parallel transport globally, we solve (3.3.1) in local charts and glue the solutions together. At the intersection of charts, we have the uniqueness of the solution for ODE since solutions obtained from two coordinate maps satisfy the same ODE.

**Question 3.3.3.** What is a parallel transport for a Lie group assuming that the metric is translation invariant? Does it agree with the group action?

## 3.4 Some analysis

In this section, we record the definition of some differential operators on $M$.

**Definition 3.4.1.** The *gradient* is an operator $\boldsymbol{\nabla} : C^\infty(M) \to \mathfrak{X}(M)$, $f \mapsto \boldsymbol{\nabla} f$, such that for any $X \in \mathfrak{X}(M)$, we have

$$\langle \boldsymbol{\nabla} f, X \rangle = X(f).$$

**Remark 3.4.2.**

1. The gradient operator can be thought of as the adjoint of the divergence operator.

2. The dual of the gradient operator is the map $C^\infty(M) \ni f \mapsto df$ since $X(f) = df(X)$. In local coordinates $\varkappa_p := (x^1, \ldots, x^n)$, we can write $df = \frac{\partial f}{\partial x^i} dx^i$. Write $\boldsymbol{\nabla} f = a^k \partial_{x^k}$. Then

$$\frac{\partial f}{\partial x^i} dx^i(\partial_{x^j}) = \langle \boldsymbol{\nabla} f, \partial_{x^j} \rangle = a^k g_{kj}.$$

   Therefore, $a^k = \partial_{x^j} f g^{jk}$ where $(g_{jk})$ is the inverse matrix of $(g_{jk})$. Therefore,

$$\boldsymbol{\nabla} f = (\partial_{x^j} f) g^{jk} \partial_{x^k}.$$

   In particular, when $g_{ij} = \delta_{ij}$, we get the standard gradient operator on $\mathbb{R}^n$.

**Definition 3.4.3.** The *Hessian* is an operator $\mathrm{Hess} : C^\infty(M) \to \mathfrak{X}^*(M) \otimes \mathfrak{X}^*(M)$, $f \mapsto \mathrm{Hess}_f$, such that for any $(X, Y) \in \mathfrak{X}(M) \times \mathfrak{X}(M)$, we have

$$\mathrm{Hess}_f(X, Y) = \langle \boldsymbol{\nabla}_X \boldsymbol{\nabla} f, Y \rangle.$$

**Lemma 3.4.4.** *For any $f \in C^\infty(M)$, $\mathrm{Hess}_f : \mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$ is a symmetric operator.*

*Proof.* Recall that the metric compatibility of the connection implies that for any $X, Y \in \mathfrak{X}(M)$ we have $\langle \boldsymbol{\nabla}_X \boldsymbol{\nabla} f, Y \rangle = X\langle \boldsymbol{\nabla} f, Y \rangle - \langle \boldsymbol{\nabla} f, \boldsymbol{\nabla}_X Y \rangle = X(Y(f)) - \boldsymbol{\nabla}_X Y(f)$. Therefore, using the symmetry of the connection, we see that

$$\mathrm{Hess}_f(X, Y) - \mathrm{Hess}_f(Y, X) = X(Y(f)) - (\boldsymbol{\nabla}_X Y)(f) - Y(X(f)) + (\boldsymbol{\nabla}_Y X)(f) = 0.$$

□

**Remark 3.4.5.** In local coordinates $\varkappa_p = (x^1, \ldots, x^n)$, we see that

$$\mathrm{Hess}_f(\partial_{x^i}, \partial_{x^j}) = \partial_{x^i}(\partial_{x^j} f) - \boldsymbol{\nabla}_{\partial_{x^i}} \partial_{x^j}(f) = \partial_{x^i}(\partial_{x^j} f) - \Gamma^k_{ij} \partial_{x^k} f.$$

In particular, $\mathrm{Hess}_f$ in $\mathbb{R}^n$ with standard metric is the second order differentiation of $f$. Note that local coordinate representation is symmetric on $i, j$ indices which checks out with Lemma 3.4.4.

**Definition 3.4.6.** The *Laplacian* $\Delta f$ of a function $f \in C^\infty(M)$ is the trace of $\mathrm{Hess}_f$ i.e.

$$\Delta f = \mathrm{Tr}(\mathrm{Hess}_f).$$

**Remark 3.4.7.**

1. Laplacian of a function is the divergence of the gradient of the function.

2. In local coordinates $\varkappa_p = (x^1, \ldots, x^n)$, we have

$$\Delta f = \mathrm{Hess}_f(\partial_{x^i}, \partial_{x^i}) = \partial_{x^i}\partial_{x^i} f - \Gamma^k_{ii}\partial_{x^k} f.$$

   Therefore, when $\Gamma^k_{ij} = 0$, we recover the Laplacian on $\mathbb{R}^n$.

# Chapter 4

# Geodesics

When we think of paths on $M$ as a trajectory of a particle, geodesics are the "purest" form of trajectories i.e. the particle traversing the trajectory is not acted upon by any force, so the "acceleration" of the particle is zero. In this chapter, we will make the idea of geodesic precise. On the other hand, recall that we defined $T_pM$ in §1.2 with a hope to linearize $M$. In §4.2 (cf. Remark 4.2.6), we will use geodesics to see when exactly the linear space $T_pM$ can encode some information about $M$. See Chapter 7 (cf. Theorem 7.3.1 and Theorem 7.4.1) to see that $T_pM$ encodes essentially every information of $M$ in certain cases.

## 4.1 Well-posedness of geodesics

**Definition 4.1.1.** A *geodesic* is a parametrized curve $\gamma(t) : [a, b] \to M$ such that $\boldsymbol{\nabla}_{\gamma_t}\gamma_t = 0$ where $\gamma_t \coloneqq \partial_t\gamma = d\gamma(\partial_t)$.

**Remark 4.1.2.**

1. In some books, geodesics are defined without a parametrization.
2. Note that $\partial_t\langle\gamma_t, \gamma_t\rangle = 2\langle\boldsymbol{\nabla}_{\gamma_t}\gamma_t, \gamma_t\rangle = 0$. Therefore, the speed, $|\gamma_t|^2$, is a constant on a geodesic.

**Example 4.1.3.** On $\mathbb{R}^n$, geodesics are straight lines parametrized at a constant speed.

Fix a coordinate $\varkappa_p \coloneqq (x^1, \ldots, x^n)$. Suppose $\gamma \coloneqq (\gamma^1, \ldots, \gamma^n)$. Then $\gamma_t \coloneqq \gamma_t^i\partial_{x^i}$. Therefore, $\gamma$ is a geodesic means that

$$0 = \boldsymbol{\nabla}_{\gamma_t}\gamma_t^i\partial_{x^i} = \gamma_t^i\boldsymbol{\nabla}_{\gamma_t}\partial_{x^i} + \gamma_t(\gamma_t^i)\partial_{x^i} = \gamma_t^i\gamma_t^j\boldsymbol{\nabla}_{\partial_{x^j}}\partial_{x^i} + \gamma_{tt}^i\partial_{x^i} = (\gamma_t^i\gamma_t^j\Gamma_{ij}^k + \gamma_{tt}^k)\partial_{x^k}.$$

In particular, the coordinates of $\gamma$ solve a system of a second-order ordinary differential equation (ODE)

$$\frac{\mathrm{d}^2\gamma^k}{\mathrm{d}t^2} + \frac{\mathrm{d}\gamma^j}{\mathrm{d}t}\frac{\mathrm{d}\gamma^k}{\mathrm{d}t}\Gamma_{ij}^k = 0. \tag{4.1.1}$$

Once we have an initial condition $\gamma(0)$ and $\gamma_t(0)$, we can uniquely solve (4.1.1) at least locally.

**Exercise 4.1.4.** Check that on a round sphere, geodesics are great circles. This amounts to writing down the metric on the sphere, finding the Christoffel symbols, and solving the ODE (4.1.1). A physical way to see that is the fact that the angular momentum of a free particle is preserved implies that a geodesic lies on a plane.

**Remark 4.1.5.**
1. From the theory of ODE, a geodesic $\gamma(t, p, v)$ starting at $p$ in a direction $v$ is smooth as a function of $t, p, v$.
2. Note that $\gamma(t, p, v) = \gamma(1, p, tv)$ i.e. starting off at $p$ and traveling for time $t$ in the direction $v$ is same as traveling for a unit time in the direction $v$. To see this, observe that both $\gamma(t, p, v)$ and $\gamma(1, p, tv)$ satisfy (4.1.1) and have the same initial condition. Therefore, the uniqueness of a solution to ODE implies equality.

## 4.2 Exponential map

Based on Remark 4.1.5, we make the following definition:

**Definition 4.2.1.** The *exponential* $\exp_p : T_pM \to M$ is a map $v \mapsto \gamma(1, p, v)$.

**Remark 4.2.2.** If a manifold is *complete* as a metric space (cf. §4.4), we will later see (cf Theorem 7.3.1) that $\exp_p$ is defined globally. In fact, completeness is equivalent to $\exp_p$ being defined on all of $T_pM$. However, if we don't assume completeness, it is only defined on a neighborhood of $0 \in T_p(M)$.

**Example 4.2.3.** Let's compute the exponential map at $0 \in \mathbb{R}^n$. Consider a geodesic with $\gamma(0) = 0$ and $\gamma_t(0) = v \in \mathbb{R}^n$. Recall that $\gamma_{tt}(t) = 0$ implies $\gamma'(t)$ is constant. Therefore, $\gamma(t) = tv$. In particular, $\gamma(1, 0, v) = v$. Therefore, $\exp_0(v) = v$. Similarly $\exp_p(v) = p + v$.

**Remark 4.2.4.** Note that in Example 4.2.3 $d\exp_p = \mathbb{1}$. Therefore, $\exp_p$ is a diffeomorphism from $T_p\mathbb{R}^n \to \mathbb{R}^n$ which means that $T_p\mathbb{R}^n$ encodes all information about $\mathbb{R}^n$. In this case, $T_p\mathbb{R}^n \cong \mathbb{R}^n$, so we don't get any extra information about the manifold. However, when the manifold is non-linear, we still expect the behavior to hold at least locally, see the following theorem.[1]

Note the differential of $\exp_p : T_pM \to M$ at $v \in T_pM$ is given by $(d\exp_p)_v : T_vT_pM \to T_pM$. With the canonical identification $T_vT_pM \cong T_pM$, we will think of $d\exp_p$ to be a map $T_pM \to T_pM$.

**Theorem 4.2.5.** *The differential of* $\exp_p$ *at* $0 \in T_pM$ *is always the identity map* $\mathbb{1}$.

[1] When the manifold is a *Lie group* (a manifold with group structure where the group operations are smooth), the tangent space at the identity (Lie algebra) encodes (via the exponential map everything about the Lie group.

*Proof.* Using homogeneity, $\gamma(1,p,tv)=\gamma(t,p,v)$, we see that

$$(d\exp_p)_0(v)=\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\exp_p(tv)=\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\gamma(1,p,tv)=\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\gamma(t,p,v)=v.$$

□

**Remark 4.2.6.**

1. The inverse function theorem applied to Theorem 4.2.5 guarantees that there is a neighborhood of $0\in T_pM$ and a neighborhood of $p\in M$ where $\exp_p$ is a diffeomorphism. In this sense, $T_pM$ encodes local information about $M$.

2. If $q\in M$ is in the neighborhood $U_p$ of $p$ where $\exp_p$ is a diffeomorphism, then there exists a unique geodesic in $U_p$ from $p$ to $q$.

   The uniqueness is strictly a local property. For instance, consider $M=\mathbb{S}^2$. then geodesics are great circles. If we take a point $p$ near the north pole $N$, the shorter arc of the great circle connecting $p$ and $N$ is a unique geodesic connecting $p$ and $N$. However, if we look globally, the longer arc is also a geodesic connecting $p$ and $N$.

3. In fact, the map $\exp_p$ is a radial isometry, see Lemma 4.3.1.

## 4.3 Metric in polar coordinates

So far, we have looked at metrics only in the Cartesian coordinate system. But when our manifold has "radial" symmetry, it is helpful to work with polar coordinates. In this section, we will discuss the implication of writing a metric in polar coordinates.

As a teaser, consider $\mathbb{R}^2$ with the standard metric $g=dx^2+dy^2$. Let $(r,\theta)$ be the polar coordinates. Recall that $x=r\cos\theta$ and $y=r\sin\theta$. Therefore,

$$\begin{aligned}
dx &= \cos\theta dr - r\sin\theta d\theta\\
dy &= \sin\theta dr + r\cos\theta d\theta\\
dx^2 &= \cos^2\theta dr^2 - 2r\cos\theta\sin\theta dr d\theta + r^2\sin^2\theta d\theta\\
dy^2 &= \sin^2\theta dr^2 + 2r\cos\theta\sin\theta dr d\theta + r^2\cos^2\theta d\theta.
\end{aligned}$$

In particular, $g=dr^2+r^2d\theta^2$ and equivalently

$$g=\begin{pmatrix}1 & 0\\ 0 & r^2\end{pmatrix} \tag{4.3.1}$$

Gauss Lemma below says that even in general manifold we can "straighten up" a metric in one direction, for instance along $\partial_r$ in (4.3.1).

**Lemma 4.3.1 (Gauss).** *For $v\in T_pM$, identify $T_vT_pM\cong T_pM$. Then*

$$\langle (d\exp_p)_v(v),(d\exp_p)_v(w)\rangle=\langle v,w\rangle. \tag{4.3.2}$$

*Proof.* Define $F(t,s) \coloneqq \exp_p(t(v+sw)) = \gamma(1,p,t(v+sw))$. Observe that $F_t(1,0) = v$ and $F_s(1,0) = w$. Therefore, (4.3.2) is equivalent to $\langle v,w\rangle = \langle F_t, F_s\rangle(1,0)$. In fact, we will prove that

$$\langle v,w\rangle = \partial_t\langle F_t, F_s\rangle(t,0). \tag{4.3.3}$$

To this end, since $\gamma(1,p,t(v+sw))$ is a geodesic for fixed $s$, $\boldsymbol{\nabla}_{F_t}F_t = 0$. Therefore,

$$\partial_t\langle F_t, F_s\rangle = \langle \boldsymbol{\nabla}_{F_t}F_t, F_s\rangle + \langle F_t, \boldsymbol{\nabla}_{F_t}F_s\rangle = \langle F_t, \boldsymbol{\nabla}_{F_t}F_s\rangle.$$

We claim that $[F_t, F_s] = 0$. Then $\boldsymbol{\nabla}_{F_t}F_s = \boldsymbol{\nabla}_{F_s}F_t$. Therefore,

$$\partial_t\langle F_t, F_s\rangle = \langle F_t, \boldsymbol{\nabla}_{F_s}F_t\rangle = \frac{1}{2}\partial_s\langle F_t, F_t\rangle.$$

On the other hand, since $\gamma(1,p,t(v+sw))$ is a geodesic, the speed, $|\gamma_t| = |F_t|$, is constant. Since $F_t(1,0) = v + sw$ we see that $|F_t|^2 = |v|^2 + 2s\langle v,w\rangle + s^2|w|^2$. Finally, $\frac{1}{2}\partial_s|F_t|^2 = \langle v,w\rangle + s|w|^2$ implies (4.3.3).

It remains to prove the commutativity of $F_s$ and $F_t$. Pick a local coordinate $\varkappa_p \coloneqq (x^1,\ldots,x^n)$ and let $\gamma \coloneqq (\gamma^1,\ldots,\gamma^n)$. Then $F_s = \partial_s\gamma^i\partial_{x^i}$ and $F_t = \partial_t\gamma^j\partial_{x^j}$. Therefore,

$$\begin{aligned}\boldsymbol{\nabla}_{F_s}F_t - \boldsymbol{\nabla}_{F_t}F_s &= \boldsymbol{\nabla}_{F_s}\partial_t\gamma^j\partial_{x^j} - \boldsymbol{\nabla}_{F_t}\partial_s\gamma^i\partial_{x^i}\\ &= \partial_s\gamma^i\partial_t\gamma^j\Gamma^k_{ij}\partial_{x^k} + F_s(\partial_t\gamma^j)\partial_{x^j} - \partial_t\gamma^j\partial_s\gamma^i\Gamma^k_{ij}\partial_{x^k} - F_t(\partial_s\gamma^i)\partial_{x^i} = 0\end{aligned}$$

since $F_t(\partial_s\gamma^i) = \partial_t\partial_s\gamma^i = \partial_s\partial_t\gamma^i = F_s(\partial_t\gamma^i)$. □

**Remark 4.3.2.** Identify $T_pM \cong \mathbb{R}^n$. Let $(r,\theta)$ be the polar coordinate in $\mathbb{R}^n$. Then define $c(t) = \exp_p(r(t)\theta(t))$ such that $r(t) = t$ and $\theta(t) = \theta_0$. Then we know that at $t = 0$, $|\partial_r|^2 = 0$ and $\langle \partial_r, \partial_\theta\rangle = 0$. Therefore, the Gauss Lemma 4.3.1 implies that the same holds true along the curve $c(t)$. In particular, $\exp_p$ gives a coordinate of $M$ in which the metric looks like (4.3.1).

In addition to providing a "radial coordinate", the exponential map allows us to prove that geodesics are locally length minimizing. We will see another proof in Chapter 8 using calculus of variation.

**Theorem 4.3.3.** *Geodesics are locally length-minimizing differentiable curves.*

*Proof.* Unless otherwise stated, we will be working in the neighborhood where $\exp_p$ is a diffeomorphism. By Remark (4.3.2), we see that any curve $c : [0,1] \to M$ starting at $p$ can be locally written as $c(t) = \exp_p(r(t)v(t))$ where $|v(t)| = 1$ is the angular component and $r(t)$ is the radial component such that $r(0) = 0$ and $r(1) = R$ for some $R > 0$. It suffices to prove that the length $L(c)$ of the curve $c$ is at least $R$ and the equality holds true for geodesic starting at $p$ and ending at $c(1)$.

To this end, using the homogeneity of $\exp_p$, we see that

$$c' = d\exp_p(r'v + rv') = r'd\exp_p(v) + rd\exp_p(v').$$

Note that $2\langle v, v'\rangle = 0$ since $|v(t)|^2 = 1$. Therefore, Gauss Lemma 4.3.1 implies that $|d\exp_p v|^2 = 1$ and $\langle d\exp_p v, d\exp_p v'\rangle = 0$. Hence

$$|c'|^2 = (r')^2|d\exp_p(v)|^2 + r^2|d\exp_p(v')|^2 \geq (r')^2.$$

In particular,

$$L(c) = \int_0^1 |\gamma'|dt \geq \int_0^1 |r'(t)|dt \geq r(t) - r(0) = R.$$

Note that we get equality when $v' = 0$ and $r$ is monotone. Since $v$ is constant, $c(t)$ is a geodesic ray. □

**Remark 4.3.4.**
1. In general, if $r(0) = R_1$ and $r(1) = R_2 > R_1$ then we see that $L(c) \geq R_2 - R_1$ in Theorem 4.3.3. Again the equality holds for geodesic rays.
2. The Theorem 4.3.3 still holds if we weaken the regularity of the curves to piece-wise differentiable.

## 4.4 Riemannian manifold as a metric space

So far, we have discussed a metric structure on the tangent spaces of $M$ and have been able to compute lengths of curves. Using the lengths of a curve, we can actually give a metric structure to $M$.

**Definition 4.4.1.** For any $p, q \in M$, the distance between $p$ and $q$ is defined as

$$d(p, q) := \inf_\gamma L(\gamma). \tag{4.4.1}$$

where $\gamma$ is a piece-wise smooth curve from $p$ to $q$.

**Proposition 4.4.2.** $(M, d)$ *is a metric space.*

*Proof.*
- **Symmetry:** Since we can reverse paths, $d(p, q) = d(q, p)$.
- **Triangle inequality:** Check that $d(p, q) \leq d(p, r) + d(r, q)$.
- **Positive definite:** Suppose that $p \neq q$. Then we can find a neighborhood of $p$ that does not contain $q$. Note that any curve $c$ from $p$ to $q$ must leave $\exp_p(B_R(0))$ for some small $R > 0$. Then Theorem 4.3.3 imply that $L(c) \geq R$. In particular, $d(p, q) \geq R$.

□

**Remark 4.4.3.** Note that the triangle inequality implies that $d$ is Lipschitz, therefore it is almost everywhere differentiable. Later in §10.5, we will prove that it is smooth everywhere outside points where either $\exp_p$ fails to be a diffeomorphism or where geodesics fail to be uniquely length minimizing.

# Chapter 5

# Curvature

In this chapter, we will define the curvature of manifolds as a measure of the failure of the commutativity of directional derivatives. Then in §5.2, we will give an equivalent definition of curvature although we won't prove things in detail. Later in §7.5, we will see how curvature measures the deviation of the metric from the Euclidean metric. Further, it measures how fast the geodesics starting at the same point but in different directions deviate (in Euclidean space they deviate at a linear speed).

