A Concise Note on Differential Geometry

Mengchen Zeng

Last Compile: October 2, 2024

Preface

I had planned to write the preface at the end of my writing. But now I think I need a temporary preface.

We will adopt Einstein's summation convention throughout this note. That is, if an index occurs both in the superscript and subscript, it means to take summation over this index. For instance, we have

$$a^i b_i = \sum_{i=1}^{n} a^i b_i.$$

If an index occurs twice in superscript or subscript, then this means something is wrong, except in the case of orthonomal frames and similar context. For latter case, when indices actually occur twice at same position, we will not omit the summation symbol.

This note is not an encyclopedia, or even a textbook. Hence many proofs will not be provided. However, I will do my best to reference these proofs.

The part on smooth manifold is kept minimal. I only want to write about what is needed in the Riemannian geometry part.

I give special thanks to my supervisor Prof. Jian Ge, whose lecture notes have become a template to me. I also give special thanks to Yifan Jin, Pengju Shao and Mingze Li, who have continuously discussed with me and provided invaluable support.

Mengchen Zeng at Wuhan, Hubei, P. R. China.
October 2, 2024.

Contents

Chapter 1

Smooth Manifolds

In this chapter we will introduce the basic object of differential geometry: smooth (or differentiable) manifolds and smooth maps. In the meanwhile, we will introduce the main idea of differential geometry: linearization. To be specific, we will introduce tangent spaces and tangent maps, i.e. the linearization of smooth maps. The notion of manifolds with boundary and submanifolds will be discussed as well.

1.1   Category of Smooth Manifolds

Feel free if you know nothing about category theory, the section title just indicates we will define the notions of smooth manifolds and smooth maps. Throughout the note, we will use "smooth" to mean being C^∞.

We first have a topological object.

Definition. Let $n \in \mathbb{N}$, a *topological manifold M^n* is a second countable Hausdorff space that is locally Euclidean, i.e. for any $p \in M$, there is an open neighborhood U of p and a map $\varphi_U : U \to \mathbb{R}^n$ that maps U to an open set in $\mathbb{R}^n$ and homeomorphically onto its image. (U, φ_U) is called a *chart*. n is called the *dimension* of the manifold, denoted by $\dim M$. We will omit the superscript that indicates dimension if the dimension is clear.

By gluing chart smoothly, we can have the notion of a smooth manifold.

Definition. Let M be a topological manifold, two charts (U, φ_U) and (V, φ_V) are called *compatible* if both $\varphi_V \circ \varphi_U^{-1}$ and $\varphi_U \circ \varphi_V^{-1}$ are smooth. If compatible charts $\{(U_i, \varphi_i)\}$ consist an open cover of M, then it is called an *atlas*. A maximal atlas is called a *smooth structure*. A topological manifold together with a smooth structure is called a *smooth manifold*.

Remark 1.1.   (1) Not all topological manifolds admit a smooth structure, see for isntance, [3].

(2) Dimension of manifolds are well-defined, that is, there is no diffeomorphisms between open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$ if $m \neq n$. The proof requires topological invariants, such as homology groups. See for instance, [4, Theorem 2.55].

(3) We will assume all the manifolds occur in our note are path connected unless we specially claim.

(4) Notice that we only need to give an atlas of a manifold to define a smooth structure.

We now present some examples of smooth manifolds.

Example 1.2.   (1) $\mathbb{R}^n$ itself is a smooth manifold, with single chart id.

(2) Any open subset U of a smooth manifold M. If $\{U_i, \varphi_i\}$ is an atlas of M, then $\{U_i \cap U, \varphi_i|_{U_i \cap U}\}$ is an atlas of U. In particular, any open set of $\mathbb{R}^n$ is a smooth manifold.

Example 1.3. Sphere $\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} : |x| = 1\}$. An atlas is given by stereographic projections from north pole and south pole. We give the formula explicitly. Let $U = \mathbb{S}^n - \{(0, \cdots, 0, 1)\}$, $V = \mathbb{S}^n - \{(0, \cdots, 0, -1)\}$, then

$$U \xrightarrow{\varphi} \mathbb{R}^n, \ (x_1, \cdots, x_{n+1}) \mapsto \left(\frac{x_1}{1 - x_{n+1}}, \cdots, \frac{x_n}{1 - x_{n+1}} \right),$$

$$V \xrightarrow{\psi} \mathbb{R}^n, \ (x_1, \cdots, x_{n+1}) \mapsto \left(\frac{x_1}{1 + x_{n+1}}, \cdots, \frac{x_n}{1 + x_{n+1}} \right).$$

It's easy to see (U, φ) and (V, ψ) are compatible.

Example 1.4.   (1) If M, N are smooth manifolds, then $M \times N$ is a smooth manifold with atlas $\{U_i \times V_j, \varphi_i \times \psi_j\}$, provided $\{U_i, \varphi_i\}$ and $\{V_j, \psi_j\}$ are atlases of M and N respectively.

(2) The torus $\mathbb{T}^n$ is defined by $\mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ for n times.

To complete the definition of category of smooth manifolds, we need to define smooth maps.

Definition. A map $f : M \to N$ is called *smooth* at $p \in M$, if there exists chart (U, φ) of p and (V, ψ) of $f(p)$, such that $\psi \circ f \circ \varphi$ is smooth on $\varphi(U)$. If f is smooth everywhere, we simply say f is smooth. If $N = \mathbb{R}$, then we call f a smooth *function*, and we denote the set (actually $\mathbb{R}$-algebra) of smooth functions on M by $C^\infty(M)$.

The isomorphisms in the category of smooth manifolds are called diffeomorphisms, and we repeat the definition here.

Definition. A smooth map $f : M \to N$ is called a *diffeomorphism* if there exists a smooth $g : N \to M$, such that $g \circ f = \mathrm{id}_M$ and $f \circ g = \mathrm{id}_N$.

At the end of the section, we discuss partition of unity.

Definition. A collection of subsets $\{X_i\}_{i \in I}$ of topological space X is called *locally finite* if for any $x \in X$, there is a neighborhood U of x such that only finite many $i \in I$ satisfy $U \cap X_i \neq \varnothing$.

Definition. The *support* of a smooth function $f \in C^\infty(M)$ is the closure of $\{p \in M : f(p) \neq 0\}$, denoted by $\mathrm{supp}\, f$.

Now we can state the partition of unity theorem here, but we will not present the long and complicated proof. For a proof, one can check [5, Theorem 2.23].

Theorem 1.5 (Partition of unity)**.** *Let M be a smooth manifold with an open cover $\{U_i\}_{i \in I}$. Then there exists a family of smooth functions $\{\varphi_i\}_{i \in I}$ satisfying*

(1) $0 \leq \varphi_i \leq 1$;

(2) $\mathrm{supp}\, \varphi_i \subset U_i$ *for each* $i \in I$;

(3) $\{\mathrm{supp}\, \varphi_i\}_{i \in I}$ *is locally finite;*

(4) $\sum_{i \in I} \varphi_i(p) = 1$ *for* $p \in M$.

Remark 1.6. By (3), the summation in (4) is a finite sum, so we will not encounter convergence problem here.

Partition of unity plays an important role in defining many geometric objects, such as integral, metric, connection and so on. But any of them is beyond our scope in this section, so we have to postpone the illustration of applications of partition of unity.

1.2 Tangent Spaces and Tangent Maps

In this section we discuss the linearization of smooth maps, the main idea of differential geometry.

We first define tangent spaces. In modern textbooks, tangent vectors are defined to be derivatives on the germ of smooth functions. This definition can show directly that tangent vectors consists a vector space. However, we will adopt traditional definition via curves on manifolds, which will make our calculation more convenient.

Definition. The *germ of smooth functions* at $p \in M$ is a ring defined by

$$\{(f, U) \mid U \subset M \text{ open}, f \in C^\infty(U)\}/\sim,$$

where $(f, U) \sim (g, V)$ if there exists a $W \subset U \cap V$ such that $f|_W = g|_W$. Ring operations are defined in natural way. Elements in $C_p^\infty(M)$ are denoted by f_p if it has representative element f.

Definition. Let $\gamma : [0, 1] \to M$ be a smooth curve, i.e. a smooth map, with $\gamma(0) = p$. Define the *tangent vector at p along γ* to be a map

$$\dot{\gamma}(0) : C_p^\infty(M) \to \mathbb{R}$$
$$f_p \mapsto (f \circ \gamma)'(0).$$

Then we define the *tangent space at p* to be

$$T_p M := \{\dot{\gamma}(0) \mid \gamma : [0, 1] \to M \text{ smooth}, \gamma(0) = p\}.$$

Remark 1.7. The assignment of curve to its tangent vector is not a bijection. For instance, on $T_0 \mathbb{R}^2$, $y = 0$ and $y - x^2 = 0$ yield same tangent vector at $(0, 0)$.

It's easy to see that

Proposition 1.8. *We have* Leibniz rule

$$\dot{\gamma}(0)(f_p g_p) = f(p)\dot{\gamma}(0)(g_p) + g(p)\dot{\gamma}(0)(f_p).$$

Therefore a tangent vector is a derivative *on $C_p^\infty(M)$.*

We will show that $T_p M$ has a real vector space structure. In fact, we have an enhanced conclusion.

Proposition 1.9. *Let M^n be a smooth manifold, (U, φ) be a chart containing p, $\varphi = (x^1, \cdots, x^n)$. Define $\left.\frac{\partial}{\partial x^i}\right|_p$ to be the tangent vector of $\sigma_i = \varphi^{-1}(\varphi(p) + te_i)$, where $e_i = (0, \cdots, 1, \cdots, 0)$ with only i-th component being 1. Then we have*

$$T_p M = \text{Span}\left\{\left.\frac{\partial}{\partial x^i}\right|_p\right\}$$

as sets, so we can equip $T_p M$ a real vector space structure.

Proof. Notice that

$$\left.\frac{\partial}{\partial x^i}\right|_p f = \frac{\partial}{\partial x^i}(f \circ \varphi^{-1})(p),$$

the partial derivative on right hand side is the usual Euclidean partial derivative. Hence by chain rule, we have

$$
\begin{aligned}
\dot{\gamma}(0)f &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f \circ \gamma)(t) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f \circ \varphi^{-1}) \circ (\varphi \circ \gamma)(t) \\
&= \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)} (f \circ \varphi^{-1}) \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} x^i(\gamma(t)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} x^i(\gamma(t)) \left.\frac{\partial}{\partial x^i}\right|_{p} f.
\end{aligned}
$$

Therefore we have

$$
\dot{\gamma}(0) = \dot{\gamma}(0)(x^i) \left.\frac{\partial}{\partial x^i}\right|_{p} \in \mathrm{Span}\left\{ \left.\frac{\partial}{\partial x^i}\right|_{p} \right\},
$$

and we obtain

$$
T_p M \subset \mathrm{Span}\left\{ \left.\frac{\partial}{\partial x^i}\right|_{p} \right\}.
$$

The reverse inclusion is trivial. $\qquad\square$

Remark 1.10. The vector space structure of tangent space does not depend on the choice of chart, since for any charts (U_1, φ_1) and (U_2, φ_2), the Jacobian of transition function $\varphi_1 \circ \varphi_2^{-1}$ is a vector space isomorphism. However, this is again rely on the well-definedness of dimension of manifolds, but we have assumed it automatically.

Next we define tangent maps.

Definition. Let $f : M \to N$ be a smooth map, $p \in M$. We define $f_{*p} : T_p M \to T_{f(p)} N$ by
$$
f_{*p}(v)(g) = v(g \circ f)
$$
for any $v \in T_p M$ and $g \in C^\infty_{f(p)} N$.

Remark 1.11. We often write f_{*p} as $\mathrm{d}f|_p$ if f is a smooth function. This coincides the notation for 1-form we will define later.

Lemma 1.12. *On charts (U, φ) of $p \in M$ and (V, ψ) of $q \in N$, the smooth map $f : M \to N$ that $f(p) = q$ satisfies*

$$
f_{*p}\left(\left.\frac{\partial}{\partial x^i}\right|_{p} \right) = \frac{\partial}{\partial x^i}(\psi \circ f \circ \varphi)^j(\varphi(p)) \left.\frac{\partial}{\partial y^j}\right|_{q}.
$$

The proof is similar to Proposition 1.9.

Moreover, we have the chain rule.

Proposition 1.13. *Let $f : M \to N$, $g : N \to P$ be two smooth maps, then for $p \in M$ we have*

$$(g \circ f)_{*p} = g_{*f(p)} \circ f_{*p}.$$

Proof. Let $v \in T_p M$, then for any $\varphi \in C^\infty_{g(f(p))}(M)$ we have

$$\begin{aligned}
(g \circ f)_{*p}(v)(\varphi) &= v(\varphi \circ (g \circ f)) \\
&= v((\varphi \circ g) \circ f) \\
&= (f_{*p}(v))(\varphi \circ g) \\
&= (g_{*f(p)}(f_{*p}(v)))(\varphi) \\
&= (g_{*f(p)} \circ f_{*p})(v)(\varphi). \qquad \square
\end{aligned}$$

Since we linearize smooth maps, we can talk about rank of a map at a point.

Definition. Let $f : M \to N$ be a smooth map and $p \in M$, the *rank* of f at p, denoted by $\mathrm{rank}_p f$, is defined by the rank of f_{*p}.

Definition. Let $f : M^m \to N^n$ be a smooth map.

(1) Assume $m \leq n$, if f_{*p} is injective for any $p \in M$, then we call f an *immersion*.

(2) Assume f is an immersion, if f maps M homeomorphically onto its image, then we call f an *embedding*.

(3) Assume $m \geq n$, if f_{*p} is surjective for any $p \in M$, then we call f a *submersion*.

Remark 1.14. An equivalent description for immersion and submersion is that they are of constant maximal rank. Also notice that a map is both an immersion and submersion if and only if it is a diffeomorphism.

Example 1.15. The definition of embedding $f : M \to N$ asks $f(M)$ is equipped with subspace topology from N, so f being injective is not enough. For example, consider the injective map

$$r : (-\pi, \pi) \to \mathbb{R}^2, \ t \mapsto (2\sin t, \sin(2t)).$$

The image of r is the zero locus of $4y^2 = x^2(4 - x^2)$, it's easy to see it is compact. But $(-\pi, \pi)$ is not compact, hence r does not map the interval homeomorphically onto its image.

1.3 Submanifolds

In this section, we discuss submanifolds of an ambient manifold. We first present the definition. Then, we will discuss rank theorem, a powerful tool to discuss submanifolds. We will use rank theorem to derive inverse function theorem and implicit function theorem, and discuss submanifolds and embedded submanifolds.

Definition. Let M be a smooth manifold, $\Sigma \subset M$ be a subset. Σ is called a *(k-dimensional) submanifold* of M if for any $p \in \Sigma$, there exists a chart (U, φ) of p in M such that

$$\varphi(U \cap \Sigma) = \varphi(U) \cap \{(x^1, \cdots, x^k, 0, \cdots, 0) \in \mathbb{R}^n \mid x^1, \cdots, x^k \in \mathbb{R}\}.$$

We define the *codimension* of Σ to be $n - k$, denoted by $\operatorname{codim} \Sigma = n - k$.

We will not linger to give any examples. On the contrary, we will discuss rank theorem at once.

Theorem 1.16 (Rank theorem). *Let $f : M^m \to N^n$ be a smooth map that has constant rank $r (\leq m)$. For any $p \in M$, there exists charts (U, φ) of p and (V, ψ) of $f(p)$, such that*

$$(\psi \circ f \circ \varphi^{-1})(x^1, \cdots, x^r, \cdots, x^m) = (x^1, \cdots, x^r, 0, \cdots, 0)$$

on $\varphi(U)$.

The proof of rank theorem is also complicated, and we just refer to [5, Theorem 4.12]. Rank theorem has important applications. First, as corollary, we have inverse function theorem and implicit function theorem.

