# Lecture Notes on Differential Geometry

Chong Song

2020 年 5 月 26 日

# Abstract

This is a lecture note originated from the course "Differentiable Manifold" taught at Xiamen University from 2017 to 2020. The course is taught in 16 weeks with three 45-minutes classes each week. The audience are usually first year graduate students and senior undergraduates with a math major.

The main objective of this note is to provide a quick view to all the basics in Differentiable Manifolds, as well as an introduction to Riemannian Geometry. This note is not self-contained, since many proofs of theorems can be easily found in standard textbooks. More efforts are made in explaining the geometric ideas lying behind the concepts, and treating the contents as a natural generalization of the classical calculus in Euclidean spaces.

Any suggestions or comments are welcome.



Email: *songchong@xmu.edu.cn*

# Contents

# Chapter 0

# Introduction and Preliminaries

## 0.1 About the course

### 0.1.1 Euclidean space revisited

The familiar Euclidean space is a model space with rich structures. We have learned that the Euclidean space $\mathbb{E}^n$ is, among other things,

- A *linear space* or *vector space*. That means after fixing a point $O$ as the origin, each point $P$ can be identified with a vector $\vec{v} = \overrightarrow{OP}$ (ie. an object with both length and direction). Then we can define *scalar multiplication* and *addition* of vectors. In this way, we think of $\mathbb{E}^n$ as a space of vectors, such that for each vector $\vec{v}, \vec{w} \in \mathbb{E}^n$, the vector $\lambda\vec{v} + \mu\vec{w}$ also belongs to $\mathbb{E}^n$, for all $\lambda, \mu \in \mathbb{R}^1$. Then we can talk about linear independent *basis* and discuss about *linear transformations* of the space.

- A *metric space*. For every two points $P, Q \in \mathbb{E}^n$, we can define the *distance function* $d(P, Q)$. The distance function $d$ then satisfies

  1. $d(P, Q) \geq 0$ and the identity holds iff $P = Q$;
  2. $d(P, Q) = d(Q, P)$;
  3. the triangle inequality $d(P, Q) + d(Q, R) \geq d(P, R)$.

- A *topological space*. The topology defines the "*openness*" of arbitrary sets. It then follows the notion of close and compact sets, neighborhoods, etc. Moreover, we can define continuous maps and homotopy, homology on topological spaces. Note that the topology can be induced by the metric structure on $\mathbb{E}^n$.

- A *differentiable space*. This means, by introducing a *coordinate system*, that we can perform differentiation on functions defined on $\mathbb{E}^n$. Then we have the notion of tangent vectors and we can push through a long way to develop a full powerful kit of calculus.

- A *geometric space*. By introducing *inner product* on $\mathbb{E}^n$, we can define length of vectors and angles between them. In particular, we have the notion of orthogonal vectors. Then we can proceed to define the length of a smooth curve, the area or volume of higher dimensional regions by means of integrations. The geometric structure further leads to a metric structure, hence a topology on $\mathbb{E}^n$.

### 0.1.2 Goal of this course

In real word applications and also motivated by physics backgrounds, we need to study spaces different from the standard Euclidean space. We are usually led to study spaces with only one or two structures listed above, often in a more abstract way. Just think of the two-dimensional sphere. It is not a linear space and we do not have a globally defined coordinate system. However, locally around any point on the sphere, it looks exactly the same as the two-dimensional plane. Imagine about the earth as a total space and the local area where we live as a local region. It turns out that we can still do analysis and study the geometry of this space.

In the course of differentiable manifolds, we will learn how to do calculus on such spaces. Intuitively, manifolds are glued together by small pieces of the Euclidean space, but usually in a very non-trivial way. Thus the global picture of the manifold can be very complicated and totally different from the familiar Euclidean space. Furthermore, we want to study the geometric structures on manifolds by assigning a general notion of inner products. This leads to the fascinating subject of Riemannian Geometry.

### 0.1.3 References

In graduate courses, you should get used to self-motivated study and learn from different resources. Try not to stick to one textbook when you get stuck! In the reference below are some textbooks we will often refer to.

### 0.1.4 Online study and exam

We might need to study online, at least for several weeks at the beginning of this semester. I will upload lecture notes and short videos that explain the main concepts to the website: *http://course.xmu.edu.cn/*. You can also download digital version of the textbooks in the reference. Read the lecture notes carefully before each class, and refer to textbooks for more details. Online classes will be devoted to proofs of theorems, discussions and Q&A.

Try to do all the exercises in this note as homework by yourself. The examples also serves as standard exercises. You can find more exercises in reference [2] and [4]. You are required to hand in your homework every Friday during the period of online teaching, by email and in pdf format. My email address is: *songchong@xmu.edu.cn*. It is preferable to prepare the pdf file in LaTex.

The final exam will be in the form of a two-hour paper at the end of this semester, covering all the topics in this lecture note. Your final score in this course will depend on your performance in class (30%), of your homework (20%) and the final exam (50%).

Ask me online if you still have any questions about the course.

## 0.2 Preliminaries

Here we list some basic concepts and theorems that we need. For self-contained proofs of the theorems, we refer to the appendix of [1] and Chapter 0 of [2]. You can also refer to any standard textbook for more details.

### 0.2.1 Basic topology

Let $X$ be an nonempty set.

**Definition 0.2.1** *A topology $\tau$ of $X$ is a family of subsets in $X$, which is called* open sets*, which satisfies the following properties:*

1. *the total space $X$ and empty set $\emptyset$ belongs to $\tau$;*
2. *the intersection of finitely many open sets is an open set;*
3. *any union of open sets is still open.*

*We call the pair $(X, \tau)$ a* topological space*.*

**Definition 0.2.2** *Let $\mathscr{B}$ be a family of subsets in $X$. If $\mathscr{B}$ satisfies*

1. *for any point $x \in X$, there exists $B \in \mathscr{B}$ such that $x \in B$;*
2. *for any two subsets $B_1, B_2 \in \mathscr{B}$, if there is a point $x \in B_1 \cap B_2$, then there exists $B_3 \in \mathscr{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.*

*Then $\mathscr{B}$ is called a* topological basis *of $X$.*

**Definition 0.2.3** *Let $(X, \tau)$ be a topological space and $x \in S$ is a point. Suppose $U$ is a set that contains $x$. If there is an open set $A \in \tau$ such that $x \in A \subset U$, then $U$ is called a* neighborhood *of $x$. If $U$ is open, then $U$ is called an open neighborhood of $x$.*

Starting from a topology define above, one can proceed to define closed sets, continuous functions similar to those in Euclidean spaces.

Finally, we also need the following concepts. Let $X$ be a topological space.

**Definition 0.2.4** *If $X$ has a countable basis of open sets, then we say $X$ is second countable.*

**Definition 0.2.5** *If for any two distinct points $x, y \in X$, there exists two open sets $U, V$, such that $x \in U, y \in V$ and $U \cap V = \emptyset$, then $X$ is called a Hausdorff space.*

### 0.2.2 Implicit and inverse function theorem

**Theorem 0.2.6 (Inverse Function Theorem)** *Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}^n$ be a $C^k$ map, $k \geq 1$. If the Jacobian matrix $Df(x_0)$ is invertible for $x_0 \in \Omega$, then there exists an open neighborhood $U$ os $x_0$ such that $f$ is a $C^k$-diffeomorphism on $U$.*

**Theorem 0.2.7 (Implicit Function Theorem)** *Suppose $U \subset \mathbb{R}^m, V \subset \mathbb{R}^n$ are two open sets and $F : U \times V \to \mathbb{R}^n$ is a $C^1$ map. If $F(x_0, y_0) = 0$ for some $(x_0, y_0) \in U \times V$ and the Jacobian determinant*

$$\det D_y F = \frac{\partial F}{\partial y}(x_0, y_0) = \frac{\partial(F^1, \cdots, F^n)}{\partial(y^1, \cdots, y^n)}(x_0, y_0) \neq 0,$$

*then there exists a neighborhood $U'$ near $x_0$, $V'$ near $y_0$ and an implicit function $f : U' \to V'$ such that the graph*

$$\{(x, f(x)) | x \in U'\} = \{(x, y) \in U' \times V' | F(x, y) = 0\}.$$

One can first prove the Inverse Function Theorem by the following contraction mapping principle, and then prove the Implicit Function Theorem.

**Theorem 0.2.8 (Contraction mapping principle)** *Let $M$ be a complete metric space and suppose $T : M \to M$ is a map such that*

$$d(Tx, Ty) \leq \theta d(x, y),$$

*where $\theta < 1$. Then $T$ has a unique fixed point.*

**Exercise 0.2.9** *Show that the Inverse Function Theorem and the Implicit Function Theorem are equivalent.*

As an application of the IFT, we have the following Rank Theorem. Recall that the rank of a function $F : U \to \mathbb{R}^n$ at $x \in U$ is defined as the rank of the Jacobian matrix $DF(x)$. We say that $F$ has rank $k$ if the rank of $F$ equals $k$ at every point $x \in U$.

**Theorem 0.2.10 (Rank Theorem)** *Let $U \subset \mathbb{R}^m, V \subset \mathbb{R}^n$ be open subsets and $F : U \to V$ be a smooth function with rank $r$. Then for any $a \in U, b = F(a) \in V$, there exists open* *neighborhoods $U' \subset U, V' \subset V$ and smooth diffeomorphisms $u : U' \to \mathbb{R}^m$, $v : V' \to \mathbb{R}^n$, such that the composed map $v \circ F \circ u^{-1}$ has the form*

$$v \circ F \circ u^{-1}(x^1, \cdots, x^m) = (x^1, \cdots, x^r, 0, \cdots, 0).$$

*Proof.* First we construct the map $u$. Since $F = (f^1, \cdots, f^n)$ has rank $r$, by rearranging indices if necessary, we may assume that the $r \times r$ matrix

$$\left(\frac{\partial f^i}{\partial x^j}\right)_{1 \le i,j \le r}$$

is non-degenerate in a neighborhood of $a$. Define the map $u$ by

$$u(x^1, \cdots, x^m) = (f^1(x), \cdots, f^r(x), x^{r+1}, \cdots, x^m).$$

Obviously, $Du$ is non-degenerate, hence by the IFT, there is a small neighborhood $U' \subset U$ on which $u$ has a smooth inverse $u^{-1}$. Then the composed map $F \circ u^{-1}$ has the form

$$F \circ u^{-1}(x^1, \cdots, x^m) = (x^1, \cdots, x^r, g^1(x), \cdots, g^{n-r}(x)).$$

Note that $F \circ u^{-1}$ also has rank $r$. It follows that the matrix

$$\left(\frac{\partial g^\alpha}{\partial x^\beta}\right)_{1 \le \alpha,\beta \le n-r}$$

vanishes. Therefore, the functions $g^\alpha, 1 \le \alpha \le n-r$ only depend on $x^1, \cdots, x^r$. (One may assume the domain is a convex ball.)

Next we construct the function $v$ by

$$v(y^1, \cdots, y^n) = \big(y^1, \cdots, y^r, y^{r+1} - g^1(y^1, \cdots, y^r), \cdots, y^n - g^{n-r}(y^1, \cdots, y^r)\big).$$

Then $Dv$ has the form

$$\begin{pmatrix} I_r & 0 \\ * & I_{n-r} \end{pmatrix}$$

and hence is a smooth diffeomorphism on a neighborhood $V' \subset V$.

It is easy to check that $u, v$ satisfies the requirements and the theorem is proved. $\square$

### 0.2.3 Sard's theorem

For a differentiable map $f : \mathbb{R}^m \to \mathbb{R}^n$, $\operatorname{rank} f(x)$ is the rank of the Jacobian matrix $Df$ at $x$.

**Theorem 0.2.11 (Sard's Theorem)** *Suppose $U \subset \mathbb{R}^m$ is an open set, $F : U \to \mathbb{R}^n$ is a smooth map. Let $A = \{x \in U | \operatorname{rank} F(x) < n\}$, then the (Lebesgue) measure of $F(A) \subset \mathbb{R}^n$ is zero.*

For a proof of the Sard's theorem and more detailed discussions, we refer to Chapter 1.30 in [2]. Intuitively, the measure of $A$ is zero because it is a low dimensional set. As an easy corollary, if $m < n$, then the image of a smooth map has measure zero, since $\operatorname{rank} F \leq \min\{m, n\}$.

### 0.2.4 Basic ODE theory

**Theorem 0.2.12** *Suppose $U \subset \mathbb{R}^n$ is an open set and $F : I \times U \to \mathbb{R}^n$ is Lipschitz in $x \in U$. Then for the ODE*

$$\frac{d}{dt}x(t) = F(t, x(t)), \quad f(0) = x_0 \in U.$$

1. *there is a unique continuous solution $x(t) : [0, T) \to \mathbb{R}^n$ for some $T > 0$;*

2. *the solution $x(t)$ is smooth if $F$ is smooth;*

3. *$x(t)$ and $T$ depends continuously on the initial value $x_0$.*

This theorem can also be proved by using the contraction mapping principle.

# Part I

# Differentiable Manifolds

# Chapter 1

# Manifolds

## 1.1 Definition of differentiable manifolds

Intuitively, a manifold is glued together by small pieces (open sets) of Euclidean spaces. We say a manifold is differentiable if the pieces are glued in a smooth way. An easy example you may keep in mind is the standard 2-sphere, which can be constructed by gluing two caps (open disks) together. Also the standard Euclidean space serves as a trivial example (only one piece!).

### 1.1.1 Differentiable structure

There are usually two ways to introduce the concept of differentiable manifold. A popular way is to start from topological manifolds an then add the differentiable structure, which appears in most textbooks (cf. [2]). But here we follow Hitchin [1] and construct the differentiable manifolds directly (which naturally induce the topology).

**Definition 1.1.1** *A* coordinate chart *on a space $X$ is a subset $U \subset X$ together with a bijection*

$$\phi : U \to \phi(U) \subset \mathbb{R}^n$$

*onto an open set $\phi(U)$ in $\mathbb{R}^n$. The coordinates of a point $p \in U$ in this chart are just the coordinates of $\phi(p) = (x^1(p), \cdots, x^n(p)) \in \mathbb{R}^n$.*

*A coordinate chart is also called a local chart and denoted by $(U, \phi)$, or $(U, \phi; x^i)$ or simply $(U; x^i)$.*

Thus a coordinate chart associates to each point of a local region (a small piece of the space) a tuple of numbers, i.e. coordinates, such that we can express the point and eventually do calculus. The Euclidean space $\mathbb{E}^n$ can be covered by one chart once we fix the axes. However, many spaces can not be covered by a single chart. Thus we need to know how different charts are glued together, or in other words, how to transform from one chart to another.

**Definition 1.1.2** *An* atlas *on $X$ is a collection of coordinate charts $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ such that*

1. *$X$ is covered by $\{U_\alpha\}_{\alpha\in I}$*

2. *for each $\alpha, \beta \in I$, $\phi_\alpha(U_\alpha \cap U_\beta)$ is open in $\mathbb{R}^n$*

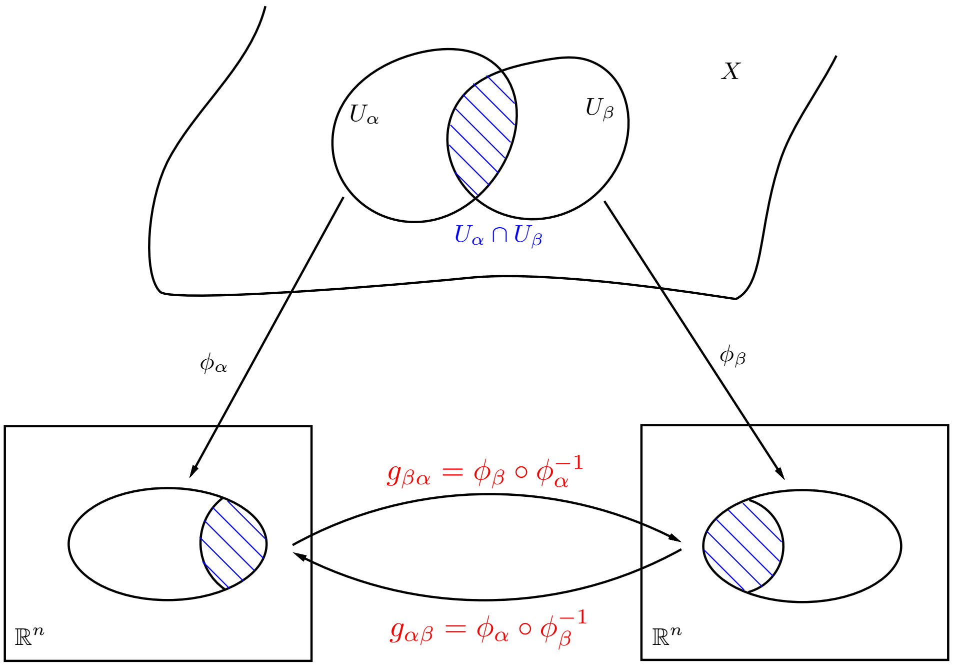


Figure 1.1: Atlas: gluing charts

3. *for each* $\alpha, \beta \in I$*, the* transition map

$$g_{\beta\alpha} := \phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$$

*is a smooth diffeomorphism.*

The transition maps $\{g_{\alpha\beta}\}$ play a key role in the construction of a manifold, and tells us how to translate the coordinates in one chart to another, see Figure 1.1 for an illustration.

We say a function $f$ is smooth if it has derivatives of all orders, and $f$ is called a smooth diffeomorphism if it has a smooth inverse. It is perfectly possible to define the atlas and thus develop the theory of manifolds with less differentiability, by only requiring the transition map belonging to $C^k$ for some integer $k$. But in our course, we will restrict ourselves to the category of smooth maps for simplicity.

There could be different atlases on a space $X$. For example, we can choose different axes on $\mathbb{E}^n$. But we would like to think of $X$ as an object independent of the choice of atlas, just like the Euclidean space does not depend on the choice of axes.

**Definition 1.1.3** *Two atlases* $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}, \{(V_\beta, \psi_\beta)\}_{\beta\in J}$ *are* compatible *or equivalent if their union is an atlas. The equivalent class of an atlas is called a* differentiable structure *on* $X$*.*

An alternative way to define a differentiable structure is to consider the maximal atlas which contains all possible compatible charts. Now we are in place to define differentiable manifolds, which we will simply call manifolds.

**Definition 1.1.4** *A* differentiable manifold *is a space* $X$ *with a differentiable structure.*

One can easily see that the notion of dimension makes sense on a manifold, at least for connected ones. The definition of a manifold takes into account the existence of many more atlases. But to prove something is a manifold, all you need is to find one atlas.

Note that there is a natural topology on a manifolds which is induced by the differentiable structure. One only need to define a subset $V \subset X$ is open if, for each $\alpha \in I$, $\phi_\alpha(V \cap U_\alpha)$ is open in $\mathbb{R}^n$. An immediate result is that the set $U_\alpha$ in each chart is open. More over, the coordinate map $\phi_\alpha : U_\alpha \to \phi_\alpha(U_\alpha)$ is a homeomorphism in the induced topology.

**Exercise 1.1.5** *Check that the differentiable structure indeed gives a topology.*

To proceed, we will always assume that *the manifold topology is Hausdorff and second countable* in this course. The assumption is very common for topological spaces, which implies that all manifolds are *paracompact*. That is, every open cover has an open refinement that is locally finite.

### 1.1.2 Basic examples

Here are some examples of manifolds. For a discussion on the topology in quotient spaces, we refer to Section 1.2 in [3].

1. Euclidean space.

2. The space of straight lines in the two dimensional plane.

3. Consider $\mathbb{R}^1$ as an additive group and the subgroup of integers $\mathbb{Z}$. The 1-dimensional torus is the quotient space $\mathbb{T}^1 = \mathbb{R}^1/\mathbb{Z}$.

4. The $n$-dimensional sphere

$$S^n = \{x \in \mathbb{R}^{n+1} || x| = 1\}.$$

5. The real projective space

$$\mathbb{R}P^n = \{n\text{-dimensional subspaces of } \mathbb{R}^{n+1}\}.$$

We can also construct new manifolds from given manifolds.

- Open subsets of a given manifold are called *open manifolds*, e.g., the general linear group

$$GL(n) = \{A \in M_{n\times n} | \det A \neq 0\}.$$

- The *product manifold* of two manifolds $M \times N$, e.g., the torus $S^1 \times S^1$.

**Exercise 1.1.6** *Construct an atlas on the 2-dimensional torus*

$$\mathbb{T}^2 = \mathbb{R}^2/(\mathbb{Z} \times \mathbb{Z}).$$

### 1.1.3 Partition of unity

The atlas and the differentiable structure of a manifold defines the way in which small pieces (local charts) are glued together. But sometimes, we need to go the opposite way, namely, we need to decompose a globally defined object on a manifold to small pieces in a specific way. The partition of unity offers such an useful tool.

**Definition 1.1.7** *A* partition of unity *is a family of smooth functions* $\{\phi_\alpha\}_{\alpha\in I}$ *such that*

1. $0 \leq \phi_\alpha \leq 1$;

2. $\{\operatorname{supp} \phi_\alpha\}_{\alpha \in I}$ *is locally finite;*

3. $\sum_{\alpha\in I} \phi_\alpha = 1$.

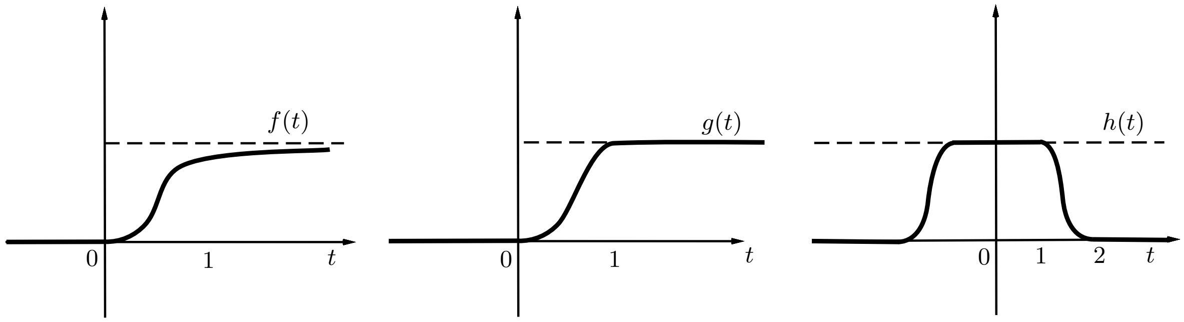


Figure 1.2: Bump function

The key property of a partition of unity lies in item (2). Here *locally finite* means each $x$ is covered by at most finitely many supports $\operatorname{supp}\phi_\alpha$'s. This ensures that the summation in item (3) make sense. To construct a partition of unity, we start from the so-called *bump functions* or *cut-off functions*.

We construct a bump function step by step (see Figure 1.2). First note that the following function is smooth (but not analytic)

$$f(t) = \begin{cases} e^{-1/t}, & t > 0 \\ 0, & t \le 0 \end{cases}.$$

Now let

$$g(t) = \frac{f(t)}{f(t) + f(1-t)}$$

so that $g$ is identically 1 when $t > 1$ and vanishes if $t \le 0$. Next write

$$h(t) = g(t+2)g(2-t).$$

The function $h(t)$ vanishes if $|t| > 2$ and equals 1 where $|t| < 1$. Finally we make an $n$-dimensional version

$$k(x_1, \cdots, x_n) = h(x_1)h(x_2)\cdots h(x_n).$$

For any $r > 0$, the function $k(x/r)$ is identically 1 in a ball of radius $r$ and 0 outside of a ball of radius $2r$. This gives a baby-version of bump functions. By a simple covering argument, we can construct the following cut-off function.

**Lemma 1.1.8** *Let $U, V \subset \mathbb{R}^n$ be two open sets where $\bar{U} \subset V$, then there exists a smooth bump function $\phi$ such that*

$$\phi(x) = \begin{cases} 1, & x \in \bar{U}; \\ 0, & x \in \mathbb{R}^n \setminus V. \end{cases}$$

*Proof.* For simplicity, we assume $\bar{U}$ is compact. For each $p \in \bar{U}$ we can find some $r_x > 0$ such that

$$x \in B_{r_x}(x) \subset B_{2r_x}(x) \subset V.$$

Then $\{B_{r_x}(x)\}_{x \in \bar{U}}$ forms an open cover of $\bar{U}$. Since $\bar{U}$ is compact, we can find a finite sub-cover $\{B_{r_i}(x_i)\}_{1 \le i \le K}$.

From previous discussion, there is a bump function $g_i$ such that $g_i = 1$ on $B_{r_i}(x_i)$ and vanishes outside $B_{2r_i}(x)$. Now we set

$$\phi(x) = 1 - \prod_{i=1}^{K} (1 - g_i(x)).$$

Then one checks this function satisfies the lemma. □

Now we are in place to construct the partition of unity. Recall that by assumption, our manifolds are Hausdorff and second countable.

**Theorem 1.1.9** *Given any open covering $\{V_\alpha\}_{\alpha \in I}$ of a manifold $M$, there exists a partition of unity $\{\phi_i\}$ subordinate to $\{V_\alpha\}_{\alpha \in I}$, i.e. for all $i$, there exists some $\alpha(i) \in I$ such that* $\operatorname{supp} \phi_i \subset V_{\alpha(i)}$.

*Proof.* *Step 1.* Exhaustion of $M$.

Since $M$ is second countable, there is a countable basis of open sets $\{O_i\}$ such that $M = \cup_i O_i$. We may assume that $\bar{O}_i$ is compact, since we can shrink $O_i$ if necessary such that each $O_i$ is homeomorphic to an open ball in the Euclidean space.

Now put $G_1 = O_1$. Then since the compact set $\bar{G}_1$ is covered by $\cup_i O_i$, there are finitely many $O_i$'s, such that

$$\bar{G}_1 \subset \cup_{i=1}^{k_1} O_i =: G_2.$$

Next we take the closure of $G_2$ and repeat the process. By induction, we will eventually get an exhaustion sequence $\{G_i\}$ such that

$$M = \cup_{i=1}^{\infty} G_i,\ \bar{G}_i \subset G_{i+1}.$$

*Step 2.* Locally finite refinement.

For each $i > 1$, consider the set

$$K_i := \bar{G}_{i+1} \setminus G_i \subset E_i := G_{i+2} \setminus \bar{G}_{i-1}.$$

Since the compact set $K_i$ is covered by $\{V_\alpha \cap E_i\}$, we can extract a finitely cover of open sets which we denote by $\{U_{il}\}_{1 \le l \le L_i}$, such that $K_i \subset \cup_{l=1}^{L_i} U_{il}$.

Clearly $\{U_{il}\}_{1 \le i \le +\infty, 1 \le l \le L_i}$ is a locally finite cover of $M$, since for each $i$, $U_{il} \cap U_{js} = \emptyset$ for sufficiently large $j$.

*Step 3.* Bump functions.

Now for any $x \in M$, there is an integer $i$ such that $x \in K_i$, thus we can find some $U_{il}$ which covers $x$. Then by Lemma 1.1.8, we can construct a bump function $\psi_x$ such that $\operatorname{supp} \psi_x \subset U_{il}$ and $\psi_x|_{W_x} = 1$ for some open neighborhood $W_x$.

Because the compact set $K_i$ is covered by $\{W_x\}_{x \in K_i}$, again there is a finite cover $\{W_{is}\}_{1 \le s \le S_i}$ with corresponding bump functions $\{\psi_{is}\}$. Finally, we let

$$\phi_{is} = \frac{\psi_{is}}{\sum_{i,s} \psi_{is}}.$$

This gives the desired partition of unity after re-indexing. □

**Exercise 1.1.10** *For a compact manifold, the construction of partition of unity can be considerably simplified. For an open covering of a compact manifold $M$, try to prove the existence of a partition of unity by yourself.*

**Exercise 1.1.11** *Suppose $U$ is an open subset of a manifold $M$ and $p \in U$ is a point. Show that for any smooth function $f : U \to \mathbb{R}^1$, there is a smooth function $\tilde{f} : M \to \mathbb{R}^1$ such that $\tilde{f}$ coincides with $f$ on a neighborhood of $p$.*

# 1.2 Maps and Submanifolds

In this section we define smooth maps between manifolds and then introduce the concept of submanifolds. Finally we will see that any manifold can be realized as a submanifold of a large dimensional Euclidean space.

