pdftitle=Assignment, colorlinks=true, linkcolor=doc!90, bookmarksnumbered=true, bookmarksopen=true

NCKU 112.1
Rudin

Eric Liu

Contents

Chapter 1 The Real and Complex Number System Page 2

1.1 Introduction 2
1.2 Ordered Fields 5
1.3 Real Numbers Field 8
1.4 Irrational Power 15
1.5 Euclidean Space 22
1.6 Complex Numbers 24
1.7 Existence of Real Numbers - Dedekind Cut* 32
1.8 Existence of Real Numbers - Decimal* 32
1.9 Uniqueness of Real Numbers* 32
1.10 Exercises 32
1.11 Complicated Exercises 34

Chapter 2 Basic Topology Page 36

2.1 ZF 36
2.2 Ordinal 43
2.3 Axiom of Choice and its equivalents 49
2.4 Unfinished Notes for Cardinal 53
2.5 Metric Space 57

Chapter 1

The Real and Complex Number System

1.1 Introduction

In this section, we will define the concept of ordered sets, and give a close look of the completeness property of real numbers, by showing the "uncompleteness" of rational numbers. First, we prove an elementary and classic theorem of rational numbers.

Theorem 1.1.1. There exists no rational p such that $p^2 = 2$

Proof. Assume there is, and we write p in the form $p = \frac{a}{b}$, where $a, b \in \mathbb{Z}$ and one of a, b is odd. Observe that $p^2 = 2 \implies a^2 = 2b^2 \implies 2$ divides $a \implies 2$ divides b CaC ■

For the sake of our discussion below, we will use $\mathbb{Q}^+$ instead of $\mathbb{Q}$ as our universal. Now, we divide the "rational numbers line" in half at the point $\sqrt{2}$, and have two subdivisions $A = \{x \in \mathbb{Q}^+ : x^2 < 2\}, B = \{x \in \mathbb{Q}^+ : x^2 > 2\}$

Axiom 1.1.2. Let S be an ordered set and X be a subset of S. Axiomatically define

$$\max X \in X \text{ and } \forall x \in X, x \leq \max X \tag{1.1}$$

$$\min X \in X \text{ and } \forall x \in X, \min X \leq x \tag{1.2}$$

Theorem 1.1.3. $\max A$ and $\min B$ both doesn't exist.

Proof. We wish to construct a function $q(p)$ on $\mathbb{Q}^+$ such that for all $p \in \mathbb{Q}^+$, we have $p^2 < 2 \implies p^2 < q^2 < 2$ and have $2 < p^2 \implies 2 < q^2 < p^2$. Notice $p < q \iff p^2 < q^2$, so we can translate the wanted property of q into

$$p^2 < 2 \implies p < q \text{ and } q^2 < 2 \tag{1.3}$$

$$p^2 > 2 \implies q < p \text{ and } 2 < q^2 \tag{1.4}$$

Let a, b be two function of p on $\mathbb{Q}^+$. To satisfy the above properties, we can let the below equation first be true and then solve for a, b.

$$q - p = \frac{2 - p^2}{a} \tag{1.5}$$

$$q^2 - 2 = \frac{p^2 - 2}{b} \tag{1.6}$$

Now we solve for a, b

$$q - p = \frac{2 - p^2}{a} \text{ and } q^2 - 2 = \frac{p^2 - 2}{b} \implies q = \frac{2 - p^2}{a} + p \text{ and } q^2 = \frac{p^2 - 2}{b} + 2 \tag{1.7}$$

$$\iff [\frac{2 - p^2}{a} + p]^2 = \frac{p^2 - 2}{b} + 2 \tag{1.8}$$

$$\iff \frac{(2 - p^2)^2}{a^2} + \frac{2p(2 - p^2)}{a} + p^2 = \frac{p^2 - 2}{b} + 2 \tag{1.9}$$

$$\iff (p^2 - 2)^2(\frac{1}{a^2}) + (p^2 - 2)(-\frac{2p}{a} + 1 - \frac{1}{b}) = 0 \tag{1.10}$$

$$\iff \frac{p^2 - 2}{a^2} - \frac{2p}{a} + 1 - \frac{1}{b} = 0 \text{ (because } p^2 - 2 \neq 0) \tag{1.11}$$

$$\iff \frac{p^2 - 2 - 2ap + a^2}{a^2} = \frac{1}{b} \tag{1.12}$$

$$\iff b = \frac{a^2}{p^2 - 2 - 2ap + a^2} = \frac{a^2}{(p - a)^2 - 2} \tag{1.13}$$

Define $a := p + c$ where $c^2 > 2$, and we are finished. ■

Now, we come back to define the concept of ordered set.

Definition 1.1.4. (Totally Ordered Set Axioms) S is an totally ordered set, sometimes simply called ordered set depend on the context, if there is a relation $\sim$ on it that satisfy

$$\forall x \in S, x \sim x \tag{1.14}$$

$$\forall x, y \in S, x \sim y \text{ or } y \sim x \tag{1.15}$$

$$\forall x, y, z \in S, x \sim y \text{ and } y \sim z \implies x \sim z \tag{1.16}$$

$$\forall x, y \in S, x \sim y \text{ and } y \sim x \implies x = y \tag{1.17}$$

From now on, we write this relation as $\leq$ and if we write $x \geq y$, we mean $y \leq x$.

To discuss the uncompleteness of rational numbers, which is an ordered set, we first define a few concepts brought by concept of ordered set.

Definition 1.1.5. Let S be an ordered set and $E \subseteq S$. E is bounded above if

$$\exists a \in S, \forall b \in E, a \geq b \tag{1.18}$$

In this case, we say a is an upper bound of E and E is bounded above by a. On the other hand, E is bounded below by c if

$$\forall b \in E, c \leq b \tag{1.19}$$

Before we give the definition of supremum, aka least upper bound, we first prove a theorem about it.

Theorem 1.1.6. If x is the smallest upper bound of a bounded above nonempty set E, then any number smaller than x is not an upper bound of E, and every upper bound of E is greater than or equal to x.

Proof. Let A be the set of upper bounds of E. Arbitrarily pick an m such that $m < x$. Assume $\forall z \in E, m > z$. We see $m \in A$, and because $x = \min A$, we see $x \leq m$ CaC . Assume $\exists n \in A, n < x$. Then $\exists z \in E, n < z$, which implies $n \notin A$ CaC ■

Corollary 1.1.7. Any number smaller than the greatest lower bound is a lower bound, and any number greater than the greatest lower bound is not a lower bound.

Definition 1.1.8. (Definition of Supremum and Infimum) Let A, B respectively be the set of upper bounds of E and the set of lower bounds of E. We define

$$\sup E := \min A \tag{1.20}$$

and

$$\inf E := \max B \tag{1.21}$$

if they ever exist.

Theorem 1.1.9. Let A be a subset of ordered set S. If $\max A$ exists, then $\max A = \sup A$. Similarly, if $\min A$ exists, then $\min A = \inf A$

Proof. $\max A$ is an upper bound and $\min A$ is an lower bound hold true by definition. Any number smaller than $\max A$ can not be an upper bound since $\max A \in A$. Similarly, and number greater than $\min A$ can not be an lower bound since $\min A \in A$ ■

Now, we look back to our subdivisions $A = \{x \in \mathbb{Q}^+ : x^2 < 2\}, B = \{x \in \mathbb{Q}^+ : x^2 > 2\}$. We first show that B is exactly the set of all upper bounds of A.

Theorem 1.1.10. Given $A = \{x \in \mathbb{Q}^+ : x^2 < 2\}, B = \{x \in \mathbb{Q}^+ : x^2 > 2\}$. We have

$$B = \{x \in \mathbb{Q}^+ : \forall y \in A, y \leq x\} \tag{1.22}$$

Proof. Arbitrarily pick $x \in B$, and we see $\forall y \in A, y^2 < 2 < x^2$. Then $\forall y \in A, y < x$. This implies $B \subseteq \{x \in \mathbb{Q}^+ : \forall y \in A, y \leq x\}$. Let x satisfy $\forall y \in A, y \leq x$. Assume $x^2 < 2$. We immediately see $x = \max A$, but $\max A$ doesn't exist CaC ■

By Theorem 1.1.3, we then can see that $\sup A$ does not exist. Now, we give this idea a name.

Definition 1.1.11. (Definition of Completed Ordered Set) An ordered set S satisfy least-upper-bound property if

$$E \subseteq S \text{ and } E \neq \varnothing \implies \sup E \text{ exists} \tag{1.23}$$

Also, we say S is completed.

$\sup A$ does not exists indicate that $\mathbb{Q}$ as an ordered set doesn't satisfy the least-upper-bound property. Before we close this section, we reveal the face of the twin brother of the least-upper-bound property. In fact, it is more like property itself in the mirror, since they are equivalent.

Theorem 1.1.12. (LUB $\iff$ GLB) S satisfy the least-upper-bound property if and only if S satisfy the greatest-lower-bound property.

Proof. From left to right, consider a bounded below set E. Let L be the set of lower bounds of E. We know $\sup L$ exists and every elements of E is an upper bound of L, so $\sup L$ is an lower bound of E. Then $\sup L = \inf E$. The other direction use the same method. ■

1.2 Ordered Fields

In this section, we first give the definition of ordered fields, and prove basic result concerning positivity. Notice in this section that x, y, z are all in $\mathbb{F}$.

Axiom 1.2.1. (Ordered Field Axioms) An ordered field $\mathbb{F}$ is a field that is not only a ordered set, but also satisfy the following axioms

$$y < z \implies x + y < x + z \tag{1.24}$$

$$x > 0 \text{ and } y > 0 \implies xy > 0 \tag{1.25}$$

Theorem 1.2.2. (Negate reverse positivity) $(x > 0 \iff -x < 0)$ and $(x < 0 \iff -x > 0)$

Proof. Observe $x > 0 \implies x+(-x) > 0+(-x) \implies -x < 0$, and $x < 0 \implies x+(-x) < 0+(-x) \implies -x > 0$. Clearly, $x = 0 \implies -x = 0$. Then the Theorem follows. ■

Theorem 1.2.3. (Multiply a negative number reverse positivity) Given $y < 0$, we have

$$(x > 0 \iff xy < 0) \text{ and } (x < 0 \iff xy > 0) \tag{1.26}$$

Proof. Observe $x > 0 \implies x(-y) > 0 \implies -xy > 0 \implies xy < 0$, and $x < 0 \implies (-x)(-y) > 0 \implies xy > 0$. Clearly, $x = 0 \implies xy = 0$. Then the Theorem follows. ■

Theorem 1.2.4. (Multiply on both side) Given $y < z$, we have

$$(x > 0 \iff xy < xz) \text{ and } (x < 0 \iff xy > xz) \tag{1.27}$$

Proof. First observe $y < z \implies z - y > 0$. Now observe $x > 0 \implies x(z - y) > 0 \implies xz - xy > 0 \implies xz > xy$. Observe $x < 0 \implies x(z - y) < 0 \implies xz < xy$. Obviously $x = 0 \implies xy = xz$. Then the Theorem follows. ■

Theorem 1.2.5. (Squares are nonnegative) $x \neq 0 \implies x^2 > 0$

Proof. If $x > 0$, then $x^2 > 0$ follows from the axiom. If $x < 0$, then $x^2 > 0$ follows from Theorem 1.2.2 ■

Corollary 1.2.6. $1 > 0$

Theorem 1.2.7. (Inverse preserve positivity) $(x > 0 \iff \frac{1}{x} > 0)$ and $(x < 0 \iff \frac{1}{x} < 0)$

Proof. If $x > 0$ but $\frac{1}{x} < 0$, then $1 < 0$. The same logic applies to $\frac{1}{x} > 0 \implies x > 0$. Because $\frac{1}{0}$ does not exist, the theorem follows. ■

If we were to define $x^{-2} := x^{-1}x^{-1}$, we must realize 0^{-1} does not exist and realize we have not yet prove some common inequalities concerning integer powers. Notice that the following inequalities require base to be positive.

Definition 1.2.8. (Definition of Inverse) For all nonzero x and naturals p, we define $x^{-p} := (x^{-1})^p$, and define $x^0 := 1$

Theorem 1.2.9. (Inequality when base is fixed) Given a positive a and two integer x, y where $x < y$, we have

$$\begin{cases} a^x < a^y \iff a > 1 \\ a^x = a^y \iff a = 1 \\ a^x > a^y \iff 0 < a < 1 \end{cases} \tag{1.28}$$

Proof. By Theorem 1.2.4, observe $1 < a \iff 1 < a < a^2 \iff 1 < a < a^2 < a^3 \iff \cdots \iff 1 < a < \cdots < a^{y-x} \iff a^x < a^y$. If $a = 1$, we know $a^x = 1 = a^y$. Again by Theorem 1.2.4, observe $0 < a < 1 \iff a^2 < a < 1 \iff \cdots \iff a^{y-x} < \cdots < 1 \iff a^y < a^x$. ■

Theorem 1.2.10. (Inequality when integer power is fixed) Given $0 < b < c$ and $z \in \mathbb{Z}$, we have

$$\begin{cases} b^z < c^z \iff 0 < z \\ b^z = c^z \iff 0 = z \\ b^z > c^z \iff 0 > z \end{cases} \tag{1.29}$$

Proof. If $z > 0$, then by Theorem 1.2.4, observe $0 < b < c \implies 0 < b^2 < bc < c^2 \implies b^3 < b^2c < bc^2 < c^3 \implies \cdots \implies b^z < c^z$. Obviously, $0 = z \implies b^z = 1 = c^z$. If $z < 0$, then $b^z = (\frac{1}{b})^{-z}$ and $c^z = (\frac{1}{c})^{-z}$ by definition. Observe $b < c \implies 1 = \frac{1}{b}b < \frac{1}{b}c = \frac{c}{b} \implies \frac{1}{c} < \frac{1}{b}$. By the logic above, we deduce $c^z = (\frac{1}{c})^{-z} < (\frac{1}{b})^{-z} = b^z$. Then the Theorem follows. ■

Theorem 1.2.11. (Positivity of integer power) If $a > 0$, then $\forall x \in \mathbb{Z}, a^x > 0$. If $a < 0$, then $a^x > 0 \iff 2|x$

Proof. The former result is a direct consequence of Axiom 1.2.1 and Theorem 1.2.7. The latter result is a direct consequence of Theorem 1.2.3 and Theorem 1.2.5 ■

Notice that if the base is negative, the corresponding inequalities can all be deduced by the above theorems with a little effort, although the results are quite messy.

Now we prove some arithmetic properties concerning nonzero base, unlike the above inequalities concerning only positive base. Notice that those properties of natural power inherited by integer power can be proven inductively, and that the base x can be any nonzero number.

Theorem 1.2.12. (Arithmetic property of integer power) For all nonzero x and integers p, q, we have

$$x^{p+q} = x^p x^q \tag{1.30}$$

Proof. If p, q are both positive, the Theorem is proven by induction. If p, q are both negative, observe $x^p x^q = (x^{-1})^{-p}(x^{-1})^{-q} = (x^{-1})^{-(p+q)} = x^{p+q}$, where the last equality hold true because $p+q < 0$. If $p > 0 > q$ and $p+q > 0$, observe $x^p x^q = x^p(x^{-1})^{-q} = x^{p-(-q)} = x^{p+q}$. If $p > 0 > q$ and $p+q < 0$, observe $x^p x^q = x^p(x^{-1})^{-q} = (x^{-1})^{-(q+p)} = x^{p+q}$, where the last equality hold true because $p + q < 0$. ■

Theorem 1.2.13. (Arithmetic property of integer power) For all nonzero x and integers p, q, we have

$$x^{pq} = (x^p)^q = (x^q)^p \tag{1.31}$$

Proof. If p, q are both positive or if any of p, q are zero, the proof is trivial. If $p < 0 < q$, observe $(x^p)^q = ((x^{-1})^{-p})^q = (x^{-1})^{-pq} = x^{pq}$, where the last equality hold true because $pq < 0$, and observe $(x^q)^p = ((x^q)^{-1})^{-p} = ((x^{-1})^q)^{-p} = (x^{-1})^{-qp} = x^{qp}$, where the second equality hold true can be proven by induction. ■

Theorem 1.2.14. (Arithmetic property of integer power) For all nonzero x, y and integer p, we have

$$x^p y^p = (xy)^p \tag{1.32}$$

Proof. If $p \geq 0$, the proof is trivial. If $p < 0$, then $x^p y^p = (x^{-1})^{-p}(y^{-1})^{-p} = (x^{-1}y^{-1})^{-p} = ((xy)^{-1})^{-p} = (xy)^p$. ■

1.3 Real Numbers Field

Although the title of this section is "Real Numbers Field", here, we will not construct the real numbers field, nor use any common property of real numbers. In fact, we will not even use the symbol $\mathbb{R}$ in this section, since we are merely proving theorems about an ordered field with least-upper-bound property. We don't know if there exists any ordered field with least-upper-property. Let's say there does; yet, we don't know if such structure is unique. Let's say it is unique; yet, we don't know if that structure have relation with $\mathbb{R}$. Here, we will use the symbol $\mathbb{F}$ to denote an ordered field with least-upper-bound property. One should realize that we can use algorithm to define a subset containing $1 \in \mathbb{F}$ that is isomorphic to $\mathbb{N}$, and thereby we abuse the notation to denote that subset $\mathbb{N}$. A subfield of $\mathbb{F}$ isomorphic to $\mathbb{Q}$ can also be defined after we define $\mathbb{Z}$, so we also thereby abuse the notation to denote that subfield $\mathbb{Q}$.