Theorem 1.17 (Inverse function theorem). *Let $f : M \to N$ be a smooth map, $p \in M$. Suppose f_{*p} is an isomorphism, then there exists a neighborhood U of p such that $f|_U$ is a diffeomorphism on U.*

Proof. Since determinant function is smooth, $\det f_{*p} \neq 0$ in a neighborhood of p. Then by rank theorem, there exists sufficiently small charts (U, φ) of p and (V, ψ) of $f(p)$ such that $\psi \circ f \circ \varphi^{-1}$ is identity on $\varphi(U)$. Thus $f|_U = \psi^{-1} \circ \varphi$ is a homeomorphism, and it is a diffeomorphism since f is smooth. $\square$

Theorem 1.18 (Implicit function theorem). *Let $f : M^{k+n} \to \mathbb{R}^k$ be a smooth map. Suppose f_{*p} is a submersion for any $p \in f^{-1}(0)$, then $f^{-1}(0)$ is an n-dimensional submanifold of M (may not be connected).*

Proof. For $p \in f^{-1}(0)$, since determinant function is smooth, there exists a neighborhood of p such that the tangent map has constant rank. Hence by rank

theorem, there exists sufficiently small charts (U, φ) of p and (V, ψ) of 0 such that

$$(\psi \circ f \circ \varphi^{-1})(x^1, \cdots, x^k, x^{k+1}, \cdots, x^{k+n}) = (x^1, \cdots, x^k).$$

Then $f^{-1}(0)$ restricts on U has coordinate

$$\varphi(f^{-1}(0) \cap U) = \{(0, \cdots, 0, x^{k+1}, \cdots, x^{k+n}) \in \mathbb{R}^{n+k}\} \cap \varphi(U).$$

Thus $f^{-1}(0)$ is an n-dimensional submanifold. $\qquad\qquad\qquad\qquad\qquad\square$

Remark 1.19.   (1) The value 0 plays an unimportant role in the proof. In fact, we can replace 0 by r_0 and consider $g(x) = f(x) - r_0$.

  (2) The condition of f_{*p} being a submersion for any $p \in f^{-1}(r_0)$ is seemingly hard to verify. However, we will prove later that such property holds *almost everywhere* in $\mathbb{R}$. (In fact, $\mathbb{R}^k$ for all $k \in \mathbb{N}$. This is Sard's Theorem, but we only need $k = 1$ case in this note.)

    We now characterize submanifolds.

Definition. Let $f : \Sigma \to M$ be an immersion, then the image of f is called an *immersed submanifold*. Moreover, if f is an embedding, the image of f is called an *embedded manifold*.

Proposition 1.20. *Any immersed manifold is locally an embedded manifold.*

Proof. Only need to show any immersion is locally an embedding, and this is clear from rank theorem. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proposition 1.21. *A subset of a smooth manifold is a submanifold if and only if it is an embedded submanifold.*

Proof. Let M be a smooth manifold, $S \subset M$ be a submanifold, then the inclusion $i : S \to M$ is an embedding. Conversely, let $f : \Sigma^k \to M^n$ be an embedding. Let $p \in f(\Sigma)$, $f(q) = p$, then by rank theorem, there exists charts (U, φ) of q and (V, ψ) of p such that

$$(\psi \circ f \circ \varphi^{-1})(x^1, \cdots, x^k) = (x^1, \cdots, x^k, 0, \cdots, 0).$$

Since $\Sigma \cong f(\Sigma)$ with latter equipped with subspace topology, $f(U) = \tilde{V} \cap f(\Sigma)$ for some open $\tilde{V} \subset M$. By shrinking $\tilde{V}$ (and U), we can assume $\tilde{V} \subset V$, then

$$\psi(\tilde{V} \cap f(\Sigma)) = \{(x^1, \cdots, x^k, 0, \cdots, 0) \in \mathbb{R}^n\} \cap \psi(\tilde{V}),$$

hence $f(\Sigma)$ is a submanifold of M. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

    Now we provide some examples of submanifolds.

Example 1.22. (1) Open subsets of a smooth manifold. This is obvious.

(2) *Surfaces* in $\mathbb{R}^3$. This means a submanifold of codimension 1. In classical differential geometry, a surface means a 2-dimensional immersed submanifold. That is, a smooth map $r : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^3$ with $r_{;1} \times r_{;2} \neq 0$. Generally, a submanifold of codimension 1 (no matter what ambient manifold) is called a *hypersurface*.

(3) *Zero locus* of $f : M \to \mathbb{R}$ such that f_{*p} has maximal rank for all $p \in f^{-1}(0)$. In particular, n-sphere $\mathbb{S}^n$ is the zero locus of $|x|^2 - 1 = \sum_{i=1}^{n}(x^i)^2 - 1$.

(4) *(Affine) algebraic manifolds*, i.e. *smooth algebraic varieties*[1]. Algebraic varieties are common zero locus of a family of polynomials $\{f^i\}_{i \in I}$. By Hilbert's basis theorem, we can assume the index set I is finite (see for example, [7, Theorem B-1.16]). Hence we can have a differential smoothness condition that $f = (f^1, \cdots, f^{|I|})$ has maximal rank everywhere, instead of using commutative algebra.

1.4 Manifolds with Boundary

Definition. A *(smooth) manifold with boundary* is (what?)

---

[1]Usually we only discuss algebraic varieties on algebraic closed fields, such as $\mathbb{C}$. But define varieties on arbitrary field does not violate the definition in algebraic geometry.

Chapter 2

Metric Structure of Riemannian Manifolds

In this chapter we introduce the metric structure on a Riemannian manifold, i.e. using the metric tensor to make the manifold into a metric space. We will prove that the topology induced by metric coincides with the topology carried by the manifold. Then we will study the length-minimizing problem: which curve minimize the distance between two points? The answer is two-sided: we will prove that length-minimizing curves are geodesics, and geodesics are *locally* length-minimizing.

2.1   Metric Structure on Riemannian Manifolds

Let (M, g) be a Riemannian manifold, $\gamma : [0, 1] \to M$ be a regular curve (i.e. γ is an immersion). We define the length functional of regular curves.

Definition. The *length functional* $L[\cdot]$ on the set of regular curves is defined by

$$L[\gamma] = \int_0^1 \sqrt{g(\dot{\gamma}(t), \dot{\gamma}(t))} \, \mathrm{d}t.$$

It is a well-known result from calculus that any regular curve can be reparametrized by arc-length. From now on the word "curve" means a regular curve.

Now we define the distance between two points.

Definition. Let $p, q \in M$ be two points on M, then we define their *distance* by

$$d(p, q) = \inf_{\gamma \in C_{p,q}} L[\gamma],$$

11

where $C_{p,q}$ is the set of all regular curves joining p and q.

Proposition 2.1. *The distance function $d : M \times M \to \mathbb{R}$ has the following properties:*

   (1) $d(p,q) \geq 0$, and $d(p,q) = 0 \iff p = q$;

   (2) $d(p,q) = d(q,p)$;

   (3) $d(p,r) \leq d(p,q) + d(q,r)$.

Thus the distance function makes M into a metric space.

Proof. Only need to show $d(p,q) = 0 \iff p = q$, all else are trivial. We assume $p \neq q$, need to show $d(p,q) > 0$. Let $\gamma : [0,1] \to M$ be any curve joining p and q. Choose a local chart (U, φ) such that $\varphi(U) = B_r(0)$, $q \neq U$. By Jordan–Brouwer Separation Theorem, γ must intersect ∂U at $s := \gamma(c)$. Then we have

$$L[\gamma] \geq L[\gamma|_{[0,c]}] = \int_0^c \sqrt{g_{ij}\dot{x}^i(\gamma(t))\dot{x}^j(\gamma(t))}\, \mathrm{d}t.$$

Regarding $g : \overline{U} \times \mathbb{S}^{n-1} \to \mathbb{R}$, g is a continuous function on a compact set, thus it attains its minimum $g(x)(v,v) \geq m$, and $m > 0$ since $v \in \mathbb{S}^{n-1} \neq 0$. Thus we have

$$L[\gamma] \geq L[\gamma|_{[0,c]}] \geq m \int_0^c |\dot{x}(\gamma(t))|\, \mathrm{d}t \geq mr > 0. \tag{2.1}$$

mr does not depend on γ, hence $d(p,q) \geq mr > 0$. $\square$

However, the metric space topology is nothing but the original topology carried by the manifold.

Proposition 2.2. *The metric space topology on M coincides with the manifold topology.*

We first need a lemma.

Lemma 2.3. *The distance function to p defined by $r(q) = d(p,q)$ is continuous with respective to the manifold topology.*

Proof. Since manifolds satisfy the second countable axiom, the Sequence Lemma holds. Then it's equivalent to show for any $q_i \to q$ in manifold topology, we have $r(q_i) \to r(q)$. Without loss of generality we can assume $\{q_i\} \subset U$ and (U, φ) is a local chart such that $\varphi(q) = 0$, $\varphi(U) = B_r(0)$. Let δ be the Euclidean metric on $B_r(0)$, then by regarding g as a continuous function on $\overline{U} \times \mathbb{S}^{n-1}$ again, we have $g \leq M\delta$ for some $M > 0$. By assumption, $q_i \to q$ in manifold topology implies $L_\delta[\psi_i] \to 0$, where $\psi_i(t) = t\varphi(q_i)$, the radial line joining $\varphi(q)$ and $\varphi(q_i)$ in $\varphi(U) = B_r(0)$. Let $\varphi^{-1}\psi_i = \gamma_i$, then we have

$$L_g[\gamma_i] = \int_0^1 \sqrt{g(\dot{\gamma}_i(t), \dot{\gamma}_i(t))}\, \mathrm{d}t \leq M \int_0^1 \sqrt{\delta(\dot{\psi}_i(t), \dot{\psi}_i(t))}\, \mathrm{d}t \leq ML_\delta[\psi_i].$$

Since r is Lipschitz, i.e. $|r(q) - r(s)| \leq d(q, s)$, we have

$$d(q_i, q) \leq L_g[\gamma_i] \leq M L_\delta[\psi_i] \to 0,$$

hence r is continuous. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of Proposition 2.2. Since distance function is continuous, metric balls are open in manifold topology. Now we prove the converse.

Let U be open with respective to manifold topology. Let $p \in U$, V be a neighborhood of p so small that $\varphi(V) = B_r(0)$ for some $r > 0$. The estimate (2.1) shows if $q \notin V$ then $d(p, q) \geq mr$ for some fixed $m > 0$, then by taking contrapositive statement, we have $q \in V$ if $d(p, q) < mr$. Therefore $B_p(mr) \subset U$, then U is open with respective to metric space topology. $\qquad\qquad\square$

2.2 Length-Minimizing Curves

This section and the next is guided by the following problem:

Problem. Let p, q be two distinct points on a Riemannian manifold (M, g), then what is the curve γ satisfying $L[\gamma] = d(p, q)$?

This section we will show if the length-minimizing curve exists, then it is a *geodesic*. Next section we will show if p, q are sufficiently closed, *the* geodesic joining p and q is length-minimizing. Finally, the existence of length-minimizing curves is related to Hopf-Rinow Theorem, which we will discuss at the last section of the chapter.

To find the length-minimizing curve, we "gather" curves with same initial and end points, which is called a variation.

Definition. Let $\gamma_0 : [0, a] \to M$ be a curve, a *variation* of γ_0 is a differentiable map $\gamma : [0, a] \times (-\varepsilon, \varepsilon) \to M$ such that $\gamma(t, 0) = \gamma_0(t)$. If $\gamma(0, s) = \gamma_0(0)$ and $\gamma(a, s) = \gamma_0(a)$ for any $s \in (-\varepsilon, \varepsilon)$, then we call the variation a *proper variation*. We call $\left. \frac{\partial}{\partial s} \right|_{s=0} \gamma(s, t) =: V(t)$ the *variation vector field*.

Now we introduce the energy functional, which is easier to calculate.

Definition. The *energy functional* on the set of curves is defined by

$$E[\gamma] = \int_0^a \frac{1}{2} |\dot{\gamma}(t)|^2 \, dt,$$

where $\gamma : [0, a] \to M$ is a regular curve.

We will prove that a curve is energy-minimizing if and only if it is length-minimizing.

Lemma 2.4. *For a curve $\gamma : [0, a] \to M$, we have*

$$L^2[\gamma] \leq 2aE[\gamma],$$

with equality holds if and only if $|\dot{\gamma}(t)| = \text{const}.$

Proof. This is Cauchy–Schwarz inequality.   □

Proposition 2.5. *If γ is length-minimizing, then it is energy-minimizing.*

Proof. Let $\tilde{\gamma}$ be another curve, then we have

$$2aE[\gamma] = L^2[\gamma] \leq L^2[\tilde{\gamma}] \leq 2aE[\tilde{\gamma}].$$ □

Our aim is to prove the converse.

Proposition 2.6. *If γ is an energy-minimizing curve, then it is length-minimizing.*

To prove this, we need to differentiate the variation.

Proposition 2.7 (First variation formula)**.** *Let $\gamma(t, s)$ be a variation, define its energy*
$E(s) = \int_0^a \frac{1}{2} \left| \frac{\partial}{\partial t} \gamma(t, s) \right|^2 \mathrm{d}t$, *then we have*

$$E'(0) = \boxed{\langle V, \dot{\gamma} \rangle |_0^a} - \int_0^a \langle V(t), \nabla_{\dot{\gamma}_0(t)} \dot{\gamma}_0(t) \rangle.$$

The boxed term is called boundary term, *and it vanishes when the variation is proper.*

Proof. This is a calculation. We have

$$\frac{\mathrm{d}}{\mathrm{d}s} E(s) = \int_0^a \left\langle \frac{\partial \gamma}{\partial t}, \nabla_{\frac{\partial}{\partial s}} \frac{\partial \gamma}{\partial t} \right\rangle \mathrm{d}t$$

$$= \int_0^a \left\langle \frac{\partial \gamma}{\partial t}, \nabla_{\frac{\partial}{\partial t}} \frac{\partial \gamma}{\partial s} \right\rangle \mathrm{d}t$$

$$= \int_0^a \frac{\partial}{\partial t} \left\langle \frac{\partial \gamma}{\partial t}, \frac{\partial \gamma}{\partial s} \right\rangle - \left\langle \frac{\partial \gamma}{\partial s}, \nabla_{\frac{\partial}{\partial t}} \frac{\partial \gamma}{\partial t} \right\rangle \mathrm{d}t.$$

Take $s = 0$, we obtain

$$E'(0) = \int_0^a \frac{\partial}{\partial t} \langle V(t), \dot{\gamma}_0(t) \rangle - \langle V(t), \nabla_{\dot{\gamma}_0(t)} \dot{\gamma}_0(t) \rangle \, \mathrm{d}t$$

$$= \langle V, \dot{\gamma}_0 \rangle |_0^a - \int_0^a \langle V(t), \nabla_{\dot{\gamma}_0(t)} \dot{\gamma}_0(t) \rangle \, \mathrm{d}t.$$ □

Now we can derive the definition of a geodesic.

Definition. A curve $\gamma : [0, a]$ is called a *geodesic* if $\nabla_{\dot{\gamma}(t)}\dot{\gamma}(t) = 0$ for $t \in [0, a]$.

Remark 2.8. Geodesics are constant speed, this can be shown by $\frac{d}{dt}|\dot{\gamma}(t)|^2 = 2\langle\dot{\gamma}(t), \nabla_{\dot{\gamma}(t)}\dot{\gamma}(t)\rangle = 0$.

Corollary 2.9. γ *is a critical value for all proper variation if and only if* $\nabla_{\dot{\gamma}(t)}\dot{\gamma}(t) = 0$, *that is,* γ *is a geodesic.*

We can give the proof of Proposition 2.6 now.

Proof of Proposition 2.6. Let $\gamma : [0, a] \to M$ be a curve such that for any $\tilde{\gamma} : [0, a] \to M$ with $\tilde{\gamma}(0) = \gamma(0)$, $\tilde{\gamma}(1) = \gamma(1)$, the inequality $E[\gamma] \leq E[\tilde{\gamma}]$ holds, we show that $L[\gamma] \leq L[\tilde{\gamma}]$. Let $\gamma(t, s)$ be any proper variation with $\gamma(t, 0) = \gamma(t)$, then γ is a critical point of $E(s)$. Hence by Corollary 2.9, γ is a geodesic. Now we can reparametrize $\tilde{\gamma}$ into arc-length, obtaining $\hat{\tilde{\gamma}}$. Therefore

$$L^2[\gamma] = 2aE[\gamma] \leq 2aE\left[\hat{\tilde{\gamma}}\right] = L^2\left[\hat{\tilde{\gamma}}\right] = L^2[\tilde{\gamma}],$$

which implies $L[\gamma] \leq L[\tilde{\gamma}]$. □

Combining all results above, we have

Proposition 2.10. *If a curve is length-minimizing, then it is a geodesic.*

2.3 Geodesics and Exponential Maps

To prove the local length-minimizing property of geodesic, we need to introduce the exponential map.