## 1.2.1 Smooth maps on manifolds

Given two manifolds $M$, $N$ and a point $p \in M$, a map $F : M \to N$ is called a *smooth map* at $p$ if there exists a local chart $(U, \phi; x)$ at $p$ and $(V, \psi; y)$ at $F(p) \in N$ such that the map

$$\psi \circ F \circ \phi^{-1} : \phi(U) \to \psi(V)$$

is smooth. If $F$ is smooth at any point in $M$, then we call it a smooth map on $M$. For simplicity, we often write $y = F(x)$ in local coordinates instead of $y = \psi \circ F \circ \phi^{-1}(x)$ if there is no need to emphasis the coordinate charts.

The *rank* of $F$ at $p$ is defined by rank of the Jacobian matrix $D(\psi \circ F \circ \phi^{-1})(p)$ and denoted by $\operatorname{rank} F_p$. It is easy to verify that the $\operatorname{rank} F_p$ is independent of choice of local charts and hence is well-defined. Obviously, if $\dim M = m$ and $\dim N = n$, then $\operatorname{rank} F_p \leq \min\{m, n\}$; if $\operatorname{rank} F_p = \min\{m, n\}$, then we say $F$ is full rank at $p$.

**Example 1.2.1** *A smooth curve is a smooth map $\gamma : I \to M$, where $I \subset \mathbb{R}^1$ is an interval. A curve is called closed if it extends to a smooth map $\gamma : S^1 \to M$.*

**Example 1.2.2** *A smooth function on $M$ is a smooth map $f : M \to \mathbb{R}^1$. Construct a smooth function by using the bump function.*

A smooth map $F : M \to N$ is called a *diffeomorphism* if it has a smooth inverse $F^{-1} : N \to M$. The role of diffeomorphism is just like that of homeomorphisms in topology. If such a diffeomorphism exists, then we say the two manifolds $M$ and $N$ are *diffeomorphic*, that it, they are equivalent in the category of differentiable manifolds.

**Example 1.2.3** *Let $M$ be the standard real line $\mathbb{R}^1$ with identity map $\phi = id$ as the coordinate chart. Let $N$ be the real line $\mathbb{R}^1$ endowed with the coordinate chart $\psi(x) = x^3$. Since* *the function $\phi \circ \psi^{-1}(x) = x^{1/3}$ is not differentiable at $x = 0$, the differentiable structures of $M$ and $N$ are not equivalent by Definition 1.1.3. But there is a diffeomorphism*

$$F : M \to N, x \to x^{1/3}.$$

*such that $\psi \circ F \circ \phi = id$. Thus the two manifolds $M$ and $N$ are in fact diffeomorphic.*

A natural question is: on a given manifold, how many differentiable structures exist that are not diffeomorphic? In 1956, J. Milnor first discovered an exotic differentiable structure on $S^7$ that is different from the standard one. It is known that the differentiable structure on $R^n, n \neq 4$ is unique in the sense of diffeomorphisms. However, the pioneering work of Freedman and Donaldson showed that there are infinitely many different differentiable structures on $R^4$. Many deep and sophisticated theories in geometry and analysis are developed to answer these questions in topology.

**Exercise 1.2.4** *Show that the 1-dimensional torus $\mathbb{T}^1 = \mathbb{R}^1/\mathbb{Z}$, real projective space $\mathbb{R}P^1$ and the 1-sphere $S^1$ are diffeomorphic.*

### 1.2.2 Subsets as manifolds

The following theorem allows us to find more manifolds by an implicit method, i.e. without constructing explicit atlas.

**Theorem 1.2.5** *Let $\Omega \subset \mathbb{R}^{m+n}$ be an open set and $F : \Omega \to \mathbb{R}^n$ be a smooth function. Taking $c \in \mathbb{R}^n$, if for each $a \in F^{-1}(c)$, $F$ is full rank, i.e. $\operatorname{rank} F_a = n$. Then $F^{-1}(c)$ has the structure of an $m$-dimensional manifold which is Hausdorff and second countable. Moreover, for each $a \in F^{-1}(c)$, there exists an neighborhood $U \subset \mathbb{R}^{m+n}$ and a diffeomorphism $G : U \to V \subset \mathbb{R}^{m+n}$ such that*

$$G(U \cap F^{-1}(c)) = \{(x, y) \in V | y = 0 \in \mathbb{R}^n\}.$$

Now we have more examples of manifolds.

**Example 1.2.6** *1. The $n$-sphere*

$$S^n = \{x \in \mathbb{R}^n | |x| = 1\}.$$

*2. The group $O(n)$ of $n \times n$ orthogonal matrices, i.e.*

$$O(n) = \{A \in M(n) | A^T A = I_n\}.$$

We also note that some manifolds are groups in the mean time. This leads to the definition of Lie groups, which provides a rich class of manifolds with nice properties.

**Definition 1.2.7** *A* Lie group *$G$ is a manifold which is also a group, such that the map*

$$G \times G \to G, (x, y) \mapsto x \cdot y^{-1}$$

*is smooth.*

In fact, for differentiable manifolds, if the group multiplication is a smooth map, then one can prove the inverse map is also smooth, see [5]. Many classical Lie groups arise from the matrix groups. For example the *general linear group*

$$GL(n) = \{A \in M(n) | \det A \neq 0\}$$

and the *special linear group*

$$SL(n) = \{A \in M(n) | \det A = 1\}$$

are both Lie groups.

**Exercise 1.2.8** *Show that $SL(n)$ is a Lie group.*

### 1.2.3 Submanifolds

The manifold $F^{-1}(c)$ in Theorem 1.2.5 is a subset of a larger manifold, which is called a *submanifold*. We can also view submanifolds as the image of a map from a given base manifold, as follows.

**Definition 1.2.9** *Suppose $M, N$ is a $m$- and $n$-dimensional manifold respectively, and $F : M \to N$ is a smooth map. If for all $p \in M$, $\operatorname{rank} F_p = m$, then $F$ is called an* immersion *of $M$ into $N$, and $(M, F)$ is called an* immersed submanifold *of $N$.*

By the Rank Theorem (Theorem 0.2.10), for any $p \in M$, there exists a local chart $(U, \phi; x)$ and $(V, \psi; y)$ such that locally the map has the form

$$y = \psi \circ F \circ \phi^{-1}(x) = (x, 0) \in \mathbb{R}^m \times \mathbb{R}^{n-m}.$$

It follows that $F : U \to F(U)$ is injective and homeomorphic. However, an immersion is not necessarily injective globally. Also, the topology of $M$ need not be the same as the induced topology on $F(M)$ from $N$.

**Example 1.2.10** *1. Consider the curve $F : \mathbb{R}^1 \to \mathbb{R}^2$ by*

$$F(t) = (\cos t, \sin t)\,.$$

*One checks $F$ is a immersion, but it is not injective since it is periodic.*

*2. Consider the curve $G : \mathbb{R}^1 \to \mathbb{R}^2$ by*

$$G(t) = \Big(2\cos(2\arctan t - \frac{\pi}{2}), \sin 2(2\arctan t - \frac{\pi}{2})\Big)\,.$$

*One checks $G$ is a injective immersion. However, the induced subset topology on $G(\mathbb{R}^1)$ from $\mathbb{R}^2$ is different from the standard one on $\mathbb{R}^1$.*

Thus we introduce the following definition. Recall that in the subset topology on $F(M) \subset N$, an set $U \subset F(M)$ is open iff it is the intersection of $F(M)$ and an open set in $N$.

**Definition 1.2.11** *Let $F : M \to N$ be a smooth immersed submanifold. If $F$ is injective and $F : M \to f(M)$ is a homeomorphism under the induced subset topology on $F(M)$, then $F$ is called an* embedding *and $(M, F)$ is called an* embedded submanifold *of $N$.*

A theorem in topology says that an injective continuous map from a compact space to a Hausdorff space is a homeomorphism. As a consequence, we have

**Theorem 1.2.12** *Any injective immersion from a compact manifold $M$ is an embedding.*

Again by the Rank Theorem 0.2.10, we have a local description of the image of an embedded submanifold. The following theorem says that, in a suitable local chart, an embedded submanifold can be realized as a hyperplane of same dimension.

**Theorem 1.2.13** *Suppose $F : M \to N$ is an injective immersion. Then $F$ is an embedding if and only if for any $p \in M$, there is a local chart $(U, \phi)$ containing $q = f(p)$ such that $\phi(q) = 0$ and*

$$F(M) \cap U = \{s \in U | \phi^\alpha(s) = 0, m+1 \leq \alpha \leq n\}.$$

**Exercise 1.2.14** *Let $f : \mathbb{R}^1 \to \mathbb{R}^2, t \to (t, kt)$ be a line through the origin where $k \in \mathbb{Q}^c$ is an irrational number. Show that $F = \pi \circ f : \mathbb{R}^1 \to \mathbb{T}^2$ is an injective immersion but not an embedding. Here $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z} \times \mathbb{Z}$ is the 2 dimensional torus and $\pi : \mathbb{R}^2 \to \mathbb{T}^2$ is the projection map.*

Using the above terminology, we can state a global version of Theorem 1.2.5 as follows.

**Theorem 1.2.15** *Suppose $F : M \to N$ is a smooth map of rank $r$, then for any $q \in F(M)$, the pre-image $F^{-1}(q)$ is a $(m - r)$ dimensional embedded submanifold of $M$.*

**Exercise 1.2.16** *Prove Theorem 1.2.15.*

### 1.2.4 Embedding theorems

Submanifolds can be very complicated. Actually, any compact manifold can be embedded into a sufficiently large dimensional Euclidean space. Therefore, the study of manifolds can be alternatively developed by studying submanifolds in Euclidean spaces.

**Theorem 1.2.17** *Suppose $M$ is a $m$-dimensional compact manifold, then there exists an integer $n$ such that $M$ can be embedded into $\mathbb{R}^n$.*

*Proof.* Since $M$ is compact, by Theorem 1.2.12, we only need to construct an injective immersion.

Choose an atlas of finitely many local charts $(U_\lambda; x^i_\lambda), 1 \leq \lambda \leq r$, and open sets $V_\lambda, W_\lambda$, such that $\bar{V}_\lambda \subset W_\lambda \subset \bar{W}_\lambda \subset U_\lambda$ and $\{V_\lambda, 1 \leq \lambda \leq r\}$ is a cover of $M$. Thus by Lemma 1.1.8, there exists smooth cut-off functions $f_\lambda$ which equals 1 on $V_\lambda$ and vanishes outside of $W_\lambda$.

Now for $n = r(m+1)$, we define a map $F : M \to \mathbb{R}^n$ by defining $n$ smooth functions as components of $F$. For $1 \leq i \leq m$ and $1 \leq \lambda \leq r$, let

$$y_\lambda^0 = f_\lambda, \quad y_\lambda^i = \begin{cases} x_\lambda^i(p) \cdot f_\lambda(p), & p \in U_\lambda \\ 0, & p \notin U_\lambda. \end{cases}$$

That is, we simply define $y_\lambda^i$ by the coordinate chart multiplied with the cut-off function, and add an extra coordinate $y_\lambda^0$ as an index. This map $F$ takes each local chart to different slices of the large Euclidean space $\mathbb{R}^n$.

It remains to check that $F$ is injective and is an immersion. First note that for any $p \in M$, there is some $V_\lambda$ around $p$ such that $y_\lambda^i = x_\lambda^i(p), i = 1, \cdots, m$. Thus the sub-matrix

$$\left( \frac{\partial y_\lambda^i}{\partial x_\lambda^j} \right)_{1 \leq i,j \leq m} = I_m$$

is invertible, hence $F$ has rank $m$ at $p$. Next if there is $p_1, p_2 \in U_\lambda$ such that $F(p_1) = F(p_2)$, then we would have

$$y_\lambda^0(p_1) = f_\lambda(p_1) = y_\lambda^0(p_2) = f_\lambda(p_2)$$

and

$$y_\lambda^i(p_1) = x_\lambda^i(p_1) \cdot f_\lambda(p_1) = y_\lambda^i(p_2) = x_\lambda^i(p_2) \cdot f_\lambda(p_2).$$

It follows $x_\lambda^i(p_1) = x_\lambda^i(p_2)$. Since $x_\lambda^i$ is an coordinate chart, we have $p_1 = p_2$. Therefore, $F$ is injective and the proof is finished. $\square$

The above theorem can be greatly improved. In fact, the assumption of compactness can be replaced by paracompact, and the dimension $n$ can be reduced to $2m + 1$. This is the famous Whitney's embedding theorem.

**Theorem 1.2.18 (Whitney's theorem)** *Every differentiable $m$-dimensional manifold can be embedded in $(2m+1)$-dimensional Euclidean space.*

For a proof of this theorem, we refer to Chapter 1.22 in [2].

# Chapter 2

# Differentiation on Manifolds

## 2.1 Tangent vector and Tangent space

One of our main goals in differential manifolds is to perform derivations (on various objects) on manifolds. Let's first recall derivatives of functions on Euclidean spaces. Suppose $f : \mathbb{R}^n \to \mathbb{R}^1$ is a smooth function, to define a derivative of $f$ at some point $x$, we first specify a direction (at $x$), which is just a vector $v \in \mathbb{R}^n$, and then define the directional derivative by

$$D_v f(x) = \lim_{h\to 0} \frac{f(x+hv) - f(x)}{h}.$$

However, on a manifold which in general is not a linear space, we don't know what a vector is. Thus the key to generalize the idea, is an equivalent yet more geometric concept of 'vectors', which represents 'directions'.

There are usually three ways to define the tangent space at a point of a manifold:

1. define the tangent vectors as equivalent classes of curves
2. define the tangent vectors as linear operators on functions
3. first define the cotangent vectors as equivalent classes of functions, then define the tangent space as the dual space of cotangent space.

### 2.1.1 Tangent vectors as equivalent classes of curves

Recall that in mechanics, a vector is usually the velocity of some particle. The velocity in turn can be regarded as an infestimal movement of that particle. Namely, if a point is moving along a path $\gamma : (a, b) \to \mathbb{R}^3$, then the velocity at time $t$ is $\gamma'(t)$, which is the limit of the average speed in a short time interval. Thus a velocity is associated with some trajectory, as a vector is tangent to some curve in the space. However, the correspondence is not unique: different curves may very well be tangent to the same vector.

Now let's apply the ideas to our manifolds. Given a point $p \in M$, we consider a particle moving through $p$. In mathematics, we represent a path passing $p$ as a curve $\gamma : (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$. Then the velocity at $t = 0$ should give us a 'vector'. We think of this in the opposite way: a 'vector' (though not defined yet) could be represented by a curve (which we already know how to define). The method should work as long as we keep one thing in

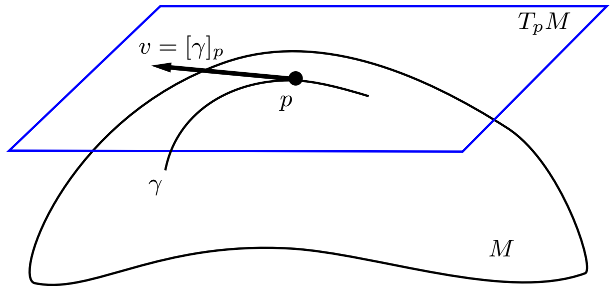


Figure 2.1: tangent space

mind: different curves passing $p$ with the same velocity represents the same vector. This leads to our first definition of a tangent vector.

We say two curves $\beta$ and $\gamma$ passing $p$ are equivalent if, under a local chart $(U, \phi)$ at $p$,

$$\frac{d}{dt}\Big|_{t=0}(\phi \circ \gamma) = \frac{d}{dt}\Big|_{t=0}(\phi \circ \beta).$$

**Definition 2.1.1** *The equivalent class $[\gamma]_p$ of a smooth curve $\gamma$ passing $p$ is called a* tangent vector *at $p \in M$. The space of all possible tangent vectors at $p$ is called the* tangent space *of $p$, and is denoted by $T_pM$.*

The equivalence relationship is independent of the choice of local charts, since the derivatives in different charts only differs by an invertible linear transformation. Thus the tangent vector given by the equivalent class $[\gamma]_p$ is well-defined. See Figure 2.1.

### 2.1.2 Tangent vectors as linear maps

The definition of tangent vectors induces a natural action of a tangent vector on a smooth function, namely, by *directional derivative*. More specifically, given a tangent vector $v = [\gamma]_p \in T_pM$ and a smooth function $f$ at $p$, we define

$$v \cdot f = \frac{d}{dt}\Big|_{t=0}(f \circ \gamma) = \lim_{t \to 0} \frac{f(\gamma(t)) - f(p)}{t}.$$

This leads to the second definition of tangent vectors as follows.

**Definition 2.1.2** *A tangent vector at $p \in M$ is a linear map $X_p : C^\infty(M) \to \mathbb{R}^1$ which satisfies the Leibnitz rule*

$$X_p(f \cdot g) = f(p) \cdot X_p g + g(p) \cdot X_p f.$$

Obviously, the directional derivative on smooth functions induced by a tangent vector defined in Definition 2.1.1 satisfies the Leibnitz rule. Therefore, for any $[\gamma]_p \in T_pM$ in Definition 2.1.1, we can define a linear map $X_p$ as in Definition 2.1.2 by

$$X_p(f) = [\gamma]_p \cdot f = \frac{d}{dt}\Big|_{t=0}(f \circ \gamma).$$

On the other hand, any linear map $X_p$ in Definition 2.1.2 can be realized as an equivalent class $[\gamma]_p$ of smooth curves through $p$ as in Definition 2.1.1. To do this, we choose a local chart $(U; x^i)$ at $p$ and find a simplest curve, i.e. a straight line, as a representative of $[\gamma]_p$. In fact, if the action of $X_p$ on the $i$-th coordinate function is

$$X_p(x^i) = a^i \in \mathbb{R}^1,$$

Then we can define a line $\gamma$ by

$$x^i \circ \gamma(t) = a^i t + x^i(p), 1 \le i \le n.$$

Finally, to show the equivalence of the two definitions of tangent vectors, we only need to show that the above two mappings are inverse of each other, which is left to the readers.

### 2.1.3 Tangent space

Given a local chart $(U, \phi)$ with $\phi(p) = 0$, there is a natural set of tangent vectors in $T_pM$ given by the coordinate axes at $p$. More specifically, for each $1 \le i \le n$, let $\gamma^i$ be the smooth curve through $p$ such that $\phi(\gamma^i)$ is just the $i$-th axis. That is, we require

$$x^j \circ \gamma^i(t) = \delta_{ij} t,\ 1 \le j \le n.$$

where

$$\delta_{ij} = \begin{cases} 0, \ i \neq j; \\ 1, \ i = j. \end{cases}$$

is the *Kronecker symbol*. Then we define the tangent vectors

$$\frac{\partial}{\partial x^i}\Big|_p := [\gamma^i]_p, \quad 1 \leq i \leq n.$$

By definition, for a smooth function $f \in C^\infty(M)$, we have

$$\begin{aligned}
\frac{\partial}{\partial x^i}\Big|_p \cdot f &= \frac{d}{dt}\Big|_{t=0}(f \circ \gamma^i) \\
&= \sum_{j=1}^n \frac{\partial}{\partial x^j}\Big|_{x=x_0}(f \circ \phi^{-1}) \cdot \frac{d}{dt}\Big|_{t=0}(x^j \circ \gamma^i) \\
&= \frac{\partial}{\partial x^i}(f \circ \phi^{-1})(x_0).
\end{aligned}$$

One checks that $\{\frac{\partial}{\partial x^i}|_p\}$ forms a basis of the tangent space $T_pM$, which is called the *natural frame*. Actually, for any tangent vector $X \in T_pM$ and $a^i = X(x^i) \in \mathbb{R}^1, 1 \leq i \leq n$, we have

$$X = a^1 \frac{\partial}{\partial x^1}\Big|_p + \cdots + a^n \frac{\partial}{\partial x^n}\Big|_p = a^i \frac{\partial}{\partial x^i}\Big|_p.$$

Here we adopt *Einstein's convention of summations*. That is, when the index $i$ appears twice (usually as subscript on the top and bottom), then we take the summation for $i = 1, \cdots n$.

Clearly, the definitions works fine on Euclidean spaces: the tangent vector $\frac{\partial}{\partial x^i}\Big|_p$ is just the $i$-th unit vector along $x^i$-axis. But on a general manifold, there is a major difference here: the tangent spaces $T_pM$ and $T_qM$ at different points $p, q \in M$ are essentially different. To be more precise, although they are linear spaces of the same dimension, there is no *canonical* way of identifying them (like the parallel translation in Euclidean space). Thus a tangent vector at $p$ can not be viewed as a tangent vector at $q \neq p$.

**Exercise 2.1.3** *Prove that the tangent space $T_pM$ is an $n$-dimensional vector space by showing that the tangent vectors $\{\frac{\partial}{\partial x^i}\Big|_p\}_{1\leq i\leq n}$ forms a basis.*

Since we often need to transform from a local chart to another, it is very important to record how the natural frames transforms under change of local charts.

Suppose $(U,\phi;x^i)$ and $(V,\psi;y^i)$ are two local charts on $M$ such that $U\cap V\neq\emptyset$. Then for $p\in U\cap V$ there are two sets of basis $\{\frac{\partial}{\partial x^i}|_p\},\{\frac{\partial}{\partial y^i}|_p\}\subset T_pM$. Recall that for a smooth function $f\in C^\infty(M)$,

$$\frac{\partial}{\partial x^i}\Big|_p\cdot f=\frac{\partial}{\partial x^i}(f\circ\phi^{-1})(x_0).$$

Similarly, for $y_0=\psi(p)$,

$$\frac{\partial}{\partial y^i}\Big|_p\cdot f=\frac{\partial}{\partial y^i}(f\circ\psi^{-1})(y_0).$$

But on $U\cap V$, we have $x=x(y)=\phi\circ\psi^{-1}(y)$. Thus

$$\begin{aligned}\frac{\partial}{\partial y^i}(f\circ\psi^{-1})(y_0)&=\frac{\partial}{\partial y^i}[(f\circ\phi^{-1})\circ(\phi\circ\psi^{-1})](y_0)\\&=\frac{\partial}{\partial x^j}(f\circ\phi^{-1})(x_0)\cdot\frac{\partial x^j}{\partial y^i}(y_0)\\&=\frac{\partial x^j}{\partial y^i}(y_0)\frac{\partial}{\partial x^j}\Big|_p\cdot f\end{aligned}$$

Since the equality holds for any $f\in C^\infty$, we obtain

$$\frac{\partial}{\partial y^i}\Big|_p=\frac{\partial x^j}{\partial y^i}(y_0)\frac{\partial}{\partial x^j}\Big|_p.$$

**Exercise 2.1.4** *Let $\{v_i\}_{i=1}^n\in T_pM$ be a basis of tangent vectors at $p$. Show that there is a local chart $(U,\phi;x^i)$ around $p$ such that*

$$v_i=\frac{\partial}{\partial x^i}\Big|_p, i=1,\cdots,n.$$

### 2.1.4 Cotangent space

Yet another method to define the tangent space is to first define its dual space by equivalent classes of smooth functions, as follows. For $f,g\in C^\infty(M)$, we define equivalent relation $f\sim g$ if in a local chart $(U,\phi)$,

$$d(f\circ\phi^{-1})|_p=d(g\circ\phi^{-1})|_p.$$

Note that the equivalent relation is well-defined since it is independent of the choice of local charts. Now we define the *cotangent space* at $p$ by

$$T_p^*M = C^\infty(M)/\sim = \{[f]_p | f \in C^\infty(M)\}.$$

Note that in the above definition, we can replace the space of global functions $C^\infty(M)$ by locally defined functions $C_p^\infty(M)$. Obviously, the cotangent space inherits a linear structure from $C^\infty(M)$, once we define

$$\lambda[f]_p + \mu[g]_p = [\lambda f + \mu g]_p, \quad \forall \lambda, \mu \in \mathbb{R}^1, f, g \in C^\infty(M).$$

In a local chart $(U, x^i)$, there is a natural choice of cotangent vectors given by the equivalent class of coordinate functions

$$dx^i|_p := [x^i]_p, \ i = 1, \cdots, n.$$

We usually denote the equivalent class of $f \in C^\infty(M)$ by $df|_p = [f]_p \in T_p^*M$. Since in a local chart

$$d(f \circ \phi^{-1}) = \frac{\partial}{\partial x^i}(f \circ \phi^{-1})(x_0)dx^i,$$

it follows that

$$df|_p = \frac{\partial}{\partial x^i}(f \circ \phi^{-1})(x_0)dx^i|_p.$$

There is a natural pairing between the cotangent space $T_p^*M$ and the tangent space $T_pM$ as follows. For any tangent vector $[\gamma]_p = v^i \frac{\partial}{\partial x^i}|_p \in T_pM$, we let

$$(df|_p, [\gamma]_p) = [\gamma]_p \cdot f = \frac{d}{dt}\Big|_{t=0}(f \circ \gamma) = v^i \frac{\partial}{\partial x^i}(f \circ \phi^{-1})(x_0).$$

**Theorem 2.1.5** *The cotangent space $T_p^*M$ is the dual space of the tangent space $T_pM$. In a local chart $(U, x^i)$, $\{dx^i|_p\} \subset T_p^*M$ is a dual basis of $\{\frac{\partial}{\partial x^i}|_p\} \subset T_pM$.*

**Exercise 2.1.6** *Prove Theorem 2.1.5.*

**Exercise 2.1.7** *Show that under a change of local charts, the natural basis of the cotangent* *space $T_p^*M$ transforms by*

$$dx^i|_p = \frac{\partial x^i}{\partial y^j}(y_0)dy^j|_p.$$

### 2.1.5 Tangent map

The tangent map is a natural generalization of the *full derivatives* of functions between Euclidean space. Recall that for a vector valued multi-variable function $F : \mathbb{R}^m \to \mathbb{R}^n$, the derivative

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

is a (Jacobian) matrix, which can also be viewed as a linear map.

A tangent vector can also act on a smooth map between manifolds, which should be compared to directional derivative of vector valued functions. Let $F : M \to N$ be a smooth map and $v = [\gamma]_p \in T_pM$ be a tangent vector. Note that $F$ maps a curve $\gamma$ passing $p$ in $M$ to a curve $F \circ \gamma$ passing $q := F(p)$ in $N$. One can verify that $F$ actually maps the equivalent class $[\gamma]_p$ to a equivalent class $[F \circ \gamma]_q$ in $T_qN$. In other words, for a smooth function $g$ at $q$, the action

$$[F \circ \gamma]_q \cdot g = \frac{d}{dt}\Big|_{t=0}(g \circ F \circ \gamma)$$

is independent of the choice of representatives in $[\gamma]_p$. Thus the map from $[\gamma]_p$ to $[F \circ \gamma]_q$, which is introduced by $F$, is well-defined. We call this map the *tangent map* of $F$ at $p$, and denote it by $(F_*)|_p$ or $dF|_p$. Therefore, we define the derivative of $v$ on $F$ by

$$v \cdot F = [\gamma]_p \cdot F = dF|_p(v) = [F \circ \gamma]_q \in T_qN.$$

That is, the tangent map $dF|_p$ *pushes forward* a tangent vector $v \in T_pM$ to a tangent vector $dF|_p(v) \in T_qN$. See Figure 2.2.