Theorem 1.3.1. $\mathbb{N}$ is unbounded above.

Proof. Assume $\mathbb{N}$ is bounded above. Because $1 > 0$, we know $\sup \mathbb{N} - 1 < \sup \mathbb{N}$. Then $\sup \mathbb{N} - 1$ is not an upper bound of $\mathbb{N}$. Arbitrarily pick any $m \in \mathbb{N}$ greater than $\sup \mathbb{N} - 1$. We see $m > \sup \mathbb{N} - 1 \implies m + 1 > \sup \mathbb{N}$, where $m + 1 \in \mathbb{N}$ CaC ■

Corollary 1.3.2. Both $\mathbb{Z}$ and $\mathbb{Q}$ are unbounded both above and below.

Corollary 1.3.3. (Divided by 1) Given any $x \in \mathbb{F}$, there exists $n \in \mathbb{Z}$ such that $n \leq x < n + 1$

Proof. If $x > 0$, let $S = \{n \in \mathbb{N} : n > x\}$. Notice $S = \varnothing$ implies $\mathbb{N}$ is bounded above by x, so S is nonempty. Then by well-ordering principle, we know $\min S$ exists. We now show $\min S - 1 \leq x < \min S$. Observe that $\min S \in S \implies x < \min S$. Assume $\min S - 1 > x$. We immediately see $\min S - 1 \in S$ CaC (done) .

If $x < 0$, let $S = \{n \in \mathbb{N} : n \geq -x\}$. Again, $S = \varnothing$ implies $\mathbb{N}$ is bounded above by $-x$, so S is nonempty. Then by well-ordering principle, we know $\min S$ exists. We now show $-\min S \leq x < -\min S + 1$. Observe that $\min S \in S \implies \min S \geq -x \implies x \geq -\min S$. Assume $-\min S + 1 \leq x$. Then $\min S - 1 \geq -x > 0$; thus

$\min S - 1 \in S$ CaC (done)

If $x = 0$, then we let $n = 0$. ■

Theorem 1.3.4. (Archimedean Property) Given $x, y \in \mathbb{F}$ and $0 < x$, there exists $n \in \mathbb{N}$ such that $nx > y$

Proof. Because $\mathbb{N}$ is unbounded above, we know $\frac{y}{x}$ can not be an upper bound of $\mathbb{N}$, so we know $\exists n \in \mathbb{N}, n > \frac{y}{x}$. Then because $x > 0$, we can deduce $nx > y$. ■

Theorem 1.3.5. ($\mathbb{Q}$ is dense in $\mathbb{F}$) Given $x, y \in \mathbb{F}$ and $x < y$, we know there exists $p \in \mathbb{Q}$ such that $x < p < y$

Proof. Every rational, positive or negative, can be expressed in the form $\frac{m}{n}$ for some integer m and naturals n. We seek to find some integer m and n such that $x < \frac{m}{n} < y$. Notice that $x < \frac{m}{n} < y \iff nx < m < ny$. Because m has to be an integer, we know for $nx < m < ny$ to hold true, we must first have $ny - nx > 1$. Because $y - x > 0$, by Archimedean Property, there exists $n \in \mathbb{N}$ such that $ny - nx = n(y - x) > 1$. By Corollary 1.3.3, we know there exists $m \in \mathbb{Z}$ such that $m \leq ny < m + 1$.

Notice $m = ny$ if and only if $y \in \mathbb{Q}$. So we can split the proof into two cases.

Case 1: $y \in \mathbb{Q}$

We see that the set $\{r \in \mathbb{Q} : r < y\}$ have supremum y, since $y \in \mathbb{Q}$. Then $x < y$ tell us x is not an upper bound of the set, then we can pick some rational r in the set greater than x, so $x < r < y$ (done) .

Case 2: $y \notin \mathbb{Q}$

We know $m < ny < m + 1$. $ny < m + 1$ tell us $nx < ny - 1 < m$, so $nx < m < ny$ (done) ■

Theorem 1.3.6. (Positive root of power uniquely exists) For all natural n and $y > 0$, there exists a unique positive x such that $x^n = y$

Proof. By Theroem 1.2.10, we know two different positive numbers $0 < x < x'$ are different when raised to the power of n, being $0 < x^n < (x')^n$, so if such positive power exists, it must be unique.

We have handled the uniqueness part of the proof. Denote $E := \{m \in \mathbb{F}^+ : m^n < y\}$ and $x := \sup E$. Now we do the existence part by proving x exists and $x^n = y$.

To show $x = \sup\{m \in \mathbb{F}^+ : m^n < y\}$ exists, we only have to show the set $\{m \in \mathbb{F}^+ : m^n < y\}$ is nonempty and bounded above. In other word, we wish to construct function $a \in \mathbb{F}^+$ and b of y such that for all positive input $y > 0$, we have $a^n < y$ and $(0 < m^n < y \longrightarrow m < b)$. In the followings, the domain of a and b are only positives.

First we construct a. By Theorem 1.2.9, we know if $a < \min\{1, y\}$, then $a^n < a < y$, so we construct a such that $0 < a < \min\{1, y\}$. Notice that a must be positive because we are constructing a number in E, where E contain only positives. Express a in the form $a = \frac{p}{q}$ where p, q are both function of y. In the process of construction, We must be careful to make sure a exists for all positive y.

To satisfy $0 < a$, we need only guarantee p, q are always of the same sign for all positive y. If such p, q exists, we can change both sign of p, q when they are negative, and get two positive function. So, we can just require p, q to be positive for all positive y.

To satisfy $a < 1$, observe $a < 1 \iff \frac{p}{q} < 1 \iff p < q$. The easiest construction is to let $q = p + c$ where c is positive.

To satisfy $a < y$, observe $a < y \iff \frac{p}{q} < y \iff p(y-1) + cy = qy - p > 0$. The easiest construction is to let $p(y-1) + cy = y^2$, which is possible, if we let $p = y$ and $c = 1$. In this case $p = y > 0$ and $q = y + 1 > 0$, and $a = \frac{y}{y+1}$. We finished proving E is nonempty. **(Notice $c = y^2, p = y^3, q = y^3 + y^2, a = \frac{y^3}{y^3+y^2}$ also do the trick)**

Now we construct b. By Theorem 1.2.10, we know if $0 < b$ and $0 < m^n < b^n$, then $m < b$, so we construct b such that $y < b^n$ which lead to $0 < m^n < y < b^n$ if $m^n < y$. Because $y > 0$, this is fairly easy. Simply let $b = y + 1$, so we have $b > 1$ and $b > y$; thus by Theorem 1.2.9, we have $b^n > b > y$, finishing proving E is bounded above, where $b = y + 1$ is an upper bound. (done)

To show $x^n = y$, we show $x^n \geq y$ and $x^n \leq y$. We will assume that $x^n < y$ or $x^n > y$, but before we do such, let's see what property from which can we possibly draw contradiction. Notice that because we just prove the existence of the supremum of E, we haven't use the fact that $x = \sup E$ in anywhere of our proof. We know

$$x = \sup E \iff \forall d > 0, \begin{cases} x + d \notin E \ (x \text{ is an upper bound}) \\ \text{and} \\ x - d \text{ is not an upper bound of } E \text{ (the } \textit{least} \text{ upper bound)} \end{cases} \tag{1.33}$$

So, you see, we wish to construct a small and positive h and k such that if we assume

$x^n < y$ or $x^n > y$ we can draw $x+h \in E$ or $x-k$ is an upper bound of E. **(We are going to assume** $\sup E$ **is smaller or greater than** $\sqrt[n]{y}$**)**

Observe $x+h \in E \iff (x+h)^n < y \iff (x+h)^n - x^n < y - x^n$, and observe $x-k$ is an upper bound of $E \iff (m^n < y \longrightarrow m < x-k) \iff (m \geq x-k \longrightarrow m^n \geq y) \iff (m \geq x-k \longrightarrow x^n - m^n \leq x^n - y)$.

Notice that the act of subtracting x^n at the both side of the inequality play an important role in our proof: not only does the act allow us to use the identity $a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + b^{n-1})$, and the act also tell us between $x+h \in E$ and $x-k$ is an upper bound of E, which contradiction statement should we draw from $x^n > y$. If $y - x^n < 0$, then $y - x^n < 0 < (x+h)^n - x^n$, so we can not possibly draw $x+h \in E$ from $x^n > y$. **(If** $\sup E > \sqrt[n]{y}$**, then it is too big, we can find a smaller upper bound of** E**)**

Assume $x^n > y$. We wish to construct positive k such that $m \geq x-k \longrightarrow x^n - m^n \leq x^n - y$, so we can draw the contradiction $x-k$ is an upper bound of E.

Notice that $m \geq x-k \implies x^n - m^n \leq x^n - (x-k)^n$, so if $x^n - (x-k)^n \leq x^n - y$, our proof at this part is finished. Now our job is to single out the k in the inequality to give an condition such that $x^n - (x-k)^n \leq x^n - y$ hold if the condition hold. It is easy to see that computing the polynomial of k on left hand side of inequality and to prove such positive k exists for all n is almost impossible. Thus, we take a possible but actually non-existing risk in our next step of the proof. Use the $a^n - b^n$ identity and that $x-k < x$ to deduce

$$x^n - (x-k)^n = k(x^{n-1} + x^{n-2}(x-k) + \cdots + (x-k)^{n-1}) \leq knx^{n-1} \tag{1.34}$$

So, if we have $knx^{n-1} \leq x^n - y$, which is equivalent to $k \leq \frac{x^n - y}{nx^{n-1}}$, our proof is partially finished. Notice that $x^n - y > 0$ and $nx^{n-1} > 0$, so $\frac{x^n - y}{nx^{n-1}} > 0$; thus the positive k exists. (done) CaC **(The reason I use the word "possible risk" is that if we use an identity that show us** $x^n - (x-k)^n$ **smaller than a quantity greater than** $x^n - y$**, the proof can not be done.)**

Assume $x^n < y$. We wish to construct positive h such that $(x+h)^n - x^n < y - x^n$, so we can draw the contradiction $x+h \in E$.

Again, we use the same identity to deduce

$$(x+h)^n - x^n = h((x+h)^{n-1} + (x+h)^{n-2}x + \cdots + x^{n-1}) < hn(x+h)^{n-1} \tag{1.35}$$

To single out the h in the $hn(x+h)^{n-1}$, notice that we can take the risk to add the constraint $h < 1$ at the end of our construction to have $(x+h)^n - x^n < hn(x+h)^{n-1} < hn(x+1)^{n-1}$.

Then, if we have $hn(x+1)^{n-1} < y - x^n$, which is equivalent to $h < \frac{y-x^n}{n(x+1)^{n-1}}$, our proof is finished. To sum up, any h satisfy $h < \min\{1, \frac{y-x^n}{n(x+1)^{n-1}}\}$ does the trick, and such h exists, since $0 < \min\{y - x^n, n(x+1)^{n-1}\}$. (done) CaC (done) ■

If you want a proof with less of my commentary, here you go.

Proof. Observe $\min\{1, \frac{y}{2}\} \in E$ and $\max\{1, y\}$ is an upper bound of E, so $\sup E$ exists. Denote $x := \sup E$. Assume $x^n > y$. Observe $x - k$ is an upper bound of $E \iff m^n < y \longrightarrow m < x - k \iff m \geq x - k \longrightarrow m^n \geq y$. We know $(x-k)^n \geq y \iff x^n - (x-k)^n \leq x^n - y$, and know $x^n - (x-k)^n \leq knx^{n-1}$. If we let $0 < k \leq \frac{x^n - y}{nx^{n-1}}$, then $x - k$ is an upper bound of E CaC . Assume $x^n < y$. Observe $x + h \in E \iff (x+h)^n - x^n < y - x^n$. We know that $(x+h)^n - x^n < hn(x+h)^{n-1}$ and that if $h < 1$, we have $hn(x+h)^{n-1} < hn(x+1)^{n-1}$. So we let $h = \min\{1, \frac{y-x^n}{n(x+1)^{n-1}}\}$, and we can see $x + h \in E$ CaC . ■

Definition 1.3.7. The number x in Theorem 1.3.6 is written $x = \sqrt[n]{y}$ or $x = y^{\frac{1}{n}}$.

The above theorem is by far the trickiest we have seen. Often the theorem is proven only in special case $\sqrt{2}$ as a classical example in the first class of analysis. Here, we prove a general result. The following theorem is to make sure the definition of rational power make sense . In last section we prove some inequalities concerning integer power. Here, we prove those inequalities are inherited by rational power. Of course, the arithmetic properties of rational power, also inherited from those of integer power, will be proven after these inequalities. About the coverage of definition, notice that we didn't and won't define $y^{\frac{1}{n}}$ when y is negative. Also, notice that for all nonzero rational s, we can define $0^s = 0$.

Theorem 1.3.8. Given $m, p \in \mathbb{Z}$ and $n, q \in \mathbb{N}$ and $a > 0$ and $\frac{m}{n} = \frac{p}{q}$, we have

$$(a^m)^{\frac{1}{n}} = (a^p)^{\frac{1}{q}} \tag{1.36}$$

Proof. Observe

$$x = (a^m)^{\frac{1}{n}} \tag{1.37}$$
$$\iff x^n = a^m \tag{1.38}$$
$$\iff (x^n)^q = (a^m)^q \tag{1.39}$$
$$\iff x^{nq} = a^{mq} = a^{np} \text{ (because } n, m, q \in \mathbb{Z}) \tag{1.40}$$
$$\iff (x^q)^n = (a^p)^n \text{ (again, because } n, p, q \in \mathbb{Z}) \tag{1.41}$$
$$\iff x^q = a^p \text{ (by Theorem 1.3.6)} \tag{1.42}$$
$$\iff x = (a^p)^{\frac{1}{q}} \text{ (by Theroem 1.3.6)} \tag{1.43}$$
$$\iff (a^m)^{\frac{1}{n}} = x = (a^p)^{\frac{1}{q}} \tag{1.44}$$

■

Definition 1.3.9. (Definition of Rational Powers) Given a rational $r = \frac{p}{q}$, where $q \in \mathbb{N}$. For all a, we define $a^r = (a^p)^{\frac{1}{q}}$

Theorem 1.3.10. (Inequality when base is fixed) Given a positive a and two rational x, y where $x < y$, we have

$$\begin{cases} a^x < a^y \iff a > 1 \\ a^x = a^y \iff a = 1 \\ a^x > a^y \iff 0 < a < 1 \end{cases} \tag{1.45}$$

Proof. Express x, y in the form $x = \frac{q}{p}, y = \frac{n}{m}$ where $q, n \in \mathbb{Z}$ and $m, p \in \mathbb{N}$. Notice $x < y \implies \frac{q}{p} < \frac{n}{m} \implies mq < np$. Observe

$$a^x < a^y \implies a^{\frac{q}{p}} < a^{\frac{n}{m}} \tag{1.46}$$
$$\implies a^q = (a^{\frac{q}{p}})^p < (a^{\frac{n}{m}})^p \text{ (by Theorem 1.2.10)} \tag{1.47}$$
$$\implies a^{mq} = (a^q)^m < ((a^{\frac{n}{m}})^p)^m = ((a^{\frac{n}{m}})^m)^p = (a^n)^p = a^{np} \tag{1.48}$$
$$\implies a > 1 \text{ (by Theorem 1.2.9)} \tag{1.49}$$

Notice that the above implication still hold true if $<$ is replaced by $>$ and $>$ is replaced by $<$, and notice that the above implication still hold true if $<$ and $>$ are all replaced by $=$, so we in fact have three implications. The Theorem follows from the three implication.