We first need to investigate the equation that determine a geodesic.

Proposition 2.11. *Given* $p \in M$ *and* $v \in T_pM$, *there exists a unique geodesic* γ *(whose domain may not be maximal) such that* $\gamma(0) = p$, $\dot{\gamma}(0) = v$.

Proof. Let (U, φ) be a local chart containing p, compose φ with γ we obtain coordinate curves x^i's. Then the geodesic equation is equivalent to

$$\ddot{x}^k(t) + \Gamma_{ij}^k(\gamma(t))\dot{x}^i(t)\dot{x}^j(t) = 0, \ k = 1, \cdots, n.$$

This is a system of second order ordinary differential equations, by the unique existence theorem of ODE, the solution is completely determined by x^i's and $\dot{x}^i$'s, that is, p and v. □

Since the solution of an ODE relies continuously on its initial value, we have the following proposition.

Proposition 2.12. *For any $p \in M$, there exists a neighborhood V of p, such that there exists $\delta > 0$, $\varepsilon > 0$ and a differentiable map $\gamma : (-\delta, \delta) \times \mathscr{U} \to M$, where $\mathscr{U} = \{(q, v) \in TM : q \in V, v \in T_q M, |v| < \varepsilon\}$, such that $\gamma(t; q, v)$ is a geodesic with $\gamma(0) = q$, $\dot{\gamma}(0) = v$.*

A proof can be found in [8, Chapter 3, Lemma 1].

Observe that $\gamma(\lambda t; p, v) = \gamma(t; p, \lambda v)$. Denote $\gamma(t; p, v)$ by $\gamma_v(t)$, then above observation can be written as $\gamma_{\lambda v}(t) = \gamma_v(\lambda t)$. Therefore, we can shorten the initial vector to lengthen the domain of geodesic.

Definition. Let $U \subset T_p M$ be a neighborhood of origin, such that for any $v \in U$, $\gamma_v(1)$ is defined (existence is guaranteed by Proposition 2.12). We define the *exponential map* at p by

$$\exp_p : U \to M$$
$$v \mapsto \gamma_v(1).$$

Remark 2.13. We can scale the initial vector and obtain

$$\exp_p(v) = \gamma_v(1) = \gamma_{v/|v|}(|v|).$$

This means the action of exponential map on v is to move forward the distance $|v|$ along the geodesic with initial direction $v/|v|$.

Proposition 2.14. $\exp_{p*}|_0 : T_0(T_p M) \to T_p M$ *is identity (we identify $T_0(T_p M)$ with $T_p M$).*

Proof. We have

$$\exp_{p*}|_0(v) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}(tv) = v. \qquad \square$$

Corollary 2.15. *There exists a ball $B_\varepsilon(0) \subset T_p M$ such that $\exp_p : B_\varepsilon(0) \to M$ is a diffeomorphism onto its image.*

Proof. Since $\exp_{p*}|_0$ is identity, it is nondegenerate, the corollary follows from inverse function theorem. $\qquad \square$

Exponential maps enjoys naturality.

Proposition 2.16. *Let $f : M \to N$ be a local isometry, $p \in M$, $q \in N$ and $f(p) = q$. Then we have the following commutative diagram*

$$
\begin{array}{ccc}
U \subset T_p M & \xrightarrow{\ f_{*p}\ } & V \subset T_q N \\
\downarrow{\scriptstyle \exp_p} & & \downarrow{\scriptstyle \exp_q} \\
M & \xrightarrow{\quad f \quad} & N.
\end{array}
$$

Proof. This is simply because local isometry preserves geodesics. □

Example 2.17. (1) We know that the geodesics on S^n are great circles, hence $\exp_p$ is defined on the whole T_pM. But $\exp_p$ is not injective, since

$$\exp_p(0) = \exp(2\pi v) = p$$

for unit vector v in T_pM.

(2) Let $M = S^1 \times \mathbb{R}$ be the cylinder. We know from elementary differential geometry that the geodesics on cylinder are directrix circles, helices and generatrix lines. Then in local charts $(e^{2\pi it}, s) \mapsto (t, s)$, we know $\exp_p$ is not injective in the direction $(1, 0)$, and injective in other directions.

We postpone the discussion on whether the exponential map can be defined on the whole tangent space, the answer is Hopf–Rinow Theorem, which will be discussed in next section.

Now we prove that geodesics are locally length-minimizing. For this, we introduce some local charts. Given a Riemannian manifold (M, g) and $p \in M$, let $\exp_p : B_\varepsilon(0) \to \exp(B_\varepsilon(0)) = B_\varepsilon(p)$ be a diffeomorphism.

Definition. We define *geodesic normal coordinate* as follows: Let $\{e_i\}$ be an orthonormal basis of Euclidean space (T_pM, δ), $\{\alpha^i\}$ be its dual basis. Then we define the coordinate by

$$q \in B_\varepsilon(p) \mapsto (\alpha^1(\exp_p^{-1}(q)), \cdots, \alpha^n(\exp_p^{-1}(q))).$$

Proposition 2.18. *Under geodesic normal coordinate, we have*

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

Proof. Since $\exp_p$ is a diffeomorphism, we have $\frac{\partial}{\partial x^i}\big|_p = \exp_{p*}|_0(e_i) = e_i$, hence $g_{ij} = \delta(e_i, e_j) = \delta_{ij}$. Moreover, let $x(t) = ty$ for $y \in T_pM - \{0\}$, then $x(t)$ is the coordinate of some geodesic in $B_\varepsilon(p)$, thus it satisfies the equation

$$\ddot{x}^k(t) + \Gamma_{ij}^k(x(t))\dot{x}^i(t)\dot{x}^j(t) = 0.$$

Since $\dot{x}^i = y^i \neq 0$, $\ddot{x}^k = 0$, we have $\Gamma_{ij}^k(ty) = 0$. Let $y \to 0$ we obtain the conclusion. □

Next we introduce the geodesic polar coordinate.

Definition. We define *geodesic polar coordinate* as follows: Let $(r, \theta^1, \cdots, \theta^{n-1})$ be a polar coordinate on Euclidean space $(T_p M, \delta)$, then we defined the coordinate by

$$q \in B_\varepsilon(p) - \{p\} \mapsto (r(\exp_p^{-1}(q)), \theta^1(\exp_p^{-1}(q)), \cdots, \theta^{n-1}(\exp_p^{-1}(q))).$$

Proposition 2.19. *Under geodesic polar coordinate, we have*

$$\left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \right\rangle = 1, \quad \left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta^i} \right\rangle = 0.$$

Proof. To make things clear, we write the inverse of geodesic polar coordinate as

$$F : (r, \omega) \mapsto \exp_p(r\omega)$$

for $r \in (0, \varepsilon)$, $\omega \in \mathbb{S}^{n-1}$. Then we use $\partial_0, \partial_1, \cdots, \partial_{n-1}$ to denote the tangent vectors in $(0, \varepsilon) \times \mathbb{S}^{n-1}$, we have

$$\frac{\partial}{\partial r} = F_*(\partial_0),$$

$$\frac{\partial}{\partial \theta^i} = F_*(\partial_i), \ i = 1, \cdots, n-1.$$

First we know that ∂_0 is the tangent vector of radial line $r\omega$, hence $\partial/\partial r$ is the tangent vector of a unit-speed radial geodesic, that is,

$$\left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \right\rangle = 1.$$

Moreover, we have

$$\frac{\partial}{\partial r} \left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta^i} \right\rangle = \left\langle \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta^i} \right\rangle + \left\langle \frac{\partial}{\partial r}, \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta^i} \right\rangle$$

$$= \left\langle \frac{\partial}{\partial r}, \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta^i} \right\rangle$$

$$= \left\langle \frac{\partial}{\partial r}, \nabla_{\frac{\partial}{\partial \theta^i}} \frac{\partial}{\partial r} \right\rangle$$

$$= \frac{1}{2} \frac{\partial}{\partial \theta^i} \left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \right\rangle$$

$$= 0,$$

hence $\left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta^i} \right\rangle$ is constant. However, if we let $r \to 0$, we have $\partial/\partial \theta^i \to 0$, therefore

$$\left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta^i} \right\rangle = 0.$$

$\square$

Corollary 2.20. *Under geodesic polar coordinate, the metric tensor has local expression*

$$g = \mathrm{d}r^2 + g_{ij}(r,\theta)\, \mathrm{d}\theta^i \otimes \mathrm{d}\theta^j,$$

where $[g_{ij}]_{i,j>0}$ is positive definite.

As an application, we prove that geodesics are locally length-minimizing as we promised.

Proposition 2.21. *Let $\gamma : [0,1] \to M$ be a geodesic contained in an open set U, where geodesic polar coordinate is defined on U. Let $\tilde{\gamma}$ be any curve contained in U with $\tilde{\gamma}(0) = \gamma(0) = p$, $\tilde{\gamma}(1) = \gamma(1) = q$. Then $L[\gamma] \leq L[\tilde{\gamma}]$.*

Proof. Let $q = \exp_p(v)$, φ is the geodesic polar coordinate. Then we have

$$\gamma(t) = \varphi(tr_0, \omega_0), \ \tilde{\gamma}(t) = \varphi(r(t), \omega(t))$$

such that $r(1) = r_0$, $\omega(t) \in \mathbb{S}^{n-1}$. Therefore

$$
\begin{aligned}
L[\gamma] &= \int_0^1 |\dot{\gamma}(t)|\, \mathrm{d}t \\
&= \int_0^1 r_0 \, \mathrm{d}t = r_0, \\
L[\tilde{\gamma}] &= \int_0^1 (|\dot{r}|^2(t) + g_{ij}\dot{\theta}^i(t)\dot{\theta}^j(t))^{1/2} \\
&\geq \int_0^1 |\dot{r}(t)|\, \mathrm{d}t \\
&\geq \int_0^1 \dot{r}(t)\, \mathrm{d}t = r_0.
\end{aligned}
$$

□

Remark 2.22. The hypothesis of Proposition 2.21 can be weakened to $\exp_p$ is an immersion on U.

2.4   Hopf–Rinow Theorem

We now answer the problem whether length-minimizing curve always exists. The answer is Hopf–Rinow Theorem.

We adopt the metric geometry version of Hopf-Rinow Theorem from [1], using geometric approach to prove the theorem and keeping the differential tools minimal. We first define

Definition. A Riemannian manifold (M, g) is called *(geodesically) complete* if M is complete as a metric space.

Our theorem is

Theorem 2.23 (Hopf–Rinow–Cohn–Vossen). *Let (M, g) be a Riemannian manifold, the following four assertions are equivalent:*

(1) *M has the Heine–Borel property, i.e. every closed geodesic ball is compact;*

(2) *M is geodesically complete;*

(3) *Every geodesic $\gamma : [0, a) \to M$ can be extended to a continuous curve $\overline{\gamma} : [0, a] \to M$;*

(4) *There is a point $p \in M$ such that every length-minimizing geodesic $\gamma : [0, a) \to M$ with $\gamma(0) = p$ can be extended to a continuous curve $\overline{\gamma} : [0, a] \to M$.*

We first establish several lemmas.

Lemma 2.24. *The length functional is lower semi-continuous in the following sense: let $\gamma_i, \gamma : [a, b] \to M$, if $\gamma_i \to \gamma$ pointwisely as $i \to \infty$, then*

$$L[\gamma] \leq \lim_{i \to \infty} L[\gamma_i].$$

Proof. Let Y be a partition of $[a, b]$ with $a = y_0 < \cdots < y_N = b$, denote

$$\Sigma(Y) = \sum_{i=1}^{N} d(\gamma(y_{i-1}), \gamma(y_i)).$$

Take $\varepsilon > 0$ and fix a partition Z such that $L[\gamma] - \Sigma(Z) < \varepsilon$. Now consider $\Sigma_j(Z)$ for curves γ_j corresponding to same partition Z. Choose j so large that the inequality $d(\gamma_j(z_i), \gamma(z_i)) < \varepsilon$ holds for all $z_i \in Z$. Then

$$L[\gamma] \leq \Sigma(Z) + \varepsilon \leq \Sigma_j(Z) + \varepsilon + (N+1)\varepsilon \leq L[\gamma_j] + (N+2)\varepsilon.$$

Since ε is arbitrary, the lemma holds. $\qquad\square$

Lemma 2.25. *Let γ_i be length-minimizing curves for $i = 1, 2, \cdots$, suppose $\gamma_i \to \gamma$ pointwisely, then γ also minimizes length.*

Proof. Let γ has end points p and q, then since $L[\gamma_i]$ equals to the distance between its end points, $L[\gamma_i] \to d(p, q)$. By the lower semi-continuity of length functional, we have

$$L[\gamma] \leq \lim_{i \to \infty} L[\gamma_i] = d(p, q) \leq L[\gamma]. \qquad\square$$

Lemma 2.26. *Let (M, g) be a compact Riemannian manifold (maybe with boundary), then any two points $p, q \in M$ can be joined by a length-minimizing curve.*

Proof. By the definition of distance, there exists a sequence of curves $\{\gamma_i\}$ with constant speed such that $\gamma_i(0) = p$, $\gamma_i(1) = q$ and $L[\gamma_i] \to d(p,q)$. Then let $L[\gamma_i] < A$ for all $i = 1, 2, \cdots$, we have $d(p,x) \le L[\gamma_i|_{[0,a]}] < A$ for $x = \gamma_i(a)$, hence the family $\{\gamma_i\}$ is uniformly bounded. Moreover, fix $\varepsilon > 0$, let $\delta = \varepsilon/A$, we have

$$d(x,y) \le L[\gamma_i|_{[a,b]}] \le A(b-a) < \varepsilon$$

provided $x = \gamma_i(a)$, $y = \gamma_i(b)$, and $b - a < \delta$ for any $i = 1, 2, \cdots$, hence the family $\{\gamma_i\}$ is equicontinuous. Then the family verifies the conditions of Arzela–Ascoli Theorem, it converges (up to a subsequence) to a curve γ. Then by the lower semi-continuity of length functional, we have

$$L[\gamma] \le \lim_{i \to \infty} L[\gamma_i] = d(p,q) \le L[\gamma].$$

Thus γ is length-minimizing. $\square$

Proof of Theorem 2.23. Implications (1) $\implies$ (2) $\implies$ (3) $\implies$ (4) are all easy, we prove (4) $\implies$ (1).

Let $R = \sup\{\overline{B_r(p)} \text{ is a compact set}\}$, then $R > 0$ since manifolds are locally compact, $\overline{B_r(p)}$ is compact for r sufficiently small. We argue by contradiction. Suppose $R < +\infty$, that is, there exists noncompact geodesic balls. The argument is divided into two steps.

1. First we prove $B_R(p)$ is sequentially compact. Let $\{p_i\} \subset B_R(p)$, set $d(p, p_i) = r_i$. We may assume $r_i \to R$ as $i \to \infty$, otherwise $\{p_i\}$ is eventually contained in a smaller geodesic ball, and it has a convergent subsequence by the definition of R.

Now let $\gamma_i : [0, r_i] \to M$ be a length-minimizing curve joining p and p_i, whose existence is guaranteed by Lemma 2.26. Notice that γ_i's are parametrized by arc-length. We can choose a subsequence of $\{\gamma_i\}$ such that the restrictions of the curves to $[0, r_1]$ converge by Arzela–Ascoli Theorem. From this subsequence, we can choose a further subsequence such that the restrictions to $[0, r_2]$ converge, and so on. Then by Cantor diagonal procedure, we have a sequence $\{\gamma_{i_n}\}$ such that for $t \in [0, R)$, $\gamma_{i_n}(t)$ is well-defined for n sufficiently large and $\gamma_{i_n}(t) \to \gamma(t)$ as $n \to \infty$. Moreover, Arzela–Ascoli Theorem asserts γ is smooth for any restriction to $[0, r]$ provided $r < R$ by uniform convergence, hence γ is smooth.