## 5.1 Curvature tensor

**Definition 5.1.1.** The *curvature tensor* $R$ is a map $\mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ such that for any $X, Y, Z \in \mathfrak{X}(M)$ we define

$$R(X,Y)Z \coloneqq \boldsymbol{\nabla}_Y \boldsymbol{\nabla}_X Z - \boldsymbol{\nabla}_X \boldsymbol{\nabla}_Y Z + \boldsymbol{\nabla}_{[X,Y]} Z. \tag{5.1.1}$$

**Remark 5.1.2.**

1. Fix a local coordinate $\varkappa_p = (x^1, \ldots, x^n)$. Then for any $Z \in \mathfrak{X}(M)$, we have

$$R(\partial_{x^i}, \partial_{x^j})Z = \boldsymbol{\nabla}_{\partial_{x^j}} \boldsymbol{\nabla}_{\partial_{x^i}} Z - \boldsymbol{\nabla}_{\partial_{x^i}} \boldsymbol{\nabla}_{\partial_{x^j}} Z. \tag{5.1.2}$$

   In particular, $R$ measures the commutativity of directional derivatives along the coordinates i.e. $R = 0$ means that directional derivative commute. In this sense, $R$ looks more like $\mathbb{R}^n$ where the directional derivatives commute. However, the global picture might be different.

2. In terms of Christoffel symbols, we see that

$$R(\partial_{x^i}, \partial_{x^j})\partial_{x^k} = (\Gamma^r_{ik}\Gamma^l_{jr} - \Gamma^r_{jk}\Gamma^l_{ir} + \partial_{x^j}(\Gamma^l_{ik}) - \partial_{x^i}(\Gamma^l_{jk}))\partial_{x^l}.$$

3. Using a natural pairing, we can see $R$ as a multi-linear map $\mathfrak{X}^*(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$ making $R$ a $(1,3)$-tensor. We leave it for the reader to check that

$R$ is $C^\infty(M)$-multilinear to guarantee that it is a tensor. Then the previous remark shows that $R$ viewed as a $(1,3)$-tensor can be written as

$$R^l_{ijk} = (\Gamma^r_{ik}\Gamma^l_{jr} - \Gamma^r_{jk}\Gamma^l_{ir} + \partial_{x^j}(\Gamma^l_{ik}) - \partial_{x^i}(\Gamma^l_{jk})). \tag{5.1.3}$$

We will think of $R$ as a $(1,3)$ tensor and as in Definition 5.1.1 interchangeably.

4. Recall that the Christoffel symbols can be written in terms of the first derivative of the metric, see (3.2.2). Therefore, $R$ is written in terms of the second-order derivatives of the metric. In fact, $R$ is the only tensor dependent on the second derivative (and lower order terms) of the metric. The proof requires representation theory which is beyond the scope of the notes.

5. Geometrically, we can view $R(X,Y)Z$ as the measure of deviation from $Z$ of the parallel transport of $Z$ along the boundary of the "rectangle" formed by $X, Y$. We won't go into the detail of the proof here.

The following proposition lists some symmetry that $R$ enjoys.

**Proposition 5.1.3.** *Write $\langle R(X,Y)Z,W\rangle := g(R(X,Y)Z,W)$. Then*

1. *Skew symmetry in the first two and the last two entries:*

$$\langle R(X,Y)Z,W\rangle = -\langle R(Y,X)Z,W\rangle$$
$$\langle R(X,Y)Z,W\rangle = -\langle R(X,Y)W,Z\rangle$$

2. *Pair symmetry:*

$$\langle R(X,Y)Z,W\rangle = \langle R(Z,W)X,Y\rangle.$$

3. *Bianchi identity:*

$$\langle R(X,Y)Z,W\rangle + \langle R(Y,Z)X,W\rangle + \langle R(Z,X)Y,W\rangle = 0.$$

4. *Second Bianchi identity:*

$$(\nabla_Z R)(X,Y) + (\nabla_X R)(Y,Z) + (\nabla_Y R)(Z,X) = 0.$$

*Proof.* We will only prove the skew symmetry in the entry $Z$ and $W$ and leave the rest of the proof to the reader. Note that it suffices to prove that $\langle R(X,Y)V,V\rangle = 0$.[1] In fact, for any bilinear form $B$ such that $B(v,v) = 0$ we have

$$B(x+y,x+y) = B(x,x) + B(y,y) + B(x,y) + B(y,x) \tag{5.1.4a}$$
$$-B(x-y,x-y) = -B(x,x) - B(y,y) + B(x,y) + B(y,x). \tag{5.1.4b}$$

[1] This is called polarization trick.

Then adding (5.1.4a) and (5.1.4b), we see that $B(x,y) = -B(y,x)$. Now set $B(\bullet,\bullet) = \langle R(X,Y)\bullet,\bullet\rangle$.

To this end, we can use metric compatibility to get

$$\begin{aligned}\langle \nabla_Y\nabla_X V, V\rangle &= Y\langle \nabla_X V, V\rangle - \langle \nabla_X V, \nabla_Y V\rangle\\ &= \frac{1}{2}Y(X|V|^2) - X\langle V, \nabla_Y V\rangle + \langle V, \nabla_X\nabla_Y V\rangle\\ &= \frac{1}{2}(YX - XY)|V|^2 + \langle V, \nabla_X\nabla_Y V\rangle\\ &= -\langle \nabla_{[X,Y]}V, V\rangle + \langle V, \nabla_X\nabla_Y V\rangle.\end{aligned}$$

Therefore, $\langle R(X,Y)V,V\rangle = 0$. □

**Remark 5.1.4.** The symmetry of $R$ reduces the number of $R^l_{ijk}$ one has to compute. For instance, in $\mathbb{S}^n$ we have $R_{1212} = R_{2121} = -R_{1221} = -R_{2112}$ and rest of the entries will be zero. Here $R_{1212}$ is called *sectional curvature*. It determines the curvature

**Question 5.1.5.** Are two manifolds with the same curvature isometric?

The answer turns out to be false in general. However, we will shortly see that the metric is determined if the curvature is constant and positive while it is not clear when the space has constant negative curvature.

Since $R$ is a $(1,3)$ tensor, we can contract them to get lower-rank tensors.

**Definition 5.1.6.** The *Ricci curvature* $\mathrm{Ric}(\bullet,\bullet)$ is the trace of Riemann curvature tensor i.e. for any $X,Z\in\mathfrak{X}(M)$, in a local coordinate $(x^1,\dots,x^n)$, we have

$$\mathrm{Ric}(X,Z) \coloneqq \langle R(X,\partial_{x^i})Z,\partial_{x^i}\rangle.$$

**Remark 5.1.7.**
1. The Ricci curvature is a $(0,2)$-tensor.
2. Note that $\mathrm{Ric}_{ij} = \mathrm{Ric}(\partial_{x^i},\partial_{x^j}) = R^l_{ikj}g_{lk}$. Now using (3.2.2) and (5.1.1) we can see that

$$\mathrm{Ric}_{ij} = -\frac{1}{2}\Delta g_{ij} + \text{lower order terms} \tag{5.1.5}$$

where $\Delta g_{ij}$ should be interpreted as the Laplacian of the function $g_{ij}$ instead of $(\Delta g)_{ij}$. The expression (5.1.5) hints that when we expand the metric $g$ at a point $p$, the Ricci-curvature encode information about the second-order perturbation of $g$ from the Euclidean metric.

Since we have a metric, we can raise an index of the Ricci curvature so that we can trace it.

**Definition 5.1.8.** The *scalar curvature* is the contraction of $(1,1)$ version of the Ricci curvature i.e.

$$S \coloneqq \mathrm{Tr}(\mathrm{Ric}) = g^{ij}\,\mathrm{Ric}_{ij}\,.$$

**Remark 5.1.9.** The Riemann curvature tensor can be decomposed into a sum of the three components: the first contains the sectional curvature, the second contains the Ricci curvature (trace) and the third part contains a trace-free part called the *Weyl tensor*. The decomposition requires some understanding of representation theory which is out of the scope of these notes.

## 5.2 Sectional curvature

There is an equivalent description of the Riemann curvature tensor. The sectional curvature is more geometric.

**Definition 5.2.1.** Let $\sigma \subset T_pM$ be a two-dimensional subspace. Then the sectional curvature $\kappa(\sigma)$ of $\sigma$ at $p$ is defined as

$$\kappa(\sigma) := \frac{\langle R(X,Y)X,Y\rangle}{|X|^2|Y|^2 - \langle X,Y\rangle^2}$$

where $X, Y \in \sigma$ are linearly independent vectors.

**Remark 5.2.2.**

1. Check that $\kappa(\sigma)$ does note depend on the choice of $X$ and $Y$. To show that it suffices to check that $\kappa(\sigma)$ is preserved under the change of basis:
   - $\{X,Y\} \to \{Y,X\}$
   - $\{X,Y\} \to \{\lambda X,Y\}$ for $\lambda \neq 0$
   - $\{X,Y\} \to \{X+\lambda Y,Y\}$.
2. Using the polarization trick, we see that the form $\langle R(X,Y)Z,W\rangle$ is uniquely determined if we know $\kappa(\sigma)$ for all $\sigma \subset T_pM$. In particular, the Riemann curvature $R(X,Y)Z$ is determined by the sectional curvature.

The characterization of $R$ via $\kappa$ implies that it $R$ can be written in terms so the metric as follows.

**Lemma 5.2.3.** *A manifold $M$ has constant section curvature equal to $K_0$ if and only if*

$$\langle R(X,Y)Z,W\rangle = K_0(\langle X,W\rangle\langle Y,Z\rangle - \langle Y,W\rangle\langle X,Z\rangle),$$

*i.e if and only if $R_{ijkl} = K_0(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})$.*

*Proof.* Left as an exercise. □

# Chapter 6

# Submanifolds

In other chapters of these notes, we have emphasized the intrinsic properties of manifolds. However, we can study extrinsic properties of a manifold $M$ once we embed it in a larger space say for instance the Euclidean space.[1] As we will see in Theorem 6.1.3, one can use the curvature of the ambient space $\overline{M}$ to compute up to some error the curvature of $M$. As an application, we will compute in §6.2 the curvature of a sphere viewed as a submanifold of $\mathbb{R}^n$. The term used in the error allows us to define a notion of (*principal*) curvature that depends on the ambient space. In §6.3 we will see that the only hypersurface that has the same principal curvature in every direction has to be a sphere or a plane (or their quotients).

On the other hand, the submanifolds of $M^n$ we have seen so far are geodesics. We saw how they are locally length-minimizing. We could instead study $m$-volume minimizing submanifolds. In §6.4, we scratch the theory of minimal submanifolds, see [CM11] for more details.

## 6.1 The second fundamental form

Consider a submanifold $M^n$ of $\overline{M}^{n+m}$. For any $p \in M \subset \overline{M}$, we can decompose $T_p\overline{M}$ as $T_pM \oplus (T_pM)^\perp$ where $(T_pM)^\perp$ is the perpendicular component in terms of the metric. Similarly, we can decompose the connection $\overline{\boldsymbol{\nabla}}$ on $\overline{M}$ into the tangential part and the normal part. First, the tangential part induces a connection $\overline{\boldsymbol{\nabla}}$ on $\overline{M}$ induces the connection $\boldsymbol{\nabla}$ on $M$ such that for $X, Y \in \mathfrak{X}(M)$,

$$\boldsymbol{\nabla}_X Y \coloneqq (\overline{\boldsymbol{\nabla}}_{\overline{X}}\overline{Y})^T \tag{6.1.1}$$

where $()^T$ is a projection to $\mathfrak{X}(M)$ and $\overline{X}, \overline{Y}$ are local extensions in $\overline{M}$ of $X, Y$. Second, the normal part turns out to be a tensor called the *second fundamental form.* The second

[1] Any manifold $M^n$ can be isometrically embedded on $\mathbb{R}^m$ for large $m$ via the Nash embedding which is out of the scope of our notes.

fundamental form, as we will see in Theorem 6.1.3, measures the difference between the curvature of $\overline{M}$ and that of $M$. And in the case of submanifolds with codimension one it measures how the normal vectors changes in the tangential direction (cf. Theorem 6.1.10)

**Definition 6.1.1.** The *second fundamental form* $A$ is a quadratic form such that for any $X, Y \in \mathfrak{X}(M)$

$$A(X,Y) \coloneqq (\overline{\boldsymbol{\nabla}}_{\overline{X}}\overline{Y})^{\perp} \coloneqq \overline{\boldsymbol{\nabla}}_{\overline{X}}\overline{Y} - \boldsymbol{\nabla}_X Y.$$

**Remark 6.1.2.** 1. Check $A$ is symmetric $C^\infty(M)$ multilinear map on $\mathfrak{X}(M) \times \mathfrak{X}(M)$.

- $\mathbb{R}$ linearity: The connection is $\mathbb{R}$ linear.
- $C^\infty(M)$ linearity: The linearity in $X$ follows from the linearity of the connection. On the other hand, say $f \in C^\infty(M)$. Since $X, Y \in \mathfrak{X}(M)$,

$$(\overline{\boldsymbol{\nabla}}_{\overline{X}} f\overline{Y})^{\perp} = (f\overline{\boldsymbol{\nabla}}_{\overline{X}}\overline{Y} + X(f)Y)^{\perp} = f(\overline{\boldsymbol{\nabla}}_{\overline{X}}\overline{Y}).$$

- Symmetry: Note that $\overline{\boldsymbol{\nabla}}_{\overline{X}}\overline{Y} - \overline{\boldsymbol{\nabla}}_{\overline{Y}}\overline{X} = [X,Y]$. Since $[X,Y] \in \mathfrak{X}(M)$, the symmetry follows.

2. We leave it to the reader to check that $A$ does not depend on the local extensions. From now onwards, we will use $X$ and $Y$ instead of $\overline{X}$ and $\overline{Y}$.

3. From the above remarks we see that for any $\eta \in \mathfrak{X}(M)^{\perp}$, $\langle A(\bullet,\bullet), \eta\rangle$ is a 2-form.

By $\overline{\boldsymbol{\nabla}}^T_X Y$ and $\overline{\boldsymbol{\nabla}}^{\perp}_X Y$ we mean $(\overline{\boldsymbol{\nabla}}_X Y)^T$ and $(\overline{\boldsymbol{\nabla}}_X Y)^{\perp}$ respectively.

**Theorem 6.1.3 (Gauss's Theorem).** *Let $\overline{R}$ and $R$ be the curvature on $\overline{M}$ and the induced curvature on $M$ respectively. Then for any $X, Y, Z, W \in \mathfrak{X}(M)$*

$$\overline{R}(X,Y,Z,W) = R(X,Y,Z,W) + \langle A(Y,Z), A(X,W)\rangle - \langle A(Y,W), A(X,Z)\rangle. \tag{6.1.2}$$

*Proof.* Decomposing $\overline{\boldsymbol{\nabla}}_X Z$ into tangential and normal components and using metric compatibility, we get

$$\begin{aligned}
\langle \overline{\boldsymbol{\nabla}}_Y \overline{\boldsymbol{\nabla}}_X Z, W\rangle &= \langle \overline{\boldsymbol{\nabla}}_Y \overline{\boldsymbol{\nabla}}^T_X Z, W\rangle - \langle \overline{\boldsymbol{\nabla}}^{\perp}_X Z, \overline{\boldsymbol{\nabla}}_Y W\rangle + Y\langle \overline{\boldsymbol{\nabla}}^{\perp}_X Z, W\rangle \\
&= \langle \overline{\boldsymbol{\nabla}}^T_Y \overline{\boldsymbol{\nabla}}^T_X Z, W\rangle - \langle \overline{\boldsymbol{\nabla}}^{\perp}_X Z, \overline{\boldsymbol{\nabla}}^{\perp}_Y W\rangle \\
&= \langle \boldsymbol{\nabla}_Y \boldsymbol{\nabla}_X Z, W\rangle - \langle A(X,Z), A(Y,W)\rangle.
\end{aligned}$$

Similarly, $\langle \overline{\boldsymbol{\nabla}}_X \overline{\boldsymbol{\nabla}}_Y Z, W\rangle = \langle \boldsymbol{\nabla}_X \boldsymbol{\nabla}_Y Z, W\rangle - \langle A(Y,Z), A(X,W)\rangle$. Further, $\langle \overline{\boldsymbol{\nabla}}_{[X,Y]} Z, W\rangle = \langle \boldsymbol{\nabla}_{[X,Y]} Z, W\rangle$. Combining everything we get (6.1.2). □

**Corollary 6.1.4.** *Let $e_1$ and $e_2$ be orthonormal vectors at $T_pM$. Suppose $\kappa_{12}$ is the sectional curvature of a surface defined by $e_i$. Then writing $A_{ij} \coloneqq A(e_i, e_j)$ we have*

$$\overline{k}_{12} = k_{12} + \det(A).$$

**Example 6.1.5.** When $\overline{M} = \mathbb{R}^n$, we have $k_{12} = \langle A_{22}, A_{11}\rangle - |A_{12}|^2$.

**Remark 6.1.6.** The second fundamental form could be different in two locally isometric submanifolds where the curvatures match up. For instance, the plane $\mathbb{R}^2$ and a cylinder $\mathbb{R} \times \mathbb{R}/\mathbb{Z}$ viewed as submanifolds of $\mathbb{R}^3$ are locally isometric and therefore the curvature vanishes in both cases. However, the plane is *minimal* (cf. Definition 6.4.3) while the cylinder is not. The vanishing condition of the curvature amounts to the determinant of $A$ being zero. However, the trace of $A$ is non-zero in the case of the cylinder.

In the rest of the section, we will prove that on $M^n \hookrightarrow \mathbb{R}^{n+1}$ the second fundamental form measures the change in the normal vectors when we move in the tangential direction.

**Definition 6.1.7.** The *Gauss map* $\vec{n} : M^n \to \mathbb{S}^n$ maps $p \in M^n$ to the unit normal vector at $p$.

**Remark 6.1.8.** For $\vec{n}$ to be well defined, we want $M$ to be orientable since the normal bundle inherits the orientation of the manifold. The Gauss map is ill-defined in the case of Möbius band.

**Example 6.1.9.**

1. If $M^n = S^n$ then the Gauss map is the identity.
2. If $M^n = \mathbb{R}^n$ or any hyperplane, the Gauss map is constant. In fact, if the Gauss map is constant then $M^n$ is a hyperplane.

**Theorem 6.1.10.** *The second fundamental form $A$ measures the change of $\vec{n}$ in tangential directions.*

*Proof.* First, note that $\langle \overline{\boldsymbol{\nabla}}_V \vec{n}, \vec{n}\rangle = \frac{1}{2}\langle \vec{n}, \vec{n}\rangle = 0$. Therefore, there is no change of $\vec{n}$ in normal direction.

Second note that after identifying $T_{\vec{n}(p)}(\mathbb{S}^n) \cong T_pM$ we can view the differential of the Gauss map as $d\vec{n}_p : T_pM \to T_pM$. For any $V \in T_pM$ choose a curve $\gamma : [0,1] \to M$ such that $\gamma(0) = p$ and $\gamma'(0) = V$. Then

$$\langle d\vec{n}_p(V), W\rangle = \left\langle \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \vec{n}(\gamma(t)), W \right\rangle = \langle \overline{\boldsymbol{\nabla}}_V \vec{n}, W\rangle = -\langle \vec{n}, \overline{\boldsymbol{\nabla}}_V W\rangle = -\langle A(V,W), \vec{n}\rangle.$$

□

## 6.2 Curvatures of the sphere

In this section, we will use (6.1.2) to compute the curvature of $\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} \mid |x|^2 = 1\}$. Since the normal for a level set of a function $f$ is given by the gradient, we see that the normal vector at $x \in \mathbb{S}^n$ is given by $x$. Therefore for any $V, W \in \mathfrak{X}(\mathbb{S}^n)$ we see that $A(V,W) = \langle A(V,W), x\rangle x$. Observe that metric compatibility and $\overline{\nabla}_V x = V$

$$\langle A(V,W), x\rangle = \langle \overline{\nabla}^{\perp}_V W, x\rangle = \langle \overline{\nabla}_V W, x\rangle = -\langle W, \overline{\nabla}_V x\rangle = -\langle W, V\rangle.$$

Therefore, (6.1.2) implies that for any $X, Y, Z, W \in \mathfrak{X}(\mathbb{S}^n)$ we have

$$R(X,Y,Z,W) = \langle Y,W\rangle\langle X,Z\rangle - \langle Y,Z\rangle\langle X,W\rangle. \tag{6.2.1}$$

**Remark 6.2.1.** The manifolds where $R(X,Y,Z,W) = \kappa\{\langle Y,W\rangle\langle X,Z\rangle - \langle Y,Z\rangle\langle X,W\rangle\}$ are called *symmetric spaces.* Here $\kappa$ is constant if and only if the manifold has constant sectional curvature equal to $\kappa$ in which case the manifold is called a *space form.* The canonical model of the space forms are $\mathbb{S}^n, \mathbb{R}^n$ and $\mathbb{H}^n$ for which $\kappa = 1, 0, -1$ respectively.

Pick an orthonormal frame $e_i$ of $\mathfrak{X}(\mathbb{S}^n)$. Then

$$\mathrm{Ric}(X,Z) = R(X,e_i,Z,e_i)\langle X,Z\rangle - \langle e_i,X\rangle\langle e_i,Z\rangle = (n-1)\langle X,Z\rangle.$$

**Remark 6.2.2.** Any manifold where $\mathrm{Ric} = \lambda g$ where $\lambda$ is a constant is called *Einstein manifold.*

The scalar curvature of the sphere is $S = \mathrm{Ric}(e_i,e_i) = n(n-1)$.

## 6.3 Normal connection

As noted in §6.1, there is a decomposition $T\overline{M} = TM \oplus (TM)^{\perp}$. In this section, we will prove that the projection on $(TM)^{\perp}$ of the curvature tensor $\overline{R}$ on $\overline{M}$ is related to the derivative of the second fundamental form (cf. Theorem 6.3.3). As an application, we will see the manifolds with the same principal curvatures, extrinsic quantities, at each point are rigid (cf. Theorem 6.3.7).

