Math Notes

AUTHOR: HSIN

E-MAIL: grammcyao@icloud.com

UPDATE: Sunday 31st December, 2023

Contents

1 Foundation **1**
- 1.1 Sets . . . 1
- 1.2 Russell's paradox . . . 6
- 1.3 Cartesian product . . . 7
- 1.4 Cartesian product . . . 7
- 1.5 Functions . . . 9

2 Differentiation in several variable calculus **11**
- 2.1 Derivatives . . . 11
  - 2.1.1 Directional derivative and differentiability . . . 11
  - 2.1.2 Partial derivative and Jacobian matrix . . . 13
- 2.2 The chain rule . . . 16
  - 2.2.1 Several calculus chain rule . . . 16
  - 2.2.2 The mean-value theorem . . . 18
- 2.3 Continuous differentiability . . . 19
- 2.4 The inverse function theorem . . . 19
  - 2.4.1 The contraction mapping theorem . . . 19
  - 2.4.2 The inverse function theorem . . . 21

3 Exterior algebra **25**
- 3.1 Tensor product . . . 25
  - 3.1.1 Tensor product of two vector spaces . . . 26
  - 3.1.2 Tensor product . . . 27
  - 3.1.3 Tensors . . . 27
- 3.2 Alternating tensors . . . 28
- 3.3 Exterior algebra . . . 30
- 3.A Dual space . . . 30
- 3.B Permutations . . . 32

4 Manifold **35**
- 4.1 Topological manifolds . . . 35
- 4.2 Smooth structures . . . 37

4.2.1 Examples of smooth manifolds 39
4.3 Smooth maps 39
4.3.1 Smooth functions on manifold 39
4.3.2 Smooth maps between manifolds 40
4.3.3 Constructing smooth maps 42
4.4 Diffeomorphisms 43

5 Tangent spaces **45**
5.1 An intuition: Geometric tangent space 45
5.2 Tangent spaces 46
5.3 Applications 48

6 Submersions, immersions, and embeddings **51**
6.1 Maps of constant rank 51

CHAPTER I

Foundation

§1.1 Sets

We present one standard way to define the sets, in terms of the axiomatic set theory.

Axiom 1.1 (Axiom of extension)**.** Two sets A and B are equal, denoted $A = B$, if and only if every element of A is an element of B and vice versa.

The axiom of extension including two fundamental concepts that of set and of element. The axiom of extension assumes that every mathematical object is a set, an element is also an object, it is meaningful to ask whether a set is an element of a set.

If x is an object, we say that x is an element of A, and we write

$$x \in A,$$

otherwise, we write

$$x \notin A.$$

Notice that we always use the *first-order logic* as our meta-language, thus the axiom of extension can be rewritten as

$$\forall A \forall B (\forall x (x \in A \iff x \in B) \iff A = B).$$

But we shall not express the statement in this form.

Since we obtain some fundamental concepts, we can define the *subsets*.

Definition 1.1.1 (Subsets)**.** Let A, B be sets. We say that A is a *subset* of B, denoted $A \subset B$, if and only if every element of A is also an element of B.

Some basic facts about subsets is given by following.

Proposition 1.1.1.1. *Let A, B and C be sets.*

(i) *(Reflexivity)* $A \subset A$.

(ii) *(Transitivity) If* $A \subset B$ *and* $B \subset C$*, then* $A \subset C$.

(iii) *(Anti-symmetry) If* $A \subset B$ *and* $B \subset A$*, then* $A = B$.

Proof. Proof omitted. $\square$

Axiom 1.2 (Axiom scheme of specification)**.** Let A be a set, and for each $x \in A$, let $P(x)$ be a property pertaining to x. Then there exists a set

$$\{x \in A : P(x) \text{ is true}\},$$

whose elements are precisely those elements x in A for which $P(x)$ is true. In other words, for any object y,

$$y \in \{x \in A : P(x) \text{ is true}\} \iff (y \in A \text{ and } P(y) \text{ is true}).$$

This axiom is also known as the *axiom scheme of separation*.

The axiom of specification assures the existence of empty set which without any elements at all.

Definition 1.1.2 (Empty set)**.** If A is a set, the *empty set* $\emptyset$ is defined to be the set

$$\emptyset := \{x \in A : x \neq x\}.$$

Since every object x must obeys $x = x$, thus for arbitrary object x we have $x \notin \emptyset$, the empty set contains no elements. The axiom of extension implies that there can be only one set with no elements.

Lemma 1.1.2.1. *If* $\emptyset_1$ *and* $\emptyset_2$ *are empty set, then* $\emptyset_1 = \emptyset_2$.

Proof. Suppose for sake of contradiction that $\emptyset_1 \neq \emptyset_2$. Then by the axiom of extension, we either have there is an element x such that $x \in \emptyset_1$ but $x \notin \emptyset_2$ or $x \in \emptyset_2$ but $x \notin \emptyset_1$. But both cases are impossible by the definition, a contradiction. $\square$

If a set is not equal to the empty set, we call it *non-empty*. The next Lemma asserts that given any non-empty set A, we are allowed to "choose" an element x of A which demonstrates this non-emptyness.

Lemma 1.1.2.2 (Single choice)**.** *Let A be a non-empty set. Then there exists an object x such that $x \in A$.*

Proof. We prove by contradiction. Suppose there does not exist any object x such that $x \in A$. Then for all objects x, $x \notin A$. Since we have $x \notin \emptyset$. Thus $x \in A$ equivalent to $x \in \emptyset$, and so $A = \emptyset$ by the axiom of extension. $\square$

Another basic fact about the empty set is that it is a subset of every set.

Lemma 1.1.2.3. *If A is a set, then $\emptyset \subset A$. In particular, we have $\emptyset \subset \emptyset$.*

Proof. Suppose for sake of contradiction that $\emptyset \subset A$ is false. This means that there is an element of $\emptyset$ which doesn't lies in A. But it is impossible for that $\emptyset$ contains no elements, a contradiction. In particular, since the emptyset is a set, hence $\emptyset \subset \emptyset$. □

Axiom 1.3 (Axiom of pairing)**.** Let a and b be objects. Then there exists a set $\{a, b\}$ whose only elements are a and b.

If A is a set such that $a \in A$ and $b \in A$ for arbitrary objects a, b, then we can apply the axiom of specification to A with the property "$x = a$ or $x = b$". We obtain a set $\{x \in A : x = a \text{ or } x = b\}$ which contains just a and b, the axiom of extension assures that such a set is unique, such a set is called the *pair*. Similarly, $\{a, a\}$ is an unordered pair, denoted $\{a\}$, and is called the *singleton*.

Axiom 1.4 (Axiom of union)**.** Let A be a set, all of whose elements are sets. Then there exists a set $\bigcup A$ whose elements are precisely those objects which are elements of the elements of A. In other words, for any object x,

$$x \in \bigcup A \iff (x \in S \text{ for some } S \in A).$$

Now we can define some important operations on sets, namely unions, intersections and difference sets.

Definition 1.1.3 (Unions)**.** The *union* $S_1 \cup S_2$ of two sets is defined to be the set

$$S_1 \cup S_2 := \bigcup \{S_1, S_2\}.$$

This equivalent to say that for each $x \in S_1 \cup S_2$ we have $x \in S_1$ or $x \in S_2$. Observe that the existence and uniqueness of $S_1 \cup S_2$ are guaranteed by the axioms of pairing, union, and extension. Here are some basic facts about unions.

Proposition 1.1.3.1 (Basic properties of unions)**.** *Let A, B be sets.*

(i) *(Minimality)* $A \cup \emptyset = A$.

(ii) *(Commutativity)* $A \cup B = B \cup A$.

(iii) *(Associativity)* $A \cup (B \cup C) = (A \cup B) \cup C$.

(iv) *(Idempotence)* $A \cup A = A$.

(v) $A \subset B$ *if and only if* $A \cup B = B$.

Proof. Proof omitted. □

From the axiom scheme of specification, we can define the intersections.

Definition 1.1.4 (Intersections)**.** The *intersection* $S_1 \cap S_2$ of two sets is defined to be the set

$$S_1 \cap S_2 := \{x \in S_1 : x \in S_2\}.$$

Here are some basic facts about unions.

Proposition 1.1.4.1 (Basic properties of intersections)**.** *Let A, B be sets.*

(i) *(Minimality) $A \cap \emptyset = \emptyset$.*

(ii) *(Commutativity) $A \cap B = B \cap A$.*

(iii) *(Associativity) $A \cap (B \cap C) = (A \cap B) \cap C$.*

(iv) *(Idempotence) $A \cap A = A$.*

(v) *$A \subset B$ if and only if $A \cap B = A$.*

Proof. Proof omitted. □

Proposition 1.1.4.2 (Distributive laws)**.** *Let A, B and C be sets. Then*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Proof. We only prove the first one, the second one is similar. If $x \in A \cap (B \cup C)$, then $x \in A$ and $x \in (B \cup C)$. Since $x \in (B \cup C)$ implies $x \in B$ or $x \in C$. If $x \in B$, then we have $x \in A$ and $x \in B$, and conclude that $x \in A \cap B$. If $x \in C$, then we have $x \in A$ and $x \in C$, and conclude that $x \in A \cap C$. Since we either have $x \in B$ or $x \in C$, thus we either have $x \in A \cap B$ or $x \in A \cap C$. Together our conclusions, we have $x \in (A \cap B) \cup (A \cap C)$. To prove the reverse, we see that either $x \in A \cap B$ or $x \in A \cap C$, this implies we either have both $x \in A$ and $x \in B$ or both $x \in A$ and $x \in C$. Since $x \in A$ is always true, we either have $x \in B$ or $x \in C$. Thus, we conclude that $x \in A$, and $x \in B$ or $x \in C$, i.e., $x \in A \cap (B \cup C)$. □

Similarly, the difference sets can be defined with the axiom scheme of specification.

Definition 1.1.5 (Difference sets)**.** The *difference set* $A \setminus B$ (or $A - B$) is defined to be the set A with any elements of B removed:

$$A \setminus B := \{x \in A : x \notin B\}.$$

The basic facts about difference sets can be described as follows.

Proposition 1.1.5.1 (Basic properties of difference sets)**.** *Let A, B, X be sets, and let A, B be subsets of X.*

(i) *$X \setminus (X \setminus A) = A$.*

(ii) *$X \setminus \emptyset = X$ and $X \setminus X = \emptyset$.*

(iii) $A \cap (X \setminus A) = \emptyset$ *and* $A \cup (X \setminus A) = X$.

(iv) $A \subset B$ *if and only if* $X \setminus B \subset X \setminus A$.

Proof. Proof omitted. □

Proposition 1.1.5.2 (De Morgan laws)**.** *Let* A, B, X *be sets, and let* A, B *be subsets of* X*. Then*

$$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B),$$
$$X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B).$$

Proof. Proof omitted. □

There is a handy operation of sets which is usually used in mathematical proofs, namely symmetric difference.

Definition 1.1.6 (Symmetric difference)**.** The *symmetric difference* $A \triangle B$ of two sets is defined to be the set

$$A \triangle B := (A \setminus B) \cup (B \setminus A).$$

There is some properties of symmetric differences.

Proposition 1.1.6.1 (Basic properties of symmetric differences)**.** *Let* A, B, C *and* X *be sets, and let* $A \subset X$*.*

(i) *(Commutativity)* $A \triangle B = B \triangle A$.

(ii) *(Associativity)* $A \triangle (B \triangle C) = (A \triangle B) \triangle C$.

(iii) $A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$.

(iv) $A \triangle \emptyset = A$ *and* $A \triangle X = X \setminus A$.

(v) $A \triangle A = \emptyset$ *and* $A \triangle (X \setminus A) = X$.

(vi) $A \triangle B = (A \cup B) \setminus (A \cap B)$.

Proof. Proof omitted. □

Axiom 1.5 (Axiom of powers)**.** Let X be a set. Then the set

$$\{Y : Y \subset X\}$$

is a set, such a set is called the *power set* of X.

Axiom 1.6 (Axiom scheme of replacement)**.** Let A be a set. For any object $x \in A$, and any object, suppose we have a statement $P(x, y)$ pertaining to x and y, such that for each $x \in A$ there is at most one y for which $P(x, y)$ is true. Then there exists a set $\{y : P(x, y) \text{ is true for some } x \in A\}$, such that for any object z,

$$z \in \{y : P(x, y) \text{ is true for some } x \in A\} \iff P(x, z) \text{ is true for some } x \in A.$$

Axiom 1.7 (Axiom of infinity)**.** There exists a set containing 0 and containing the successor of each of its elements.

The axiom of infinity introduces the most basic example of an infinite set, namely the set of natural numbers. we shall return to it when we set about constructing the natural numbers.

§1.2 Russell's paradox

Pseudo-Axiom (Axiom of comprehension)**.** Suppose for every object x we have a property $P(x)$ pertaining to x. Then there exists a set $\{x : P(x) \text{ is true}\}$ such that for every object y,

$$y \in \{x : P(x) \text{ is true}\} \iff P(y) \text{ is true}.$$

This axiom is a pseudo-axiom because it creates a logical contradiction known as *Russell's paradox*, discovered by the philosopher and logician Bertrand Russell (1872 - 1970) in 1901. The paradox runs as follows.

Let $P(x)$ be the statement "x is a set, and $x \notin x$", i.e., $P(x)$ is true only when x is a set which does not contain itself. If we let S be the set of all sets (this is possible from the axiom of comprehension), then since S is itself a set, it is an element of S, and so $P(x)$ is false. Now use the axiom of comprehension to create the set $\Omega := \{x : P(x)\}$. Ω is the set of all sets which do not contain themselves. We wonder that is $\Omega \in \Omega$ true? If Ω did contain itself, then by definition this means that $P(\Omega)$ is true, i.e., Ω is a set and $\Omega \notin \Omega$. Otherwise, if Ω did not contain itself, then $P(\Omega)$ would be true, and hence $\Omega \in \Omega$. Thus in either case we have both $\Omega \in \Omega$ and $\Omega \notin \Omega$, which is absurd.

We shall simply postulate an axiom which ensures that absurdities such as Russell's paradox do not occur.

Axiom 1.8 (Axiom of regularity)**.** If A is a non-empty set, then there is at least one element x of A which is either not a set, or is disjoint from A, i.e., $x \cap A = \emptyset$.

One particular consequence of this axiom is that sets are no longer allowed to contain themselves.

Lemma 1.2.0.1. *If A is a set, then $A \notin A$. Furthermore, if A and B are two sets, then either $A \notin B$ or $B \notin A$.*

Proof. Suppose that $A \neq \emptyset$ and $A \in A$. Then we have $A \in \{A\}$. Since A is a set, by the axiom of regularity, we only have $A \cap \{A\} = \emptyset$. But since $A \in A$ and $A \in \{A\}$, this implies that $A \cap \{A\} = A$ which is not non-empty, a contradiction. While if $A = \emptyset$ which contains no elements, thus $A \notin A$.

Furthermore, suppose that A, B are two sets and we both have $A \in B$ and $B \in A$. Then $A, B \in \{A, B\}$. By the axiom of regularity, we either have $A \cap \{A, B\} = \emptyset$ or $B \cap \{A, B\} = \emptyset$. Since $A \in B$, we have $A \cap \{A, B\} = B$, but B is non-empty for that $A \in B$. With similar reasoning, $B \cap \{A, B\}$ is empty implies $B \cap \{A, B\}$ is non-empty, a contradiction. □

Remark. The axioms of set theory that we have introduced (Axioms 1.1-1.8) are known as the *Zermelo-Fraenkel axioms* (ZF axioms) of set theory, after Ernest Zermelo (1871-1953) and Abraham Fraenkel (1891-1965). There is one further axiom we will eventually need, the famous *axiom of choice* (see Section **??**), giving rise to the *Zermelo-Fraenkel-Choice axioms* (ZFC axioms) of set theory, bu we will not need this axiom for some time.

§1.3 Cartesian product

Another fundamental operation on sets, the *Cartesian product*. We first define the *ordered pair*.