Now by Lemma 2.25, γ is a length-minimizing curve, hence by the hypothesis (iv), γ can be extended to a continuous curve $\overline{\gamma} : [0, R] \to M$. Then let $q = \overline{\gamma}(R)$, fix $\varepsilon > 0$, take n sufficiently large such that $d(p_{i_n}, \gamma_{r_{i_n}}) < \varepsilon/2$ and $d(\gamma(r_{i_n}, q)) = R - r_{i_n} < \varepsilon/2$, we have

$$d(p_{i_n}, q) \le d(p_{i_n}, \gamma(r_{i_n})) + d(\gamma(r_{i_n}), q) < \varepsilon.$$

Hence $p_{i_n} \to q$ as $n \to \infty$, we have $B_R(p)$ is sequentially compact.

2. Since $B_R(p)$ is sequentially compact, we have $\overline{B_R(p)}$ is compact. Now we show $\overline{B_{R+\varepsilon}(p)}$ is compact for some $\varepsilon > 0$. Since M is locally compact, for every $x \in \overline{B_R(0)}$ there is an $r(x) > 0$ such that $\overline{B_{r(x)}(x)}$ is compact. Then we can choose finite $x_i \in \overline{B_R(p)}$ such that $\{B_{r(x_i)}(x_i)\}$ covers $\overline{B_R(p)}$. The union of these geodesic balls is sequentially compact and contains the geodesic ball $B_{R+\varepsilon}(p)$ for $0 < \varepsilon < \min\{r(x_i)\}$. Hence $\overline{B_{R+\varepsilon}(p)}$ is compact, this contradicts the choice of R. $\qquad\square$

We have two corollaries.

Corollary 2.27 (Hopf–Rinow Theorem). *Let (M, g) be a Riemannian manifold, the following are equivalent:*

(1) M is geodesically complete;

(2) $\exp_p$ is defined on whole T_pM for any $p \in M$;

(3) $\exp_p$ is defined on whole T_pM for one $p \in M$.

Proof. (1) $\implies$ (2): Let $S = \{r \in \mathbb{R}_{>0} : \gamma : [0, r)$ can be extended to $r\}$. We prove S is both open and closed, then the implication holds. Let $r \in S$, for any $q \in \partial B_r(p)$, define $r(q)$ as follows: let unit-speed geodesic $\gamma : [0, r] \to M$ joins p and q, then there exists a geodesic $\tilde{\gamma} : [r, r + r(q))$ such that $\tilde{\gamma}(r) = q$, $\dot{\tilde{\gamma}}(r) = \dot{\gamma}(r)$, that is, γ can be extended to $[0, r + r(q))$. Cover $\partial B_r(p)$ by finite many $B_{r(q_i)}(q_i)$, then for $0 < \varepsilon < \min\{r(q_i)\}$, $B_{r+\varepsilon}(p)$ is in the union of $B_r(p)$ and $B_{r(q_i)}(q_i)$, hence any geodesics $\sigma : [0, r + \varepsilon)$ can be extended to $\sigma : [0, r + \varepsilon] \to M$. This implies $(r - \varepsilon, r + \varepsilon) \subset S$, S is open.

Now assume $\{r_i\} \subset S$ converges to r. Let $\gamma : [0, r)$ be a geodesic, then define $\gamma(r) = \lim_{i \to \infty} \gamma(r_i)$, the limit exists since M is complete. Hence $r \in S$, S is closed.

(2) $\implies$ (3) is trivial.

(3) $\implies$ (1): If $\gamma : [0, a)$ is a length-minimizing geodesic with $\gamma(0) = p$, then $\gamma(t) = \exp_p(tv)$ for some v, and clearly it can be extended to a. $\qquad\square$

Corollary 2.28. *If M is geodesically complete, then any two points can be joined by a length-minimizing geodesic.*

Proof. Let $p, q \in M$, $R > d(p, q)$. Then $q \in \overline{B_R(p)}$, the latter one is bounded and closed, hence by Theorem 2.23, it is compact. By Lemma 2.26, there is a length-minimizing geodesic joining p and q. $\qquad\square$

Chapter 3

Curvature

3.1 Curvature Tensor and Curvaure Endomorphism

Riemann Curvature Tensor

We calculated in Proposition 2.14 that $\exp_{p*}|_0$ is identity, naturally we have the following problem:

Problem. Calculate $\exp_{p*}|_v : T_v(T_pM) \to T_{\exp_p(v)}M$.

Solution. To evaluate $\exp_{p*}|_v(\xi)$, we choose a line $v + s\xi$, and then

$$\exp_{p*}|_v(\xi) = \left.\frac{d}{ds}\right|_{s=0} \exp_p(v + s\xi).$$

Now we can introduce the one parameter family of geodesics

$$\gamma(t,s) = \exp_p(t(v + s\xi)),$$

and denote $\gamma(t) = \gamma(t,0)$. We calculate the variation vector field $J(t)$ of γ, and obtain the result by taking $t = 1$. Let $J_s(t) = \frac{\partial}{\partial s}\gamma(t,s)$, then $\dot{J}_s(t) = \nabla_{\dot{\gamma}_s(t)}\frac{\partial\gamma}{\partial s}$. Since $\nabla_{\dot{\gamma}_s(t)}\frac{\partial\gamma}{\partial t} = 0$, we have

$$\ddot{J}_s(t) = \nabla_{\dot{\gamma}_s(t)}\nabla_{\dot{\gamma}_s(t)}\frac{\partial\gamma}{\partial s}$$

$$= \nabla_{\dot{\gamma}_s(t)}\nabla_{\frac{\partial\gamma}{\partial s}}\frac{\partial\gamma}{\partial t}$$

$$= \nabla_{\frac{\partial\gamma}{\partial t}}\nabla_{\frac{\partial\gamma}{\partial s}}\frac{\partial\gamma}{\partial t} - \nabla_{\frac{\partial\gamma}{\partial s}}\nabla_{\frac{\partial\gamma}{\partial t}}\frac{\partial\gamma}{\partial t}.$$

Moreover, we have $[\partial_t, \partial_s] = 0$, then we denote

$$R\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial s}\right) = \nabla_{\frac{\partial}{\partial s}}\nabla_{\frac{\partial}{\partial t}} - \nabla_{\frac{\partial}{\partial t}}\nabla_{\frac{\partial}{\partial s}} + \nabla_{[\frac{\partial}{\partial t}, \frac{\partial}{\partial s}]},$$

we obtain

$$\ddot{J}_s(t) + R\left(\frac{\partial\gamma}{\partial t}, \frac{\partial\gamma}{\partial s}\right)\frac{\partial\gamma}{\partial t} = 0.$$

Take $s = 0$, we have

$$\ddot{J}(t) + R(\dot{\gamma}(t), J(t))\dot{\gamma}(t) = 0. \tag{3.1}$$

We will show later that (3.1) is a system of ordinary differential equations, hence by solving the system with given initial value and taking $t = 1$, we obtain the answer of the problem.   □

We make it into a definition.

Definition. The *Riemann curvature tensor* $R : \Gamma(TM) \times \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ is defined by

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

Remark 3.1. Many authors define the Riemann curvature tensor as the negative of above definition, e.g. in [6]. Please be careful with the sign of the tensor.

We need to explain the name "tensor", so we must show R is truly tensorial.

Lemma 3.2. *R is a tensor.*

Proof. $R(X, Y)Z$ is clearly tensorial in X and Y, we show that $R(X, Y)(fZ) = fR(X, Y)Z$ for $f \in C^\infty(M)$. We have

$$\nabla_Y\nabla_X(fZ) = \nabla_Y((Xf)Z + f\nabla_X Z)$$
$$= (YXf)Z + (Xf)\nabla_Y Z + (Yf)\nabla_X Z + f\nabla_Y\nabla_X Z$$
$$-\nabla_X\nabla_Y(fZ) = -(XYf)Z - (Yf)\nabla_X Z - (Xf)\nabla_Y Z - f\nabla_X\nabla_Y Z$$
$$\nabla_{[X,Y]}(fZ) = ([X,Y]f)Z + f\nabla_{X,Y}Z,$$

add all three equalities and we reach the conclusion.   □

Now we can look at Riemann curvature tensor locally. A rather complicated calculation shows:

Lemma 3.3. *Let* $R = R_{ijk}{}^l \otimes dx^i \otimes dx^j \otimes dx^k \otimes \partial_l$, *then*

$$R_{ijk}{}^l = \partial_j\Gamma_{ik}^l - \partial_i\Gamma_{jk}^l + \Gamma_{ik}^m\Gamma_{jm}^l - \Gamma_{jk}^m\Gamma_{im}^l.$$

We also define a $(0,4)$-tensor by lowering the l index of $Rlijk$, that is:

Definition. $R(X, Y, Z, W) := \langle R(X, Y)Z, W \rangle$ is also called *Riemann curvature tensor*.

Example 3.4. Euclidean space $(\mathbb{R}^n, \delta)$ has $R \equiv 0$. Any metric admits zero curvature is called *flat*.

Curvature Endomorphism

Curvature also appears in another scene. Let us consider the second covariant differential of a tensor.

Definition. Let $T(\cdots)$ be a tensor, denote $(\nabla_{X,Y} T)(\cdots) := \nabla(\nabla T)(\cdots, Y, X)$.

Proposition 3.5. *For (r,s)-tensor T, we have*

$$\nabla_{X,Y} T = \nabla_X(\nabla_Y T) - \nabla_{\nabla_X Y} T. \tag{3.2}$$

Proof. Since covariant derivative commutes with contraction, we have

$$\begin{aligned}
\nabla_X(\nabla_Y T) &= \nabla_X(\text{tr}_{1,s+2} Y \otimes \nabla T) \\
&= \text{tr}_{1,s+2}(\nabla_X(Y \otimes \nabla T)) \\
&= \text{tr}_{1,s+2}((\nabla_X Y) \otimes \nabla T + Y \otimes (\nabla_X \nabla T)) \\
&= \nabla_{\nabla_X Y} T + \nabla(\nabla T)(\cdots, Y, X) \\
&= \nabla_{\nabla_X Y} T + \nabla_{X,Y} T,
\end{aligned}$$

then the result follows. $\qquad \square$

Definition. Define the *curvature endomorphism* $R(X, Y)$ on (r,s)-tensors by

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

Remark 3.6. We need to show $R(X, Y)T$ is tensorial in T so that $R(X, Y)$ is a well-defined endomorphism. This is similar to Lemma 3.2.

Proposition 3.7. *For any (r,s)-tensor T, we have the following* Ricci identity*:*

$$\nabla_{Y,X} T - \nabla_{X,Y} T = R(X, Y)T.$$

Moreover, we have a explicit formula

$$\begin{aligned}
&(R(X, Y)T)(\omega_1, \cdots, \omega_r, X_1, \cdots, X_s) \\
&= -\sum_{i=1}^{r} T(\omega_1, \cdots, R(X, Y)\omega_i, \cdots, \omega_r, X_1, \cdots, X_s) \\
&\quad - \sum_{j=1}^{s} T(\omega_1, \cdots, \omega_r, X_1, \cdots, R(X, Y)X_j, \cdots, X_s)
\end{aligned} \tag{3.3}$$

Proof. Using equation (3.2), we have

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

the second equality is torsion-freeness. Since $R(X,Y)$ clearly satisfies Leibniz Law and commutes with contraction, we can deduce the formula in the same way as covariant derivative. $\square$

Ricci identity shows the curvature appears when we interchange the second covariant differential.

Properties of Curvature

Proposition 3.8. *Riemann curvature tensor has following symmetric properties: For* $X, Y, Z, W \in \Gamma(TM)$, *we have*

(1) $R(X,Y,Z,W) = -R(Y,X,Z,W) = -R(X,Y,W,Z)$;

(2) $R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0$ *(First Bianchi identity)*;

(3) $R(X,Y,Z,W) = R(Z,W,X,Y)$;

(4) $(\nabla_X R)(Y,Z) + (\nabla_Y R)(Z,X) + (\nabla_Z R)(X,Y) = 0$ *(Second Bianchi identity).*

Proof. (1) The first equality is evident. We show the second equality. Consider the Hessian of $g(Z,W)$, then we have

$$\nabla^2 g(Z,W)(X,Y) = \langle \nabla^2_{Y,X}Z, W\rangle + \langle \nabla_X Z, \nabla_Y W\rangle + \langle \nabla_Y Z, \nabla_X W\rangle + \langle Z, \nabla^2_{Y,X}W\rangle.$$

Interchange X, Y, we have

$$\nabla^2 g(Z,W)(Y,X) = \langle \nabla^2_{X,Y}Z, W\rangle + \langle \nabla_Y Z, \nabla_X W\rangle + \langle \nabla_X Z, \nabla_Y W\rangle + \langle Z, \nabla^2_{X,Y}W\rangle.$$

These two equations must equal, hence

$$\langle \nabla^2_{Y,X}Z, W\rangle - \langle \nabla^2_{X,Y}Z, W\rangle = \langle Z, \nabla^2_{X,Y}W\rangle - \langle Z, \nabla^2_{Y,X}W\rangle,$$

this is equivalent to

$$R(X,Y,Z,W) = R(Y,X,W,Z).$$

(2) Since $R(X,Y)Z$ is tensorial, we can assume X, Y, Z are frames. Then all Lie bracket between X, Y, Z vanish, we have

$$\sum_{\text{cyc}} R(X,Y)Z = \nabla_Y\nabla_X Z - \nabla_X\nabla_Y Z + \nabla_Z\nabla_Y X - \nabla_Y\nabla_Z X$$

$$+ \nabla_X \nabla_Z Y - \nabla_Z \nabla_X Y$$
$$= \nabla_Y [X, Z] + \nabla_Z [Y, X] + \nabla_X [Z, Y]$$
$$= 0.$$

(3) By (1) and (2) we have

$$
\begin{aligned}
R(X, Y, Z, W) = & - R(Z, X, Y, W) - R(Y, Z, X, W) \\
= & R(Z, X, W, Y) + R(Y, Z, W, X) \\
= & - R(W, Z, X, Y) - R(X, W, Z, Y) \\
& - R(W, Y, Z, X) - R(Z, W, Y, X) \\
= & 2R(Z, W, X, Y) + R(X, W, Y, Z) + R(W, Y, X, Z) \\
= & 2R(Z, W, X, Y) - R(Y, X, W, Z) \\
= & 2R(Z, W, X, Y) - R(X, Y, Z, W),
\end{aligned}
$$

which implies $2R(X, Y, Z, W) = 2R(Z, W, X, Y)$.

(4) By the definition of covariant derivative, we have

$$
\begin{aligned}
(\nabla_X R)(Y, Z) = & [\nabla_X, R(Y, Z)] - R(\nabla_X Y, Z) - R(Y, \nabla_X Z) \\
= & [\nabla_X, \nabla_{[Y,Z]}] - [\nabla_X, [\nabla_Y, \nabla_Z]] - R(\nabla_X Y, Z) - R(Y, \nabla_X Z) \\
= & \nabla_{[X,[Y,Z]]} - [\nabla_X, [\nabla_Y, \nabla_Z]] - R(X, [Y, Z]) \\
& - R(\nabla_X Y, Z) - R(Y, \nabla_X Z).
\end{aligned}
$$

Take summation cyclically, we see that terms involving ∇ vanish because of Jacobi identity. Moreover, since $\nabla_X Y - \nabla_Y X = [X, Y]$, the terms involving R also vanish. $\qquad \square$

3.2 Sectional, Ricci and Scalar Curvature

We now define some special curvatures. Fix a Riemannian manifold (M, g).

Definition. Let $p \in M$, $\pi \subset T_p M$ be a 2-plane, $\pi = \mathrm{Span}\{u, v\}$. Then define the *sectional curvature* of π at p to be

$$\mathrm{Sect}_p(\pi) = \frac{R_p(u, v, u, v)}{|u|^2 |v|^2 - \langle u, v \rangle^2}.$$

Remark 3.9. One can show that sectional curvature does not depend on the choice of basis. For a proof, see [2, Proposition 3.1].