To make sense of taking the derivative of the second fundamental form, we will define a connection on normal vector fields $\mathfrak{X}(M)^{\perp}$.

**Definition 6.3.1.** The *normal connection* $\overline{\nabla}^{\perp}$ on $M \hookrightarrow \overline{M}$ is a map $\mathfrak{X}(M) \times \mathfrak{X}(M)^{\perp} \to \mathfrak{X}(M)^{\perp}$. such that for any $V \in \mathfrak{X}(M)$ and $\eta \in \mathfrak{X}(M)^{\perp}$

$$\overline{\nabla}^{\perp}_V \eta = (\overline{\nabla}_V \eta)^{\perp}. \tag{6.3.1}$$

**Remark 6.3.2.** 1. Check that $\overline{\nabla}^{\perp}$ satisfies the properties of a connection, i.e. it satisfies

- $C^{\infty}(M)$ linearity in $V$
- Leibniz rule in $\eta$.

2. The connection is metric compatible since $\overline{\nabla}$ is. However, it is not symmetric.

We can extend the definition of normal connection to define the covariant derivative of the second fundamental form such that for any $V, X, Y \in \mathfrak{X}(M)$

$$(\overline{\nabla}^{\perp}_V A)(X,Y) \coloneqq \overline{\nabla}^{\perp}_V(A(X,Y)) - A(\nabla_V X, Y) - A(X, \nabla_V Y). \tag{6.3.2}$$

**Theorem 6.3.3 (Codazzi equation).** *Consider a submanifold $M^n \hookrightarrow \overline{M}^{n+m}$. For any $V, W, X \in \mathfrak{X}(M)$*

$$(\overline{R}(v,w)x)^{\perp} = (\overline{\nabla}^{\perp}_W A)(V,X) - (\overline{\nabla}^{\perp}_V A)(W,X). \tag{6.3.3}$$

*Proof.* Note that

$$\begin{aligned}(\overline{\nabla}^{\perp}_W A)(V,X) - (\overline{\nabla}^{\perp}_V A)(W,X) =& \overline{\nabla}^{\perp}_W(A(V,X)) - A(\nabla_W V, X) - A(V, \nabla_W X) \\ &- \overline{\nabla}^{\perp}_V(A(W,X)) + A(\nabla_V W, X) + A(W, \nabla_V X).\end{aligned}$$

On the other hand,

$$\begin{aligned}(\overline{R}(V,W)X)^{\perp} =& \overline{\nabla}^{\perp}_W \overline{\nabla}_V X - \overline{\nabla}^{\perp}_V \overline{\nabla}_W X + \overline{\nabla}^{\perp}_{[V,W]} X \\ =& \overline{\nabla}^{\perp}_W(A(V,X)) + A(W, \nabla_V X) - \overline{\nabla}^{\perp}_V(A(W,X)) - A(V, \nabla_W X) \\ &+ A(\nabla_V W - \nabla_W V, X).\end{aligned}$$

□

**Remark 6.3.4.** 1. Suppose $\overline{M} = \mathbb{R}^{n+1}$. Then (6.3.3) implies that $\nabla_{\bullet} A(\bullet,\bullet)$ is symmetric in all of the entries. A 3-tensor that is symmetric in all three entries is called a *Codazzi tensor*. Similarly, $\nabla_{\bullet} A(\bullet,\bullet)$ for submanifolds of $\mathbb{S}^{n+1}$ is also a Codazzi tensor. In fact, $\langle \overline{R}(V,W)X, Z\rangle = \langle W,Z\rangle\langle V,X\rangle - \langle W,X\rangle\langle V,Z\rangle$. If three of the vectors are tangent and one is perpendicular then $\overline{R} = 0$. The same is true for submanifolds of a hyperbolic manifold.

Recall that for any $\eta \in \mathfrak{X}(M)^{\perp}$, $A_\eta(\bullet,\bullet) \coloneqq \langle A(\bullet,\bullet), \eta\rangle$ is a $(0,2)$-tensor. Then $g^{ik}(A_\eta)_{ij}$ is $(1,1)$ tensor and can viewed as a linear map $\mathfrak{X}(M) \to \mathfrak{X}(M)$ which allows us to talk about its eigenvalues.

**Definition 6.3.5.** A submanifold is *umbilic* if $A_\eta$ has one eigenvalue at each point i.e. $g^{ik}(A_\eta)_{ij} = f\delta^k_i$ for some $f \in C^\infty(M)$.

**Remark 6.3.6.** 1. On an umbilic hypersurface, $\langle A(x,y), \vec{n}\rangle = f\langle x,y\rangle$. In the case of $M^2 \hookrightarrow \mathbb{R}^3$, an umbilic surface the two principal curvatures at each point are the same.

In the following theorem, we will prove a rigidity theorem for umbilic submanifolds of $\mathbb{R}^{n+1}$.

**Theorem 6.3.7.** *Let $n \geq 2$. If $M^n \hookrightarrow \mathbb{R}^{n+1}$ is umbilic and connected then $M$ is either a sphere or a plane (or their quotients).*

*Proof.* Note that an umbilic hypersurface satisfies $\langle A, \vec{n} \rangle = fg$. It suffices to prove that $f \equiv c$ for some constant $c$. In fact, if $c = 0$ then $\langle A, \vec{n} \rangle = 0$ which implies that $A \equiv 0$. Then Theorem 6.1.10 implies that $\vec{n}$ is constant. Since $M$ is connected it has to be a plane of its quotients. On the other hand, suppose $c \neq 0$. Then for any $X, Y \in \mathfrak{X}(M)$ we know that $\langle \overline{\boldsymbol{\nabla}}_X \vec{n}, Y \rangle = -\langle \vec{n}, \overline{\boldsymbol{\nabla}}_X Y \rangle = \langle \vec{n}, A(X,Y) \rangle = cg(X,Y)$. In particular, $\langle \overline{\boldsymbol{\nabla}}_X \vec{n}, \vec{n} \rangle = 0$ and $\overline{\boldsymbol{\nabla}}_X \vec{n} = -cX$. Now define a map $F : M^n \to \mathbb{R}^{n+1}$ such that

$$F(p) \coloneqq p + \frac{1}{c} \vec{n}(p).$$

Note that for any $X \in \mathfrak{X}(M)$ we have $\overline{\boldsymbol{\nabla}}_X F = \overline{\boldsymbol{\nabla}}_X p + \frac{1}{c} \overline{\boldsymbol{\nabla}}_X \vec{n} = X + \frac{1}{c}(-cX) = 0$. Since $M$ is connected, $F$ is constant everywhere. In particular, $M$ is a sphere with the center at $F(p)$ (or its quotients).

It remains to prove that $f$ is constant. To this end, let $X, Y, Z \in \mathfrak{X}(M)$. Note that $X(f)g = X(fg) = X\langle A, \vec{n} \rangle = \langle \overline{\boldsymbol{\nabla}}_X A, \vec{n} \rangle = \langle \overline{\boldsymbol{\nabla}}^{\perp}_X A, \vec{n} \rangle$. Therefore, $X(f)g(Y,Z) = \langle (\boldsymbol{\nabla}_X A)(Y,Z), \vec{n} \rangle$. Using Theorem 6.3.3 we know that $X(f)g(Y,Z)$ fully symmetric.

For the sake of contradiction, assume that $\boldsymbol{\nabla} f \neq 0$. Choose a frame $\{e_i\}_{1 \leq i \leq n}$ of $\mathfrak{X}(M)$ such that $e_1 = \frac{\boldsymbol{\nabla} f}{|\boldsymbol{\nabla} f|}$. Note that $e_1(f) = \langle \boldsymbol{\nabla} f, e_1 \rangle = |\boldsymbol{\nabla} f|$. On the other hand, the symmetry of $X(f)g(Y,Z)$ implies that $e_1(f)g(e_2,e_2) = e_2(f)g(e_1,e_2) = 0$ which means $|\boldsymbol{\nabla} f| = 0$, a contradiction. Therefore, $f$ is a constant. $\square$

**Remark 6.3.8.** We need $n \geq 2$ because in $n = 1$ the second fundamental form is just a function i.e. any curve is umbilic, so there is no structure theorem.

## 6.4 Minimal submanifolds

In this section, we define minimal submanifolds of $\mathbb{R}^n$ although we bypass the fact that minimal submanifolds are critical points of the volume functional. See [CM11] for an in-depth study of minimal submanifolds.[2] Here we will prove that the coordinate functions on minimal submanifolds of $\mathbb{R}^n$ are harmonic.

**Definition 6.4.1.** The *mean curvature tensor* $H$ of a submanifold is defined as

$$H \coloneqq -\operatorname{Tr}(A). \tag{6.4.1}$$

**Remark 6.4.2.** In literature, the mean curvature is defined as $\frac{1}{n}\operatorname{Tr}(A)$. People working in minimal surface use (6.4.1).

[2]Later in Chapter 8 we will see how geodesics can be viewed as one-dimensional submanifolds that minimize length.

**Definition 6.4.3.** A submanifold $M$ is *minimal* when the mean curvature tensor vanishes i.e. $H \equiv 0$.

Consider a surface $M^2 \hookrightarrow \mathbb{R}^3$. Suppose $\{e_1, e_2\}$ is an orthonormal frame on $M$ and $\vec{n}$ a frame of normal vectors. Then

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \vec{n}$$

is a symmetric matrix. The eigenvalues of $A$ are known as *principle curvatures.*

If $M$ is minimal then $A_{11} = -A_{22}$. Therefore Corollary 6.1.4 implies that

$$k_{12} = -\frac{1}{2}|A|^2 \leq 0.$$

**Fun Fact 6.4.4.** Lebesgue's theory of integration is a failure to solve the Plateau problem (minimal surface area problem).

**Theorem 6.4.5.** *On a minimal submanifold $M^n \hookrightarrow \mathbb{R}^{n+m}$ the coordinate functions $x^i$ are harmonic i.e.*

$$\Delta x^i = 0.$$

*Here $\Delta$ is the Laplacian on $M$.*

*Proof.* Consider an orthonormal frame $\{e_j\}_{1 \leq j \leq n}$ on $\mathfrak{X}(M)$. Then

$$\begin{aligned}
\Delta x^i &= \langle \overline{\boldsymbol{\nabla}}^T_{e_j} \overline{\boldsymbol{\nabla}} x^i, e_j \rangle - \langle \overline{\boldsymbol{\nabla}}^T_{e_j} \overline{\boldsymbol{\nabla}}^\perp x^i, e_j \rangle \\
&= \langle \overline{\boldsymbol{\nabla}}_{e_j} \overline{\boldsymbol{\nabla}} x^i, e_j \rangle - \langle \overline{\boldsymbol{\nabla}}_{e_j} \overline{\boldsymbol{\nabla}}^\perp x^i, e_j \rangle \\
&= \overline{\Delta} x^i - \langle \overline{\boldsymbol{\nabla}}^\perp x^i, \overline{\boldsymbol{\nabla}}_{e_j} e_j \rangle \\
&= -\langle \overline{\boldsymbol{\nabla}}^\perp x^i, \overline{\boldsymbol{\nabla}}^\perp_{e_j} e_j \rangle \\
&= -\langle \overline{\boldsymbol{\nabla}}^\perp x^i, H \rangle \\
&= 0
\end{aligned}$$

where we used the fact that $\overline{\Delta} x^i = 0$. □

**Remark 6.4.6.** The harmonicity of the coordinate functions is the acceleration-free criterion in the case of geodesics.

# Chapter 7

# Jacobi fields

In this chapter, we will start setting up calculus on the space of curves on $M$, focusing on geodesics. To see a calculus on the space of general curves, see Chapter 8. The idea is to think about a geodesic as an element in a larger space rather than thinking of it in isolation. The hope is to infer some properties of the underlying manifold by studying the space of geodesics. In fact, in §7.3, we will prove that all geodesics can be infinitely extended on $M$ if and only if it is a complete metric space. As an application, we will prove in §7.4, that $M$ can be "linearized" if it is complete with non-positive sectional curvature and simply connected.

On the other hand, using the tangent field at a geodesic, we will see in §7.5 how curvature measures the deviation of geodesics starting in different directions as well as the deviation of the metric from the Euclidean metric.

## 7.1 Tangent at a geodesic

Consider a function $\gamma : [-\epsilon, \epsilon] \times [0, t] \to M$ such that $\gamma(s, \bullet)$ is a geodesic. Define $\gamma_s := d\gamma(\partial_s)$ and $\gamma_t = d\gamma(\partial_t)$. Since $\gamma(s, \bullet)$ is a geodesic, we have $\boldsymbol{\nabla}_{\gamma_t}\gamma_t = 0$. Therefore, for all $s$ we have $\boldsymbol{\nabla}_{\gamma_s}\boldsymbol{\nabla}_{\gamma_t}\gamma_t = 0$. As we saw in the proof of Gauss Lemma 4.3.1, $\gamma_t$ and $\gamma_s$ commute. Recall that $R(\gamma_t, \gamma_s)\gamma_t = \boldsymbol{\nabla}_{\gamma_s}\boldsymbol{\nabla}_{\gamma_t}\gamma_t - \boldsymbol{\nabla}_{\gamma_t}\boldsymbol{\nabla}_{\gamma_s}\gamma_t + \boldsymbol{\nabla}_{[\gamma_t,\gamma_s]}\gamma_t$. Therefore,

$$0 = \boldsymbol{\nabla}_{\gamma_s}\boldsymbol{\nabla}_{\gamma_t}\gamma_t = \boldsymbol{\nabla}_{\gamma_t}\boldsymbol{\nabla}_{\gamma_t}\gamma_s + R(\gamma_t, \gamma_s)\gamma_t. \tag{7.1.1}$$

Note that $\gamma_s$ is a perturbation of $\gamma$ in the direction perpendicular to $\gamma_t$. Therefore, (7.1.1) can be viewed as an equation for a tangent field at a geodesic. Motivated by (7.1.1), we make the following definition:

**Definition 7.1.1.** A *Jacobi field* $J$ along $\gamma$ is a vector field that satisfies the *Jacobi equation:*

$$\boldsymbol{\nabla}_{\gamma_t}\boldsymbol{\nabla}_{\gamma_t}J = -R(\gamma_t, J)\gamma_t. \tag{7.1.2}$$

The Jacobi equation (7.1.2) is a second-order linear ODE along $\gamma$. In fact, choose a parallel orthonormal frame $\{e_i\}_{1\leq i\leq n}$ along $\gamma$ such that $e_1 = \frac{\gamma_t}{|\gamma_t|}$. Writing $J := J^k e_k$, we get $\nabla_{\gamma_t}\nabla_{\gamma_t} J = \frac{\mathrm{d}^2 J^k}{\mathrm{d}t^2} e_k$. Further, $R(\gamma_t, J)\gamma_t = |\gamma_t|^2 J^i R^k_{1i1} e_k$. Therefore, for a unit speed geodesic, (7.1.2) becomes

$$\frac{\mathrm{d}^2 J^k}{\mathrm{d}t^2} = -J^i R^k_{1i1}. \tag{7.1.3}$$

In particular, if we specify $J$ and $J'$ at $t = 0$, we get a unique solution. Therefore, the dimension (over $\mathbb{R}$) of Jacobi field along $\gamma$ is $2n$.

**Example 7.1.2.**

1. Note that $J := \gamma_s$ is a Jacobi field. Set $\gamma(s,t) := \exp_{\gamma(0)}(t(\gamma_t|_0)+sw)$. Then

$$J(t) = \gamma_s = (\mathrm{d}\exp_{\gamma(0)})_{t\gamma'(0)}(tw) \tag{7.1.4}$$

satisfies (7.1.2) with $J(0) = 0$ and $J'(0) = w$.

2. Note that $\mathrm{Ric} \equiv 0$ on $\mathbb{R}^n$. Therefore, (7.1.3) implies that $J$ has to be linear in $t$. In particular, geodesics starting at a point in two different directions will diverge linearly in $t$.

3. On $\mathbb{S}^n$ we have $R_{ijkl} = \delta_{ik}\delta_{jl} - \delta_{jk}\delta_{il}$. In particular, $R_{1i1j} = \delta_{ij} - \delta_{i1}\delta_{j1} = \begin{pmatrix} 0 & 0 \\ 0 & I_{n-1} \end{pmatrix}$. Therefore, $J^1$ is linear. On the other hand, for any $k \geq 2$

$$\frac{\mathrm{d}^2 J^k}{\mathrm{d}t^2} = -J^k.$$

Then $J^k(t) = a\cos t + b\sin t$. If we want $J^k(0) = 0$ then $a = 0$ in which case $J^k(t) = b_k \sin t$. Note that $J^k = 0$ at $t = \pi$. Physically, two geodesics from the north pole meet at the south pole. Further, $\sin t \sim t$ for small $t$. It means that for small $t$ the divergence of geodesics starting at the same point behaves like that in $\mathbb{R}^n$.

4. On $\mathbb{H}^n$ we have $R_{ijkl} = -\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}$. Thus $R_{1i1j} = \delta_{ij} - \delta_{i1}\delta_{j1} = \begin{pmatrix} 0 & 0 \\ 0 & -I_{n-1} \end{pmatrix}$. As above $J^1$ is linear. However, when $k \geq 2$ we get

$$\frac{\mathrm{d}^2 J^k}{\mathrm{d}t^2} = J^k.$$

In particular, $J^k = a\sinh t + b\cosh t$. If we want $J^k(0) = 0$, we have $J^k = a\sinh t$. Note that for small $t$, $\sinh t$ behaves like $t$, so the geodesics diverge linearly. However, for large $t$ they start to diverge exponentially.

**Remark 7.1.3.** Say that $J$ is a Jacobi field along $\gamma$ and $f$ is a function on $M$. Then

$$\nabla_{\gamma_t}\nabla_{\gamma_t}(fJ) = f''J + 2f'\nabla_{\gamma'}J + f\nabla_{\gamma_t}\nabla_{\gamma_t}J = f''J + 2f'\nabla_{\gamma'}J - R(\gamma_t, fJ)\gamma_t.$$

Therefore, $fJ$ is a Jacobi field if $f''J + 2f'\nabla_{\gamma'}J = 0$.

## 7.2 Conjugate points

**Definition 7.2.1.** Suppose $\gamma : [0, l] \to M$ is a unit speed geodesic. We say that $\gamma(0)$ and $\gamma(l)$ are *conjugate points* if there exists a Jacobi field such that $J(0) = J(l) = 0$ and $J \neq 0$. The *multiplicity* of the conjugate points is the dimension of such $J's$.

**Example 7.2.2.**

1. On $\mathbb{R}^n$, there are no conjugate points as we saw in Example 7.1.2 the Jacobi fields are linear function. And the only linear function that vanishes at two points is 0. Therefore, the multiplicity is also 0.

2. On $\mathbb{S}^n$, if $J(0) = 0$ then $J(t) = \sum_{i=2}^n a_i \sin t e_i$ where $\{e_i\}_{1 \le i \le n}$ is an orthogonal parallel frame with $e_1 = \gamma'$. Note that sin vanishes when $t = m\pi$ for any $m \in \mathbb{Z}$. Therefore, the conjugate points are antipodal points. Further, a point is conjugate to itself. Moreover, the multiplicity is $n - 1$ because we can choose parameters of $e_i$ for $i \ge 2$.

3. There are no conjugate points on $\mathbb{H}^n$. For the sake of contradiction assume that a non-trivial Jacobi field vanishes at $t = 0$ and $t = l$. Then $\frac{\mathrm{d}}{\mathrm{d}t}\langle \frac{\mathrm{d}J}{\mathrm{d}t}, J\rangle = \langle \frac{\mathrm{d}^2 J}{\mathrm{d}t^2}, J\rangle + \langle \frac{\mathrm{d}J}{\mathrm{d}t}, \frac{\mathrm{d}J}{\mathrm{d}t}\rangle = \langle J, J\rangle + \langle \frac{\mathrm{d}J}{\mathrm{d}t}, \frac{\mathrm{d}J}{\mathrm{d}t}\rangle \ge 0$. Since $\langle \frac{\mathrm{d}J}{\mathrm{d}t}, J\rangle(0) = \langle \frac{\mathrm{d}J}{\mathrm{d}t}, J\rangle(l) = 0$ we see that $\langle \frac{\mathrm{d}J}{\mathrm{d}t}, J\rangle \equiv 0$. On the other hand, $\frac{\mathrm{d}}{\mathrm{d}t}\langle J, J\rangle = 2\langle \frac{\mathrm{d}J}{\mathrm{d}t}, J\rangle = 0$. Therefore, $|J|$ is constant. But $|J|(0) = 0$. In particular, $J \equiv 0$ which is a contradiction.

The following theorem relates the multiplicity to the failure of the exponential map to be a diffeomorphism.