It is easy to see that the tangent map $dF|_p$ of $F : M \to N$ is a linear map from $T_pM$ to $T_qN$. Given two basis of the two tangent spaces, $dF|_p$ can be realized as a $m \times n$ matrix. In fact, for two local charts $(U, \phi; x^i)$ on $M$ and$(V, \psi; y^\alpha)$ on $N$ containing $p$ and $q$ respectively, we have

$$\frac{\partial}{\partial x^i}\Big|_p \cdot F = [F \circ \gamma^i]_q = a_i^\alpha \frac{\partial}{\partial y^\alpha}\Big|_q$$

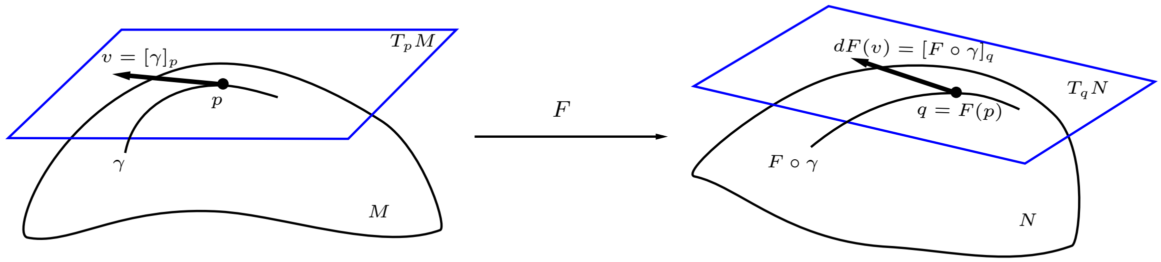


Figure 2.2: Tangent map

for some numbers $a_i^\alpha \in \mathbb{R}^1$. Since for $g \in C^\infty(N)$, by the chain rule,

$$\begin{aligned}
[F \circ \gamma^i]_q \cdot g &= \frac{d}{dt}\Big|_{t=0}(g \circ F \circ \gamma^i) \\
&= \frac{\partial}{\partial x^i}(g \circ F \circ \phi^{-1})(x_0) \\
&= \frac{\partial}{\partial x^i}[(g \circ \psi^{-1}) \circ (\psi \circ F \circ \phi^{-1})](x_0) \\
&= \frac{\partial}{\partial y^\alpha}(g \circ \psi^{-1})(y_0) \cdot \frac{\partial}{\partial x^i}(\psi \circ F \circ \phi^{-1})^\alpha(x_0)
\end{aligned}$$

If we denote the function $y = y(x) = \psi \circ F \circ \phi^{-1}(x)$, then we arrive at

$$dF|_p\left(\frac{\partial}{\partial x^i}\Big|_p\right) = \frac{\partial}{\partial x^i}\Big|_p \cdot F = \frac{\partial y^\alpha}{\partial x^i}(x_0)\frac{\partial}{\partial y^\alpha}\Big|_q.$$

and the matrix is given by

$$A = (a_i^\alpha) = \left(\frac{\partial y^\alpha}{\partial x^i}(x_0)\right), 1 \leq i \leq m, 1 \leq \alpha \leq n.$$

We call the rank of the matrix $A$ the *rank* of $F$ at $p$, which is well-defined since it is invariant under change of coordinate charts.

### 2.1.6 Pull-back map

Recall that for any linear map between two vector spaces $L : V \to W$, there is a dual linear map $L^* : W^* \to V^*$ defined by the identity

$$(Lv, w^*) = (v, L^* w^*), \quad \forall v \in V, w^* \in W^*.$$

The derivative of a smooth map $F : M \to N$ also induces a linear map between cotangent spaces. More precisely, let $p \in M$ and $q = F(p) \in N$, then for any $f \in C^\infty(N)$, $f \circ F$ defines a smooth function on $M$. This leads to a linear map

$$F^*|_p : T^*_q N \to T^*_q M, [f]_q \to [f \circ F]_p.$$

The map $F^*$ is called a pull-back map, since it pulls a cotangent vector in $T^*_q N$ to a cotangent vector in $T^* pM$.

By a similar computation as the push-forward map, we find that, in local coordinates $(U; x^i)$ in $M$ and $(V; y^\alpha)$ in $N$,

$$F^*_p(dy^\alpha|_q) = \frac{\partial y^\alpha}{\partial x^i}(x_0) dx^i|_p.$$

## 2.2 Tangent bundle and Vector fields

In last section, we define the tangent vector and tangent space $T_pM$ at a fixed point $p \in M$. Now we want to let $p$ vary and consider a tangent vector field $X$ on $M$. That is, $X$ assigns a tangent vector in $T_pM$ for every point $p \in M$.

### 2.2.1 Tangent bundle

Consider the *disjoint* union of all tangent spaces of a manifold

$$TM = \cup_{p\in M} T_pM.$$

There is a natural topological and differentiable structure on $TM$ induced by that of $M$. First we consider the projection

$$\pi : TM \to M, \pi(X_p) = p, \forall X_p \in T_pM.$$

Obviously, $\pi^{-1}(p) = T_pM$ for all $p \in M$. For a coordinate chart $(U, \phi; x^i)$ on $M$, there is a so-called *local trivialization*

$$\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^n, X_p \to (p, X_p(x^1), \cdots, X_p(x^n)).$$

That is, if $X_p = a^i \frac{\partial}{\partial x^i}|_p \in T_pM$ in the natural local frame, then

$$\Phi(X_p) = (p, a^1, \cdots, a^n).$$

This is an bijection and thus induces a topology on the tangent bundle $TM$, such that $\Phi$ is a homeomorphism. In other words, we say $V \subset TM$ is an open set in $TM$ if the set $\Phi(V \cap \pi^{-1}(U))$ is open in $U \times \mathbb{R}^n$ for all local charts on $M$.

**Exercise 2.2.1** *Check that the induce topology on $TM$ is second countable and Hausdorff.*

In fact, the tangent bundle is a $2n$-dimensional differentiable manifold. To see this, we only need to construct an atlas. Let $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ be an atlas of $M$, and $\Phi_\alpha$ be the local trivialization corresponding to $(U_\alpha, \phi_\alpha)$. Then one checks that an atlas on $TM$ is given by

$$\{(\tilde{U}_\alpha, \tilde{\Phi}_\alpha) = (\pi^{-1}(U_\alpha), (\phi_\alpha, id) \circ \Phi_\alpha)\}_{\alpha \in I}.$$

Indeed, for two local charts $(U, \phi; x)$ and $(V, \psi; y)$ on $M$. Recall that at any point $p \in U \cap V$, we have

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

Thus for $(x, v) \in \tilde{\Phi}(\tilde{U} \cap \tilde{V}) \subset \mathbb{R}^n \times \mathbb{R}^n$, we have

$$\tilde{\Psi} \circ \tilde{\Phi}^{-1}(x, v) = \tilde{\Psi}\left(v^i \frac{\partial}{\partial x^i}\right) = \tilde{\Psi}\left(v^i \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}\right) = (y(x), g_x(v)),$$

where $y(x) = \psi \circ \phi^{-1}(x)$ and the transition map

$$g_x = \left(\frac{\partial y^j}{\partial x^i}(x)\right)_{1 \le i, j \le n} \in GL(n), \forall x \in U \cap V.$$

**Definition 2.2.2** *The differentiable manifold $TM$ defined as above is called the tangent bundle of $M$.*

In a similar way, we can define the *cotangent bundle* $T^*M = \cup_{p \in M} T^*M$, which is also a $2n$-dimensional differentiable manifold.

**Exercise 2.2.3** *Construct an atlas on the cotangent bundle $T^*M$.*

### 2.2.2 Tangent fields

One should be familiar with the notion of vector fields, which appears in classical physics. By definition, a vector field on $\Omega \subset \mathbb{R}^n$ is just a vector valued function $X : \Omega \to \mathbb{R}^n$. For example the velocity field of fluids, the electric field and the magnetic field in $\mathbb{R}^3$. Here we can regard tangent vector fields on manifolds as a generalization of vector fields on Euclidean spaces. An example is the velocity field of wind on the earth, which is a vector field on the 2-sphere $S^2$.

**Definition 2.2.4** *A tangent vector field is a smooth map $X : M \to TM$ such that*

$$\pi \circ X(p) = p, \forall p \in M.$$

*The space of all tangent vector field is denoted by $\mathfrak{X}(M) = \Gamma(TM)$.*

Intuitively, a tangent vector field assigns a tangent vector to each point in a smooth way. Given a local chart $(U, \phi; x^i)$, the coordinate functions give a natural local frame

$$\partial_i := \frac{\partial}{\partial x^i} : p \to \frac{\partial}{\partial x^i}\Big|_p \in T_pM, i = 1, \cdots, n.$$

Thus the restriction of a tangent vector field $X \in \mathfrak{X}(M)$ on $U$ can be written as

$$X|_U = X^1 \frac{\partial}{\partial x^1} + \cdots X^n \frac{\partial}{\partial x^n} = X^i \partial_i,$$

where $X^i : U \to \mathbb{R}^1$ are smooth functions.

Via the directional derivative, a tangent vector field $X \in \mathfrak{X}(M)$ acts on a smooth function $f \in C^\infty(M)$ by

$$X(f) : p \to X(p) \cdot f \in \mathbb{R}^1, \quad \forall p \in M,$$

which gives a linear map

$$X : C^\infty(M) \to C^\infty(M), f \mapsto X(f).$$

Obviously, this map satisfies the Leibnitz rule

$$X(f \cdot g) = X(f) \cdot g + f \cdot X(g), \quad \forall f, g \in C^\infty(M).$$

Conversely, a linear map satisfying the Leibnitz rule can be realized as a tangent vector field.

A tangent vector field $X \in \mathfrak{X}(M)$ can also act on a smooth map $F : M \to N$ by

$$F_*(X) = X(F) : p \to X|_p \cdot F = dF(X_p) \in T_{F(p)}N.$$

We call $F_*$ the tangent map of $F$, or *push-forward*. In local coordinates $(U; x^i)$ on $M$ and $(V; y^\alpha)$ on $N$, we have

$$F_*(X) = dF\left(X^i \frac{\partial}{\partial x^i}\right) = X^i \frac{\partial F^\alpha}{\partial x^i} \frac{\partial}{\partial y^\alpha}.$$

Similarly, we can define cotangent vector fields in the cotangent bundle $T^*M$. Then $F$ induces a *pull-back* map $F^*$ on cotangent vector fields.

Form now on, we will simply call a tangent vector field by vector field. Unlike the Euclidean space, generally there does not exist a "constant" vector field on a non-trivial tangent bundle. For example, the famous *Poincaré-Hopf theorem* says that there is no non-vanishing tangent vector field on $S^2$.

### 2.2.3 Vector bundles

The concept of tangent bundles can be greatly generalized to general vector bundles. Intuitively, a vector bundle is just a family of vector spaces attached to a manifold, which is glued together in a specific way. More precisely, we define

**Definition 2.2.5** *A vector bundle $(E, M, \pi)$ of rank $r$ consists of a bundle space $E$, a base manifold $M$ and a projection $\pi : E \to M$ such that*

1. *For all $p \in M$, the fiber $E_p := \pi^{-1}(p)$ is isomorphic to $\mathbb{R}^r$*
2. *For all $p \in M$, there exists a neighborhood $U \subset M$ and a local trivialization, i.e. a diffeomorphism*
$$\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^r$$
*where the restriction of $\Phi$ on $E_p$ is linear.*
3. *The map $\Phi$ satisfies*
$$\pi \circ \Phi^{-1}(p, v) = p, \forall p \in M, v \in \mathbb{R}^r.$$

Suppose there are two local charts $(U, \phi)$ and $(V, \psi)$ on $M$, and two local trivialization maps

$$\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^n, \Psi : \pi^{-1}(V) \to V \times \mathbb{R}^n.$$

Then they are glued together by

$$\Psi \circ \Phi^{-1} = (\pi_1, g) : (U \cap V) \times \mathbb{R}^n \to (U \cap V) \times \mathbb{R}^n.$$

Here $\pi_1$ is the projection to the first variable, and the transition map $g_p = g(p, \cdot) : \mathbb{R}^n \to \mathbb{R}^n$ is linear for all $p \in U \cap V$.

A key point in the definition of vector bundles is the linear structure of the fiber space, and that the transition maps keep the linear structure. Actually, one can alternatively define a vector bundle by using the transition maps.

**Theorem 2.2.6** *Suppose $M$ is a manifold with an open cover $\{U_\alpha\}_{\alpha \in I}$. If for all $\alpha, \beta \in I$, there is a smooth map*

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(r),$$

*such that for all $\alpha, \beta, \gamma \in I$,*

- $g_{\alpha\alpha} = I_r$;
- $g_{\alpha\beta} \cdot g_{\beta\gamma} = g_{\alpha\gamma}$.

*then there exists a vector bundle $(E, M, \pi)$ whose transition maps are exactly $\{g_{\alpha\beta}\}_{\alpha,\beta \in I}$.*

It follows from the definition that $g_{\alpha\beta} = g_{\beta\alpha}^{-1}, \forall \alpha, \beta \in I$.

**Example 2.2.7** *1. The tangent bundle of $S^1$ is a trivial bundle $TS^1 = S^1 \times \mathbb{R}^1$.*

*2. The Möbius band $\mathcal{M}$ is a non-trivial bundle on $S^1$, which is constructed as follows. First identify $S^1 \simeq \mathbb{T}^1 = \mathbb{R}^1/\mathbb{Z}^1$. Then take $U = [(0, 2\pi)], V = [(-\pi, \pi)] \subset S^1$ and let $\tilde{U} = U \times \mathbb{R}^1, \tilde{V} = V \times \mathbb{R}^1$. Define the transition map $g : U \cap V \to GL(1)$ by*

$$g(\theta) = \begin{cases} 1, & \theta \in [(0, \pi)] \\ -1, & \theta \in [(\pi, 2\pi)]. \end{cases}$$

*Then the Möbius band is glued by $\mathcal{M} = (\tilde{U} \cup \tilde{V})/g$. Namely, we identify points $(\theta_1, v_1) \in \tilde{U}$ and $(\theta_2, v_2) \in \tilde{V}$ if $\theta_1 = \theta_2$ and $v_2 = g(\theta_1) \cdot v_1$.*

**Definition 2.2.8** *A section of a vector bundle $(E, M, \pi)$ is a smooth map $s : M \to E$ such that $\pi \circ s = id_M$. The space of all sections is denoted by $\Gamma(E)$.*

In other words, for any $p \in M$, a section $s$ takes value $s(p) \in E_p$. In a local trivialization, $s$ is equivalent to a smooth map $u : U \to \mathbb{R}^n$ such that

$$\Phi \circ s(p) = (p, u(p)), \forall p \in U.$$

From the definition, we can see that locally, a fiber bundle has a product space structure by the local trivialization map. However, in general, a vector bundle is not a product space globally. If there is a global trivialization $\Phi : E \to M \times \mathbb{R}^n$, then $E$ is called a *trivial* bundle. In fact, we have

**Theorem 2.2.9** *A vector bundle $(E, M, \pi)$ of rank $r$ is trivial if and only if there exists $r$ linear independent sections $\{X_1, \cdots, X_r\} \subset \Gamma(E)$.*

**Exercise 2.2.10** *Try to prove that there is no non-vanishing section on the Mobiüs band $\mathcal{M}$ (viewed as a vector bundle over $S^1$). That is, for all section $s \in \Gamma(\mathcal{M})$, there exists $\theta \in S^1$ such that $s(\theta) = 0$.*

## 2.3 Lie derivative

### 2.3.1 1-parameter group of diffeomorphisms

Think of wind velocity (assuming it is constant in time) on the surface of the earth as a vector field on the sphere $S^2$ . There is another interpretation we can make by tracking trajectories of particles. A particle at position $x \in S^2$ moves after time $t$ seconds to a position $\varphi_t(x) \in S^2$ . After a further $s$ seconds it is at

$$\varphi_{t+s}(x) = \varphi_s \circ \varphi_t(x).$$

What we get this way is a homomorphism of groups: from the additive group $\mathbb{R}^1$ to the group of diffeomorphisms of $S^2$ under the operation of composition. The technical definition is the following:

**Definition 2.3.1** *A one-parameter group of diffeomorphisms of a manifold $M$ is a smooth map*

$$\varphi : M \times \mathbb{R}^1 \to M, (x,t) \mapsto \varphi_t(x)$$

*such that*

- $\varphi_t : M \to M$ *is a diffeomorphism for every* $t$
- $\varphi_0 = id_M$
- $\varphi_{s+t} = \varphi_s \circ \varphi_t$.

We will show that vector fields generate one-parameter groups of diffeomorphisms locally, and vice versa. Thus we can interpret vector fields as "infinitesimal diffeomorphisms" rather than as abstract derivations of functions.

First we want to generate a vector field from an one-parameter groups of diffeomorphisms $\varphi_t$. For any $x \in M$, the map $\gamma_x(t) = \varphi_t(x) : \mathbb{R}^1 \to M$ gives a smooth curve through $x$. Then by our first definition of tangent vectors, $\gamma_x$ induces a tangent vector $[\gamma_x] \in T_xM$ such that for any smooth function $f \in C^\infty(M)$,

$$[\gamma_x] \cdot f = \frac{\partial}{\partial t}\Big|_{t=0} f(\gamma_x(t)).$$

So as $x$ varies we have a vector field $X \in \Gamma(TM)$ with $X(x) = [\gamma_x]$. In local coordinates, if we write $y = y(x,t) = \varphi_t(x)$, then

$$X(f)(x) = \frac{\partial}{\partial t}\Big|_{t=0} f(\varphi_t(x)) = \frac{\partial f}{\partial y^i}(y)\frac{\partial y^i}{\partial t}(x)\Big|_{t=0}.$$

Since $y = \varphi_0(x) = x$ at $t = 0$, it follows

$$X(x) = \frac{d}{dt}(x^i \circ \varphi_t(x))\Big|_{t=0}\frac{\partial}{\partial x^i} := \frac{d}{dt}\varphi_t(x)\Big|_{t=0}.$$

### 2.3.2 Integral curve

Next we want to reverse the above: go from a vector field $X \in \Gamma(TM)$ to the diffeomorphism. We start from tracking the "trajectory" of a single particle along the vector field.

**Definition 2.3.2** *An integral curve of a vector field $X$ is a smooth curve $\gamma : I \to M$ such that*

$$d\gamma\left(\frac{d}{dt}\right) = \frac{d\gamma}{dt}(t) = X(\gamma(t)), \forall t \in I.$$

That is, $X$ coincides with the tangent vector generated by the curve $\gamma$. By solving the above ordinary differential equation in a local chart, we can get a local integral curve for the vector field. Next we patch local solutions together to get a maximal one.

**Theorem 2.3.3** *Given a vector field $X \in \Gamma(TM)$ and $p \in M$, there exists an integral curve $\gamma_p : I \to M$ through $p$ with maximal defining interval $I$.*

*Proof.* First consider a local chart $(U, \phi)$ around $p$. Suppose that in this chart, we have $X = X^i\partial_i$. Then the equation of integral curve can be written as

$$\frac{d}{dt}x(t) = X(x(t)).$$

Then the basic theorem of ODE asserts that there is a unique solution $x(t)$ on an interval $[0, T)$ with initial data $x(0) = \phi(p)$. This gives an integral curve in the chart $U$.

Now suppose $\gamma_p : I \to M$ is any integral curve with $\gamma_p(0) = p$ (not necessarily the same as the one we just constructed). For each $\tau \in I$, the subset $\gamma_p([0, \tau]) \subset M$ is compact, so we can cover it by finite number of coordinate charts, in each of which we again apply the basic theorem of ODE. Uniqueness implies that all possible solutions agree with $\gamma_p$ on any subinterval containing 0.

Finally we take the integral curve with maximal defining open interval. $\square$

Note that the maximal interval might not be the whole $\mathbb{R}^1$ if the manifold is not compact (e.g. a open disk $D^2 \subset \mathbb{R}^2$). To find the one-parameter group of diffeomorphisms we now let $p \in M$ vary.

**Theorem 2.3.4** *Let $X \in \Gamma(TM)$ be a vector field and for $x \in M$, let $\varphi_t(x) = \gamma_x(t)$ be the maximal integral curve of $X$ through $x$. Then*

1. *the map $(t, x) \to \varphi_t(x)$ is smooth*
2. *$\varphi_t \circ \varphi_s = \varphi_{t+s}$ wherever the maps are defined*
3. *if $M$ is compact, then $\varphi_t(x)$ is defined on $\mathbb{R}^1 \times M$ and gives a one-parameter group of diffeomorphism.*

*Proof.*

1. The smoothness follows directly from the smooth dependence of the solution of ODEs on the initial value by the theorem of ODE.

2. First we check that for any $x \in M$, the maps $\varphi_t \circ \varphi_s(x)$ and $\varphi_{t+s}$ both gives an integral curve through $y = \varphi_s(x)$. Then the conclusion follows from the uniqueness part of the ODE theorem.

3. Since $M$ is compact, we can find a finite number of open sets $\{U_i\}_{i=1}^K$ such that a smaller compact set $V_i \subset U_i$ still covers $M$. For each $x \in V_i$, the defining interval $I_x$ of the integral curve $\tilde{\varphi}_t(x) = \gamma_x(t)$ depends smoothly on $x$. Thus we can find a positive number $T_i > 0$ such that $\tilde{\varphi}_t$ is defined on $[-T_i, T_i] \times V_i$. Let $T = \min_{1 \leq i \leq K}\{T_i\}$, then the map $\tilde{\varphi}_t$ is well-defined on $[-T, T] \times M$.

   Finally we want to extend $\tilde{\varphi}_t$ to the whole $\mathbb{R}^1$. To do this, for any $t \in \mathbb{R}^1$, choose $n \in \mathbb{Z}$ such that $|t/n| \leq T$, then we define

$$\varphi_t(x) = \tilde{\varphi}_{t/n} \circ \cdots \circ \tilde{\varphi}_{t/n}(x) = [\tilde{\varphi}_{t/n}]^n(x).$$

One checks that this map still satisfies property (2).

□

### 2.3.3 Lie bracket

Given two vector fields $X, Y \in \mathfrak{X}(M)$ viewed as linear maps on $C^\infty(M)$, we can consider the composition $X \circ Y$ and $Y \circ X$. In other words, we consider the double derivatives of a function along two vector fields. However, the composition of two vector fields are not vector fields, since

$$X \circ Y(fg) = X(f(Yg) + (Yf)g) = (Xf)(Yg) + f(XYg) + (Xg)(Yf) + g(XYf),$$
$$Y \circ X(fg) = Y(f(Xg) + (Xf)g) = (Yf)(Xg) + f(YXg) + (Yg)(Xf) + g(YXf).$$

Namely, they do not satisfy the Leibnitz rule. However, it is easy to see that their difference $X \circ Y - Y \circ X$ does. This leads to the following definition:

**Definition 2.3.5** *The* Lie bracket *of vector fields is a linear map*

$$[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$$

*given by*

$$[X, Y] = X \circ Y - Y \circ X.$$

In local coordinates, if $X = X^i\partial_i, Y = Y^i\partial_i$, then by definition

$$[X, Y]f = X^i\partial_i(Y^j\partial_j f) - Y^j\partial_j(X^i\partial_i f) = X^i\partial_i Y^j\partial_j f - Y^j\partial_j X^i\partial_i f.$$

Thus we get

$$[X, Y] = (X^i\partial_i Y^j - Y^i\partial_i X^j)\partial_j.$$

In particular, we have $[\partial_i, \partial_j] = 0, \forall 1 \leq i, j \leq n$.

**Exercise 2.3.6** *Suppose $F : M \to N$ is a diffeomorphism and $X, Y \in \mathfrak{X}(M)$ are tangent vector fields on $M$. Prove that*

$$F_*[X, Y] = [F_*X, F_*Y].$$

One checks that the Lie bracket is bi-linear and satisfies

- skew-symmetry
$$[X, Y] = -[Y, X]$$

- the *Jacobi identity*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

A vector space with a Lie bracket operation satisfying the two properties is called a *Lie algebra*. An elementary example is the familiar $\mathbb{R}^3$ with Lie bracket given by the cross product. A rich source of Lie algebra comes from the tangent space of Lie groups.

### 2.3.4 Lie derivative

The composition of two vector fields or double derivatives of a function raises the problem of derivation of a vector field on a manifold. With the help of 1-parameter group of diffeomorphisms, we can naturally generalizes the usual concept of derivation of vector fields (i.e. vector valued functions) in Euclidean spaces to manifolds.

**Definition 2.3.7** *Let $X, Y \in \mathfrak{X}(M)$ be two vector fields, and $\varphi_t$ be the 1-parameter group of diffeomorphisms generated by $X$. The Lie derivative of $Y$ with regard to $X$ is defined by*

$$\mathcal{L}_X Y = \frac{d}{dt}\Big|_{t=0} (\varphi_{-t})_* \circ Y \circ \phi_t = \lim_{t \to 0} \frac{(\varphi_{-t})_* \circ Y(\varphi_t(x)) - Y(x)}{t}.$$

It turns out that the Lie derivative is exactly the Lie bracket, i.e. the commutator, of two vector fields.

**Theorem 2.3.8** *For any vector fields $X, Y \in \mathfrak{X}(M)$, we have*

$$\mathcal{L}_X Y = [X, Y].$$

*Proof.* [Proof 1.] In local coordinates, suppose $X = a^i \frac{\partial}{\partial x^i}$ and $Y = b^i \frac{\partial}{\partial x^i}$. Let $\varphi_t$ by the 1-parameter group of diffeomorphisms generated by $X$ and $y = \varphi_t(x)$. Then by definition

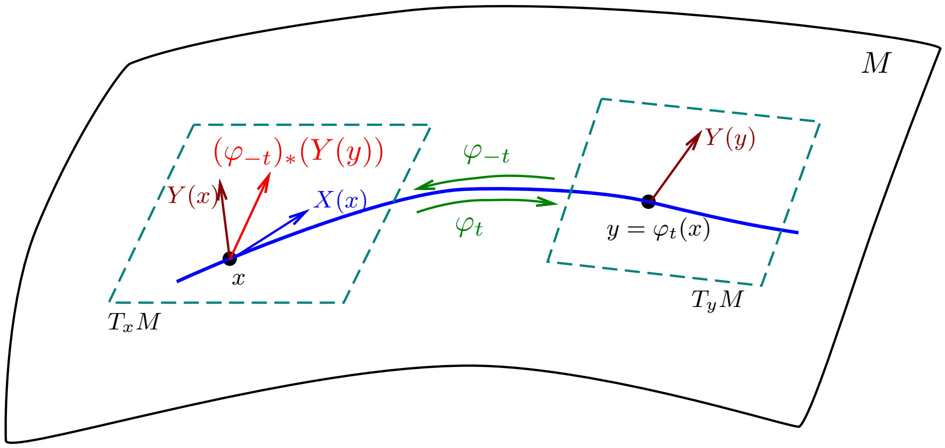


Figure 2.3: Lie Derivative

$\partial_t \varphi_t|_{t=0} = X$ and

$$\begin{aligned}
\mathcal{L}_X Y(x) &= \frac{d}{dt}\Big|_{t=0} (\varphi_{-t})_*(b^i(y)\frac{\partial}{\partial y^i}) \\
&= \frac{d}{dt}\Big|_{t=0} b^i(y)\frac{\partial \varphi_{-t}}{\partial y^i} \\
&= \frac{\partial b^i}{\partial y^j}\frac{dy^j}{dt}\frac{\partial \varphi_{-t}}{\partial y^i}\Big|_{t=0} + b^i(y)\frac{d}{dt}\left(\frac{\partial \varphi_{-t}}{\partial y^i}\right)\Big|_{t=0} \\
&= \frac{\partial b^i}{\partial x^j} a^j \frac{\partial}{\partial x^i} - b^i \frac{\partial a^j}{\partial x^i}\frac{\partial}{\partial x^j}.
\end{aligned}$$

Here in the first term, we use the fact that $\varphi_0 = id$, while we interchange the derivatives of the second term. $\square$

*Proof.* [Proof 2.] For any $x \in M$, let $y = \varphi_t(x)$ and

$$Y_t(x) := (\varphi_{-t})_* \circ Y(y) \in T_xM$$

be a time-dependent tangent vector at $x$, then by definition

$$\mathcal{L}_X Y(x) = \frac{d}{dt}\Big|_{t=0} Y_t(x),$$

But for a smooth function $f$ near $x$,

$$Y_t(x) \cdot f = Y(y) \cdot \varphi_{-t}^* f = \big[Y \cdot (f \circ \varphi_{-t})\big](y)$$

Differentiating the above equation at $t = 0$, we get

$$\mathcal{L}_X Y \cdot f = X \cdot (Y \cdot f) + Y \cdot (-X \cdot f) = [X, Y] f.$$

$\square$

**Remark 2.3.9** *The Lie derivative satisfies the following Leibnitz rule*

$$X \cdot (Y \cdot f) = \mathcal{L}_X Y \cdot f + Y \cdot (X \cdot f).$$

*Compare to general affine connection,*

$$\nabla_X \nabla_Y f = \nabla_X \langle \nabla f, Y \rangle = \langle \nabla_X \nabla f, Y \rangle + \langle \nabla f, \nabla_X Y \rangle = \nabla^2_{X,Y} f + \nabla_{\nabla_X Y} f.$$

*It follows that*

$$\begin{aligned} \nabla^2_{X,Y} f - \nabla^2_{Y,X} f &= \nabla_X \nabla_Y f - \nabla_Y \nabla_X f - \nabla_{\nabla_X Y} f + \nabla_{\nabla_Y X} f \\ &= X \circ Y \cdot f - Y \circ X \cdot f - (\nabla_X Y) \cdot f + (\nabla_Y X) \cdot f \\ &= [X, Y] \cdot f - (\nabla_X Y - \nabla_Y X) \cdot f \end{aligned}$$

*Thus the connection is torsion free if and only if the Hessian of any function is symmetric.*

## 2.4 Frobenius theorem

### 2.4.1 Distributions

The integral curves can be viewed as the integration of vector fields, which can be compared to the indefinite integral of functions on Euclidean spaces. That is, we integrate the derivative (tangent vector field) to get a family of anti-derivatives (integral curves), which is also equivalent to the one-parameter group of diffeomorphisms.