Definition 1.3.1 (Ordered pair)**.** If a and b are any objects, we define the *ordered pair* (a, b) to be the set

$$(a, b) := \{\{a\}, \{a, b\}\}.$$

We prove that two pairs are equal if and only if both their components match.

Lemma 1.3.1.1. *If (a, b) and (x, y) are ordered pairs, then $(a, b) = (x, y)$ if and only if $a = x$ and $b = y$.*

Proof. We first show that (a, b) is singleton if and only if $a = b$. If (a, b) is singleton, then $\{a\} = \{a, b\}$, so that $b \in \{a\}$, and hence $b = a$. Conversely, if $a = b$, then $(a, b) = \{\{a\}\}$ which is singleton.

Now we return to the lemma. If $a = b$, then both (a, b) and (x, y) are singletons, so that $x = y$. Since $\{x\} \in (a, b)$ and $\{a\} \in (x, y)$, this implies that a, b, x and y are all equal.

If $a \neq b$, then (a, b) and (x, y) contain exactly one singleton, i.e., $\{a\}$ and $\{x\}$. Thus $a = x$. Furthermore, (a, b) and (x, y) contain exactly one pair, i.e., $\{a, b\}$ and $\{x, y\}$, and thus $\{a, b\} = \{x, y\}$. Since $a = x$ and $b \in \{x, y\}$, we only have $b = y$.

For the other direction, suppose that $a = x$ and $b = y$. We only need to show that case of $a \neq b$. By the hypothesis, we have $\{a\} = \{x\}$ and $\{a, b\} = \{x, y\}$. Thus $(a, b) = (x, y)$ is hold for that $\{\{a\}, \{a, b\}\} = \{\{x\}, \{x, y\}\}$ is hold. This complete the proof. □

Now we can introduce the *Cartesian product*.

Definition 1.3.2 (Cartesian product)**.** If X and Y are sets, then we define the *Cartesian product* $X \times Y$ to be the collection of ordered pairs, whose first component lies in X and second component lies in Y, i.e.,

$$X \times Y := \{(x, y) : (x, y) \text{ for some } x \in X \text{ and } y \in Y\}.$$

Here are some properties of Cartesian product.

Proposition 1.3.2.1. *Let A, B, X, Y be sets.*

(i) $(A \cup B) \times X = (A \times X) \cup (B \times X)$.

(ii) $(A \cap B) \times (X \cap Y) = (A \times X) \cap (B \times Y)$.

(iii) $(A \setminus B) \times X = (A \times X) \setminus (B \times X)$.

(iv) $(X \times Y) \setminus (A \times B) = ((X \setminus A) \times Y) \cup (X \times (Y \setminus B))$.

(v) $A \times B = \emptyset$ *if and only if* $A = \emptyset$ *or* $B = \emptyset$.

(vi) *If* $A \subset X$ *and* $B \subset Y$*, then* $A \times B \subset X \times Y$.

(vii) *Let* $A \times B \neq \emptyset$*. If* $A \times B \subset X \times Y$*, then* $A \subset X$ *and* $B \subset Y$.

(viii) *If* $A = X$ *and* $B = Y$*, then* $A \times B = X \times Y$*. Conversely, if* $A \times B \neq \emptyset$ *and* $A \times B = X \times Y$*, then* $A = Y$ *and* $B = Y$.

Proof. We only prove (v), (vi) and (vii).

(v) For the "if" part, assume that $A \times B \neq \emptyset$. Then there is $(x, y) \in A \times B$, which means $x \in A$ and $y \in B$. So that $A \neq \emptyset$ and $B \neq \emptyset$, a contradiction. Conversely, assume that neither A nor B is empty, then there is $x \in A$ and $y \in B$. This implies $(x, y) \in A \times B$, a contradiction.

(vi) Suppose $A \subset X$ and $B \subset Y$, then for every $x \in A$ and every $y \in B$, we have $(x, y) \in A \times B$. Since $x \in X$ and $y \in Y$, we implies that $(x, y) \in X \times Y$. Thus $A \times B \subset X \times Y$.

(vii) Suppose that $(x, y) \in A \times B$. Suppose for sake of contradiction that there exists $a \in A$ but $a \notin X$. Then $(a, y) \in A \times B$ and $(a, y) \notin X \times Y$, but it is impossible for that $A \times B \subset X \times Y$. Thus $A \subset X$. A same proof shows that $B \subset Y$.

Note that we requires $A \times B$ be non-empty, this equivalently requires that both A and B are non-empty from (v). If $A = \emptyset$, we must have $A \times B = \emptyset \subset X \times Y$ for arbitrary B. We can let $B \supset Y$ and $B \neq Y$, then the premise is true but the conclusion is false. □

Definition 1.3.3. Relations A *relation* R is a subset of a Cartesian product $X \times Y$. We usually write xRy instead of $(x, y) \in R$.

We define the *domain* $\operatorname{dom} R$ of R and the *range* $\operatorname{ran} R$ of R to be the sets

$$\operatorname{dom} R := \{x : xRy \text{ for some } y\}$$

and

$$\operatorname{ran} R := \{y : xRy \text{ for some } x\}.$$

§1.4 Functions

Definition 1.4.1 (Functions)**.** Let X, Y be sets. A *function from X to Y* is a relation f such that $\operatorname{dom} f = X$ and that for each $x \in X$ there is a unique $y \in Y$ such that $(x, y) \in f$. In other words, for each $x \in X$, if $(x, y) \in f$ and $(x, z) \in f$, then $y = z$. If f is a function from X to Y, we write $f : X \to Y$.

The domain of a function f from X to Y is equal to X, but its range need not be equal to Y. The range consists of those elements $y \in Y$ for which there exists an $x \in X$ such that $f(x) = y$.

There are certain special types of functions which are important to deserve a name.

Definition 1.4.2 (Injective, surjective, and bijective functions)**.** Let $f : X \to Y$ be a function.

- A function f is *surjective* or *onto* if $\operatorname{ran} f = Y$.
- A function f is *injective* or *one-to-one* if different elements map to different elements, i.e.,

$$f(x) = f(x') \implies x = x'.$$

- A function f is *bijective* or *invertible* if it is surjective and injective.

CHAPTER II

Differentiation in several variable calculus

§2.1 Derivatives

2.1.1 Directional derivative and differentiability

First, let us recall how the derivative in one-dimension is defined.

Definition 2.1.1. Let E be a subset of $\mathbf{R}$, $f : E \to \mathbf{R}$ be a function, $L \in \mathbf{R}$, and let a be an interior point of E. We say that f is *differentiable at a with derivative $f'(a) := L$* if we have

$$\lim_{h\to 0} \frac{f(a+t) - f(a)}{t} \tag{2.1}$$

converges to L. This is equivalent to

$$\lim_{h\to 0} \frac{f(a+t) - f(a) - L \cdot t}{t} = 0. \tag{2.2}$$

The following facts are an immediate consequence:

(i) Differentiable functions are continuous.

(ii) Composites of differentiable functions are differentiable.

We seek now to define the derivative of a function $f : E \to \mathbf{R}^n$ for $E \subset \mathbf{R}^n$. We cannot simply replace a and t in the definition just given by points of $\mathbf{R}^m$ in (2.1), for we cannot divide a point of $\mathbf{R}^m$ by a point of $\mathbf{R}^n$ when $m > 1$. Here is a first attempt at a definition.

Definition 2.1.2 (Directional derivative)**.** Let $E \subset \mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, let a be an interior point of E, and let $v \in \mathbf{R}^n$. If the limit

$$\lim_{t\to 0} \frac{f(a+tv) - f(a)}{t}$$

exists, we say that f is *differentiable in the direction v at a*, and denote the above limit by $D_v f(a)$.

The directional derivative is not a suitable generalization of the concept of derivative, since it would not follow that differentiability implies continuity or that composites of differentiable functions are differentiable. However, it is closely related to the partial derivative, which we will define later.

Observe the equation (2.2) that L is a linear map from $\mathbf{R}$ to $\mathbf{R}$. This leads to the following definition.

Definition 2.1.3 (Differentiability)**.** Let $E \subset \mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, let a be an interior point of E, and let $L : \mathbf{R}^n \to \mathbf{R}^m$ be a linear map. We say that f is *differentiable at a with derivative $Df(a) := L$* if we have

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - Lh\|}{\|h\|} = 0.$$

Before we proceed further, we have to check a basic fact, which is that a function can have *at most one* derivative at any point of its domain.

Lemma 2.1.3.1 (Uniqueness of derivatives)**.** *Let $E \subset \mathbf{R}^n$. If $f : E \to \mathbf{R}^m$ is differentiable at a, then there is a unique linear map $L : \mathbf{R}^n \to \mathbf{R}^m$ such that*

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - Lh\|}{\|h\|} = 0.$$

Proof. Suppose for sake of contradiction that there are two linear maps L_1, L_2 such that $L_1 \neq L_2$ and both are derivatives of f, then there is non-zero element $v \in \mathbf{R}^n$ such that $L_1 v \neq L_2 v$. Let t be a positive real number, we have $tv \to 0$ as $t \to 0$. Let $F(h) = f(a+h) - f(a)$, we have

$$\begin{aligned}
\lim_{t \to 0} \frac{\|L_1(tv) - L_2(tv)\|}{\|tv\|} &= \lim_{t \to 0} \frac{\|L_1(tv) - F(tv) + F(tv) - L_1(tv)\|}{\|tv\|} \\
&\leq \lim_{t \to 0} \frac{\|L_1(tv) - F(tv)\|}{\|tv\|} + \lim_{t \to 0} \frac{\|F(tv) - L_1(tv)\|}{\|tv\|} = 0.
\end{aligned}$$

Then

$$0 = \lim_{t \to 0} \frac{\|L_1(tv) - L_2(tv)\|}{\|tv\|} = \lim_{t \to 0} \frac{\|L_1 v - L_2 v\|}{\|v\|},$$

implies $L_1 v = L_2 v$, a contradiction. $\square$

We now show that this definition is stronger than the directional derivative, and that it is indeed a suitable definition of differentiability. Specifically, we verify the following facts:

(i) Differentiability of f at a implies the existence of all the directional derivatives of f at a.

(ii) Differentiable functions are continuous.

(iii) Composites of differentiable functions are differentiable.

For the first fact, we have following important lemma.

Lemma 2.1.3.2. *Let $E \subset \mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, let a be an interior point of E, and let $v \in \mathbf{R}^n$. If f is differentiable at a, then f is also differentiable in the direction v at a, and*

$$D_v f(a) = Df(a)v.$$

Proof. If $v = 0$, we have $D_v f(a) = Df(a)v = 0$. Let $v \neq 0$, we have $\|v\| > 0$. By the definition, we have

$$\lim_{t \to 0} \frac{\|f(a+tv) - f(a) - Df(a)(tv)\|}{\|tv\|} = 0.$$

If $t > 0$, above limit implies by multiply $\|v\|$

$$\lim_{t \to 0} \left\| \frac{f(a+tv) - f(a)}{t} - Df(a)v \right\| = 0.$$

If $t < 0$, we reach the same conclusion by multiply $-\|v\|$. Thus $D_v f(a) = Df(a)v$. □

For the second fact, we show that the differentiable functions are continuous.

Proposition 2.1.3.1. *Let $E \subset \mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, and let a be an interior point of E. If f is differentiable at a, then f is continuous at a.*

Proof. Use the definition and the triangular inequality, we have

$$\|f(a+h) - f(a)\| \leq \|h\| Df(a) + Lh.$$

If f is differentiable at a and $h \to 0$, we have

$$\lim_{h \to 0} \|f(a+h) - f(a)\| = 0.$$

Thus f is continuous at a. □

We shall deal with composites of differentiable functions in the next section.

2.1.2 Partial derivative and Jacobian matrix

Now we show how to calculate $Df(a)$, provided it exists. We first introduce the notion of the partial derivatives of a function. We define the partial derivative to be the directional derivative in the direction e_j, where e_j are the standard basis of $\mathbf{R}^n$.

Definition 2.1.4 (Partial derivative)**.** Let $E \subset \mathbf{R}^n$, let $f : E \to \mathbf{R}^m$ be a function, let a be an interior point of E, and let $e_1, \cdots, e_n$ be a standard basis of $\mathbf{R}^n$. Then the *partial derivative of f with respect to e_j at a*, denoted $D_j f(a)$, is defined by

$$D_j f(a) := \lim_{t \to 0} \frac{f(a + te_j) - f(a)}{t}$$

provided of course that the limit exists. Otherwise, we leave $D_j f(a)$ undefined. The function f

is called *partial differentiable at a* if every $D_1 f(a), \cdots, D_n f(a)$ exist.

Remark. Informally, given $f(x_1, \cdots, x_j, \cdots, x_n)$, the partial derivative $D_j f(a)$ can be obtained by holding all the variables other than x_j fixed, and then applying the single-variable calculus derivative in the x_j variable.

The Lemma 2.1.3.2 has an important corollary that ensures one can write directional derivatives in terms of partial derivatives.

Corollary 2.1.4.1. *Let $E \subset \mathbf{R}^n$, $f : \mathbf{R}^n \to \mathbf{R}^m$ be a function, let a be an interior point of E. If f is differentiable at a, then for $v = (v_1, \cdots, v_n) \in \mathbf{R}^n$ we have*

$$Df(a)v = \sum_{j=1}^{n} v_j D_j f(a).$$

Proof. By Lemma 2.1.3.2 and the definition of partial derivative, we have

$$Df(a)v = Df(a) \sum_{j=1}^{n} v_j e_j = \sum_{j=1}^{n} v_j Df(a) e_j = \sum_{j=1}^{n} v_j D_j f(a),$$

as desired. □

Lemma 2.1.4.1. *Let $E \subset \mathbf{R}^n$, let $f : E \to \mathbf{R}^m$ with $f = (f_1, \cdots, f_m)$, and let a be an interior point of E. The function f is differentiable at a if and only if each coordinate function f_j for $1 \leq j \leq m$ is differentiable at a.*

Proof. Since

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - Lh\|}{\|h\|} = 0$$

is equivalent to for every $1 \leq j \leq m$,

$$\lim_{h \to 0} \frac{\|f_j(a+h) - f_j(a) - Lh\|}{\|h\|} = 0,$$

as desired. □

Remark. This lemma tells us that

$$Df(a) = \begin{pmatrix} Df_1(a) \\ \vdots \\ Df_m(a) \end{pmatrix}$$

when f is differentiable at a with the form of column vector in $\mathbf{R}^m$.

It is often convenient to consider the matrix which is uniquely determined by the linear map $Df(a) : \mathbf{R}^n \to \mathbf{R}^m$ and the standard basis of $\mathbf{R}^n$ and $\mathbf{R}^m$. This $m \times n$ matrix is called the *Jacobian matrix of f at a*, following lemma gives the form of Jacobian matrix.

Lemma 2.1.4.2. *Let $E \subset \mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, and let a be an interior point of E. If f is differentiable at a, then*

$$Df(a) = \begin{pmatrix} D_1 f_1(a) & D_2 f_1(a) & \cdots & D_n f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) & \cdots & D_n f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f_m(a) & D_2 f_m(a) & \cdots & D_n f_m(a) \end{pmatrix}$$

This matrix is called the Jacobian matrix *of f at a.*

Proof. We already know that there is an $m \times n$ matrix determined by $Df(a)$ with respect to the standard basis of $\mathbf{R}^n$ and $\mathbf{R}^m$. Let $e_1, \cdots, e_n$ and $e_1, \cdots, e_m$ be standard bases of $\mathbf{R}^n$ and $\mathbf{R}^m$, respectively. Then we have $Df(a)e_j = \sum_{j=1}^{m} a_{ij} e_j$, we only need to verify that $a_{ij} = D_i f_j(a)$ for every $1 \leq i \leq m$, $1 \leq j \leq n$.