**Proposition 7.2.3.** *Consider a unit speed geodesic $\gamma : [0, l] \to M$. Then $\gamma(0)$ and $\gamma(l)$ are conjugate points if and only if $l\gamma'(0)$ is a critical point for $\exp_{\gamma(0)}$. Further, the multiplicity is equal to $\dim \ker(\mathrm{d}\exp_{\gamma(0)})_{l\gamma'(0)}$.*

*Proof.* Recall from Example 7.1.2 that a field with $J(0) = 0$ and specified $J'(0)$ is given by $J(t) = (\mathrm{d}\exp_{\gamma(0)})_{t\gamma'(0)}(tJ'(0))$. Therefore, $J(l) = 0$ if and only if $(\mathrm{d}\exp_{\gamma(0)})_{l\gamma'(0)}(lJ'(0)) = 0$. Note that if $J'(0) = J(0) = 0$ then $J \equiv 0$. Therefore, for non-trivial Jacobi field, $l\gamma'(0)$ is the critical point of $\exp_{\gamma(0)}$. The second part is left as an exercise. □

## 7.3 Hopf–Rinow

Recall that $\exp_p$ is a diffeomorphism from the neighborhood of the origin in $T_pM$ to a neighborhood of $p$. However, it is not always clear if $\exp_p$ is defined everywhere. In fact, consider $\mathbb{R}^2 \setminus \{0\}$ and $g_{ij} = \delta_{ij}$. A geodesic heading towards the origin can't be defined past 0. In particular, $\exp_p$ is not defined everywhere. Similarly, consider an infinite cylinder $S^1 \times \mathbb{R}$ with metric $\mathrm{d}\theta^2 + \frac{1}{1+y^4}\mathrm{d}y^2$ and coordinate map $(y, \theta)$. Let $\gamma(t) = (t, 0)$. Then $L(\gamma) = \int_{-\infty}^{\infty} \frac{1}{1+t^4}\mathrm{d}t$. It means that after finite time we will exhaust all of the distance there is to travel in $y$ direction. In other words, we can't extend the geodesic until infinite time. In the following theorem, we will characterize the manifolds where $\exp_p$ is defined everywhere as well as where geodesics can be continued forever.

**Theorem 7.3.1 (Hopf–Rinow).** *Suppose $M$ is connected and $p \in M$. The following are equivalent:*

1. $\exp_p : T_pM \to M$ *is defined everywhere.*
2. *Every closed and bounded set $K \subset M$ is compact.*
3. $(M, d)$ *is a complete metric space.*
4. *$M$ is geodesically complete i.e. all geodesics $\gamma(t)$ are defined on $[0, \infty]$.*
5. *There exists compact exhaustion of $M$ (i.e) there are compact sets $K_n \Subset M$ such that $K_n \subset K_{n+1}$ such that $d(p, M \setminus K_n) \to \infty$.*

   *Moreover, any of the above implies the following:*

6. *Given any $q$ there exists a geodesic from $p$ to $q$.*

**Remark 7.3.2.** In an open disk, 6 does not imply 5. Therefore, 6 is not equivalent to the rest of the statements.

**Lemma 7.3.3.** *Suppose $B_t(x)$ is normal and $y \notin \bar{B}_t(x)$. Let $z \in \partial B_t(x)$ be the closest point to $y$. Then $d(y, z) + t = d(x, y)$.*

*Proof.* Note that $d(x, y) \leq d(x, z) + d(z, y) = t + d(y, z)$. On the other hand, let $\sigma$ be any curve from $x$ to $y$. Note that $\sigma$ has to hit $\partial B_t(x)$. Let $w \in \partial B_t(x) \cap \sigma$. Then $L(\sigma) \geq d(x, w) + d(w, y) = t + d(w, y) \geq t + d(z, y)$ where we used the fact that $z$ is the closest point in $\partial B_t(x)$ to $y$. Therefore, $t + d(y, z) \leq d(x, y)$. $\square$

*Proof of Theorem 7.3.1.* 1 $\implies$ 6: Assume that $\exp_p$ is defined on all of $T_pM$. Then for any $q \in M$ we need to construct geodesic $\gamma$ from $p$ to $q$.

Consider a normal ball $B_\delta(p)$ for some $\delta > 0$. Since the distance function $d(\bullet, q)$ is continuous and $\partial B_\delta(p)$ is compact there is some $x_0 \in \partial B_\delta(p)$ that is closest to $q$. Let $\gamma$ be a geodesic from $p$ through $x_0$ with $\gamma'(0) = v \in T_pM$ with $|v| = 1$. Note that $x_0 = \exp_p(\delta v)$. Set $r = d(p, q)$. We claim that $\exp_p(rv) = q$. It suffices to prove that $d(q, \exp_p(rv)) = 0$. To this end, define

$$A \coloneqq \{s \in [0, r] \mid d(\exp_p(sv), q) = r - s\}.$$

Note that $A$ is non-empty since $0 \in A$. Since $A$ is closed it suffices to prove that $A$ is open which would imply $A = [0, r]$.

We claim that $[0, s] \subset A$ whenever $s \in A$, i.e. $d(\gamma(s), q) = r - s$. Say $t < s$. Note that the triangle inequality implies $d(\gamma(0), q) \leq d(\gamma(0), \gamma(t)) + d(\gamma(t), q)$ which implies $r - t \leq d(\gamma(t), q)$. On the other hand, $d(q, \gamma(t)) \leq d(q, \gamma(s)) + d(\gamma(s), \gamma(t)) \leq r - s + s - t = r - t$. Therefore, $d(\gamma(t), q) = r - t$.

To this end, assume that $s \in A$. Consider a normal ball $B_t(\gamma(s))$ and let $z$ be the closest point in $\partial B_t(\gamma(s))$ to $q$. Then Lemma 7.3.3 implies that $d(q,z) + t = d(q, \gamma(s)) = r - s$ i.e. $d(z,q) = r - (s+t)$. We claim that $z = \gamma(t+s)$ which implies $d(\gamma(s+t), q) = r - (s+t)$ i.e. $[0, s+t] \in A$. Therefore, $A$ is open.

It remains to prove that $z = \gamma(s+t)$. Note that $r = d(p,q) \leq d(p,z) + d(z,q) = d(p,z) + r - (s+t)$. Therefore, $d(p,z) \geq s+t$. On the other hand, $d(p,z) \leq d(p,\gamma(s)) + d(\gamma(s), z) = s+t$. Therefore, $d(p,z) = s+t$. On the other hand, let $\sigma$ be the curve that joins $p$ and $z$ obtained joining $\gamma[0,s]$ and a geodesic connecting $\gamma(s)$ and $z$. Note that $L(\sigma) = s+t$. Since $d(p,z) = L(\sigma)$, $\sigma$ is a length-minimizing curve which implies that it is a geodesic. By the uniqueness of the geodesic with specified boundary conditions, we see that $\sigma = \gamma[0, s+t]$. In particular, $z = \gamma(s+t)$ which finishes the proof of the implication 1 $\implies$ 6.

1 $\implies$ 2: Suppose $K \subset M$ is closed and bounded. Note that boundedness implies that $K \subset B_R(p)$ for some $R > 0$ i.e. $K \subset \exp_p(\overline{B_R(0)})$. Since the exponential map is continuous $\exp_p(\overline{B_R(0)})$ is compact. In particular, $K$ is compact.

2 $\implies$ 3: Consider a Cauchy sequence $\{p_n\}$. Since the sequence is Cauchy the points are contained in a bounded set and therefore a compact set. Thus there is a subsequence $\{p_{n'}\}$ that converges. However, a Cauchy sequence with a convergent subsequence converges.

3 $\implies$ 4: Suppose $\gamma$ is defined on $[0,l)$ but not past $l$. Assuming that $\gamma$ has unit speed $L(\gamma) = l$. Suppose $\{t_j\}$ converges to $l$. Then $d(\gamma(t_i), \gamma(t_j)) \leq |t_i - t_j|$. Therefore, $\gamma(t_i)$ forms a Cauchy sequence which has a limit say $q$. We can choose a normal ball around $q$ and extend the geodesic $\gamma$ past $l$ which is a contradiction.

4 $\implies$ 1: It follows from the definition of the exponential map.

2 $\iff$ 5: It is a fact in topology. □

## 7.4 Cartan–Hadamard

In this section, we will study the topological implication of having negative sectional curvature. For a review of fundamental groups see [Mun00].

**Theorem 7.4.1 (Cartan–Hadamard).** *Suppose $(M,d)$ is a complete manifold without boundary and non-positive sectional curvature. Further, $\pi_1(M) = \{0\}$. Then $\exp_p$ is a global diffeomorphism. In particular, $M$ is diffeomorphic to $\mathbb{R}^n$.*

**Corollary 7.4.2.** *If $(M,g)$ is complete and sectional is less than $0$ then the universal cover $\tilde{M}$ is diffeomorphic to $\mathbb{R}^n$.*[1]

[1] $M$ is a quotient of $\mathbb{R}^n$. In general, one can study quotients under a group action. For a detailed study of the spaces of negative sectional curvature (hyperbolic geometry), see [BH13].

**Corollary 7.4.3.** *For any $n \geq 2$, $\mathbb{S}^n$ does not admit a metric with non-positive sectional curvature.*

**Remark 7.4.4.** We need $\pi_1(M) = \{0\}$. In fact, let $M = \mathbb{S}^1$ for which $\pi(M) = \mathbb{Z}$. Identify the tangent space at a point in $M$ with $\mathbb{R}$. Note that $\mathbb{R} \to \mathbb{S}^1$ is given by $e^{i\theta}$ which is not a global diffeomorphism.

**Lemma 7.4.5.** *Suppose $F : M \to N$ is a local diffeomorphism between complete manifolds. If $F$ is a local isometry, then $F$ is a covering map.*

*Proof of Theorem 7.4.1.* Note that Example 7.2.2 (3) implies that there are no conjugate points on $M$. Therefore, Proposition 7.2.3 implies that the exponential map is a local diffeomorphism. In light of Lemma 7.4.5, we claim that $\exp_p : T_pM \to M$ is a local isometry between complete manifolds. Therefore, $\exp_p$ is a covering map. Since $\pi_1(T_pM) = 0 = \pi_1(M)$, $T_pM$ has to be diffeomorphic to $M$ by the universal property of a universal cover.

To prove that $\exp_p$ is a local isometry between complete manifolds, we put the pullback metric $\exp_p^*(g)$ on $T_pM$ defined as follows: for any $w_1, w_2 \in T_v(T_pM)$ where $v \in T_pM$ we have

$$\langle w_1, w_1 \rangle_{\exp_p^*(g)_v} := g_{\exp_p(v)}((d\exp_p)_v(w_1), (d\exp_p)_v(w_2)).$$

The definition of $\exp_p^*(g)$ makes sense because $d\exp_p$ does not have any kernel. Further, $\exp_p$ pulls back geodesic rays on $M$ to rays on $(T_pM, \exp_p^*(g))$. Since $(M, g)$ is geodesically complete so is $(T_pM, \exp_p^*(g))$. Therefore, Theorem 7.3.1 implies that $(T_pM, \exp_p^* g)$ is complete. By construction, $\exp_p$ is a local isometry. □

## 7.5 Curvature revisited

In this section, we will list some statements that shed light on how curvature measures the deviation of a manifold from Euclidean space.

**Proposition 7.5.1.** *Consider the Jacobi field*

$$J(t) = \gamma_s = (\mathrm{d}\exp_{\gamma(0)})_{t\gamma'(0)}(tw) \tag{7.5.1}$$

*along a unit speed geodesic $\gamma$ starting at $p \in M$ where $w \in T_{\gamma'(0)}(T_pM)$ is a unit vector. Then the Taylor expansion of $|J(t)|^2$ about $t = 0$ is given by*

$$|J(t)|^2 = t^2 - \frac{1}{3}\langle R(\gamma'(0), w)\gamma'(0), w\rangle t^4 + \epsilon(t) \tag{7.5.2}$$

*where* $\lim_{t\to 0} \frac{\epsilon(t)}{t^4} = 0$.

*Proof.* Left as an exercise. □

In particular, if $\sigma \subset T_p(M)$ is the subspace spanned by $w$ and $\gamma'(0)$ then

$$|J(t)| = t - \frac{1}{3}\kappa(\sigma)t^3 + \tilde{\epsilon}(t) \tag{7.5.3}$$

where $\lim_{t\to 0} \frac{\tilde{\epsilon}(t)}{t^3} = 0$.

**Remark 7.5.2.** Here $|J(t)|$ is the speed at which the geodesic starting in the $\gamma'(0)$ and $w$ start to deviate for a small time. As hinted before, the sectional curvature measures how fast the geodesics diverge.

**Proposition 7.5.3.** *In a normal coordinate $(x^1, \ldots, x^n)$ at a point $p \in M$ the Taylor expansion of the metric $g$ is*

$$g_{ij}(x^1, \ldots, x^n) = \delta_{ij} - \frac{1}{3}R_{ikjl}x^k x^l + o(\|x\|^3). \tag{7.5.4}$$

**Corollary 7.5.4.** *Writing the normal coordinate $(x^1, \ldots, x^n)$ as $tv$ where $v \in T_pM$ with $|v| = 1$ we get*

$$\det(g_{ij}(tv)) = 1 + \frac{1}{3}\operatorname{Ric}(v, v)t^2 + o(t^3). \tag{7.5.5}$$

# Chapter 8

# Calculus of variation

In this chapter, we will formalize the idea in Theorem 4.3.3 that geodesics are locally length-minimizing curves. Indeed, consider the length $L : PM \to \mathbb{R}$ as a function on $PM$, the space of paths on $M$. Then Theorem 4.3.3 means that geodesics are minimizer (critical points in general) of $L$. However, we have to make sense of the minimization problem by developing a calculus (tangents, derivatives, etc) on $PM$ a precursor to which is given in §7.1.[1] The issue is $PM$ is an infinite-dimensional manifold. Recall that we built calculus on $M$ by considering paths $\gamma$ starting at $p \in M$ and said that $T_pM$ consists of $\gamma'(0)$. Along the same line, our first goal is to develop a concept of paths (*variation*) in $PM$. We hope that an infinitesimal change of a path will give us a notion of a tangent at $\gamma \in PM$. Once we have tangent at a path $\gamma$, we can take *derivative* of functionals which allows us to talk about their critical points.[2]

On the other hand, in §8.3 we will see how studying the energy functional on $PM$ provides global (topological) information about the underlying manifold $M$ with some knowledge of local information (positivity of Ricci curvature).

## 8.1 First variation

**Definition 8.1.1.** A *variation of a path* $\gamma$ is a smooth map $F(s,t) : [-\epsilon, \epsilon] \times [0, l] \to M$ such that $F(s, \bullet)$ is a path in $M$ such that $F(0,t) = \gamma(t)$. We say that the variation is *proper* if $F(s, 0) = \gamma(0)$ and $F(s,l) = \gamma(l)$.

Note that a tangent at a point $p \in M$ is a tangent vector, so intuitively a tangent at $\gamma \in PM$

[1] A full-fledged analysis of $PM$ is beyond the scope of these notes. For a probabilistic study of $PM$, see [Str00].

[2] Note that paths are one-dimensional spaces and length is a one-dimensional measure. One can carry out a similar study on the space of $k$-dimensional spaces and study the critical points of $k$-dimensional measure. The study is beyond the scope of these notes. See [CM11] for details on minimal surfaces.

should be a vector field along $\gamma$. In fact, a tangent at $\gamma(t) = F(0,t)$ is a vector field $\partial_s F(0,t)$. Then a *variation* is a perturbation of $\gamma$ in the direction $dF(\partial_s)$. With this interpretation, proper variation is a perturbation that fixes endpoints of $\gamma$ i.e. $\partial_s F(s,0) = \partial_s F(s,l) = 0$. We say that $V(t) \coloneqq \partial_s F(0,t)$ is a *variational field* along $\gamma$.

**Proposition 8.1.2.** *Any smooth vector field $V(t)$ along $\gamma$ arises as a variational field.*

*Proof.* Note that $\gamma[0,l] \subset M$ is compact. Therefore, it is possible to find $\delta > 0$ such that $\exp_{\gamma(t)}$ is defined for all $t \in [0,l]$ and $v \in T_{\gamma(t)}M$ with $|v| < \delta$. Fix $\epsilon < \frac{\delta}{\max_{t\in[0,l]}|V(t)|}$. Then $F : [-\epsilon,\epsilon] \times [0,l] \to M$ with $F(s,t) \coloneqq \exp_{\gamma(t)}(sV(t))$ is a well defined variation of $\gamma$. On the other hand, $dF(\partial_s) = F_s = V(t)$. $\square$

**Remark 8.1.3.** One often allows $\gamma$ and $F_s$ to be piece-wise smooth i.e. there exist $0 = t_0 < t_1 < \cdots < t_k < l = t_{k+1}$ such that $\gamma$ and $F_s$ are smooth in $(t_i, t_{i+1})$. We can allow jump discontinuities on $t_i$. Unless otherwise stated, we use $t_i$ to note the point do jump discontinuity of a curve.

The functional $PM \to \mathbb{R}$ of interest to us are *length* $L(\gamma) = \int |\gamma'(t)| \mathrm{d}t$. However, it is easier to work with the *energy* $E(\gamma) = \int |\gamma'(t)|^2 \mathrm{d}t$. Note that $L(\gamma)^2 \leq lE(\gamma)$ In fact,

$$\begin{aligned}
0 &\leq \int_0^l \left( l|\gamma'| - \int |\gamma'| \right)^2 \\
&= \int_0^l \left( l^2|\gamma|^2 - 2l|\gamma'| \int |\gamma'| + \left( \int |\gamma'| \right)^2 \right) \\
&= 2l^2 \int |\gamma'|^2 - 2l \left( \int |\gamma| \right)^2 .
\end{aligned}$$

The equality holds if and only if $|\gamma'| = \frac{1}{l} \int |\gamma'| = \frac{L}{l}$ i.e. when the speed is constant or equivalently $\gamma$ is parametrized by arc length.

An advantage of working with energy is that we don't have to deal with parametrization when it comes to minimizers. Note that when $\gamma$ is a geodesic $lE(\gamma) = L(\gamma)^2 \leq L(c)^2 \leq lE(c)$ where $c$ is any curve joining $\gamma(0)$ and $l$. Therefore, any minimizer of $E$ is a minimizer of $L$. Further, the minimizer of $E$ is parametrized by arclength.

**Proposition 8.1.4 (First variation of energy).** *Suppose $F(s,t)$ is a variation of a piece-wise differentiable curve $\gamma$. Define the energy $E(s) \coloneqq E(F(s,\bullet))$. Let $V(t)$ be the variational field. Then*

$$\frac{1}{2}E'(0) = -\int_0^l \langle V, \boldsymbol{\nabla}_{\gamma'}\gamma' \rangle dt - \sum_{i=1}^k \langle V(t_i), \gamma'(t_i^+) - \gamma'(t_i^-) \rangle - \langle V(0), \gamma'(0) \rangle + \langle V(l), \gamma'(l) \rangle. \tag{8.1.1}$$

*Proof.* Write $F_t := \partial_t F(s,t)$ and $F_s := \partial_s F(s,t)$ Then $E(s) = \sum_{i=0}^{k} \int_{t_i}^{t_{i+1}} \langle F_t, F_t \rangle dt$.

To this end, using metric compatibility, $[F_s, F_t] = 0$ (see the proof of Lemma 4.3.1), we get

$$\partial_s \langle F_t, F_t \rangle = 2\langle \nabla_{F_s} F_t, F_t \rangle = 2\langle \nabla_{F_t} F_s, F_t \rangle = 2\partial_t \langle F_s, F_t \rangle - 2\langle F_s, \nabla_{F_t} F_t \rangle. \tag{8.1.2}$$

Now integrating (8.1.2), using the fundamental theorem of calculus, and $\nabla_{F_t} F_t = \nabla_{\gamma'}\gamma'$ at $s = 0$, we get (8.1.1). □

**Definition 8.1.5.** A curve $\gamma \in PM$ is a *critical point* of $E$ is $E'(0) = 0$ for any proper variation of $\gamma$.

**Theorem 8.1.6.** *A curve $\gamma \in PM$ is a critical point of $E$ if and only if $\gamma$ is a geodesic.*

*Proof.* If $\gamma$ is a geodesic, we can use (8.1.1) to see that $E'(0) = 0$. Since the variation is proper, to prove the converse, it suffices to prove $\nabla_{\gamma'}\gamma' = 0$ and there are no break points in $\gamma$ i.e. $\gamma'(t_i^+) = \gamma'(t_i^-)$ for all $1 \le i \le k$.

Let $w_i := \gamma'(t_i^+) - \gamma'(t_i^-)$. Say $w_i \neq 0$. Take bump functions $\eta_i$ supported around each point $t_i$ and consider the variational field $V := \sum_i \eta_i w_i$. Then $\frac{1}{2}E'(0) = \sum_{i=1}^{k} \eta_i w_i^2 > 0$ which is a contradiction. Therefore, $w_i = 0$.

Similarly, assume that $\nabla_{\gamma'}\gamma' \neq 0$. Without loss of generality we can consider a neighborhood $U$ that does not contain break points. Further, in some direction $e_j$ $\langle \nabla_{\gamma'}\gamma', e_j \rangle > 0$ in $U$. Choose a bump function $\eta$ that vanishes outside of $U$. Define a variational field $V = \eta e_j$. $E'(0) < 0$. Therefore, a critical point satisfies $\nabla_{\gamma'}\gamma' = 0$. □

## 8.2 Second variation

In this section, we will compute the second derivative of the energy functional to gain more information about the critical points just like how one characterizes critical points in calculus as local maxima, minima or saddle points using the second derivative tests.