■

Theorem 1.3.11. (Inequality when rational power is fixed) Given $0 < b < c$ and $z \in \mathbb{Q}$, we have

$$\begin{cases} b^z < c^z \iff 0 < z \\ b^z = c^z \iff 0 = z \\ b^z > c^z \iff 0 > z \end{cases} \tag{1.50}$$

Proof. Express $z = \frac{q}{p}$, where q is an integer and p is a natural. Notice q and z are of the same sign. Observe by Theorem 1.2.10, we have $b^{\frac{q}{p}} < c^{\frac{q}{p}} \implies b^q = (b^{\frac{q}{p}})^p < (c^{\frac{q}{p}})^p = c^q \implies 0 < q$.

Notice that the above implication still hold true if $<$ is replaced by $>$ and $>$ is replaced by $<$, and notice that the above implication still hold true if $<$ and $>$ are all replaced by $=$, so we in fact have three implications. The Theorem follows from the three implication. ■

Now we prove the arithmetic properties of rational power concerning only positive base. Notice there is no definition of a negative raised to the power of a rational. From Here

Theorem 1.3.12. (Arithmetic property of rational power) Given $r, s \in \mathbb{Q}$ and $a > 0$, we have

$$a^{r+s} = a^r a^s \tag{1.51}$$

Proof. Express r, s in the form $r = \frac{p}{q}, s = \frac{m}{n}$ where $q, n \in \mathbb{N}$ and $p, m \in \mathbb{Z}$. Observe

$$(a^{r+s})^{nq} = (a^{\frac{np+mq}{nq}})^{nq} \tag{1.52}$$
$$= a^{np+mq} \tag{1.53}$$

Then observe

$$(a^r a^s)^{nq} = (a^{\frac{p}{q}} a^{\frac{m}{n}})^{nq} \tag{1.54}$$
$$= (a^{\frac{p}{q}})^{nq} (a^{\frac{m}{n}})^{nq} \text{ (by Theorem 1.2.14)} \tag{1.55}$$
$$= ((a^{\frac{p}{q}})^q)^n ((a^{\frac{m}{n}})^n)^q \tag{1.56}$$
$$= (a^p)^n (a^m)^q \tag{1.57}$$
$$= a^{np+mq} = (a^{r+s})^{nq} \text{ (Theorem 1.2.12 and Theorem 1.2.13)} \tag{1.58}$$

Then by Theorem 1.3.6, we can deduce $a^r a^s = a^{r+s}$ ■

Theorem 1.3.13. (Arithmetic property of rational power) Given $r, s \in \mathbb{Q}$ and $a > 0$, we have

$$(a^r)^s = a^{rs} \tag{1.59}$$

Proof. Express r, s in the form $r = \frac{p}{q}, s = \frac{m}{n}$ where $q, n \in \mathbb{N}$ and $p, m \in \mathbb{Z}$. Observe

$$(a^{rs})^{nq} = (a^{\frac{mp}{nq}})^{nq} \tag{1.60}$$
$$= a^{mp} \tag{1.61}$$

Then observe

$$((a^r)^s)^{nq} = (((a^{\frac{p}{q}})^{\frac{m}{n}})^n)^q \tag{1.62}$$
$$= ((a^{\frac{p}{q}})^m)^q \tag{1.63}$$
$$= (a^{\frac{p}{q}})^{mq} \tag{1.64}$$
$$= ((a^{\frac{p}{q}})^q)^m \tag{1.65}$$
$$= (a^p)^m = a^{mp} = (a^{rs})^{nq} \tag{1.66}$$

■

Corollary 1.3.14. $a^{\frac{q}{p}} = (a^{\frac{1}{p}})^q$

Theorem 1.3.15. (Arithmetic property of rational power) Given $r \in \mathbb{Q}$ and $a, b > 0$, we have

$$a^r b^r = (ab)^r \tag{1.67}$$

Proof. Express r in the form $r = \frac{q}{p}$, and observe $(a^r b^r)^p = (a^r)^p (b^r)^p = (a^{\frac{q}{p}})^p (b^{\frac{q}{p}})^p = a^q b^q = (ab)^q = ((ab)^r)^p$. ■

1.4 Irrational Power

After the rational power, we now try to define irrational power. In most of the textbooks, irrational power as a rigorous definition often come after the definition of Euler number, but here, we define the irrational power with a technique not so advanced and to be fair quite cumbersome compared to the approaches in most textbooks. Of course, the common inequalities and arithmetic properties of irrational power will be presented, although their order of presentation is different to how we present in the section before.

Lemma 1.4.1. For $b, x \in \mathbb{F}$, define $B(x) := \{b^t : t \in \mathbb{Q}, t \leq x\}$. Then for all $x \in \mathbb{F}$, $b > 1$ implies $\sup B(x)$ exists, and $0 < b \leq 1$ implies $\inf B(x)$ exists.

Proof. For all x, we know $B(x)$ is nonempty, since if not, $\mathbb{Q}$ is bounded below. Because $\mathbb{Q}$ is nonbounded above, we can pick a rational y greater than x. If $b > 1$, we deduce $\forall b^t \in B(x), b^t \leq b^y$; thus b^y is an upper bound of $B(x)$. If $0 < b < 1$, we deduce $\forall b^t \in B(x), b^y \leq b^t$; thus b^y is a lower bound of $B(x)$. If $b = 1$, just notice $B(x) = \{1\}$. ■

Definition 1.4.2. (Irrational Power Definition) For all $b, x \in \mathbb{F}$, define $B(x)$ as in Lemma 1.4.1, and for all $x \in \mathbb{F}$, if $b > 1$, define $b^x := \sup B(x)$, and if $0 < b \leq 1$, define $b^x := \inf B(x)$

The above definition is in fact not very appropriate, since we have already define b^x where $x \in \mathbb{Q}$. We don't know if the above definition is consistent with our old definition. The next theorem is to show it is.

Theorem 1.4.3. (Consistency of power definitions) Given $r \in \mathbb{Q}$, if $b > 1$, then $b^r = \sup B(r)$, and if $b < 1$, then $b^r = \inf B(r)$

Proof. If $b > 1$, deduce $b^r = \max B(r) = \sup B(r)$. If $b < 1$, deduce $b^r = \min B(r) = \inf B(r)$. ■

Now, we are going to prove one common inequalities. The other will be proven after some work.

Theorem 1.4.4. (Inequality when base is fixed) Given a positive b and $x < y$, we have

$$\begin{cases} b^x < b^y \iff b > 1 \\ b^x = b^y \iff b = 1 \\ b^x > b^y \iff 0 < b < 1 \end{cases} \tag{1.68}$$

Proof. By Theorem 1.3.5, we can pick a rational q such that $x < q < y$. We have $b^q \in B(y)$. We now consider three possible situations.

If $b > 1$, then $b^q \leq \sup B(y)$. Because $\forall b^t \in B(x), t \leq x < q$, we can deduce $\forall b^t \in B(x), b^t < b^q$; that is, b^q is an upper bound of $B(x)$. Notice that there exists rational between $x < q$, so b^q can not be the least upper bound of $B(x)$. Then we can deduce $\sup B(x) < b^q \leq \sup B(y)$.

If $b = 1$, then $b^x = 1 = b^y$.

If $b < 1$, then $\inf B(y) \leq b^q$. Because $\forall b^t \in B(x), t \leq x < q$, we can deduce $\forall b^t \in B(x), b^t > b^q$; that is, b^q is an lower bound of $B(x)$. Notice that there exists rational between $x < q$, so b^q can not be the greatest lower bound of $B(x)$. Then we can deduce $\inf B(y) \leq b^q < \inf B(x)$.

The Theorem follows from the three implications of the three situation. ■

Lemma 1.4.5. Let S, U and be two bounded above subset of $\mathbb{F}$, containing only nonnegative numbers. Define $SU := \{su : s \in S, u \in U\}$. We have

$$\sup SU = \sup S \sup U \tag{1.69}$$

and have

$$\inf SU = \inf S \inf U \tag{1.70}$$

where the infimum part hold true even if S, U are not bounded above

Proof. Because S, U contain only nonnegative numbers, we know $s \leq \sup S$ and $u \leq \sup U$ implies $su \leq \sup S \sup U$, so $\sup S \sup U$ is an upper bound of SU. Now we show $\sup S \sup U$ is the *least* upper bound of SU.

Assume there is an upper bound x of AB smaller than $\sup A \sup B$. Express x in the form $x = \sup S \frac{x}{\sup S}$. Because $x < \sup S \sup U$, we know $\frac{x}{\sup S} < \sup U$, which implies there is a number $u \in U$ greater than $\frac{x}{\sup S}$. Observe $u > \frac{x}{\sup S} \implies \frac{x}{u} < \sup S$, which implies that there is a number $s \in S$ greater than $\frac{x}{u}$. Observe $\frac{x}{u} < s \implies x < su$ CaC (done)

Clearly, we know $s \geq \inf S$ and $u \geq \inf U$ implies $su \geq \inf S \inf U$, so $\inf S \inf U$ is an lower bound of SU. Now we show $\sup S \sup U$ is the *greatest* lower bound of SU.

Assume there is an lower bound x of AB greater than $\sup A \sup B$. Express x in the form $x = \inf S \frac{x}{\inf S}$. Because $x > \inf S \inf U$, we know $\frac{x}{\inf S} > \inf U$, which implies there is a number $u \in U$ less than $\frac{x}{\inf S}$. Observe $u < \frac{x}{\inf S} \implies \frac{x}{u} > \inf S$, which implies that there is a number $s \in S$ less than $\frac{x}{u}$. Observe $\frac{x}{u} > s \implies x > su$ CaC (done) ■

Theorem 1.4.6. (Arithmetic property) Given $r, s \in \mathbb{F}$ and $0 < a$, we have

$$a^{r+s} = a^r a^s \tag{1.71}$$

Proof. We prove $A(r+s) = \{xy : x \in A(r) \text{ and } y \in A(s)\}$. Denote the set on right side E. Observe that $xy \in E \implies \exists q \leq r \in \mathbb{Q}, x = a^q$ and $\exists m \leq s \in \mathbb{Q}, y = a^m \implies xy = a^{q+m}$ where $q + m \leq r + s \implies xy \in A(r+s)$. Then we know $E \subseteq A(r+s)$. Given $a^d \in A(r+s)$, we know $d \leq r + s$, so if we express $a^d = a^r a^{d-r}$, we are sure $a^{d-r} \in A(s)$ and $a^r \in A(r)$, which implies $a^d \in E$. Then we can deduce $A(r+s) \subseteq E$. (done)

By Lemma 1.4.5, our proof is finished. ■

Now, we introduce an idea to aid our proof of inequality.

Definition 1.4.7. (Definition of order homomorphism) Let S be an subset of $\mathbb{F}$, the function $\phi : S \to \mathbb{F}$ is an *order homomorphism*, if for all $a, b \in S$, we have

$$a < b \implies \phi(a) < \phi(b) \tag{1.72}$$

Two subset U, V are said to be *order isomorphic*, if there exists an bijective order homomorphism from one to another.

Lemma 1.4.8. Let S, U be two order isomorphic bounded above subset of $\mathbb{F}$, containing only nonnegative numbers, where $\phi : S \to U$ is an order isomorphism. Let $US_\phi = \{s\phi(s) : s \in S\}$. We have

$$\sup S \sup U = \sup US_\phi \tag{1.73}$$

and have

$$\inf S \inf U = \inf US_\phi \tag{1.74}$$

where the infimum part hold true even if S, U are not bounded above.

Proof. For all $s\phi(s)$ in US_ϕ, we know $0 \le s \le \sup S$ and $0 \le \phi(s) \le \sup U$, so we can deduce $s\phi(s) \le \sup S \sup U$. We have proven that $\sup S \sup U$ is an upper bound of US_ϕ. We now prove it is the least. Assume there exists an upper bound x of US_ϕ smaller than $\sup S \sup U$. Because $\frac{x}{\sup U} < \sup S$, we know there exists $s \in S$ such that $x < (\sup U)s$. Then because $\frac{x}{s} < \sup U$, we know there exists $\phi(s') \in U$ such that $x < s\phi(s')$. If $s < s'$, then we see $x < s\phi(s') < s'\phi(s') \in US_\phi$. If $s > s'$, then we see $x < s\phi(s') < s\phi(s)$. If $s = s'$, then we see $x < s\phi(s') = s\phi(s)$ CaC

For all $s\phi(s)$ in US_ϕ, we know $s \ge \inf S$ and $\phi(s) \ge \inf U$, so we can deduce $s\phi(s) \ge \inf S \inf U$. We have proven that $\inf S \inf U$ is an lower bound of US_ϕ. We now prove it is the greatest. Assume there exists an lower bound x of US_ϕ greater than $\inf S \inf U$. Because $\frac{x}{\inf U} > \inf S$, we know there exists $s \in S$ such that $x > (\inf U)s$. Then because $\frac{x}{s} > \inf U$, we know there exists $\phi(s') \in U$ such that $x > s\phi(s')$. If $s < s'$, then we see $x > s\phi(s') > s\phi(s) \in US_\phi$. If $s > s'$, then we see $x > s\phi(s') > s'\phi(s')$. If $s = s'$, then we see $x > s\phi(s') = s\phi(s)$ CaC ■

Theorem 1.4.9. (Arithmetic property) Given $b, c > 0$ and $z \in \mathbb{F}$, we have

$$b^z c^z = (bc)^z \tag{1.75}$$

Proof. From now, we will use $C(z)$ to denote $\{c^u : u \in \mathbb{Q}, u \le z\}$. Define $\phi : B(z) \to C(z)$ as $b^t \mapsto c^t$. Observe that if b, c are both greater than 1, then $b^t < b^m \implies t < m \implies c^t < c^m$, and that if b, c are both less than 1, then $b^t < b^m \implies t > m \implies c^t < c^m$. So If b, c are both greater than 1 or both less than 1, then ϕ is an order isomorphism. Notice that $\{(bc)^u : u \in \mathbb{Q}, u \le z\} = \{b^t\phi(b^t) : t \in \mathbb{Q}, t \le z\} = C(z)B(z)_\phi$. Then by Lemma 1.4.8, our proof is finished. We now show $(b^{-1})^z b^z = 1$.

WOLG, let $b < 1$. Denote $H := \{b^t : t \in \mathbb{Q}, t \le z\}, E := \{b^{-t} : t \in \mathbb{Q}, t \le z\}$, so $b^z = \inf H$ and $(b^{-1})^z = \sup E$. Assume $(\sup E)(\inf H) > 1$. Notice $\sup E > \frac{1}{\inf H}$ implies that there exists $t \le z$ such that $b^{-t} > \frac{1}{\inf H}$, which leads to $\inf H > b^t$ CaC . Assume $(\sup E)(\inf H) < 1$. Notice $\inf H < \frac{1}{\sup E}$ implies that there exists $t \le z$ such that $b^{-t} < \frac{1}{\sup E}$, which leads to $\sup E < b^t$ CaC (done)

If $b < 1 < c$ and $bc > 1$, observe that $(bc)^z(b^{-1})^z = c^z \implies (bc)^z = (bc)^z(b^{-1})^z b^z = b^z c^z$. If $b < 1 < c$ and $bc < 1$, observe that $(bc)^z(c^{-1})^z = b^z \implies (bc)^z = (bc)^z(c^{-1})^z c^z = b^z c^z$ ■

Theorem 1.4.10. (Inequality when power is fixed) Given $0 < b < c$ and $z \in \mathbb{F}$, we have

$$\begin{cases} b^z < c^z \iff 0 < z \\ b^z = c^z \iff 0 = z \\ b^z > c^z \iff 0 > z \end{cases} \tag{1.76}$$

Proof. By Theorem 1.4.4, $z > 0 \implies (\frac{c}{b})^z > (\frac{c}{b})^0 = 1 \implies c^z > b^z$, and $z = 0 \implies b^z = 1 = c^z$, and $z < 0 \implies (\frac{c}{b})^z < (\frac{c}{b})^0 = 1$ ■

Before the last arithmetic property, we first prove the existence of logarithm, which play an important role in the proof of last arithmetic property.

Theorem 1.4.11. (Existence of logarithm) For all positive b, y where $b \neq 1$, there exists an unique $x \in \mathbb{F}$ such that $b^x = y$

Proof. The uniqueness part have been handled by Theorem 1.4.4. Now we split the proof into two cases: $b > 1$ and $b < 1$

Case 1: $b > 1$

Define $A := \{w : b^w < y\}$ and $x = \sup A$.