Proposition 3.10. *The sectional curvature determines the curvature tensor.*

Proof. Let R, R' have same sectional curvature, denote $\tilde{R} = R - R'$, then $\widetilde{\text{Sect}} = 0$. We show that $\tilde{R} = 0$. First, we have

$$
\begin{aligned}
\tilde{R}(X, Y, X, W) &= \tilde{R}(X, Y - W, X, W) + \tilde{R}(X, W, X, W) \\
&= \tilde{R}(X, Y - W, X, W - Y) + \tilde{R}(X, Y - W, X, Y) \\
&= \tilde{R}(X, Y, X, Y) - \tilde{R}(X, W, X, Y) \\
&= \tilde{R}(X, Y, X, W),
\end{aligned}
$$

hence $\tilde{R}(X, Y, X, W) = 0$. Then we have

$$
\begin{aligned}
\tilde{R}(X, Y, Z, W) &= \tilde{R}(X, Y, Z, W - X) + \tilde{R}(X, Y, Z, X) \\
&= \tilde{R}(X - W, Y, Z, W - X) + \tilde{R}(W, Y, Z, W - X) \\
&= \tilde{R}(W, Y, Z, W) - \tilde{R}(W, Y, Z, X) \\
&= -\tilde{R}(W, Y, Z, X).
\end{aligned}
$$

By the same reasoning, we have

$$
\begin{aligned}
\tilde{R}(X, Y, Z, W) &= \tilde{R}(Z, W, X, Y) \\
&= -\tilde{R}(Y, W, X, Z) \\
&= -\tilde{R}(X, Z, Y, W).
\end{aligned}
$$

Passing the result to the $(1, 3)$-tensor, by first Bianchi identity, we have

$$
\tilde{R}(X, Z)Y + \tilde{R}(Y, Z)X + \tilde{R}(Z, X)Y = 0,
$$

which implies $\tilde{R}(Y, Z)X = 0$. Then

$$
\begin{aligned}
0 &= \tilde{R}(Y, Z, X, W) \\
&= -\tilde{R}(W, Z, X, Y) \\
&= -\tilde{R}(X, Y, W, Z) \\
&= \tilde{R}(X, Y, Z, W).
\end{aligned}
$$

Thus we proved $\tilde{R} = 0$, and the result follows. $\square$

We mention here a little observation.

Lemma 3.11. *If* Sect_p *is constant for all* 2-*planes in* $T_p M$, *say* K_p, *then we have*

$$
R_p(X, Y, Z, W) = K_p(\langle X, Z \rangle \langle Y, W \rangle - \langle X, W \rangle \langle Y, Z \rangle).
$$

Then we can prove the following Schur's Theorem.

Theorem 3.12 (Schur). *Let (M^n, g) be a Riemannian manifold with $n \geq 3$. If $\mathrm{Sect}_p(\pi)$ is independent from $\pi \subset T_p M$ for all $p \in M$, then M has constant sectional curvature.*

Proof. Since the tensor

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

evidently satisfies the Proposition 3.8, Proposition 3.10 shows $R = fR'$ for some $f \in C^\infty(M)$. We show that f is constant. Since $n \geq 3$, there exists three orthonormal vectors X, Y, Z. Take W arbitrary, then by second Bianchi identity, we have

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

Since $\nabla_X R = \nabla_X(fR') = (Xf)R' + f(\nabla_X R')$, by taking summation cyclically we obtain

$$(Xf)R'(Y, Z)W + (Yf)R'(Z, X)W + (Zf)R'(X, Y)W = 0.$$

Since the sectional curvature is constant for any $\pi \subset T_p M$, the "sectional curvature" corresponding to R' is also constant, hence we can use Lemma 3.11 to obtain

$$\begin{aligned}
0 =& (Xf)K(\langle Y, W \rangle Z - \langle Z, W \rangle Y) \\
& + (Yf)K(\langle Z, W \rangle X - \langle X, W \rangle Z) \\
& + (Zf)K(\langle X, W \rangle Y - \langle Y, W \rangle X),
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
0 =& ((Xf)\langle Y, W \rangle - (Yf)\langle X, W \rangle)Z \\
& + ((Yf)\langle Z, W \rangle - (Zf)\langle Y, W \rangle)X \\
& + ((Zf)\langle X, W \rangle - (Xf)\langle Z, W \rangle)Y.
\end{aligned}$$

Since X, Y, Z are orthonormal, the coefficient of X, Y, Z must all equal to 0. Thus by taking $W = Y$, we obtain

$$Xf = (Xf)\langle Y, Y \rangle = (Yf)\langle X, Y \rangle = 0.$$

Since X is arbitrary, we have $f \equiv$ const. This deduces M has constant sectional curvature. $\qquad\square$

Definition. Let $R_{ijk}{}^l$ be the local expression of Riemann curvature tensor, then we define the *Ricci curvature* by

$$\mathrm{Ric}_j{}^l = \mathrm{tr}_{13} R_{ijk}{}^l.$$

Equivalently, let $\{e_i\}$ be an orthonormal basis at p, then

$$\mathrm{Ric}_p(X) = \sum_{i=1}^{n} R(e_i, X)e_i.$$

Clearly, Ric is a self-adjoint linear transformation by Proposition 3.8 (3).

Definition. We define the *scalar curvature* by taking trace of Ric, or equivalently

$$\mathrm{Scal}(p) = \sum_{i=1}^{n} \langle \mathrm{Ric}_p(e_i), e_i \rangle$$

for an orthonormal basis $\{e_i\}$ at p.

Definition. A Riemannian manifold (M, g) is called an *Einstein manifold* if $\mathrm{Ric} = \lambda\delta$ for some $\lambda \in C^\infty(M)$.

We have another theorem of Schur on Einstein manifolds.

Theorem 3.13 (Schur)**.** *Let M^n be an Einstein manifold with $n \geq 3$, then M has constant scalar curvature.*

First we need a lemma.

Lemma 3.14. *For any metric g, we have*

$$2\,\mathrm{tr}_{13}\,\nabla\,\mathrm{Ric} = \mathrm{d}\,\mathrm{Scal}.$$

Proof. We check the equation locally. The second Bianchi identity can be written as

$$R_{ijm}{}^l{}_{;k} + R_{kim}{}^l{}_{;j} + R_{jkm}{}^l{}_{;i} = 0,$$

or equivalently

$$R_{ijm}{}^l{}_{;k} - R_{ikm}{}^l{}_{;j} + R_{jkm}{}^l{}_{;i} = 0.$$

Multiply g^{im} we obtain

$$0 = \mathrm{Ric}_j{}^l{}_{;k} - \mathrm{Ric}_k{}^l{}_{;j} + R_{jk}{}^{il}{}_{;i}.$$

Let $l = j$, we obtain

$$\begin{aligned}
0 &= \mathrm{Ric}_j{}^j{}_{;k} - \mathrm{Ric}_k{}^j{}_{;j} - R_{kj}{}^{ij}{}_{;i} \\
&= \mathrm{Scal}_{;k} - \mathrm{Ric}_k{}^j{}_{;j} - \mathrm{Ric}_k{}^i{}_{;i} \\
&= \mathrm{Scal}_{;k} - 2\mathrm{Ric}_k{}^i{}_{;i}.
\end{aligned}$$

Then the result follows. $\square$

Proof of Theorem 3.13. Let $\mathrm{Ric} = \lambda \delta$, then by the lemma,

$$
\begin{aligned}
\mathrm{d\,Scal} &= 2\,\mathrm{tr}_{1,3}\,\nabla\,\mathrm{Ric} \\
&= 2\,\mathrm{tr}_{1,3}\,\nabla(\lambda\delta) \\
&= 2\,\mathrm{tr}_{1,3}(\delta \otimes \mathrm{d}\lambda) \\
&= 2\,\mathrm{d}\lambda.
\end{aligned}
$$

However, we have

$$
\begin{aligned}
\mathrm{d\,Scal} &= \mathrm{d}\,\mathrm{tr}_{1,2}\,\mathrm{Ric} \\
&= \mathrm{d}\lambda\,\mathrm{tr}_{1,2}\,\delta \\
&= n\,\mathrm{d}\lambda.
\end{aligned}
$$

This means $(n-2)\,\mathrm{d\,Scal} = 0$, which implies $\mathrm{Scal} \equiv \mathrm{const}$ for $n \geq 3$. $\qquad\square$

3.3 Jacobi Fields

At the beginning of Section 3.1, we introduced the one parameter family of geodesics $\gamma(t,s) = \exp_p(t(v + s\xi))$. It is a variation of curve $\gamma(t) = \gamma(t,0)$, and its variation vector field J satisfies (3.1). Motivated by this, we give the definition of Jacobi fields.

Definition. Let γ be a geodesic, a vector field J along γ is called a *Jacobi field*, if $\ddot{J}(t) + R(\dot{\gamma}(t), J(t))\dot{\gamma}(t) = 0$.

Let $\{e_i(t)\}$ be a parallel frame along γ, and $J(t) = f^i e_i$. Define $a^i_j(t) = R(\dot{\gamma}(t), e_i(t), \dot{\gamma}(t), e_j(t))$, then the equation of Jacobi field is equivalent to

$$
\ddot{f}^i(t) + a^i_j(t) f^j(t) = 0, \ i = 1, \cdots, n.
$$

Hence (3.1) is indeed a system of ODE. By ODE theory, given $f^i(0), \dot{f}^i(0) \ i = 1, \cdots, n$, the $f^i(t)$'s are uniquely determined. Translate into geometric language, a Jacobi field J is determined by $J(0)$ and $\dot{J}(0)$.

We summarize above discussion.

Proposition 3.15. *Let $\mathscr{J}\gamma$ be the vector space of all Jacobi fields along γ, then*

$$
\dim \mathscr{J}(\gamma) = 2n.
$$

The proposition below shows only Jacobi fields that perpendicular to γ are interesting.

Proposition 3.16. *Let γ be a geodesic, J is a Jacobi field along γ. Then we have the decomposition*

$$J(t) = J^{\perp}(t) + (at + b)\dot{\gamma}(t),$$

where $J^{\perp}(t) \perp \dot{\gamma}(t)$. If a Jacobi field is perpendicular to γ, then we call it a normal Jacobi field.

Proof. We have

$$\frac{d^2}{dt^2}\langle J(t), \dot{\gamma}(t)\rangle = \langle \ddot{J}(t), \dot{\gamma}(t)\rangle$$
$$= -\langle R(\dot{\gamma}(t), J(t))\dot{\gamma}(t), \dot{\gamma}(t)\rangle$$
$$= 0. \qquad \square$$

At the beginning of Section 3.1, we see that a one parameter family of geodesics gives rise to a Jacobi field. The converse is also true.

Proposition 3.17. *Let J be a Jacobi field, then J is the variation field of some one parameter family of geodesics.*

Proof. Given a Jacobi field J, it is determined by $J(0)$ and $\dot{J}(0)$. Let $\zeta(s)$ be the geodesic with initial tangent vector $J(0)$, and $T(s)$ be its tangent vector field. Let $W(s)$ be parallel along ζ and $W(0) = \dot{J}(0)$. Now set $\gamma(t,s) = \exp_{\zeta(s)}(t(T(s) + sW(s)))$. Then

$$\left[\frac{\partial}{\partial t}, \frac{\partial}{\partial s}\right] = \gamma_*[\partial_t, \partial_s] = 0,$$

hence we can interchange the partial derivative. Let U be the variation field of $\gamma(t,s)$, then U is a Jacobi field. We verify U and J has same initial value. We have

$$U(0) = \left.\frac{\partial}{\partial s}\right|_{s=0} \gamma(0,s)$$
$$= \left.\frac{\partial}{\partial s}\right|_{s=0} \exp_{\zeta(s)}(0)$$
$$= \left.\frac{\partial}{\partial s}\right|_{s=0} \zeta(s)$$
$$= \dot{\zeta}(0) = J(0),$$

and

$$\dot{U}(0) = \left.\frac{\partial}{\partial t}\right|_{t=0} \left.\frac{\partial}{\partial s}\right|_{s=0} \exp_{\zeta(s)}(t(T(s) + sW(s)))$$

$$= \frac{\partial}{\partial s}\Big|_{t=0} \frac{\partial}{\partial t}\Big|_{s=0} \exp_{\zeta(s)}(t(T(s) + sW(s)))$$

$$= \frac{\partial}{\partial s}\Big|_{s=0} \exp_{\zeta s*}|_0(T(s) + sW(s))$$

$$= \frac{\partial}{\partial s}\Big|_{s=0} (T(s) + sW(s))$$

$$= W(0) = \dot{J}(0). \qquad \square$$

Now we can completely answer the calculation problem of differential of exponential map.

Proposition 3.18. *Let a Jacobi field J along $\gamma(t) = \exp_p(tv)$ satisfy $J(0) = 0$, $\dot{J}(0) = \xi$, then $J(t) = \exp_{p*}|_{tv}(t\xi)$.*

Proof. Let J be the variation vector field of $\gamma(t, s) = \exp_p(t(v + s\xi))$. Then

$$J(t) = \frac{\partial}{\partial s} \exp_p(t(v + s\xi))$$

$$= \exp_{p*}|_{tv}(t\xi).$$

Now consider the initial condition for J, we have

$$J(0) = \exp_{p*}|_0(0) = 0,$$

and

$$\dot{J}(0) = \frac{\partial}{\partial t}\Big|_{t=0} \frac{\partial}{\partial s}\Big|_{s=0} \exp_p(t(v + s\xi))$$

$$= \dot{J}(0) = \frac{\partial}{\partial s}\Big|_{s=0} \frac{\partial}{\partial t}\Big|_{t=0} \exp_p(t(v + s\xi))$$

$$= \frac{\partial}{\partial s}\Big|_{s=0} \exp_{p*}|_0(v + s\xi)$$

$$= \xi.$$

Then the conclusion follows by the uniqueness of Jacobi fields. $\qquad \square$

Proposition 3.19 (Gauss Lemma). $\langle \exp_{p*}|_v(\xi), \dot{\gamma}_v(1)\rangle = \langle \xi, v\rangle.$

Proof. Let J be a Jacobi field along γ_v with $J(0) = 0$, $\dot{J}(0) = \xi$. Then by Proposition 3.18, we have

$$\langle \exp_{p*}|_v(\xi), \dot{\gamma}_v(1)\rangle = \langle J(1), \dot{\gamma}_v(1)\rangle.$$

We differentiate above inner product, obtaining

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle J(t), \dot{\gamma}_v(t)\rangle = \langle \dot{J}(t), \dot{\gamma}_v(t),\rangle$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \dot{J}(t), \dot{\gamma}_v(t)\rangle = \langle \ddot{J}(t), \dot{\gamma}_v(t)\rangle$$
$$= -\langle R(\dot{\gamma}_v, J)\dot{\gamma}_v, \dot{\gamma}_v\rangle$$
$$= 0.$$

Hence $\langle \dot{J}(t), \dot{\gamma}_v(t)\rangle$ is constant. Clearly $\langle \dot{J}(0), \dot{\gamma}_v(0)\rangle = \langle \xi, v\rangle$, then we have

$$\langle J(1), \dot{\gamma}_v(1)\rangle - \langle J(0), \dot{\gamma}_v(0)\rangle = \int_0^1 \langle \xi, v\rangle \,\mathrm{d}t$$
$$= \langle \xi, v\rangle.$$

Notice that $\langle J(0), \dot{\gamma}_v(0)\rangle = 0$ and we obtain the result.   □

Finally we consider the case where $\exp_{p*}$ degenerates.

Definition. Let γ be a geodesic starting at p, if $\exp_{p*}|_{\gamma(t_0)}$ is degenerate at $\gamma(t_0)$, then $\gamma(t_0)$ is called a *conjugate point* of p.

A simple criterion for conjugate points is

Proposition 3.20. $\exp_{p*}$ *degenerates at* $\gamma(t_0)$ *if and only if there is a nontrivial Jacobi field* J *with* $J(0) = J(t_0) = 0$.

Proof. $\exp_{p*}|_{\gamma(t_0)}(\xi) = 0$ if and only if Jacobi field $J(t) = \frac{\partial}{\partial s}\exp_p(t(v + s\xi))$ satisfies $J(0) = J(t_0) = 0$.   □

Remark 3.21. This proposition shows conjugate is symmetric.

3.4   Index Form

Using the first variation of energy, we conclude that a length-minimizing curve must be a geodesic. However, the converse is generally not true. For instance, a geodesic on sphere starting from the north pole and crossing the south pole is not length-minimizing. We need to find a sufficient condition for a geodesic to be length-minimizing. Calculus hints to us that we should differentiate energy twice. This lead to the second variation formula.