By Lemma 2.1.3.2, Corollary 2.1.4.1, Lemma 2.1.4.1, and the linearity of $Df(a)$, we have

$$\begin{aligned} Df(a)e_j &= (Df_1(a), \cdots, Df_m(a))e_j \\ &= (Df_1(a)e_j, \cdots, Df_m(a)e_j) \\ &= (D_j f_1(a), \cdots, D_j f_m(a)) \\ &= \sum_{i=1}^{m} D_j f_i(a) e_i. \end{aligned}$$

Therefore, $a_{ij} = D_j f_i(a)$, as desired. □

Lemma 2.1.3.2 shows that the differentiability of f at a point guarantees the existence of the directional derivatives, but the converse is false. However, if the partial derivatives are continuous, we can recover the differentiability of f. We given the following handy theorem:

Theorem 2.1.4.1. *Let $E \subset \mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, a be an interior point of $F \subset E$. f is differentiable at a if and only if all the partial derivatives $D_j f(a)$ exist on F and are continuous at a.*

Proof. The "only if" part follows easily from Lemma 2.1.3.2. We show the "if" part.

Let $L : \mathbf{R}^n \to \mathbf{R}^m$ be the linear map

$$Lh := \sum_{j=1}^{n} h_j D_j f(a)$$

where $h = (h_1, \cdots, h_n) \in \mathbf{R}^n$. We have to prove that

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - Lh\|}{\|h\|} = 0.$$

Let $\varepsilon > 0$. It will suffice to find a radius $\delta > 0$ such that

$$\|f(a+h) - f(a) - Lh\| \leq \varepsilon \|h\|$$

for every $a+h \in B(a,\delta)$.

Because a is an interior point of F, there exists a ball $B(a,r) \subset F$. Because each partial derivative $D_j f(a)$ is continuous on F, there thus exists an $0 < \delta_j < r$ such that $\|D_j f(a+h) - D_j f(a)\| \leq \varepsilon/nm$ for every $a+h \in B(a,\delta_j)$. If we take $\delta = \min(\delta_1, \cdots, \delta_n)$, then we thus have $\|D_j f(a+h) - D_j f(a)\| \leq \varepsilon/nm$ for every $a+h \in B(a,\delta_j)$ and every $1 \leq j \leq n$.

Let $a+h \in B(a,\delta)$, and write $h = h_1 e_1 + \cdots + h_n e_n$. Our task is to show that

$$\left\| f(a+h_1e_1+\cdots+h_ne_n) - f(a) - \sum_{j=1}^n h_j D_j f(a) \right\| \leq \varepsilon \|h\|.$$

Write f in components as $f = (f_1, \cdots, f_m)$. From the mean value theorem, we have

$$f_i(a+h_1e_1) - f_i(a) = D_1 f_i(a+t_ie_1)h_1$$

for some $t_i \in (0,h_1)$. Then by the triangular inequality,

$$\begin{aligned}
\|f(a+h_1e_1) - f(a) - D_1 f(a)h_1\| &\leq \sum_{i=1}^m |f_i(a+h_1e_1) - f_i(a) - D_1 f_i(a)h_1| \\
&= \sum_{i=1}^m |D_1 f_i(a+t_ie_1) - D_1 f_i(a)| \cdot |h_1| \\
&\leq \sum_{i=1}^m \|D_1 f(a+t_ie_1) - D_1 f(a)\| \cdot \|h\| \\
&\leq \|h\| \sum_{i=1}^m \varepsilon/nm \\
&= \varepsilon \|h\|/n.
\end{aligned}$$

A similar argument gives

$$\|f(a+h_1e_1+h_2e_2) - f(a+h_1e_1) - D_2 f(a)h_2\| \leq \varepsilon\|h\|/n,$$

and so forth up to

$$\|f(a+h_1e_1+\cdots+h_ne_n) - f(a+h_1e_1+\cdots+h_{n-1}e_{n-1}) - D_n f(a)h_2\| \leq \varepsilon\|h\|/n.$$

If we sum these n inequalities and use the triangle inequality $\|x+y\| \leq \|x\| + \|y\|$, we obtain

$$\left\| f(a+h_1e_1+\cdots+h_ne_n) - f(a) - \sum_{j=1}^n h_j D_j f(a) \right\| \leq \varepsilon \|h\|$$

as desired. □

§2.2 The chain rule

2.2.1 Several calculus chain rule

We are now ready to prove the fact that composites of differentiable functions are differentiable, which is called the several variable calculus chain rule, or the chain rule for short.

Theorem 2.2.0.1 (Several calculus chain rule)**.** *Let $E \subset \mathbf{R}^n$, and $F \subset \mathbf{R}^m$. Let $f : E \to F$ be a function, and let $g : F \to \mathbf{R}^p$ be another function. Let a be interior point of E. Suppose that f is differentiable at a, and that $f(a)$ is the interior point of F. Suppose also that g is differentiable at $f(a)$. Then $f \circ g : E \to \mathbf{R}^p$ is also differentiable at a, and we have the formula*

$$D(g \circ f)(a) = Dg(f(a))Df(a).$$

Proof. Let $b = f(a)$, let $S = Df(a)$, $T = Dg(f(a))$, and let

$$\begin{aligned} F(h) &:= f(a+h) - f(a) - Sh, \\ G(k) &:= g(b+k) - g(b) - Tk. \end{aligned}$$

From the definition of differentiability, we have

$$\lim_{h\to 0} \frac{\|F(h)\|}{\|h\|} = 0, \quad \text{and} \quad \lim_{k\to 0} \frac{\|G(k)\|}{\|k\|} = 0.$$

We want to show that

$$\lim_{h\to 0} \frac{\|(g\circ f)(a+h) - (g\circ f)(a) - (TS)h\|}{\|h\|} = 0.$$

Since

$$\begin{aligned} (g\circ f)(a+h) - (g\circ f)(a) - (TS)h &= g(f(a+h)) - g(b) - T(Sh) \\ &= g(f(a+h)) - g(b) - T(f(a+h) - f(a) - F(h)) \\ &= g(f(a+h)) - g(b) - T(f(a+h) - f(a)) + T(F(h)). \end{aligned}$$

Let $k = f(a+h) - f(a)$, we have $T(f(a+h) - f(a)) = g(f(a+h)) - g(b) - G(f(a+h) - f(a))$, then

$$(g\circ f)(a+h) - (g\circ f)(a) - (TS)h = G(f(a+h) - f(a)) + T(F(h)).$$

It suffices to show that

$$\lim_{h\to 0} \frac{\|G(f(a+h) - f(a))\|}{\|h\|} = 0 \quad \text{and} \quad \lim_{h\to 0} \frac{\|T(F(h))\|}{\|h\|} = 0.$$

We use the fact that linear map is bounded, i.e., for a linear map $T : \mathbf{R}^n \to \mathbf{R}^m$, there is a number $M > 0$ such that $\|Tx\| \le M\|x\|$ for every $x \in \mathbf{R}^n$. Then for the first part, we see that

$$\begin{aligned} \lim_{h\to 0} \frac{\|G(f(a+h) - f(a))\|}{\|h\|} &= \lim_{h\to 0} \frac{\|G(f(a+h) - f(a))\|}{\|f(a+h) - f(a)\|} \frac{\|f(a+h) - f(a)\|}{\|h\|} \\ &\le \lim_{h\to 0} \frac{\|G(f(a+h) - f(a))\|}{\|f(a+h) - f(a)\|} \Big(\frac{\|F(h)\|}{\|h\|} + M \Big) \\ &= 0. \end{aligned}$$

And for the second part,

$$\lim_{h\to 0} \frac{\|T(F(h))\|}{\|h\|} = \lim_{h\to 0} \frac{\|T(F(h))\|}{\|F(h)\|} \frac{\|F(h)\|}{\|h\|} \le \lim_{h\to 0} M \frac{\|F(h)\|}{\|h\|} = 0,$$

as desired. □

Remark. In the linear algebra, we usually write $Dg(f(a))Df(a)$ instead of $Dg(f(a)) \circ Df(a)$ for the composite of linear maps. Recall that the composite of linear maps is corresponding to the multiplication of matrices, thus we have $D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$ when the derivatives are given by the linear maps, and $D(g \circ f)(a) = Dg(f(a))Df(a)$ when the derivatives are given by the Jacobian matrices.

2.2.2 The mean-value theorem

As an application of the chain rule, we generalize the mean-value theorem of single-variable calculus.

Theorem 2.2.0.2 (Mean-value theorem)**.** *Let E be a convex open subset of $\mathbf{R}^n$, let $f : E \to \mathbf{R}^m$ be differentiable on E. Suppose that there is a real number M such that $\|Df(x)\| \leq M$ for every $x \in E$. Then*

$$\|f(x) - f(y)\| = M\|x - y\|$$

for every $x, y \in E$.

Proof. Let $F(t) := f(tx + (1-t)y)$ for $t \in [0,1]$. Since E is convex, we see that $tx + (1-t)y \in E$ for $t \in [0,1]$. By the chain rule, F is differentiable and $F'(t) = Df(tx + (1-t)y))(x - y)$. By the mean-value theorem of single-variable calculus, there is a $t_0 \in [0,1]$ such that

$$\|f(x) - f(y)\| = \|F(1) - F(0)\| = \|F'(t_0)\| \leq M\|x - y\|,$$

as desired. □

Proposition 2.2.0.1.

(i) *Let U be a connected, open subset of $\mathbf{R}^n$, $f : U \to \mathbf{R}^m$ be a function. If f is differentiable on U and $Df(x) = 0$ for every $x \in U$, then f is a constant function.*

(ii) *Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear map. Then T is continuously differentiable at every point of $\mathbf{R}^n$, and we have $DT(x) = T$ for every $x \in \mathbf{R}^n$.*

Proof. (i) Let $c \in \mathbf{R}^m$, and let

$$S := \{x \in U : f(x) = c\}.$$

We want to show that $S = U$. If we can prove that S is closed and open subset of U, then the claim follows since the subset of connected set is open and closed if and only if it equals to U or $\emptyset$. Since the differentiability implies the continuity, and $\{c\}$ is closed, then $S = f^{-1}(\{c\})$ is also closed. For the other hand, let $a \in S$, there is $\varepsilon > 0$ such that $B(a, \varepsilon) \subset U$. Because $Df(x) = 0$ for every $x \in U$, by Theorem 2.2.0.2, we have $\|f(x) - f(a)\| \leq 0$ for every $x \in B(a, \varepsilon)$, this implies that $f(x) = f(a) = c$. Thus $B(a, \varepsilon) \subset S$ and S is open. Since S is nonempty, we only have $S = U$, as desired.

(ii) For every $x \in \mathbf{R}^n$, we have

$$\lim_{h \to 0} \frac{\|T(x+h) - Tx - Th\|}{\|h\|} = \lim_{h \to 0} \frac{\|Tx + Th - Tx - Th\|}{\|h\|} = 0.$$

Thus by definition, T is differentiable at x with derivative $DT(x) = T$. □

§2.3 Continuous differentiability

We now investigate what happens if one differentiates a function twice.

Definition 2.3.1 (Continuous differentiability)**.** Let E be an open subset of $\mathbf{R}^n$, and let $f : E \to \mathbf{R}^m$ be a function. We say that f is *continuously differentiable* if the partial derivatives $D_1 f, \cdots, D_n f$ exists and are continuous on E. We say that f is C^k or *k-times continuously differentiable* if all the partial derivatives of f of order less than or equal to k exist and are continuous function on E. A function f is called C^∞ or *smooth* or *infinitely differentiable* if f is C^k for all $k \geq 0$.

There is an important theorem in calculus concerning continuous differentiability.

Theorem 2.3.1.1 (Clairaut-Schwartz's theorem)**.** *Let E be an open subset of $\mathbf{R}^n$, and let $f : E \to \mathbf{R}^m$ be C^2 on E. Then we have*

$$D_j D_i f(a) = D_i D_j f(a)$$

for all $1 \leq i, j \leq n$.

Proof. □

Theorem 2.3.1.2 (Taylor's theorem)**.**

§2.4 The inverse function theorem

2.4.1 The contraction mapping theorem

Before we turn to the next topic - namely, the inverse function theorem - we need to develop a useful fact from the theory of complete metric spaces, namely the contraction mapping theorem.

Definition 2.4.1 (Contraction)**.** Let (X, d) be a metric space, and let $f : X \to X$ be a map. We say that f is a *contraction* if there exists a constant $0 < c < 1$ such that

$$d(f(x), f(y)) \leq c d(x, y)$$

for every $x, y \in X$.

Definition 2.4.2 (Fixed points)**.** Let $f : X \to X$ be a map, and $x \in X$. We say that x is a *fixed point* of f if $f(x) = x$.

We now give the contraction mapping theorem, it is also called the *Banach's fixed point theorem.*

Theorem 2.4.2.1 (Contraction mapping theorem)**.** *Let (X, d) be a complete metric space, and $f : X \to X$ be a contraction. Then f has* exactly *one fixed point.*

Proof. The uniqueness is trivial, for if $f(x) = x$ and $f(y) = y$, then $d(x, y) \leq cd(x, y)$ if and only if $d(x, y) = 0$.

We prove the existence of a fixed point. Pick an arbitrary point $x_0 \in X$, and we define the sequence of x_n recursively that $x_n = f(x_{n-1})$ for $n \geq 1$. Choose $0 < c < 1$ such that $d(x_{n+1}, x_n) \leq d(x_n, x_{n-1})$ for $n \geq 1$. Use the induction we have

$$d(x_{n+1}, x_n) \leq c^n d(x_1, x_0)$$

for $n \geq 1$. For arbitrary $1 \leq n < m$, we have

$$d(x_n, x_m) \leq \sum_{k=n}^{m-1} d(x_{k+1}, x_k) \leq c^k d(x_1, x_0) \leq \frac{c^k}{1-c} d(x_1, x_0).$$

Thus $(x_n)_{n=0}^{\infty}$ is a Cauchy sequence. Since X is complete, $\lim_{n\to\infty} x_n = x$ for some $x \in X$. This shows that $f(x) = x$, as desired. □

Remark. The contraction mapping theorem is one example of a *fixed point theorem* - a theorem which guarantees, assuming certain conditions, that a map will have a fixed point.

We shall give one consequence of the contraction mapping theorem which is important for our application to the inverse function theorem.

Lemma 2.4.2.1. *Let $B(0, r) \subset \mathbf{R}^n$, and let $g : B(0, r) \to \mathbf{R}^n$ be a map such that $g(0) = 0$ and*

$$\|g(x) - g(y)\| \leq \frac{1}{2}\|x - y\|$$

for all $x, y \in B(0, r)$. Then the function $f : B(0, r) \to \mathbf{R}^n$ defined by $f(x) := x + g(x)$ is one-to-one, and furthermore the image $f(B(0, r))$ of this map contains the ball $B(0, r/2)$.

Proof. We first show that f is one-to-one. Suppose for sake of contradiction that we have two different point $x, y \in B(0, r)$ such that $f(x) = f(y)$. This means that $x + g(x) = y + g(y)$ and $\|g(x) - g(y)\| = \|x - y\|$. From the hypothesis, this requires that $\|x - y\| = 0$, i.e., $x = y$, a contradiction. Thus f is one-to-one.

We now show that $B(0, r/2) \subset f(B(0, r))$. Let $y \in B(0, r/2)$ be arbitrary. If we can find a point $x \in B(0, r)$ such that $f(x) = y$, i.e., $y - g(x) = x$, and if we can prove that x is a fixed point of the map $x \mapsto y - g(x)$, then the claim follows.