We first prove the identity for all positive integer m and $c > 1$, we have $c - 1 \geq (c^{\frac{1}{m}} - 1)m$

Let $d = c^{\frac{1}{m}}$, so we have $c - 1 = d^m - 1 = (d-1)(d^{m-1} + d^{m-1} + \cdots + 1) \geq (d-1)m = (c^{\frac{1}{m}} - 1)m$ (done)

Assume $b^x > y$. Let $t = \frac{b^x}{y}$ and $\frac{b-1}{t-1} < n = m$ and $b = c$ to use the identity, so we have $b - 1 \geq n(b^{\frac{1}{n}} - 1)$. Then we have $\frac{b-1}{n} \geq b^{\frac{1}{n}} - 1$. Notice that $b^x > y \implies t - 1 > 0$, so because $n > \frac{b-1}{t-1}$, we then have $t - 1 > \frac{b-1}{n} \geq b^{\frac{1}{n}} - 1$, which implies $t > b^{\frac{1}{n}}$.

Observe $b^{x-\frac{1}{n}} = \frac{b^x}{b^{\frac{1}{n}}} > \frac{b^x}{t} = y$. Then $x - \frac{1}{n}$ is an upper bound of A CaC .

Assume $b^x < y$. Let $t = \frac{y}{b^x}$ and let $\frac{b-1}{t-1} < n = m$ and $b = c$ to use the identity, so we have $b - 1 \geq n(b^{\frac{1}{n}} - 1)$. Then we have $\frac{b-1}{n} \geq b^{\frac{1}{n}} - 1$. Notice that $b^x < y \implies t - 1 > 0$, so because $n > \frac{b-1}{t-1}$, we then have $t - 1 > \frac{b-1}{n} \geq b^{\frac{1}{n}} - 1$, which implies $t > b^{\frac{1}{n}}$.

Observe $b^{x+\frac{1}{n}} = b^x b^{\frac{1}{n}} < b^x t = y$. Then $x + \frac{1}{n} \in A$ CaC (done)

Case 2: $b < 1$

Define $B := \{w : b^w < y\}$ and $x := \inf B$.

We first prove the identity for all positive integer m and $0 < c < 1$, we have $c-1 \leq (c^{\frac{1}{m}}-1)m$

Let $d = c^{\frac{1}{m}}$, so we have $c - 1 = d^m - 1 = (d-1)(d^{m-1} + d^{m-2} + \cdots + 1) \leq (d-1)m = (c^{\frac{1}{m}} - 1)m$ (done) .

Assume $b^x > y$. Let $t = \frac{y}{b^x}$ and $\frac{b-1}{t-1} < n = m$ and $b = c$ to use the identity, so we have $b - 1 \leq (b^{\frac{1}{n}} - 1)n$. Then we have $\frac{b-1}{n} \leq b^{\frac{1}{n}} - 1$. Notice that $b^x > y \implies t - 1 < 0$, so because $n > \frac{b-1}{t-1}$, we then have $t - 1 < \frac{b-1}{n} \leq b^{\frac{1}{n}} - 1$, which implies $t < b^{\frac{1}{n}}$.

Observe $b^{x+\frac{1}{n}} = b^x b^{\frac{1}{n}} > b^x t = y$. Then $x + \frac{1}{n}$ is an lower bound of B CaC .

Assume $b^x < y$. Let $t = \frac{b^x}{y}$ and $\frac{b-1}{t-1} < n = m$ and $b = c$ to use the identity, so we have $b - 1 \leq (b^{\frac{1}{n}} - 1)n$. Then we have $\frac{b-1}{n} \leq b^{\frac{1}{n}} - 1$. Notice that $b^x < y \implies t - 1 < 0$, so because $n > \frac{b-1}{t-1}$, we then have $t - 1 < \frac{b-1}{n} \leq b^{\frac{1}{n}} - 1$, which implies $t < b^{\frac{1}{n}}$.

Observe $b^{x-\frac{1}{n}} = \frac{b^x}{b^{\frac{1}{n}}} < \frac{b^x}{t} = y$. Then $x - \frac{1}{n} \in B$ CaC (done) ■

Definition 1.4.12. (Definition of logarithm) Define $\log_b y := x$, where x, b, y are in Theorem 1.4.11.

Lemma 1.4.13. Given $0 < x < rs$ and $0 < r$ and $0 < s$, there exist positive rational $u < s$ and positive rational $t < r$ such that $x < tu < rs$.

Proof. $0 < x < rs \implies \frac{x}{s} < r$. Then there exists rational t such that $\frac{x}{s} < t < r$, which implies $\frac{x}{t} < s$. Then there exists rational u such that $\frac{x}{t} < u < s$. Then $x < tu < rs$. ■

Theorem 1.4.14. (Arithmetic property) Given $r, s \in \mathbb{F}$ and $a > 0$, we have

$$(a^r)^s = a^{rs} \tag{1.77}$$

Proof. The proof is very very long. We first denote three important sets:

$$T := \{a^t : t \in \mathbb{Q}, t \leq r\} \tag{1.78}$$

$$U := \{(a^r)^u : u \in \mathbb{Q}, u \leq s\} \tag{1.79}$$

$$X := \{a^x : x \in \mathbb{Q}, x \leq rs\} \tag{1.80}$$

So, we have

$$a^r = (\sup \text{ or } \inf)T \tag{1.81}$$

$$(a^r)^s = (\sup \text{ or } \inf)U \tag{1.82}$$

$$a^{rs} = (\sup \text{ or } \inf)X \tag{1.83}$$

Notice that from now, when variables t, u, x and their variants t', u', x' are mentioned, they are rational and respectively less than or equal to r, s, rs.

Now, we split the proof into sixteen cases: a may be smaller or greater than 1; r, rs may be positive or negative; rs is rational or irrational.

We first do all 8 cases of $rs \notin \mathbb{Q}$. Notice that $rs \notin \mathbb{Q} \implies x < rs$.

The easiest case first: $a > 1$ and $r > 0$ and $s > 0$ and $rs \notin \mathbb{Q}$.

Assume $a^{rs} > (a^r)^s$; that is, $\sup X > \sup U$. Then $\exists x, \forall u, a^x > (a^r)^u = (\sup T)^u$. Let $u > 0$, so we have $\forall t, a^{\frac{x}{u}} > a^t$. By Lemma 1.4.13 CaC . Assume $a^{rs} < (a^r)^s$; that is, $\sup X < \sup U$. Then $\exists u, \forall x, (\sup T)^u = (a^r)^u > a^x$. Because $a^r > 1$, we know $u > 0$, since $u < 0 \implies (a^r)^u < 1$. Then we have $\exists u, \forall x, \sup T > a^{\frac{x}{u}}$, so we have $\exists u, \forall x, \exists t, a^t > a^{\frac{x}{u}}$. But we see that for all u, we can let $x > ru$, so $a^{\frac{x}{u}} > a^r$. CaC (done)

The second case: $a > 1$ and $r > 0$ and $s < 0$ and $rs \notin \mathbb{Q}$.

Assume $a^{rs} > (a^r)^s$; that is, $\sup X > \sup U$. Then $\exists x, \forall u, a^x > (a^r)^u = (\sup T)^u$. Because $u \leq s < 0$, we have $\exists x, \forall u, a^{\frac{x}{u}} < \sup T$. Then, we have $\exists x, \forall u, \exists t, a^{\frac{x}{u}} < a^t$. Express $x = rm$, and let $0 > s > u > m$, so we have $\frac{x}{u} = \frac{rm}{u} > r$ CaC . Assume $a^{rs} < (a^r)^s$; that is, $\sup X < \sup U$. Then $\exists u, \forall x, a^x < (a^r)^u = (\sup T)^u$. Because $u \leq s < 0$, we have $\exists u, \forall x, a^{\frac{x}{u}} > \sup T$. Then, we have $\exists u, \forall x, \forall t, a^{\frac{x}{u}} > a^t$. If $s \notin \mathbb{Q}$, then we notice that because $u < s < 0$, we have $r(\frac{s}{u}) < r$. Pick t such that $r(\frac{s}{u}) < t < r$, so we have $rs > tu$. Then pick x such that $0 > rs > x > tu$, and observe $a^{\frac{x}{u}} < a^t$ CaC . If $s \in \mathbb{Q}$, then it is possible to happen $u = s$. Then $a^{\frac{x}{u}} = a^{\frac{x}{s}} > a^{\frac{rs}{s}} = a^r \geq a^t$ CaC (done)

The third case: $a > 1$ and $r < 0$ and $s > 0$ and $rs \notin \mathbb{Q}$

Assume $a^{rs} > (a^r)^s$; that is, $\sup X > \inf U$. Then $\exists x, \exists u, a^x > (a^r)^u = (\sup T)^u$. Notice that $u \leq 0 \implies (a^r)^u \geq 1 > a^x$, so u must be positive. Then we have $\exists x, \exists u, a^{\frac{x}{u}} > \sup T$. Then, we have $\exists x, \exists u, \forall t, a^{\frac{x}{u}} > a^t$. Notice that $\frac{x}{u} < \frac{rs}{u} < r$, so there exists t such that $a^{\frac{x}{u}} < a^t$ CaC . Assume $a^{rs} < (a^r)^s$; that is, $\sup X < \inf U$. Let $\sup X < m < \inf U$, and let $n = \log_a m$, so $a^{rs} < a^n < \inf U$. Pick a rational k so that $a^{rs} < a^k < a^n < \inf U$. Then we know for all positive u, we have $a^{\frac{k}{u}} < a^r = \sup T$. This give us $\forall u > 0, \exists t, a^k < a^{tu}$. Because $rs < k$, we know $s > \frac{k}{r}$. Let $s > u' > \frac{k}{r}$. Then we see $\forall t, u't < u'r < k$ CaC (done)

The forth case: $a > 1$ and $r < 0$ and $s < 0$ and $rs \notin \mathbb{Q}$

Assume $a^{rs} > (a^r)^s$; that is, $\sup X > \inf U$. Then $\exists x, \exists u, a^x > (a^r)^u$. Because $u \leq s < 0$, we have $\exists x, \exists u, a^{\frac{x}{u}} < a^r = \sup T$. Then $\exists x, \exists u, \exists t, a^x > a^{tu}$. Observe $tu \geq ru \geq rs > x$ CaC . Assume $a^{rs} < (a^r)^s$; that is, $\sup X < \inf U$. Let $\sup X < m < \inf U$, and let $n = \log_a m$, so $a^{rs} < a^n < \inf U$. Pick a rational k so that $a^{rs} < a^k < a^n < \inf U$. Then, $\forall u, a^{\frac{k}{u}} > a^r = \sup T$. This implies $\forall u, \forall t, a^k < a^{tu}$. Because $rs < k$, we know $s > \frac{k}{r}$. Let $s > u' > \frac{k}{r}$. Then $u'r < k$, which implies $r > \frac{k}{u'}$. Let $r > t' > \frac{k}{u'}$. Then we see $t'u' < k$ CaC (done)

The next four cases where $a < 1$ use the same logic, and reference to Theorem 1.4.9.

The forth to eighth cases: $a < 1$ and $rs \notin \mathbb{Q}$

By the above four cases, we know $((\frac{1}{a})^r)^s = (\frac{1}{a})^{rs}$. Observe $a^{rs}(\frac{1}{a})^{rs} = 1^{rs} = 1$, and observe $(a^r)^s((\frac{1}{a})^r)^s = (a^r(\frac{1}{a})^r)^s = ((1)^r)^s = 1$ (done) .

The final eight cases: $rs \in \mathbb{Q}$

If r, s are both rational, the proof reference to Theorem 1.3.13. Notice that if only one of r, s is rational, then rs is irrational, so we can assume r, s are both irrational. Observe by the eight cases above, $(a^{rs})^{\frac{1}{s}} = a^r$, so by Definition 1.3.9, $(a^r)^s = ((a^{rs})^{\frac{1}{s}})^s = a^{rs}$ (done)

Notice that outside of the sixteen cases above, if $a = 1$ or $s = 0$ or $s = 0$, the proof is trivial.

■

1.5 Euclidean Space

In this section, we will use the notion of vector space to define Euclidean space, so we will first prove a simple and fundamental theorem about vector space.

Theorem 1.5.1. Let V be a vector space over arbitrary filed $\mathbb{F}$ and of dimension n. We have that V is isomorphic to $\mathbb{F}^n$

Proof. For people who don't know, $\mathbb{F}^n$ in most textbooks is defined as a vector space consisting of n-tuples, and n-tuples can be defined as follows for any natural n.

$$(c,d) := \{\{c\},\{c,d\}\} \tag{1.84}$$
$$(b,c,d) := (b,(c,d)) = \{\{b\},\{b,(c,d)\}\} = \{\{b\},\{b,\{\{c\},\{c,d\}\}\}\} \tag{1.85}$$
$$(a,b,c,d) := (a,(b,c,d)) = \cdots \tag{1.86}$$

Let $\{v_1,\ldots,v_n\}$ be a basis for V, and define $\phi : \{v_1,\ldots,v_n\} \to \mathbb{F}^n$ by $c_1v_1 + \cdots + c_nv_n \mapsto (c_1,\ldots,c_n)$. Notice that ϕ can verified to be linear and bijective by using the theorem that every vector in V can be uniquely expressed as a linear combination of $v_1,\ldots,v_n$. ■

Notice that the existence and uniqueness of real numbers field will be handled in Section 1.7,1.8,1.9, and from now on, let's just pretend that there exists an unique completed ordered field, which we call Real Numbers Field and denote $\mathbb{R}$.

In this section, we will introduce the definition of Euclidean space and Cauchy-Schwarz inequality in Euclidean space.

Definition 1.5.2. (Definition of absolute value) Let $x \in \mathbb{R}$. We define the absolute value $|x|$ of x by

$$|x| := \begin{cases} x & \text{if } x > 0 \\ -x & \text{if } x \leq 0 \end{cases} \tag{1.87}$$

Theorem 1.5.3. $\forall x \in \mathbb{R}, |x| = \sqrt{x^2}$

Proof. If $x > 0$, then $|x| > 0$ and $|x|^2 = x^2$. If $x \leq 0$, then $|x| \geq 0$ and $|x|^2 = (-x)^2 = x^2$. ■

Definition 1.5.4. (Definition of L_p norm, or, p-norm) Let $p \geq 1 \in \mathbb{R}$, and let $\mathbf{x} \in \mathbb{R}^n$

$$\|\mathbf{x}\|_p := (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}} \tag{1.88}$$

Definition 1.5.5. (Definition of Euclidean n-space) When we use the word *Euclidean n-Space*, we mean $\mathbb{R}^n$ as a vector space equipped with a function $\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ called Euclidean inner product defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^{n} x_i y_i \tag{1.89}$$

and equipped with 2-norm

$$\|\mathbf{x}\| := \|\mathbf{x}\|_2 = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}} \tag{1.90}$$

2-norm is also called *Euclidean norm* or L_2 norm. Most time we use the notation $\mathbf{x} \cdot \mathbf{y}$ in place of $\langle \mathbf{x}, \mathbf{y} \rangle$ and the notation $|\mathbf{x}|$ in place of $\|\mathbf{x}\|$, for abbreviation. Notice that by Theorem 1.5.3, when $\mathbf{x} = (x_1) \in \mathbb{R}^1$, we have $|\mathbf{x}| = |x_1|$, so we don't have to worry about the compatibility of abbreviation. Lastly, we can use Theorem 1.5.3, to verify $|\mathbf{x}| = (\mathbf{x} \cdot \mathbf{x})^{\frac{1}{2}}$ in Euclidean Space.

Although the common usage of notation $\mathbb{R}^n$ only refer to the set without any structure, but from now, we will use $\mathbb{R}^n$ to denote Euclidean n-space.

Notice that Euclidean 3-space is a great tool to describe the physical world, and the idea of Euclidean norm and Euclidean inner product captures the essence of length and angle really well as we will explain in later section and chapter.

Lastly, we close this section with a proof of Cauchy-Schwarz inequality in Euclidean space, and leave some important properties of Euclidean norm and Euclidean inner product to next section.