Here we want to generalized the integration to higher dimensions. Namely, we want to consider the integration of a collection of vector fields.

**Definition 2.4.1**

- *A rank $k$* distribution *$\mathcal{D}$ of a manifold $M$ is a map $\mathcal{D}$ which assigns a $k$-dimensional subspace $\mathcal{D}(p) \subset T_pM$ at every point $p \in M$.*
- *If for all $p \in M$, there exists a neighborhood $U$ and local smooth vector fields $X_1, \cdots, X_k$ such that $\mathcal{D}(q)$ is spanned by $X_1(q), \cdots, X_k(q), \forall q \in U$, then we say $\mathcal{D}$ is* smooth.
- *For a vector field $X \in \Gamma(TM)$, if $X(p) \in \mathcal{D}(p)$ for all $p \in M$, then we denote $X \in \mathcal{D}$. If for any $X, Y \in \mathcal{D}$, we have $[X, Y] \in \mathcal{D}$, then $\mathcal{D}$ is called* involutive *or* integrable.
- *Let $N \subset M$ be an imbedded submanifold, if for all $p \in N$, we have*

$$T_pN = \mathcal{D}(p),$$

  *then $N$ is called an* integral manifold *of $\mathcal{D}$.*

We will always assume the smoothness of distributions.

Obviously, $k$ linear independent vector fields generates a rank $k$ distribution. Note that for a smooth rank $k$ smooth distribution $\mathcal{D}$, in general we can not find global smooth vector fields $X_1, \cdots, X_k$ which generate $\mathcal{D}$. An easy example is the tangent bundle of $S^2$. Thus a rank 1 distribution $\mathcal{D}$ is not necessarily generated by a vector field $X$. But when it is, the integral curves of $X$ are integral manifolds of $\mathcal{D}$.

### 2.4.2 Frobenius theorem

It is easy to see that if for all $p \in M$, there exists an integral manifold of $\mathcal{D}$ through $p$, then $\mathcal{D}$ is involutive. Conversely, we have the following Frobenius theorem.

**Theorem 2.4.2 (Frobenius Theorem)** *Suppose $\mathcal{D}$ is a rank $k$ smooth distribution on a manifold $M$. Then $\mathcal{D}$ is involutive if and only if there is a local integral manifold of $\mathcal{D}$ at each point.*

It is easy to see that if $\mathcal{D}$ is generated by integral manifolds, then it is involutive. To prove the inverse statement, we first need two lemmas.

**Lemma 2.4.3** *Suppose $\mathcal{D}$ is an involutive distribution of rank $k$, then for any $p \in M$, there exists a neighborhood $U$ and $k$ vector fields $X_1, \cdots, X_k \in \mathcal{D}$ which span $\mathcal{D}$ and*

$$[X_a, X_b] = 0, \forall 1 \leq a, b \leq k.$$

*Proof.* In a local chart around $p$, choose $Y_a = Y_a^i \partial_i, a = 1, \cdots, k$ which spans $\mathcal{D}$. Since $Y_1, \cdots, Y_k$ are linear independent, the matrix $(Y_a^i)_{k \times m}$ has rank $k$. Without loss of generality, we may assume its submatrix $(Y_a^b)_{k \times k}$ is invertible with inverse $(A_a^b)_{k \times k}$. Now let

$$X_b = A_b^a Y_a = \partial_b + A_b^a Y_a^\alpha \partial_\alpha,$$

where $\alpha$ is an index from $k+1$ to $m$. We claim that $[X_a, X_b] = 0, \forall 1 \leq a, b \leq k$.

A simple computation shows

$$[X_a, X_b] = C_{ab}^\alpha \partial_\alpha, \forall 1 \leq a, b \leq k,$$

for some $C_{ab}^\alpha$. On the other hand, since $\mathcal{D}$ is involutive and the linear independent $X_1, \cdots, X_k \in \mathcal{D}$ also forms a basis of $\mathcal{D}$, for any $1 \leq a, b \leq k$, we can find coefficients $D_{ab}^c, 1 \leq c \leq k$ such that

$$[X_a, X_b] = D_{ab}^c X_c = D_{ab}^c (\partial_c + A_c^a X_a^\alpha \partial_\alpha).$$

Comparing the above two identities, we find $D_{ab}^c = 0$ and the lemma is proved. $\square$

**Lemma 2.4.4** *Suppose $\varphi_t$ and $\sigma_s$ are one-parameter groups of diffeomorphisms which are generated by vector fields $X, Y$ respectively. Then $[X, Y] = 0$ if and only if*

$$\varphi_t \circ \sigma_s = \sigma_s \circ \varphi_t.$$

*Proof.* Firstly, if the identity holds, then by taking derivatives on both side with respect to $t = 0$ and $s = 0$, we find $[X, Y] = 0$.

Conversely, suppose $[X, Y] = 0$, we want to prove the identity, which is equivalent to

$$\sigma_s = (\varphi_t)^{-1} \circ \sigma_s \circ \varphi_t = \varphi_{-t} \circ \sigma_s \circ \varphi_t.$$

Note that, for any fixed $t$, both sides of the above identity are one-parameter groups of diffeomorphisms (w.r.t. parameter $s$). Thus to prove the identity, we only need to show their corresponding vector fields coincide, i.e.

$$Y = (\varphi_{-t})_* \circ Y \circ \varphi_t.$$

Denote $Y_t = (\varphi_{-t})_* \circ Y \circ \varphi_t$. By assumption, and definition of Lie derivative,

$$[X, Y] = \mathcal{L}_X Y = \frac{d}{dt}\Big|_{t=0} Y_t = 0.$$

By the group property of $\varphi_t$, we have

$$Y_{\tau+t} = (\varphi_{-\tau})_* \circ (\varphi_{-t})_* \circ Y \circ \varphi_t \circ \varphi_\tau = (\varphi_{-\tau})_* \circ Y_t \circ \varphi_\tau.$$

Thus we compute for arbitrary $\tau$,

$$\begin{aligned} \frac{d}{dt}\Big|_{t=\tau} Y_t &= \frac{d}{dt}\Big|_{t=0} Y_{\tau+t} \\ &= \frac{d}{dt}\Big|_{t=0} (\varphi_{-\tau})_* \circ Y_t \circ \varphi_\tau \\ &= (\varphi_{-\tau})_* \circ \mathcal{L}_X Y \circ \varphi_\tau = 0 \end{aligned}$$

Therefore, $Y_t = Y_0 = Y$ is constant and the proof is finished. □

Now we are ready to prove the Frobenius Theorem.

*Proof.* [Proof of Theorem 2.4.2] Suppose $\mathcal{D}$ is an involutive distribution of rank $k$, we need to show that near each point $p \in M$, there is an integral submanifold of $\mathcal{D}$ .

Let $(U, \phi; x^i)$ be a local chart around $p$ such that $\phi(p) = 0$. By Lemma 2.4.3, there exists local vector fields $X_1, \cdots, X_k \in \mathcal{D}$ such that $[X_a, X_b] = 0$ and

$$\{X_1, \cdots, X_k, \partial_{k+1}, \cdots, \partial_m\}$$

forms a local basis of $TU$. Let $\varphi^a$ be the one-parameter group of diffeomorphisms generated by $X_a, 1 \leq a \leq k$. Then we define a map $\Phi : I^k \times \Omega \to U$ by

$$\Phi(t_1, \cdots, t_k, x^{k+1}, \cdots, x^m) = \varphi^1_{t_1} \circ \cdots \circ \varphi^k_{t_k}(q),$$

where $q = \phi^{-1}(0, \cdots, 0, x^{k+1}, \cdots, x^n)$ and $\Omega \subset \mathbb{R}^{m-k}$ is an open set.

Applying Lemma 2.4.4, we can change the order of $\varphi^a$'s arbitrarily in the definition of $\Phi$. Consequently, for $1 \leq a \leq k$, we have

$$\frac{\partial}{\partial t_a}\Phi(p) = X_a(p).$$

It follows

$$\Phi_*|_p \left(\frac{\partial}{\partial t_1}, \cdots, \frac{\partial}{\partial t_k}, \frac{\partial}{\partial x^{k+1}}, \cdots, \frac{\partial}{\partial x^m}\right) = \left(X_1, \cdots, X_k, \frac{\partial}{\partial x^{k+1}}, \cdots, \frac{\partial}{\partial x^m}\right)$$

and $\Phi_*|_p$ is non-degenerate.

Therefore, by the IFT, there is a small neighborhood $V \subset U$ on which $\Phi$ is a diffeomorphism. Namely, $(t_1, \cdots, t_k, x^{k+1}, \cdots, x^m)$ defines a local coordinate chart on $V$. Obviously, the submanifold

$$N_p = \{q \in V | x^\alpha(q) = 0, k+1 \leq \alpha \leq m\}$$

gives a local integral manifold of $\mathcal{D}$. $\square$

**Remark 2.4.5** *There is a dual version of Frobenius theorem, cf. Section 3.7 of [2] or Theorem 3.2, page 198 of [4]. For a rank $k$ distribution $\mathcal{D}$, we can find $l := m-k$ differential local 1-forms $\omega_1, \cdots, \omega_l \in \Gamma(T^*U)$ annihilating $\mathcal{D}$, i.e.*

$$X \in \mathcal{D} \Longleftrightarrow \omega_1(X) = \cdots = \omega_l(X) = 0.$$

*Then the requirement of involutive distribution in the Frobenius theorem can be replaced by* *the following equivalent condition:*

$$d\omega_i \wedge \omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_l = 0, \forall 1 \leq i \leq l.$$

### 2.4.3 Foliation

The above Frobenius theorem only gives local integral submanifold of an involutive distribution. Next, we can extend the local submanifolds to maximal ones (or glue them together), just like in the case of integral curves, which yields a global structure known as a foliation.

**Definition 2.4.6** *A $k$-dimensional foliation $\mathcal{F}$ on an $m$-dimensional manifold $M$ is a decomposition of $M$ into a union of disjoint connected submanifolds $\{\mathcal{F}_\alpha\}_{\alpha \in I}$, called the leaves of the foliation, with the following property:*

> *For every point in $M$, there is a local chart $(U, \phi; x^i)$, such that for each leaf $\mathcal{F}_\alpha$, the components of $U \cap \mathcal{F}_\alpha$ is given by*
>
> $$\{q \in U \cap \mathcal{F}_\alpha | x^s(q) = c^s \textit{ is constant}, k+1 \leq s \leq m\}.$$

As a standard model, imagine the decomposition of the Euclidean space $\mathbb{R}^m$ into the cosets $x + \mathbb{R}^k$ of the standardly embedded subspace $R^k$. Another explicit example is the foliation of a 2-torus by a family of cross circles. However, the decomposition of a sphere by level sets of $z$-axis is not foliation because of the the singularities at the poles. In general, the global structure of a foliation can be unexpectedly complicated.

**Theorem 2.4.7** *Suppose $\mathcal{D}$ is an involutive distribution of rank $k$ on $M$, then the set of maximal integral submanifolds of $\mathcal{D}$ gives a $k$-dimensional foliation of $M$.*

**Exercise 2.4.8** *Suppose $G$ is a Lie group, and $H$ is a Lie subgroup. Show that $G$ is foliated by cosets of $H$.*

# Chapter 3

# Integration of Forms

## 3.1 Tensor algebra

### 3.1.1 Multi-linear functions

Let $V_1, \cdots, V_r$ be $r$ vector spaces, a *multi-linear function* is a map $\Phi : V_1 \times \cdots \times V_r \to \mathbb{R}^1$ satisfying

$$\Phi(v_1, \cdots, \alpha v_i + \beta v_i', \cdots, v_r) = \alpha\Phi(v_1, \cdots, v_i, \cdots, v_r) + \beta\Phi(v_1, \cdots, v_i', \cdots, v_r)$$

for all $i = 1, 2, \cdots, r$, $v_i, v_i' \in V_i$ and $\alpha, \beta \in \mathbb{R}^1$. We denote the space of multi-linear functions on $V_1 \times \cdots \times V_r$ by $\mathscr{L}(V_1, \cdots, V_r; \mathbb{R}^1)$.

For example, a linear functions of $V$ is just an element in the dual space $V^*$. A *bi-linear function* $B \in \mathscr{L}(V, W; \mathbb{R}^1)$ is a linear map on $V \times W$ which satisfies

$$B(\alpha v_1 + \beta v_2, w) = \alpha B(v_1, w) + \beta B(v_2, w),$$
$$B(v, \alpha w_1 + \beta w_2) = \alpha B(v, w_1) + \beta B(v, w_2).$$

### 3.1.2 Tensor product

Now we define an operation on multi-linear functions which acts like products. We start with linear functions. Let $\xi \in V^*, \eta \in W^*$ be two linear functions, we define the *tensor product* $\xi \otimes \eta$ to be a bi-linear function given by

$$\xi \otimes \eta(v, w) = \xi(v)\eta(w), \ \forall v \in V, w \in W.$$

It is easy to verify that $\xi \otimes \eta \in \mathscr{L}(V, W; \mathbb{R}^1)$. Then we can go on to define multi-linear functions by using tensor product on multiple linear functions. For example, for $\xi_i \in V_i^*, i = 1, \cdots, r$, we have $\xi_1 \otimes \cdots \otimes \xi_r \in \mathscr{L}(V_1, \cdots, V_r; \mathbb{R}^1)$.

It is easy to see that the tensor product, as an operation, is bi-linear on its components and is associative. However, the tensor product is *not* commutative.

Recall that for a finite dimensional vector space, we have $V = (V^*)^*$, which means we can regard $V$ as the space of linear functions of $V^*$. Thus the tensor product can be as well defined for any vector space.

### 3.1.3 Tensor spaces

All possible tensor products of elements in $V^*$ and $W^*$ naturally span a tensor product space, denoted by

$$V^* \otimes W^* := \operatorname{span}\{\xi \otimes \eta | \xi \in V^*, \eta \in W^*\}.$$

If $\{\xi_i\}_{i=1}^m$ is a basis of $V^*$ and $\{\eta_\alpha\}_{\alpha=1}^n$ is a basis of $W^*$, then an induced basis of $V^* \otimes W^*$ is $\{\xi_i \otimes \eta_\alpha\}_{1\le i\le m, 1\le \alpha\le n}$. Hence the dimension of $V^* \otimes W^*$ is $mn$. Note that a typical element in $V^* \otimes W^*$ is not always decomposable, that is, it is not in a tensor product form $\xi \otimes \eta$. For example, for a two dimensional space $V^*$ with basis $\xi_1, \xi_2$, the bi-linear function $B = \xi_1 \otimes \xi_1 + \xi_2 \otimes \xi_2$ is not decomposable.

Now for a vector space $V$ and non-negative integers $r, s$, we define the $(s, r)$-tensor space of $V$ to be

$$T^r_s V := \underbrace{V \otimes \cdots \otimes V}_{s} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{r}.$$

If $\{e_i\}_{i=1}^m$ is a basis of $V$ and $\{e_i^*\}_{i=1}^m$ is its dual basis of $V^*$, then a basis of $T^r_s V$ is given by

$$e_{i_1} \otimes \cdots \otimes e_{i_s} \otimes e^*_{j_1} \otimes \cdots \otimes e^*_{j_r}, 1 \le i_1, \cdots, i_s, j_1, \cdots, j_r \le m.$$

Thus it is a linear space of dimension $n^{r+s}$ and an element $L \in T^r_s V$ can be expressed by

$$L = a^{i_1\cdots i_s}_{j_1\cdots j_r} e_{i_1} \otimes \cdots \otimes e_{i_s} \otimes e^*_{j_1} \otimes \cdots \otimes e^*_{j_r}.$$

By convention, we set $T^0_0 V = \mathbb{R}^1$. The total tensor space of $V$ is

$$\mathscr{T} V = \oplus_{r,s=0}^\infty T^r_s V.$$

It is easy to see that the tensor product space $V^* \otimes W^*$ is exactly the space of bi-linear functions $\mathscr{L}(V, W; \mathbb{R}^1)$. Similarly, using the tensor product, one can identify the space of linear maps $\mathcal{L}(V; W)$ with the tensor space $V^* \otimes W$. Namely, if for a basis $\{e_i\}_{i=1}^m \subset V$ and $\{f_\alpha\}_{\alpha=1}^n \subset W$, a linear map $L \in \mathcal{L}(V; W)$ satisfies

$$L(\xi_i) = a_i^\alpha f_\alpha, \quad i = 1, \cdots, m,$$

then $L$ can be identified with $a_i^\alpha e_i^* \otimes f_\alpha \in V^* \otimes W$, where $e_i^*$ is the dual basis of $e_i$.

**Example 3.1.1** *For two manifolds $M, N$ and a smooth map $F : M \to N$, the tangent map of $F$ at $x \in M$ is a linear map*

$$dF(x) = (F_*)|_x : T_xM \to T_{F(x)}N.$$

*In local coordinates, since*

$$(F_*)|_x \left(\frac{\partial}{\partial x^i}\right) = \frac{\partial F^\alpha}{\partial x^i}(x)\frac{\partial}{\partial y^\alpha},$$

*it can be written as*

$$dF(x) = \frac{\partial F^\alpha}{\partial x^i}(x)dx^i \otimes \frac{\partial}{\partial y^\alpha} \in T_x^*M \otimes T_{F(x)}N.$$

### 3.1.4 Tensor operations

A tensor $\Phi \in T^rV := T_0^rV$ is called *symmetric* if $\forall 1 \le i, j \le r$,

$$\Phi(v_1, \cdots, v_i, \cdots, v_j, \cdots, v_r) = \Phi(v_1, \cdots, v_j, \cdots, v_i, \cdots, v_r),$$

and is called *anti-symmetric* (or *alternating*) if for all $i, j$,

$$\Phi(v_1, \cdots, v_i, \cdots, v_j, \cdots, v_r) = -\Phi(v_1, \cdots, v_j, \cdots, v_i, \cdots, v_r).$$

We denote the space of symmetric and anti-symmetric multi-linear functions by $S^r(V)$ and $A^r(V)$, respectively.

Let $P_r$ be the space of permutations of $r$ numbers. Recall that a permutation $\sigma \in P_r$ is a product of transpositions that only interchange two numbers. The sign of a permutation $\operatorname{sgn}\sigma = (-1)^k$, where $k$ is the number of transpositions, is well-defined. A permutation $\sigma \in P_r$ can act on $\Phi \in T^rV$ by letting

$$(\sigma\Phi)(v_1, \cdots, v_r) = \Phi(v_{\sigma(1)}, \cdots, v_{\sigma(r)}).$$

The *symmetrizing mapping* is the projection $\mathscr{S} : T^rV \to S^r(V)$ given by

$$\mathscr{S}(\Phi) = \frac{1}{r!}\sum_{\sigma \in P_r} \sigma\Phi.$$

Similarly, the *anti-symmetrizing mapping* is a projection $\mathscr{A} : T^r V \to A^r(V)$ given by

$$\mathscr{A}(\Phi) = \frac{1}{r!} \sum_{\sigma \in P_r} \operatorname{sgn} \sigma \cdot \sigma \Phi.$$

It is easy to see that both mappings are linear. Note that the constant is needed to ensure that $\mathscr{S}$ and $\mathscr{A}$ are indeed projections, i.e. $\mathscr{S}^2 = \mathscr{S}$ and $\mathscr{A}^2 = \mathscr{A}$.

There is another very useful operation for tensors, namely, the contraction mapping. It is defined as a linear map $C_{qp} : T^r_s V \to T^{r-1}_{s-1} V$ for $r, s \geq 1$ by letting

$$\begin{aligned} &C_{qp}(a^{i_1 \cdots i_s}_{j_1 \cdots j_r} e_{i_1} \otimes \cdots \otimes e_{i_s} \otimes \delta^{j_1} \otimes \cdots \otimes \delta^{j_r}) \\ &\quad = (\delta^{j_p}, e_{i_q}) a^{i_1 \cdots i_s}_{j_1 \cdots j_r} e_{i_1} \otimes \cdots \otimes \hat{e}_{i_q} \otimes \cdots \otimes e_{i_s} \otimes \delta^{j_1} \otimes \cdots \otimes \hat{\delta}^{j_p} \otimes \cdots \otimes \delta^{j_r}. \end{aligned}$$

**Example 3.1.2** *A linear map from a vector space $V$ to itself can be viewed as a tensor $\Phi \in T^1_1 V$. Fixing a basis $\{e_i\}_{i=1}^n$ of $V$ and its dual basis $\{\delta^i\}_{i=1}^n$ of $V^*$, we can write*

$$\Phi = a^i_j e_i \otimes \delta^j \in T^1_1 V.$$

*Or equivalently, we represent $\Phi \in \mathcal{L}(V; V)$ by a matrix $A := (a^i_j)$. Then the symmetrizing mapping takes $A$ to $\mathscr{S}(A) = \frac{1}{2}(A + A^t)$, while the anti-symmetrizing mapping maps $A$ to $\mathscr{A}(A) = \frac{1}{2}(A - A^t)$. Note that $A = \mathscr{S}(A) + \mathscr{A}(A)$. Moreover, the contraction map gives the trace* $\operatorname{tr} A = C_{11}(A) = a^i_i$.

### 3.1.5 Exterior product

The anti-symmetric multi-linear functions is particularly important in differential geometry. There is natural product operation in the category of anti-symmetric multi-linear functions.

Suppose $\Phi \in A^r(V)$ and $\Psi \in A^s(V)$, the *exterior product* or *wedge product* is a map

$$\wedge : A^r(V) \times A^s(V) \to A^{r+s}(V)$$

defined by

$$\Phi \wedge \Psi = \frac{(r+s)!}{r!s!} \mathscr{A}(\Phi \otimes \Psi).$$

The exterior product is obviously bi-linear on its components. It is also associative in view of the following lemma. The constant in the definition of wedge product is chosen such that the identity in the lemma is valid.

**Lemma 3.1.3** *For $\Phi_i \in A^{r_i}(V)$, $i = 1, \cdots, p$,*

$$\Phi_1 \wedge \cdots \wedge \Phi_p = \frac{(r_1 + \cdots + r_p)!}{r_1! \cdots r_p!} \mathscr{A}(\Phi_1 \otimes \cdots \otimes \Phi_p).$$

*In particular, for $\xi_i \in A^1(V) = V^*$, $i = 1, \cdots, p$, we have*

$$\xi_1 \wedge \cdots \wedge \xi_p = p! \mathscr{A}(\xi_1 \otimes \cdots \otimes \xi_p).$$

The most useful property is that it is skew-symmetric in the following sense.

**Lemma 3.1.4** $\Phi \wedge \Psi = (-1)^{rs} \Psi \wedge \Phi$.

*Proof.* By linearity, it suffices to show the equality for $\Phi = \xi^{i_1} \wedge \cdots \wedge \xi^{i_r}$ and $\Psi = \eta^{j_1} \wedge \cdots \wedge \eta^{j_s}$. Using the fact that $\xi \wedge \eta = -\eta \wedge \xi$ for $\xi, \eta \in V^*$, we have

$$\begin{aligned}
\Phi \wedge \Psi &= \xi^{i_1} \wedge \cdots \wedge \xi^{i_r} \wedge \eta^{j_1} \wedge \cdots \wedge \eta^{j_s} \\
&= -\xi^{i_1} \wedge \cdots \wedge \eta^{j_1} \wedge \xi^{i_r} \wedge \cdots \wedge \eta^{j_s} \\
&= (-1)^r \eta^{j_1} \wedge \xi^{i_1} \wedge \cdots \wedge \xi^{i_r} \wedge \eta^{j_2} \wedge \cdots \wedge \eta^{j_s} \\
&= ((-1)^r)^s \eta^{j_1} \wedge \cdots \wedge \eta^{j_s} \wedge \xi^{i_1} \wedge \cdots \wedge \xi^{i_r} \\
&= (-1)^{rs} \Psi \wedge \Phi.
\end{aligned}$$

$\square$

By a same process as in the definition of tensor product space, we define the exterior product space

$$\wedge^r V^* = \underbrace{V^* \wedge \cdots \wedge V^*}_{r} = \text{span}\{\xi^1 \wedge \cdots \wedge \xi^r | \xi^i \in V^*, 1 \le i \le r\}.$$

It is easy to see that $A^r(V) = \wedge^r V^*$. Suppose $\{\delta^i\}_{i=1}^n$ is a basis of $V^*$, then one checks that a basis of $A^r(V)$ is

$$\delta^{i_1} \wedge \cdots \wedge \delta^{i_r}, \quad 1 \le i_1 < \cdots < i_r \le n.$$

Hence the dimension of $A^r(V)$ is $C_n^r$. In particular, $A^n(V)$ has dimension 1 and is simply the linear space generated by $\delta^1 \wedge \cdots \wedge \delta^n$.

**Remark 3.1.5** *Similarly, one can also define a product on symmetric multi-linear functions* $\cdot : S^r(V) \times S^s(V) \to S^{r+s}(V)$ *by letting*

$$\Phi \cdot \Psi = \frac{(r+s)!}{r!s!} \mathscr{S}(\Phi \otimes \Psi).$$

*For example, given two vectors* $\omega^1, \omega^2 \in A^1(V) = V^*$, *we have*

$$\omega^1 \wedge \omega^2 = \omega^1 \otimes \omega^2 - \omega^2 \otimes \omega^1,$$
$$\omega^1 \cdot \omega^2 = \omega^1 \otimes \omega^2 + \omega^2 \otimes \omega^1.$$

Here are more useful properties for exterior products.