Let $F : B(0, r) \to B(0, r)$ to be the function $F(x) := y - g(x)$. If $x \in B(0, r)$, then

$$\|F(x)\| \leq \|y\| + \|g(x)\| < \frac{r}{2} + \frac{1}{2}\|x\| < \frac{r}{2} + \frac{r}{2} = r,$$

so F does indeed map $B(0, r)$ to itself. The same argument shows that for a sufficiently small $\varepsilon > 0$, F maps the closed ball $\overline{B(0, r - \varepsilon)}$ to itself. For any $x, x' \in B(0, r)$ we have

$$\|F(x) - F(x')\| = \|g(x') - g(x)\| \leq \frac{1}{2}\|x' - x\|.$$

Thus F is a contraction on $B(0, r)$, and hence on the complete space $\overline{B(0, r-\varepsilon)}$. By the contraction mapping theorem, F has a fixed point, as desired. □

2.4.2 The inverse function theorem

Let U be an open subset of $\mathbf{R}^n$, $a \in U$, and let $f : U \to \mathbf{R}^m$ be differentiable at a. If f is injective and $V = f(U)$ be open in $\mathbf{R}^m$, then f is bijective and its inverse function $f^{-1} : V \to U$ exists. But this would not necessarily imply that f^{-1} is differentiable at $f(a)$. In fact, if f has a differentiable inverse f^{-1}, by the chain rule and Proposition 2.2.0.1(ii), we have

$$Df^{-1}(f(a))Df(a) = D(f^{-1} \circ f)(a) = DI(a) = I(a),$$

where $I : \mathbf{R}^n \to \mathbf{R}^m$ is the identity and linear map. This means $Df^{-1}(f(a))$ is an inverse of $Df(a)$. Thus a necessary condition of f to have a differentiable inverse f^{-1} is that $Df(a)$ is invertible, in other words, $\mathbf{R}^n$ and $\mathbf{R}^m$ have same dimension, i.e., $n = m$. We now prove that the derivative $Df(a)$ is invertible is sufficient for f to have a differentiable inverse, at least locally. This result is called the *inverse function theorem*.

Theorem 2.4.2.2 (Inverse function theorem). *Let E be open subset of $\mathbf{R}^n$, let $f : E \to \mathbf{R}^n$ be a function which is continuously differentiable on E, let $a \in E$. Suppose $Df(a)$ is invertible. Then there exists an open set $U \subset E$ containing a, and an open set $V \subset \mathbf{R}^n$ containing $f(a)$, such that*

(i) *f is a bijection from U to V.*

(ii) *The inverse map $f^{-1} : V \to U$ is differentiable at $f(a)$ with derivative*

$$Df^{-1}(f(a)) = (Df(a))^{-1}.$$

Proof. Observe claim (ii) that if the inverse map f^{-1} is differentiable, the formula $Df^{-1}(f(a)) = (Df(a))^{-1}$ is automatic. To see this, notice that $I = f^{-1} \circ f$ on U, where $I : \mathbf{R}^n \to \mathbf{R}^n$ is the identity map. Then use the chain rule at a, we have

$$DI(a) = Df^{-1}(f(a))Df(a).$$

By Proposition 2.2.0.1(ii) that $DI = I$, we thus have $Df^{-1}(f(a)) = (Df(a))^{-1}$.

We are going to simplify the problem by making three additional assumptions and showing that it suffices to prove the theorem under these assumptions.

We first assume that $f(a) = 0$. To see that it suffices to prove the theorem under this assumption, let $F(x) := f(x) - f(a)$. If there is an open set $U \subset E$ containing a, and an open set $V' \subset \mathbf{R}^n$ containing $F(a) = 0$ such that $F : U \to V'$ is bijective. Let $V := V' + f(a)$, and let $\sigma(x) \mapsto x + f(a)$ be a function, which is obviously bijective, from V' to V. Then $\sigma \circ F : U \to V$ is also bijective and $\sigma \circ F = f$.

Next, one can assume that $a = 0$. To see that it suffices to prove the theorem under this assumption, let $F(x) := f(x + a)$. If there is an open set $U' \subset E$ containing 0, and an open set $V \subset \mathbf{R}^n$ containing $F(a)$ such that $F : U' \to V$ is bijective. Let $U := U' - a$, and let $\iota(x) \mapsto x - a$

be a bijective function from U' to U. Then $F \circ \iota : U \to V$ is also bijective and $F \circ \iota = f$. Thus we now can assume that $f(0) = 0$ and that $f'(0)$ is invertible.

Finally, one can assume that $Df(0) = I$, where $I : \mathbf{R}^n \to \mathbf{R}^n$ is the identity map. To see this assumption, let $F(x) := (Df(0))^{-1} f(x)$. Without loss of generality, we have $F(0) = 0$ and that

$$DF(0) = (Df(0))^{-1} Df(0) = I.$$

If there exists an open set U' containing 0, and an open set V' containing 0 such that F is a bijection from U' to V', and that $F^{-1} : V' \to U'$ is differentiable at 0 with derivative I. Then by $f(x) = Df(0)F(x)$, we have f is a bijection from U' to V, for that if we let $V = Df(0) \cdot V'$ and $\tau(x) := Df(0)x$ which is bijection, and hence $f = F \circ \tau$.

Together our assumptions, it suffices to consider the special case $a = 0$, $f(a) = f(0) = 0$, and $Df(0) = I$. All we have to do now is prove the inverse function theorem under these three assumptions.

We show that f is locally bijective around 0. We use Lemma 2.4.2.1 to show that f is one-to-one. Let $g : E \to \mathbf{R}^n$ be the function defined as $g(x) := f(x) - x$. Then $g(0) = 0$ and $Dg(0) = 0$ (this is comes from our simplification). In particular, $D_j g(0) = 0$ for $j = 1, \cdots, n$. Since g is continuously differentiable, there exists a ball $B(0, r) \subset E$ such that

$$\|D_j g(x)\| \leq \frac{1}{2n^2}$$

for every $x \in B(0, r)$. In particular, for any $x \in B(0, r)$ and $v = (v_1, \cdots, v_n) \in \mathbf{R}^n$ we have

$$\|D_v g(x)\| = \left\| \sum_{j=1}^{n} v_j D_j g(x) \right\| \leq \sum_{j=1}^{n} |v_j| \|D_j g(x)\| \leq \sum_{j=1}^{n} \|v\| \frac{1}{2n^2} \leq \frac{1}{2n} \|v\|.$$

By the fundamental theorem of calculus, for every $x, y \in B(0, r)$ we have

$$\begin{aligned} g(y) - g(x) &= \int_0^1 \frac{d}{dt} g(x + t(y - x)) dt \\ &= \int_0^1 D_{y-x} g(x + t(y - x)) dt \\ &\leq \int_0^1 \frac{1}{2n} \|y - x\| dt \\ &= \frac{1}{2n} \|y - x\|. \end{aligned}$$

Since every component of $g(y) - g(x)$ has magnitude at most $\frac{1}{2n}\|y - x\|$, and hence $g(y) - g(x)$ itself has magnitude at most $\frac{1}{2}\|y - x\|$. Then by Lemma 2.4.2.1, $f(x) = x + g(x)$ is one-to-one on $B(0, r)$, and $B(0, r/2) \subset f(B(0, r))$. In particular, we have an inverse map $f^{-1} : B(0, r/2) \to B(0, r)$.

Applying the contraction bound with $y = 0$ we obtain that $\|g(x)\| \leq \frac{1}{2}\|x\|$ for every $x \in B(0, r)$, and so by the triangle inequality

$$\frac{1}{2}\|x\| \leq \|f(x)\| \leq \frac{3}{2}\|x\|$$

for every $x \in B(0, r)$.

Set $V := B(0, r/2)$ and $U := f^{-1}(B(0, r/2))$. Then by construction f is a bijection from U to V. V is clearly open, and $U = f^{-1}(V)$ is also open since f is continuous.

Now we want to show that $f^{-1} : V \to U$ is differentiable at 0 with derivative $I^{-1} = I$. In other words, we wish to show that

$$\lim_{h\to 0} \frac{\|f^{-1}(0+h) - f^{-1}(0) - Ih\|}{\|h\|} = \lim_{h\to 0} \frac{\|f^{-1}(h) - h\|}{\|h\|} = 0,$$

since $f(0) = 0$ implies $f^{-1}(0) = 0$.

Above limit holds, if for any sequence $(x_n)_{n=1}^{\infty}$ in V converges to 0, we have

$$\lim_{n\to\infty} \frac{\|f^{-1}(x_n) - x_n\|}{\|x_n\|} = 0.$$

Write $y_n := f^{-1}(x_n)$. Then $y_n \in B(0, r)$ and $x_n = f(y_n)$. In particular, we have

$$\frac{1}{2}\|y_n\| \leq \|x_n\| \leq \frac{3}{2}\|y_n\|.$$

Since $\|x_n\|$ goes to 0, $\|y_n\|$ also goes to 0, and their ratio remains bounded. It will thus suffice to show that

$$\lim_{n\to\infty} \frac{\|y_n - f(y_n)\|}{\|y_n\|} = 0.$$

But since $y_n \to 0$, and f is differentiable at 0, we have

$$\lim_{n\to\infty} \frac{\|y_n - f(y_n)\|}{\|y_n\|} = \lim_{n\to\infty} \frac{\|f(y_n) - f(0) - Df(0)y_n\|}{\|y_n\|} = 0$$

as desired. □

CHAPTER III

Exterior algebra

§3.1 Tensor product

The motivation for our next topic comes from wanting to form the product of two vectors. This product is called the tensor product, and will be an element of some new vector space.

Definition 3.1.1 (Multilinear map)**.** Let $V_1, \cdots, V_k$ be vector spaces over $\mathbf{F}$. A function $f : V_1 \times \cdots \times V_m \to \mathbf{F}$ is called a *multilinear map* (in specific, *k-linear map*) if for every $1 \leq j \leq k$ we have

$$\begin{aligned} f(v_1, \cdots, v_j + w_j, \cdots, v_k) &= f(v_1, \cdots, v_j, \cdots, v_k) + f(v_1, \cdots, w_j, \cdots, v_k), \\ f(v_1, \cdots, \lambda v_j, \cdots, v_k) &= \lambda f(v_1, \cdots, v_j, \cdots, v_k), \end{aligned}$$

for every $v_j, w_j \in V_j$ and every $\lambda \in \mathbf{F}$. We denote the set of all k-linear maps from $V_1 \times \cdots \times V_k$ to $\mathbf{F}$ by $\mathcal{L}(V_1 \times \cdots \times V_k, \mathbf{F})$.

Remark. 1-linear map is called the linear map, and 2-linear map is called the bilinear map. This is easy to see that $\mathcal{L}(V_1 \times \cdots \times V_k, \mathbf{F})$ is a vector space over $\mathbf{F}$ with two operations, addition $+$ and multiplication $\cdot$, which defined by

$$\begin{aligned} (f + g)(v_1, \cdots, v_k) &= f(v_1, \cdots, v_k) + g(v_1, \cdots, v_k), \\ (\lambda f)(v_1, \cdots, v_k) &= \lambda f(v_1, \cdots, v_k) \end{aligned}$$

for every $f, g \in \mathcal{L}(V_1 \times \cdots \times V_k, \mathbf{F})$ and every $\lambda \in \mathbf{F}$.

From the dual space, we have a commutative diagram, which describes the relationship between linear map $T \in \mathcal{L}(V, W)$, linear functional $f \in W^*$, and dual map $T^* f \in V^*$:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[dr, "T^* f"'] & W \arrow[d, "f"] \\ & \mathbf{F} \end{tikzcd}$$

You may probably noticed that the space of multilinear maps $\mathcal{L}(V_1 \times \cdots \times V_k, \mathbf{F})$ is just the dual space $(V_1 \times \cdots \times V_k)^*$, this inspire us to generalize above commutative diagram to the multilinear map. Appendix 3.A gives a quick review of the dual space.

3.1.1 Tensor product of two vector spaces

We begin by the special case of the tensor product of two vector spaces.

Definition 3.1.2 (Tensor product of two vector spaces)**.** Let V, W be vector spaces over $\mathbf{F}$, let $f \in V^* = \mathcal{L}(V, \mathbf{F})$ and $g \in W^* = \mathcal{L}(W, \mathbf{F})$. The *tensor product* $f \otimes g$, of 1-linear maps, is defined by $f \otimes g(v, w) = f(v) \cdot g(w)$ for every $v \in V$ and $w \in W$. We denote $V^* \otimes W^*$ to be the vector space with elements of the form $f \otimes g$ for $f \in V^*$ and $g \in W^*$.

This is easy to see that $f \otimes g$ is a bilinear map on $V \times W$, i.e., $f \otimes g \in \mathcal{L}(V \times W, \mathbf{F})$.

Lemma 3.1.2.1. *Let V, W be vector spaces over $\mathbf{F}$, and let $(v_i)_{1\le i\le n}$ and $(w_j)_{1\le j\le m}$ be bases of V and W, respectively. Suppose that $(v_i^*)_{1\le i\le n}$ and $(w_j^*)_{1\le j\le m}$ are bases of V^* and W^*, respectively. Then $(v_i^* \otimes v_j^*)_{1\le i\le n, 1\le j\le m}$ is a basis of $V^* \otimes W^*$. Furthermore, we have $\dim(V^* \otimes W^*) = (\dim V^*)(\dim W^*)$ and $V^* \otimes W^* = \mathcal{L}(V \times W, \mathbf{F})$.*

Proof. For arbitrary $f \in V^*$ and $g \in W^*$ such that $f \otimes g \in V^* \otimes W^*$, we have

$$\begin{aligned} f \otimes g(v, w) &= \left(\sum_{i=1}^{n} f(v_i) v_i^*(v)\right)\left(\sum_{j=1}^{m} g(w_j) w_j^*(w)\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} f(v_i) g(w_j) v_i^*(v) w_j^*(w) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} f(v_i) g(w_j) v_i^* \otimes w_j^*(v, w) \end{aligned}$$

for every $v \in V$, $w \in W$. Thus every element of $V^* \otimes W^*$ can be represented as a linear combination of $(v_i^* \otimes w_j^*)_{1\le i\le n, 1\le j\le m}$.

We now show that $(v_i^* \otimes w_j^*)_{1\le i\le n, 1\le j\le m}$ is linearly independent. Let $\lambda_{ij} \in \mathbf{F}$ for $1 \le i \le n, 1 \le j \le m$ such that

$$\sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_{ij} v_i^* \otimes w_j^* = 0.$$

Since $(\sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_{ij} v_i^* \otimes w_j^*)(v_i, w_j) = \lambda_{ij}$ for every $1 \le i \le n, 1 \le j \le m$, this requires that λ_{ij} all equals to zero. Thus $(v_i^* \otimes w_j^*)_{1\le i\le n, 1\le j\le m}$ is a basis of $V^* \otimes W^*$, and we have $\dim(V^* \otimes W^*) = nm = (\dim V^*)(\dim W^*)$.

Since $V^* \otimes W^*$ is a subspace of $\mathcal{L}(V \times W, \mathbf{F})$ and $\dim \mathcal{L}(V \times W, \mathbf{F}) = nm$, this implies that $V^* \otimes W^* = \mathcal{L}(V \times W, \mathbf{F})$. This complete the proof. □

Remark. Since V and W are the dual spaces of V^* and W^*, respectively, we can analogously define the tensor product $V \otimes W$. Then we have $V \otimes W = \mathcal{L}(V^* \times W^*, \mathbf{F})$.

A typical element of $V \otimes W$ is a linear combination of elements of the form $v \otimes w$. In other words, given any bilinear map f from $V \times W$ to a space Y, we can find a linear map T from $V \otimes W$ to Y such that $f(v, w) = T(v \otimes w)$ for every v, w.

Theorem 3.1.2.1. *Suppose $T : V \times W \to V \otimes W$ is the bilinear map obtain from the tensor product $\otimes$, i.e., for $v \in V$, $w \in W$,*

$$T(v, w) = v \otimes w.$$

Then for any bilinear map $g : V \times W \to \mathbf{F}$, there exists a unique linear map $f : V \otimes W \to \mathbf{F}$ such that $g = f \circ T : V \times W \to \mathbf{F}$.