Theorem 1.5.6. (Special Case of Schwarz inequality) Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. We have

$$|\mathbf{x}||\mathbf{y}| \geq |\mathbf{x} \cdot \mathbf{y}| \tag{1.91}$$

The equality hold only if $\exists \lambda \in \mathbb{R}, \mathbf{x} = \lambda \mathbf{y}$

Proof. Let $f(t) = \sum_{i=1}^{n} (x_i t - y_i)^2$, so we have

$$\begin{aligned} f(t) &= \sum (x_i t - y_i)^2 && (1.92)\\ &= \sum x_i^2 t^2 - 2x_i y_i t + y_i^2 && (1.93)\\ &= t^2 \sum x_i^2 - 2t \sum x_i y_i + \sum y_i^2 && (1.94)\\ &= |\mathbf{x}|^2 t^2 - 2(\mathbf{x} \cdot \mathbf{y})t + |\mathbf{y}|^2 && (1.95) \end{aligned}$$

$f(t)$ is an nonnegative quadratic polynomial of t, since $f(t)$ is a sum of squares. Then $D(f) \leq 0$; that is $4(\mathbf{x} \cdot \mathbf{y})^2 - 4|\mathbf{x}|^2|\mathbf{y}|^2 \leq 0$, which implies $(\mathbf{x} \cdot \mathbf{y})^2 \leq |\mathbf{x}|^2|\mathbf{y}|^2$, and further implies $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$

If $|\mathbf{x}||\mathbf{y}| = |\mathbf{x} \cdot \mathbf{y}|$, Then $D(f) = 0$, that is, there exists a unique $t' \in \mathbb{R}$ such that $f(t') = 0$. Then we see $0 = f(t') = \sum (x_i t' - y_i)^2$, which implies $\forall i, y_i = x_i t'$ ■

1.6 Complex Numbers

Theorem 1.6.1. If we define vector multiplication in $\mathbb{R}^2$ as

$$\mathbf{xy} = (x_1 y_1 - x_2 y_2, x_1 y_2 + x_2 y_1) \tag{1.96}$$

, then $\mathbb{R}^2$ become a field.

Proof. The whole proof is long and tedious, here we only present some that are worth mentioning.

$$\mathbf{yx} = (y_1x_1 - y_2x_2, y_1x_2 + y_2x_1) = (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1) = \mathbf{xy} \tag{1.97}$$

$$\begin{aligned}
(\mathbf{xy})\mathbf{z} &= (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1)(z_1, z_2) && (1.98)\\
&= (x_1y_1z_1 - x_2y_2z_1 - x_1y_2z_2 - x_2y_1z_2, x_1y_1z_2 - x_2y_2z_2 + x_1y_2z_1 + x_2y_1z_1) && (1.99)\\
&= (x_1(y_1z_1 - y_2z_2) - x_2(y_2z_1 + y_1z_2), x_1(y_1z_2 + y_2z_2) + x_2(y_1z_1 - y_2z_2)) && (1.100)\\
&= (x_1, x_2)(y_1z_1 - y_2z_2, y_1z_2 + y_2z_1) = \mathbf{x}(\mathbf{yz}) && (1.101)
\end{aligned}$$

$$(1, 0)\mathbf{x} = (1, 0)(x_1, x_2) = (x_1, x_2) \text{ and } \mathbf{x}(1, 0) = (x_1, x_2)(1, 0) = (x_1, x_2) \tag{1.102}$$

$$\begin{aligned}
\mathbf{x}(\frac{x_1}{x_1^2 + x_2^2}, \frac{-x_2}{x_1^2 + x_2^2}) &= (\frac{x_1^2 + x_2^2}{x_1^2 + x_2^2}, \frac{-x_1x_2 + x_1x_2}{x_1^2 + x_2^2}) && (1.103)\\
&= (1, 0) && (1.104)
\end{aligned}$$

$$\begin{aligned}
\mathbf{x}(\mathbf{y} + \mathbf{z}) &= (x_1, x_2)(y_1 + z_1, y_2 + z_2) && (1.105)\\
&= (x_1y_1 + x_1z_1 - x_2y_2 - x_2z_2, x_2y_1 + x_2z_1 + x_1y_2 + x_1z_2) && (1.106)\\
&= (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1) + (x_1z_1 - x_2z_2, x_2z_1 + x_1z_2) && (1.107)\\
&= \mathbf{xy} + \mathbf{xz} && (1.108)
\end{aligned}$$

■

The fact that the field in Theorem 1.6.1 is isomorphic to the customary $\mathbb{C}$ can be easily verified by checking $(a, b) \mapsto a + bi$ is field isomorphism. We now define $\mathbb{C}$ in this way. Notice that this definition is an abuse of notation, since in this way, $\mathbb{R} \notin \mathbb{C}$. The motivation behind this definition is to emphasize the geometric nature of $\mathbb{C}$.

Definition 1.6.2. (Definition of Complex Numbers) We define $\mathbb{C}$ as the field in Theorem 1.6.1. We write (a, b) as $a + bi$ and $(a, -b)$ as $a - bi$. We define $\mathrm{Re}(a + bi) := a$ and $\mathrm{Im}(a + bi) := b$. We say the real part of $a + bi$ is a and the imaginary part of $a + bi$ is b. We define the absolute of value of complex number as its length when it is treated as a vector in complex plane, which in our definition is a Euclidean space. Precisely, we define $|a + bi| := (a^2 + b^2)^{\frac{1}{2}} = |(a, b)|$. Moreover, we define the conjugate as $\overline{a + bi} = a - bi$. We abuse the notation so that $a + 0i \in \mathbb{R} \subseteq \mathbb{C}$.

Theorem 1.6.3. Let $z, w \in \mathbb{C}$. We have

$$\overline{(\overline{z})} = z \tag{1.109}$$

$$\overline{z+w} = \overline{z} + \overline{w} \tag{1.110}$$

$$\overline{zw} = (\overline{z})(\overline{w}) \tag{1.111}$$

$$\overline{z^n} = (\overline{z})^n \tag{1.112}$$

$$z + \overline{z} = 2\mathrm{Re}(z) \text{ and } z - \overline{z} = 2i\mathrm{Im}(z) \tag{1.113}$$

$$|\overline{z}| = |z| \tag{1.114}$$

$$|zw| = |z||w| \tag{1.115}$$

$$|z^n| = |z|^n \tag{1.116}$$

Proof. We prove only the following. Let $z = z_1 + z_2 i$ and $w = w_1 + w_2 i$. We have

$$\overline{(\overline{z})(\overline{w})} = \overline{(z_1 - z_2 i)(w_1 - w_2 i)} \tag{1.117}$$

$$= \overline{z_1 w_1 - z_2 w_2 - (z_2 w_1 + z_1 w_2)i} \tag{1.118}$$

$$= z_1 w_1 - z_2 w_2 + (z_2 w_1 + z_1 w_2)i \tag{1.119}$$

$$= zw \tag{1.120}$$

$$|\overline{z}| = (z_1^2 + (-z_2)^2)^{\frac{1}{2}} = (z_1^2 + z_2^2)^{\frac{1}{2}} = |z| \tag{1.121}$$

$$|zw| = |z_1 w_1 - z_2 w_2 + (z_2 w_1 + z_1 w_2)i| \tag{1.122}$$

$$= [(z_1 w_1 - z_2 w_2)^2 + (z_2 w_1 + z_1 w_2)^2]^{\frac{1}{2}} \tag{1.123}$$

$$= [(z_1 w_1)^2 + (z_2 w_2)^2 - 2 z_1 z_2 w_1 w_2 + (z_2 w_1)^2 + (z_1 w_2)^2 + 2 z_1 z_2 w_1 w_2]^{\frac{1}{2}} \tag{1.124}$$

$$= [z_1^2 w_1^2 + z_2^2 w_1^2 + z_1^2 w_2^2 + z_2^2 w_2^2]^{\frac{1}{2}} \tag{1.125}$$

$$= [(z_1^2 + z_2^2)(w_1^2 + w_2^2)]^{\frac{1}{2}} = (z_1^2 + z_2^2)^{\frac{1}{2}} (w_1^2 + w_2^2)^{\frac{1}{2}} = |z||w| \tag{1.126}$$

■

In last section we define the Euclidean norm. Here, we give the axioms for norm function and verify that Euclidean norm satisfy all of them. Notice that the four axioms of norm function each describe a property of length in Euclidean space, which is so "obviously true" if we use a non-rigorous approach in geometry.

Axiom 1.6.4. (Axioms for norm function) Let V be a vector space over $\mathbb{F} \subseteq \mathbb{C}$. A norm, or norm function, $p : V \to \mathbb{R}$ is a function that satisfy

$$\forall \mathbf{x},\mathbf{y} \in V, p(\mathbf{x}+\mathbf{y}) \leq p(\mathbf{x}) + p(\mathbf{y}) (\text{ Triangle inequality}) \tag{1.127}$$

$$\forall \mathbf{x} \in V, \forall \lambda \in \mathbb{F}, p(\lambda \mathbf{x}) = |\lambda| p(\mathbf{x}) \text{ (Absolute homogenity)} \tag{1.128}$$

$$\forall \mathbf{x} \in V, p(\mathbf{x}) \geq 0 \text{(Nonnegativity)} \tag{1.129}$$

$$\forall \mathbf{x} \in V, p(\mathbf{x}) = 0 \implies \mathbf{x} = \mathbf{0} \text{ (Positive definiteness)} \tag{1.130}$$

Theorem 1.6.5. Euclidean norm satisfy the four axioms.

Proof. We started from the last sentence of Definition 1.5.5, for which we have $|\mathbf{x} + \mathbf{y}| = ((\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}))^{\frac{1}{2}}$. Observe

$$|\mathbf{x} + \mathbf{y}|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) \tag{1.131}$$

$$= \mathbf{x} \cdot (\mathbf{x} + \mathbf{y}) + \mathbf{y} \cdot (\mathbf{x} + \mathbf{y}) \tag{1.132}$$

$$= |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2(\mathbf{x} \cdot \mathbf{y}) \tag{1.133}$$

and observe

$$(|\mathbf{x}| + |\mathbf{y}|)^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2|\mathbf{x}||\mathbf{y}| \tag{1.134}$$

by Theorem 1.5.6, we then have $|\mathbf{x} + \mathbf{y}|^2 \leq (|\mathbf{x}| + |\mathbf{y}|)^2$. The triangle inequality follows.

Observe

$$|\lambda \mathbf{x}| = (\sum (\lambda x_i)^2)^{\frac{1}{2}} \tag{1.135}$$

$$= (\lambda^2 \sum x_i^2)^{\frac{1}{2}} \tag{1.136}$$

$$= |\lambda| (\sum x_i^2)^{\frac{1}{2}} \tag{1.137}$$

$$= |\lambda||\mathbf{x}| \tag{1.138}$$

Notice that $|\mathbf{x}| = \sum x_i^2 \geq 0$ and notice that $|\mathbf{x}| = 0 \implies \sum x_i^2 = 0 \implies \forall i, x_i = 0 \implies \mathbf{x} = \mathbf{0}$ ■

In later chapters, we will introduce norms defined in different ways and spaces. For example, in Euclidean space we can define $\|\mathbf{x}\| := \max(|x_1|, |x_2|, \ldots, |x_n|)$. One can check this definition does satisfy four axioms. Also, notice that because $\mathbb{C}$ is an Euclidean 2-space, by Theorem 1.6.5, the four axioms of norm function also apply to the absolute of complex numbers.

Next, we introduce the three axioms for inner product.

Axiom 1.6.6. (Axioms for inner product) Let V be a vector space over $\mathbb{F}$, where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. An inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ is a function that satisfy the following three axioms:

$$\forall \mathbf{x},\mathbf{y} \in V, \langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle} \text{ (Conjugate Symmetry)} \tag{1.139}$$

$$\forall \mathbf{x,y,z} \in V, \forall a, b \in \mathbb{F}, \langle a\mathbf{x} + b\mathbf{y}, \mathbf{z}\rangle = a\langle \mathbf{x}, \mathbf{z}\rangle + b\langle \mathbf{y}, \mathbf{z}\rangle \text{ (Linearity in first argument)} \tag{1.140}$$

$$\mathbf{x} \neq \mathbf{0} \longrightarrow \langle \mathbf{x}, \mathbf{x}\rangle > 0 \text{ (Positive Definitness)} \tag{1.141}$$

Notice that to satisfy the third axiom, we first have to guarantee that $\forall \mathbf{x} \in V, \langle \mathbf{x}, \mathbf{x}\rangle \in \mathbb{R}$. This is implied by the first axiom, since $\langle \mathbf{x}, \mathbf{x}\rangle = \overline{\langle \mathbf{x}, \mathbf{x}\rangle}$

Theorem 1.6.7. Euclidean inner product satisfy the three axioms.

Proof. Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, and $a, b \in \mathbb{R}$. Observe

$$\mathbf{x} \cdot \mathbf{y} = \sum x_i y_i = \sum y_i x_i = \mathbf{y} \cdot \mathbf{x} = \overline{\mathbf{y} \cdot \mathbf{x}} \tag{1.142}$$

$$(a\mathbf{x} + b\mathbf{y}) \cdot \mathbf{z} = \sum (ax_i + by_i) z_i = a \sum x_i z_i + b \sum y_i z_i = a(\mathbf{x} \cdot \mathbf{z}) + b(\mathbf{y} \cdot \mathbf{z}) \tag{1.143}$$

$$\mathbf{x} \cdot \mathbf{x} = \sum x_i^2 \geq 0 \tag{1.144}$$

■

In 2 or 3 dimension, Euclidean inner product captures the essence of angle really well, as one may remember $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos\theta$. However, to rigorously explain why $\sum x_i y_i = |\mathbf{x}||\mathbf{y}| \cos\theta$ is for now impossible, since we haven't define cos. Notice that one can define cos using Taylor series.

Normally, one doesn't use the notation $\mathbf{x} \cdot \mathbf{y}$ in place of $\langle \mathbf{x}, \mathbf{y}\rangle$, since these two notations have completely two different meaning in modern mathematics. The former is called dot product, defined only in Euclidean space and defined only by $\mathbf{x} \cdot \mathbf{y} := \sum x_i y_i$. The latter is a wide notation of every function that satisfy the three axioms. Obviously the former is only an example of the latter, and we only use $\mathbf{x} \cdot \mathbf{y}$ in place of $\langle \mathbf{x}, \mathbf{y}\rangle$, when the latter in context is equivalent to the former, e.g. Euclidean inner product.

A quite meaningless example of an inner product is $\langle \mathbf{x}, \mathbf{y}\rangle := 0$. A more complicated example of an inner product is for space of continuous complex valued function from real interval $[a, b]$ defined by $\langle f, g\rangle := \int_a^b f(t)\overline{g(t)}dt$

In last section, we "verified" that in Euclidean space, we have $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x}\rangle^{\frac{1}{2}}$. Now, we prove that every inner product induce a norm, using Parallelogram Law. However, here we have to emphasize that *even though every inner product space come with a norm, not any normed space come with an inner product.*

Definition 1.6.8. (Parallelogram Law) We say a normed space V satisfy Parallelogram Law if for all $\mathbf{x}, \mathbf{y} \in V$, we have

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) \tag{1.145}$$

Lemma 1.6.9. (Arising from inner product $\longrightarrow$ Satisfying Parallelogram Law) For an inner product space, if we define $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{\frac{1}{2}}$, then we have

Proof. Notice that $\langle \mathbf{a}, \mathbf{b} + \mathbf{c} \rangle = \langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{a}, \mathbf{c} \rangle$ and observe

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle && (1.146) \\ &= \langle \mathbf{x}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{x} - \mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle && (1.147) \\ &= 2\langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{x} \rangle && (1.148) \\ &= 2\langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{y}, \mathbf{y} \rangle = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) && (1.149) \end{aligned}$$

■

Theorem 1.6.10. Every inner product give rise to a definition of norm $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{\frac{1}{2}}$

Proof. Let V be an inner product space and define $\|\mathbf{x}\| := \langle \mathbf{x}, \mathbf{x} \rangle^{\frac{1}{2}}$. Assume there exists $\mathbf{x}, \mathbf{y}$ such that $\|\mathbf{x} + \mathbf{y}\| > \|\mathbf{x}\| + \|\mathbf{y}\|$. Observe

$$\|\mathbf{x} + \mathbf{y}\|^2 > \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| \tag{1.150}$$

Notice $\langle -\mathbf{a}, -\mathbf{a} \rangle = -\langle \mathbf{a}, -\mathbf{a} \rangle = \overline{-\langle -\mathbf{a}, \mathbf{a} \rangle} = \overline{\langle \mathbf{a}, \mathbf{a} \rangle} = \langle \mathbf{a}, \mathbf{a} \rangle$ and observe

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 &> \|\mathbf{x}\|^2 + \| - \mathbf{y}\|^2 + 2\|\mathbf{x}\|\| - \mathbf{y}\| && (1.151) \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| && (1.152) \end{aligned}$$

So we have

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 > 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) + 4\|\mathbf{x}\|\|\mathbf{y}\| \geq 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) \tag{1.153}$$

This CaC to Lemma 1.6.8.