**Proposition 8.2.1 (Second variation of energy).** *Suppose $F(s,t)$ is a smooth proper variation of a geodesic $\gamma$. Define the energy $E(s) := E(F(s,\bullet))$. Let $V(t)$ be the variational field. Then*

$$\frac{1}{2}E''(0) = -\int_0^l \langle \nabla_{\gamma'}\nabla_{\gamma'} V + R(\gamma', V)\gamma', V \rangle dt \tag{8.2.1}$$

*Proof.* Recall that $\frac{1}{2}E'(s) = \int \partial_t \langle F_s, F_t \rangle - \langle F_s, \nabla_{F_t} F_t \rangle dt$. Note that the first term is zero since $F_s(s,0) = F_s(s,0) = 0$ for proper variation. Therefore,

$$\begin{aligned} \frac{1}{2}E''(s) &= \int \partial_s \langle \nabla_{F_t} F_s, F_t \rangle dt \\ &= \int \langle \nabla_{F_s} F_s, \nabla_{F_t} F_t \rangle + \int \langle F_s, \nabla_{F_s} \nabla_{F_t} F_t \rangle \end{aligned}$$

Note that $\nabla_{F_t} F_t = 0$ at $s = 0$ since $F_t = \gamma'$ and $\gamma$ is a geodesic. Note that $R(\gamma', V)\gamma' = \nabla_V \nabla_{\gamma'} \gamma' - \nabla_{\gamma'} \nabla_V \gamma'$ since $[\gamma', V] = [F_t, F_s] = 0$ at $s = 0$. Therefore,

$$\begin{aligned}\frac{1}{2} E''(0) &= -\int \langle V, \nabla_V \nabla_{\gamma'} \gamma' \rangle \\ &= -\int \langle \nabla_{\gamma'} \nabla_V \gamma' + R(\gamma', V)\gamma', V \rangle \\ &= -\int \langle \nabla_{\gamma'} \nabla_{\gamma'} V + R(\gamma', V)\gamma', V \rangle\end{aligned}$$

□

**Remark 8.2.2.** If the variation field is a Jacobi field, then (8.2.1) implies that $E''(0) = 0$. It also follows because the curves along the variation are geodesics i.e. critical points of $E$, so $E' \equiv 0$.

## 8.3 Bonnet–Myers

So far, we have focused mostly on the local properties of manifolds which unsurprisingly turns out to be similar to that of $\mathbb{R}^n$. For instance, the divergence of geodesics starting in two directions is evident only when we leave a small neighborhood of the starting point. Then we start to see conjugate points, the existence of which measure the failure of $M$ being geodesically complete. However, it is when we look globally we see geometric and topological properties of $M$ different than that of $\mathbb{R}^n$. Since we only have microscopic snapshots of manifolds in the form of coordinate charts, all we can do to study global picture is to restrict certain properties on each snapshot and ask whether the microscopic snapshots together give meaningful information about the global picture. In this section, we will scratch the philosophy and study global diameter bounds on complete manifolds with lower bounds on the Ricci curvature.[3]

Here, $\operatorname{diam}(M) \coloneqq \sup_{p,q \in M} d(p, q)$. Further, $\operatorname{Ric} \geq \lambda$ means that $\operatorname{Ric}(V, V) \geq \lambda |V|^2$ for any $V \in \mathfrak{X}(M)$.

**Theorem 8.3.1 (Bonnet–Myers).** *Suppose $(M^n, g)$ is complete with a Ricci curvature bound $\operatorname{Ric} \geq \frac{n-1}{r^2}$. Then $M$ is compact and $\operatorname{diam}(M) \leq \pi r$. The inequality is sharp on $\mathbb{S}^n$.*

**Remark 8.3.2** (Historical)**.** Bonnet proved this theorem for a lower bound on the sectional curvature while Myers proved it for a lower bound on the Ricci curvature.

**Corollary 8.3.3.** *The fundamental group of $(M^n, g)$ is finite.*

*Proof.* Consider the universal cover $F : \tilde{M} \to M$ and a pullback metric $F^*g$. Then Theorem 8.3.1 applied to $(\tilde{M}, F^*g)$ implies that $\tilde{M}$ is compact. Since $\pi_1(\tilde{M}) = \{0\}$, the index of

[3] For more details on what the Ricci curvature has to say about global geometry of a manifold see [BCJ].

$F_*(\pi_1(\tilde{M}))$ in $\pi_1(M)$ is the $|\pi_1(M)|$. Note that the index of $F_*(\pi_1(\tilde{M}))$ is given by the number of sheets of the cover (cf. Proposition 1.32 in [Hat05]). On the other hand, the universal cover has finite sheets because $\tilde{M}$ is compact. In particular, $|\pi_1(M)|$ is finite. □

*Proof of Theorem 8.3.1.* Using Theorem 7.3.1, we know that $M$ is compact once we get a bound on the diameter. Since $M$ is geodesically complete, we get a diameter bound if we prove that any length-minimizing geodesic $\gamma : [0, l] \to M$ satisfies $L(\gamma) \le \pi r$.

Recall that a minimizer of length is also a minimizer of energy i.e. for any geodesic if $\gamma$ is length minimizing then $E''(0) \ge 0$. Therefore, it is enough to prove that $E''(0) \ge 0$ for all proper variation field $V$ implies that $L(\gamma) \le \pi r$.

Fix a orthonormal parallel frame $\{e_i\}_{1\le i\le n}$ such that $e_1 = \gamma'$ where $|\gamma'| = 1$. Define a proper variation $V_j := \sin\left(\frac{\pi t}{l}\right)e_j$ for $j \ge 2$. Note that $\boldsymbol{\nabla}_{\gamma'}V_j = \frac{\pi}{l}\cos\left(\frac{\pi t}{l}\right)e_j$ and $\boldsymbol{\nabla}_{\gamma'}\boldsymbol{\nabla}_{\gamma'}V_j = -\left(\frac{\pi}{l}\right)^2 V_j$. Moreover, $\langle R(\gamma', V_j)\gamma', V_j\rangle = \langle R(e_1, e_j)e_1, e_j\rangle \sin^2\frac{\pi t}{l}$. Let $E_j$ to mean the energy of curves generated by the variation $V_j$. Then applying (8.2.1) to $E_j$ and summing over $j$ we get implies that

$$0 \le \sum_{j=2}^{n} E_j''(0) = -\sum_{j=2}^{n}\left(\int_0^l -\left(\frac{\pi}{l}\right)^2 \sin^2\frac{\pi t}{l} + R_{1j1j}\sin^2\frac{\pi t}{l}dt\right).$$

Since $R_{1111} = 0$ we see that $\sum_{j=2}^{n} R_{1j1j} = \mathrm{Ric}_{11} \ge \frac{n-1}{r^2}$. Therefore,

$$\frac{n-1}{r^2}\int \sin^2\frac{\pi t}{l} \le (n-1)\frac{\pi^2}{l^2}\int \sin^2\frac{\pi t}{l}$$

whence $l \le \pi r$. Since $|\gamma'| = 1$, $l = L(\gamma)$.

Note that on $S^n$, $\sin\frac{\pi t}{l}e_j$ is a Jacobi field. Therefore, $E'' \equiv 0$ i.e. $l = \pi r$. □

## 8.4 Poincaré inequality

In this section, we will digress to talk about Poincaré inequality on $\mathbb{S}^n$ using the second variation formula for $E$. Note that the sectional curvature equals some constant $k > 0$ on $\mathbb{S}^n$. Consider a proper variation field $V = ue$ along a geodesic $\gamma : [0, l] \to M$ where $e$ is a parallel vector field and is orthogonal to $\gamma'$ and $u \in C^\infty(M)$. Then (8.2.1) becomes

$$\frac{1}{2}E'' = -\int_0^l (u''u + ku^2)\mathrm{d}t \tag{8.4.1}$$

Since $V$ is proper $u(0) = u(l) = 0$. Therefore, upon integration by parts $\int_0^l u''u = -\int_0^l |u'|^2$. Therefore, $\frac{1}{2}E'' = \int(|u'|^2 - ku^2)$. Since $\gamma$ is a geodesic $E'' \ge 0$ and therefore

$$k\int u^2 \le \int |u'|^2. \tag{8.4.2}$$

(8.4.2) is called Poincaré inequality.

One is often interested in finding the best constant $c$ such that $c\int u^2 \leq \int |u'|^2$. Note that we need Dirichlet condition $u(0) = u(l) = 0$ otherwise we could take a constant function $u$. In the case of sphere, since $\gamma$ is a great circle, we can use Fourier analysis to get a best constant. In fact, we have a Fourier series expansion $u(t) = \sum_i a_i \sin \frac{i\pi t}{l}$. We get a best constant when $i = 1$ and $c = \left(\frac{\pi}{l}\right)^2$. In general,

$$\int u^2 \leq \frac{l^2}{\pi^2} \int (u')^2.$$

In particular, $\frac{l^2}{\pi^2} \leq \frac{1}{k}$. For the round sphere $k = \frac{1}{r^2}$ where $r$ is the radius. In particular, we recover Theorem 8.3.1 in the case of $\mathbb{S}^n$.

One is also interested in the minimizer of $\lambda := \inf_u \frac{\int (u')^2}{\int u^2}$. Say that $\lambda$ is achieved for a function $u$ satisfying the Dirichlet condition. Suppose $v$ is another function satisfying the Dirichlet condition. For $s \in \mathbb{R}$, define

$$f(s) := \frac{\int (u' + sv')^2}{\int (u + sv)^2}.$$

Then assuming $\int u^2 = 1$ we see that

$$0 = f'(0) = 2\int u'v' - 2\int |u'|^2 \int uv = -2\int \left(u'' + u\int |u'|^2\right) v.$$

Therefore, $u'' = cu$ where $c = -\int |c'|^2$. In particular, $u$ has to be a sinusoidal function.

# Chapter 9

# Morse index theory

The Hessian of a function viewed as a bilinear form often sheds light on the nature of critical points of the function. Recall that in §8.2 we computed the second variation of the energy. In this section, we will look at the second variation of the energy in terms of the associated "Hessian." The number of negative eigenvalues of the Hessian turns out to encode the number conjugate points along a geodesic (critical point) counted with multiplicity.[1]

## 9.1 Review of linear algebra

Let $W$ be an $n$-dimensional $\mathbb{R}$-vector space with an inner product $\langle \bullet, \bullet \rangle$. Let $B : W \to W$ be a linear operator. The *transpose* $B^T$ of $B$ is defined by the action $\langle B^T v, w\rangle \coloneqq \langle v, Bw\rangle$ for any $v, w \in W$. We say that an operator is *symmetric* if $B^T = B$. We say that an operator is *positive (negative) definite* if $\langle v, Bv\rangle > 0$ ($< 0$) for all non-zero $v \in W$. An *eigenvalue* $\lambda$ of $B$ is defined to be a number such that $Bv_\lambda = \lambda v_\lambda$ for some non-zero $v_\lambda \in W$. For a symmetric operator $B$, there exists an orthonormal basis of eigenvectors $v_{\lambda_i}$ with eigenvalues $\lambda_i$ i.e. $B_{ij} = \lambda_i \delta_{ij}$ where $\lambda_i \leq \lambda_{i+1}$.

**Definition 9.1.1.** The *index* $\operatorname{ind}(B)$ of a symmetric operator $B$ is defined as the number of negative eigenvalues. The *nullity* $\operatorname{nullity}(B)$ of $B$ is the number of zero eigenvalues.

**Remark 9.1.2.** For a positive definite operator $\operatorname{ind}(B) = \operatorname{nullity}(B) = 0$.

**Lemma 9.1.3.** *Suppose $B$ is a symmetric linear operator with the smallest eigenvalue $\lambda_1$. Then for any non-zero $v \in W$*

$$\langle Bv, v\rangle \geq \lambda_1 \langle v, v\rangle \tag{9.1.1}$$

*and the equality holds if and only if $v = v_{\lambda_1}$.*

[1] We used the energy functional and its critical points to study $PM$. More generally, one can study functions on manifolds to probe the information about the manifolds. See [Mil63] for more details.

*Proof.* Diagonalize $B$ to prove (9.1.1). It is clear that the equality holds if $v = v_\lambda$. Conversely, assume that the equality holds for some $v \in W$ with $|v| = 1$. Then $B_{ij}v_iv_j = \lambda_1 v_iv_j$. Differentiating both sides with respect to $v_i$ we get $B_{ij}v_j = \lambda_1 v_j$. In particular, $v = v_{\lambda_1}$. □

Lemma 9.1.3 gives a variational characterization of the lowest eigenvalue of $B$ as $\lambda_1 = \min_{v\neq 0} \frac{\langle Bv,v\rangle}{\langle v,v\rangle}$. In the following Proposition, we give a variational characterization of the index.

**Proposition 9.1.4.** *Suppose $B$ is a symmetric linear operator on $W$. Then*

$$\operatorname{ind}(B) = \max\{\dim V \mid V \subset W, \langle Bv, v\rangle < 0 \textit{ for all } v \in V \setminus \{0\}\}. \tag{9.1.2}$$

*Proof.* Order the eigenvalues of $B$ as before. Define $V \coloneqq \operatorname{span}\{v_1, \dots v_{\operatorname{ind}(B)}\}$ where $Bv_i = \lambda_i v_i$. Note that $v = \sum_{i=1}^{\operatorname{ind}(B)} a_i v_i$ for any $v \in V$. Therefore, $\langle Bv, v\rangle = \sum a_i^2 \lambda_i < 0$. In particular, $\operatorname{ind}(B) \leq \max\{\dim V \mid V \subset W, \langle Bv, v\rangle < 0 \text{ for all } v \in V \setminus \{0\}\}$.

To prove the converse, consider $V \subset W$ such that $\dim V > I$. Define $P : V \to \mathbb{R}^{\operatorname{ind}(B)}$ to be "projection" map such that $P(v)_i = \langle v, v_i\rangle$. The rank nullity theorem implies that $\dim \operatorname{Im} + \dim \ker = \dim V$. Therefore, $\dim \ker > 0$ i.e. there is a non-zero $v \in V$ such that $P(v) = 0$. In particular, we can write $v = \sum_{i=\operatorname{ind}(B)+1}^{n} a_i v_i$. Therefore, $\langle Bv, v\rangle \geq 0$. Therefore, $\max\{\dim V \mid V \subset W, \langle Bv, v\rangle < 0 \text{ for all } v \in V \setminus \{0\}\} \leq \operatorname{ind}(B)$. □

## 9.2 Index lemma

In this section, we will get a lower bound on the smallest eigenvalue of the *Jacobi operator* defined below. Further, we will prove (cf. Lemma 9.2.2) that Jacobi fields are the critical points of the "Hessian" of the energy functional, a result we will use in the next section to compute the index and nullity of the index form.

For the rest of the section, fix a curve $\gamma : [0, l] \to M$. Consider the vector space $\mathcal{V}$ of piecewise differentiable variation fields along $\gamma$ vanishing at the endpoints. We write $0 = t_0 < t_1 < \cdots < t_k < l = t_{k+1}$ to denote the breakpoints of elements in $\mathcal{V}$. For any $V, W \in \mathcal{V}$, we define the inner product:

$$\langle V, W\rangle_{\mathcal{V}} \coloneqq \int_0^l \langle V, W\rangle.$$

With respect to the inner-product $\langle \bullet, \bullet\rangle_{\mathcal{V}}$ observe that the Jacobi operator $J$ is a symmetric linear operator on $\mathcal{V}$.

First, note that we can write $\langle JV, V\rangle_{\mathcal{V}} = \int |\boldsymbol{\nabla}_{\gamma'} V|^2 - \langle R(\gamma', V)\gamma', V\rangle$. Here, the first term is positive and the second term remains bounded in terms of the sectional curvature. Therefore, if $\lambda_1$ is the smallest eigenvalue of $J$ then $\lambda_1 \geq \min(-\kappa)$.

Instead of dealing with the Jacobi operator directly, we work with an associated bilinear form defined below.

**Definition 9.2.1.** The *index form* $I$ *along a curve* $\gamma$ is a symmetric bilinear form on any variation fields $V, W$ along $\gamma$ such that

$$I(V,W) := \int_0^l \langle V', W'\rangle - \langle R(\gamma', V)\gamma', W\rangle \tag{9.2.1}$$

where $V' := \boldsymbol{\nabla}_{\gamma'} V$.

Using the fundamental theorem of calculus for $\partial_t \langle V', W\rangle$ we see that

$$I(V,W) = -\int \langle JV, W\rangle + \langle V', W\rangle(l) - \langle V', W\rangle(0) - \sum_{j=1}^k \langle V'(t_j^+) - V'(t_j^-), W(t_j)\rangle. \tag{9.2.2}$$

We can think of the index form as the Hessian of the energy functional. We saw that $E''(0) = 0$ for a Jacobi field. In fact, the following lemma says that a Jacobi field variation along $\gamma$ minimizes the index form among variation fields that vanishes at $\gamma(0)$ and agree at $\gamma(l)$.

**Lemma 9.2.2 (Index lemma).** *Suppose* $\gamma : [0,l] \to M$ *is a geodesic without conjugate points on* $(0,l]$*. Suppose* $W$ *is a Jacobi field variation and* $V$ *is a peice-wise differentiable variation field along* $\gamma$ *such that* $W(0) = V(0) = 0$ *and* $W(l) = V(l)$*. Then*

$$I(V,V) \geq I(W,W)$$

*and the equality holds if and only if* $V \equiv W$ *on* $[0,l]$*.*

*Proof.* Since $\gamma$ has no conjugate point we can choose a basis of Jacobi field $\{v_i\}_{1\leq i\leq n}$ with the initial condition $V_i(0) = 0$ and $V_i'(0) = e_i$ where $e_i$ is a unit orthonormal basis at $\gamma(0)$. Here we used the fact that the initial condition of $J$ specifies it uniquely. Moreover, $v_i$ doesn't vanish on $(0,l]$ since there are no conjugate points. In fact, any linear combination of $V_i$ (which is also a Jacobi field) does not vanish on $(0,l]$. In particular, any vector field $V$ with $V(0) = 0$ can be written as a sum $f_i(t)V_i(t)$ for some functions $f_i$ along $\gamma$. Then the Jacobi field that equals $V$ at end point is given by $W = f_i(l)V_i(t)$.

Note that the Jacobi equation implies that

$$I(W,W) = \int_0^l (f_i f_j \langle V_i', V_j\rangle)' = f_i f_j \langle V_i', V_j\rangle(l).$$

Further,

$$\begin{aligned}(f_i f_j \langle V_i', V_j\rangle)' &= f_i' f_j \langle V_i', V_j\rangle + f_i f_j' \langle V_i', V_j\rangle + f_i f_j \langle V_i', V_j'\rangle + f_i f_j \langle V_i'', V_j\rangle \\ &= f_i' f_j \langle V_i, V_j'\rangle + f_i f_j' \langle V_i', V_j\rangle + f_i f_j \langle V_i', V_j'\rangle - f_i f_j \langle R(\gamma', V)i)\gamma', V_j\rangle\end{aligned}$$

where we used the Jacobi equation for $V_i$ to get $(\langle V_i', V_j\rangle - \langle V_i, V_j'\rangle)' = 0$. Since $\langle V_i', V_j\rangle - \langle V_i, V_j'\rangle = 0$ at 0 we see that $\langle V_i', V_j\rangle = \langle V_i, V_j'\rangle$ on $[0, l]$.

On the other hand,

$$\begin{aligned}
I(V,V) &= \int \langle f_i' V_i + f_i V_i', f_j' V_j + f_j V_j'\rangle - \langle R(\gamma', f_i V_i)\gamma', f_j V_j\rangle \\
&= \int f_i' f_j' \langle V_i, V_j\rangle + f_i' f_j \langle V_i, V_j'\rangle + f_i f_j' \langle V_i', V_j\rangle + f_i f_j \langle V_i', V_j'\rangle - f_i f_j \langle R(\gamma', V_i)\gamma', V_j\rangle \\
&= I(W,W) + \int f_i' f_j' \langle V_i, V_j\rangle \\
&\geq I(W,W)
\end{aligned}$$

since $\int f_i' f_j' \langle V_i, V_j\rangle = \int |V|^2 \geq 0$. And the equality holds if and only if $f_i' V_i = 0$ i.e. when $f_i' = 0$. Therefore, $V = f_i(l) V_i = W$. □

## 9.3 Morse index theorem

In this section, we will compute the index and nullity of the index form acting on $\mathcal{V}$.

**Theorem 9.3.1 (Morse Index Theorem).** *The index* $\operatorname{ind}(I)$ *of the index form is the number of conjugate points along* $\gamma$ *in* $[0, l]$ *counted with multiplicity.*

To this end, define $I_t$ to be the restriction of $I$ on $[0, t] \subset [0, l]$. The Morse index theorem amounts to understanding the function $\operatorname{ind} : [0, l] \to \mathbb{N}$ where $\operatorname{ind}(t) := \operatorname{ind}(I_t)$. Note that Proposition 9.1.4 implies that $\operatorname{ind}(t)$ is a non-decreasing function. As a first step, we prove in Lemma 9.3.2 the finiteness of $\operatorname{ind}(t)$ (as well as of nullity $\operatorname{nullity}(t) := \operatorname{nullity}(I_t)$). Then we will prove (9.3.1) and (9.3.2) to quantify the jump.

Consider $t_1, \ldots, t_k \in [0, l]$ such that $\gamma\big|_{(t_i, t_{i+1}]}$ has no conjugate points. Define a subspace $\mathcal{V}^+$ of the variation vector fields with elements $V \in \mathcal{V}^+$ such that $V(t_j) = 0$. Note that $I$ is positive definite on $\mathcal{V}^+$. In fact, Lemma 9.2.2 implies that $I\big|_{[t_i, t_{i+1}]}(V, V) \geq I\big|_{[t_i, t_{i+1}]}(W, W)$ where $W$ is a Jacobi field on $[t_i, t_{i+1}]$ that matches with $V$ at enpoints. Since $V(t_j) = 0$, $W \equiv 0$. Therefore, $I\big|_{[t_i, t_{i+1}]}(V, V) > 0$ for non-trivial $V \in \mathcal{V}^+$. Therefore, only the complement of $\mathcal{V}^+$ contributes to the non-negative-definiteness of $I$. The next lemma shows that $\mathcal{V}^+$ has finite codimension in $\mathcal{V}$.