Proof. Let $(v_i)_{1 \le i \le m}$ and $(w_j)_{1 \le j \le n}$ be the bases of V and W, respectively. Define a linear map $f : V \otimes W \to \mathbf{F}$ as following

$$f(v_i \otimes w_i) = T(v_i, w_i)$$

for every $1 \le i \le m, 1 \le j \le n$. Let $v = \sum_{i=1}^m a_i v_i \in V$ and $w = \sum_{j=1}^n b_j w_j \in W$. Then

$$f(v \otimes w) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j f(v_i \otimes w_j) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j T(v_i, w_j) = T(v, w).$$

Therefore $g = f \circ T : V \times W \to \mathbf{F}$, and it is clear that f is unique. □

3.1.2 Tensor product

Definition 3.1.3 (Tensor product)**.** Let $f \in \mathcal{L}(V_1 \times \cdots \times V_r, \mathbf{F})$ and $g \in \mathcal{L}(V_1 \times \cdots \times V_s, \mathbf{F})$. The *tensor product* $f \otimes g$ is defined by

$$f \otimes g(v_1, \cdots, v_r, w_1, \cdots, w_s) = f(v_1, \cdots, v_r) \cdot g(w_1, \cdots, w_s)$$

for every $v_i \in V_i$, $1 \le i \le r$, and every $w_j \in W_j$, $1 \le j \le s$. Obviously, $f \otimes g$ is a $(r+s)$-linear map on $V_1 \times \cdots \times V_r \times W_1 \times \cdots \times W_s$.

Theorem 3.1.3.1 (Associativity of tensor product)**.** *Let $f \in \mathcal{L}(U_1 \times \cdots \times U_r, \mathbf{F})$, $g \in \mathcal{L}(V_1 \times \cdots \times V_s, \mathbf{F})$, and $h \in \mathcal{L}(W_1 \times \cdots \times W_t, \mathbf{F})$. Then $(f \otimes g) \otimes h = f \otimes (g \otimes h)$.*

3.1.3 Tensors

Definition 3.1.4 (Tensor)**.** Let V be an n-dimensional vector space over $\mathbf{F}$ with dual space V^*. The elements in the tensor product

$$V_s^r := \underbrace{V \otimes \cdots \otimes V}_{r \text{ terms}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{s \text{ terms}}$$

are called (r,s)-*type tensors*, where r is the *contravariant order* and s is the *covariant order*. We denote $\mathcal{T}^r(V) := V_0^r$.

Remark. In particular, the elements in V_0^r are called r-*contravariant tensors*, and those in V_s^0 are called s-*covariant tensors*. We also have $V_0^0 = \mathbf{F}$, $V_0^1 = V$, and $V_1^0 = V^*$. The elements of V are called *contravectors*, and the elements of V^* are called *covectors*.

By Lemma 3.1.2.1, we have $\dim V_s^r = n^{r+s}$ if $\dim V = n$, and

$$V_s^r = \mathcal{L}(\underbrace{V^* \times \cdots \times V^*}_{r \text{ terms}} \times \underbrace{V \times \cdots \times V}_{s \text{ terms}}, \mathbf{F}).$$

Thus a $(r+s)$-linear map on V_s^r is also a $(r+s)$-type tensor. In particular, a r-linear map $f \in \mathcal{L}(V_0^r, \mathbf{F})$ is a r-contravariant tensor.

§3.2 Alternating tensors

Denote the permutation group of the set $\{1, \cdots, r\}$ by $\mathcal{S}_r$.

Definition 3.2.1 (Permutation action)**.** Let $f \in \mathcal{T}^r(V)$. We define the *permutation action* of f by

$$\sigma f(v_1^*, \cdots, v_r^*) = f(v_{\sigma(1)}^*, \cdots, v_{\sigma(r)}^*),$$

where $v_i^* \in V^*$.

- We say that f is *alternating*, if for every $\sigma \in \mathcal{S}_r$ we have $\sigma f = \operatorname{sgn} \sigma \cdot f$.
- We say that f is *symmetric*, if for every $\sigma \in \mathcal{S}_r$ we have $\sigma f = f$.

We see that f is a r-contravariant tensor, and we denote the set of all alternating r-contravariant tensors by $\Lambda^r(V)$, and the set of all symmetric r-contravariant tensors by $P^r(V)$.

Lemma 3.2.1.1. *Let $\sigma, \tau \in \mathcal{S}_r$, and $f, g \in \mathcal{T}^r(V)$. Then*

(i) *(Linearity) $\sigma(af + bg) = a\sigma f + b\sigma g$ for $a, b \in \mathbf{F}$.*

(ii) *(Composite) $\tau(\sigma f) = (\tau \circ \sigma) f$.*

Proof. (i) Since $af + bg \in \mathcal{T}^r(V)$, we have

$$\begin{aligned} a\sigma f + b\sigma g &= a\sigma f(v_1^*, \cdots, v_r^*) + b\sigma g(v_1^*, \cdots, v_r^*) \\ &= af(v_{\sigma(1)}^*, \cdots, v_{\sigma(r)}^*) + bg(v_{\sigma(1)}^*, \cdots, v_{\sigma(r)}^*) \\ &= (af + bg)(v_{\sigma(1)}^*, \cdots, v_{\sigma(r)}^*) \\ &= \sigma(af + bg)(v_1^*, \cdots, v_r^*) \\ &= \sigma(af + bg). \end{aligned}$$

(ii) For every $v_1^*, \cdots, v_r^* \in V^*$, we have

$$\tau(\sigma f) = \tau f(v_{\sigma(1)}^*, \cdots, v_{\sigma(r)}^*)$$

$$\begin{aligned} &= f(w_1^*, \cdots, w_r^*) \\ &= f(v_{\tau(\sigma(1))}^*, \cdots, v_{\tau(\sigma(r))}^*) \\ &= f(v_{(\tau\circ\sigma)(1)}^*, \cdots, v_{(\tau\circ\sigma)(r)}^*) \\ &= (\tau\circ\sigma)f, \end{aligned}$$

where $w_i^* = v_{\sigma(i)}^*$ for $1 \le i \le r$. □

Definition 3.2.2. Let $f \in \mathcal{T}^r(V)$. We define the *alternating map* $A_r : \mathcal{T}^r(V) \to \mathcal{T}^r(V)$ as

$$A_r(f) := \frac{1}{r!} \sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}\sigma \cdot \sigma f.$$

We define the *symmetric map* $S_r : \mathcal{T}^r(V) \to \mathcal{T}^r(V)$ as

$$S_r(f) := \frac{1}{r!} \sum_{\sigma\in\mathcal{S}_r} \sigma f.$$

Following lemma shows that $A_r(f)$ and $S_r(f)$ is indeed alternating and symmetric, respectively.

Lemma 3.2.2.1. *Suppose $f \in \mathcal{T}^r(V)$. Then $A_r(f)$ is alternating and $S_r(f)$ is symmetric.*

Proof. Let $\tau \in \mathcal{S}_r$. By Lemma 3.2.1.1, Lemma 3.B.3.1, and $\operatorname{sgn}\tau = (\operatorname{sgn}\tau)^{-1}$ we have

$$\begin{aligned} \tau(A_r(f)) &= \tau\Big(\frac{1}{r!}\sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}\sigma\cdot\sigma f\Big) \\ &= \frac{1}{r!}\sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}\sigma\cdot(\tau\circ\sigma) f \\ &= (\operatorname{sgn}\tau)^{-1}\frac{1}{r!}\sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}(\tau\circ\sigma)\cdot(\tau\circ\sigma) f \\ &= (\operatorname{sgn}\tau)\frac{1}{r!}\sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}(\tau\circ\sigma)\cdot(\tau\circ\sigma) f \\ &= (\operatorname{sgn}\tau)A_r(f), \end{aligned}$$

the last equation holds for that $\tau\circ\sigma$ runs through $\mathcal{S}_r$ as σ does. Similarly, we have

$$\begin{aligned} \tau(S_r(f)) &= \tau\Big(\frac{1}{r!}\sum_{\sigma\in\mathcal{S}_r} \sigma f\Big) \\ &= \frac{1}{r!}\sum_{\sigma\in\mathcal{S}_r} (\tau\circ\sigma) f \\ &= S_r(f). \end{aligned}$$

This complete the proof. □

§3.3 Exterior algebra

Due to E. Cartan's systematic development of the method of exterior differentiation, the alternating tensors have played an important role in the study of manifolds. An alternating r-contravariant tensor is also called an *exterior vector of degree* r or an *exterior* r*-vector*. The space $\Lambda^r(V)$ is called the *exterior space* of V of degree r.

More important, there exists an operation, the exterior (wedge) product, for exterior vectors such that the product of two exterior vectors is another exterior vector.

Definition 3.3.1 (Exterior product)**.** Let $\xi \in \Lambda^k(V)$ and $\zeta \in \Lambda^\ell(V)$. We define the *exterior product* (also is called the *wedge product*) of ξ and ζ to be a $(k+\ell)$-vector by

$$\xi \wedge \zeta := A_{k+\ell}(\xi \otimes \zeta),$$

where $A_{k+\ell}$ is the alternating map.

Following lemma gives the properties of exterior product.

Lemma 3.3.1.1. *Let $\xi, \xi_1, \xi_2 \in \Lambda^k(V), \eta, \eta_1, \eta_2 \in \Lambda^l(V), \zeta \in \Lambda^h(V)$. Then*

(i) *(Distributive law)* $(\xi_1 + \xi_2) \wedge \eta = \xi_1 \wedge \zeta + \xi_2 \wedge \zeta$, *and* $\xi \wedge (\eta_1 + \eta_2) = \xi \wedge \eta_1 + \xi \wedge \eta_2$.

(ii) *(Anticommutativity)* $\xi \wedge \eta = (-1)^{kl} \eta \wedge \xi$.

(iii) *(Associativity)* $(\xi \wedge \eta) \wedge \zeta = \xi \wedge (\eta \wedge \zeta)$.

§3.A Dual space

This section gives a quick review of dual space. We begin by the definition of dual space, which is a special vector space.

Definition 3.A.1 (Dual space)**.** Let V be a vector space over $\mathbf{F}$, let $f : V \to \mathbf{F}$ be a linear map (it has a special name, linear functional). The set of all linear functionals on V, denote $V^* := \mathcal{L}(V, \mathbf{F})$, is called the *dual space* of V.

Lemma 3.A.1.1. *If V is an n-dimensional vector space over $\mathbf{F}$, then V^* is also an n-dimensional vector space over $\mathbf{F}$.*

Proof. Let $v_1, \cdots, v_n$ be a basis of V. For every $v \in V$ there are $a_1, \cdots, a_n \in \mathbf{F}$ such that

$$v = \sum_{i=1}^n a_i v_i.$$

Then for linear functional $f \in \mathcal{L}(V, \mathbf{F})$, which is an element of V^*, we have

$$f(v) = \sum_{i=1}^n a_i f(v_i).$$

This means that the linear functional f is determined by its values $f(v_i)$ on the basis. We may define linear functionals $v_i^* \in V^*$ such that

$$v_i^*(v_j) = \delta_j^i$$

for every $1 \leq j \leq n$, where δ_j^i is the Kronecker δ-symbol. Then $v_i^*(v) = a_i$, we thus write $f(v)$ in the form

$$f(v) = \sum_{i=1}^{n} f(v_i) v_i^*(v)$$

for $f(v_i) \in \mathbf{F}$, and thus

$$f = \sum_{i=1}^{n} \lambda_i v_i^*$$

where $\lambda_i = f(a_i)$ for some $a_i \in \mathbf{F}$. This says that every element of V^* can be expressed as a linear combination of $v_1^*, \cdots, v_n^*$. Thus $\operatorname{span}(v_1^*, \cdots, v_n^*) = V^*$.

Now, the only task is to show that $v_1^*, \cdots, v_n^*$ is linearly independent. Let $a_1, \cdots, a_n \in \mathbf{F}$ such that

$$a_1 v_1^* + \cdots + a_n v_n^* = 0.$$

Since $(a_1 v_1^* + \cdots + a_n v_n^*)(v_i) = a_i$ for every $1 \leq i \leq n$, this implies $a_i = 0$ for every $1 \leq i \leq n$, as desired. □

From the proof of Lemma 3.A.1.1, the dual basis is well-defined.

Definition 3.A.2 (Dual bais)**.** Let V be a vector space over $\mathbf{F}$, and let V^* be a dual space of V. If $v_1, \cdots, v_n$ is a basis of V, then the *dual basis* $v_1^*, \cdots, v_n^*$ of $v_1, \cdots, v_n$ is defined by

$$v_i^*(v_j) := \delta_j^i$$

for every $1 \leq j \leq n$, where δ_j^i is the Kronecker δ-symbol. Furthermore, the dual basis of a basis of V is a basis of V^*.

In the definition below, note that if T is a linear map from V to W then T^* is a linear map from W^* to V^*.

Definition 3.A.3 (Dual map)**.** Let V, W be vector spaces over $\mathbf{F}$, and let V^* and W^* be dual spaces of V and W, respectively. If $T \in \mathcal{L}(V, W)$, then the *dual map* of T is the linear map $T^* \in \mathcal{L}(W^*, V^*)$ defined by $T^* f := f \circ T$ for $f \in W^*$.

This is easy to show that T^* is indeed a linear map from V to $\mathbf{F}$, i.e., $T^* f \in V^*$, and the proof is omitted here.

Following lemma gives the properties of dual map.

Lemma 3.A.3.1. *Let U, V and W be vector spaces over $\mathbf{F}$. Then*

(i) $(S+T)^* = S^* + T^*$ *for every* $S, T \in \mathcal{L}(V, W)$.

(ii) $(\lambda T)^* = \lambda T^*$ *for every* $\lambda \in \mathbf{F}$ *and every* $T \in \mathcal{L}(V, W)$.

(iii) $(ST)^* = T^* S^*$ *for every* $T \in \mathcal{L}(U, V)$ *and every* $S \in \mathcal{L}(V, W)$.

Proof. The proof omitted □

§3.B Permutations

In this section we introduce the particular kind of tensors with which we shall be concerned - the alternating tensors - and derive some of their properties. In order to do this, we need some basic facts about permutations.

Definition 3.B.1 (Permutation)**.** A *permutation* of a finite set X is a bijection $\sigma : X \to X$. We denote the set of all permutations on X by S_X. We abbreviate S_X as S_n for $X = \{1, \cdots, n\}$.

Definition 3.B.2 (Elementary permutation)**.** Let $1 \leq i \leq n$, let $\alpha_i \in S_n$ is defined by $\alpha_i(i) = i+1$ and $\alpha_i(i+1) = i$, and $\alpha_i(j) = j$ for $j \neq i$ and $j \neq i+1$. We call α_i an *elementary permutation*.

Following lemma shows that every permutation can be described by the elementary permutations, in other words, we can obtain arbitrary permutation with elementary permutations step by step.

Lemma 3.B.2.1. *If $\sigma \in S_n$, then σ equals a composite of elementary permutations.*

Proof. Given $0 \leq i \leq n$. We say that σ fixes the first i integers if $\sigma(j) = j$ for every $0 \leq j \leq i$. We see that if σ is an identity, then $\sigma = \alpha_j \circ \alpha$ for every $0 \leq j < n$.

We first show that if σ fixes the first $i-1$ integers, then σ can be written as the composite $\sigma = \pi \circ \sigma'$, where π is a composite of elementary permutations and σ' fixes the first i integers.

Since σ fixes the first $i-1$ integers, $\sigma(i)$ taking the value $\sigma(i) = i$ or $\sigma(i) = j > i$. If $\sigma(i) = i$, let $\sigma' = \sigma$ and let π be identity, we obtain the claim. While if $\sigma(i) = j$ for $j > i$, we set $\sigma' = \alpha_i \circ \cdots \circ \alpha_{j-1} \circ \sigma$. Obviously, σ' fixes the first $i-1$ integers, and also fixes i for that $\sigma(i) = j$ and $(\alpha_i \circ \cdots \circ \alpha_{j-1})(j) = i$ by definition of elementary permutation. Since $\alpha_j = \alpha_j^{-1}$ for every $1 \leq j \leq n$. We have $\sigma = \alpha_{j-1} \circ \cdots \circ \alpha_i \circ \sigma$.