**Proposition 3.1.6** *1)* $v_1, \cdots, v_r \in V^*$ *is linearly dependent if and only if*

$$v_1 \wedge \cdots \wedge v_r = 0$$

*2) If* $A = (a_i^j)$ *is a* $r \times r$ *matrix and* $w_i = a_i^j v_j, 1 \le i \le r$, *then*

$$w_1 \wedge \cdots \wedge w_r = \det A \cdot v_1 \wedge \cdots \wedge v_r.$$

*3) Let* $F : V \to W$ *be a linear map, then for* $\Phi \in A^r(V)$ *and* $\Psi \in A^s(V)$,

$$F^*(\Phi \wedge \Psi) = F^*\Phi \wedge F^*\Psi.$$

*Here the pull-back map* $F^*$ *is defined by*

$$F^*\Phi(v_1, \cdots, v_r) := \Phi(Fv_1, \cdots, Fv_r), \quad \forall v_1, \cdots, v_r \in V.$$

**Remark 3.1.7** *Item 2 of the above proposition is related to the formula of change of variables in multi-integrals. That is, the determinant automatically appear in wedge products, which is the key in defining integration on manifolds in a way that is independent on choice of coordinates.*

### 3.1.6 Tensor fields and forms

For a manifold $M$, consider the tangent space $T_pM$ at a point $p \in M$. Define the tensor space

$$T_p^{(s,r)}M := T_s^r(T_pM) = \underbrace{T_pM \otimes \cdots \otimes T_pM}_{s} \otimes \underbrace{T_p^*M \otimes \cdots T_p^*M}_{r}.$$

Then the $(s,r)$-tensor bundle is the total space

$$T^{(s,r)}M := \cup_{p\in M} T_p^{(s,r)}M,$$

which is itself a manifold in a natural way. Indeed, it is a vector bundle of rank $n^{r+s}$ on $M$ and in local coordinates, has a natural basis

$$\frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_s}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_r}.$$

In particular, $T^{(1,0)}M$ is the tangent bundle and $T^{(0,1)}M$ is the cotangent bundle of $M$.

A section of $T^{(s,r)}M$ is called a $(s,r)$-type tensor field, or simply $(s,r)$-tensor, on $M$. As usual, we always assume the section, hence the tensor field is smooth.

**Exercise 3.1.8** *Deduce the transformation formula of coefficients of a $(r,s)$-tensor under the change of local coordinates.*

In a similar manner, we consider the space of anti-symmetric multi-linear functions $A^r(T_pM)$ and construct the $r$-th exterior bundle on $M$ given by

$$\wedge^r T^*M = \cup_{p\in M} \wedge^r T_p^*M = \cup_{p\in M} A^r(T_pM).$$

A section of $\wedge^r T^*M$ is called a $r$-form, and the space of all $r$-forms is denoted by $\Lambda^r(M)$. By convention, 0-forms are just smooth functions $\Lambda^0(M) = C^\infty(M)$ and 1-forms are cotangent vector fields $\Lambda^1(M) = \Gamma(T^*M)$. The total space of all forms is denoted by

$$\Lambda(M) := \oplus_{r=0}^m \Lambda^r(M).$$

As the point in the underlying manifold varies, a $(s,r)$-tensor $\tau \in \Gamma(T^{(s,r)}M)$ defines a linear map

$$\tau : \underbrace{\Lambda^1(M) \times \cdots \times \Lambda^1(M)}_{s} \times \underbrace{\mathfrak{X}(M) \times \cdots \mathfrak{X}(M)}_{r} \to C^\infty(M).$$

In fact, the map is $C^\infty$-linear, i.e. for $\omega^1, \cdots, \omega^s \in \Lambda^1(M)$, $X_1, \cdots, X_r \in \mathfrak{X}(M)$, and any $f \in C^\infty(M)$ and $1 \le i \le s, 1 \le j \le r$, we have

$$\begin{aligned} \tau(\omega^1, \cdots, f\omega^i, \cdots, \omega^s, X_1, \cdots, X_r) &= \tau(\omega^1, \cdots, \omega^s, X_1, \cdots, fX_j, \cdots, X_r) \\ &= f \cdot \tau(\omega^1, \cdots, \omega^s, X_1, \cdots, X_r). \end{aligned}$$

Similarly, a $r$-form $\omega \in \Lambda^r(M)$ defines an anti-symmetric $C^\infty$-linear map

$$\omega : \underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{r} \to C^\infty(M).$$

**Exercise 3.1.9** *Show that any $C^\infty(M)$-linear map*

$$\tau : \underbrace{\Lambda^1(M) \times \cdots \times \Lambda^1(M)}_{s} \times \underbrace{\mathfrak{X}(M) \times \cdots \mathfrak{X}(M)}_{r} \to C^\infty(M)$$

*can be identified with a $(s, r)$-tensor field.*

**Remark 3.1.10** *More generally, we can define the tensor product bundle $E \otimes F$ on $M$ of two vector bundles $E$ and $F$ on $M$. If $E$ and $F$ has rank $r, s$ with fiber space $V$ and $W$ respectively, then $E \otimes F$ has rank $rs$ with fiber space $V \otimes W$. More precisely, if $\{v_i\}_{i=1}^r$ is a local frame of $E$ and $\{w_\alpha\}_{\alpha=1}^s$ is a local frame of $F$, then $\{v_i \otimes w_\alpha\}$ is a local frame of $E \otimes F$. For two overlapped open sets $U, V \subset M$, suppose the transition maps of $E$ and $F$ are $g_E : U \cap V \to GL(r)$ and $g_F : U \cap V \to GL(s)$ respectively. Then the corresponding transition map of $E \otimes F$ is simply*

$$\tilde{g} : U \cap V \to GL(r+s), \quad \tilde{g}(x) = g_E(x) \otimes g_F(x).$$

*such that*

$$\tilde{g}(v_i \otimes w_\alpha) = g_E(v_i) \otimes g_F(w_\alpha).$$

*In particular, the tensor product and wedge product operation also applies to a vector bundle $E$ and its dual bundle $E^*$.*

## 3.2 Exterior derivative

### 3.2.1 Exterior derivatives

The exterior derivative extends the differential of a function to a globally defined operator on forms on manifolds.

First we define the exterior derivative locally. Suppose $U \subset \mathbb{R}^n$ is an open set with coordinates $x^i, i = 1, \cdots, n$. Consider the space of $p$-forms $\Lambda^p(U)$, which is spanned by the basis

$$dx^I := dx^{i_1} \wedge \cdots \wedge dx^{i_p}, 1 \leq i_1 < \cdots < i_p \leq n.$$

Then a local $p$-form on $U$ has the form

$$\omega = a_I dx^I = a_{i_1 \cdots i_p}(x) dx^{i_1} \wedge \cdots \wedge dx^{i_p}.$$

Then the local *exterior derivative* on $U$ is defined by

$$d_U : \Lambda^p(U) \to \Lambda^{p+1}(U), \quad d_U \omega = da_I \wedge dx^I,$$

where $da_I$ is the usual differential of function $a_I$. In particular, for a smooth function $f \in C^\infty(U) = \Lambda^0(U)$, $d_U f = df$ is the full differential of $f$.

**Proposition 3.2.1** *The local exterior derivative $d_U : \Lambda(U) \to \Lambda(U)$ satisfies*

*1) $d_U$ is linear*

*2) $d_U^2 = d_U \circ d_U = 0$.*

*3) $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^p \omega \wedge d\eta$ if $\omega \in \Lambda^p(U)$.*

*Proof.* The linearality is obvious. The second property follows from the fact that $d_U^2 f = 0$ for any smooth function $f$ on $U$. For the third property, it suffices to check for simples forms where $\omega = f dx^I$ and $\eta = g dx^J$. In this simple case, we have

$$\begin{aligned} d(\omega \wedge \eta) &= d(fg dx^I \wedge dx^J) = d(fg) \wedge dx^I \wedge dx^J \\ &= df \wedge dx^I \wedge g dx^J + (-1)^p f dx^I \wedge dg \wedge dx^J \\ &= d\omega \wedge \eta + (-1)^p \omega \wedge d\eta. \end{aligned}$$

□

Now for a differentiable manifold $M$, we have a local exterior derivative on every local chart as above. If the definition agrees on overlapped areas of different charts, then we can patch them together to give a globally defined operator, which is independent of the choice of coordinates. Suppose the restriction of a $p$-form $\omega \in \Lambda^p(M)$ on two overlapped charts $(U; x)$ and $(V; y)$ are given by

$$\omega|_U = a_I(x) dx^I, \quad \omega|_V = b_J(y) dy^J.$$

By definition, we have

$$d_U \omega|_U = d_x a_I \wedge dx^I, \quad d_V \omega|_V = d_y b_J \wedge dy^J.$$

We need to show that the above two forms coincide on the overlapped area $U \cap V$.

To see this, consider $y = y(x)$ and $b_J(y) = b_J(y(x))$ as functions of $x$ on $U \cap V$. Then by Proposition 3.2.1, we have

$$\begin{aligned} d_U(a_I dx^I) &= d_U(b_J dy^J) \\ &= d_x b_J \wedge dy^J + b_J d_U \circ d_U y^J \\ &= d_y b_J \wedge dy^J = d_V(b_J dy^J). \end{aligned}$$

Here the equality $d_x b_J = d_y b_J$ is due to the invariance of 1st-order differential of a function.

Therefore, we get a globally defined exterior derivative $d : \Lambda(M) \to \Lambda(M)$ on the manifold $M$. Namely, for a $p$-form $\omega \in \Lambda^p(M)$, we require

$$(d\omega)|_U = d_U \omega_U.$$

Alternatively, we can find a partition of unity $\{\phi_k\}$ subordinating to an atlas $\{U_k\}$ and write $\omega = \sum \phi_k \omega$. Then $d$ is given by

$$d\omega = \sum d_{U_k}(\phi_k \omega).$$

Obviously, the global operator $d$ also satisfies Proposition 3.2.1. Moreover, for a smooth function $f \in \Lambda^0(M)$, $df \in \Lambda^1(M)$ is exactly the tangent map of $f$. Namely, for any tangent vector field $X \in \mathfrak{X}(M)$, we have

$$df(X) = X \cdot f.$$

In fact, the above identity and those properties in Proposition 3.2.1 uniquely determines the exterior derivative $d$ on a given manifold (exercise).

**Proposition 3.2.2** *Suppose $F : M \to N$ is a smooth map and $\omega \in \Lambda^p(N)$, then*

$$F^*(d\omega) = d(F^*\omega).$$

*Proof.* We first show the proposition holds for smooth functions $f \in \Lambda^0(N)$, i.e.

$$F^*(df) = d(F^*f) = d(f \circ F).$$

This follows from the identity that for $\forall X \in \mathfrak{X}(M)$,

$$(F^*(df), X) = (df, F_*(X)) = F_*(X) \cdot f = X \cdot (f \circ F).$$

The general case follows by applying the above identity in local charts. Again note that since $F^*$ and $d$ are both linear operators, it suffices to prove the proposition for simple forms. The details are left to the readers. $\square$

**Example 3.2.3** *The exterior operator $d$ can be regarded as a generalization of differential operators (grad, curl and div) in $\mathbb{R}^3$.*

### 3.2.2 Invariant formula

**Theorem 3.2.4** *Suppose $\omega \in \Lambda^p(M)$ is a $p$-form and $V_1, \cdots, V_{p+1} \in \mathfrak{X}(M)$ are vector fields, then*

$$\begin{aligned} d\omega(V_1, \cdots, V_{p+1}) = &\sum_{i=1}^{p+1} (-1)^{i-1} V_i\big(\omega(V_1, \cdots, \hat{V_i}, \cdots, V_{p+1})\big) \\ &+ \sum_{1 \le i < j \le p+1} (-1)^{i+j} \omega([V_i, V_j], V_1, \cdots, \hat{V_i}, \cdots, \hat{V_j}, \cdots, V_{p+1}). \end{aligned}$$

*Proof.* First note that the left hand side is $C^\infty$-linear (i.e. tensorial) on $V_i$. One can verify that the right hand side is also $C^\infty$-linear on $V_i$. Thus we only need to prove the theorem for a set of basis $V_i = \partial_{k_i}, i = 1, \cdots, p+1$.

In a local chart, we assume $\omega = a_I dx^I$, then $d\omega = da_I \wedge dx^I$. Then for $V_i = \partial_{k_i}$, we have

$$\begin{aligned} d\omega(V_1, \cdots, V_{p+1}) &= da_I \wedge dx^I(\partial_{k_1}, \cdots, \partial_{k_{p+1}}) \\ &= \sum_{i=1}^{p+1} (-1)^{i-1} \partial_{k_i} a^I dx^I(\partial_{k_1}, \cdots, \hat{\partial_{k_i}}, \cdots, \partial_{k_{p+1}}). \end{aligned}$$

Since $[\partial_i, \partial_j] = 0$, the theorem follows. □

In particular, for a 1-form $\omega \in \Lambda^1(M)$, we have

$$d\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]).$$

### 3.2.3 Lie derivative of forms

First recall the definition of Lie derivative of vector fields. Suppose $\phi_t$ is the 1-parameter group of diffeomorphisms generated by a vector field $X \in \mathfrak{X}(M)$. By using the pull-back $\phi_t^*$ and the push-forward $(\phi_{-1})^*$, the Lie derivative with regard to $X$ can be defined on any tensor field. In particular, we can define the Lie derivative of a $p$-form $\omega \in \Lambda^p(M)$ by

$$\mathcal{L}_X \omega := \frac{d}{dt}\Big|_{t=0} \phi_t^* \omega \in \Lambda^p(M).$$

From the property $\phi_t^*(\eta \wedge \omega) = (\phi_t^* \eta) \wedge (\phi_t^* \omega)$, we get

$$\mathcal{L}_X(\eta \wedge \omega) = (\mathcal{L}_X \eta) \wedge \omega + \eta \wedge (\mathcal{L}_X \omega).$$

Similarly, from $\phi_t^*(d\omega) = d(\phi_t^* \omega)$, we have

$$\mathcal{L}_X(d\omega) = d(\mathcal{L}_X \omega).$$

There is an *interior product* opertator $\iota : \mathscr{X}(M) \times \Lambda^p(M) \to \Lambda^{p-1}(M)$ defined by

$$\iota_X \omega(X^1, \cdots, X^{p-1}) = \omega(X, X^1, \cdots, X^{p-1}).$$

Obviously, $\iota$ is $C^\infty$-linear both in $X$ and $\omega$. In particular, if $\omega \in \Lambda^1(W)$ is a 1-form ,then $\iota_X\omega = \omega(X)$. Moreover, for $\omega \in \Lambda^p(M)$, we have

$$\iota_X(\omega \wedge \eta) = (\iota_X\omega) \wedge \eta + (-1)^p\omega \wedge (\iota_X\eta).$$

**Theorem 3.2.5 (Cartan's identity)** *The Lie derivative of a p-form satisfies*

$$\mathcal{L}_X\omega = d(\iota_X\omega) + \iota_X(d\omega).$$

*Proof.* We first demonstrate an easy prove for 1-forms. Suppose $\omega$ is a 1-form, then for $X, Y \in \mathscr{X}(M)$, we have the Leibnitz rule

$$(\mathcal{L}_X\omega)(Y) = \mathcal{L}_X(\omega(Y)) - \omega(\mathcal{L}_XY).$$

By Theorem 3.2.4, we have

$$d\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]).$$

Therefore,

$$(\mathcal{L}_X\omega)(Y) = Y(\omega(X)) + d\omega(X, Y).$$

For the general case, define

$$T(X, \omega) = \mathcal{L}_X\omega - d(\iota_X\omega) - \iota_X(d\omega).$$

We will show $T(X, \omega) \equiv 0$ for all $X$ and $\omega$. First note that $T$ is tensorial in $\omega$ (actually it is tensorial on both arguments). Indeed, for any $f \in C^\infty(M)$, we compute

$$\begin{aligned}
\mathcal{L}_X(f\omega) &= f\mathcal{L}_X\omega + \mathcal{L}_Xf \cdot \omega, \\
d(\iota_X(f\omega)) &= d(f\iota_X\omega) = df \wedge \iota_X\omega + fd(\iota_X\omega), \\
\iota_Xd(f\omega) &= \iota_X(df \wedge \omega + fd\omega) = df(X) \cdot \omega - df \wedge \iota_X\omega + f\iota_X(d\omega).
\end{aligned}$$

Since $\mathcal{L}_Xf = df(X)$, it follows $T(X, f\omega) = fT(x, \omega)$. Thus we only need to prove $T(X, \omega) =$

0 for a basis $\omega = dx^1 \wedge \cdots \wedge dx^p$. Suppose $X = X^i \partial_i$, then we have $d\omega = 0$ and

$$\begin{aligned} T(X,\omega) &= \mathcal{L}_X(dx^1 \wedge \cdots \wedge dx^p) - d(\iota_X(dx^1 \wedge \cdots \wedge dx^p)) \\ &= \sum_{i=1}^{p} dx^1 \wedge \cdots \wedge \mathcal{L}_X dx^i \wedge \cdots \wedge dx^p - \sum_{i=1}^{p} d((-1)^{i-1} dx^1 \wedge \cdots \iota_X dx^i \cdots \wedge dx^p) \\ &= (-1)^{i-1}(\mathcal{L}_X dx^i - d(\iota_X dx^i)) \wedge dx^1 \wedge \cdots \wedge \hat{dx^i} \wedge \cdots \wedge dx^p. \end{aligned}$$

Finally, note that $\mathcal{L}_X dx^i = d(\mathcal{L}_X x^i) = d(X(x^i)) = dX^i$ and $\iota_X dx^i = dx^i(X) = X^i$ and the proof is finished. □

### 3.2.4 de Rham cohomology

The property $d^2 = 0$ induces a natural cohomology on a differentiable manifold, which is called the *de Rham cohomology*.

Write the exterior derivative on $\Lambda^p M$ as $d_p : \Lambda^p M \to \Lambda^{p+1} M$, then $d_p \circ d_{p-1} = 0$ implies $\mathrm{Im} d_{p-1} \subset \mathrm{Ker}\, d_p$. Therefore we can defined the *p-th de Rham cohomology group* as the quotient space

$$H_d^p(M) = \mathrm{Ker}\, d_p / \mathrm{Im} d_{p-1}.$$

We call elements in $\mathrm{Ker}\, d$ *closed forms* and those in $\mathrm{Im} d$ *exact forms*. Thus an equivalent class $[\omega] \in H_d^p(M)$ of a closed form $\omega \in \mathrm{Ker}\, d_p$ is given by

$$[\omega] = \{\omega + d\eta | \eta \in \Lambda^{p-1}(M)\}.$$

By convention, we define $H_d^0(M) = \mathrm{Ker}\, d_0$ to be the space of constant functions, which is just $\mathbb{R}^1$ for connected manifolds.

**Example 3.2.6** *The de Rham cohomology of the sphere $S^n$.*

## 3.3 Integration of forms

### 3.3.1 Orientation

First let's recall the definition of orientation in basic analysis, which is defined extrinsically. For any point in a embedded surface $\Sigma$ in the Euclidean space $\mathbb{R}^3$, there are two opposite choices of unit normal vectors. An orientation then is a continuous way of assigning a normal vector at each point $\vec{n}: \Sigma \to \mathbb{R}^3$. A surface which supports such a map $\vec{n}$ is called orientable and has two opposite orientations $\pm\vec{n}$.

On the other hand, assigning an normal vector $\vec{n}$ at a point is equivalent to assigning an ordered basis $\{\vec{e_1}, \vec{e_2}\}$ of the tangent space, such that $\{\vec{e_1}, \vec{e_2}, \vec{n}\}$ forms an oriented basis of $\mathbb{R}^3$. For example, if the tangent plane is the $x-y$ plane, then choosing $\partial z$ as the normal direction is equivalent, by the right hand principle, to choosing $\{\partial x, \partial y\}$ as the ordered basis. Then (locally) an orientation is just a continuous choice of the ordered basis.

This motivates an intrinsic method of defining an orientation on a $n$-dimensional manifold $M$. The ordered basis $\{e_1, \cdots, e_n\} \subset T_pM$ at a point $p \in M$ is represented by the wedge product

$$e^1 \wedge \cdots \wedge e^n \in \wedge^n(T_pM)\backslash\{0\}$$

Recall that the top exterior space $\wedge^n(T_pM)$ has dimension one. Thus if we regard different sets of ordered basis, i.e. different elements in $\wedge^n(T_pM)\backslash\{0\}$ as the same when they differ by a positive number, then we get exactly two different choices. Thus an orientation of the manifold is (the equivalent class of) a continuous section $\omega: M \to \Lambda^n(M)$, which dose not vanish anywhere.

**Definition 3.3.1**

- *A smooth manifold is called* orientable *if there exists a non-vanishing $n$-form $\omega \in \Lambda^n(M)$.*
- *An* orientation *of $M$ is a the equivalent class of non-vanishing $n$-forms, where $\omega \sim \omega'$ if $\omega = f\omega'$ for a positive function $f > 0$.*

Obviously, there are two orientations for every connected orientable manifold, and a manifold is orientable if and only if $\wedge^n(TM)$ is a trivial bundle.

There is an equivalent definition by using local charts. Suppose we have an orientation represented by a non-vanishing $\omega \in \Lambda^n(M)$, then in each local chart $(U, x)$, we may arrange the indices of coordinates, such that

$$\omega|_U = f dx^1 \wedge \cdots \wedge dx^n,$$

where $f$ is a positive smooth function on $U$. We call such a chart *compatible* with the orientation $\omega$. Then if two compatible charts overlap, we have

$$\omega|_{U\cap V} = f dx^1 \wedge \cdots \wedge dx^n = g dy^1 \wedge \cdots \wedge dy^n,$$

where $f, g$ are both positive. Recall that under the change of coordinates, if we denote the Jacobi matrix $A = (\partial x^i/\partial y^j)$,

$$dx^1 \wedge \cdots \wedge dx^n = \det A dy^1 \wedge \cdots \wedge dy^n.$$

Therefore, $\det A = g/f$ is positive. Conversely, if we have an atlas where all the Jacobian matrices of transformation maps have positive determinant, then we can easily construct a non-vanishing $n$-form by partition of unity, which gives a compatible orientation. We will call such an atlas an *oriented atlas* and the corresponding $n$-form the induced orientation. This leads to the following.

**Definition 3.3.2** *An orientation is a choice of oriented atlas such that the Jacobian determinant of all transition maps are positive.*

**Example 3.3.3** *Orientable manifolds:*

*1) The orientation of the Euclidean space $\mathbb{R}^n$ is given by the standard form*

$$\Omega^n = dx^1 \wedge \cdots \wedge dx^n.$$

*2) (Hyper-surfaces as level sets) Suppose $f$ is a smooth function on $\mathbb{R}^{n+1}$ and $c \in \mathbb{R}^1$ is a regular value of $f$. Then the level set $M := \{x \in \mathbb{R}^{n+1} | f(x) = c\}$ is an orientable manifold. Typical examples are the $n$-sphere $S^n$.*

*Since $c$ is a regular value, $N = \nabla f$ gives a non-vanishing normal vector field on $M$. Thus we can construct a non-vanishing $n$-form $\omega = \iota_N \Omega^{n+1}$ on $M$, where $\Omega^{n+1}$ is the* *standard form on* $\mathbb{R}^{n+1}$*. More explicitly, we have* $\nabla f = \partial_i f \partial_i$*, thus*

$$\iota_N \Omega = (-1)^{i-1} \partial_i f dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^{n+1}.$$

*Note that the n-form is originally defined in* $\mathbb{R}^{n+1}$*, but can be pulled-back/restricted to* $M$ *via the embedding* $M \hookrightarrow \mathbb{R}^{n+1}$*, which is still non-vanishing.*

*Alternatively, we can chose a n-form* $\omega'$ *by requiring*

$$df \wedge \omega' = \Omega.$$

*More precisely, if in a neighborhood* $U \subset M$*,* $\partial_j f \neq 0$ *for some* $j$*, then*

$$(x_1, \cdots, x_{j-1}, x_{j+1}, \cdots, x_{n+1})$$

*forms a local coordinate system on* $U$*. Then we define* $\omega'$ *on* $U$ *by*

$$\omega' = (-1)^{j-1} \frac{1}{\partial_j f} dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^{n+1}.$$

*Then one checks that* $\omega'$ *coincides on different coordinate charts and gives a globally defined non-vanishing n-form. Actually,* $\omega = |\nabla f|^2 \omega'$*.*

**Exercise 3.3.4** *By generalizing the above construction, show that a higher co-dimensional submanifold given by the level set of a rank* $k$ *vector-valued function* $F : \mathbb{R}^{n+k} \to \mathbb{R}^k$ *is orientable.*

**Example 3.3.5** *Non-orientable manifolds:*

1) *(Möbius band) Recall that the Möbius band is constructed from the 2 dimensional strip* $S = \{(x, y) \in \mathbb{R}^2 | -1 < x < 1, -1 < y < 1\}$ *by identifying the points* $(x, y)$ *and* $(x-1, -y)$ *for all* $0 < x < 1, -1 < y < 1$*.*

   *There is a projection* $\pi : S \to M$ *where* $M$ *is the Möbius band. Assume that* $M$ *is orientable and supports a non-vanishing two-form* $\omega$*, then* $\pi^*\omega$ *gives a non-vanishing two form on* $S$*. Since there is a canonical two-form* $w_0 = dx \wedge dy$ *on* $S$*, there exists a non-zero smooth function* $f$ *such that* $\pi^*\omega = f\omega_0$*. However, this is impossible since we must have* $f(x, y) = -f(x-1, -y)$*.*

*2) (The real projective space $\mathbb{R}P^n$) Let $\sigma : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ be the reflection $\sigma(x) = -x$, then the restriction of $\sigma$ on $S^n$ gives the antipodal map. There is a natural projection $\pi : S^n \to \mathbb{R}P^n$ defined by $\pi(x) = [x]$ where $[x]$ is the line through $x$. Obviously, $\pi \circ \sigma = \pi$ and $\sigma^2 = id$. Actually, $S^n$ is a double cover of $\mathbb{R}P^n$.*

*The standard orientation of $S^n$ is given by $\omega = \iota_N \Omega^{n+1}$, where $N$ is the outer unit normal of the sphere $S^n$. Under the antipodal map $\sigma$, we have for any $x \in S^n$,*

$$\sigma^*\omega(x) = \iota_{\sigma_*(N(-x))}\sigma^*(\Omega^{n+1}(-x)) = \iota_{N(x)}(-1)^{n+1}\Omega^{n+1}(x) = (-1)^{n+1}\omega(x).$$

*Now if $n$ is even, assume there is a non-vanishing $n$-form $\eta$ on $\mathbb{R}P^n$. Since $\pi$ is a local diffeomorphism, the pull-back $\pi^*\eta$ gives a non-vanishing $n$-form on $S^n$. Thus $\pi^*\eta = f\omega$ for some non-zero function $f$ on $S^n$. But*

$$f\omega = \pi^*\eta = (\pi \circ \sigma)^*\eta = \sigma^* \circ \pi^*\eta = \sigma^*(f\omega) = \sigma^* f \cdot \sigma^*\omega = f \circ \sigma \cdot (-1)^{n+1}\omega.$$

*It follows $f = (-1)^{n+1} f \circ \sigma$, thus $f$ must change sign when $n$ is even, which contradicts to the fact that $f$ is positive.*

*On the other hand, if $n$ is odd, then $\sigma^*\omega = \omega$. Hence by projecting the standard $n$-form on $S^n$, the local diffeomorphism $\pi$ induces a well-defined non-vanishing $n$-form on $\mathbb{R}P^n$. More precisely, for any $[x] \in \mathbb{R}P^n$, let $\pi_1$ and $\pi_2$ be the two diffeomorphisms defined by restricting $\pi$ on a neighborhood of $x$ and $-x$, respectively. Then we need to check $(\pi_1^{-1})^*\omega(x)$ and $(\pi_2^{-1})^*\omega(-x)$ coincide. But*

$$(\pi_2^{-1})^*\omega \circ \sigma(x) = (\pi_2^{-1})^* \circ \sigma^*\omega(x) = (\sigma \circ \pi_2^{-1})^*\omega(x) = (\pi_1^{-1})^*\omega(x).$$

*The last equality holds since $\pi_1 = \pi_2 \circ \sigma$ and $\sigma \circ \pi_2^{-1} = \pi_1^{-1}$.*

*Therefore $\mathbb{R}P^n$ is orientable iff $n$ is odd.*

**Exercise 3.3.6** *Show that the Klein bottle is non-orientable.*

### 3.3.2 Integration

Now we can define the integration of $n$-forms on an orientable manifold. The idea originates from the classical change of variables formula of integrations. Namely, under a change of variables $x \to y$ in an open set $U \in \mathbb{R}^n$, we have

$$\int_{\phi(U)} f(y)dy^1 \cdots dy^n = \int_U f(y(x))\Big|\frac{\partial(y^1, \cdots, y^n)}{\partial(x^1, \cdots, x^n)}\Big|dx^1 \cdots dx^n.$$

Notice that we take the absolute value of the Jacobi determinant.