**Lemma 9.3.2.** *The codimension of* $\mathcal{V}^+$ *is at most* $nk$ *i.e.* $\operatorname{ind}(t) + \operatorname{nullity}(t) \leq nk$.

*Proof.* For the sake of contradiction assume that there is an $nk+1$-dimensional vector space $\mathcal{W}$ such that $I(V, V) \leq 0$ for any $V \in \mathcal{W} \setminus \{0\}$. Define a linear map $\psi : \mathcal{W} \to \mathbb{R}^{nk}$ as $V \mapsto (V(t_1), \ldots, V(t_k))$. The rank-nullity theorem implies that $\dim(\operatorname{Im}\psi) + \dim(\ker\psi) = nk + 1$. Therefore, $\dim\ker \geq 1$ i.e. there exists $V \neq 0$ such that $V \in \mathcal{W} \cap \mathcal{V}^+$ which is a contradiction. □

Since $\operatorname{ind}(t)$ is finite, it is a step function. Therefore, proving Theorem 9.3.1 amounts to proving that

- the jumps of $\operatorname{ind}(t)$ are at conjugate points and
- the jumps are the multiplicities of the conjugate points.

To this end, we claim that if $\operatorname{nullity}(t) > 0$ then there exists $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0)$,

$$\operatorname{ind}(t+\epsilon) = \operatorname{ind}(t) + \operatorname{nullity}(t). \tag{9.3.1}$$

We will break up the proof of (9.3.1) into two inequalities (cf. Lemma 9.3.5 and Lemma 9.3.7).

Further, given $t \in [0, l]$ we claim that (cf. Lemma 9.3.8) there exists $\epsilon_0$ such that for all $\epsilon \in (0, \epsilon_0)$

$$\operatorname{ind}(t-\epsilon) = \operatorname{ind}(t). \tag{9.3.2}$$

Note that (9.3.2) implies "continuity" of the index and (9.3.1) says that the jump of width $\operatorname{nullity}(t)$ occurs at conjugate points. Since $\operatorname{nullity}(t)$ is the multiplicity of conjugate point, we get Theorem 9.3.1.

**Remark 9.3.3.** In a finite-dimensional setting, the claim in (9.3.1) and (9.3.2) follow from continuity but they need some justification in our infinite-dimensional setting. First, the index form acts on different space for different time so the continuity of operators don't even make sense. Second, the interaction between nullity and index at different $t$ in infinite-dimensional setting is different than that in a finite dimensional setting. For instance, define $A(t) := \begin{pmatrix} 1 & 0 \\ 0 & 1-t \end{pmatrix}$. Note that (9.3.1) and (9.3.2) hold by continuity of $(1-t)$. In contrast, consider a family of infinite-dimensional matrix $A(t) := (\frac{1}{i} - t)\delta_{ij}$. Note that $\operatorname{ind} A(\epsilon) \neq \operatorname{ind} A(0) + \operatorname{nullity} A(0)$.

**Proposition 9.3.4.** *The kernel of $I$ consists of Jacobi fields that vanish at endpoints of $\gamma$.*

*Proof.* Note that $V \in \ker(I)$ if and only if $I(V, W) = 0$ for all $W$. Using bump function technique in (9.2.2) similar to that in Theorem 8.1.6 we can prove that $JV \equiv 0$ and $V$ vanishes at endpoints. $\square$

**Lemma 9.3.5.** *Suppose* $\operatorname{nullity}(t) > 0$*. Then there exists $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0)$ we have* $\operatorname{ind}(t+\epsilon) \geq \operatorname{ind}(t) + \operatorname{nullity}(t)$.

*Proof.* Pick $\{t_i\}_{1 \leq i \leq k}$ such that $\gamma\big|_{(t_i, t_{i+1}]}$ has no conjugate points. Let $\mathcal{V}(t)$ denote the space of variation vector fields along $\gamma\big|_{[0,t]}$ that vanishes at endpoints and let $\mathcal{W}(t)$ denote the space of variation vector field along $\gamma\big|_{[0,t]}$ where $I$ is non-negative definite. To compute $\operatorname{ind}(t+\epsilon)$, we need to look at the dimension of $\mathcal{W}(t+\epsilon)$.

To this end, extend $V \in \mathcal{V}(t)$ using Jacobi fields as:

$$\tilde{V}(s) = \begin{cases} V(s) & s \leq t_k \\ J & \text{on } [t_k, t+\epsilon] \\ 0 & \text{at } t+\epsilon \end{cases}$$

where $J$ is a Jacobi field that is $J(t+\epsilon) = 0$ and $J(t_k) = V(t_k)$. Note that $V \mapsto \tilde{V}$ is a linear map. Further, $I(\tilde{V}, \tilde{V}) \leq I(V, V)$. In fact, they are equal on $[0, t_k]$. On the other hand, extend $V$ on $[t_k, t+\epsilon]$ by 0. Then $\tilde{V}$ is a Jacobi field that matches $V$ at end points on $[t_k, t+\epsilon]$. Therefore, Lemma 9.2.2 implies that $I(\tilde{V}, \tilde{V}) \leq I(V, V)$ on $[t_k, t+\epsilon]$. The equality holds if and only if $V$ is a Jacobi field i.e. $V \equiv 0$ on $[t, t+\epsilon]$.

Therefore, if $V \in \mathcal{W}(t)$ then $\tilde{V} \in \mathcal{W}(t+\epsilon)$ i.e.

$$\operatorname{ind}(t+\epsilon) + \operatorname{nullity}(t+\epsilon) \geq \operatorname{ind}(t) + \operatorname{nullity}(t) \tag{9.3.3}$$

which is close to what we want. However, if we prove that $I(\tilde{V}, \tilde{V}) = 0$ if and only if $\tilde{V} = 0$ $\tilde{V} \in \mathcal{W}(t+\epsilon)$ then we get $\operatorname{ind}(t+\epsilon) \geq \operatorname{ind}(t) + \operatorname{nullity}(t)$. In fact, $0 = I(\tilde{V}, \tilde{V}) \leq I(V, V) \leq 0$ implies that $I(V, V) = 0$. Therefore, Proposition 9.3.4 implies that $V$ is a Jacobi field. Since $V \equiv 0$ on $[t_k, t]$ the same holds true on $[0, t]$. Therefore, $\tilde{V} \equiv 0$. □

**Corollary 9.3.6.** *If $\gamma(t)$ is the first conjugate point along $\gamma$ then $\gamma$ is no longer minimizing after $t$.*

*Proof.* The index form measures the second variation of the energy. □

**Lemma 9.3.7.** *Given some $t \in [0, l]$ there exists $\epsilon > 0$ such that* $\operatorname{ind}(t+\epsilon) \leq \operatorname{ind}(t) + \operatorname{nullity}(t)$.

**Lemma 9.3.8.** *Given $t$ there exists $\epsilon_0 > 0$ such that for any $\epsilon \in (0, \epsilon_0)$ we have* $\operatorname{ind}(t-\epsilon) = \operatorname{ind}(t)$.

**Remark 9.3.9.** The Lemma 9.3.5 says that the negative eigenvalues of an operator at $t+\epsilon$ is controlled by the zero eigenvalues appearing at $t$ and the behavior we were concerned about for $A(t) := (\frac{1}{i} - t)\delta_{ij}$ does not happen. The idea of the proof of Lemma 9.3.5 is to boil down the infinite dimensional space of variation vector field to a fixed finite-dimensional space and use continuity argument.

In fact, fix some $t \in [0, l]$. Then pick $\{t_i\}_{1 \leq i \leq k}$ such that $t_k \leq t$ and $\gamma\big|_{(t_i, t_{i+1}]}$ does not have any conjugate points. Define $\mathcal{V}^-(t)$ to be the space of piece-wise Jacobi field with breaks at $t_i$ and vanishing at 0 and $t$. Define $\mathcal{V}^+(t)$ be the space of variation fields with $V(t_k) = 0$. Note that the space of variation vector field can be decomposed as $\mathcal{V}^+ \oplus \mathcal{V}^-$ since $\mathcal{V}^+ \cap \mathcal{V}^- = \emptyset$ and $\operatorname{span}\{\mathcal{V}^+, \mathcal{V}^-\} = \mathcal{V}$. Therefore, any variation vector field $V$ can be written as $V = V^+ + V^-$ where $V^\pm \in \mathcal{V}^\pm(t)$. Since (9.2.2) implies that $I(V^+, V^-) = 0$, we have

$$I(V, V) = I(V^+, V^+) + I(V^-, V^-). \tag{9.3.4}$$

Further, $I(V^+, V^+) > 0$ for non-trivial $V^+ \in \mathcal{V}^+$, we see that the index of $I$ restricted to $\mathcal{V}$ is same as that of $I$ restricted to $\mathcal{V}^-$.

Since the Jacobi field in a normal neighborhood are specified by its value at endpoints, we see that $\mathcal{V}^-(t) \cong T_{\gamma(t_1)}M \oplus T_{\gamma(t_2)}M \oplus \cdots \oplus T_{\gamma(t_k)}M =: S$.

Further, $\mathcal{V}^-(t)$ depends continuously on $t$. In fact, the space of Jacobi fields before $t_k$ remain the same under perturbation, and since the solution to the Jacobi equation vary continuously on $t$ the fields on $[t_k, t]$ depend continuously on $t$.

Therefore, $I(t) := I\big|_{\mathcal{V}^-(t)}$ is an $nk \times nk$ matrix acting on $S$. Note that $I(t)$ is continuous on $t$. Therefore, the eigenvalues of $I(t)$ are continuous in $t$. Then by the continuity argument, both Lemma 9.3.7 and Lemma 9.3.8 follow immediately.

# Chapter 10

# Comparison geometry

In Riemannian geometry, the hyperbolic spaces, Euclidean space, and sphere are the models with constant curvatures we know a lot about. In practice, we restrict the property of curvature of a manifold and ask how other geometric quantities compare with those of the model manifolds. For instance, one can study geometric quantities like diameter, volume, eigenvalues of $\Delta$, and harmonic functions on manifolds with $\mathrm{Ric} \geq 0$ and compare it with $\mathbb{R}^n$ where $\mathrm{Ric} = 0$ and with $\mathbb{S}^n$ where $\mathrm{Ric} = 1$.[1] For more details on comparison geometry see [CE75].

## 10.1 Bochner's identity

The answers to a lot of questions in comparison geometry involve Bochner's identity:

$$\frac{1}{2}\Delta|\boldsymbol{\nabla} u|^2 = |\mathrm{Hess}_u|^2 + \langle \boldsymbol{\nabla} u, \boldsymbol{\nabla}\Delta u\rangle + \mathrm{Ric}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} u). \tag{10.1.1}$$

*Proof.* First we prove

$$\frac{1}{2}\Delta|\boldsymbol{\nabla} u|^2 = |\mathrm{Hess}_u|^2 + \langle \boldsymbol{\nabla} u, \Delta\boldsymbol{\nabla} u\rangle. \tag{10.1.2}$$

Then we prove the commutator relation between $\Delta$ and $\boldsymbol{\nabla}$ :[2]

$$[\Delta, \boldsymbol{\nabla}]u = \mathrm{Ric}(\boldsymbol{\nabla} u). \tag{10.1.3}$$

Note that (10.1.2) and (10.1.3) imply (10.1.1).

[1] It is out of scope of these notes but one can ask if a bound on Ricci curvature imposes any restriction on topological properties. See a forthcoming book [BCJ] for more detail.

[2] The commutator of differential operators has lower degree than that of operators we are commuting. In our case, $[\Delta, \boldsymbol{\nabla}]$ is a zeroth order operator.

To this end, by taking geodesic normal coordinates, we choose an orthonormal frame $e_i$ such that $\boldsymbol{\nabla}_{e_i} e_j = 0$ at a point. Then

$$\begin{aligned}\Delta|\boldsymbol{\nabla} u|^2 = \langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla}|\boldsymbol{\nabla} u|^2, e_i\rangle &= e_i\langle \boldsymbol{\nabla}|\boldsymbol{\nabla} u|^2, e_i\rangle = e_i(e_i(|\boldsymbol{\nabla} u|^2))\\ &= 2e_i[\langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, \boldsymbol{\nabla} u\rangle]\\ &= 2[\langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, \boldsymbol{\nabla} u\rangle + \langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u\rangle]\\ &= 2\langle \boldsymbol{\nabla} u, \Delta\boldsymbol{\nabla} u\rangle + |\mathrm{Hess}_u|^2\end{aligned}$$

where we used $|\mathrm{Hess}_u|^2 = |\langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, e_j\rangle|^2 = |\boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u|^2$ in the last line. This gives us (10.1.2). On the other hand,

$$\langle \boldsymbol{\nabla} u, \boldsymbol{\nabla}\Delta u\rangle = \boldsymbol{\nabla} u(\Delta u) = \boldsymbol{\nabla} u\langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, e_i\rangle = (\langle \boldsymbol{\nabla}_{\boldsymbol{\nabla} u}\boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, e_i\rangle + \langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, \boldsymbol{\nabla}_{\boldsymbol{\nabla}_u} e_i\rangle). \tag{10.1.4}$$

Note that the second summand in (10.1.4) is 0. Now the idea is to switch $\boldsymbol{\nabla}_{\boldsymbol{\nabla} u}$ and $\boldsymbol{\nabla}_{e_i}$. Recall that $R$ measures how well covariant differentiation along two vector fields commute. In particular,

$$R(e_i, \boldsymbol{\nabla} u)\boldsymbol{\nabla} u = \boldsymbol{\nabla}_{\boldsymbol{\nabla} u}\boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u - \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla}_{\boldsymbol{\nabla} u}\boldsymbol{\nabla} u + \boldsymbol{\nabla}_{[e_i,\boldsymbol{\nabla} u]}\boldsymbol{\nabla} u,$$

where$[e_i, \boldsymbol{\nabla} u] = \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u - \boldsymbol{\nabla}_{\boldsymbol{\nabla} u} e_i = \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u$ since the second term is 0. Therefore, (10.1.4) becomes

$$\langle \boldsymbol{\nabla}_{\boldsymbol{\nabla} u}\boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, e_i\rangle = \langle R(e_i, \boldsymbol{\nabla} u)\boldsymbol{\nabla} u, e_i\rangle + \langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla}_{\boldsymbol{\nabla} u}\boldsymbol{\nabla} u, e_i\rangle - \langle \boldsymbol{\nabla}_{\boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u}\boldsymbol{\nabla} u, e_i\rangle. \tag{10.1.5}$$

Using the symmetry of $R$, we see that the first summand in (10.1.5) contributes $-\operatorname{Ric}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} u)$. However, the second summand contributes 0 since we chose geodesic frames. Finally, using the symmetry of Hessian for the third summand in (10.1.5), we get

$$\langle \boldsymbol{\nabla}_{\boldsymbol{\nabla} u}\boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, e_i\rangle = -\operatorname{Ric}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} u) + e_i\langle \boldsymbol{\nabla}_{\boldsymbol{\nabla} u}\boldsymbol{\nabla} u, e_i\rangle - \langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u\rangle. \tag{10.1.6}$$

Using the symmetry of Hessian in the second summand of (10.1.6), we get that $e_i\langle \boldsymbol{\nabla}_{\boldsymbol{\nabla} u}\boldsymbol{\nabla} u, e_i\rangle = e_i\langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, \boldsymbol{\nabla} u\rangle = \langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, \boldsymbol{\nabla} u\rangle + \langle \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, \boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u\rangle$. Therefore, (10.1.6) becomes

$$\langle \boldsymbol{\nabla}_{\boldsymbol{\nabla} u}\boldsymbol{\nabla}_{e_i}\boldsymbol{\nabla} u, e_i\rangle = -\operatorname{Ric}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} u) + \langle \Delta\boldsymbol{\nabla} u, \boldsymbol{\nabla} u\rangle. \tag{10.1.7}$$

Combining (10.1.7) with (10.1.4), we get (10.1.3). □

## 10.2 Eigenvalue estimates

In this section, we will see an application of Bochner's identity (10.1.1) on the estimation of the eigenvalues of $\Delta$ on $M$, a closed compact manifold without boundary. Note that $\Delta$ acts on $L^2$ space of functions on $M$ as a self-adjoint operator. Therefore, $\Delta$ has a discrete spectrum $\{\lambda_i\}_{i=1}^{\infty}$ i.e. there exists a sequence $\{u_i \in L^2\}$ such that $\Delta u_i = -\lambda_i u_i$.

We are interested in getting an estimate on $\lambda_1$ as an application of which we get the Poincaré inequality[3]

$$\int u^2 \leq C \int |\nabla u|^2 \tag{10.2.1}$$

for $u$ in the space perpendicular to the quotient functions i.e. $\int u = 0$. Define the Rayleigh constant $R_u := \frac{\int |\nabla u|^2}{\int u^2}$. Note that $R_u = 0$ for constant functions. To get bounds interesting bounds, we look at the functions $u$ orthogonal to 1 i.e. $\int u = 0$. Define

$$\lambda_1(M, g) \coloneqq \inf_{\int u = 0} \frac{\int |\nabla u|^2}{\int u^2}. \tag{10.2.2}$$

Note that $\lambda_1(M, g)$ is realized for $u_1$ such that $\Delta u_1 = -\lambda_1 u_1$ where $\lambda_1$ is the lowest eigenvalue. Note that getting a lower bound on $\lambda_1$ gives us (10.2.1).

**Example 10.2.1.** Consider $M = \mathbb{S}^1$. Then $u_1 = \sin\theta$ or $\cos\theta$ and $\lambda_1 = 1$. Therefore, for $u$ such that $\int u = 0$, we get

$$\int u^2 \leq \int |\nabla u|^2.$$

A bound on $\lambda_1$ helps us to study the behaviour of solutions to the heat equation:

$$u_t = u_{\theta\theta}. \tag{10.2.3}$$

Note that any solution to (10.2.3) on $\mathbb{S}^1$ has the form

$$u(\theta, t) = \sum_k a_k e^{-k^2 t} \sin(k\theta) + \sum_k b_k e^{-k^2 t} \cos k\theta.$$

Therefore, the decay of $u$ is governed mostly by the term $k = 1$. In general, we expect to decompose solutions of the heat equation in terms of the eigenfunctions and eigenvalues. Further, the bound for $\lambda_1$ will give us an estimate for the rate of decay of the solution.

Without further ado, we will prove a lower bound on $\lambda_1$ under certain geometric conditions.

**Theorem 10.2.2 (Lichnerowicz).** *Suppose $n \geq 2$ and $M$ is a complete $n$ dimensional manifold without boundary such that $\mathrm{Ric}_M \geq (n-1)$ then $\lambda_1 \geq n$. Further, the inequality is sharp on $\mathbb{S}^n$.*

*Proof.* Suppose $u_1$ is an eigenfunction associated to $\lambda_1$ i.e. $\Delta u_1 = -\lambda_1 u_1$. Applying Bochner's formula to $u_1$, we get

$$\begin{aligned} \frac{1}{2}\Delta|\nabla u_1|^2 &= |\mathrm{Hess}_{u_1}|^2 + \langle \nabla u_1, \nabla \Delta u_1 \rangle + \mathrm{Ric}(\nabla u_1, \nabla u_1) \\ &\geq |\mathrm{Hess}_{u_1}|^2 - \lambda_1 |\nabla u_1|^2 + (n-1)|\nabla u_1|^2. \end{aligned}$$

[3]The Poincaré inequality gives a bound on the variance of a function in terms of the gradient energy.

Recall that Bonnet–Myers theorem implies that $M$ is compact without boundary. Therefore, the Stokes theorem implies that $\int \Delta u = 0$. In particular,

$$0 = \frac{1}{2}\int \Delta|\boldsymbol{\nabla} u_1|^2 \geq \int |\mathrm{Hess}_{u_1}|^2 + (n-1-\lambda_1)|\boldsymbol{\nabla} u_1|^2. \tag{10.2.4}$$

Observe that (10.2.4) already implies $\lambda_1 \geq n-1$. To improve the lower bound to $n$, we diagonalize $\mathrm{Hess}_{u_1}$ with its eigenvalues say $\sigma_i$. Note that $|\mathrm{Tr}(\mathrm{Hess}_{u_1})|^2 \leq n|\mathrm{Hess}_{u_1}|^2$.[4] In fact, let $v = (\sigma_1, \ldots, \sigma_n)$ and $w = (1, \ldots, 1)$. Then Cauchy–Schwarz implies that $|v \cdot w| \leq |v||w|$. In particular, $|\mathrm{Hess}_u|^2 = \sum \sigma_i \geq \frac{1}{n}(\sum \sigma_i)^2$. on the other hand, $\Delta u_1 = \sum \sigma_i$. Therefore,

$$|\mathrm{Hess}_{u_1}|^2 \geq \frac{1}{n}(\Delta u_1)^2 = \frac{1}{n}(\lambda_1^2 u_1^2). \tag{10.2.5}$$

Therefore, (10.2.4) and (10.2.5) together with (10.2.2) imply that

$$(\lambda_1 - n + 1)\int |\boldsymbol{\nabla} u_1|^2 \geq \frac{\lambda_1}{n}\int |\boldsymbol{\nabla} u_1|^2.$$

In particular, we get $\lambda_1(1-\frac{1}{n}) \geq (n-1)$ which implies $\lambda_1 \geq n$ for $n \geq 2$.