Proposition 3.22 (Second variation formula). *Let $\gamma : [0, a] \times (-\varepsilon, \varepsilon)$ be a variation, and $\gamma_0(t) = \gamma(t, 0)$ is a geodesic, then*

$$E''(0) = \int_0^a |\dot{V}(t)|^2 - \langle R(\dot{\gamma}_0(t), V(t))\dot{\gamma}_0(t), V(t) \rangle \, \mathrm{d}t + \boxed{\langle \nabla_{V(t)} V(t), \dot{\gamma}_0(t) \rangle |_0^a}.$$

The boxed term is called boundary term, *and it vanishes when the variation is proper.*

Proof. We take the expression of $E'(s)$ in the first variation formula and differentiate, we have

$$E''(s) = \int_0^a \frac{\partial}{\partial s} \left\langle \frac{\partial \gamma}{\partial t}, \nabla_{\frac{\partial \gamma}{\partial s}} \frac{\partial \gamma}{\partial t} \right\rangle \mathrm{d}t$$

$$= \int_0^a \left\langle \nabla_{\frac{\partial \gamma}{\partial s}} \frac{\partial \gamma}{\partial t}, \nabla_{\frac{\partial \gamma}{\partial s}} \frac{\partial \gamma}{\partial t} \right\rangle + \left\langle \frac{\partial \gamma}{\partial t}, \nabla_{\frac{\partial \gamma}{\partial s}} \nabla_{\frac{\partial \gamma}{\partial s}} \frac{\partial \gamma}{\partial t} \right\rangle \mathrm{d}t,$$

where

$$\left\langle \nabla_{\frac{\partial \gamma}{\partial s}} \frac{\partial \gamma}{\partial t}, \nabla_{\frac{\partial \gamma}{\partial s}} \frac{\partial \gamma}{\partial t} \right\rangle = \left\langle \nabla_{\frac{\partial \gamma}{\partial t}} \frac{\partial \gamma}{\partial s}, \nabla_{\frac{\partial \gamma}{\partial t}} \frac{\partial \gamma}{\partial s} \right\rangle = |\dot{V}(t)|^2$$

when $s = 0$, and

$$\nabla_{\frac{\partial \gamma}{\partial s}} \nabla_{\frac{\partial \gamma}{\partial s}} \frac{\partial \gamma}{\partial t} = \nabla_{\frac{\partial \gamma}{\partial s}} \nabla_{\frac{\partial \gamma}{\partial t}} \frac{\partial \gamma}{\partial s} = \nabla_{\frac{\partial \gamma}{\partial t}} \nabla_{\frac{\partial \gamma}{\partial t}} \frac{\partial \gamma}{\partial s} - R\left(\frac{\partial \gamma}{\partial s}, \frac{\partial \gamma}{\partial t} \right) \frac{\partial \gamma}{\partial s}.$$

Thus when $s = 0$, we have

$$E''(0) = \int_0^a |\dot{V}(t)|^2 - \langle R(\dot{\gamma}_0, V)\dot{\gamma}_0, V \rangle + \frac{\partial}{\partial t} \langle \dot{\gamma}_0, \nabla_V V \rangle - \langle \nabla_{\dot{\gamma}_0} \dot{\gamma}_0, \nabla_V V \rangle,$$

the last term is 0 since γ_0 is a geodesic, and we obtain the formula. $\square$

Second variation formula gives rise to a symmetric bilinear form on the vector space of vector fields along a geodesic γ, which we will call it index form, to deduce whether a geodesic is locally length minimizing.

Definition. Let $\gamma : [0, a] \to M$ be a geodesic. The *index form* of γ is a bilinear form on $\Gamma(TM, \gamma)$ defined by

$$I(X, Y) = \int_0^a \langle \dot{X}, \dot{Y} \rangle - \langle R(\dot{\gamma}, X)\dot{\gamma}, Y \rangle \, \mathrm{d}t.$$

Clearly index form is symmetric.

Lemma 3.23. *Let $U \in \Gamma_0(TM, \gamma) = \{Y \in \Gamma(TM, \gamma) : Y(0) = Y(a) = 0\}$, then U is a Jacobi field if and only if $I(U, Y) = 0$ for any $Y \in \Gamma_0(TM, \gamma)$.*

Proof. First we observe that by integrate by parts, we have

$$I(U, Y) = -\int_0^a \langle \ddot{U} + R(\dot{\gamma}, U)\dot{\gamma}, Y \rangle.$$

If U is a Jacobi field, then $\ddot{U} + R(\dot{\gamma}, U)\dot{\gamma} = 0$, which implies $I(U, Y) = 0$. Conversely, if $I(U, Y) = 0$ for any $Y \in \Gamma_0(TM, \gamma)$, choose a smooth function f with $f(0) = f(a) = 0$ and positive elsewhere, let $Y = f(t)(\ddot{U} + R(\dot{\gamma}, U)\dot{\gamma})$, then

$$0 = -f \int_0^a |\ddot{U} + R(\dot{\gamma}, U)\dot{\gamma}|^2 \, dt,$$

this implies $\ddot{U} + R(\dot{\gamma}, U)\dot{\gamma} = 0$, that is, U is a Jacobi field. □

We observe if a Jacobi field J does not vanish at the starting or ending point of the geodesic, the above integration by parts is changed into

$$I(Y, J) = \langle Y, \dot{J} \rangle|_0^a - \int_0^a \langle \ddot{J} + R(\dot{\gamma}, J)\dot{\gamma}, Y \rangle \, dt$$
$$= \langle Y, \dot{J} \rangle|_0^a.$$

The next proposition shows the positive definiteness of I is related to conjugate points.

Proposition 3.24. *Let $\gamma : [0, a] \to M$ be a geodesic, I be its index form.*

(1) *If γ contains no conjugate points of $\gamma(0)$, then I is positive definite on $\Gamma_0(TM, \gamma)$.*

(2) *If $\gamma(a)$ is the only conjugate point of $\gamma(0)$, then I is positive semidefinite but not positive definite.*

(3) *If there is a $t_0 < a$ such that $\gamma(t_0)$ is conjugate to $\gamma(0)$, then there is a $U \in \Gamma_0(TM, \gamma)$ such that $I(U, U) < 0$.*

Proof of Proposition 3.24 (1). Let $\gamma(0) = p$, $\tilde{\gamma}$ is the radial line in T_pM defined by $\tilde{\gamma}(t) = t\dot{\gamma}(0)$. Since γ contains no conjugate points of $\gamma(0)$, $\exp_p$ is nondegenerate on the whole $\tilde{\gamma}$. Hence there is a neighborhood U of $\tilde{\gamma}([0, a])$ such that $\exp_p : U \to M$ is an immersion. By Proposition 2.21 and the subsequent remark, γ is length-minimizing. Hence by Proposition 2.5, γ also minimizes energy. Let $\gamma(t, s) : [0, a] \times (-\varepsilon, \varepsilon) \to M$ be any proper variation, by taking ε small enough, we can assume the variation is contained in U. Then we have

$$E''(0) = \lim_{s \to 0} \frac{E(-s) + E(s) - 2E(0)}{s^2} \geq 0.$$

Since $E''(0)(V) = I(V,V)$ for any variation vector field, we have $I(V,V) \geq 0$ for any $V \in \Gamma_0(TM, \gamma)$. Now we must show $I(V,V) = 0$ implies $V = 0$. Let $I(V,V) = 0$, $X \in \Gamma_0(TM, \gamma)$ and $\delta > 0$, we have

$$0 \leq I(V + \delta X, V + \delta X) = I(V,V) + 2\delta I(V,X) + \delta^2 I(X,X),$$

this implies

$$2I(V,X) + \delta I(X,X) \geq 0.$$

Let $\delta \to 0$, we obtain $I(V,X) \geq 0$ for any $X \in \Gamma_0(TM, \gamma)$. Similarly, consider $I(V - \delta X, V - \delta X)$, we have $I(V,X) \leq 0$ for any $X \in \Gamma_0(TM, \gamma)$. This means $I(V,X) = 0$ for any $X \in \Gamma_0(TM, \gamma)$. Now by Lemma 3.23, V is a Jacobi field. But γ has no conjugate points of $\gamma(0)$, V must identically equal to 0. This proves I is positive definite. $\qquad \square$

Before proving the rest of Theorem 3.24, we need the following *Index Lemma*.

Proposition 3.25 (Index Lemma). *Assume $\gamma : [0,a] \to M$ is a geodesic without conjugate points. Let $U \in \Gamma(TM, \gamma)$ with $U(0) = 0$, J be a Jacobi field such that $J(0) = 0$, $J(a) = U(a)$. Then $I(J,J) \leq I(U,U)$, the equality holds if and only if $U = J$.*

Proof. Since $U - J \in \Gamma_0(TM, \gamma)$, by Theorem 3.24 (1), we have $I(U - J, U - J) \geq 0$ with equality holds iff $U = J$. Then we have

$$0 \leq I(U - J, U - J) = I(U,U) - 2I(U,J) + I(J,J).$$

But $I(U,J) = \langle U, \dot{J} \rangle|_0^a = \langle J, \dot{J} \rangle|_0^a = I(J,J)$, thus

$$I(U,U) \geq I(J,J). \qquad \square$$

Proof of the rest of Proposition 3.24. (2) For any $0 < t_0 < a$, let $I_{[0,t_0]}$ denote the index form of $\gamma|_{[0,t_0]}$, i.e.,

$$I_{[0,t_0]}(X,Y) = \int_0^{t_0} \langle \dot{X}, \dot{Y} \rangle + \langle R(\dot{\gamma}, X)\dot{\gamma}, Y \rangle \, \mathrm{d}t.$$

Then for $X \in \Gamma_0(TM, \gamma|_{[0,t_0]})$, by (1) we have $I_{[0,t_0]}(X,X) \geq 0$. We now construct $\tau_{t_0}(U)$ for any $U \in \Gamma_0(TM, \gamma)$. Let $\{e_i(t)\}$ be a parallel frame along γ, $U = f^i(t)e_i(t)$. Then we define

$$\tau_{t_0}(U)(t) = f^i\left(\frac{a}{t_0}t\right) e_i\left(\frac{a}{t_0}t\right) \in \Gamma_0(TM, \gamma|_{[0,t_0]}).$$

Thus we have

$$I_{[0,t_0]}(\tau_{t_0}(U), \tau_{t_0}(U)) \geq 0,$$

let $t_0 \to a$ we obtain the conclusion. Moreover, since $\gamma(a)$ is conjugate to $\gamma(0)$, there exists a nontrivial Jacobi field J with $J(0) = J(a) = 0$, and $I(J, J) = \langle J, \dot{J} \rangle|_0^a = 0$. This shows I is positive semidefinite but not positive definite.

(3) Since $\gamma(0)$ is conjugate to $\gamma(t_0)$, there exists a Jacobi field $\tilde{J}_1$ along $\gamma|_{[0,t_0]}$ such that $\tilde{J}_1(0) = \tilde{J}_1(t_0) = 0$. Choose a smooth $m_1 : [0, a] \to \mathbb{R}$ with supp $m_1 = [0, t_0]$, and set $J_1 = m_1 \tilde{J}_1$. Let

$$V(t) = \begin{cases} J_1(t), & 0 \le t \le t_0, \\ 0, & t_0 \le t \le a, \end{cases}$$

then $V \in \Gamma(TM, \gamma)$ is smooth. Let $\delta > 0$ so small that $\gamma_{[t_0-\delta,t_0+\delta]}$ is contained in a normal neighborhood of $\gamma(t_0)$. Then there exists a Jacobi field $\tilde{J}_2$ along $\gamma|_{[t_0-\delta,t_0+\delta]}$ with $\tilde{J}_2(t_0 - \delta) = J_1(t_0 - \delta)$, $\tilde{J}_2(t_0 + \delta) = 0$. Let $m_2 : [t_0 - \delta, t_0 + \delta]$ satisfy

$$\dot{J}_1 = \nabla_{\dot{\gamma}} m_2 \tilde{J}_2 = \frac{\mathrm{d}m_2}{\mathrm{d}t} \tilde{J}_2 + m_2 \dot{\tilde{J}}_2$$

and

$$R(\dot{\gamma}(t_0 - \delta), J_1(t_0 - \delta))\dot{\gamma}(t_0 - \delta) = m_2(t_0 - \delta)R(\dot{\gamma}(t_0 - \delta), \tilde{J}_2(t_0 - \delta))\dot{\gamma}(t_0 - \delta),$$

then by ODE theory, m_2 is the solution of a linear ODE of order 1, hence it exists and is smooth. Now set $J_2 = m_2 \tilde{J}_2$, and

$$U(t) = \begin{cases} J_1(t), & 0 \le t \le t_0 - \delta, \\ J_2(t), & t_0 - \delta \le t \le t_0 + \delta, \\ 0, & t_0 + \delta \le t \le a, \end{cases}$$

then up to a third mollification, $U(t) \in \Gamma(TM, \gamma)$. Therefore by Index Lemma, we have

$$\begin{aligned} I(U, U) &= I_{[0,t_0-\delta]}(J_1, J_1) + I_{[t_0-\delta,t_0+\delta]}(J_2, J_2) + 0 \\ &< I_{[0,t_0-\delta]}(V, V) + I_{[t_0-\delta,t_0+\delta]}(V, V) \\ &= I_{[0,t_0]}(V, V) + I_{[t_0,a]}(V, V) \\ &= \langle J_1, \dot{J}_1 \rangle|_0^{t_0} + 0 \\ &= 0. \qquad \qquad \square \end{aligned}$$

Translating Proposition 3.24 to the language of geometry, we have the following theorem.

Theorem 3.26 (Jacobi)**.** *Let $\gamma : [0, a] \to M$ be a geodesic, then*

(1) *If γ has no conjugate points of $\gamma(0)$, then γ is length-minimizing under any small proper variation.*

(2) *If there is a $t_0 \in (0, a)$ such that $\gamma(t_0)$ is conjugate to $\gamma(0)$, then there is a proper variation γ_s such that $L[\gamma_s] < L[\gamma]$ for any s.*

3.5 Bochner Formula

We introduce a useful calculation tool. We will later apply it to distance function.

Theorem 3.27 (Bochner formula). *Let M be a Riemannian manifold, $f \in C^\infty(M)$. Then we have*

$$\Delta \frac{1}{2} |\nabla f|^2 = |\nabla^2 f|^2 + \langle \nabla \Delta f, \nabla f \rangle + \langle \mathrm{Ric}(\nabla f), \nabla f \rangle.$$

We first need a lemma.

Lemma 3.28. *Under geodesic coordinate around p, the Ricci identity (Proposition 3.7) of 1-forms is equivalent to*

$$f_{;kij} - f_{;kji} = -f_{;l} R_{ijk}{}^{l},$$

where $f_{;i}$ means $\partial_i f$.