We first define the integration of a local $n$-form on a local chart $(U, \phi; x)$. Suppose $\omega$ is a locally supported $n$-form in $U$ given by

$$\omega = f dx^1 \wedge \cdots \wedge dx^n,$$

where supp$f \subset\subset U$. Then we define the integral of $\omega$ by

$$\int_U \omega = \int_{\phi(U)} f dx^1 \cdots dx^n.$$

Next on the orientable manifold $M$ with an oriented atlas $\mathscr{A}$. Given a globally defined $n$-form $\omega \in \Lambda^n(M)$, we first decompose it into local ones by a partition of unity $\{\psi_k\}$ subordinating to the oriented atlas $\mathscr{A} = \{(U_k, \phi_k)\}$, such that

$$\omega = \sum_k \psi_k \omega = \sum_k f_k dx_k^1 \wedge \cdots \wedge dx_k^n,$$

and supp$f_k \subset\subset U_k$. Then the integral of $\omega$ on $M$ is given by

$$\int_M \omega = \sum_k \int_{U_k} \psi_k \omega = \sum_k \int_{\phi_k(U_k)} f_k dx_k^1 \cdots dx_k^n.$$

This integration is well-defined since the Jacobian determinants of compatible charts are positive and consistent to the change of variable formula. Note that the integral differs by a sign if we choose the opposite orientation.

More globally, if we fix a non-vanishing $n$-form $\Omega$ on an oriented manifold $M$ as a background (volume) form. Then any $n$-form $\omega \in \Lambda^n(M)$ can be represented by $\omega = f\Omega$ for some $f \in C^\infty(M)$, and the integral of $\omega$ is

$$\int_M \omega = \int_M f\Omega.$$

## 3.4 Stokes' theorem

### 3.4.1 Manifold with boundary

Let $H^n := \{x \in \mathbb{R}^n | x_n \geq 0\} \subset \mathbb{R}^n$ be the $n$-dimensional upper half-plane, which is equipped with the subset topology induced from $\mathbb{R}^n$. The boundary of $H^n$ is the $(n-1)$-dimensional plane

$$\partial H^n = \{x \in \mathbb{R}^n | x^n = 0\}.$$

For open sets $U, V \subset H^n$, we say a map $f : U \to V$ is smooth if it is the restriction of a smooth map from an open set in $\mathbb{R}^n$ to $\mathbb{R}^n$. Similarly, $f$ is a diffeomorphism if it is the restriction of a diffeomorphism between two open sets in $\mathbb{R}^n$.

**Definition 3.4.1** *A manifold with boundary is a manifold where each local chart is homeomorphic to an open set in the half space $H^n$. A point $p \in M$ is called an interior point if there is a local chart such that $\phi(p) \notin \partial H^n$. Otherwise, we call it a boundary point. The boundary of $M$ is the set of all boundary points and denoted by $\partial M$.*

It is easy to see that the definition of boundary points is independent of choice of local charts. A point $p \in M$ is a boundary point iff in any local chart it corresponds to a boundary point of the upper half-plane $\phi(p) \in \partial H^n$. So locally, the boundary points are exactly the set $\phi^{-1}(U \cap \partial H^n)$.

The restriction of an atlas of $M$ gives an atlas on $\partial M$. Thus $\partial M$ is a $(n-1)$ dimensional embedded manifold of $M$. Note that a manifold could have (even infinitely) many disconnected boundary components.

Typical examples include the half-plane, unit ball, semi-sphere and cylinders. In particular, a manifold with an open set removed yields a manifold with boundary.

### 3.4.2 Induced orientation on boundary manifold

**Theorem 3.4.2** *If $M$ is an orientable manifold with boundary, then its boundary $\partial M$ is orientable. In particular, the restriction of an oriented atlas on $M$ gives an oriented atlas on $\partial M$.*

*Proof.* Suppose $(U, x), (V, y)$ are two charts near boundary on $M$ such that the Jacobian matrix $\partial y / \partial x$ has positive determinant. By definition, for any boundary point $p \in U \cap V$, we have $x^n(p) = y^n(p) = 0$ and in the interior $(U \cap V) \setminus \partial M$, we have $x^n > 0, y^n > 0$. So the transition map $y = y(x)$ has the property that $y^n(x^1, \cdots, x^{n-1}, 0) = 0$ and $y^n(x) > 0$ if $x^n > 0$. Therefore

$$\frac{\partial y^n}{\partial x^1}\Big|_p = \cdots = \frac{\partial y^n}{\partial x^{n-1}}\Big|_p = 0, \quad \frac{\partial y^n}{\partial x^n}\Big|_p > 0.$$

Let $U' = U \cap \partial M, V' = V \cap \partial M$ and $x' = (x^1, \cdots, x^{n-1}), y' = (y^1, \cdots, y^{n-1})$ be the restriction of $(U, x), (V, y)$ on $\partial M$. Then we have

$$\det \frac{\partial y}{\partial x} = \det \frac{\partial y'}{\partial x'} \cdot \det \frac{\partial y^n}{\partial x^n} > 0.$$

It follows that $\det \frac{\partial y'}{\partial x'} > 0$ and $(U', x'), (V', y')$ are compatible charts on $\partial M$.

Therefore, if there is an oriented atlas on $M$, then its restriction gives an oriented atlas on $\partial M$. This finishes the proof. □

Next, we specify the induced orientation on the boundary manifold, which is the "outer normal" orientation defined as follows.

First observe that there exist a non-vanishing *outer normal* vector field $N$ on $\partial M$. In each local chart $(U, x)$ at the boundary, a local outer normal vector field is given by $-\frac{\partial}{\partial x_n}$. Then $N$ is constructed by gluing them together via partition of unity. One verifies that $N$ is indeed non-vanishing on $\partial M$.

Now for an orientation on $M$ represented by a non-vanishing $n$-form $\Omega \in \Lambda^n(M)$, the induced orientation $\Omega'$ of $\partial M$ is defined by

$$\Omega' = \big(\iota_N \Omega\big)|_{\partial M}.$$

Note that in a local chart $(U, x)$ which is compatible with $\Omega$, the induced orientation $\Omega'$ is

$$\iota_{-\frac{\partial}{\partial x^n}}\big(dx^1 \wedge \cdots \wedge dx^n\big) = (-1)^n dx^1 \wedge \cdots \wedge dx^{n-1}.$$

That is, the restricted chart $(\partial U, (x^1, \cdots, x^{n-1}))$ on $\partial M$ differs from the induced orientation $\Omega'$ by $(-1)^n$.

**Remark 3.4.3** *Obviously there is an (opposite) induced orientation on the boundary, which corresponds to the inner normal vector field. We choose the outer normal orientation to make the following Stokes' theorem more elegant.*

**Remark 3.4.4** *We are not saying that the boundary of a non-orientable manifold is not orientable. For example, the boundary of a Möbius band is $S^1$, which is of course orientable.*

### 3.4.3 Stokes' formula

Since $\iota : \partial M \to M$ is a submanifold, if $\omega \in \Lambda^p(M)$ is a form on $M$, its restriction $\omega|_{\partial M} := \iota^*\omega$ gives a form on $\partial M$.

**Theorem 3.4.5** *Suppose $M$ is an $n$-dimensional oriented manifold with boundary $\partial M$, which is endowed with the induced orientation. Then for any $(n-1)$-form $\omega \in \Lambda^{n-1}(M)$,*

$$\int_M d\omega = \int_{\partial M} \omega|_{\partial M}.$$

*Proof.* Choose a partition of unity subordinating to an oriented atlas on $M$. By definition

$$\int_M d\omega = \sum \int_U \psi d\omega = \int_M d\left(\sum \psi\omega\right) = \sum \int_U d(\psi\omega).$$

Locally on each chart $(U, \phi)$, we have

$$\psi\omega = \sum_{i=1}^n f_i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n,$$

and

$$d(\psi\omega) = \sum_{i=1}^n (-1)^{i-1} \partial_i f_i dx^1 \wedge \cdots \wedge dx^n.$$

Since $f_i$ has compact support on $U$, if $U$ does not intersect with the boundary, we have

$$\int_U d(\psi\omega) = \sum_{i=1}^n \int_{\mathbb{R}^n} (-1)^{i-1} \partial_i f_i dx^1 \cdots dx^n = 0.$$

Otherwise,

$$\begin{aligned}\int_U d(\psi\omega) &= (-1)^{n-1} \int_{\mathbb{R}^{n-1}} dx^1 \cdots dx^{n-1} \int_0^{+\infty} \partial_n f_n dx^n \\ &= (-1)^n \int_{\mathbb{R}^{n-1}} f_n(x^1, \cdots, x^{n-1}, 0) dx^1 \cdots dx^{n-1}.\end{aligned}$$

On the other hand, at the boundary of $U$, we have

$$\begin{aligned}(\psi\omega)|_{\partial U} &= \psi'\omega|_{\partial M} = f_n|_{x^n=0}dx^1\wedge\cdots\wedge dx^{n-1}\\ &= (-1)^n f_n(x^1,\cdots,x^{n-1},0)\cdot(-1)^n dx^1\wedge\cdots\wedge dx^{n-1}.\end{aligned}$$

Since the induced orientation on $\partial M$ is $(-1)^n dx^1\wedge\cdots\wedge dx^{n-1}$, we get

$$\int_{\partial M}\omega|_{\partial M} = \sum\int_{U'}\psi'\omega|_{\partial M} = \sum\int_{\mathbb{R}^{n-1}}(-1)^n f_n(x^1,\cdots,x^{n-1},0)dx^1\cdots dx^{n-1}.$$

This completes the proof. $\square$

As an easy corollary, if the manifold $M$ is closed, i.e. $\partial M=\emptyset$, then $\int_M d\omega = 0$.

**Exercise 3.4.6** *Show that the above Stokes' theorem is a generalization of the classical Green formula, Gauss formula and Stokes formula in multi-variable calculus.*

## 3.5 de Rham cohomology

### 3.5.1 Definition

The property $d^2 = 0$ induces a natural cohomology on a differentiable manifold, which is called the *de Rham cohomology*.

Write the restriction of the exterior derivative on $\Lambda^p(M)$ as $d_p : \Lambda^p(M) \to \Lambda^{p+1}(M)$, then $d_p \circ d_{p-1} = 0$ implies $\mathrm{Im} d_{p-1} \subset \mathrm{Ker}\, d_p$. Therefore we can defined the *p-th de Rham cohomology group* as the quotient space

$$H_d^p(M) = \mathrm{Ker}\, d_p / \mathrm{Im} d_{p-1}.$$

We call elements in $\mathrm{Ker}\, d$ *closed forms* and those in $\mathrm{Im} d$ *exact forms*. Thus an equivalent class $[\omega] \in H_d^p(M)$ of a closed form $\omega \in \mathrm{Ker}\, d_p$ is given by

$$[\omega] = \{\omega + d\eta | \eta \in \Lambda^{p-1}(M)\}.$$

By convention, we define $H_d^0(M) = \mathrm{Ker}\, d_0$ to be the space of constant functions, which is just $\mathbb{R}^1$ for connected manifolds.

**Remark 3.5.1** *Although we can it the cohomology group, it is simply a real vector space.*

It is straight forward to check that the de Rham cohomology groups of a manifold $M$ of dimension $n$ satisfies the following properties:

- $H^p(M) = 0$ if $p > n$.
- for $a \in H^p(M)$ and $b \in H^q(M)$, there is a bilinear product $a \wedge b \in H^{p+q}(M)$ which satisfies
$$a \wedge b = (-1)^{pq} b \wedge a.$$
- if $F : M \to N$ is a smooth map, then it defines a natural linear map
$$F^* : H^p(N) \to H^p(M)$$
where commutes with the wedge product.

**Theorem 3.5.2** *Let $F : M \times [0,1] \to N$ be a smooth map. Set $F_t(x) = F(x,t)$ and consider the induced map $F_t^* : H^p(N) \to H^p(M)$. Then*

$$F_1^* = F_0^*.$$

*Proof.* Represent $a \in H^p(N)$ by a closed $p$-form $\alpha \in \Lambda^p(N)$ and consider the pull-back form $F^*\alpha$ on $M \times [0,1]$. We can decompose it uniquely in the form

$$F^*\alpha = \beta + dt \wedge \gamma$$

where $\beta(t) \in \Lambda^p(M)$ and $\gamma(t) \in \Lambda^{p-1}(M)$ for all $t \in [0,1]$. More explicitly,

$$\beta = F_t^*\alpha, \quad \gamma = \iota_{\frac{\partial}{\partial t}} F^*\alpha.$$

Since $\alpha$ is closed, we have

$$0 = d\alpha = d_M\beta + dt \wedge \partial_t\beta - dt \wedge d_M\gamma,$$

where $d_M$ is the exterior derivative of $M$. It follows that $\partial_t\beta = d_M\gamma$.

Now integrating with respect to $t$, we obtain

$$F_1^*\alpha - F_0^*\alpha = \int_0^1 \partial_t\beta dt = \int_0^1 d_M\gamma dt = d_M \int_0^1 \gamma dt.$$

So the closed forms $F_1^*\alpha$ and $F_0^*\alpha$ differ by an exact form and

$$F_1^* a = F_0^* a.$$

□

**Example 3.5.3** $H^p(\mathbb{R}^n) = 0$ *for* $p > 0$.

*Proof.* Consider the family of maps

$$F : \mathbb{R}^n \times [0,1] \to \mathbb{R}^n, \quad (x,t) \mapsto tx.$$

Then $F_1(x) = x$ is identity and $F_0(x) = 0$. It follows that $F_1^* : H^p(\mathbb{R}^n) \to H^p(\mathbb{R}^n)$ is identity while $F_0^* : H^p(\mathbb{R}^n) \to H^p(\mathbb{R}^n)$ vanishes. Then if follows from Theorem 3.5.2 that

$$H^p(\mathbb{R}^n) = F_1^*(H^p(\mathbb{R}^n)) = F_0^*(H^p(\mathbb{R}^n)) = 0.$$

□

Similar arguments show that the $H^p(U) = 0$ for a star-shaped set $U \subset \mathbb{R}^n$ and $p > 0$. This is usually called the *Poincaré Lemma*.

**Theorem 3.5.4 (Poincaré Lemma)** *Any closed form on a star-shaped set $U \subset \mathbb{R}^n$ is exact.*

**Example 3.5.5** *$H^p(S^n) = \mathbb{R}^1$ if $p = 0$ or $p = n$, and vanishes otherwise.*

Refer to Hitchin [1] for more details and applications.

### 3.5.2 de Rham theorem

The *differentiable singular homology* is constructed by requiring the continuous map in the definition of singular simplices to be differentiable. Namely, a differentiable singular $p$-simplex in $M$ is a *smooth* map $\sigma$ from the standard $p$-simplex $\Delta^p \subset \mathbb{R}^p$ to $M$, where

$$\Delta^p := \{x \in \mathbb{R}^p | \sum_{i=1}^{p} x^i = 1 \text{ and } x_i \geq 0\}.$$

Then the integration of a $p$-form $\omega \in \Lambda^p M$ on the simplex $\sigma$ is defined as

$$\int_\sigma \omega := \int_{\sigma(\Delta^p)} \omega = \int_{\Delta^p} \sigma^* \omega.$$

This integration can be naturally generalized to $p$-chains and $p$-cycles.

**Theorem 3.5.6 (de Rham theorem)** *Suppose $M$ is an $n$-dimensional differentiable manifold, then for each $p = 0, 1, \cdots, n$, there exists an isomorphism between the $p$-th de Rham cohomology $H_d^p(M)$ and the $p$-th (differentiable) singular cohomology $H_p^*(M; \mathbb{R}^1)$, which is given by*

$$[\omega] \cdot [c] = \int_c \omega.$$

*Here $\omega$ is a closed $p$-form and $c$ is a (differentiable) $p$-cycle.*

The above action is well-defined since for any representative $\omega' = \omega + d\eta \in [\omega]$ and $c' = c + \partial b \in [c]$, we have

$$\int_{c'} \omega' = \int_{c+\partial b} (\omega + d\eta) = \int_c \omega + \int_c d\eta + \int_{\partial b} \omega'.$$

Recall that by definition $\partial c = 0$ and $d\omega' = 0$. It follows from Stoke's theorem that

$$\int_c d\eta = \int_{\partial c} \eta = 0$$

and

$$\int_{\partial b} \omega' = \int_b d\omega' = 0.$$

In fact, one can prove the differential singular homology is the same as the (continuous) singular homology. Thus the de Rham cohomology is a topological invariant. Moreover, the exterior product induces an ring structure on the de Rham cohomology, and there is a ring structure on the singular cohomology given by the *cup product*. The map constructed in the de Rham theorem is actually an algebra isomorphism, i.e. it preserves the ring structure.

# Part II

# Riemannian Geomery

# Chapter 4

# Riemannian Manifolds

## 4.1 Riemannian metrics

We have now seen that with the differential structure on a differentiable manifold, we can perform both differentiation and integration on the manifold. The basic idea is to mimic locally what we do on Euclidean space, and then patch everything together in a well-defined global way. There is another important structure other than the differential structure, namely, the inner produce structure, that can be also generated to differentiable manifolds in a similar way. In this section, we will see that for any differentiable manifold $M$, we can equip a metric on $M$, which is represented by a $(0, 2)$-tensor. With this Riemannian metric, we can finally discuss the geometry of a differentiable manifold.

### 4.1.1 Inner product on a vector space

Let's first investigate the notion of a metric on a (finitely dimensional) linear space, i.e. the Euclidean space. A metric space is given by a distance function which satisfies the triangle inequality. For a linear space $V$, a norm gives a distance function by letting

$$d(v, w) = |v - w|, \quad \forall x, y \in V.$$

In particular, an inner product on $V$ will provide us both a norm $|v| = \langle v, v\rangle^{\frac{1}{2}}$ and the notion of an angle, by

$$\langle v, w\rangle = |v| \cdot |w| \cdot \cos\angle(v, w).$$

For a finitely dimensional space, to define an inner product is equivalent to specify a set of orthonormal basis, which is in turn equivalently defined by a positive definite matrix under a choice of basis. For us, however, the most convenient way is to view an inner product as a $(0, 2)$-tensor $g \in T^{(0,2)}V$ such that

$$g(v, w) = \langle v, w\rangle, \quad \forall x, y \in V.$$

where $g$ is symmetric and positive definite.

### 4.1.2 Riemannian metric

Now for a differentiable manifold $M$, at each point $p \in M$ the infinitesimal local structure is given by the tangent space $T_pM$, which is indeed a linear space. Thus we first define an inner product on $T_pM$ at each point $p \in M$, then we require it varies smoothly on $M$.

**Definition 4.1.1** *A Riemannian metric on a manifold $M$ is a symmetric positive definite $(0,2)$-tensor $g \in \Gamma(T^{(0,2)}M)$. A manifold endowed with a Riemannian metric is called a Riemannian manifold*

In local coordinates, we can write

$$g|_U(x) = g_{ij}(x)dx^i \otimes dx^j =: g_{ij}(x)dx^i dx^j,$$

where $dx^i dx^j = \mathscr{S}(dx^i \otimes dx^j)$ is the symmetric tensor product. Again the matrix $(g_{ij})$ is symmetric and positive definite.

The existence of a Riemannian metric dose not impose any restrictions on the manifold. Indeed, for any manifold, one can construct a Riemannian metric by patching local Euclidean metrics together by a partition of unity.

**Example 4.1.2** *The Euclidean metric on a Euclidean space can be written as*

$$g_{\mathbb{R}^n} = \delta_{ij}dx^i dx^j = \sum_i (dx^i)^2$$

*in Cartesian coordinates. In particular, in the polar coordinates on $\mathbb{R}^2$, the standard metric is*

$$g_{\mathbb{R}^2} = dr^2 + r^2 d\theta^2.$$

An immersed submanifold of a Riemannian manifold is again a Riemannian manifold. More precisely, if $(M, g)$ is a Riemannian manifold and $\iota : N \hookrightarrow M$ is an immersion, then $N$ is a Riemannian manifold with the pull-back metric $\iota^* g$.

**Example 4.1.3** *The two sphere $S^2$ is naturally embedded in the Euclidean space $\mathbb{R}^3$. Denote the north and south pole of $S^2$ by $N$ and $S$ respectively. Recall that the spherical coordinates*

*$(r,\theta)\in(0,\pi)\times[0,2\pi)$ is well-defined on $S^2\setminus\{N,S\}$. The embedding is given by $\iota: S^2\to\mathbb{R}^3$ where*

$$(x^1,x^2,x^3)=\iota(r,\theta)=(\sin r\cos\theta,\sin r\sin\theta,\cos r).$$

*The standard metric of $S^2$ is the pull-back metric $g=\iota^*g_0=dr^2+\sin^2 r d\theta^2$.*

*Using the stereographic projection at $S$, we can write the metric in another coordinate. Recall the stereographic projection $\mathcal{P}_S: S^2\setminus\{S\}\to\mathbb{R}^2$ is a bijection given by*

$$(\xi^1,\xi^2)=\mathcal{P}_S(x^1,x^2,x^3)=\left(\frac{x^1}{1+x^3},\frac{x^2}{1+x^3}\right).$$

*Then one verifies that the metric has the form*

$$g=\mathcal{P}_S^*g_0=\frac{4}{(1+\sum_i(\xi^i)^2)^2}\delta_{ij}d\xi^i d\xi^j.$$

**Example 4.1.4** *(The two dimensional hyperbolic space)*

- *The half-plane model*

$$\mathbb{H}^2=\{(x,y)\in\mathbb{R}^2|y>0\}$$

*with*

$$g_1=\frac{1}{y^2}(dx^2+dy^2).$$

- *The Poincaré disk model*

$$\mathbb{D}^2=\{(\xi^1,\xi^2)\in\mathbb{R}^2|\xi^2+\eta^2<1\}$$

*with*

$$g_2=\frac{4}{1-\sum_i(\xi^i)^2}\delta_{ij}d\xi^i d\xi^j.$$

- *The hyperboloid model*

$$\mathcal{H}^2=\{(x^1,x^2,x^3)\in\mathbb{R}^{2+1}|(x^1)^2+(x^2)^2-(x^3)=-1,x^3>0\}$$

*with*

$$g_3=dr^2+\sinh^2 r d\theta^2,$$

*where we use the hyperbolic coordinates*

$$(r, \theta) \to (x^1, x^2, x^3) = (\sinh r \cos\theta, \sinh r \sin\theta, \cosh r).$$

**Definition 4.1.5** *Let $(M, g)$ and $(N, h)$ be two Riemannian manifolds and $F : M \to N$ be a diffeomorphism. The map $F$ is called an* isometry *if the pull-back metric satisfies $F^*h = g$. If such an isometry exists, we also say $(M, g)$ and $(N, h)$ is* isometric.

We may view the half-plane $\mathbb{H}^2$ and the poincaré disk $\mathbb{D}^2$ as subsets in the complex plane $\mathbb{C}^1$. Then the isometry of $\mathbb{H}^2$ and $\mathbb{D}^2$ is given by the Mobius tranformation

$$\mathbb{H}^2 \to \mathbb{D}^2, \quad z \to i\frac{z-i}{z+1}.$$

The isometry between the Poincaré disk $\mathbb{D}^2$ and the hyperboloid $\mathcal{H}^2$ is given by the stereographic projection.

**Exercise 4.1.6** *For an open neighborhood $U \subset M$, show that there exists a local orthonormal basis $\{\omega_i\} \subset \Gamma(T^*U)$ such that $g|_U = \sum(\omega^i)^2$.*

### 4.1.3 Manipulating indices of tensors

At each point $p \in M$, the Riemannian metric $g$ induces an isomorphism $\Phi_g : T_pM \to T_p^*M$ as follows. For each $v \in T_pM$, we can find a unique dual $v^* \in T_p^*M$ by letting

$$v^*(w) = g(v, w), \forall w \in T_pM.$$

Similarly, for each $\omega \in T_p^*M$, we have a dual $w^* \in T_pM$ such that

$$\omega(w) = g(\omega^*, w), \forall w \in T_pM.$$

Locally, suppose $\{e_i\}$ is a basis of $T_pM$ and $\{\delta^i\} \subset T_p^*M$ is its dual. If $v = v^i e_i$, then we have $v^* = v_i\delta^i$, where $v_i = g_{ij}v^j$. Similarly, for $\omega = a_i\delta^i$ and $\omega^* = a^i e_i$, we have $a^i = g^{ij}a_i$ where $(g^{ij})$ is the inverse of $(g_{ij})$. We call this process lowering the index of $v$ and raising the index of $\omega$, respectively.

By varying the point, the above process can be easily defined on a vector field or an 1-form. For example, a vector field $V = v^i \frac{\partial}{\partial x^i}$ can be identified with a 1-form $\omega = g_{ij} v^j dx^i$. In fact, we may extend the isomorphism $\Phi_g$ to any tensor space, and identify a $(r,s)$-tensor $T$ with a $(r+1, s-1)$-tensor $\bar{T}$ or a $(r-1, s+1)$-tensor $\tilde{T}$ by raising or lowering the index, respectively. In local coordinates, this means

$$\bar{T}^{i_1 \cdots i_r k}_{j_1 \cdots j_{s-1}} = g^{lk} T^{i_1 \cdots i_r}_{j_1 \cdots j_{s-1} l}, \quad \tilde{T}^{i_1 \cdots i_{r-1}}_{j_1 \cdots j_s k} = g_{lk} T^{i_1 \cdots i_{r-1} l}_{j_1 \cdots j_s}.$$

**Exercise 4.1.7** *For a $(0,2)$-tensor $h \in \Gamma(T^{(0,2)}M)$, identify it with a $(1,1)$-tensor and a $(2,0)$-tensor by raising the indices. In particular, what are the corresponding tensors for the Riemannian metric $g$?*

### 4.1.4 Induced metric on tensors

The Riemannian metric $g$ naturally induces a metric $g^*$ on 1-forms by

$$g^*(\omega, \eta) = g(\omega^*, \eta^*), \forall \omega, \eta \in \Lambda^1(M).$$

Then the isomorphism $\Phi_g$ defined above becomes an isometry.

Next we can extend $g$ to any tensor space on $M$ by induction as follows. Suppose $\alpha_1, \alpha_2 \in T^{(r,s)}M, \beta_1, \beta_2 \in T^{(k,l)}M$, then we define

$$g(\alpha_1 \otimes \beta_1, \alpha_2 \otimes \beta_2) = g(\alpha_1, \alpha_2) \cdot g(\alpha_2, \beta_2).$$

For example, for the $(0,2)$-tensor $g$ itself, we can compute

$$\langle g, g \rangle = \langle g_{ij} dx^i \otimes dx^j, g_{kl} dx^k \otimes dx^l \rangle = g_{ij} g_{kl} g^{ik} g^{jl} = \delta^k_i \delta^l_k = n.$$

Since any $p$-form is just an anti-symmetric tensor, we also obtain an inner produce for $p$-forms. Indeed, for $\alpha_i, \beta_j \in \Lambda^1(M), 1 \le i, j \le p$, we have

$$g(\alpha_1 \wedge \cdots \wedge \alpha_p, \beta_1 \wedge \cdots \wedge \beta_p) = \det\big(g(\alpha_i, \beta_j)\big).$$

**Exercise 4.1.8** *Prove the above identity.*

### 4.1.5 Volume form

Suppose $(M, g)$ is an oriented Riemannian manifold. Since the space $\Lambda^n(M)$ of $n$-forms is of one dimensional, there exists a unique $n$-form $dv$ of unit length, which represents the orientation. We call $dv$ the *volume form* of $(M, g)$.