Now we use the induction. When σ fixes the first n integers, the claim is trivial. Inductively suppose that σ fixes the first 0 integers, σ will taking the form $\sigma = \pi_0 \circ \cdots \circ \pi_n \circ \sigma'$, where σ' fixes the first n integers and $\pi_1, \cdots, \pi_n$ are composites of elementary permutations. This shows that every permutation equals a composite of elementary permutations. □

Definition 3.B.3 (Sign of permutation)**.** Let $\sigma \in S_n$. A pair (i, j) is called an *inversion* in σ if we have $\sigma(i) > \sigma(j)$ and $i < j$. We define the sign of permutation σ by

$$\operatorname{sgn} \sigma := \begin{cases} -1 & \text{if the number of inversions in } \sigma \text{ is odd,} \\ 1 & \text{if the number of inversions in } \sigma \text{ is even.} \end{cases}$$

We call σ an *odd permutation* if $\operatorname{sgn} \sigma = -1$, and *even permutation* if $\operatorname{sgn} \sigma = 1$.

Lemma 3.B.3.1. *Let $\sigma, \tau \in S_n$.*

(i) *If σ equals a composite of m elementary permutations, then* $\operatorname{sgn} \sigma = (-1)^m$.

(ii) $\operatorname{sgn}(\sigma \circ \tau) = (\operatorname{sgn} \sigma) \cdot (\operatorname{sgn} \tau)$.

(iii) $\operatorname{sgn} \sigma^{-1} = \operatorname{sgn} \sigma$.

(iv) *If $p \neq q$, and if τ is the permutation that exchanges p and q and leaves all other integers fixed, then* $\operatorname{sgn} \tau = -1$.

CHAPTER IV

Manifold

What is manifold? A good example to keep in mind is the surface of a smooth ball. If you look at a small portion of it from very close up, then it looks like a portion of a flat plane. More generally, an n-dimensional manifold, is a geometrical object that looks "locally" like n-dimensional *Euclidean space.*

§4.1 Topological manifolds

We first introduce topological manifolds, the most basic type of manifolds. A topological manifold is a space whose each point is surrounded by a region that can be identified with parts of Euclidean space. We may hope that there is a "nice" one-to-one map φ from the neighborhood into $\mathbf{R}^n$. Then φ allows us to use the coordinates in $\mathbf{R}^n$ to label points in the neighborhood: if p belongs to the neighborhood, then one can label it with the coordinates of $\varphi(p)$. This condition is given by the locally Euclidean. The function φ is called a coordinate chart of the neighborhood.

Definition 4.1.1 (Topological manifold)**.** Suppose M is a topological space. We say that M is an *n-dimensional topological manifold* (or a *topological n-manifold*) if it has the following properties:

(i) M is *Hausdorff*: for every distinct points $x, y \in M$, there exists disjoint open sets U, V such that $x \in U$ and $y \in V$.

(ii) M is *second-countable*: M has a countable basis.

(iii) M is *locally Euclidean of dimension n*: for every $x \in M$, there exists a neighborhood U of x such that U is homeomorphic to an open set in $\mathbf{R}^n$.

Remark. For the dimension of a topological manifold to be well-defined, we need to know that for $n \neq m$ an open subset of $\mathbf{R}^n$ is not homeomorphic to an open subset of $\mathbf{R}^m$. This fact, called *invariance of dimension*, is indeed true but is not easy to prove directly. Thus, we will not prove the theorem and just leave it here.

Theorem 4.1.1.1 (Topological invariance of dimension)**.** *A nonempty n-dimensional topological manifold cannot be homeomorphic to an m-dimensional manifold unless $m = n$.*

Now, here is a straightforward but central observation. Suppose that M is a topological manifold of dimension n, U and V are two neighborhoods in M that intersect, and suppose that functions $\varphi : U \to \mathbf{R}^n$ and $\psi : V \to \mathbf{R}^n$ are used to give them each a coordinate chart. Then the intersection $U \cap V$ is given two coordinate charts, and this gives us an identification between the open regions $\varphi(U \cap V)$ and $\psi(U \cap V)$ of $\mathbf{R}^n$: given a point p in the first region, the corresponding point in the second is $\psi(\varphi^{-1}(p))$. This composition of maps tells us how the coordinates from one of the charts on the intersecting region relate to those of the other.

Definition 4.1.2 (C^∞-compatible chart)**.** Let M be a topological n-manifold.

- A *coordinate chart* (or *chart*) on M is a pair (U, φ), where U is an open subset of M and $\varphi : U \to \varphi(U)$ is a homeomorphism from U to an open subset $\varphi(U) \subset \mathbf{R}^n$. We say that (U, φ) is *centered* at $p \in U$ if $\varphi(p) = 0$.
- Two charts (U, φ) and (V, ψ) of a topological manifold are C^∞*-compatible* if the two maps

$$\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V) \quad \text{and} \quad \psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$$

  are C^∞. This two maps are called the *transition maps* between the charts.

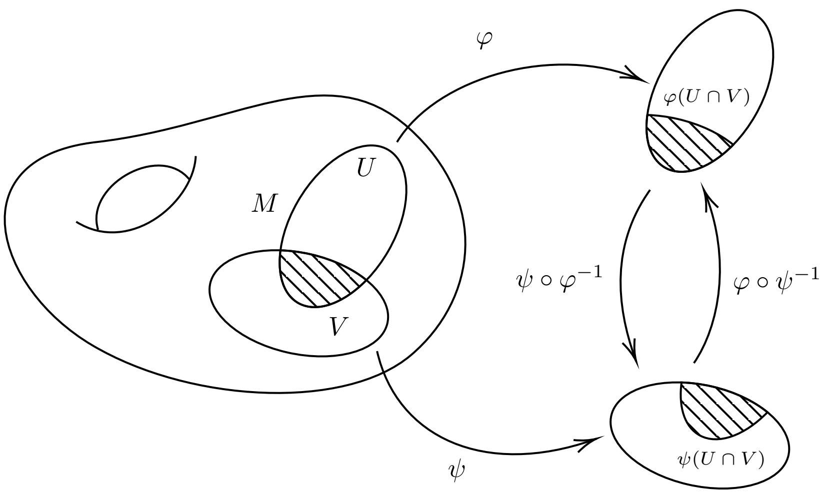

Figure 4.1: Transition maps (see Definition 4.1.2)

Given a topological n-manifold M, and a chart (U, φ) containing $p \in M$. Since φ is a homeomorphism, for any $p \in U$, we can define the coordinate of p to be the coordinate of $\varphi(p) \in \mathbf{R}^n$. Specifically, let $y_1, \cdots, y_n$ be the coordinates on $\mathbf{R}^n$, we let $x_i = y_i \circ \varphi$ be the ith component of φ and write $\varphi(p) = (x_1(p), \cdots, x_n(p))$. Thus, for $p \in U$, $(x_1(p), \cdots, x_n(p))$ is a point in $\mathbf{R}^n$. The functions $x_1, \cdots, x_n$ are called *local coordinates* on U.

Here is some simple examples of topological manifolds.

Example 4.1.2.1 (Euclidean sapce)**.** The Euclidean space $\mathbf{R}^n$ is covered by a single chart $(\mathbf{R}^n, \mathbf{1}_{\mathbf{R}^n})$, where $\mathbf{1}_{\mathbf{R}^n}$ is the identity map of $\mathbf{R}^n$. Every open subset U of $\mathbf{R}^n$ is also a topological manifold with chart $(U, \mathbf{1}_U)$.

Example 4.1.2.2 (Graphs of continuous functions)**.** Let $U \subset \mathbf{R}^n$ be an open subset, and let $f : U \to \mathbf{R}^k$ be a continuous function. The graph of f is the subset of $\mathbf{R}^n \times \mathbf{R}^k$ defined by

$$\Gamma(f) := \{(x, y) \in \mathbf{R}^n \times \mathbf{R}^k : x \in U, y = f(x)\},$$

with the subspace topology. Let $\pi_1 : \mathbf{R}^n \times \mathbf{R}^k \to \mathbf{R}^n$ denote the projection onto the first component.

Example 4.1.2.3 (Spheres)**.** For each $n \geq 0$, the unit n-sphere $\mathbf{S}^n$ is Hausdorff and second-countable because it is a topological subspace of $\mathbf{R}^{n+1}$. To show that it is locally Euclidean, for each $1 \leq i \leq n+1$, let U_i^+ denote the subset of $\mathbf{R}^{n+1}$ where the ith coordinate is positive:

$$U_i^+ := \{(x_1, \cdots, x_{n+1}) \in \mathbf{R}^{n+1} : x_i > 0\}.$$

Similarly, U_i^- is the set where $x_i < 0$.

Let $\mathbf{B}^n$ be the set of open unit balls in $\mathbf{R}^n$, and let $f : \mathbf{B}^n \to \mathbf{R}$ be the continuous function

$$f(u) = \sqrt{1 - \|u\|^2}.$$

Example 4.1.2.4 (Projective spaces)**.**

§4.2 Smooth structures

There are no constraints on the coordinate chart maps on a topological manifold except that they should be continuous. However, some continuous functions are quite unpleasant, so one typically introduces extra constraints that the coordinate chart maps are infinitely differentiable. If a manifold has a collection of charts for which all the coordinate chart maps are C^∞, then it is said to have a *smooth differentiable structure*, and it is called a *smooth manifold*. Smooth manifolds are especially interesting because they are the natural arena for calculus. Roughly speaking, they are the most general context in which the notion of differentiation to any order makes intrinsic sense.

Definition 4.2.1 (C^∞-differentiable structure)**.** Let M be a topological n-manifold. If given a collection of charts $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha)\}$ on M satisfies the following conditions, then we call $\mathfrak{A}$ a C^∞*-differentiable structure* on M:

(i) $\mathfrak{A}$ covers M, i.e., $M = \bigcup_\alpha U_\alpha$.

(ii) Any two charts in $\mathfrak{A}$ are C^∞-compatible.

(iii) $\mathfrak{A}$ is *maximal*, i.e., if a chart (V, ψ) is C^∞-compatible with all charts in $\mathfrak{A}$, then $(V, \psi) \in \mathfrak{A}$.

The collection of charts $\mathfrak{A}$ is called a C^∞-*atlas* for M if it satisfying condition (i) and (ii). Therefore, A C^∞-*differentiable structure* on M is a maximal atlas on M.

Definition 4.2.2 (Smooth manifold)**.** Let M be a topological n-manifold. If a C^∞-differentiable structure is given on M, then M is called a C^∞-*manifold of dimension* n (or a *smooth* n-*manifold*).

In practice, to check that a topological manifold M is a smooth manifold, it is not necessary to exhibit a maximal atlas. The existence of any atlas on M will do, because of the following proposition.

Proposition 4.2.2.1. *Let M be a topological manifold. Every atlas $\mathfrak{A} = \{(U_\alpha, \varphi_\alpha)\}$ for M is contained in a unique maximal atlas.*

Proof. Let $\mathfrak{A}$ be a C^∞-atlas for M, and let $\mathfrak{M}$ be the set of all charts that are C^∞-compatible with every chart in $\mathfrak{A}$. We first show that $\mathfrak{M}$ is a C^∞-atlas, we need to show that any two charts in $\mathfrak{M}$ are C^∞-compatible with each other. Let $(U, \varphi), (V, \psi) \in \mathfrak{M}$, we need to show that $\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ is C^∞.

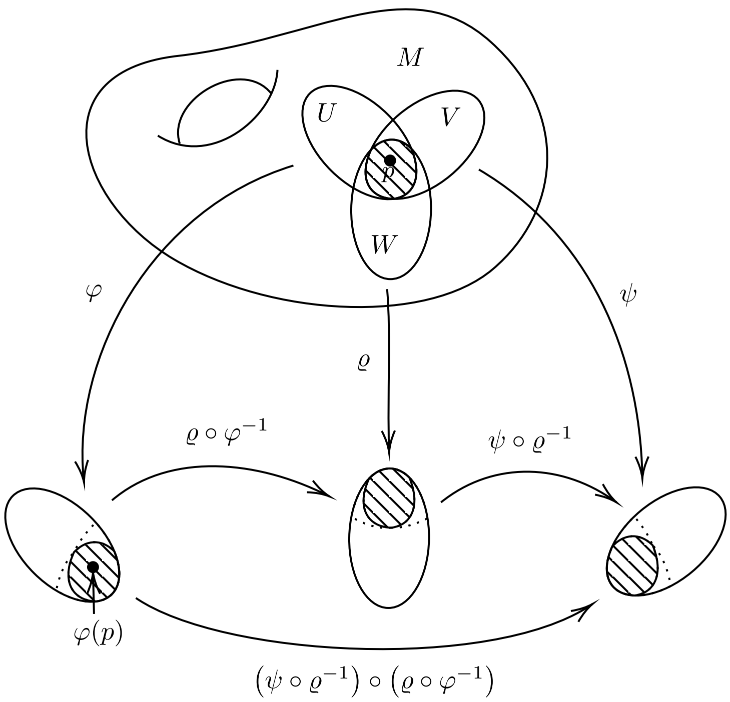

Figure 4.2: Proof of Proposition 4.2.2.1

Let $\varphi(p) \in \varphi(U \cap V)$ be arbitrary. Since $\bigcup \mathfrak{A} = M$, there is a chart $(W, \varrho) \in \mathfrak{A}$ such that $p \in W$. Since every chart in $\mathfrak{M}$ is C^∞-compatible with (W, ϱ), both of the maps $\varrho \circ \varphi^{-1}$ and $\psi \circ \varrho^{-1}$ are C^∞.

Since $p \in U \cap V \cap W$, we have $\psi \circ \varphi^{-1} = (\psi \circ \varrho^{-1}) \circ (\varrho \circ \varphi^{-1})$ is C^∞ on a neighborhood of $\varphi(p)$. Since $\varphi(p)$ is arbitrary, $\psi \circ \varphi^{-1}$ is C^∞ in a neighborhood of each point of $\varphi(U \cap V)$. This is similar to show that $\varphi \circ \psi^{-1}$ is C^∞ on $\psi(U \cap V)$. Thus $\mathfrak{M}$ is a C^∞-atlas.

The existence of the maximal atlas $\mathfrak{M}$ is given the definition of $\mathfrak{M}$ and $\mathfrak{A} \subset \mathfrak{M}$. To check that $\mathfrak{M}$ is unique, suppose that $\mathfrak{N}$ be any other maximal atlas that containing $\mathfrak{A}$. Then each charts in $\mathfrak{N}$ is C^∞-compatible with each charts in $\mathfrak{A}$, hence $\mathfrak{N} \subset \mathfrak{M}$. Other hand, since $\mathfrak{N}$ is maximal, we have $\mathfrak{N} \supset \mathfrak{M}$. Thus, $\mathfrak{N} = \mathfrak{M}$. $\square$

Since a C^∞-atlas for M covers M, Proposition 4.2.2.1 shows that it suffices to verify M to be a smooth manifold if it satisfying

(i) M is Hausdorff and second countable.

(ii) M has a C^∞-atlas (not necessarily to be maximal).

Notation 4.2.2.1. When we say a chart on a manifold, we mean a chart in the differentiable structure of manifold.

4.2.1 Examples of smooth manifolds

Example 4.2.2.1 (Euclidean spaces)**.** The Euclidean space $\mathbf{R}^n$ is a smooth n-manifold with the smooth structure determined by the atlas consisting of the single chart $(\mathbf{R}^n, \mathbf{1}_{\mathbf{R}^n})$. We call this the *standard smooth structure on* $\mathbf{R}^n$ and the resulting coordinate map *standard coordinates*.

Example 4.2.2.2 (Open submanifolds)**.** Let U be any open subset of $\mathbf{R}^n$. Then U is a topological n-manifold, and the single chart $(U, \mathbf{1}_U)$ defines a smooth structure on U.

More generally, let M be a smooth n-manifold, and let $U \subset M$ be any open subset. Define an atlas on U by

$$\mathfrak{U} := \{(V, \varphi) : V \subset U \text{ and } (V, \varphi) \text{ is a chart on } M\}.$$

We verify this definition makes sense. Every point $p \in U$ is contained in the domain of some chart (W, φ) on M. If we set $V = W \cap U$, then $(V, \varphi|_U)$ is a chart in $\mathfrak{U}$ who containing p. Therefore, U is covered by $\mathfrak{U}$. Thus any open subset of M is itself a smooth n-manifold. We call any open subset an *open submanifold of* M.