Let's first prove that the inequality is sharp on $\mathbb{S}^n$. Note that

$$\Delta x_k = \langle \boldsymbol{\nabla}_{e_i}\partial_k^T, e_i\rangle = -\langle \boldsymbol{\nabla}_{e_i}\partial_k^\perp, e_i\rangle = \langle \partial_k^\perp, \boldsymbol{\nabla}_{e_i}^\perp e_i\rangle = \langle \partial_k, \boldsymbol{\nabla}_{e_i}^\perp e_i\rangle$$

which is the mean curvature $H$ on $\mathbb{S}^n$. Recall that $H = -n\vec{x}$ on $\mathbb{S}^n$. Therefore, $\Delta x_k = -nx_k$. In particular, $\lambda_1 \leq n$. However, we proved that $\lambda_1 \geq n$. Therefore, $\lambda_1(\mathbb{S}^n) = n$. □

**Remark 10.2.3** (Historical)**.** Bochner initially used (10.1.1) for forms to prove that if $\mathrm{Ric} > 0$ then the rank of de Rham cohomology is at most $n$ where the equality holds for the torus.

## 10.3 Obata's theorem

In this section, we will prove the equality $\lambda_1(M) = n$ in Theorem 10.2.2 holds if and only if $M = \mathbb{S}^n$ up to isometry.

First, note that we get equality in (10.2.5). In particular, $\mathrm{Hess}_u = cg$ for some $c$. Since $\mathrm{Tr}\,\mathrm{Hess}_u = \Delta u$ and $\lambda_1 = n$ we see that $c = -u$ i.e. $\mathrm{Hess}_u = -ug$. The following theorem due to Obata [Oba62] implies that $M = \mathbb{S}^n$.

**Theorem 10.3.1.** *Suppose $n \geq 2$ and $M$ is a complete $n$ dimensional manifold without boundary such that $\mathrm{Ric}_M \geq (n-1)$. Then $M$ is isometrically a sphere if there exists a non-trivial function such that*

$$\mathrm{Hess}_u = -ug. \tag{10.3.1}$$

[4] In fact, for any $n \times n$ symmetric matrix $B$ we have $|\mathrm{Tr}(B)|^2 \leq n|B|^2$. To see that, observe $|B|^2 = \frac{\mathrm{Tr}(B)^2}{n} + |B_0|^2$ where $B_0 := B - \frac{\mathrm{Tr}\,B}{n}\delta_{ij}$ is the trace-free part of $B$. In the case when $B$ does not have a full rank, we can even prove that $|\mathrm{Tr}\,B|^2 \leq (\mathrm{rank}\,B)|B|^2$.

*Proof.* Suppose $u$ satisfies (10.3.1) then for any $X \in \mathfrak{X}(M)$ we have

$$\begin{aligned} X(u^2 + |\nabla u|^2) &= 2\langle X, u\nabla u\rangle + 2\langle \nabla_X \nabla u, \nabla u\rangle \\ &= 2u\langle X, \nabla u\rangle + 2\operatorname{Hess}_u(X, \nabla u) \\ &= 2u\langle X, \nabla u\rangle - 2u\langle X, \nabla u\rangle = 0. \end{aligned}$$

Therefore, $u^2 + |\nabla u|^2$ is constant. Note that Bonnet–Myeres Theorem 8.3.1 implies that $M$ is closed and diam $\leq \pi$. In particular, $u$ attains a maximum and a minimum. Dividing $u$ by $\max u$, we can set the constant to be 1. Therefore, $u^2 + |\nabla u|^2 = 1$. In particular, the maximum of $u$ is 1 and the minimum is $-1$.

To this end, let $p \in M$ be such that $u(p) = -1$. Consider a unit speed geodesic $\gamma$ starting at $p$. Define $f(t) := u(\gamma(t))$. Then $f' = \langle \nabla u, \gamma'\rangle$ and $f'' = \langle \nabla'_\gamma \nabla u, \gamma'\rangle = \operatorname{Hess}_u(\gamma', \gamma')$. Therefore, (10.3.1) implies that

$$f'' = -f \tag{10.3.2}$$

which implies that $f(t) = a\sin t + b\cos t$. Since $f(0) = -1$ is a minimum we see that $f(t) = -\cos t$. Therefore, along the geodesics $f$ varies from $-1$ to 1 as $t$ runs from 0 to $\pi$.

Note that (10.3.1) implies these critical points are non-singular and therefore isolated. On the other hand, $M_\pi := \{p \in M \mid u(p) = 1\}$ is a continuous image of the sphere $|v| = \pi$ in $T_pM$, so $M_\pi$ consists of one point say $p_+$ and since diam $M \leq \pi$ we see that $\gamma(\pi) = p_+$. In particular, all geodesics starting at $p_-$ such that $u(p_-) = -1$ meet at $p_+$. One can similarly show that there is one minimum. Note that the geodesics starting at the minimum point are uniquely minimizing for all $t < \pi$ and all of them meet at the maximum point i.e. $M$ is diffeomorphic to a sphere.

To prove that it is also isometric to the sphere, we look at $g$ in the exponential coordinates and Jacobi fields. Consider a family of geodesics $F(t, s) := \exp_p(t(v + sw))$ where $v \perp w$. Then $(d\exp_p)_v(w) = F_s(1, 0)$. Further, $\nabla_{F_t} F_t = 0$ and $\nabla_{F_s} F_t = \nabla_{F_t} F_s$. Recall from Example 3 we have $|F_s|^2 = |w|^2 \sin^2 |v|$ on a sphere. It suffices to prove that the same holds true on $M$.

To this end, note that $u$ is radial, so $\nabla u$ is a radial vector. In particular, we can write $F_t = \frac{\nabla u}{|\nabla u|}|F_t|$. Therefore,

$$\begin{aligned} \partial_t |F_s|^2 &= 2\langle \nabla_{F_t} F_s, F_s\rangle \\ &= 2\langle \nabla_{F_s} F_t, F_s\rangle \\ &= 2\langle \nabla_{F_s} \frac{\nabla u}{|\nabla u|}|F_t|, F_s\rangle \\ &= 2\frac{|F_t|}{|\nabla u|}\operatorname{Hess}_u(F_s, F_s) \\ &= -2u\frac{|F_t|}{|\nabla u|}|F_s|^2. \end{aligned}$$

Integrating with respect to $t$ we get that $|F_s|^2 = |w|^2 \sin^2 |v|$. $\square$

**Remark 10.3.2.** We actually have if and only if condition in Theorem 10.3.1. In fact, write $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ as the level set of the function $|x|^2$. Then the height function $u$ that sends $x \mapsto x^1$ satisfies $\mathrm{Hess}_u = -ug$.

## 10.4 Laplacian comparison

Consider $\mathbb{R}^n$ equipped with the Euclidean metric. Let $r(x)$ be the distance function from the origin to $x$. Except at the origin, we have $|\boldsymbol{\nabla} r| = 1$ and $\Delta r^2 = 2n$. Since $\Delta r^2 = 2r\Delta r + 2|\boldsymbol{\nabla} r|^2$, we get $\Delta r = \frac{n-1}{r}$. In this section, we will use Bochner's formula to prove a similar result on a manifold.

**Theorem 10.4.1.** *Consider a manifold $M^n$ such that* $\mathrm{Ric} \geq 0$*. Let $r(x) := d(x,p)$ for some $p \in M$. Then wherever $r$ is smooth*

$$\Delta r \leq \frac{n-1}{r} \qquad \text{and} \qquad \Delta r^2 \leq 2n. \tag{10.4.1}$$

*Proof.* Let $\gamma(t)$ be a length minimizing geodesic connecting $x$ and $p$ such that $|\gamma'| = 1$. Then Gauss's lemma 4.3.1 implies that $\boldsymbol{\nabla} r = \gamma'$. In fact, say that $\tilde{r}$ is a radial coordinate in $T_pM$. Let $d_{T_pM}(0, \bullet)$ be the distance function in $T_pM$. Note that $\boldsymbol{\nabla} r = d\exp_p(\boldsymbol{\nabla} d_{T_pM})$. Check that $\boldsymbol{\nabla} d_{T_pM} = \partial_{\tilde{r}}$. Since $\exp_p$ is a radial isometry, we see that $\boldsymbol{\nabla} r = d\exp_p(\partial_{\tilde{r}}) = \gamma'$. Therefore, $|\boldsymbol{\nabla} r| = 1$ which together with (10.1.1) implies that

$$0 = \frac{1}{2}\Delta|\boldsymbol{\nabla} r|^2 = |\mathrm{Hess}_r|^2 + \langle \boldsymbol{\nabla} r, \boldsymbol{\nabla}\Delta r\rangle + \mathrm{Ric}(\boldsymbol{\nabla} r, \boldsymbol{\nabla} r). \tag{10.4.2}$$

To this end, for any $Y \in \mathfrak{X}(M)$ we have

$$\mathrm{Hess}_r(\boldsymbol{\nabla} r, Y) = \langle \boldsymbol{\nabla}_{\boldsymbol{\nabla} r}\boldsymbol{\nabla} r, Y\rangle = \langle \boldsymbol{\nabla}_{\gamma'}\gamma', Y\rangle = 0.$$

In particular, $\operatorname{rank}\mathrm{Hess}_r \leq n-1$. Therefore, the remark in footnote 4 implies that $|\mathrm{Hess}_r|^2 \geq \frac{(\Delta r)^2}{n-1}$ which together with $\mathrm{Ric} \geq 0$ and (10.4.2) implies that

$$0 \geq \frac{(\Delta r)^2}{n-1} + \langle \boldsymbol{\nabla} r, \boldsymbol{\nabla}\Delta r\rangle. \tag{10.4.3}$$

Setting $f(t) = \Delta r(\gamma(t))$ we get the *Riccati equation*

$$0 \geq \frac{f^2}{n-1} + f'. \tag{10.4.4}$$

Note that (10.4.4) implies that $f$ is decreasing unless $f = 0$. We can assume that $f(s) > 0$ for all $s < t$ since $f(s) \leq 0$ immediately gives (10.4.1).

Assuming that $f > 0$, observe that (10.4.4) is equivalent to $\left(\frac{1}{f}\right)' = -\frac{f'}{f^2} \geq \frac{1}{n-1}$. If we integrate on both sides on $\epsilon < s \leq t$ we get $\frac{1}{f(t)} - \frac{1}{f(\epsilon)} \geq \frac{t-\epsilon}{n-1}$. In particular,

$$f(t) \leq \frac{n-1}{t-\epsilon}.$$

Now taking $\epsilon \to 0$ we get $\Delta r \leq \frac{n-1}{r}$ which implies that $\Delta r^2 \leq 2n$. □

**Remark 10.4.2.**

1. In the above proof, we secretly used that $r$ is smooth at $\gamma(t)$ implies that it is smooth at all $\gamma(s)$ for $s < t$ which follows because $\gamma$ is length minimizing, so there are no conjugate points nor cut points. In the next section, we will prove that such is the case.

2. Calabi proved that (10.4.2) holds even when $r$ is not smooth.

## 10.5 Smoothness of the distance function

In Theorem 10.4.1, we assumed that the distance function $r(x) \coloneqq d(x,p)$ where $p \in M$ is smooth. Note that there exists a normal neighborhood $N_p$ of $p$ where geodesics are minimized and $r(x)$ is smooth in $N_p$. However, $r(x)$ might not be smooth once one leaves $N_p$. In this section, we will study the smoothness of $r$ along geodesics assuming that $(M,d)$ is a complete metric space.

By the Hopf–Rinow theorem 7.3.1 we know that for any $v \in \mathbb{S}^{n-1} \subset T_pM$ we can define $\gamma_v(t) \coloneqq \exp_p(tv)$ for all $t \in \mathbb{R}_+$. Define

$$\mathbf{R}(v) \coloneqq \sup\{T : \gamma_v(s) \text{ is minimizing for all } 0 \leq s \leq T\}.$$

Observe that when $t < \mathbf{R}(v)$ we have $d(p, \gamma_v(t)) = t$ which means that $d$ is smooth along $\gamma_v$ at least up until $t < \mathbf{R}(v)$.

**Definition 10.5.1.** The *cut locus at p* is the set $\mathrm{Cut}(p) \coloneqq \{\gamma_v(\mathbf{R}(v)) \mid v \in \mathbb{S}^{n-1} \text{ such that } \mathbf{R}(v) < \infty\}$. We say that $q \in \mathrm{Cut}(p)$ is a *cut point.*

**Example 10.5.2.**

1. In $\mathbb{S}^n$, $\mathbf{R}(v) = \pi$ for all $v \in \mathbb{S}^{n-1}$ and $\mathrm{Cut}(p) = \{-p\}$.

2. In an infinite cylinder $\mathbb{S}^1 \times \mathbb{R}$, the geodesic in the direction of the axis is always length-minimizing. These geodesics are called *rays.* However, a geodesic in any other direction fails to be length-minimizing after it hits the line opposite the rays. Indeed, embed $\mathbb{S}^1$ in $\mathbb{C}$. Then for any $(p,x) \in \mathbb{S}^1 \times \mathbb{R}$ we have $\mathrm{Cut}((p,x)) = \{\bar{p} \times \mathbb{R}\}$.

In both cases, the cut points are either conjugate points or have more than one length-minimizing geodesics. The observation is true in general.

**Theorem 10.5.3.** *Suppose* $q \coloneqq \gamma(\mathbf{R}(v)) \in \mathrm{Cut}(p)$ *for some* $v \in \mathbb{S}^{n-1}$*. Then*

1. *either $q$ is the first conjugate point along $\gamma$*
2. *or there exist two length-minimizing geodesics connecting $p$ and $q$.*

*Proof.* Unless otherwise stated, the geodesics have unit speed. Observe that $\gamma$ satisfies the following properties:

- $\gamma(t)$ is length minimizing up to and including $q$.
- $\gamma(t)$ is uniquely length-minimizing up to but not necessarily including $q$.
- $\gamma(t)$ is not length minimizing for $t > \mathbf{R}(v)$.
- There are no conjugate points before $q$. This follows from Morse Index Lemma.

Therefore, it suffices to assume that $q$ is not the first conjugate point along $\gamma$.

Consider $q_i$ to be points along $\gamma(t)$ for $t > \mathbf{R}(v)$ such that $q_i$ converge to $q$. Since $\gamma$ is not length-minimizing past $q$, there exist length-minimizing geodesics $\sigma_i : [0, T_i] \to M$ connecting $p$ and $q_i$. Since the distance function is continuous, we see that $L(\sigma_i)$ converges to $d(p, q)$.

To this end, since $q$ is not a conjugate point, there is some $\delta > 0$ and a neighborhood $U_q$ of $q$ such that $\exp_p$ is a diffeomorphism from $B_\delta(\mathbf{R}(v)v)$ and $U_q$. For large enough $i$, we can see that $T_i\sigma_i'(0)$ lie outside of $B_\delta(\mathbf{R}(v)v)$. Note that $T_i$ are close to $\mathbf{R}(v)$ and remain bounded. Therefore, $T_i$ converges to some $T_\infty$. On the other hand, by the compactness unit sphere in $T_pM$, there exists a subsequence $\sigma_i'(0)$ that converges to some $\sigma'(0)$ and such $T_\infty\sigma'(0)$ lies outside of $B_\delta(\mathbf{R}(v)v)$. Note that $L(\sigma) = d(p, q)$ and $\sigma(T_\infty) = q$. Further, $\sigma \neq \gamma$ because $\sigma'(0) \neq \gamma'(0)$. □

**Remark 10.5.4.**

1. The cut locus $\mathrm{Cut}(p)$ is closed. In fact, the set of conjugate points is closed since it is defined by a closed condition $\det\big(d\exp_p\big) = 0$. On the other hand, if we have a sequence of points $q_i = \gamma(\mathbf{R}(v_i)) \in \mathrm{Cut}(p)$ with at least two length-minimizing geodesics. Suppose $q_i$ converges to $q$. If $q$ is not a conjugate point, one can use a similar argument to prove that there are two distinct length-minimizing geodesics connecting $p$ and $q$ which along with the fact that geodesics are uniquely length-minimizing before $q_i$ shows that $q \in \mathrm{Cut}(p)$.

2. A similar argument show that for any $z \coloneqq \gamma(t)$ for $t < \mathbf{R}(v)$ there is a small neighborhood $B_\delta(z)$ of $z$ such that geodesics are uniquely length-minimizing. In particular, the distance function is smooth in $B_\delta(z)$ since it is the length of along the exponential map which is a diffeomorphic onto $\exp_p^{-1}(B_\delta(z))$.

# Chapter 11

# Harmonic functions

## 11.1 Maximum principle

Consider a function $f \in C^2([a,b];\mathbb{R})$. If $f$ has a local maximum at $x \in (a,b)$ then we know that $f'(x) = 0$ and $f''(x) \leq 0$. Therefore, if $f'' \geq 0$ then $f$ attains maximum on the boundary and if such function attains a maximum inside it has to be a constant function. In this section, we will prove similar results for functions on manifolds.

**Theorem 11.1.1 (Weak maximum principle).** *Let $\Omega \Subset M$ be a compact domain with boundary $\partial\Omega$. If $\Delta u > 0$ on $\Omega$ then*

$$\max_{\bar{\Omega}} u = \max_{\partial\Omega} u$$

*Proof.* Suppose $u$ attains a local maximum at $x \in \Omega$ in the interior. Let $\gamma$ be a unit speed geodesic on $M$ with $\gamma(0) = x_0$. Then $f(t) = u(\gamma(t))$ has a local maximum at 0. Note that

$$\begin{aligned} f'(t) &= \langle \boldsymbol{\nabla} u(\gamma(t)), \gamma'(t) \rangle \\ f''(t) &= \langle \boldsymbol{\nabla}_{\gamma'} \boldsymbol{\nabla} u, \gamma' \rangle + \langle \boldsymbol{\nabla} u, \boldsymbol{\nabla}_{\gamma'} \gamma' \rangle = \mathrm{Hess}_u(\gamma', \gamma') \end{aligned}$$

Therefore, the first and second derivative tests imply that at $t = 0$, $\langle \boldsymbol{\nabla} u, \gamma' \rangle = 0$ and $\mathrm{Hess}_u(\gamma', \gamma') \leq 0$. This is true for any geodesic $\gamma$ through $x_0$, particularly for geodesics $\gamma^1, \ldots, \gamma^n$ such that $\gamma^i_t(0) = \partial_{x^i}$, an orthonormal frame at $x_0$. Hence $|\boldsymbol{\nabla} u| = 0$ and $\Delta u \leq 0$.

Since $\Delta u > 0$ everywhere on $\Omega$, there is no interior maximum. Using compactness we see that the maximum is attained in the boundary. $\square$

**Theorem 11.1.2 (Hopf's strong maximum principle).** *If $\Delta u \geq 0$ on $\Omega \Subset M$ is a connected domain whose closure is compact then*

1. $\max_{\bar{\Omega}} u = \max_{\partial\Omega} u$ *and*

2. *If max is achieved in the interior then $u$ is a constant function.*

*Proof.* Set $m \coloneqq \max_{\bar{\Omega}} u$. Assume that there exists an interior point $x_0 \in \Omega$ with $u(x_0) = m$. We claim that $u \equiv m$ in $\bar{\Omega}$.

Define a set $S \coloneqq \{x \in \Omega \mid u(x) = m\}$. Note that $x_0 \in S$ so this set is nonempty. Since $u$ is continuous $S$ is closed. It suffices to prove that $S$ is open.

To this end, take $x \in S$ and consider small $t > 0$ such that $B_t(x) \subset \Omega$ and $B_{2t}(z) \subset \Omega$ for all $x \in B_t(x)$. For the sake of contradiction, say $y \in B_t(x)$ is not in $S$. Define $s = \inf\{s_0 \mid B_{s_0}(y) \cap S \neq \emptyset\}$. Note that $0 < s < tr$ as well as there exists $z \in \partial B_s(y) \cap S = B_s(y) \cap S \subset \Omega$. Then $u(z) = m$ but $u < m$ inside $B_s(y)$. Since $z$ is in the interior of $\Omega$ and attains a maximum we have $|\boldsymbol{\nabla} u|(z) = 0$. To prove the openness of $S$, it is enough to prove that $\boldsymbol{\nabla} u(z) \neq 0$.

The idea is to proceed by constructing a “barrier” of $u$. Define $A \coloneqq B_s(y) \setminus B_{\frac{s}{2}}(y)$. We say that $v$ is a *barrier* of $u$ if

1. $v \geq u$ on $\partial A = \partial B_s(y) \cup \partial B_{\frac{s}{2}}(y)$,
2. $\Delta v < 0$ on $A$,
3. $\frac{\partial v}{\partial r} > 0$, and
4. $u(z) = v(z)$.

If such $v$ exists, then 2 implies that $\Delta(u - v) = \Delta u - \Delta v > 0$ on $A$. Then Theorem 11.1.1 implies that $u - v$ attains its maximum on $\partial A$ which together with 1 implies that $u - v \leq 0$ on $A$. Since $u(z) = v(z)$ and $\partial_r v > 0$ we get that $\frac{\partial u}{\partial r} > 0$ at $r = s$. In particular $\boldsymbol{\nabla} u(z) \neq 0$.