Locally in a compatible chart, if $dv = a(x)dx^1 \wedge \cdots dx^n$, then

$$\langle dv, dv\rangle = a^2 \det(g^{ij}) = 1.$$

Thus $dv = \sqrt{\det G}dx^1 \wedge \cdots dx^n$ where $G = \det(g_{ij})$.

With the volume form, we can define the integration of any smooth function $f \in C^\infty(M)$ by

$$\int_M f dv = \int_M f(x)\sqrt{\det G}dx^1 \wedge \cdots dx^n.$$

### 4.1.6 Distance function on a Riemannian manifold

Using the Riemannian metric, we can define the length of any tangent vector. Namely, we define the norm

$$|v| = \langle v, v\rangle^{\frac{1}{2}} = g(v, v)^{\frac{1}{2}}, \forall v \in T_pM.$$

Since a smooth curve can be approximated by a piecewise linear curve, we then can define the length of the curve by taking a limit. More precisely, suppose $\gamma : [a, b] \to M$ is a smooth curve, then we define the length of $\gamma$ by

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

Finally, for any $p, q \in M$, we consider the space $\mathcal{C}(p, q)$ of all possible curves on $M$ that connects $p$ and $q$, and define the *distance function* by

$$d(p, q) = \inf_{\gamma\in\mathcal{C}(p,q)} \mathcal{L}(\gamma).$$

One can verify that $d$ is indeed a distance function, i.e. it satisfies

1) $d(p, q) = d(q, p) \geq 0$ for any $p, q \in M$,

2) $d(p, q) = 0$ iff $p = q$,

3) $d(p,q) \leq d(p,r) + d(r,q)$ for any $p, q, r \in M$.

It takes some efforts to prove item 2). One needs to show that since the metric is smooth, locally the metric $g = g_{ij}dx^i dx^j$ and the standard one $g_0 = \sum_i dx^i dx^i$ are equivalent. In fact, we have the following theorem.(cf. [3], page 94)

**Theorem 4.1.9** *The topology of the manifold $M$ coincides with the topology induced by the distance function $d$.*

# 4.2 Hodge theorem

## 4.2.1 Hodge star operator

For an $n$-dimensional vector space $V$, since the dimension of $\Lambda^p V$ and $\Lambda^{n-p}V$ are equal $C_n^p = C_n^{n-p}$, one may wonder if there is an canonical isomorphism between these two linear spaces. This is indeed the case if $V$ is endowed with an inner product structure.

Fixing an orientation and hence a volume form $dv$, the *Hodge star operator* is the linear operator

$$* : \Lambda^p V \to \Lambda^{n-p}V, \quad \eta \mapsto *\eta$$

such that the $n$-form

$$\omega \wedge (*\eta) = \langle \omega, \eta \rangle dv, \quad \forall \omega \in \Lambda^p V.$$

Note that we need an orientation on $V$ to assign the volume form. As an easy corollary, we have $*1 = dv$ and

$$\omega \wedge (*\omega) = \langle \omega, \omega \rangle dv, \quad \forall \omega \in \Lambda^p V.$$

More explicitly, if $\{e_1, \cdots, e_n\} \in V^*$ is an oriented orthonormal basis of $V^*$ such that $dv = e_1 \wedge \cdots \wedge e_n$. Then we have

$$*(e_1 \wedge \cdots \wedge e_p) = e_{p+1} \wedge \cdots \wedge e_n.$$

It is easy to verify that

$$* \circ * = (-1)^{p(n-p)} \mathrm{id}_p.$$

**Exercise 4.2.1** *Prove the above identity.*

## 4.2.2 Co-differential operator

On an oriented Riemannian manifold $M$ with a metric $g$, the Hodge star operator is point-wisely well-defined on each cotangent space $T_p^*M, \forall p \in M$. Thus we have a global operator $* : \Lambda^p(M) \to \Lambda^{n-p}(M)$.

We can also introduce an inner product on $\Lambda^p(M)$ by letting

$$(\omega, \eta) := \int_M \langle \omega, \eta \rangle dv = \int_M \omega \wedge (*\eta).$$

Recall that there is an exterior differential operator $d : \Lambda^p(M) \to \Lambda^{p+1}(M)$. Using the star operator, one can define an adjoint operator $\delta : \Lambda^p(M) \to \Lambda^{p-1}(M)$ by setting

$$\delta = (-1)^{n(p+1)+1} * \circ d \circ *.$$

**Theorem 4.2.2** *For a compact Riemannian manifold $M$ without boundary and $\omega \in \Lambda^p(M)$ and $\eta \in \Lambda^{p+1}(M)$,*

$$(d\omega, \eta) = (\omega, \delta\eta).$$

*Proof.* First, we have

$$d(\omega \wedge *\eta) = d\omega \wedge *\eta + (-1)^p \omega \wedge (d \circ *\eta).$$

Applying Stoke's theorem, we get

$$(d\omega, \eta) = \int_M d\omega \wedge *\eta = (-1)^{p+1} \int_M \omega \wedge (d \circ *\eta).$$

But for $d \circ *\eta \in \Lambda^{n-p}(M)$,

$$d \circ *\eta = (-1)^{p(n-p)} * \circ * \circ d \circ *\eta = (-1)^{-p^2+1} * \circ \delta\eta.$$

Since $p(p+1)$ is even, the theorem follows. □

**Remark 4.2.3** *Actually, we have an invariant formula for $\omega \in \Lambda^p(M)$ as follows*

$$\delta\omega(\cdot) = -div\omega(\cdot) = -\nabla_{e_i}\omega(e_i, \cdot),$$

*where $\nabla$ is the Levi-Civita connection. In particular, for a one form $\alpha = a_i dx^i$, we have*

$$\delta\alpha = -div\alpha = -g^{ij} a_{i,j}.$$

*Applying the Hodge star operator, we get*

$$* \circ \delta\alpha = -d \circ *\alpha = - * div\alpha = -div\alpha dv.$$

*Therefore, by the Stoke's formula, we have*

$$\int_M div\alpha dv = \int_{\partial M} *\alpha.$$

### 4.2.3 Hodge theorem

The *Hodge Laplacian* operator $\Delta : \Lambda^p(M) \to \Lambda^p(M)$ for $p$-forms is defined by

$$\Delta = \delta \circ d + d \circ \delta.$$

A $p$-form $\omega$ is called *harmonic* if $\Delta\omega = 0$. Since

$$(\Delta\omega, \omega) = (d\omega, d\omega) + (\delta\omega, \delta\omega),$$

we see that $\omega$ is harmonic iff $d\omega = 0$ and $\delta\omega = 0$. Moreover, the Laplacian operator is self-adjoint, i.e.

$$(\Delta\omega, \eta) = (\omega, \Delta\eta).$$

**Remark 4.2.4** *For 0-forms $f \in \Lambda^0(M) = C^\infty(M)$, the Hodge Laplacian is*

$$\Delta f = \delta \circ df,$$

*which differs from the trace Laplacian by a minus sign. Thus the Hodge Laplacian is a generalization of classical Laplacian operator on forms.*

**Theorem 4.2.5 (Hodge theorem)** *There is a unique harmonic $p$-form in each equivalent class in the de Rham cohomology group $H^p_d(M), \forall 1 \le p \le n$.*

In fact, the exterior space have the following decomposition

$$\Lambda^p(M) = \mathrm{Im}(\Delta_p) \oplus \mathrm{Ker}(\Delta_p) = \mathrm{Im}(d_{p-1}) \oplus \mathrm{Im}(\delta_{p+1}) \oplus \mathrm{Ker}(\Delta_p).$$

Therefore, if we define the Hodge cohomology

$$\mathcal{H}^p(M) = \mathrm{Ker}(\Delta_p) = \{\omega \in \Lambda^p(M) : \Delta\omega = 0\},$$

then we have an isomorphism between the Hodge cohomology and the de Rham cohomology.

**Remark 4.2.6** *Recall that the de Rham cohomology is isomorphic to the singular cohomology, which is entirely a topological invariant. Thus, for a differentiable manifold, we can apply the powerful machinery of a Riemannian metric to investigate the topological information. This is perhaps one of the most important goals of Riemannian geometry.*

# Chapter 5

# Connections

# 5.1 Linear connection

## 5.1.1 Directional derivative and linear connection

We first review the classical directional derivatives on Euclidean space, then introduce the concept of linear connections on manifolds.

Given a point $x \in \mathbb{R}^n$ and a vector $v \in T_x\mathbb{R}^n = \mathbb{R}^n$, the directional derivative of a vector field (that is, a vector-valued function) $Y$ is given by

$$D_vY(x) = \frac{d}{dt}\Big|_{t=0} Y(x+tv) = \lim_{t\to 0}\frac{Y(x+tv)-Y(x)}{t}.$$

The operator $D$ satisfies Leibniz's rule for $Y$, i.e.

$$D_v(fY) = D_vf \cdot T + f \cdot D_vT,$$

and is linear on $v$.

One way of taking derivatives of two vector fields on a differentiable manifold $M$ is the Lie derivative. Recall that the Lie derivative of two vector fields $X, Y \in \mathfrak{X}(M)$ by

$$\mathcal{L}_XY(x) = \frac{d}{dt}\Big|_{t=0}(\phi_{-t})_* \circ Y(t) = \lim_{t\to 0}\frac{(\phi_{-t})_* \circ Y(\phi_t(x)) - Y(x)}{t} = [X,Y],$$

where $\phi_t$ is the one-parameter group of diffeomorphisms generated by $X$. The Lie derivative satisfies the Leibniz's rule for $Y$, i.e.

$$\mathcal{L}_X(fY) = X(f)Y + f\mathcal{L}_XY.$$

However, to define the Lie derivative $\mathcal{L}_XY(x)$, we need $X$ to be defined on a neighborhood of $x$. The problem is that we use $X$ to define $\phi_t$, and hence the mapping $(\phi_t)_*$ between different tangent spaces. Thus the Lie derivative is essentially different from the classical directional derivative.

To solve the problem, the key is to find a way to *connect* different tangent spaces. There are many different ways to assign such a connection. One short-cut is to put the most desired properties directly in our definition, as follows.

**Definition 5.1.1** *A* linear connection *is a linear operator*

$$D : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M),$$
$$(X, Y) \mapsto D_X Y$$

*which satisfies, for any $f \in C^\infty(M)$,*

*(1) Leibniz's rule on $Y$:*

$$D_X(fY) = X(f) \cdot Y + f \cdot D_X Y.$$

*(2) $C^\infty$-linear (or* tensorial*) on $X$:*

$$D_{fX} Y = f D_X.$$

Property (2) in the above definition guarantees that $D_X Y(x)$ only depends on $X(x)$ (instead of the value of $X$ in a neighborhood). It also allows us to view $D$ as a *covariant derivative* operator

$$D : \mathfrak{X}(M) \to \mathfrak{X}(M) \otimes \Lambda^1(M),$$
$$Y \mapsto DY,$$

such that $(DY)(X) = D_X Y$.

For a smooth function $f \in C^\infty(M)$, we will denote $Df = df$ and

$$D_X f = df(X) = X(f).$$

The connection naturally extends to 1-forms by requiring the Leibniz's rule

$$D_X(\omega(Y)) = D_X \omega(Y) + \omega(D_X Y).$$

Or equivalently,

$$D_X \omega = D_X \circ \omega - \omega \circ D_X = [D_X, \omega].$$

Moreover, we can extend $D$ to a differential operator on tensors

$$D : \mathfrak{X}(M) \times \mathscr{T}(M) \to \mathscr{T}(M)$$

by letting

$$
\begin{aligned}
&D_X(V_1 \otimes \cdots \otimes V_r \otimes \omega^1 \otimes \cdots \otimes \omega^s) \\
&= D_X V_1 \otimes \cdots \otimes V_r \otimes \omega^1 \otimes \cdots \otimes \omega^s + \cdots + V_1 \otimes \cdots \otimes V_r \otimes \omega^1 \otimes \cdots \otimes D_X \omega^s.
\end{aligned}
$$

**Exercise 5.1.2** *Show the existence of at least one linear connection on a manifold. (Hint: first define trivial linear connections in local charts, then use partition of unity.)*

### 5.1.2 Connection coefficients and connection forms

Now we discuss local expressions of a linear connection $D$. Locally in a neighborhood $x \in U \subset M$, suppose $\{e_i\} \subset \Gamma(TU)$ is a local frame and $\{\delta^i\}$ its dual. Suppose $X = a^i e_i$ and $Y = b^j e_j$ are two vector fields, then

$$D_X Y = a^i D_i(b^j e_j) = a^i(\partial_i b^j e_j + b^j D_i e_j),$$

where use the notation $D_i = D_{e_i}$ and $\partial_i = \partial_{e_i}$. Thus the linear connection $D$ is completely determined by $D_i e_j$. Since $D_i e_j$ is still a tangent vector, we can write

$$D_i e_j = A_{ij}^k e_k, \tag{5.1}$$

where $\{A_{ij}^k\}$ are locally defined smooth functions. Alternatively, we can write

$$De_j = D_i e_j \otimes \delta^i = A_{ij}^k \delta^i \otimes e_k.$$

In this way, we locally identify $D$ with a matrix-valued *connection 1-form*

$$A = A_i \delta^i = (A_{ij}^k)\delta^i$$

where $A_i = (A_{ij}^k)_{1 \le j,k \le n}$ is $n \times n$ matrix for each $i = 1, \cdots, n$. Thus locally we can simply denote $D = d + A$ such that

$$DX = D(X^i e_i) = (dX^k + X^j A_{ij}^k \delta^i) \otimes e_k = (d + A)X.$$

Next we study the transformation law of connection 1-forms under a change of local frames. Suppose for a different choice of local frame $\bar{e}_j = s^i_j e_i$, the connection 1-form of $D$ becomes $\bar{A} = \bar{A}_i \bar{\delta}^i$. If we write in matrix form $\bar{e} = Se$ where $S = (s^i_j) \in GL(n)$, then

$$D\bar{e} = \bar{A}\bar{e} = D(Se) = dSe + SAe = dS \circ S^{-1}\bar{e} + S \circ A \circ S^{-1}\bar{e}.$$

It follows

$$\bar{A} = -S \circ dS^{-1} + S \circ A \circ S^{-1}. \tag{5.2}$$

Therefore a linear connection $D$ is equivalent to a set of local connection 1-forms $\{A_\alpha\}$ subordinating to an open cover $\{U_\alpha\}$, which satisfies the above transformation law (5.2) on overlap area. In particular, the connection form $A$ is *not* a globally defined tensor, but the difference of two connection forms is.

**Remark 5.1.3** *Note that the locally expressions of $D$ can vary for different types of tensors. For example, since*

$$0 = D_i(\delta^k(e_j)) = (D_i\delta^k)(e_j) + \delta^k(D_i e_j),$$

*we get for 1-forms,*

$$D_i\delta^k = -A^k_{ij}\delta^j.$$

### 5.1.3 Parallel transport

Now we explain the relation of linear connections with the classical directional derivatives.

Let $D$ be a linear connection on $M$. Suppose $\gamma$ is a curve on $M$ with $\gamma(0) = p \in M$, and $X \in \mathfrak{X}(M)$ is a vector field. We say $X$ is *parallel* along $\gamma$ if

$$\nabla_t X = \nabla_{\gamma'} X(\gamma(t)) = 0.$$

In local coordinates, if $X(t) = a^i(t)\partial_i$ and $\gamma'(t) = b^j\partial_j$, then we require

$$\nabla_t X = \left(\frac{da^i}{dt} + a^j b^k \Gamma^i_{jk}\right)\partial_i = 0.$$

This is a first-order ordinary equation and have a unique solution for any initial data $X(0) = v \in T_pM$.

Now given a curve $\gamma$ starting from $p \in M$, we can define the *parallel transport*

$$P_t : T_pM \to T_{\gamma(t)}M, \quad v \mapsto P_t(v)$$

by requiring $P_t(v)$ to be parallel to $v$ along $\gamma$. Obviously, $P_t$ is a linear in view of the ODE theory. Moreover, the parallel transport along $\gamma$ in the opposite direction gives an inverse of $P_t$. Thus $P_t$ is a linear isomorphism. Note the dependence of $P_t$ on the curve $\gamma$ here, which is essentially different from the Euclidean space.

**Theorem 5.1.4** *The linear connection $D$ satisfies*

$$D_vY(x) = \frac{d}{dt}\Big|_{t=0} P_t^{-1} \circ Y(t) = \lim_{t\to 0} \frac{P_t^{-1} \circ Y(\gamma(t)) - Y(p)}{t}. \tag{5.3}$$

*Proof.* Let $\gamma$ be a smooth curve with $\gamma(0) = x$ and $\gamma'(0) = v$. Let $\{e_i\} \subset T_xM$ be a set of frame at $x$. By parallel transport along $\gamma$, we get a moving frame $\{\bar{e}_i(t) := P_t(e_i)\}$ on $\gamma$ which is satisfies $D_t\bar{e}_i = 0$.

Now suppose $Y(t) = b^i(t)\bar{e}_i(t)$ along $\gamma$. Then we have

$$D_vY(x) = D_t(b^i\bar{e}_i)|_{t=0} = \partial_t b^i(0)e_i.$$

One the other hand, we have

$$P_t^{-1} \circ Y(t) = b^i(t)P_t^{-1}(\bar{e}_i(t)) = b^i(t)e_i.$$

Therefore, we get

$$\frac{d}{dt}\Big|_{t=0} P_t^{-1} \circ Y(t) = \partial_t b^i(0)e_i,$$

and the theorem follows. □

**Remark 5.1.5** *Actually, given a parallel transportation map which satisfies suitable conditions, then we can define the linear connection by (5.3). (Cf. Wu)*

### 5.1.4 Torsion tensor

First recall that the *Hessian* of a smooth function $f$ on $\mathbb{R}^n$ is simply defined as the symmetric matrix

$$D^2 f = \left(\frac{\partial^2 f}{\partial x^i \partial x^j}\right).$$

The Hessian of a function $f \in C^\infty(M)$ with regard to a linear connection $D$ is defined by the $(0,2)$-tensor $D^2 f = D \circ Df$. Given $X, Y \in \mathfrak{X}(M)$, by the Leibniz's rule,

$$D_Y(Df(X)) = (D_Y Df)(X) + Df(D_Y X).$$

It follows

$$D^2 f(X,Y) := (D_Y Df)(X) = D_Y D_X f - D_{D_Y X} f = Y \circ X(f) - (D_Y X)(f).$$

Interchanging $X$ and $Y$, we get

$$D^2 f(Y,X) = X \circ Y(f) - (D_X Y)(f).$$

The difference of the above two terms is

$$D^2 f(X,Y) - D^2 f(Y,X) = (D_X Y - D_Y X - [X,Y])f.$$

Thus we are led to define the operator $T : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ by

$$T(X,Y) = D_X Y - D_Y X - [X,Y],$$

it is easy to verify that $T$ is a skew-symmetric $(1,2)$-tensor. We call $T$ the *torsion tensor* of the linear connection $D$. Obviously, the Hessian $D^2 f$ is symmetric iff the torsion tenor $T$ vanishes.

**Remark 5.1.6** *Note that the following Leibniz's rule is false*

$$D_Y(X(f)) = Y \circ X(f) \neq (D_Y X)f + X(D_Y(f)).$$

*In fact,*

$$D_Y(X(f)) = D_Y(Df, X) = (D_Y Df, X) + (Df, D_Y X).$$

**Exercise 5.1.7** *In a local frame $\{e_i\}$, find the local expression of the torsion tensor $T$ in terms of the connection coefficients $\{A_{ij}^k\}$ of $D$.*

## 5.2 Levi-Civita connection

### 5.2.1 Definition

There are infinitely many (even torsion free) linear connections on a given manifold. If the manifold is endowed with a Riemannian metric, then we will see that there is a canonical linear connection on the Riemannian manifold.

**Definition 5.2.1** *The Levi-Civita connetion of a Riemannian manifold $(M, g)$ is a linnear connection $\nabla$ satisfying*

*1) (torsion free) $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] = 0, \forall X, Y \in \mathfrak{X}(M)$,*

*2) (compatible with the metric) $\nabla g = 0$.*

Note that the torsion-free property gives

$$\nabla_X Y - \nabla_Y X = [X, Y].$$

And compatibility with $g$ is equivalent to

$$\nabla_X \langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle,\ \forall X, Y, Z \in \mathfrak{X}(M).$$

Equivalently, if we choose an orthonormal frame $\{e_i\}$ such that $\nabla_i e_j = A_{ij}^k e_k$, then

$$0 = \nabla_i \langle e_j, e_k\rangle = \langle \nabla_i e_j, e_k\rangle + \langle e_j, \nabla_i e_k\rangle = A_{ij}^k + A_{ik}^j.$$

Thus $A_{ij}^k = -A_{ik}^j$ is a skew-symmetric matrix for each $i$. In other words, the connection form in an orthonormal frame is a $\mathfrak{so}(n)$-valued 1-form.

The Levi-Civita connection is particularly important because of the following theorem.

**Theorem 5.2.2** *There exists a unique Levi-Civita connection $\nabla$ on every Riemannian manifold $(M.g)$.*

*Proof.* It suffices to verify that a Levi-Civita connection is determined by the formula

$$\begin{aligned} 2\langle \nabla_X Y, Z\rangle =& X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle \\ & + \langle [X,Y], Z\rangle - \langle [Y,Z], X\rangle + \langle [Z,X].Y\rangle. \end{aligned} \tag{5.4}$$

□

As before, the Levi-Civita connection and the covariant derivative can be extended to any tensor

$$\nabla : \mathscr{T}(M) \to \mathscr{T}(M) \otimes \Lambda^1(M), \quad T \mapsto \nabla T = \nabla_i T \otimes \delta^i.$$

**Exercise 5.2.3** *Let $(M, g)$ be a Riemannian manifold and $D$ be a linear connection. Show that the parallel transport induced by $D$ is an isometry if and only if $D$ is compatible with $g$.*

### 5.2.2 Christoffel symbols

In local coordinates, the connection coefficients of the Levi-Civita connection in natural frame $\{\partial_i := \frac{\partial}{\partial x^i}\}$, which is also known as the *Christoffel symbols*, are given by

$$\nabla_i \partial_j = \Gamma_{ij}^k \partial_k.$$

Then for 1-forms, we have

$$\nabla_i dx^k = -\Gamma_{ij}^k dx^j.$$

By torsion-freeness, we have

$$0 = [\partial_i, \partial_j] = \nabla_i \partial_j - \nabla_j \partial_i = (\Gamma_{ij}^k - \Gamma_{ji}^k)\partial_k.$$

By compatibility with the metric, we have

$$\partial_i g_{jk} = \nabla_i \langle \partial_j, \partial_k\rangle = \langle \nabla_i \partial_j, \partial_k\rangle + \langle \partial_j, \nabla_i \partial_k\rangle = g_{kl}\Gamma_{ij}^l + g_{jl}\Gamma_{ik}^l.$$

It follows that the Christoffel symbol is symmetric for fixed $k$, i.e. $\Gamma_{ij}^k = \Gamma_{ji}^k$ and

$$\Gamma_{ij}^k = \frac{1}{2} g^{kl} \left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\right). \tag{5.5}$$

Note that the connection 1-form $\Gamma = \Gamma_{ij}^k dx^i$ changes by the transformation law (5.2) under changes of coordinates.

Using the Christoffel symbol, we can write down the Hessian of a function

$$\nabla^2 f = (\partial_i \partial_j f - \Gamma_{ij}^k \partial_k f) dx^i \otimes dx^j.$$

Taking the trace of the Hessian, we get the *Laplacian*

$$\Delta f = \mathrm{tr}_g \nabla^2 f = g^{ij}(\partial_i \partial_j f - \Gamma_{ij}^k \partial_k f).$$

### 5.2.3 Examples

Here we present some basic examples of Riemannian manifolds. Note that the metric and the Levi-Civita connection has different coefficients in different coordinates.

**Example 5.2.4** *For the two dimensional Euclidean space $\mathbb{R}^2$, the Levi-Civita connection is just the standard derivative. In the Cartesian coordinates we have $g_0 = dx^2 + dy^2$ and the Christoffel symbol vanishes. However, if we choose the polar coordinates $(r, \theta)$, then the metric becomes $g_0 = dr^2 + r^2 d\theta^2$. With a simple computation by using (5.5), it follows that the non-vanishing Christoffel symbols are*

$$\Gamma_{r\theta}^\theta = \Gamma_{\theta r}^\theta = \frac{1}{r},\ \Gamma_{\theta\theta}^r = -r.$$

*Hence the Laplacian of a smooth function $f$ is*

$$\Delta f = \partial_r^2 f + \frac{1}{r}\partial_r f + \frac{1}{r^2}\partial_\theta^2 f.$$

Generally, if a metric is given by $g = dr^2 + \phi^2(r) d\theta^2$, which is sometimes called *wrapped metric*. Then the Christoffel symbols are given by

$$\Gamma_{r\theta}^\theta = \Gamma_{\theta r}^\theta = \frac{\phi'}{\phi},\ \Gamma_{\theta\theta}^r = -\phi'\phi.$$

Hence the Laplacian of a smooth function $f$ is

$$\Delta f = \partial_r^2 f + \frac{\phi'}{\phi}\partial_r f + \frac{1}{\phi^2}\partial_\theta^2 f.$$

**Remark 5.2.5** *Here is another explanation for the above example. It is easy to see that* $X = \frac{1}{\phi}\partial_\theta$ *is parallel along* $\partial_r$*. Thus*

$$\nabla_r \partial_\theta = \nabla_r(\phi X) = \phi' X = \frac{\phi'}{\phi}\partial_\theta.$$

*On the other hand,*

$$\Gamma^r_{\theta\theta} = \langle \nabla_\theta \theta, \partial_r \rangle = -\langle \partial_\theta, \nabla_\theta \partial_r \rangle = -\phi^2 \Gamma^\theta_{\theta r}.$$

**Definition 5.2.6** *Let* $(M, g)$ *be a Riemannian manifold, a metric* $h$ *is called (point-wisely)* conformal *to* $g$ *if* $h = \lambda g$ *for some positive function* $\lambda \in C^\infty(M)$.

**Definition 5.2.7** *Let* $(M, g)$ *and* $(N, h)$ *be two Riemannian manifolds and* $F : M \to N$ *be a map.* $F$ *is called a* conformal map *if the pull-back metric* $F^*h = \lambda g$ *is conformal to* $g$.

Now suppose $h = \lambda g$ is conformal to $g$ and denote the Christoffel symbols of $h$ and $g$ by $\tilde{\Gamma}^k_{ij}$ and $\Gamma^k_{ij}$ respectively. It is easy to verify that

$$\tilde{\Gamma}^k_{ij} = \Gamma^k_{ij} + \frac{1}{2}(\delta^k_j \partial_i \lambda + \delta^k_i \partial_j \lambda - g_{ij} g^{kl} \partial_l \lambda).$$

# 5.3 Geodesics

## 5.3.1 Definition and examples

**Definition 5.3.1** *A smooth curve $\gamma : [a,b] \to M$ is called a geodesic if it satisfies the equation*

$$\nabla_{\gamma'}\gamma' = 0.$$

Locally, set $\gamma(t) = (x^1(t), \cdots, x^n(t))$, the equation of geodesic has the form

$$\frac{d^2x^k}{dt^2} + \Gamma_{ij}^k \frac{dx^i}{dt}\frac{dx^j}{dt} = 0,$$

where $\Gamma_{ij}^k$ are the Christoffel symbols. This is a second order ODE system. By standard ODE theory, it has a unique local solution $\gamma$ on an interval $[0,T)$ if it is supplemented by initial data

$$\gamma(0) = p, \gamma'(0) = v \in T_pM.$$

Moreover, the solution $\gamma$ and the survival time $T$ depends smoothly on the initial data.