§4.3 Smooth maps

4.3.1 Smooth functions on manifold

On a smooth manifold, the concept of a smooth function is well-defined.

Definition 4.3.1 (Smooth function)**.** Let M be a smooth manifold of dimension n, and let $f : M \to \mathbf{R}^d$ be a function. If $p \in M$, We say that f is C^∞ or *smooth at* p if there exists a chart

(U, φ) containing p such that $f \circ \varphi^{-1}$ is C^∞ at $\varphi(p)$ on the open set $\varphi(U) \subset \mathbf{R}^n$. We say that f is *smooth* if it is smooth at every point of M. We denote the set of all smooth functions on M by $C^\infty(M)$.

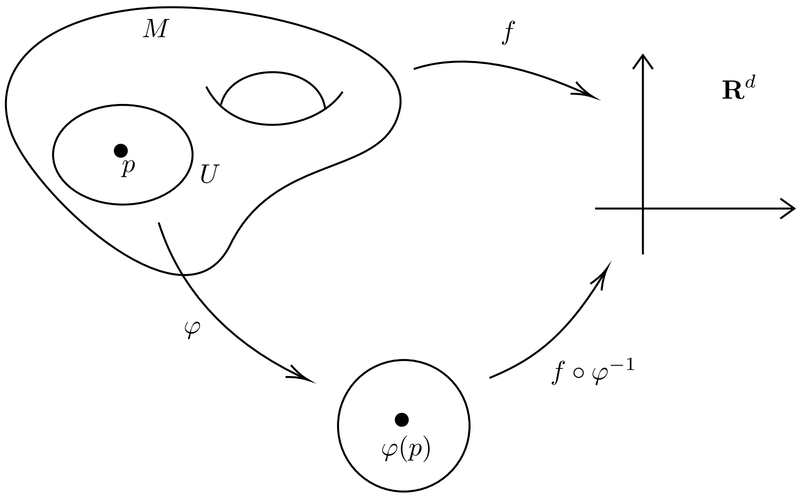

Figure 4.3: Smooth functions

Remark. The smoothness of a function is independent of the choice of charts and is therefore well-defined (see Proposition 4.3.1.1).

Smooth real-valued functions are important special case of smooth functions, we give the equivalent characterisation of smooth real-valued functions.

Proposition 4.3.1.1 (Characterisation of smooth real-valued functions)**.** *Let M be a smooth manifold of dimension n, and let $f : M \to \mathbf{R}$ be a real-valued function. Then following are equivalent:*

(i) *f is C^∞.*

(ii) *The manifold M has an atlas such that for every chart (U, φ) in the atlas, one has $f \circ \varphi^{-1} : \varphi(U) \to \mathbf{R}$ is C^∞.*

(iii) *For every chart (V, ψ) on M, $f \circ \psi^{-1} : \psi(V) \to \mathbf{R}$ is C^∞.*

Proof. (i) $\Rightarrow$ (iii). Let (V, ψ) be arbitrary chart on M and $p \in V$. By definition, there exists a chart (U, φ) containing $p \in U$ such that $f \circ \varphi^{-1}$ is C^∞ at $\varphi(p) \in \varphi(U)$. We see that $f \circ \psi^{-1} = (f \circ \varphi^{-1}) \circ (\varphi \circ \psi^{-1})$ is C^∞ at $\psi(p) \in \psi(U \cap V)$. Since $p \in V$ is arbitrary, $f \circ \psi^{-1}$ is C^∞ on $\psi(V)$.

(iii) $\Rightarrow$ (ii). This is trivial.

(ii) $\Rightarrow$ (i). This is immediately comes from the definition. $\square$

4.3.2 Smooth maps between manifolds

The definition of smooth functions generalizes easily to maps between manifolds.

Definition 4.3.2 (Smooth map)**.** Let M and N be smooth manifolds of dimension n and m, respectively. Let $F : M \to N$ be a map. If $p \in M$, we say that F is C^∞ or *smooth at p* if there exist charts (U, φ) containing $p \in M$ and (V, ψ) containing $F(p) \in N$ such that $F(U) \subset V$ and the composite $\psi \circ F \circ \varphi^{-1}$ is C^∞ at $\varphi(p)$ from $\varphi(U)$ to $\psi(V)$. We say that F is *smooth* if it is smooth at every point of M.

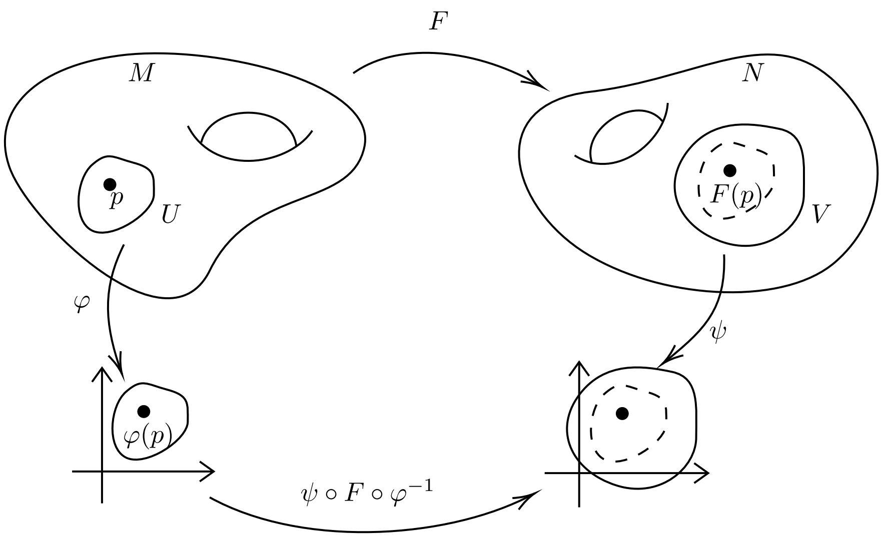

Figure 4.4: Smooth maps

There are some tractable characterisations of smoothness.

Proposition 4.3.2.1 (Characterisation of smooth maps)**.** *Let M and N be smooth manifolds, and let $F : M \to N$ be a map. Then following are equivalent:*

(i) *F is C^∞.*

(ii) *For every $p \in M$, there exist charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $U \cap F^{-1}(V)$ is open in M and the composite map $\psi \circ F \circ \varphi^{-1}$ is C^∞ from $\varphi(U \cap F^{-1}(V))$ to $\psi(V)$.*

(iii) *F is continuous and there exist atlases $\mathfrak{A}$ for M and $\mathfrak{B}$ for N such that for every chart $(U_\alpha, \varphi_\alpha) \in \mathfrak{A}$ and $(V_\beta, \psi_\beta) \in \mathfrak{B}$, $\psi_\beta \circ F \circ \varphi_\alpha^{-1}$ is C^∞ from $\varphi_\alpha(U_\alpha \cap F^{-1}(V_\beta))$ to $\psi_\beta(V_\beta)$.*

Proof. (i) $\Rightarrow$ (ii). Suppose F is C^∞. Let $p \in M$ be arbitrary, then there exist charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $F(U) \subset V$ and the composite $\psi \circ F \circ \varphi^{-1}$ is C^∞ at $\varphi(p)$ from $\varphi(U)$ to $\psi(V)$. Since for every $x \in M$ and every chart (V, ψ) such that $F(p) \in V$, we have $x \in U$ and $F(U) \subset V$, F is continuous. This implies that $F^{-1}(V)$ is open in M for V is open in N, hence $U \cap F^{-1}(V)$ is open in M. Obviously, $p \in U \cap F^{-1}(V)$ for that $F(p) \in V$. Since $\varphi(U \cap F^{-1}(V)) \subset \varphi(U)$, $\psi \circ F \circ \varphi^{-1}$ is also C^∞ at $\varphi(p)$ from $\varphi(U \cap F^{-1}(V))$ to $\psi(V)$. Because $p \in M$ is arbitrary, the claim follows.

(ii) $\Rightarrow$ (iii). We only need to show that F is continuous, the remaining part is trivial. Since for every $p \in M$ there is a chart (V, ψ) such that $F(p) \in V$. Then for $p \in U \cap F^{-1}(V)$, we have $F(U \cap F^{-1}(V)) \subset F(U) \cap F(F^{-1}(V)) \subset F(U) \cap V \subset V$. Thus F is continuous.

(iii) $\Rightarrow$ (i). Suppose (iii) and let $U = U_\alpha \cap F^{-1}(V_\beta) \subset U_\alpha$. Then we obtain $(U, \varphi_\alpha|_U)$ such that (iii) holds which is easy to prove. From the continuity of F, we have $F(U) \subset V_\beta$ is open. This complete the proof. $\square$

Previous proposition shows that smoothness automatically implies continuity. Following corollary shows that the smoothness of maps is independent of the choice of charts.

Corollary 4.3.2.1. *Let M, N be smooth manifolds, and let $F : M \to N$ be a smooth map. If (U, φ) is arbitrary chart containing $p \in M$ and (V, ψ) is arbitrary chart containing $F(p) \in N$, then $\psi \circ F \circ \varphi^{-1}$ is C^∞ on M.*

Remark. The map $\psi \circ F \circ \varphi^{-1}$ is usually called the *coordinate representation of F* with respect to the given charts. It maps the set $\varphi(U \cap F^{-1}(V))$ to $\psi(V)$.

Proof. Let $(\widetilde{U}, \widetilde{\varphi})$ and $(\widetilde{V}, \widetilde{\psi})$ be arbitrary charts for M. Suppose that $\widetilde{U} \cap F^{-1}(\widetilde{V}) \neq \emptyset$. Since F is smooth, for $p \in \widetilde{U} \cap F^{-1}(\widetilde{V}) \subset M$, there exist charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $\psi \circ F \circ \varphi^{-1}$ is C^∞ at $\varphi(p)$ on $\varphi(U \cap F^{-1}(V))$. By the C^∞-compatibility of charts in a differentiable structure, we have $\widetilde{\psi} \circ F \circ \widetilde{\varphi}^{-1} = (\widetilde{\psi} \circ \psi^{-1}) \circ (\psi \circ F \circ \varphi^{-1}) \circ (\varphi \circ \widetilde{\varphi}^{-1})$ is C^∞ at $\widetilde{\varphi}(p)$ on $\widetilde{\varphi}(\widetilde{U} \cap F^{-1}(\widetilde{V}))$. Since $(\widetilde{U}, \widetilde{\varphi})$ and $(\widetilde{V}, \widetilde{\psi})$ are arbitrary, the claim follows. $\square$

4.3.3 Constructing smooth maps

How does one go about constructing smooth maps from one smooth manifold to another? Following theorem gives a number of rules for constructing smooth maps.

Theorem 4.3.2.1 (Construction of smooth maps)**.** *Let M, N and P be smooth manifolds.*

(i) *(Constant map) Every constant map $f : M \to N$ is smooth.*

(ii) *(Identity) The identity map $\mathbf{1}_M$ is smooth.*

(iii) *(Inclusion) If $U \subset M$ is an open submanifold, then the inclusion map $\iota : U \hookrightarrow M$ is smooth.*

(iv) *(Composites) If $F : M \to N$ and $G : N \to P$ are smooth, then $G \circ F : M \to P$ is smooth.*

(v) *(Restriction) $F : M \to N$ is smooth if and only if every point $p \in M$ has a neighborhood U such that $F|_U : U \to N$ is smooth.*

Proof. (i) Define $f : M \to N$ as $f(p) = c$. Let $p \in M$ be arbitrary, there exist charts (U, φ) containing $p \in M$ and (V, ψ) containing c. We see that $U \cap f^{-1}(V) = U \cap M = U$ is open in M and $\psi \circ f \circ \varphi^{-1}$ is C^∞ from $\varphi(U)$ to $\psi(V)$. Thus, f is C^∞.

(ii) Let $p \in M$ be arbitrary, there exist (U, φ) containing p. Then $U \cap \mathbf{1}_M^{-1}(U) = U$ is open in M and $\varphi \circ \mathbf{1}_M \circ \varphi^{-1}$ is C^∞ from $\varphi(U)$ to $\varphi(U)$. Thus, $\mathbf{1}_M$ is C^∞.

«««< HEAD (iii) Let $p \in U$ be arbitrary. Let (W, φ) be a chart on U containing p. Then (W, φ) is also a chart on M containing $\iota(p) = p$. Then, $\varphi \circ \iota \circ \varphi^{-1} : \varphi(U) \to \varphi(U)$ is the identity map, which is smooth. Hence ι is smooth.

(iv) By definition, there exist charts (V, ϱ) containing $F(p)$ and (W, ψ) containing $G(F(p))$ such that $G(V) \subset W$ and $\psi \circ G \circ \varrho^{-1}$ is C^∞ from $\varrho(V)$ to $\psi(W)$. Since F is continuous and $F^{-1}(V)$ is a neighborhood of $p \in M$, so there is a chart (U, φ) for M such that $p \in U \subset F^{-1}(V)$. By Corollary 4.3.2.1, $\psi \circ F \circ \varphi^{-1}$ is C^∞ from $\varphi(U)$ to $\psi(V)$. Then we have $(G \circ F)(U) \subset G(U) \subset W$, and $\varrho \circ (G \circ F) \circ \varphi^{-1} = (\varrho \circ G \circ \psi^{-1}) \circ (\psi \circ F \circ \varphi^{-1})$ is C^∞ from $\varphi(U)$ to $\varrho(W)$.

(v) Suppose that F is smooth. Let $p \in M$ be arbitrary. Since F is C^∞, there are charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $\psi \circ F \circ \varphi^{-1}$ is C^∞. By Example 4.2.2.2, U is an open submanifold. Thus for every $p \in U$, there exist (U, φ) containing p and (V, ψ) containing $F(p)$ such that $\psi \circ F|_U \circ \varphi^{-1}$ is C^∞. Thus $F|_U$ is C^∞. ======= (iii) By definition, there exist charts (V, ϱ) containing $F(p)$ and (W, ψ) containing $G(F(p))$ such that $G(V) \subset W$ and $\psi \circ G \circ \varrho^{-1}$ is C^∞ from $\varrho(V)$ to $\psi(W)$. Since F is continuous and $F^{-1}(V)$ is a neighborhood of $p \in M$, so there is a chart (U, φ) for M such that $p \in U \subset F^{-1}(V)$. By Corollary 4.3.2.1, $\psi \circ F \circ \varphi^{-1}$ is C^∞ from $\varphi(U)$ to $\psi(V)$. Then we have $(G \circ F)(U) \subset G(U) \subset W$, and $\varrho \circ (G \circ F) \circ \varphi^{-1} = (\varrho \circ G \circ \psi^{-1}) \circ (\psi \circ F \circ \varphi^{-1})$ is C^∞ from $\varphi(U)$ to $\varrho(W)$.

(iv) Suppose that F is smooth. Let $p \in M$ be arbitrary. Since F is C^∞, there are charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $\psi \circ F \circ \varphi^{-1}$ is C^∞. By Example 4.2.2.2, U is an open submanifold. Thus for every $p \in U$, there exist (U, φ) containing p and (V, ψ) containing $F(p)$ such that $\psi \circ F|_U \circ \varphi^{-1}$ is C^∞. Thus $F|_U$ is C^∞. »»»> 53c256507d2a398e47311a8dc90d4ac4dc69f63c

Conversely, suppose that for every $p \in M$ there exists a neighborhood U such that $F|_U$ is C^∞. Then by definition, there exist charts (V, φ) containing p and (W, ψ) containing $F|_U(p)$ such that $V \cap (F|_U)^{-1}(W)$ is open in U and $\psi \circ F|_U \circ \varphi^{-1}$ is C^∞. Since V is an open subset of U, and U is open in M. Thus V is open subset of M. Similarly, $V \cap (F|_U)^{-1}(W)$ is open in M. Thus F is C^∞. □

§4.4 Diffeomorphisms

Definition 4.4.1 (Diffeomorphism)**.** Let M and N be smooth manifolds. A *diffeomorphism from M to N* is a bijective smooth map $F : M \to N$ whose inverse F^{-1} is also smooth. We say that M and N are *diffeomorphic* if there exists a diffeomorphism between them.