Observe

$$\|\lambda \mathbf{x}\| = (\langle \lambda \mathbf{x}, \lambda \mathbf{x} \rangle)^{\frac{1}{2}} = (\lambda \overline{\lambda} \langle \mathbf{x}, \mathbf{x} \rangle)^{\frac{1}{2}} = (\lambda \overline{\lambda})^{\frac{1}{2}} (\langle \mathbf{x}, \mathbf{x} \rangle)^{\frac{1}{2}} = |\lambda| \|\mathbf{x}\| \tag{1.154}$$

Nonnegativity and Positive definiteness of norm function follows from Positive definiteness of inner product and $\langle \mathbf{0}, \mathbf{0} \rangle = 0$. ■

For the last sentence of the last paragraph, we here state it in precision: Let $f(x,y)$ be an inner product, then the function $g(x) = \sqrt{f(x,x)}$ satisfy the norm axiom, but if let $l(x)$ be a norm function, it doesn't always exists a function $h(x,y)$ such that $l(x) = \sqrt{h(x,x)}$ and h satisfy the inner product axioms.

An amazing fact is that for a normed space V over $\mathbb{R}$ or $\mathbb{C}$, if V satisfy Parallelogram Law, then we can define an inner product on V by $\langle \mathbf{x}, \mathbf{y} \rangle := \begin{cases} \frac{\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2}{4} & \text{if over } \mathbb{R} \\ \frac{\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2}{4} + i\frac{\|i\mathbf{x}-\mathbf{y}\|^2 - \|i\mathbf{x}+\mathbf{y}\|^2}{4} & \text{if over } \mathbb{C} \end{cases}$ so that this inner product not only satisfy all the axioms for inner product, we also have $\|\mathbf{x}\| = (\langle \mathbf{x}, \mathbf{x} \rangle)^{\frac{1}{2}}$.

So, in other word, a norm is induced by an inner product if and only if the norm satisfy the Parallelogram Law. Isn't this amazing? For the proof of only if part, we put it in complicated exercises.

We close this section with a special case of Cauchy-Schwarz inequality and the most general Cauchy-Schwarz inequality.

Theorem 1.6.11. (Cauchy-Schwarz inequality in $\mathbb{C}^n$) Let $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{C}^n$ and $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{C}^n$

$$|\sum v_j \overline{w_j}| \leq (\sum |v_j|^2)^{\frac{1}{2}} (\sum |w_j|^2)^{\frac{1}{2}} \tag{1.155}$$

and the equality hold if and only if $\exists \lambda \in \mathbb{C}, \mathbf{w} = \lambda \mathbf{v}$

Also, if we define inner product on $\mathbb{C}^n$ as $\langle \mathbf{v}, \mathbf{w} \rangle := \sum v_j \overline{w_j}$ and induce norm as $\|\mathbf{v}\| := \langle \mathbf{v}, \mathbf{v} \rangle^{\frac{1}{2}}$, the special case of Cauchy-Schwarz inequality in $\mathbb{C}^n$ can also be shortened to

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq (\|\mathbf{v}\|)(\|\mathbf{w}\|) \tag{1.156}$$

Proof. Define $A := \sum |v_j|^2 = \|\mathbf{v}\|^2, B := \sum |w_j|^2 = \|\mathbf{w}\|^2, C := \sum v_j \overline{w_j} = \langle \mathbf{v}, \mathbf{w} \rangle$

Notice that $B = 0$ implies $\forall j, w_j = 0$, then two sides of inequality are both 0 and $\lambda = 0$. We have proven the case of $B = 0$, now we prove the case of $B > 0$. Keep in mind that

$A, B \in \mathbb{R}$ and observe

$$\sum |Bv_j - Cw_j|^2 = \sum (Bv_j - Cw_j)\overline{(Bv_j - Cw_j)} \tag{1.157}$$
$$= \sum (Bv_j - Cw_j)(B\overline{v_j} - \overline{C}\overline{w_j}) \tag{1.158}$$
$$= \sum B^2|v_j|^2 - BC\overline{v_j}w_j - B\overline{C}v_j\overline{w_j} + C\overline{C}|w_j|^2 \tag{1.159}$$
$$= B^2A - BC\overline{C} - B\overline{C}C + C\overline{C}B \tag{1.160}$$
$$= B^2A - BC\overline{C} \tag{1.161}$$
$$= B(AB - |C|^2) \tag{1.162}$$

Because $B > 0$, then we can deduce

$$\sum |v_j|^2 \sum |w_j|^2 - |\sum v_j\overline{w_j}|^2 = AB - |C|^2 = \frac{1}{B}\sum |Bv_j - Cw_j|^2 \geq 0 \tag{1.163}$$

So we can deduce

$$\sum |v_j|^2 \sum |w_j|^2 \geq |\sum v_j\overline{w_j}|^2 \tag{1.164}$$

Square both side then the Theorem follows.

Notice that the equality hold true if and only if $\sum |Bv_j - Cw_j|^2 = 0$, which is equivalent to $\forall j, w_j = \frac{B}{C}v_j$, and equivalent to $\mathbf{w} = \frac{B}{C}\mathbf{v}$, where $\lambda = \frac{B}{C} = \frac{\|\mathbf{w}\|^2}{\langle \mathbf{v}, \mathbf{w}\rangle}$ ■

Theorem 1.6.12. (General Cauchy-Schwarz inequality) For all inner product space V where the norm is induced by the inner product, we have

$$|\langle \mathbf{v}, \mathbf{w}\rangle| \leq (\|\mathbf{v}\|)(\|\mathbf{w}\|) \tag{1.165}$$

and the equality hold if and only if $\mathbf{w} = \lambda\mathbf{v}$

Proof. Define $A := \|\mathbf{v}\|^2, B := \|\mathbf{w}\|^2, C := \langle \mathbf{v}, \mathbf{w}\rangle$

Notice that $B = 0$ implies $\mathbf{w} = \mathbf{0}$, which further implies $\langle \mathbf{v}, \mathbf{w}\rangle = \overline{\langle \mathbf{0}, \mathbf{v}\rangle} = \overline{\langle \mathbf{a} - \mathbf{a}, \mathbf{v}\rangle} = \overline{\langle \mathbf{a}, \mathbf{v}\rangle - \langle \mathbf{a}, \mathbf{v}\rangle} = 0$, so two side of the inequality are both 0 and the equality hold true, where $\lambda = 0$. We have proven the case of $B = 0$. Now we prove the case of $B > 0$

Keep in mind that $A, B \in \mathbb{R}$ and observe

$$\|B\mathbf{v} - C\mathbf{w}\|^2 = \langle B\mathbf{v} - C\mathbf{w}, B\mathbf{v} - C\mathbf{w}\rangle \tag{1.166}$$
$$= \langle B\mathbf{v}, B\mathbf{v}\rangle - \langle C\mathbf{w}, B\mathbf{v}\rangle - \langle B\mathbf{v}, C\mathbf{w}\rangle + \langle C\mathbf{w}, C\mathbf{w}\rangle \tag{1.167}$$
$$= B^2\langle \mathbf{v}, \mathbf{v}\rangle - CB\langle \mathbf{w}, \mathbf{v}\rangle - B\overline{C}\langle \mathbf{v}, \mathbf{w}\rangle + C\overline{C}\langle \mathbf{w}, \mathbf{w}\rangle \tag{1.168}$$
$$= B^2A - CB\overline{C} - B\overline{C}C + C\overline{C}B \tag{1.169}$$
$$= B(AB - |C|^2) \tag{1.170}$$

Because $B > 0$, we can deduce

$$\|\mathbf{v}\|^2\|\mathbf{w}\|^2 - |\langle \mathbf{v}, \mathbf{w}\rangle|^2 = AB - |C|^2 = \frac{\|B\mathbf{v} - C\mathbf{w}\|^2}{B} \geq 0 \tag{1.171}$$

So we can deduce

$$\|\mathbf{v}\|^2\|\mathbf{w}\|^2 \geq |\langle \mathbf{v}, \mathbf{w}\rangle|^2 \tag{1.172}$$

Square both side then the Theorem follows. ■

Corollary 1.6.13. (Verification of Triangle Inequality) Let V be an inner product space where the norm is induced by the inner product, we have

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\| \tag{1.173}$$

where the equality hold true only when $\langle \mathbf{v}, \mathbf{w}\rangle \geq 0$ and $\mathbf{v}, \mathbf{w}$ are linearly dependent.

Proof. Observe

$$\begin{aligned}
\|\mathbf{v} + \mathbf{w}\|^2 &= \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}\rangle && (1.174)\\
&= \|\mathbf{v}\|^2 + \langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle + \|\mathbf{w}\|^2 && (1.175)\\
&= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + \langle \mathbf{v}, \mathbf{w}\rangle + \overline{\langle \mathbf{v}, \mathbf{w}\rangle} && (1.176)\\
&= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\mathrm{Re}(\langle \mathbf{v}, \mathbf{w}\rangle) && (1.177)\\
&\leq \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2|\langle \mathbf{v}, \mathbf{w}\rangle| && (1.178)\\
&\leq \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2(\|\mathbf{v}\|)(\|\mathbf{w}\|) && (1.179)\\
&= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2 && (1.180)
\end{aligned}$$

■

1.7 Existence of Real Numbers - Dedekind Cut*

1.8 Existence of Real Numbers - Decimal*

1.9 Uniqueness of Real Numbers*

1.10 Exercises

Question 1

Given a nonzero rational r and an irrational x, prove that $r + x$ and rx are irrational.

Question 2

Prove that no rational r satisfy $r^2 = 12$

Question 3

Let E be a nonempty subset of an ordered set; suppose α is a lower bound and β is an upper bound of E. Prove $\alpha < \beta$

Question 4

Let A be a nonempty subset of real numbers which is bounded below. Define $-A := \{-x : x \in A\}$. Prove that

$$\inf A = -\sup(-A) \tag{1.181}$$

Question 5

Prove that no ordered relation can be defined on $\mathbb{C}$ so that $\mathbb{C}$ become an ordered field.

Question 6

Let $w = u + vi$ and $a = (\frac{|w|+u}{2})^{\frac{1}{2}}$ and $b = (\frac{|w|-u}{2})^{\frac{1}{2}}$

Verify that if $v \geq 0$, then $w = (a + bi)^2$, and that if $v < 0$, then $w = (a - bi)^2$

Prove that every complex number have at least 2 square roots.

Question 7

Let $z_1, \ldots, z_n \in \mathbb{C}$. Prove that

$$|\sum z_j| \leq \sum |z_j| \tag{1.182}$$

Question 8

Let $x, y \in \mathbb{C}$. Prove that

$$||x| - |y|| \leq |x - y| \tag{1.183}$$

Question 9

Let $z \in \mathbb{C}$ and $|z| = 1$. Compute

$$|1 + z|^2 + |1 - z|^2 \tag{1.184}$$

Question 10

Let $\mathbf{x} \in \mathbb{R}^k$ and $\mathbf{x} \neq \mathbf{0}$. When $k = 1$, prove that there does not exists $\mathbf{y} \in \mathbb{R}^k$, such that $\mathbf{x} \cdot \mathbf{y}$. When $k \geq 2$, prove that there exists infinitely amount of $\mathbf{y} \in \mathbb{R}^k$ such that $\mathbf{x} \cdot \mathbf{y} = 0$

Question 11

Let $k \geq 3, \mathbf{x}, \mathbf{y} \in \mathbb{R}^k, |\mathbf{x} - \mathbf{y}| = d > 0$ and $r > 0$. Prove that if $2r < d$, then there exists no $z \in \mathbb{R}^k$ such that $|\mathbf{z} - \mathbf{x}| = |\mathbf{z} - \mathbf{y}| = r$. Prove that if $2r = d$, there exists exactly one $\mathbf{z} \in \mathbb{R}^k$ that satisfy $|\mathbf{z} - \mathbf{x}| = |\mathbf{z} - \mathbf{y}| = r$. Prove that if $2r > d$, then there are infinitely many $\mathbf{z} \in \mathbb{R}^k$ that satisfy $|\mathbf{z} - \mathbf{x}| = |\mathbf{z} - \mathbf{y}| = r$. Prove that if $k = 2$ and $2r > d$, then there exists exactly 2 unique $\mathbf{z} \in \mathbb{R}^k$ such that $|\mathbf{z} - \mathbf{x}| = |\mathbf{z} - \mathbf{y}| = r$. Prove that if $k = 1$ and $2r > d$ then there exists no $\mathbf{z} \in \mathbb{R}^k$ that satisfy $|\mathbf{z} - \mathbf{x}| = |\mathbf{z} - \mathbf{y}| = r$

1.11 Complicated Exercises

In Axiom 1.2.1, we present the two ordered field axioms:

$$y < z \longrightarrow x + y < x + z \tag{1.185}$$

$$x > 0 \text{ and } y > 0 \longrightarrow xy > 0 \tag{1.186}$$

If we define the order for $\mathbb{R}$ completely in reverse, that is; for any x, y, where originally we have $x \leq y$, we now define $x \geq y$, then we can see the new order relation does not satisfy the second axiom, i.e. $x > 0$ and $y > 0 \to xy > 0$, by observing $-1 > 0 \implies 1 = (-1)^2 < 0$

Question 12: Uniquely Orderd

Prove that $\mathbb{Q}$ and $\mathbb{R}$ are uniquely ordered field; that is, any order relation defined on $\mathbb{Q}$ and $\mathbb{R}$ must be exactly the same as how we usually define it to satisfy the two ordered field axioms. Notice that you have to first come up with a way to describe our usual ways for ordering $\mathbb{Q}$ and $\mathbb{R}$ in your proof and make sure that the description do let us tell weather $x < y$ or $y < x$ for all x, y.

Next question refers to Definition 1.6.8

Question 13

Give an example of a normed vector space V such that for all inner product that can be defined on V, there exists $\mathbf{x} \in V$ such that $\|\mathbf{x}\| \neq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$, and show that this normed space does not satisfy Parallelogram Law.

Prove a norm is induced by some inner product if and only if the norm satisfy Parallelogram Law.

Question 14

Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ and $m \in \mathbb{R}^+$. Find $\mathbf{c} \in \mathbb{R}^n$ and $r > 0$ such that

$$|\mathbf{x} - \mathbf{a}| = m|\mathbf{x} - \mathbf{b}| \iff |\mathbf{x} - \mathbf{c}| = r \tag{1.187}$$

Question 15

State and prove the Pythagorean Law in Euclidean n-space, where $n \geq 2$

Chapter 2

Basic Topology

2.1 ZF

This section is written as a confirmation of knowledge. For a even more rigorous treatment, one can look for another note where the discussion of ZF is put after that of formal system and zeroth and first order logic.

Notice that in our discussion, the domain of discourse are sets. Plain speaking, every variable should be interpreted as set. In this section, we never interpret symbol x as a real number (we never use symbol x to denote a real number), and we never interpret symbol G as a group (we never use symbol G to denote a group). Whatever variable symbol we use, we interpret the symbol a set (we use the symbol to denote a set), albeit x, G or h.

However, if you wish, you can still interpret the symbol x as a number or structure or any mathematical object or even cats, but I doubt there exists any interpretation that isn't set won't make people feel weird.

Notice that ZF is a first order logic theory. We must first declare the following

Definition 2.1.1. (Atomic Formula) In our discussion in this section, there are precisely two atomic formulas:

$$x \in A \tag{2.1}$$

and

$$A = B \tag{2.2}$$

The above definition may be confusing if one has no basic knowledge for formal system. I here highly recommend the first two chapters of the excellent book *Elementary Formal System* written by Smullyan, and one can understand what does the words "formula" and "axiom" mean.

Now we start our discussion of ZF.

Axiom 2.1.2. (Axiom of Extension)

$$\forall x(x \in A \longleftrightarrow x \in B) \longrightarrow A = B \tag{2.3}$$

TBH, I have no clue why this axiom is named axiom of extension. This axiom axiomatically define equality (between sets).

The practice of putting the word *between sets* in a parenthesis in the above paragraph serve as a reminder that the domain of discourse of our discussion should be interpreted as the class of all sets.

ZF in our discussion is a theory where the only two non-logical symbol is $\in$ and $=$. In our discussion, whenever the symbols $\subseteq, \cup, \cap, \neq$ appear, one should immediately realize the sentence in which the symbols appear is an interpretation.

In ZF, we never define $A \subseteq B$. We never define $A \neq B$. We never define $A \cup B$. We merely construction a formal sentence that one feel it make sense to interpret as $A \subseteq B$ or the English sentence "A is a subset of B.".

Definition 2.1.3. (Interpretation) We can interpret the formal sentence $\forall x(x \in A \longrightarrow x \in B)$ as informal sentence $A \subseteq B$. If we ever say $A \subseteq B$, we mean $\forall x(x \in A \longrightarrow x \in B)$

Definition 2.1.4. (Interpretation) We can interpret the formal sentence $\exists x((x \in A \wedge \neg x \in B) \vee (x \in B \wedge \neg x \in A))$ as the sentence $A \neq B$.