Proof. On the one hand, since all Christoffel symbols vanish at p, we have

$$\begin{aligned}
\nabla_{\partial_i} \nabla_{\partial_j} \, \mathrm{d}f &= \nabla_{\partial_i} \nabla_{\partial_j} f_{;k} \, \mathrm{d}x^k \\
&= \nabla_{\partial_i} (\partial_j f_{;k} \, \mathrm{d}x^k + f_{;k} \nabla_{\partial_j} \, \mathrm{d}x^k) \\
&= \nabla_{\partial_i} (\partial_j f_{;k} \, \mathrm{d}x^k - f_{;k} \Gamma^k_{jl} \, \mathrm{d}x^l) \\
&= \partial_i \, \partial_j f_{;k} \, \mathrm{d}x^k + \partial_j f_{;k} \nabla_{\partial_i} \, \mathrm{d}x^k \\
&= f_{;kji} - \partial_j f_{;k} \Gamma^k_{il} \, \mathrm{d}x^l \\
&= f_{;kji} \, \mathrm{d}x^k,
\end{aligned}$$

then we have

$$\begin{aligned}
(R(\partial_i, \partial_j) \, \mathrm{d}f)(\partial_k) &= (\nabla_{\partial_j} \nabla_{\partial_i} - \nabla_{\partial_i} \nabla_{\partial_j} + \nabla_{[\partial_i, \partial_j]}) \, \mathrm{d}f(\partial_k) \\
&= f_{;kij} - f_{;kji}.
\end{aligned}$$

On the other hand, we have

$$(R(\partial_i, \partial_j) \, \mathrm{d}f)(\partial_k) = -\, \mathrm{d}f(R(\partial_i, \partial_j) \, \partial_k)$$

$$= -\,\mathrm{d}f(R_{ijk}{}^{l}\,\partial_{l})$$
$$= -f_{;m}\,\mathrm{d}x^{m}(R_{ijk}{}^{l}\,\partial_{l})$$
$$= -f_{;l}R_{ijk}{}^{l},$$

hence the equality holds.   □

Proof of Bochner formula. Under geodesic normal coordinate we have the following calculation:

$$\Delta\frac{1}{2}|\nabla f|^{2} = \sum_{i,j}\left(\frac{1}{2}f_{;j}^{2}\right)_{;ii}$$
$$= \sum_{i,j}(f_{;j}f_{;ji})_{;i}$$
$$= \sum_{i,j}(f_{;ji}^{2} + f_{;j}f_{;jii})$$
$$= \sum_{i,j}(f_{;ij}^{2} + f_{;j}f_{;iji})$$
$$= \sum_{i,j}f_{;ij}^{2} + \sum_{i,j}f_{;j}(f_{;iij} - f_{;k}R_{jii}{}^{k})$$
$$= \sum_{i,j}f_{;ij}^{2} + \sum_{i,j}f_{;j}f_{;iij} + \sum_{i,j}f_{;j}f_{;k}R_{iji}{}^{k}$$
$$= \sum_{i,j}f_{;ij}^{2} + \sum_{i,j}f_{;j}f_{;iij} + \sum_{j}f_{;j}f_{;k}\mathrm{Ric}_{j}{}^{k}$$
$$= |\nabla^{2}f|^{2} + \langle\nabla f,\nabla\Delta f\rangle + \langle\mathrm{Ric}(\nabla f),\nabla f\rangle. \qquad\qquad □$$

Chapter 4

Submanifolds

4.1 Second Fundamental Form

We first set up our stage. Let $f : \Sigma^k \to (M^n, g)$ be an immersion, equip Σ with pullback metric f^*g (still denoted by g for simplicity). There is a decomposition $\nabla_X Y = (\nabla_X Y)^\top + (\nabla_X Y)^\perp$ for $X, Y \in \Gamma(T\Sigma)$, and simple observation shows

Proposition 4.1. *Let ∇^Σ be the Levi–Civita connection on (Σ, g), then ∇^Σ is given by*

$$\nabla^\Sigma_X Y = (\nabla_X Y)^\top$$

for $X, Y \in \Gamma(T\Sigma)$.

Now we need the concept of normal bundle.

Definition. Let Σ be a submanifold of M, then for any $p \in M$ we have the decomposition

$$T_p M = T_p \Sigma \oplus N_p \Sigma,$$

where $N_p \Sigma$ is the orthogonal complement of $T_p \Sigma$. Then define the *normal bundle* of Σ to be

$$N\Sigma = \bigsqcup_{p \in M} N_p \Sigma.$$

It is similar to tangent bundle to make $N\Sigma$ into a vector bundle.

Thus we can define the second fundamental form of a submanifold.

Definition. Let Σ be a submanifold of M, then we define the *second fundamental form* of Σ to be

$$\mathrm{II} : \Gamma(T\Sigma) \times \Gamma(T\Sigma) \to \Gamma(N\Sigma)$$
$$(X, Y) \mapsto (\nabla_X Y)^\perp.$$

Lemma 4.2. *The second fundamental form is symmetric.*

Proof. We have

$$\text{II}(X, Y) - \text{II}(Y, X) = (\nabla_X Y)^\perp - (\nabla_Y X)^\perp = ([X, Y])^\perp = 0. \qquad \square$$

Moreover, we can define the shape operator.

Definition. For a fixed $\xi \in \Gamma(N\Sigma)$, we define

$$S_\xi : \Gamma(T\Sigma) \to \Gamma(T\Sigma)$$
$$X \mapsto \nabla_X \xi - (\nabla_X \xi)^\perp.$$

We have the following Weingarten formula.

Proposition 4.3. *For $\xi \in \Gamma(N\Sigma)$ and $X, Y \in \Gamma(T\Sigma)$, we have*

$$\langle S_\xi(X), Y \rangle = \langle \text{II}(X, Y), -\xi \rangle.$$

In particular, shape operator is self-adjoint.

Proof. We have the calculation

$$
\begin{aligned}
\langle S_\xi(X), Y \rangle &= \langle \nabla_X \xi - (\nabla_X \xi)^\perp, Y \rangle \\
&= \langle \nabla_X \xi, Y \rangle - \langle (\nabla_X \xi)^\perp, Y \rangle \\
&= X \langle \xi, Y \rangle - \langle \xi, \nabla_X Y \rangle \\
&= \langle -\xi, \nabla_X Y \rangle \\
&= \langle -\xi, (\nabla_X Y)^\perp \rangle \\
&= \langle \text{II}(X, Y), -\xi \rangle. \qquad \square
\end{aligned}
$$

For a two-sided hypersurface $\Sigma^{n-1} \subset M^n$, choose a unit normal vector field N, we can introduce the *second fundamental form tensor h*, which is defined by $h(X, Y) := \langle \text{II}(X, Y), N \rangle$. In other words, $\text{II}(X, Y) = h(X, Y)N$, and the Weingarten formula is $h(X, Y) = -\langle S(X), Y \rangle$.

4.2   Gauss–Codazzi Equations

In this section, our goal is to prove the following Gauss–Codazzi equations on the relation of curvature between a hypersurface and ambient manifold. After this, we give a brief discussion on principal curvatures.

In the following contents, we assume $\Sigma^{n-1} \subset M^{n-1}$, Σ is two-sided. We will use $D_X Y$ to denote the Levi–Civita connection of M, and $\nabla_X Y$ for Σ. N will denote the unit normal vector field of Σ.

Theorem 4.4. *Under above settings, for $X, Y, Z, W \in \Gamma(T\Sigma)$ and $N \in \Gamma(N\Sigma)$, we have*

- *Gauss equation:*

$$R^\Sigma(X, Y, Z, W) = R^M(X, Y, Z, W) + h(X, Z)h(Y, W) - h(X, W)h(Y, Z);$$

- *Codazzi equation:*

$$R^M(X, Y, Z, N) = (\nabla_Y h)(X, Z) - (\nabla_X h)(Y, Z).$$

Proof. We first calculate $R^\Sigma(X, Y)Z$. We have

$$
\begin{aligned}
R^\Sigma(X, Y)Z =& \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z \\
=& D_Y \nabla_X Z - h(\nabla_X Z, Y)N - D_X \nabla_Y Z + h(\nabla_Y Z, X)N \\
& + D_{[X,Y]} Z - h([X, Y], Z)N \\
=& D_Y D_X Z - D_Y h(X, Z)N - h(\nabla_X Z, Y)N \\
& - D_X D_Y Z + D_X h(Y, Z)N + h(\nabla_Y Z, X)N \\
& + D_{[X,Y]} Z - h(\nabla_X Y - \nabla_Y X, Z)N \\
=& R^M(X, Y)Z - D_Y h(X, Z)N + h(\nabla_Y X, Z) + h(X, \nabla_Y Z)N \\
& + D_X h(Y, Z)N - h(\nabla_X Y, Z)N - h(Y, \nabla_X Z)N \\
=& R^M(X, Y)Z \\
& - Y h(X, Z)N - h(X, Z)D_Y N + h(\nabla_Y X, Z) + h(X, \nabla_Y Z)N \\
& + X h(Y, Z)N + h(Y, Z)D_X N - h(\nabla_X Y, Z)N - h(Y, \nabla_X Z)N.
\end{aligned}
$$

First we consider $\langle R^\Sigma(X, Y)Z, W \rangle$, by Weingarten formula we have

$$
\begin{aligned}
\langle R^\Sigma(X, Y)Z, W \rangle &= \langle R^M(X, Y)Z, W \rangle + \langle h(Y, Z)D_X N, W \rangle - \langle h(X, Z)D_Y N, W \rangle \\
&= \langle R^M(X, Y)Z, W \rangle + h(X, Z)h(Y, W) - h(X, W)h(Y, Z).
\end{aligned}
$$

Next we consider $\langle R^\Sigma(X, Y)Z, N \rangle$, notice that $\langle D_X N, N \rangle = \frac{1}{2} X \langle N, N \rangle = 0$, we have

$$
\begin{aligned}
0 =& \langle R^M(X, Y)Z, N \rangle + X h(Y, Z) - h(\nabla_X Y, Z) - h(Y, \nabla_X Z) \\
& - Y h(X, Z) + h(\nabla_Y X, Z) + h(X, \nabla_Y Z) \\
=& \langle R^M(X, Y)Z, N \rangle + (\nabla_X h)(Y, Z) - (\nabla_Y h)(X, Z),
\end{aligned}
$$

which is equivalent to Codazzi equation. $\square$

We now turn to principal curvatures.

Definition. Let $p \in \Sigma^{n-1} \subset M^n$, the *principal curvatures* at p are the eigenvalues of shape operator S. Usually we denote κ_i, $i = 1, \cdots, n-1$ for principal curvatures.

Principal curvatures are extrinsic geometry quantities, but on 2-dimensional surfaces, their product happens to be intrinsic. This is the *Theorema Egregium*, a glorious achievement of classical differential geometry. At the era of Gauss, who discovered the theorem, Theorema Egregium needs tedious calculation to be proved. But using modern language, Theorema Egregium is almost self-evident. We now state and prove this theorem.

Theorem 4.5 (Theorema Egregium). *On a surface $\Sigma^2 \subset \mathbb{R}^3$, we define* Gaussian curvature $\kappa = \kappa_1\kappa_2$, *the product of principal curvatures. The Gaussian curvature is equal to the sectional curvature on a point, hence Gaussian curvature is intrinsic, i.e. only depends on metric.*

Proof. Since $S^i_j = -g^{ik}h_{kj}$, we have

$$\kappa = \kappa_1\kappa_2 = \det(S^i_j) = \frac{\det(h_{ij})}{\det(g_{ij})}.$$

However, by Gauss equation, we have

$$\text{Sect}_p(T_p\Sigma) = \frac{\langle R(\partial_i, \partial_j)\,\partial_i, \partial_j \rangle}{g_{ii}g_{jj} - g_{ij}^2}$$

$$= \frac{1}{\det(g_{ij})}(0 + h_{ii}h_{jj} - h_{ij}^2)$$

$$= \kappa. \qquad \square$$

Finally we discuss totally umbilical hypersurfaces.

Definition. A point $p \in \Sigma^{n-1} \subset M^n$ is called *umbilical* if $h_p = cg_p$ for some $c \in \mathbb{R}$. Σ^{n-1} is called *totally umbilical* if $h = fg$ for some $f \in C^\infty(M)$.

Lemma 4.6. *A point is umbilical if and only if all the principal curvatures at this point are equal.*

Proof. This follows by $S^i_j = -g^{ik}h_{kj} = -fg^{ik}g_{kj} = -f\delta^i_j$. $\qquad \square$

Proposition 4.7. *If Σ is a totally umbilical hypersurface in $(\mathbb{R}^n, \delta)$, i.e. $h = f\delta$, then f must be constant.*

Proof. By Codazzi equation, at $p \in \Sigma$, for $u, w, v \in T_p\Sigma$, we have

$$0 = \langle R(u,v)w, n \rangle = (\nabla h)(u,w,v) - (\nabla h)(v,w,u). \tag{4.1}$$

Since $h = f\delta$, we have

$$\nabla h = \nabla(f\delta) = \delta \otimes df + f\nabla\delta = \delta \otimes df.$$

Hence equation 4.1 is equivalent to

$$0 = \langle u, w \rangle \, df_p(v) - \langle v, w \rangle \, df|_p(u).$$

Choose v arbitrarily, $w \neq 0$ satisfying $\langle v, w \rangle = 0$ and $u = w$, we have $df|_p(v) = 0$. This shows $df|_p = 0$. Since p is arbitrary, $df = 0$ and this implies $f \equiv$ const. $\qquad \square$

Remark 4.8. Some authors mistakenly think this proposition holds for general ambient manifold. We have a counterexample as follows. Let's consider the warped product $\mathbb{R} \times \mathbb{R}^2$ with metric $g = (dx^1)^2 + F^2(x^1)((dx^2)^2 + (dx^3)^2)$, where F is a nonconstant function of x^1. If $\{\partial_i\}$ is coordinate vector field, then $N = \partial_1$ is orthonormal to $\Sigma = \mathbb{R}^2$. We have

$$\Gamma^1_{13} = \frac{1}{2}g^{11}(\partial_3 g_{11} + \partial_1 g_{13} - \partial_1 g_{13}) = \frac{\partial_3 F}{F} = \Gamma^2_{23},$$

and

$$\Gamma^2_{13} = \frac{1}{2}g^{22}(\partial_3 g_{12} + \partial_1 g_{23} - \partial_2 g_{13}) = 0.$$

Therefore we have

$$h(\partial_1, \partial_1) = -\langle \nabla_{\partial_1} \partial_3, \partial_1 \rangle = -F^2 \Gamma^1_{13} = -F \partial_3 F = h(\partial_2, \partial_2),$$

and

$$h(\partial_1, \partial_2) = -\langle \nabla_{\partial_1} \partial_3, \partial_2 \rangle = -F^2 \Gamma^2_{13} = 0.$$

Hence we have

$$h = -\frac{\partial_3 F}{F} g =: fg,$$

f is not constant as F is not constant.

If we assume a priori the classification theorem of constant sectional curvature spaces, we can have the following corollary.

Corollary 4.9. *A complete totally umbilical hypersurface Σ in $(\mathbb{R}^n, \delta)$ is either a hyperplane or a sphere.*

Proof. Let $p \in \Sigma$, $u, v \in T_p\Sigma$, then by Gauss equation, we have

$$\text{Sect}_p(\text{Span}\{u, v\}) = 0 + \frac{h(u,u)h(v,v) - h^2(u,v)}{\delta(u,u)\delta(v,v) - \delta^2(u,v)}$$

$$= f^2 \frac{\delta(u,u)\delta(v,v) - \delta^2(u,v)}{\delta(u,u)\delta(v,v) - \delta^2(u,v)}$$

$$= f^2$$

$$\equiv \text{const} \geq 0.$$

When the constant is 0, Σ is a hyperplane; when the constant is positive, Σ is a sphere. □

4.3 Minimal Surfaces

We now give a brief introduction to minimal surfaces. We will use the word "area" to denote the volume of a submanifold.

Definition. Let $\Sigma^k \subset M^n$ be a submanifold (maybe with boundary), the *area functional* is defined as

$$A[\Sigma] = \int_\Sigma \mathrm{d}A_\Sigma,$$

where A_Σ is the volume element on Σ.

The guiding problem of this section is

Problem (Generalized Plateau problem)*.* Given a fixed boundary submanifold $B^{k-1} \subset M^n$ ($k < n$), find the area-minimizer in all submanifolds Σ^k with $\partial\Sigma = B$.

We use variation to deduce a necessary condition for a submanifold to be area-minimizing. To describe the first variation formula of area, we introduce the notion of mean curvature.

Definition. Let $\Sigma^k \subset M$, define the *mean curvature vector* at a point $p \in \Sigma$ to be

$$\mathbf{H}_p = \sum_{i=1}^k \mathrm{II}_p(e_i, e_i),$$

where $\{e_i\}$ is an orthonormal basis of $T_p\Sigma$. When $k = n - 1$, i.e. Σ is a hypersurface, we define *mean curvature* $H = \text{tr}\, h$, the trace of second fundamental form tensor.