Following proposition gives the relation between diffeomorphism and chart.

Proposition 4.4.1.1. *Let M be a smooth manifold. If (U, φ) is a chart on M, then $\varphi : U \to \varphi(U)$ is a diffeomorphism. Conversely, if $\varphi : U \to \varphi(U)$ is a diffeomorphism on the open subset $U \subset M$, then (U, φ) is a chart in the differentiable structure of M.*

Proof. Suppose that (U, φ) is a chart on M, then by definition, $\varphi : U \to \varphi(U)$ is a homeomorphism. We need to show that φ and φ^{-1} is C^∞. Consider the atlases $\{(U, \varphi)\}$ and $\{(\varphi(U), \mathbf{1}_{\varphi(U)})\}$ which contains single chart. Since φ and φ^{-1} are continuous, and $\mathbf{1}_{\varphi(U)} \circ \varphi \circ \varphi^{-1}$ and $\varphi \circ \varphi^{-1} \circ \mathbf{1}_{\varphi(U)}$ are identity maps on $\varphi(U)$ and U, respectively. Thus, by Theorem 4.3.2.1(ii) and Proposition 4.3.2.1(iii), φ and φ^{-1} are C^∞.

Conversely, suppose that $U \subset M$ is open and $\varphi : U \to \varphi(U)$ is diffeomorphism. Then by definition, φ and φ^{-1} are C^∞. For any chart (V, ψ) in maximal atlas $\mathfrak{A}$ of M, we have ψ and ψ^{-1}

are C^∞ from above conclusion. Then by Theorem 4.3.2.1(iii), $\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$ and $\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ are C^∞. This means that (U, φ) is compatible with (V, ψ). Since (V, ψ) is arbitrary, $(U, \varphi) \in \mathfrak{A}$, as desired. $\square$

Proposition 4.4.1.2. *Let M, N and P be smooth manifolds.*

(i) *(Composites) If $F : M \to N$ and $G : N \to P$ are diffeomorphisms, then $G \circ F : M \to P$ is diffeomorphism.*

(ii) *(Restriction) If $U \subset M$ is an open submanifold and $F : M \to N$ is a diffeomorphism, then $F|_U : U \to F(U)$ is a diffeomorphism.*

CHAPTER V

Tangent spaces

§5.1 An intuition: Geometric tangent space

Definition 5.1.1 (Geometric tangent space)**.** Let $p \in \mathbf{R}^n$. We define the *geometric tangent space* of $\mathbf{R}^n$ at p to be the set of all pairs (p, v) for $v \in \mathbf{R}^n$, and denote it by $\mathbf{R}^n_p$. The element of $\mathbf{R}^n_p$ is called the *geometric tangent vector*, we abbreviate (p, v) as v_p.

We think the tangent vector v_p as an arrow goes from the initial point p to the end point $p+v$. Obviously, $\mathbf{R}^n_p$ is a vector space with natural operations

$$v_p + w_p := (v + w)_p \quad \text{and} \quad cv_p := (cv)_p,$$

where $v_p, w_p \in \mathbf{R}^n_p$ and $c \in \mathbf{R}$.

For each geometric tangent vector $v_p \in \mathbf{R}^n_p$, the directional derivative at p gives a map of vector spaces

$$D_v : C^\infty_p \to \mathbf{R},$$

which takes a differentiable function $f \in C^\infty_p$ to a real number $D_v f(p) \in \mathbf{R}$. If we denote $D_v|_p f := D_v f(p)$, then we have the properties

$$D_v|_p(af + bg) = aD_v|_p f + bD_v|_p g$$

and

$$D_v|_p(fg) = g(p)D_v|_p f + f(p)D_v|_p g,$$

where $f, g \in C^\infty_p$ and $a, b \in \mathbf{R}$. We generalize this to following definition.

Definition 5.1.2 (Derivation at point)**.** Let $p \in \mathbf{R}^n$. A linear map $w : C^\infty(\mathbf{R}^n) \to \mathbf{R}$ is called a *derivation* at p of $C^\infty(\mathbf{R}^n)$ if it satisfies

$$w(fg) = f(p)wg + g(p)wf,$$

for every $f, g \in C^\infty(\mathbf{R}^n)$. The set of all derivations at p of $C^\infty(\mathbf{R}^n)$ is denoted by $T_p\mathbf{R}^n$.

Clearly, $T_p\mathbf{R}^n$ is a vector space under the operations

$$(v+w)f := vf + wf \quad \text{and} \quad (cw)f := c(wf),$$

where $v, w \in T_p\mathbf{R}^n$ and $c \in \mathbf{R}$.

Proposition 5.1.2.1. *Let $(e_i)_{1\le i\le n}$ be a standard basis of $\mathbf{R}^n$. Then for every $p \in \mathbf{R}^n$, $(D_j|_p)_{1\le j\le n}$ is a basis of $T_p\mathbf{R}^n$. In particular, we have $\dim T_p\mathbf{R}^n = n$.*

§5.2 Tangent spaces

Definition 5.2.1 (Tangent space)**.** Let M be a smooth manifold, let $p \in M$. A linear map $v : C^\infty(M) \to \mathbf{R}$ is called a *derivation at p* if it satisfies

$$v(fg) = f(p)vg + g(p)vf$$

for every $f, g \in C^\infty(M)$. The set of all derivations of $C^\infty(M)$ at p, denoted by T_pM, is a vector space called the *tangent space to M at p*. The element of T_pM is called a *tangent vector at p*.

Tangent vectors have following properties.

Lemma 5.2.1.1. *Let M be a smooth manifold, let $p \in M$ and $v \in T_pM$, and let $f, g \in C^\infty(M)$.*

(i) *If f is a constant function, then $vf = 0$.*

(ii) *If $f(p) = g(p) = 0$, then $v(fg) = 0$.*

Proof. (i) It suffices to prove that $f(x) = 1$ implies $vf = 0$ since $v(c \cdot 1) = cv(1)$. Let $f \equiv 1$, from the definition, we have $vf = v(ff) = f(p)vf + f(p)vf = 2vf$. This implies that $vf = 0$, as desired.

(ii) From the definition, $v(fg) = f(p)vg + g(p)vf = 0$. □

In the case of a smooth map between Euclidean spaces, the total derivative of the map at a point is a linear map that represents the "best linear approximation" to the map near the given point. In the manifold case there is a similar linear map, but it makes no sense to talk about a linear map between manifolds. Instead, it will be a linear map between tangent spaces.

Definition 5.2.2 (Differential)**.** Let M and N be smooth manifolds, and let $F : M \to N$ be a smooth map. The *differential of F at p* is a map $dF_p : T_pM \to T_{F(p)}N$ as follows. Given $v \in T_pM$, we let $dF_p(v)$ be the derivation at $F(p)$ that acts on $f \in C^\infty(N)$ by the rule

$$dF_p(v)(f) := v(f \circ F).$$

Remark. Note that if $f \in C^\infty(N)$, then $f \circ F : M \to \mathbf{R}$ is also smooth (thanks for Theorem 4.3.2.1). Thus $f \circ F \in C^\infty(M)$, and $v(f \circ F)$ makes sense.

Some properties of differential are given by following.

Proposition 5.2.2.1. *Let M, N and P be smooth manifolds, and let $F : M \to N$ and $G : N \to P$ be smooth maps. Let $p \in M$. Then*

(i) $dF_p : T_pM \to T_{F(p)}N$ *is indeed a derivation at* $F(p)$.

(ii) $dF_p : T_pM \to T_{F(p)}N$ *is linear.*

(iii) $d(G \circ F)_p = dG_{F(p)} \circ dF_p : T_pM \to T_{G\circ F(p)}P$.

(iv) $d(\mathbf{1}_M)_p = \mathbf{1}_{T_pM} : T_pM \to T_pM$.

(v) *If F is a diffeomorphism, then $dF_p : T_pM \to T_{F(p)}N$ is an isomorphism, and* $(dF_p)^{-1} = d(F^{-1})_{F(p)}$.

Proof. (i) Let $v \in T_pM$ and let $f, g \in C^\infty(N)$. Then

$$\begin{aligned} dF_p(v)(fg) &= v((fg)\circ F) \\ &= v((f\circ F)(g\circ F)) \\ &= (f\circ F)(p)v(g\circ F) + (g\circ F)(p)v(f\circ F) \\ &= f(F(p))dF_p(v)(g) + g(G(p))dF_p(v)(f). \end{aligned}$$

(ii) Let $v \in T_pM$ and let $f, g \in C^\infty(N)$ and $\alpha, \beta \in \mathbf{R}$. Then

$$\begin{aligned} dF_p(v)(\alpha f + \beta g) &= v((\alpha f + \beta g)\circ F) \\ &= v(\alpha(f\circ F) + \beta(g\circ F)) \\ &= \alpha v(f\circ F) + \beta v(g\circ F) \\ &= \alpha dF_p(v)(f) + \beta dF_p(v)(g). \end{aligned}$$

(iii) Let $v \in T_pM$ and let $f \in C^\infty(P)$ be a smooth map at $G(F(p))$. Then

$$\begin{aligned} d(G\circ F)_p(v)(f) &= v(f\circ(G\circ F)) \\ &= v((f\circ G)\circ F) \\ &= dF_p(v)(f\circ G) \\ &= dG_{F(p)}(dF_p(v))(f) \\ &= (dG_{F(p)}\circ dF_p)(v)(f). \end{aligned}$$

(iv) Let $v \in T_pM$ and $f \in C^\infty(M)$. Then

$$d(\mathbf{1}_M)_p(v)(f) = v(f\circ \mathbf{1}_M) = vf = (\mathbf{1}_M\, v)f.$$

(v) Suppose $F : M \to N$ is diffeomorphism, then it has an inverse $F^{-1} : N \to M$ such that $F\circ F^{-1} = \mathbf{1}_N$ and $F^{-1}\circ F = \mathbf{1}_M$. Then by (iii) and (iv) shown in above, we have

$$dF_p \circ d(F^{-1})_{F(p)} = d(F\circ F^{-1})_{F(p)} = d(\mathbf{1}_N)_{F(p)} = \mathbf{1}_{T_pN},$$

and

$$d(F^{-1})_{F(p)} \circ dF_p = d(F^{-1}\circ F)_p = d(\mathbf{1}_M)_p = \mathbf{1}_{T_pM}\,.$$

Thus, dF_p is an isomorphism, and $(dF_p)^{-1} = d(F^{-1})_{F(p)}$. □

We now can prove a weak version of *invariance of dimension* (see Theorem 4.1.1.1).

Lemma 5.2.2.1. *If an open set $U \subset \mathbf{R}^n$ is diffeomorphic to an open set $V \subset \mathbf{R}^m$, then $n = m$.*

Proof. Let $F : U \to V$ be a diffeomorphism and let $p \in U$. By Proposition 5.2.2.1(v), dF_p is an isomorphism of tangent spaces. Since there are vector space isomorphisms $T_pU \simeq \mathbf{R}^n$ and $T_{F(p)}V \simeq \mathbf{R}^m$, we must have that $m = n$. □

Theorem 5.2.2.1 (Diffeomorphism invariance of dimension)**.** *A nonempty m-dimensional smooth manifold cannot be diffeomorphic to an n-dimensional smooth manifold unless $m = n$.*

Proof. Suppose M is a nonempty smooth m-manifold, N is a nonempty n-manifold, and $F : M \to N$ is a diffeomorphism. Choose any point $p \in M$, and let (U, φ) and (V, ψ) be smooth coordinate charts containing p and $F(p)$, respectively. Then $\psi \circ F \circ \varphi^{-1}$ is a diffeomorphism from an open subset of $\mathbf{R}^m$ to an open subset of $\mathbf{R}^n$, so it follows from Lemma 5.2.2.1 that $m = n$. □

§5.3 Applications

Bases

Let M be a smooth n-manifold, and let (U, φ) be a smooth coordinate chart on M. By Proposition 4.4.1.1, $\varphi : U \to \varphi(U)$ is a diffeomorphism, then by Proposition 5.2.2.1(v), we see that $d\varphi_p : T_pM \to T_{\varphi(p)}\mathbf{R}^n$ is a vector space isomorphism.

By Proposition 5.1.2.1, $(D_j|_{\varphi(p)})_{1\leq j\leq n}$ is a basis of $T_p\mathbf{R}^n$. Thus, T_pM has same dimension n with $T_p\mathbf{R}^n$. Therefore, the preimage of $(D_j|_{\varphi(p)})_{1\leq j\leq n}$ under the isomorphism form a basis of T_pM. Specifically, we denote

$$D_j|_p := (d\varphi_p)^{-1}(D_j|_{\varphi(p)}).$$

Then for $f \in C^\infty(\mathbf{R}^n)$, we have

$$\begin{aligned} d\varphi_p(D_j|_p)f &= D_j|_p(f \circ \varphi) \\ &= (d\varphi_p)^{-1}(D_j|_{\varphi(p)})(f \circ \varphi) \\ &= d(\varphi^{-1})_{\varphi(p)}(D_j|_{\varphi(p)})(f \circ \varphi) \\ &= D_j|_{\varphi(p)}(f \circ \varphi \circ \varphi^{-1}) \\ &= D_j|_{\varphi(p)}(f). \end{aligned}$$

Jacobian matrix

Consider a smooth map $F : U \to V$, where $U \subset \mathbf{R}^n$ and $V \subset \mathbf{R}^m$ are open subsets of Euclidean spaces. For any $p \in U$, we determine the matrix of $dF_p : T_p\mathbf{R}^n \to T_{F(p)}\mathbf{R}^m$ in terms of the standard bases. For every $f \in C^\infty(V)$, we use the chain rule and obtain

$$dF_p(D_j|_p)f = D_j|_p(f \circ F) = \sum_{i=1}^m D_i|_{F(p)}f D_j|_pF_i = \Big(\sum_{i=1}^m D_j|_pF_iD_i|_{F(p)}\Big)f.$$

Thus we have

$$dF_p(D_j|_p) = \sum_{i=1}^{m} D_j|_p F_i D_i|_{F(p)}$$

and the Jacobian matrix is given by

$$\begin{pmatrix} D_1|_p F_1 & D_2|_p F_1 & \cdots & D_n|_p F_1 \\ D_1|_p F_2 & D_2|_p F_2 & \cdots & D_n|_p F_2 \\ \vdots & \vdots & \ddots & \vdots \\ D_1|_p F_m & D_2|_p F_m & \cdots & D_n|_p F_m \end{pmatrix}.$$

CHAPTER VI

Submersions, immersions, and embeddings

§6.1 Maps of constant rank

Definition 6.1.1. Let M and N be smooth manifolds. Let $F : M \to N$ be a smooth map, and let $p \in M$.

- We define the *rank of F at p* to be the rank of the linear map $dF_p : T_pM \to T_{F(p)}N$. If F has same rank r at every point of M, then we say that it has *constant rank*, and denote $\operatorname{rank} F = r$.
- The smooth map F is called a *smooth submersion* if F_p is surjective at every point of M, i.e., $\operatorname{rank} F = \dim N$.
- The smooth map F is called a *smooth immersion* if F_p is injective at every point of M, i.e., $\operatorname{rank} F = \dim M$.

Theorem 6.1.1.1 (Rank theorem)**.** *Let M and N be smooth manifold of dimension m and n, respectively. Let $F : M \to N$ be a smooth map with constant rank r. For every $p \in M$ there exist charts (U, φ) for M centered at p and (V, ψ) for N centered at $F(p)$ such that $F(U) \subset V$ and*

$$\psi \circ F \circ \varphi^{-1}(x_1, \cdots, x_r, x_{r+1}, \cdots, x_m) = (x_1, \cdots, x_r, \underbrace{0, \cdots, 0}_{m \text{ terms}}).$$