Axiom 2.1.5. (Axiom of Subset) Let P be a predicate. The axiom states

$$\forall A \exists B \forall x(x \in B \longleftrightarrow x \in A \wedge P(x)) \tag{2.4}$$

Because of axiom of subset, when we informally says $B = \{x \in A : P(x)\}$ in our argument, we know B exists if P is a predicate. Notice that Russell's paradox rely on the construction of such set: $\{y : P(y)\}$ where predicate $P(x)$ is defined by $\neg x = x$. The axiom of subset only allow us to construct "subset" (not set) filled with elements satisfying certain predicate.

Also, the axiom of subset allow us to construct intersection of two sets. Consider the set $\{x \in A : P(x)\}$ where $P(x)$ is defined by $x \in B$.

Notice that one at this point can trivially and formally deduce the existence of two sets that are ought to be interpreted as $A \cap B$ and $B \cap A$, and formulate a formal sentence that is ought to be interpreted as "$A \cap B$ is unique" and another formal sentence that is ought to be interpreted as "$A \cap B = B \cap A$", and deduce these two formal sentences.

One can also formulate and deduce the formal sentence of "intersection operation is associative" as an exercise.

Now, before we introduce the next axiom, we first declare that the symbol $\varnothing$ is in our alphabet.

Axiom 2.1.6. (Axiom of Existence of Empty Set)

$$\exists \varnothing \forall y \neg (y \in \varnothing) \tag{2.5}$$

In some people's formation of ZF, the axiom of existence of empty set isn't necessary, since it is possible to construct the empty set with axiom of subset. If we already know that the universe is nonempty, we can arbitrarily pick a set A and construct $\{x \in A : P(x)\}$ where $P(x)$ is defined $\neg x = x$.

At this point, one can try to formulate the formal sentence of "empty set is unique and is a subset of every set" and formally deduce such using axiom of extension.

We now introduce an axiom to construct new set.

Axiom 2.1.7. (Axiom of Pairing)

$$\forall x \forall y \exists A \forall v (v \in A \longleftrightarrow v = x \vee v = y) \tag{2.6}$$

By axiom of pairing, one can deduce the formal sentence of "$\{x, y\}$ exists if x, y exist".

Notice that we can also deduce the formal sentence of "$\{x\}$ exists if x exists".

Axiom 2.1.8. (Axiom of Union)

$$\forall A \exists B \forall x (x \in B \longleftrightarrow \exists C (C \in A \wedge x \in C)) \tag{2.7}$$

By axiom of union, one can deduce the formal sentence of "$\bigcup X$ exists, if X exist".

Also, one can try to formulate and deduce the formal sentence of "For all set X, we know $\bigcup X$ is unique", that of "For all sets x, y, we know $x \cup y = y \cup x$" and that of "For all sets x, y, z, we know $(x \cup y) \cup z = x \cup (y \cup z)$".

Now we introduce a new axiom to construct even more sets.

Axiom 2.1.9. (Axiom of Power Set)

$$\forall A \exists B \forall x (x \in B \longleftrightarrow \forall y (y \in x \longrightarrow y \in A)) \tag{2.8}$$

The axiom of power set can be interpreted as that every set has at least one power set (and it can be proved that there exists only one for every set and that if two set have the same power set then the two set are the same). Normally and informally, we denote the power set of x as $\mathcal{P}(x)$

Now we introduce the standard way to "construct a model for natural numbers", that is, to "formally formulate sets that can be interpreted as natural numbers". The construction is straight forward: we interpret $\varnothing$ as 0, and while interpreting x as a natural, interpret $x \cup \{x\}$ as $x + 1$.

This construction seems to satisfy Peano axioms, but a problem will emerge when one trying to verify: one can construct any natural number, for instance 3, by constructing $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$, but one can not construct the set of all natural numbers, that is, informally writing, $\mathbb{N} = \{0, 1, 2, 3, \ldots\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, \ldots\}$.

The construction of the set $\mathbb{N}$, rely on the next axiom (while the construction of no element in $\mathbb{N}$ rely on the next axiom). In fact, the formulation of any infinite set rely on the next axiom. Without this axiom, we can not formulate any infinite set (after we construct the predicate $P(x)$ to be interpreted as "$P(x)$ is true if x is infinite").

Axiom 2.1.10. (Axiom of Infinity)

$$\exists I (\varnothing \in I \wedge \forall y (y \in I \longrightarrow y \cup \{y\} \in I)) \tag{2.9}$$

Notice that every set interpreted as a natural number can be proved to belong to the set I in Axiom 2.1.10. In other (informal) words, the set interpreted as $\mathbb{N}$ is a subset of I.

Before we formulate such subset of I, we first state the five Peano Axioms, outside of our formal system. One can and should read the following five axioms algebraically, not formally.

Axiom 2.1.11. (Peano Axioms, outside of our formal system) We say $\langle U, S\rangle$, where S is a function, is a model of Peano axioms if

$$0 \in U \text{ (First Peano axiom)} \tag{2.10}$$

$$a \in U \implies S(a) \in U \text{ (Second Peano axiom)} \tag{2.11}$$

$$\forall x \in U, S(x) \neq 0 \text{ (Third Peano axiom)} \tag{2.12}$$

$$S(a) = S(b) \implies a = b \text{ (Forth Peano axiom)} \tag{2.13}$$

$$0 \in X \text{ and } \forall u \in U, u \in X \longrightarrow S(u) \in X \implies U \subseteq X \text{ (Induction axiom)} \tag{2.14}$$

Before reading the proof of the following theorem, notice that the set $\{y \in \bigcup x : \forall z(z \in x \longrightarrow y \in x)\}$ is what we mean by $\bigcap x$, and notice that the set $\{a \in x : \neg a \in y\}$ is what we mean by $x \setminus y$.

One can try to completely formalize $\bigcap x$ and $x \setminus y$

Theorem 2.1.12. (Existence of Model of Natural Numbers in ZF) We first define five predicate resembling the five Peano axioms.

$$P_1(x) := \varnothing \in x \tag{2.15}$$

$$P_2(x) := \forall a(a \in x \longrightarrow a \cup \{a\} \in x) \tag{2.16}$$

$$P_3(x) := \forall a(a \in x \longrightarrow a \cup \{a\} \neq \varnothing) \tag{2.17}$$

$$P_4(x) := \forall a \forall b(a \in x \wedge b \in x \wedge a \cup \{a\} = b \cup \{b\} \longrightarrow a = b) \tag{2.18}$$

$$P_5(x) := \forall a([\forall b(b \in a \longrightarrow b \in x) \wedge \varnothing \in a \wedge \forall b(b \in a \longrightarrow b \cup \{b\} \in a)] \longrightarrow x = a) \tag{2.19}$$

Reminder: Notice that the notation $\cup$ is not in our formal language, I use the notation only for abbreviation.

The theorem states:

$$\exists A(P_1(A) \wedge P_2(A) \wedge P_3(A) \wedge P_4(A) \wedge P_5(A)) \tag{2.20}$$

Proof. Let I be the set in Axiom 2.1.10. Define $\mathcal{L} := \{X \in \mathcal{P}(I) : \varnothing \in X \wedge \forall a(a \in X \longrightarrow a \cup \{a\} \in X)\}$. From now on, we will call $a \cup \{a\}$ the successor $S(a)$ of a in our proof.

Then $\mathcal{L}$ can be precisely described as the set that contain all subsets X of I such that the set interpreted as 0 is in X and that the successor of each element in X belong to X.

Let $Y = \bigcap \mathcal{L}$. We seek to show that

$$P_1(Y) \wedge P_2(Y) \wedge P_3(Y) \wedge P_4(Y) \wedge P_5(Y) \tag{2.21}$$

$P_1(Y)$ and $P_2(Y)$ are trivial. (done) (done)

$P_3(Y)$ is also trivial. (done)

We now prove $P_5(Y)$

Let $E \subseteq \bigcap \mathcal{L}, \varnothing \in E$ and $\forall b(b \in E \longrightarrow b \cup \{b\} \in E)$. Notice that $E \subseteq \bigcap \mathcal{L} \implies \forall X \in \mathcal{L}, E \subseteq X \subseteq I$, so $E \in \mathcal{L}$. Then $\bigcap \mathcal{L} \subseteq E$ (done)

We now prove $P_4(Y)$

Before reading the following, one may want to bear in mind that $\forall x(x \in x \cup \{x\}$ and $x \subseteq x \cup \{x\})$

Define $S := \{n \in \bigcap \mathcal{L} : \forall m(m \in n \longrightarrow m \subseteq n)\}$. We see that $S \subseteq \bigcap \mathcal{L}$ and that $\{\varnothing\} \in S$, trivially, so by $P_5(Y)$, we only have to show $n \in S \longrightarrow n \cup \{n\} \in S$ to show $S = \bigcap \mathcal{L}$. To show such, observe that if $n \in S$ and $m \in n \cup \{n\}$, then we know $m \in n$ or $m \in \{n\}$, which is equivalent to $m \in n$ or $m = n$, which implies $m \subseteq n$, since by the definition of S, both $m \in n$ and $m = n$ implies $m \subseteq n$. Then we know $n \in S$ and $m \in n \cup \{n\} \longrightarrow m \subseteq n \subseteq n \cup \{n\}$.

The above prove that every element of element n of $\bigcap \mathcal{L}$ is a subset of n. Assume $m \neq n \in \bigcap \mathcal{L}$ but $m \cup \{m\} = n \cup \{n\}$. Because $n \in n \cup \{n\} = m \cup \{m\}$ and $n \neq m$, we can deduce $n \in m$, which implies $n \subseteq m$. Using the same method we can also deduce $m \subseteq n$ CaC (done) ■

Notice that we didn't prove Peano axioms give a unique structure up to isomorphism. I personally haven't proved it yet, but I guess to prove such isn't trivial and will require to show structure like $\{0, 1, 2, \ldots; a, b\}$ where $S(a) = b$ and $S(b) = a$ does not satisfy the induction axiom.

Say, we have proved the uniqueness of model of Peano axiom. We still haven't show in the universe of ZF, our choice of model for natural number is unique. For all we know, maybe the construction $\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\{\{\varnothing\}\}\}, \ldots$ also works. (Hint: Let's say there are other construction for natural numbers, there are still reasons we pick $\bigcap \mathcal{L}$ as our model, since later you will see $a < b \in \mathbb{N} \iff a \in b$ and $a \subset b$)

So, it seems like Theorem 2.1.12 is relatively weak. This is not the case however. The fact is that, we don't really require anything extra of Theorem 2.1.12 to do further discussion.

The core concept of Axiom is that we don't care about what a mathematical object is, *we only care about what a mathematical object does*. One should *never* view axiom as something that is apparently true and can not be proved. One should know: in most domains of mathematics, individually, axioms are simply a convention between one mathematician to another. *Axioms are properties we wish a object to have so that we can make knowledge that is truly important using the property of the object.*

For instance, we never care about what are sets. We never care about what is $\mathbb{N}$. We never care about what is $\mathbb{R}$. We only care about that $\mathbb{R}$ is completed and we can do induction on $\mathbb{N}$. It doesn't matter if $\mathbb{N}$ contain a cat or Homer Simpson, as long as it satisfy Peano axiom and we can deduce all the Theorem.

Notice the construction of natural number: $\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \ldots$ can each be formally proved to belong to $\bigcap \mathcal{L}$. We now introduce the last axiom.

Axiom 2.1.13. (Axiom of Foundation)

$$\forall X(X \neq \varnothing \longrightarrow \exists Y(Y \in X \wedge \forall z(z \in Y \longrightarrow \neg z \in X))) \tag{2.22}$$

In English, the axiom of foundation states that every nonempty set X has an element Y such that $X \cap Y = \varnothing$. This axiom is sometimes discarded by people who use ZFC for mathematics not so related to set theory, since in the construction of most mathematical object, this axiom is unnecessary. This is also the reason this axiom is introduced last in this section. But of course, this axiom have purpose. We now introduce some results.

Theorem 2.1.14. (No set belong to itself)

$$\forall x, \neg x \in x \tag{2.23}$$

Proof. Assume $\exists x, x \in x$. Observe that by the axiom of foundation we have $x \cap \{x\} = \varnothing$, but we also have $x \in x$ and $x \in \{x\}$, so $x \in x \cap \{x\}$ CaC ■

Theorem 2.1.15. (No infinitely descending chain of sets) Informally, we can say that the following situation will never happen

$$\cdots \in x_3 \in x_2 \in x_1 \tag{2.24}$$

Formally, we say

$$\neg \exists X (\forall x (x \in X \longrightarrow \exists y (y \in X \wedge y \in x))) \tag{2.25}$$

which reads "There does not exist a set X such that for all $x \in X$ there exists a set $y \in X$ belong to x"

Proof. Assume there exists a set X such that for all $x \in X$ there exists $y \in X \cap x$, we immediately see that for all $x \in X$, the intersection $x \cap X$ is nonempty CaC . ■

2.2 Ordinal

In previous section, we employed a purely set-theoretic formal language to delve into Zermelo-Fraenkel Set Theory. While this methodology is undoubtedly elegant for discussing set theory, it becomes notably cumbersome in mathematical domains with limited connections to set theory. This is mainly because many of the tools we typically rely on (such as functions, groups, real numbers, relations, and more) become unavailable. To utilize them, we must first construct all the sets that are interpreted as these tools, leading to extensive and convoluted formulae.

In this section, our goal isn't to rigorously prove everything using a formal language. Instead, we'll explain our proofs in sufficient detail so that, if desired, they can be formalized with minimal additional effort, albeit taking a bit more time.

Please note that this section primarily serves as a foundational toolkit for the subsequent sections. We will introduce all the necessary tools, and that's the extent of it.

In last section, we use a very philosophical and logical way to discuss. Here, we back to our usual algebraic and analytical way to discuss. We firs introduce axioms for ordered sets.

Axiom 2.2.1. (Axiom for Poset) Let $\leq$ be a relation on set S. We say $\langle S, \leq \rangle$ is partially ordered if

$$\forall x \in S, x \leq x \text{ (Reflexive)} \tag{2.26}$$

$$\forall x, y, z \in S, x \leq y \text{ and } y \leq z \implies x \leq z \text{ (Transitive)} \tag{2.27}$$

$$\forall x, y \in S, x \leq y \text{ and } y \leq x \implies x = y \text{ (Antisymmetric)} \tag{2.28}$$

Notice that poset dose not require every two element to be comparable (trichotomy). Totally ordered set require such.

Axiom 2.2.2. (Axiom for Totally Ordered Set) Let $\leq$ be a relation on set S. We say $\langle S, \leq \rangle$ is totally ordered (or linearly ordered) if

$$\forall x \in S, x \leq x \text{ (Reflexive)} \tag{2.29}$$

$$\forall x, y \in S, x \leq y \text{ or } y \leq x \text{ (Connected)} \tag{2.30}$$

$$\forall x, y, z \in S, x \leq y \text{ and } y \leq z \implies x \leq z \text{ (Transitive)} \tag{2.31}$$

$$\forall x, y \in S, x \leq y \text{ and } y \leq x \implies x = y \text{ (Antisymmetric)} \tag{2.32}$$

We now give the axioms for strictly ordered sets.

Axiom 2.2.3. (Axiom for Strictly Totally and Partially Ordered Set) Let $<$ be a relation on set S. We say $\langle S, < \rangle$ is strictly totally ordered if

$$\forall x \in S, x \not< x \text{ (Irreflexive)} \tag{2.33}$$

$$\forall x, y \in S, x < y \implies y \not< x \text{ (Asymmetric)} \tag{2.34}$$

$$\forall x, y, z \in S, x < y \text{ and } y < z \implies x < z \text{ (Transitive)} \tag{2.35}$$

$$\forall x, y \in S, x \neq y \implies x < y \text{ or } y < x \text{ (Connected)} \tag{2.36}$$

$\langle S, < \rangle$ is strictly partially ordered if satisfy irreflexive, asymmetric and transitive axiom.

Notice that if S is totally ordered, then we can let S be strictly totally ordered by defining $a < b$ if $a \leq b$ and $a \neq b$, and if S is strictly totally ordered, we can let S be totally ordered by defining $a \leq b$ if $a < b$ or $a = b$.

Again, if S is partially ordered, then we can let S be strictly partially ordered by defining $a < b$ if $a \leq b$ and $a \neq b$, and if S is strictly partially ordered, we can let S be partially ordered by defining $a \leq b$ if $a < b$ or $a = b$.

We now give definition of some basic notions.

Definition 2.2.4. (Definition of Minimal and Maximal) A maximal element a of a subset X of poset Y or totally ordered set Y is an element that no element $b \in X$ satisfy $a < b$. Minimal element is defined in similar fashion.