**Lemma 5.3.2** *Suppose $\gamma$ is a smooth curve on $(M,g)$, and $V, W$ are parallel vector fields along $\gamma$, then*

$$\frac{d}{dt}\langle V, W\rangle = 0.$$

*Proof.* Since $V, W$ are parallel, $\nabla_{\gamma'}V = \nabla_{\gamma'}W = 0$. It follows from the compatibility of Levi-Civita connection that

$$\frac{d}{dt}\langle V, W\rangle = \langle \nabla_{\gamma'}V, W\rangle + \langle V, \nabla_{\gamma'}W\rangle = 0.$$

□

**Corollary 5.3.3** *If $\gamma$ is a geodesic, then $\gamma'$ is parallel along $\gamma$ and $|\gamma'| = c$ is constant.*

From the above proposition, we can see that if a curve is a geodesic, then its parametrization has to be proportional to its arclength parameter. But this does not add extra constraints on the curve, since any curve can be parameterized by its arclength. In fact, for $\gamma : [a,b] \to M$, we can let

$$s := a + \int_a^t |\gamma'| dt.$$

Then under the new parameter $s$, $\gamma$ is a map $\gamma : [0, l] \to M$ with $l$ its length, and

$$\left|\frac{d\gamma}{ds}\right| = \left|\frac{d\gamma}{dt}\right|\left|\frac{dt}{ds}\right| = 1.$$

**Example 5.3.4** 1. *Euclidean space* $(\mathbb{R}^n, g_0)$*:*

*Since the standard connection $D$ on $\mathbb{R}^n$ is trivial, the Christoffel symbols vanish, and the equation is reduced to*

$$\frac{d^2x^k}{dt^2} = 0, x^i(0) = b^i, \frac{dx^i}{dt}(0) = a^i.$$

*It has a unique solution which is just a linear function $\gamma(t) = a \cdot x + b$. Thus geodesics in $\mathbb{R}^n$ are just straight lines.*

2. *Torus* $(T^n = \mathbb{R}^n/(\mathbb{Z}^2)^n, g_0)$*:*

   *The metric are still flat and geodesics are straight lines. However, since for each $p, q \in T^n$, there are infinitely many preimages in the covering space $\mathbb{R}^n$, there are infinitely many lines which connects $p$ and $q$.*

3. *Sphere* $(S^n, g_0)$*:*

   *Using the embedding $S^n \subset \mathbb{R}^{n+1}$, the Levi-Civita connection on $S^n$ can be expressed by*

$$\nabla_X Y = (D_X Y)^\top, \forall X, Y \in (X)(S^n).$$

   *Then one can easily verify that geodesics on $S^n$ are great circles. Observe that there are infinitely many great circles connecting antipodal points $p, -p \in S^n$. Otherwise, for general two points $p, q \in S^n$, there is a unique great circle connecting $p$ and $q$, which is divided into two segments.*

4. *Hyperbolic space $(\mathbb{H}, g)$ (where $\mathbb{H}$ is the half plane and $g = \frac{1}{(x^n)^2}\sum_{i=1}^n (dx^i)^2$):*

   *Geodesics in the half plane model are semi-circles which are orthogonal to the hyperplane $\{x^n = 0\}$.*

### 5.3.2 First variation formula

We have seen that geodesics on Riemannian manifolds are natural generalization of straight lines in Euclidean spaces. In fact, we will show that the shortest curve between two points are geodesics.

Given a smooth curve $\gamma : [a, b] \to M$, its length is defined by

$$L(\gamma) = \int_a^b |\gamma'| dt.$$

Then we define the distance of $p, q \in M$ by

$$dist(p, q) = \inf\{L(\gamma) | \gamma : [a, b] \to M, \gamma(a) = p, \gamma(b) = q\}.$$

Thus if $\gamma$ attains the distance of $p, q$, then it must be a critical point of the length functional $L$.

Thus we choose a variation $\beta : [a, b] \times (-\epsilon, \epsilon) \to M$, i.e. is a family of smooth curves $\gamma_s(t) = \beta(t, s) : [a, b] \to M$. Then we compute

$$\begin{aligned}
\frac{d}{ds} L(\gamma_s) &= \frac{d}{ds} \int_a^b |\partial_t \beta| dt \\
&= \int_a^b \frac{1}{|\partial_t \beta|} \langle \nabla_s \nabla_t \beta, \nabla_t \beta \rangle dt \\
&= \int_a^b \frac{1}{|\partial_t \beta|} \langle \nabla_t \nabla_s \beta, \nabla_t \beta \rangle dt \\
&= \int_a^b \frac{1}{|\partial_t \beta|} \left( \partial_t \langle \nabla_s \beta, \nabla_t \beta \rangle - \langle \nabla_s \beta, \nabla_t \nabla_t \beta \rangle \right) dt
\end{aligned}$$

Therefore, if $|\gamma'| = l$ is constant, and we denote $V := \partial_s \beta|_{s=0}$, which is a vector fields along $\gamma$, then

$$\left.\frac{d}{ds}\right|_{s=0} L(\gamma_s) = \frac{1}{l} \left( \langle V, \gamma' \rangle|_a^b - \int_a^b \langle V, \nabla_{\gamma'} \gamma' \rangle dt \right).$$

**Theorem 5.3.5** *Suppose $\gamma : [0, l] \to M$ is a shortest curve connecting $q, a \in M$ and is parameterized by its arclength, then $\gamma$ is a geodesic.*

*Proof.* Let $\gamma_s(t) = \beta(t,s) : [0,l] \times (-\epsilon,\epsilon) \to M$ be a family of smooth curves connecting $p$ and $q$. Then $\beta(0,s) = p, \beta(l,s) = q$, hence $V(0) = 0, V(1) = 0$. Then by the first variational formula, we have

$$\frac{d}{ds}\Big|_{s=0} L(\gamma_s) = \int_a^b \langle V, \nabla_{\gamma'}\gamma' \rangle dt = 0.$$

Since the equality holds for any $V$, it follows $\nabla_{\gamma'}\gamma' = 0$. □

## 5.4 Exponential map

### 5.4.1 Definition of exponential map

Suppose for $v \in T_pM$, there is a radial geodesic $\gamma : [0,1] \to M$ such that $\gamma(0) = p, \gamma'(0) = v$, then we can define a map from $v$ to $\gamma(1) \in M$. The following lemma, which is a direct conclusion from ODE theory, says this map is always well-defined in a small neighborhood.

**Lemma 5.4.1** *Given a fixed point $p_0 \in M$, there exists a neighborhood $U$ and a small number $\epsilon > 0$, such that for all $p \in U$ and $v \in T_pM$, $|v| \leq \epsilon$, there is a unique geodesic $\gamma_v : (-\epsilon, \epsilon) \to M$ satisfying $\gamma_v(0) = p$ and $\gamma_v'(0) = v$.*

**Definition 5.4.2** *For each $p \in M$, the exponential map $\exp_p : U \to M$ is a map defined on a neighborhood $U \subset T_pM$ such that $\exp_p(v) = \gamma_v(1), \forall v \in U$.*

An easy observation is that (using uniqueness of geodesics)

$$\exp_p(tv) = \gamma_{tv}(1) = \gamma_v(t), \forall t \in [0,1], v \in T_pM.$$

By the above lemma, the exponential map is well-defined in a small $\epsilon$-ball centered at the origin in $T_pM$. Therefore, we can use the exponential map to define a *normal coordinate* in a neighborhood of $p$. Namely, we have a local chart $(U, \exp_p^{-1}; y^i)$ where $(y^1, \cdots, y^n)$ is the Cartesian coordinates on $T_pM$.

However, $\epsilon$ is dependent on the base point $p$. Also, in general it is not known if it exists in a large domain. A simple example is the incomplete space $\mathbb{R}^n \setminus \{O\}$.

**Example 5.4.3** *1. exponential map for matrices.*

$$\exp_I(A) = \exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k, \forall A \in M_n.$$

*2. exponential map for $S^1$.*

$$\exp_0 x = \exp(ix), \forall x \in \mathbb{R}^1.$$

### 5.4.2 Gauss lemma

**Lemma 5.4.4** *The radial geodesics from a point $p \in M$ are orthogonal to distance spheres around $p$.*

*Proof.* It suffices to show that the radial geodesic is orthogonal to any curve on the distance sphere.

Let $v(s)$ be a curve on the distance sphere of radius $a > 0$, namely, $|v(s)| = a$. Define $f(r,s) = \exp_p(rv(s))$. Then by definition, $\gamma_s(r) = f(r,s)$ is a radial geodesic and $\nabla_r f = \gamma_s' = v(s)$. Then we compute

$$\begin{aligned}\partial_r \langle \nabla_r f, \nabla_s f\rangle &= \langle \nabla_r \nabla_r f, \nabla_s f\rangle + \langle \nabla_r f, \nabla_f \nabla_s f\rangle \\ &= \langle \nabla_r f, \nabla_s \nabla_r f\rangle = \frac{1}{2}|\nabla_r f|^2 = 0.\end{aligned}$$

Thus $\langle \nabla_r f, \nabla_s f\rangle$ is constant along $r$.

On the other hand, at the origin where $r = 0$, since $f(0,s) = p$, we have $\nabla_s f(0,s) = 0$. Therefore,

$$\langle \nabla_r f, \nabla_s f\rangle(r,s) = \langle \nabla_r f, \nabla_s f\rangle(0,s) = 0.$$

□

As a consequence of the Gauss lemma, we can find a *geodesic polar coordinate* around a point $p \in M$ such that

$$g = dr^2 + \sum_{i,j=1}^{n-1} g_{ij}(r,\theta)d\theta^i d\theta^j.$$

**Lemma 5.4.5** *Suppose $B_\epsilon(0) \to M$ is an embedding, then*

1. *For all $v \in B_\epsilon(0)$, let $\gamma_v(t) = \exp_p(tv) : [0,1] \to M$ and $q = \gamma_v(1)$. Then $\gamma_v$ is the unique curve satisfying*
$$L(\gamma_v) = d(p,q) = |v|.$$
*In particular, $\gamma_v$ is the unique geodesic connecting $p$ and $q$.*

2. *For all $q \notin B_\epsilon(p)$, then exists $q' \in \partial B_\epsilon(0)$ such that*
$$d(p,q) = \epsilon + d(q',q) = d(p,q') + d(q',q).$$

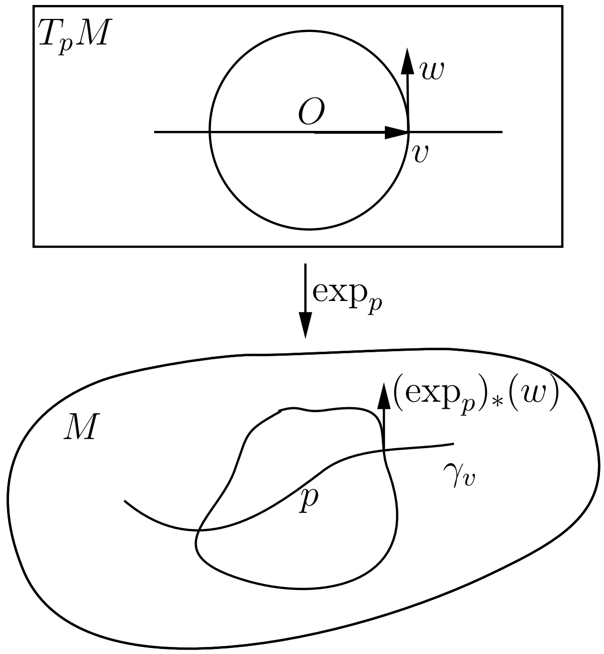


Figure 5.1: Gauss Lemma

*Proof.* (1) Let $\sigma : [0,1] \to M$ be a curve connecting $p$ and $q$. Denote the distance function $r(x) = d(p,x)$, . Since $|\nabla r| = |\partial_r| = 1$, we have

$$|\sigma'| \geq \langle \sigma', \nabla r \rangle = \frac{d}{dt} r \circ \sigma(t).$$

It follows □

### 5.4.3 Hopf-Rinow theorem

### 5.4.4 Normal coordinates

In Euclidean space, there is a natural choice of coordinates, namely, the Cartesian coordinate, by fixing a set of orthonormal basis at the origin. Then in this Cartesian coordinates, lines are represented by linear functions. Here we want to find an analog in a Riemannian manifold. It turns out that the exponential map provides us a natural choice.

**Normal coordinates**

Let $(M, g)$ be a Riemannian manifold and $p \in M$. Suppose $U \in T_pM$ is a neighborhood of the origin such that $\exp_p : U \to V$ is a diffeomorphism for some neighborhood $V \subset M$ of $p$. The tangent space is linear space and we can define a Cartesian coordinate by fixing an orthonormal basis $\{e_i\}_{i=1}^n \subset T_pM$. Namely, we define a linear map $\sigma : T_pM \to \mathbb{R}^n$ by letting $\sigma(e_i) = (0, \cdots, 1, \cdots, 0)$. Then for any $v = v^i e_i \in T_pM$, we have$\sigma(v) = (v^1, \cdots, v^n)$. Set $\phi := \sigma \circ \exp_p^{-1}$, then $(V, \phi; x^i)$ is a local chart around $p$, which is called the *normal coordinate.*

**Lemma 5.4.6** *If $g(x) = g_{ij}(x)dx^i dx^j$ in the normal coordinate, then at the base point $p$,*

1. $g_{ij}(0) = \delta_{ij}$;
2. $\Gamma_{ij}^k(0) = 0$;
3. $\partial_k g_{ij}(0) = 0$.

*Proof.*

1. $g_{ij}(0) = g(\partial_i, \partial_j) = g(e_i, e_j) = \delta_{ij}$.
2. In the normal coordinate, for any $v = v^i e_i \in T_pM$, the radial geodesic $\gamma_v(t) = \exp_p(tv)$ has the expression $\phi \circ \gamma_v(t) = (tv^1, \cdots, tv^n)$. Using the equation of geodesics, we have
$$0 = \nabla_t^2 \gamma = \frac{d^2 x^k}{dt^2} + \Gamma_{ij}^k(x) \frac{dx^i}{dt} \frac{dx^j}{dt}.$$
Since $x^i = tv^i$ is linear, we get $\Gamma_{ij}^k(tv)v^i v^j = 0$. Evaluating at $t = 0$ we have $\Gamma_{ij}^k(0)v^i v^j = 0$. Since it holds for all $v$, we get $\Gamma_{ij}^k(0) = 0$.
3. It follows directly from the formula
$$\partial_k g_{ij} = \partial_k g(\partial_i, \partial_j) = g(\nabla_k \partial_i, \partial_j) + g(\partial_i, \nabla_k \partial_j) = \Gamma_{ki}^l g_{lj} + \Gamma_{kj}^l g_{il}.$$

□

It follows from the above lemma and Taylor expansion that, in the normal coordinates
$$g(x) = g_{\mathbb{R}^n} + O(|x|^2).$$

More precisely, we have

$$g_{ij}(x) = \delta_{ij} + \frac{1}{2}\partial_k\partial_l g_{ij} x^k x^l + O(|x|^3).$$

We will see that the second order term is determined by the curvature tensor in Section...

### Geodesic polar coordinate

If we choose the polar coordinate on $T_pM$ instead of the Cartesian coordinate, then we get the geodesic polar coordinate. More precisely, let $\bar{\sigma} : T_pM\backslash\{0\} \to \mathbb{R}^1_+ \times S^{n-1}, \bar{\sigma}(v) = (r, \theta)$ where $r = |v|$ and $\theta = (\theta^1, \cdots, \theta^{n-1})$ are standard spherical coordinates in $S^{n-1}$. Let $\bar{\phi} = \bar{\sigma} \circ \exp_p^{-1}$, then $(V, \bar{\phi}; (r, \theta))$ is a local chart around $p$ and is called the *geodesic polar coordinate.*

**Lemma 5.4.7** *In the geodesic polar coordinates, the metric has the form*

$$g(r, \theta) = dr^2 + \sum_{i,j=1}^{n-1} g_{ij}(r, \theta) d\theta^i d\theta^j.$$

*Proof.* It suffices to prove that $g_{r\theta^i} = 0$ and $g_{rr} = 1$. The first identity follows directly from the Gauss Lemma. For the second one, observe that any point $q = \exp_p(v) \in V$ where $v = \bar{\sigma}^{-1}(r_0, \theta_0)$, let $e = v/|v|$ be the corresponding unit vector and $\gamma_e(t) = \exp_p(te)$ be the radial geodesic. Then the tangent vector $\partial_r$ at $q$ is just the derivative of $\gamma_e$ at $t = r_0$. Therefore, we have

$$|\partial_r| = |\gamma_e'(r_0)| = |\gamma_e'(0)| = |(\exp_p)_*|_0(e)| = |e| = 1.$$

□

A important consequence of the lemma is $\nabla r = \partial_r$. Since $r(q) = d(q) = dist(p, q)$ we know that $|\nabla d| = |\partial r| = 1$ (in a injective area).

From the fact that $g(x) = g_{\mathbb{R}^n} + O(|x|^2)$ and $g_{\mathbb{R}^n} = dr^2 + r^2 g_{S^{n-1}}$, we find that the second term in the above has the form

$$\sum_{i,j=1}^{n-1} g_{ij}(r, \theta) d\theta^i d\theta^j = r^2 g_{S^{n-1}} + O(r^2).$$

## 5.5 Curvature

### 5.5.1 Definition and symmetries of curvature tensor

Recall that the double covariant derivative is defined by

$$\nabla^2_{X,Y} = \nabla_x \nabla_Y - \nabla_{\nabla_X Y}$$

and is applicable to any tensor. Since the Levi-Civita is torsion free, the double derivative of a function commutes, i.e.

$$\nabla^2_{X,Y} f = \nabla^2_{Y,X} f, \forall X, Y \in \mathscr{X}(M).$$

However, this is not the case for general tensors, and the curvature is just the communicator.

**Definition 5.5.1** *The curvature of a connection is a* $(1,3)$*-tensor defined by*

$$R \left| \begin{array}{ll} \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) & \to \mathfrak{X}(M) \\ (X,Y,Z) & \mapsto R(X,Y)Z = \nabla^2_{X,Y} Z - \nabla^2_{Y,X} Z \end{array} \right.$$

*By definition of the double derivative, it is equivalent to define*

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

It is easy to verify that $R$ is indeed a tensor. By raising the index, it also gives a $(0,4)$-tensor. However, it differs from the $(1,3)$-tensor by a minus sign.

**Definition 5.5.2** *The curvature* $(0,4)$*-tensor is*

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

In local coordinates, we set $R(\partial_i, \partial_j)\partial_k = R^l_{ijk} \partial_l$ and

$$R_{ijkl} = R(\partial_i, \partial_j, \partial_k, \partial_l) = -\langle R(\partial_i, \partial_j)\partial_k, \partial_l \rangle = -R^m_{ijk} g_{ml}.$$

Pay attention to the minus sign here, or one will get negative curvature for spheres. By the definition of the curvature, one can verify that

$$R^l_{ijk} = \partial_i \Gamma^l_{jk} - \partial_j \Gamma^l_{ik} + \Gamma^l_{im}\Gamma^m_{jk} - \Gamma^l_{jm}\Gamma^m_{ik}. \tag{5.6}$$

In particular, in normal coordinates of a fixed point $p \in M$, we have $g_{ij}(p) = \delta_{ij}, \partial_k g_{ij}(p) = 0, \Gamma^k_{ij}(p) = 0$. Recall that

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl}(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}),$$

it follows that at $p$,

$$-R_{ijkl} = R^l_{ijk} = \partial_i \Gamma^l_{jk} - \partial_j \Gamma^l_{ik} = \frac{1}{2}(\partial_i\partial_k g_{jl} - \partial_i\partial_l g_{jk} - \partial_j\partial_k g_{il} + \partial_j\partial_l g_{ik}). \tag{5.7}$$

Thus $R$ is completely determined by the second order derivatives of $g_{ij}$. In fact, the converse is also true. Namely, we have

$$\frac{1}{2}\partial_k\partial_l g_{ij} = \frac{1}{3}(R_{kijl} + R_{lijk}).$$

This is exactly the way Riemann originally defined the curvature tensor.

The the symmetry of $R$ is obvious from the expression (5.7), namely, we have

$$R_{ijkl} = -R_{jikl} = -R_{ijlk} = R_{klij}$$

and the first Bianch identity

$$R^l_{ijk} + R^l_{jki} + R^l_{kij} = 0. \tag{5.8}$$

Or equivalently, we have

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

and

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

Actually, the covariant derivative $\nabla R$ also another symmetry. To see this, we compute in normal coordinate that

$$\nabla_j R^l_{jkm} = \partial_i \partial_j \Gamma^l_{km} - \partial_i \partial_k \Gamma^l_{jm}.$$

Thus we get the second Bianchi identity:

$$\nabla_i R^l_{jkm} + \nabla_j R^l_{kim} + \nabla_k R^l_{ijm} = 0. \tag{5.9}$$

or equivalently,

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

### 5.5.2 Other curvatures

The *Ricci curvature* is a $(1,1)$-tensor defined by taking trace of $R$, i.e. for an orthonormal basis $\{E_i\}$,

$$Rc(X) = \sum_{i=1}^n R(X, E_i)E_i.$$

It also gives a symmetric $(0,2)$-tensor by

$$Ric(X,Y) = \sum_{i=1}^n R(E_i, X, E_i, Y) = \langle Ric(X), Y \rangle.$$

In local coordinates,

$$Ric = R_{ij} = g^{kl} R_{kilj} = R^k_{kij}, Rc = R^j_i = g^{jk} R_{ik}.$$

The *scalar curvature* is the trace of Ricci curvature, i.e.

$$R = \sum_{i=1}^n Ric(E_i, E_i) = \sum_{i,j=1}^n R(E_i, E_j, E_i, E_j).$$

In local coordinates,

$$R = \operatorname{tr} Ric = R^i_i = g^{ij} R_{ij} = g^{ij} g^{kl} R_{ikjl}.$$

Taking trace w.r.t. $i$ and $l$ in (5.9), we get

$$\nabla_i R^i_{jkm} - \nabla_j R_{km} + \nabla_k R_{jm} = 0.$$

Taking another trace w.r.t. $k$ and $m$, we get

$$\nabla_i R^i_j - \nabla_j R + \nabla_k R^k_j = 0.$$

Here we used the fact that the covariant derivative commutes with trace, i.e. $\nabla \circ \mathrm{tr} = \mathrm{tr} \circ \nabla$. Therefore, we arrive at

$$\mathrm{div} Ric = \frac{1}{2} dR.$$

The *sectional curvature* assigns a real number to each 2-plane in the tangent space by

$$S(X,Y) = \frac{R(X,Y,X,Y)}{|X \wedge Y|^2} = \frac{\langle R(X,Y)Y, X\rangle}{|X|^2|Y|^2 - \langle X,Y\rangle^2}.$$

Obviously, $S(X,Y)$ only depends on the 2-plane $\Pi$ spanned by $X, Y$. Evidently, it is the Gauss curvature of the 2-submanifold that is tangent to $\Pi(X,Y)$. In fact, the sectional curvatures determines the full curvature tensor.

**Proposition 5.5.3** *Let $R$ and $R'$ be two $(0,4)$-curvature tensors which satisfies $S(\Pi) = S'(\Pi)$ for all 2-planes $\Pi \subset T_pM$. Then $R = R'$.*

### 5.5.3 Example

As an example, we compute the curvature of the 2 dimensional manifold with *wrapped product metric* $g = dr^2 + \phi^2(r)d\theta^2$.

**Example 5.5.4** *Suppose $(M,g)$ is a two manifold where $g = dr^2 + \phi^2(r)d\theta^2$ (in geodesic polar coordinate).*

*First we compute the Christoffel symbols. Since $\partial_r$ is parallel with $|\partial_r| = 1$, we have $\Gamma^r_{rr} = \Gamma^\theta_{rr} = 0$. Let $X = \frac{1}{\phi}\partial_\theta$, then $|X| = \frac{1}{\phi^2}\langle\partial_\theta, \partial_\theta\rangle = 1$. By Gauss lemma $\langle\partial_r, X\rangle = 0$, thus $\nabla_r X = 0$ and $X$ is also parallel along $r$. Then we can compute*

$$\nabla_r \partial_\theta = \nabla_r(\phi X) = \phi' X = \frac{\phi'}{\phi}\partial_\theta.$$

*It follows* $\Gamma^\theta_{r\theta} = \phi'/\phi$ *and* $\Gamma^r_{r\theta} = 0$*. On the other hand,*

$$\langle \nabla_\theta \partial_\theta, \partial_\theta \rangle = \frac{1}{2}\partial_\theta \langle \partial_\theta, \partial_\theta \rangle = \frac{1}{2}\partial_\theta \phi^2 = 0,$$

*and*

$$\langle \nabla_\theta \partial_\theta, \partial_r \rangle = -\langle \partial_\theta, \nabla_\theta \partial_r \rangle = -\phi^2 \Gamma^\theta_{\theta r}.$$

*It follows* $\Gamma^\theta_{\theta\theta} = 0$ *and* $\Gamma^r_{\theta\theta} = -\phi'\phi$*.*

*Now we can compute the curvature. By (5.6), we have*

$$\begin{aligned} R^\theta_{r\theta r} &= \partial_r \Gamma^\theta_{\theta r} - \partial_\theta \Gamma^\theta_{rr} + \Gamma^\theta_{rp}\Gamma^p_{\theta r} - \Gamma^\theta_{\theta p}\Gamma^p_{rr} \\ &= \partial_r \Gamma^\theta_{\theta r} + \Gamma^\theta_{r\theta}\Gamma^\theta_{\theta r} \\ &= \partial_r \left(\frac{\phi'}{\phi}\right) + \left(\frac{\phi'}{\phi}\right)^2 \\ &= \frac{\phi''}{\phi}. \end{aligned}$$

*It follows* $R_{r\theta r\theta} = -g_{\theta\theta} R^\theta_{r\theta r} = -\phi''\phi$*. Therefore*

$$S(\partial_r, \partial_\theta) = \frac{R_{r\theta r\theta}}{|\partial_r|^2 |\partial_\theta|^2 - \langle \partial_r, \partial_\theta \rangle^2} = -\frac{\phi''}{\phi}.$$

As a consequence, we now know the curvature of the standard space forms:

1. Euclidean space $(\mathbb{R}^2, g = dr^2 + r^2 d\theta^2)$. Since $\phi(r) = r$, $\phi''(r) = 0$, it has constant curvature 0.

2. Standard sphere $(S^2, g = dr^2 + \sin^2 r d\theta^2)$. Since $\phi(r) = \sin r$, $\phi''(r) = -\sin r$, it has constant curvature 1.

3. Hyperbolic space $(B_1, g = dr^2 + \sinh^2 r d\theta^2)$ (disk model). Since $\phi(r) = \sinh r$, $\phi''(r) = \sinh r$, it has constant curvature $-1$.

# Bibliography


[1] N. Hitchin, *Differentiable manifolds*, Lecture notes(2012).

[2] T. Aubin, *A course in differential geometry*, GSM vol. 27(2001).

[3] 白正国, 沈一兵, 水乃翔, 郭孝英, 《黎曼几何初步(第二版)》, 高等教育出版社(2004).

[4] 陈维桓, 《微分流形初步(第二版)》, 高等教育出版社(2001).

[5] 梅加强, 《流形与几何初步》, 科学出版社(2019).

[6] 伍鸿熙, 沈纯理, 虞言林, 《黎曼几何初步》, 北京大学出版社(2003).