Proposition 4.10 (First variation of area). *Let M^n be a Riemannian manifold, Σ^k be a manifold with boundary with $k < n$, $f : \Sigma \times (-\varepsilon, \varepsilon) \to M$ be a family of immersions, denote $\Sigma_t = f(\Sigma, t)$. Suppose $f(p, t) = p$ for $p \in \partial\Sigma$ and $t \in (-\varepsilon, \varepsilon)$, the variation vector field $X := \frac{\partial}{\partial t}\big|_{t=0} f \in \Gamma(TM, \Sigma_0)$. Then we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} A[\Sigma_t] = -\int_{\Sigma_0} \langle \mathbf{H}, X \rangle \, \mathrm{d}A_{\Sigma_0}.$$

Proof. Denote $g_t = f_t^* g$ and $A_t = A_{\Sigma_t}$. Let

$$\mathscr{J}(p, t) = \frac{\sqrt{\det g_t}}{\sqrt{\det g_0}},$$

then we have $A_t(p) = \mathscr{J}(p, t) A_0(p)$. Now we can calculate (omitting p)

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathscr{J}(t) &= \frac{1}{\sqrt{\det g_0}} \cdot \frac{1}{2\sqrt{\det g_0}} \cdot \det g_0 \cdot g_0^{ij} \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} g_{ij} \\
&= \frac{1}{2} g_0^{ij} \left(\left\langle \nabla_{\partial_t} \frac{\partial f}{\partial x^i}, \frac{\partial f}{\partial x^j} \right\rangle + \left\langle \frac{\partial f}{\partial x^i}, \nabla_{\partial_t} \frac{\partial f}{\partial x^j} \right\rangle \right) \\
&= g_0^{ij} \left\langle \nabla_{\partial_t} \frac{\partial f}{\partial x^i}, \frac{\partial f}{\partial x^j} \right\rangle \\
&= g_0^{ij} \left\langle \nabla_{\partial_i} \frac{\partial f}{\partial t}, \frac{\partial f}{\partial x^j} \right\rangle \\
&= g_0^{ij} \left\langle \nabla_{\partial_i} X^\perp, \frac{\partial f}{\partial x^j} \right\rangle + g_0^{ij} \left\langle \nabla_{\partial_i} X^\top, \frac{\partial f}{\partial x^j} \right\rangle \\
&= -\left\langle g_0^{ij} \,\mathrm{II}\left(\frac{\partial f}{\partial x^i}, \frac{\partial f}{\partial x^j} \right), X^\perp \right\rangle + \operatorname{div} X^\top \\
&= -\langle \mathbf{H}, X^\perp \rangle + \operatorname{div} X^\top.
\end{aligned}
$$

Notice that $-\langle \mathbf{H}, X^\perp \rangle = -\langle \mathbf{H}, X \rangle$, take integral we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} A[\Sigma_t] = -\int_{\Sigma_0} \langle \mathbf{H}, X \rangle \, \mathrm{d}A_{\Sigma_0} + \int_{\Sigma_0} \operatorname{div} X^\top \, \mathrm{d}A_{\Sigma_0}.$$

Since the variation fixes boundary, we have $X|_{\partial\Sigma_0} = 0$, then by divergence theorem, let N denote the outer unit normal vector field of $\partial\Sigma_0$, we have

$$\int_{\Sigma_0} \operatorname{div} X^\top \, \mathrm{d}A_{\Sigma_0} = \int_{\partial\Sigma_0} \langle X^\top, N \rangle \, \mathrm{d}A_{\partial\Sigma_0} = 0.$$

Hence we proved the first variation formula of area. $\qquad\square$

Thus we have a necessary condition for area-minimizing submanifold.

Corollary 4.11. *If Σ is a solution of generalized Plateau problem, then $\mathbf{H} = 0$.*

The corollary gives rise to the notion of minimal submanifolds.

Definition. Let Σ be a submanifold of M, if the mean curvature vector field $\mathbf{H} \equiv 0$, then we call Σ a *minimal submanifold*.

Remark 4.12. We do not demand a minimal submanifold is with boundary or actually minimzes area. But if the minimal submanifold is a graph, then it minimzes area.

Lemma 4.13. *Let $u : \Omega \subset \mathbb{R}^{n-1} \to \mathbb{R}$ be a smooth function, then its graph*

$$G[u] := \{(x^1, \cdots, x^{n-1}, u(x^1, \cdots, x^{n-1}))\}$$

is a minimal surface if and only if it satisfies the following minimal surface equation *(MSE):*

$$\mathrm{div}_{\mathbb{R}^{n-1}} \left(\frac{\mathrm{D}u}{\sqrt{1 + |\mathrm{D}u|^2}} \right) = 0,$$

where D is the Euclidean connection, and norm is Euclidean norm.

Proof. Let

$$r : \Omega \subset \mathbb{R}^{n-1} \to \mathbb{R}^n$$
$$(x^1, \cdots, x^{n-1}) \mapsto (x^1, \cdots, x^{n-1}, u(x^1, \cdots, x^{n-1})),$$

$r(\Omega) = G[u]$ has a unit normal vector field

$$N = \frac{(\mathrm{D}u, -1)}{\sqrt{1 + |\mathrm{D}u|^2}}.$$

Then we have

$$g_{ij} = \delta_{ij} + \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^j},$$

or $g = I + \mathrm{D}u(\mathrm{D}u)^t$. Hence we have formally

$$(I + \mathrm{D}u(\mathrm{D}u)^t)^{-1} = I - \mathrm{D}u(\mathrm{D}u)^t + (\mathrm{D}u(\mathrm{D}u)^t)^2 - (\mathrm{D}u(\mathrm{D}u)^t)^3 + \cdots$$
$$= I - \mathrm{D}u(\mathrm{D}u)^t(1 - |\mathrm{D}u|^2 + |\mathrm{D}u|^4 + \cdots)$$
$$= I - \frac{1}{1 + |\mathrm{D}u|^2}\mathrm{D}u(\mathrm{D}u)^t,$$

and one can check we indeed obtain the inverse matrix. Moreover, we have second fundamental form tensor

$$h_{ij} = \langle r_{;ij}, N \rangle = -\frac{1}{\sqrt{1 + |Du|^2}} \cdot \frac{\partial^2 u}{\partial x^i \, \partial x^j}.$$

Thus by taking trace, we have

$$
\begin{aligned}
H &= g^{ij} h_{ij} \\
&= -\sum_{i,j} \left(\delta_{ij} - \frac{1}{1 + |Du|^2} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^j} \right) \left(\frac{1}{\sqrt{1 + |Du|^2}} \cdot \frac{\partial^2 u}{\partial x^i \, \partial x^j} \right) \\
&= -\sum_{i,j} \frac{\partial_i(\partial_i u) \sqrt{1 + |Du|^2} - \partial_j u \, \partial_i u \, \partial_i(\partial_j u)(1 + |Du|^2)^{-1/2}}{1 + |Du|^2} \\
&= -\sum_{i,j} \frac{\partial}{\partial x^i} \left(\frac{\partial_i u}{\sqrt{1 + |Du|^2}} \right) \\
&= -\operatorname{div}_{\mathbb{R}^{n-1}} \left(\frac{Du}{\sqrt{1 + |Du|^2}} \right),
\end{aligned}
$$

where we used the fact

$$
\begin{aligned}
\frac{\partial}{\partial x^i} \sqrt{1 + |Du|^2} &= \frac{\partial_i(1 + \sum_j (\partial_j u)^2)}{2\sqrt{1 + |Du|^2}} \\
&= \sum_j \frac{\partial_i(\partial_j u) \, \partial_j u}{\sqrt{1 + |Du|^2}}.
\end{aligned}
$$

Hence we proved MSE. $\qquad\square$

Remark 4.14. When $n = 3$, MSE was found by Lagrange in 1762.

Proposition 4.15. *A minimal graph in $\mathbb{R}^n$ is area-minimizing.*

Proof. We need to show if Σ_u is a minimal graph, Σ is a surface such that $\partial \Sigma_u = \partial \Sigma$, then $A[\Sigma_u] \le A[\Sigma]$. Define a vector field $X \in \Gamma(T\mathbb{R}^n)$ such that

$$X = \frac{(-Du, 1)}{\sqrt{1 + |Du|^2}}.$$

Since X does not depend on x^n, by MSE we have

$$\operatorname{div}_{\mathbb{R}^n} X = \operatorname{div}_{\mathbb{R}^{n-1}} \left(\frac{-Du}{\sqrt{1 + |Du|^2}} \right) = 0.$$

Let Ω be the domain enclosed by Σ_u and Σ, N_{Σ_u} and N_Σ be the outward unit normal vector field of Σ_u and Σ respectively. Then by divergence theorem, we have

$$0 = \int_\Omega \operatorname{div}_{\mathbb{R}^n} X \, \mathrm{d} \operatorname{Vol}_\Omega$$
$$= \int_{\Sigma_u} \langle X, N_{\Sigma_u} \rangle \, \mathrm{d} A_{\Sigma_u} + \int_\Sigma \langle X, N_\Sigma \rangle \, \mathrm{d} A_\Sigma.$$

Since $N_{\Sigma_u} = \frac{(Du, -1)}{\sqrt{1+|Du|^2}}$, we have $\langle X, N_{\Sigma_u} \rangle = -1$, hence by Cauchy–Schwarz inequality,

$$A[\Sigma_u] = \int_\Sigma \langle X, N_\Sigma \rangle \, \mathrm{d} A_\Sigma \le \int_\Sigma |X| |N_\Sigma| \, \mathrm{d} A_\Sigma = A[\Sigma],$$

where $|X| = |N_\Sigma| = 1$.   □

Finally we have a characterization of immersed minimal surfaces in Euclidean space.

Proposition 4.16. *Let* $f : \Sigma^k \to \mathbb{R}^n$ *($k < n$) be an immersion, denote* $\Delta f = (\Delta f^1, \cdots, \Delta f^n)$, *then we have*
$$\Delta f = \mathbf{H}.$$

Proof. We use D to denote the Levi–Civita connection of $\mathbb{R}^n$ (i.e. directional derivative) and ∇ for Σ. First, we notice that this proposition is local, thus we can prove it on a sufficiently small coordinate neighborhood W on which f is an embedding. Moreover, notice that we have an identification

$$f_* Y = Y f = Y f^i \frac{\partial}{\partial x^i} = (Y f^1, \cdots, Y f^n) \in T\mathbb{R}^n,$$

thus we have $Y = Yf \in \Gamma(Tf(\Sigma))$. Now let $\{X_i\}$ be an orthonormal frame on W, we have

$$\Delta f = \sum_i X_i(X_i f) - (\nabla_{X_i} X_i) f$$
$$= \sum_i X_i(X_i) - \nabla_{X_i} X_i$$
$$= \sum_i D_{X_i} X_i - \nabla_{X_i} X_i$$
$$= \sum_i \mathrm{II}(X_i, X_i)$$
$$= \mathbf{H}.$$ □

Corollary 4.17. *An submanifold in Euclidean space is a minimal surface if and only if its every coordinate function is harmonic.*

Corollary 4.18. *There is no compact minimal submanifold without boundary in Euclidean space.*

Proof. Since the submanifold is minimal, all its coordinate functions are harmonic. But the submanifold is compact, the coordinate functions reach their maximum and minimum. Hence by Liouville's theorem, every coordinate functions is constant. This is impossible. □

Chapter 5

Curvature and Topology I

In this chapter, we will disucss some topological results derived by curvature restrictions. We will follow the order of Ricci curvature with positive lower bound, nonpositive sectional curvature and constant sectional curvature. The corresponding results are Bonnet–Myers Theorem, Cartan–Hadamard Theorem and classification theorem of space forms. More results on topology by curvature restrictions involves comparison theorem, which will be discussed later.

From this chapter, we will assume that all Riemannian manifolds are complete.

5.1   Bonnet–Myers Theorem

Bonnet–Myers Theorem asserts that a Riemannian manifold with positive lower bound for Ricci curvature is bounded, hence compact. Precisely, we have

Theorem 5.1 (Bonnet–Myers)**.** *Let (M^n, g) be a complete Riemannian manifold with* $\text{Ric} \geq (n-1)Kg > 0$ *(here we lowered the index of Ricci tensor), then* $\text{diam}(M) \geq \pi/\sqrt{K}$*. In particular, M is compact.*

Proof. We may scale the metric to make $K = 1$. Use contradiction argument, we may assume there exists $p, q \in M$ with $d(p, q) > \pi$. By completeness, there exists a unit speed geodesic $\gamma : [0, a] \to M$ such that $\gamma(0) = p$, $\gamma(a) = q$, so $d(p, q) = a > \pi$, γ is length-minimizing. Let $\{E_i(t)\}$ be a parallel orthonormal frame along γ with $E_n(t) = \dot{\gamma}(t)$, define

$$U_i(t) = \sin\left(\frac{\pi t}{a}\right).$$

Let I be the index form on γ, then we have

$$
\begin{aligned}
\sum_{i=1}^{n-1} I(U_i, U_i) &= \sum_{i=1}^{n} \int_0^a \langle \dot{U}_i, \dot{U}_i \rangle - \langle R(\dot{\gamma}, U_i)\dot{\gamma}, U_i \rangle \, \mathrm{d}t \\
&= \sum_{i=1}^{n-1} \int_0^a -\langle \ddot{U}_i, U_i \rangle - \langle R(U_i, \dot{\gamma})U_i, \dot{\gamma} \rangle \, \mathrm{d}t \\
&= \sum_{i=1}^{n-1} \int_0^a \frac{\pi^2}{a^2} \sin^2 \left(\frac{\pi t}{a} \right) - \sin^2 \left(\frac{\pi t}{a} \right) \langle R(E_i, \dot{\gamma})E_i, \dot{\gamma} \rangle \, \mathrm{d}t \\
&= \int_0^a \left((n-1)\frac{\pi^2}{a^2} - \langle \mathrm{Ric}(\gamma), \gamma \rangle \right) \sin^2 \left(\frac{\pi t}{a} \right) \\
&\leq \int_0^a (n-1) \left(\frac{\pi^2}{a^2} - 1 \right) \sin^2 \left(\frac{\pi t}{a} \right) \mathrm{d}t \\
&< 0,
\end{aligned}
$$

the second equality is integration by parts. But γ is a length-minimizing geodesic, I must be positive definite, hence

$$
\sum_{i=1}^{n-1} I(U_i, U_i) \geq 0,
$$

this is a contradiction. $\square$

Remark 5.2.   (1) If we know a priori Cheng's maximal diameter theorem, the above proof can be regarded as a comparison between Jacobi fields on M and sphere. Thus this is a prototype of our proof to Rauch's comparison theorem.

(2) The condition cannot be weakened to $K = 0$, in fact, Sect > 0 is not enough. The surface $z = x^2 + y^2$ in $\mathbb{R}^3$ is an counterexample.

Corollary 5.3. *Under above settings, the universal covering $\pi : \tilde{M} \to M$ is compact. Moreover, $\pi_1(M)$ is finite.*

Proof. We can lift the metric of M to make π a Riemannian covering, then $\pi^* g$ has also a Ricci lower bound, and we can apply Bonnet–Myers Theorem. For the next claim, let $p \in M$, then $\pi^{-1}(p)$ is a discrete closed set in $\tilde{M}$, hence must be finite since $\tilde{M}$ is compact. Then the covering map is of finite sheet, $\pi_1(\tilde{M})$ has finite index in $\pi_1(M)$. But $\pi_1(\tilde{M})$ is trivial, hence $\pi_1(M)$ is finite. $\square$

5.2 Cartan–Hadamard Theorem

Cartan–Hadamard Theorem gives the topological characterization of a Riemannian manifold with nonpositive sectional curvature.

Theorem 5.4 (Cartan–Hadamard). *Let M^n be a complete Riemannian manifold with Sect ≤ 0, then for any $p \in M$ the exponential map $\exp_p : T_pM \to M$ is a universal covering.*

Bibliography

[1] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[2] Manfredo Perdigão do Carmo. *Riemannian geometry. Translated from the Portuguese by Francis Flaherty*. Boston, MA etc.: Birkhäuser, 1992.

[3] Michel A. Kervaire. A manifold which does not admit any differentiable structure. *Comment. Math. Helv.*, 34:257–270, 1960.

[4] John M. Lee. *Introduction to topological manifolds*, volume 202 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2011.

[5] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

[6] Peter Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*. Springer, Cham, third edition, 2016.

[7] Joseph J. Rotman. *Advanced modern algebra. Part 1*, volume 165 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, third edition, 2015.

[8] Hung-Hsi Wu, Chunli Shen, and Yanlin Yu. *Introduction to Riemannian Geometry*. Higher Education Press, 2014.