It remains to find a barrier $v$ of $u$. We claim that $v(x) \coloneqq m + \epsilon^2(e^{-as^2} - e^{-ar^2})$ where $r \coloneqq d(x, y)$ works for any tiny $\epsilon > 0$ and appropriately chosen $a > 0$. Observe that 3 holds. Since $v(z) = m$ on $\partial B_s(y)$, 4 comes for free. Further, $u < m$ inside $A$, therefore we can pick $\delta > 0$ such that $u \leq m - \delta$. Then choose $\epsilon$ small enough such that $-\epsilon^2(e^{-as^2} - e^{-a\frac{s^2}{4}}) \leq \delta$ which implies 1. Finally, to prove 2, we need to show that $\Delta e^{-ar^2} > 0$ on $A$.

Note that $|\boldsymbol{\nabla} r| = 1$ and $\delta r^2 \leq 2n$ (cf. Theorem 10.4.1) imply that

$$
\begin{aligned}
\Delta e^{-ar^2} &= \boldsymbol{\nabla} \cdot \left(-a \boldsymbol{\nabla} r^2 e^{-ar^2}\right) \\
&= -a(\Delta r^2 e^{-ar^2} - a|\boldsymbol{\nabla} r^2|^2 e^{-ar^2}) \\
&= 4a^2 r^2 e^{-ar^2} - a\Delta r^2 e^{-ar^2} \\
&\geq 4a^2 r^2 e^{-ar^2} - 2nae^{-ar^2} \\
&\geq e^{-ar^2}(a^2 s^2 - 2na)
\end{aligned}
$$

where we used $r \geq \frac{s}{2}$ in the last line. Therefore, for large $a$ $\Delta e^{-ar^2} > 0$. $\square$

**Remark 11.1.3.** The results in this section hold for general differential operators $L$ that behave like $\Delta$. For more details look for elliptic differential operators.

## 11.2 Gradient estimate

**Theorem 11.2.1 (Cheng–Yau).** *Suppose $M^n$ is complete and $\operatorname{Ric} \geq 0$. Then there exists a constant $C$ such that if $\Delta u = 0$ on $B_{2R}(p) \subset M$ then*

$$\sup_{B_R(p)} |\nabla u| \leq \frac{C}{R} \sup_{B_{2R}(p)} |u|. \tag{11.2.1}$$

**Remark 11.2.2.**

1. The estimate in (11.2.1) is known as *interior estimate* and is common for second-order elliptic operators like $\Delta$.
2. On the left side of (11.2.1), we can use $B_{\delta R}$ for $0 < \delta < 1$, however we can't use $\delta = 1$ since the constant $C$ blows up as $\delta$ approaches 1. To see this, consider the case of holomorphic functions on $\mathbb{C}$. Note that the real and imaginary parts of holomorphic functions are harmonic. For simplicity consider $z^n$ on $B_1(0)$. Note that $|z^n| \leq 1$ on $B_1(0)$ and $|\nabla z^n| = n|z|^{1-n}$ is uniformly (in $n$) bounded on $B_{1/2}(0)$. However, $|\nabla z^n| = n$ on $\partial B_1$. Since we can take arbitrarily large $n$, there is no uniform $C$ such (11.2.1) holds if we replace $B(1/2)$ with $B_1$.
3. The equation (11.2.1) implies that any harmonic function on $M$ with sublinear growth has to be constant, a result known as the *Liouville theorem.* The result is in the same spirit as the Liouville theorem in complex analysis which says that a bounded entire function has to be constant.
4. On Euclidean space, we can iterate the bound for higher-order derivatives. It is called *bootstraping* in PDE.
5. There is a sharper version of the gradient estimate called the *differential Harnack inequality* for positive functions.

**Lemma 11.2.3 (Kato's inequality).** *Whenever $|\nabla u| \neq 0$ we have*

$$|\nabla |\nabla u|| \leq |\operatorname{Hess}_u|. \tag{11.2.2}$$

*Proof.* Write $\partial_{x^i} := \partial_i$. Then

$$\begin{aligned}
|\nabla |\nabla u||^2 &= g^{ij}\partial_i |\nabla u| \partial_j |\nabla u| \\
&= g^{ij} g^{kl} g^{mn} \partial_i \partial_k u \frac{\partial_l u}{|\nabla u|} \partial_j \partial_m u \frac{\partial_n u}{|\nabla u|} \\
&= \left\langle \operatorname{Hess}_u\left(\frac{\nabla u}{|\nabla u|}\right), \operatorname{Hess}_u\left(\frac{\nabla u}{|\nabla u|}\right)\right\rangle \\
&\leq \langle \operatorname{Hess}_u(\partial_i), \operatorname{Hess}_u(\partial_i)\rangle \\
&= |\operatorname{Hess}_u|^2.
\end{aligned}$$

In the above proof, we abused the notation and thought of the Hessian as an operator $\mathrm{Hess}_u : \mathfrak{X}(M) \to \mathfrak{X}(M)$. $\square$

**Lemma 11.2.4 (Peter–Paul inequality).** [1] *Suppose $a, b \geq 0$. Then*

$$2ab \leq \epsilon^2 a^2 + \frac{1}{\epsilon^2} b^2. \tag{11.2.3}$$

*Proof.* Expand $\left(\epsilon a - \frac{b}{\epsilon}\right)^2$. $\square$

*Proof of Theorem 11.2.1.* Define $F \coloneqq \phi^2 |\boldsymbol{\nabla} u|^2 + a u^2$ for some $\phi$ and a constant $a$. We will prove that $\Delta F \geq 0$ for some appropriately chosen constant $a \sim cR^2$ and $\phi$ such that

1. $\phi = 0$ on $\partial B_R(p)$ and $\phi \geq \frac{3}{4} R^2$ on $B_{\frac{R}{2}}(p)$,
2. $\Delta \phi^2 \geq -4nR^2$, and
3. $|\boldsymbol{\nabla} \phi| \leq 2R$.

For such $a$ and $\phi$ we see that

$$\frac{9}{16} R^4 \sup_{B_{\frac{R}{2}}(p)} |\boldsymbol{\nabla} u|^2 \leq \sup_{B_{R/2}(p)} F \leq \sup_{B_R(p)} F = \sup_{\partial B_R(p)} F = a \sup_{\partial B_R(p)} u^2$$

which implies (11.2.1).

First, let's prove $\Delta F \geq 0$. To this end,

$$\Delta(\phi^2 |\boldsymbol{\nabla} u|^2) = \phi^2 \Delta |\boldsymbol{\nabla} u|^2 + |\boldsymbol{\nabla} u|^2 \Delta \phi^2 + 2\langle \boldsymbol{\nabla} \phi^2, \boldsymbol{\nabla} |\boldsymbol{\nabla} u|^2 \rangle.$$

Using (10.1.1) and $\mathrm{Ric} \geq 0$ we see that $\Delta |\boldsymbol{\nabla} u|^2 = 2|\mathrm{Hess}_u|^2 + 2\,\mathrm{Ric}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} u) \geq 2|\mathrm{Hess}_u|^2$. On the other hand,

$$\begin{aligned} 2\langle \boldsymbol{\nabla} \phi^2, \boldsymbol{\nabla} |\boldsymbol{\nabla} u|^2 \rangle &\geq -8|\phi||\boldsymbol{\nabla} \phi||\boldsymbol{\nabla} u||\boldsymbol{\nabla}|\boldsymbol{\nabla} u|| \\ &\geq -16R|\phi||\boldsymbol{\nabla} u||\boldsymbol{\nabla}|\boldsymbol{\nabla} u|| \\ &\geq -16R|\phi||\boldsymbol{\nabla} u||\mathrm{Hess}_u| \\ &\geq -\phi^2 |\mathrm{Hess}_u|^2 - 64R^2 |\boldsymbol{\nabla} u|^2 \end{aligned}$$

where we used Cauchy–Schwarz inequality in the first line, $|\boldsymbol{\nabla} \phi| \leq 2R$ in the second line, Kato's inequality (11.2.2) in the third line and Peter–Paul inequality (11.2.3) applied to $a \coloneqq 8R|\boldsymbol{\nabla} u|$ and $b = |\phi||\mathrm{Hess}_u|$ and $\epsilon = 1$ in the fourth line. Therefore,

$$\Delta(\phi^2 |\boldsymbol{\nabla} u|^2) \geq (-4nR^2 - 64R^2)|\boldsymbol{\nabla} u|^2 + \phi^2 |\mathrm{Hess}_u|^2 \geq (-4nR^2 - 64R^2)|\boldsymbol{\nabla} u|^2 \tag{11.2.4}$$

[1] Say $\epsilon^2 a^2$ is Peter and $\frac{1}{\epsilon^2} b^2$ is Pal. If we rob Peter, we have to pay Pal.

On the other hand, $\Delta u^2 = 2|\boldsymbol{\nabla} u|^2$ which together with (11.2.4) implies that

$$\Delta F = \Delta\phi^2|\boldsymbol{\nabla} u|^2 + a\Delta u^2 \geq |\boldsymbol{\nabla} u|^2(2a - 4nR^2 - 64R^2).$$

Setting $2a = 4nR^2 + 64R^2$, we get $\Delta F \geq 0$. Therefore,

$$\frac{9}{16}R^4 \sup_{B_{\frac{R}{2}}(p)} |\boldsymbol{\nabla} u|^2 \leq (2n+32)R^2 \sup_{B_R(p)} u^2.$$

The final missing ingredient is $\phi$. Define $r(x) \coloneqq d(x,p)$. We claim that $\phi \coloneqq R^2 - r^2$ works. Note that 1 holds. Further, 3 holds because $|\boldsymbol{\nabla}\phi| = |-2r\boldsymbol{\nabla} r| \leq 2R$ since $|\boldsymbol{\nabla} r| = 1$. Finally, 2 holds because

$$\begin{aligned}
\Delta\phi^2 &= 2\phi\Delta\phi + 2|\boldsymbol{\nabla}\phi|^2 \\
&\geq 2\phi\Delta\phi \\
&= 2(R^2 - r^2)\Delta(R^2 - r^2) \\
&= -2(R^2 - r^2)\Delta r^2 \\
&\geq -4(R^2 - r^2)n \\
&\geq -4nR^2
\end{aligned}$$

where we used Theorem 10.4.1 in the last line. □

**Remark 11.2.5.** Note that $F$ is not smooth everywhere since $\phi$ is not smooth which follows because $r$ is not smooth as we saw in §10.5. However, one can still get the maximum principle when $r$ is not smooth using "viscosity solution methods", which was originally done by Calabi (cf [Cal58]) and later rediscovered by Pierre-Louis Lions.

Note that for $r(x) \coloneqq d(x,p)$ to be smooth we need the geodesic from $p$ to $x$ to be unique and minimizing and it has no conjugate points. Recall that $r$ fails to be smooth at the cut locus $\mathrm{Cut}(p)$. Further, $y \in \mathrm{Cut}(p)$. Say $\gamma$ is a geodesic starting at $p$ and ending at $y$. Note that $\gamma$ is length minimizing but not uniquely when $y$ is a cut point and is not minimizing past $y$ if it is a conjugate point. However, for any small $\epsilon$ the geodesic from $\gamma(\epsilon)$ to $y$ is uniquely minimizing. Therefore, $r_\epsilon(y) \coloneqq \epsilon + d(\gamma(\epsilon), y)$ is smooth. Further, $r(y) = r_\epsilon(y)$ and $r \sim r_\epsilon$ for points close to $y$. Then we run our proof with $r$ replaced by $r_\epsilon$ wherever $r$ is not smooth.

## 11.3 What's next?

In this section, we will list some topics that might be natural to read after reading these notes. We apologize for not providing an extensive bibliography for two reasons: one is the limitation in the writer's knowledge and the second is that we hope that one can quickly get references online if they type the terms we have listed here.

- **Minimal submanifolds:** We can study minimal submanifolds of general manifolds. See [CM11], [Law80], [Oss13].
- **Ricci curvature:** One can study Einstein manifolds where $\mathrm{Ric} = \lambda g$. See [Bes07] and [LW99]. In general, one can study manifolds with $\lambda_2 g \leq \mathrm{Ric} \leq \lambda_1 g$. See [BCJ].
- **Geometric Flows:** In these notes, we have mostly focused on static spaces. One can consider time-dependent families of manifolds or geometric quantities. The Ricci flow is an example of evolution of metric under the "heat equation." It has been used to prove the Poincaré conjecture. The Ricci flow in dimensions greater than 3 is still an active area of research. See [CLN06] and [Top06] for introduction. There are other flows like the mean curvature flow which is a dynamic version of minimal submanifolds. There is also geodesic flow.
- **Geometric Relativity:** Instead of considering a positive definite (Riemannian) metric, we can consider a Lorentz metric whose one eigenvalue is negative.
- **Ricci-limit spaces:** Note that the limit of smooth functions does not have to be smooth. One can consider studying non-smooth manifolds by taking the limits of smooth manifolds. For instance, a cone (which is not smooth at its tip) can be thought of as a limit of manifolds formed by smudging the tip of the cone. Just like how Sobolev space provides a class of solutions to certain PDEs, one can hope that the Ricci-limit spaces provide solutions to geometric PDEs.

# Appendix A

# Universal property of tensor product

Fix a commutative ring $\mathcal{R}$. Let $A, B, C$ denote $\mathcal{R}$-modules. Denote $A^* \coloneqq \mathbf{Hom}(A; \mathcal{R})$, the $\mathcal{R}$-module of $\mathcal{R}$-linear maps from $A$ to $\mathcal{R}$. Define $\mathbf{Hom}^2(A^* \times B^*; \mathcal{R})$ to be the $\mathcal{R}$-module of $\mathcal{R}$-bilinear maps $A^* \times B^* \to \mathcal{R}$ where a map $h : A \times B \to C$ is *$\mathcal{R}$-bilinear* if it is linear in each variable i.e. for all $a, a_1, a_2 \in A$, $b, b_1, b_2 \in B$ and $r_1, r_2 \in \mathcal{R}$ we have

$$h(a, r_1 b_1 + r_2 b_2) = r_1 h(a, b_1) + r_2 h(a, b_2)$$
$$h(r_1 a_1 + r_2 a_2, b) = r_1 h(a_1, b) + r_2 h(a_2, b).$$

**Definition A.0.1.** *The* tensor product $A \otimes B$ using an equivalence relation $\sim$ on the free $\mathcal{R}$-module generated by $A \times B$ such that for any $r \in \mathcal{R}$ and $(a, b), (c, d) \in A \times B$:

- **Scalar multiplication:** $r(a, b) \sim (ra, b) \sim (a, rb)$,
- **Distribution:** $(a, b + d) \sim (a, b) + (a, d)$ and $(a + c, b) \sim (a, b) + (c, d)$..

The Proposition A.0.3 below says that we can view elements of $A \otimes B$ as $\mathcal{R}$-bilinear maps $A^* \times B^* \to \mathcal{R}$.

**Remark A.0.2.** Setting $\mathcal{R} = C^\infty(M)$, $A = \mathfrak{X}(M)$, and $B = \mathfrak{X}^*(M)$, we can provide the legitimacy of the definition in (2.1.2) of $(1, 1)$-tensors as $C^\infty(M)$–bilinear maps $\mathfrak{X}^*(M) \times \mathfrak{X}(M) \to C^\infty(M)$. We can easily generalize the idea to $(r, s)$ tensors.

**Proposition A.0.3.** *There is a unique $\mathcal{R}$-module isomorphism $A \otimes B \cong \mathbf{Hom}^2(A^* \times B^*; \mathcal{R})$ where $a \otimes b$ gets sent to a map with the same notation and is defined as*

$$(a \otimes b)(a^*, b^*) = a^*(a) b^*(b).$$

In the rest of the section, we will provide a bigger stroke painting of proof of Proposition A.0.3.

We claim that there is always a *universal $\mathcal{R}$-bilinear map* out of $A \times B$ for every $\mathcal{R}$-modules $A$ and $B$. By that we mean there exists an $\mathcal{R}$-module $T$ and an $\mathcal{R}$-bilinear map $b : A \times B \to T$ that satisfies the universal property:

$$\begin{tikzcd} A \times B \arrow[r, "b"] \arrow[dr, "\forall k\ \mathcal{R}\text{-bilinear}"'] & T \arrow[d, dashed, "\exists!\ \bar{k}\ \mathcal{R}\text{-linear}"] \\ & W \end{tikzcd}$$

for any $\mathcal{R}$-module $W$. Here, the figure means that if there is an $\mathcal{R}$-bilinear map $k : A \times B \to W$ then there exists a unique $\mathcal{R}$-linear map $\bar{k} : T \to W$ such that $k = \bar{k} \circ b$. In other words, $k$ factors through $T$. Before proving existence, let's prove "uniqueness."

**Exercise A.0.4.** Suppose $b : A \times B \to T$ and $b' : A \times B \to T'$ are universal maps out of $A \times A$. Then there is a unique isomorphsim $j : T \to T'$ such that $j \circ b = b'$.

We call $T$ *the* tensor product of $A$ and $B$ and denote it by $A \otimes B$ with a justification given by Exercise A.0.5. The Exercise A.0.4 implies that no matter how we construct $T$, if we require it to satisfy the universal property, it is unique up to a unique isomorphism.

**Exercise A.0.5.** Check that Definition A.0.1 of tensor product with the quotient map $A \times B \to A \otimes B$ satisfies the universal property.

**Exercise A.0.6.** Define a map $b : A \times B \to \mathbf{Hom}^2(A^* \times B^*; \mathcal{R})$ that sends $(a, b) \in A \times B$ to an $\mathcal{R}$-bilinear map $A^* \times B^* \to \mathcal{R}$ such that for any $(a^*, b^*) \in A^* \times B^*$:

$$(a, b)(a^*, b^*) = a^*(a)b^*(b).$$

Check that $b$ has the universal property.

**Exercise A.0.7.** Prove Proposition A.0.3.

**Remark A.0.8.** The universality of the tensor product fits into a bigger story in the category theory of defining structures using their universal properties. For instance, suppose $S$ is a set. Then the discrete topology $\mathcal{T}$ on $S$ (all points are open) has the following universal property:

$$\begin{tikzcd} S \arrow[r, hook, "i"] \arrow[dr, "\forall k \text{ function}"'] & \mathcal{T} \arrow[d, dashed, "\exists!\ \bar{k} \text{ continuous}"] \\ & X \end{tikzcd}.$$

Here $i$ is the inclusion map and $X$ is a topological space. The figure means that for any function $f : S \to X$, there exists a unique continuous function $\bar{k} : \mathcal{T} \to X$ such that $\bar{k} \circ i = k$.

# Bibliography


[BCJ] Richard Bamler, Bennett Chow, and Wenshuai Jiang, *Ricci curvature in Riemannian geometry*. In preparation.

[Bes07] Arthur L Besse, *Einstein manifolds*, Springer Science & Business Media, 2007.

[BH13] Martin R Bridson and André Haefliger, *Metric spaces of non-positive curvature*, Vol. 319, Springer Science & Business Media, 2013.

[Cal58] Eugenio Calabi, *An extension of E. Hopf's maximum principle with an application to Riemannian geometry*, Duke Mathematical Journal **25** (1958), no. 1, 45–56.

[CE75] Jeff Cheeger and David G Ebin, *Comparison theorems in Riemannian geometry*, Vol. 9, North-Holland Amsterdam, 1975.

[CLN06] Bennett Chow, Peng Lu, and Lei Ni, *Hamilton's ricci flow*, Vol. 77, American Mathematical Soc., 2006.

[CM11] Tobias H Colding and William P Minicozzi, *A course in minimal surfaces*, Vol. 121, American Mathematical Soc., 2011.

[dC16] Manfredo P do Carmo, *Differential geometry of curves and surfaces: revised and updated second edition*, Courier Dover Publications, 2016.

[dC92] Manfredo Perdigão do Carmo, *Riemannian geometry*, Mathematics. Theory & applications, Birkhäuser, Boston, 1992.

[DFN84] BA Dubruvin, AT Fomenko, and SP Novikov, *Modern geometry—methods and application*, Springer-Verlag, New York Inc New York, 1984.

[Eti22] Pavel Etingof, *Lie groups and Lie algebras*, arXiv preprint arXiv:2201.09397 (2022).

[Hat05] Allen Hatcher, *Algebraic topology*, Cambridge University Press, 2005.

[KN63] Shoshichi Kobayashi and Katsumi Nomizu, *Foundations of differential geometry*, Vol. 1, New York, London, 1963.

[Law80] H Blaine Lawson, *Lectures on minimal submanifolds*, Vol. 1, Publish or Perish, 1980.

[LW99] Claude LeBrun and McKenzie Yuen-kong Wang, *Essays on Einstein manifolds*, International Press of Boston Incorporated, 1999.

[Mil63] John Milnor, *Morse theory*, Annals of mathematics studies, vol. 51, Princeton University Press, 1963.

[Mun00] James R Munkres, *Topology*, Vol. 2, Prentice Hall Upper Saddle River, 2000.

[Oba62] Morio Obata, *Certain conditions for a Riemannian manifold to be isometric with a sphere*, Journal of the Mathematical Society of Japan **14** (1962), no. 3, 333–340.

[Oss13] Robert Osserman, *A survey of minimal surfaces*, Courier Corporation, 2013.

[Str00] Daniel W Stroock, *An introduction to the analysis of paths on a Riemannian manifold*, American Mathematical Soc., 2000.

[SY94] Richard M Schoen and Shing-Tung Yau, *Lectures on differential geometry*, Vol. 2, International press Cambridge, MA, 1994.

[Tau11] Clifford Henry Taubes, *Differential geometry: Bundles, connections, metrics and curvature*, Vol. 23, OUP Oxford, 2011.

[Top06] Peter Topping, *Lectures on the ricci flow*, Vol. 325, Cambridge University Press, 2006.