Definition 2.2.5. (Definition of Well Ordered Set) A totally ordered set S is said to be well ordered if every nonempty subset X of S has a minimal element x.

Notice that $\mathbb{N}$ is well ordered, but all $\mathbb{Z}, \mathbb{Q}, \mathbb{Q}^+, \{0\} \cup \mathbb{Q}^+, \mathbb{R}$ is not well ordered.

We now give the definition of ordinal, which is the core of this section.

Notice that there are other ways to define ordinal, but Von Neumann ordinal is the most elegant that can be formalized by our system in section 2.1.

Definition 2.2.6. (Definition of Von Neumann Ordinal) A set S is an ordinal if S is a strictly totally well ordered set with respect to $\in$ and every element of S is a subset of S. Precisely, we say a set S is an ordinal if satisfy the following:

$$\forall x \in S, x \notin x \text{ (Irreflexive)} \tag{2.37}$$

$$\forall x, y \in S, x \in y \implies y \notin x \text{ (Asymmetric)} \tag{2.38}$$

$$\forall x, y, z \in S, x \in y \text{ and } y \in z \implies x \in z \text{ (Trnasitive)} \tag{2.39}$$

$$\forall x, y \in S, x \neq y \implies x \in y \text{ or } y \in x \text{ (Connected)} \tag{2.40}$$

$$\forall V \subseteq S, \exists x \in V, \forall y \in V, x \in y \text{ (Well Order)} \tag{2.41}$$

$$\forall x \in S, x \subseteq S \text{ (Subset)} \tag{2.42}$$

In other words, S is an ordinal if $\langle S, \in \rangle$ is strictly totally well ordered and every element of S is a subset of S.

Von Neumann's definition of ordinal may seem simple and straight forward, but is in fact very strong. Here we prove properties of Von Neumann ordinal we intend to have by definition.

Notice that in ZF, irreflexive and antisymetric is trivially satisfied by axiom of foundation.

Theorem 2.2.7. (Every Element of Ordinal is an Ordinal) If S be an ordinal, then $A \in S$ is an ordinal.

Proof. Notice that if $A = \varnothing$, then the proof is all trivial, so we only have to consider $A \neq \varnothing$. Because $A \subseteq S$, so transitive, connected and well ordered is trivially proved.

We now prove $\forall x \in A, x \subseteq A$.

Let $x \in A$, if x are all empty, the proof is trivial. Let $\alpha \in x$, so we have $\alpha \in x \in A$

Because $A \in S$, we know $A \subseteq S$, so we know $x \in A \subseteq S$. Then we know $x \subseteq S$, so we know $\alpha \in x \subseteq S$. So we know α, x, A all belong to S, then because S is transitive, we conclude $\alpha \in A$ (done) ■

Theorem 2.2.8. (Ordinal Successor) If a is an ordinal, then $a \cup \{a\}$ is an ordinal.

Proof. Notice that $b \in a \cup \{a\} \implies b \in a$ or $b \in \{a\} \implies b \in a$ or $b = a$. If $b = a$ and $b \in a$, we see $a \in a$, so we know $b \in a \cup \{a\} \implies b \in a$ exclusively or $b = a$.

The proofs for subset part and connected part are relatively simple. Let $x, y \in a \cup \{a\}$. If $x \in a$, then $x \subseteq a \subseteq a \cup \{a\}$. If $x = a$, then $x = a \subseteq a \cup \{a\}$. Let $x \neq y$. If x, y are both in a, then the proof is finished. If, WOLG, $y = a$, then $x \in a = y$.

We now prove the transitive part.

Let $x, y, z \in a \cup \{a\}$ and $x \in y$ and $y \in z$. By axiom of foundation, it is impossible more than one of x, y, z equals to a. If $x = a$, then we have $y \in a$, so we see $\cdots a = x \in y \in a \in y \in a$, which implies it is impossible $x = a$. Similarly, we can deduce $y \neq a$. So far, we have proven $x \in a$ and $y \in a$. If $z \in a$, then because a is an ordinal, our proof if done. If $z = a$, we see $x \in a = z$ (done)

We now prove the well ordered part.

Let $V \subseteq a \cup \{a\}$. By axiom of foundation, we know a and $\{a\}$ are disjoint, so if $a \notin V$, then $V \subseteq a$, and by the definition of a, our proof is finished. If $a \in V$, we see either a is the only element of V or there exist element other than a in V. If the situation is former, then the proof is trivially finished. If the situation is latter, then we observe $V \setminus \{a\} \subseteq a$, so $V \setminus \{a\}$ has a minimal element x, and $x \in a$ indicate x is also the minimal element of V (done) ■

Theorem 2.2.9. (Basic Property of Ordinals) If a, b are ordinals, we have

$$a \in b \iff a \subset b \tag{2.43}$$

Proof. From left to right is relatively simple. If $a \in b$, then we have $a \subseteq b$. By axiom of foundation, we know $a \in b \implies a \neq b$, so we have $a \subset b$.

($\longleftarrow$)

Let $a \subset b$. We wish to prove $a = x := \min b \setminus a$, so that $a = \min b \setminus a \in b$. We now prove $x \subseteq a$ and prove $a \subseteq x$.

By definition of x, we have $\forall y \in b \setminus a, y \notin x$. This tell us $b \setminus a$ and x are disjoint. Because $x \subseteq b$, the fact that $b \setminus a$ and x are disjoint tell us that $x \subseteq a$ (done)

Let $z \in a$, we wish to prove $z \in x$. Notice that $z \in a \subset b$, so because b is an ordinal and $x \in b$, we know either $z \in x$ or $z = x$ or $x \in z$. Assume $x = z$. We have $\min b \setminus a = x = z \in a$ CaC . Assume $x \in z$. Because a is an ordinal and $z \in a$, we have $\min b \setminus a = x \in z \subseteq a$ CaC . Now we have $z \in x$ as desired (done) ■

Lemma 2.2.10. If A, B are ordinals, then $A \cap B$ is an ordinal.

Proof. Transitive and connected and well ordered of $A \cap B$ trivially inherit from that of A. Observe $x \in A \cap B \implies x \in A$ and $x \in B \implies x \subseteq A$ and $x \subseteq B \implies x \subseteq A \cap B$, and we are finished. ■

Theorem 2.2.11. (Trichotomy of Ordinals) If a, b are ordinals, we have

$$a \subset b \text{ or } a = b \text{ or } b \subset a \tag{2.44}$$

Proof. Assume $a \not\subset b$ and $a \neq b$ and $b \not\subset a$. We have $a \setminus b \neq \varnothing$ and $b \setminus a \neq \varnothing$, since, say, if $a \setminus b = \varnothing$, we have $a \subseteq b$.

By Lemma 2.2.10, we know $a \cap b$ is an ordinal. Because $a \setminus b \neq \varnothing$, we know $a \cap b \subset a$, and then by Theorem 2.2.9 we have $a \cap b \in a$. Similarly, we have $a \cap b \in b$. Then we have $a \cap b \in a \cap b$ CaC ■

Corollary 2.2.12. (Trichotomy of Ordinals) If a, b are ordinals, we have

$$a \in b \text{ or } a = b \text{ or } b \in a \tag{2.45}$$

Corollary 2.2.13. (The Class of All Ordinal is Totally Ordered) The class of all ordinals is totally ordered by set-membership $\in$.

Proof. The irreflexive and assymmetric part is trivially satisfied by axiom of foundation. The transitive part is simple: $x \in y$ and $y \in z \implies x \in y \subseteq z$. The connected part is just proven by Corollary 2.2.12. ■

Theorem 2.2.14. (The Class of All Ordinals is Well Ordered) Every set of ordinals has a $\in$-minimal element.

Proof. If not, then it contradicts to Theorem 2.1.15 ■

Theorem 2.2.15. (Burali-Forti Paradox) All ordinals do not form a set.

Formally speaking, define $P(S)$ are true if and only if S is an ordinal. Then we have

$$\neg\exists A \forall x (x \in A \longleftrightarrow P(x)) \tag{2.46}$$

Proof. Assume Ω is the set of all ordinals. By Theorem 2.2.14, we see Ω satisfy at least irreflexive, asymmetric, trnasitive, connected and well ordering requirement of ordinal.

Let $x \in \Omega$, and $y \in x$. By Theorem 2.2.7, we know $y \in \Omega$, so in fact we have $x \subseteq \Omega$, which prove Ω is an ordinal. Then we have $\Omega \in \Omega$ CaC ■

Whewww, we have proved a lot of intended properties of ordinal. Notice that although there does not exist a set consist precisely of all ordinals. We still can mention the class of all ordinal by Ord.

Notice that Ord is in fact an extension of $\mathbb{N}$, and it goes like this

$$0, 1, 2, 3, \ldots; \omega, \omega + 1, \omega + 2, \ldots \tag{2.47}$$

Theorem 2.2.16. (All Naturals are Ordinals) All element of $\bigcap \mathcal{L}$ are ordinals.

Proof. We prove by induction, which is the $P_5(\bigcap \mathcal{L})$ of Theorem 2.1.12.

Notice that $\varnothing$ trivially satisfy all property of ordinals, which finish the proof for base step, and notice that by Theorem 2.2.8, the proof for induction step is also finished. ■

Theorem 2.2.17. ($\mathbb{N}$ as a set is an ordinal) $\bigcap \mathcal{L}$ is an ordinal.

Proof. By Theorem 2.2.16, we have proved that $\cap\mathcal{L}$ is a strictly totally ordered set. Notice that every subset of a well ordered set is well order, so it only left to prove the subset part of definition of ordinal.

We prove by induction. Let $X := \{x \in \bigcap \mathcal{L} : x \subseteq \bigcap \mathcal{L}\}$. Base case: $\varnothing \in X$ is trivial. Let $a \in X$. Then we have $a \in \bigcap \mathcal{L}$ and $a \subseteq \bigcap \mathcal{L}$, so we have $a \cup \{a\} \subseteq \bigcap \mathcal{L}$ (done) ■

Theorem 2.2.18. (Transfinite Induction) Let X be a nonempty ordinal. For all $x \in X$, we define

$$S_x := \{z \in X : z \in x\} \tag{2.48}$$

Let $Y \subseteq X$. If Y satisfy the property:

$$\forall x \in X(S_x \subseteq Y \longrightarrow x \in Y) \tag{2.49}$$

we have $Y = X$

Proof. Assume Y satisfy the property, but $Y \neq X$. We first show Y is nonempty. Because X is an ordinal, we can pick an element $r \in X$ such that $\forall u \in X, r \in u$, so we have $S_r = \varnothing \subseteq Y$. Then because Y satisfy the property, we have $r \in Y$ (done)

Because X is an ordinal, we know there exists $x_0 \in X \setminus Y$ such that $\forall x \neq x_i \in X \setminus Y, x_0 \in x$.

Let $a \neq x_0 \in X$, we know either $a \in Y$ or $a \in X \setminus Y$, we now prove $a \in x_0 \implies a \in Y$. Assume $a \notin Y$, we then have $a \in X \setminus Y$, so we have $x_0 \in a$ CaC (done) .

The blue part show that $S_{x_0} \subseteq Y$, so by the property of Y, we have $x_0 \in Y$, but $x_0 \in X \setminus Y$ CaC .

■

2.3 Axiom of Choice and its equivalents

The axiom of choice in practice allow us to carry out constructions that require infinite sequence of choice, each of which depending on the preceding one, so that one does not know initially just what choice are to made in what order.
In this section, we let P be a nonempty poset.

Definition 2.3.1. (Definition of Chain) We say $X \subseteq P$ is a chain if X is totally ordered.

Definition 2.3.2. (Definition of Well Ordered Set) Let $X \subseteq P$. We say X is well ordered only if X is totally ordered.

Definition 2.3.3. (Definition of Upper Bound) Let $X \subseteq P$. We say X is bounded above by t if

$$\forall x \in X, x \leq t \tag{2.50}$$

Definition 2.3.4. (Definition of Initial Segment) Let C be a chain of P. We say T is an initial segment of C if

$$T \subseteq C \text{ and } \forall v \in T, \forall u \in C, u \leq v \implies u \in T \tag{2.51}$$

In this section, we denote the statement: T is an initial segment of C by $T \subseteq' C$, and we denote $T \subseteq' C$ and $T \neq C$ by $T \subset' C$

Lemma 2.3.5. An initial segment of a chain is also a chain

Lemma 2.3.6. Let $S \subseteq P$ be well ordered and t be an upper bound of S. We have $S \cup \{t\}$ is also well ordered.

Proof. Let $V \subseteq S \cup \{t\}$ and $V \neq \varnothing$. If $V \cap S$ is empty, then we know $V = \{t\}$ and the proof is trivially finished. If $V \cap S$ is nonempty, let $x = \min V \cap S$. We know $y \in V \implies y \in V \cap S$ or $y = t$. If $y \in V \cap S$, then $x \leq y$ by definition of x. If $y = t$, then $x \leq t = y$ by $x \in S$ and the definition of t ■

Lemma 2.3.7. If $\mathcal{C}$ is a set of well ordered chains in P where

$$X, Y \in \mathcal{C} \implies X \text{ is an initial segment of } Y \text{ or } Y \text{ is an initial segment of } X \tag{2.52}$$

then $\bigcup \mathcal{C}$ is a well ordered chain

Proof. Notice that the definition of $\mathcal{C}$ implies

$$X, Y \in \mathcal{C} \implies X \subseteq Y \text{ or } Y \subseteq X \tag{2.53}$$

Let $a, b \in \bigcup \mathcal{C}$, and let $a \in A \in \mathcal{C}, b \in B \in \mathcal{C}$. Observe $A \subseteq B$ or $B \subseteq A \implies a, b \in A$ or $a, b \in B$, because A, B are chains, we deduce a, b are comparable. All the other properties of totally ordered set inherit from P, so we have proven $\bigcup \mathcal{C}$ is a chain.

We first prove $\langle \mathcal{C}, \subseteq' \rangle$ is totally ordered.

Connected part and antisymmetric part are implied by the definition of $\mathcal{C}$. One can verify the reflexive part trivially hold true. Let $A \subseteq' B \subseteq' C$ and $a \in A, c \in C$ and $c \leq a$. Because $a \in B$ and $B \subseteq' C$, we can deduce $c \in B$ from $c \leq a$. Then we can deduce $c \in A$ from $c \in B$ and $A \subseteq' B$, finishing the proof for transitive part (done)

Let $V \subseteq \bigcup \mathcal{C}$. We now prove V contain a minimal element.

Arbitrarily pick $X \in \mathcal{C}$ so that $V \cap X$ is nonempty. Because $V \cap X \subseteq X$ and X is well ordered, we know $\min V \cap X$ exists. Assume $\exists v' \in V, v' < \min V \cap X$. We know $\min V \cap X \in X \in \langle \mathcal{C}, \subseteq' \rangle$. Let $S = \{Z \in \mathcal{C} : v' \in Z\}$. If $X \in S$, then $v' \in V \cap X$ contradicting to $v' < \min V \cap X$, so we know $X \notin S$. Observe $v' \in Z \subseteq' X \implies v' \in X$ causing the same contradiction, so we know every element in S is greater than X with respect to $\subseteq'$. We know S is non empty, since $v' \in V \subseteq \bigcup \mathcal{C}$. Then we can arbitrarily pick $Z \in S$. By definition of initial segment, we can deduce $v' \in X$ from $v' < \min V \cap X$ and $Z \subseteq' X$ CaC (done) ■

Lemma 2.3.8. The class of initial segments of a chain is totally ordered by $\subseteq'$

Proof. Antisymmetric, reflexive and transitive parts are implied by definition of $\subseteq'$, as shown in the proof of Lemma 2.3.7. Now we let X be a chain and let $A \subseteq' X$ and $B \subseteq' X$, and prove $A \subseteq' B$ or $B \subseteq' A$

Assume $A \not\subseteq B$ and $B \not\subseteq A$. Arbitrarily pick $x \in A \setminus B$ and $y \in B \setminus A$. We know $x \neq y$, since $y \notin A \setminus B$. WOLG, let $x < y$. Because $x \in A \subseteq X$, where $y \in B$ and $B \subseteq' X$, so we deduce $x \in B$ CaC

WOLG, let $A \subseteq B$. Assume $A \not\subseteq' B$. Then there exists $a \in A, b \in B$ such that $b < a$ and $b \notin A$. Yet, we see $A \subseteq' X, b \in X$ and $a \in A$, so we can deduce $b \in A$ CaC ■

Lemma 2.3.9. A class of chain $\mathcal{C}$ is totally ordered by $\subseteq'$ if and only if $\bigcup \mathcal{C}$ is a chain.

It may seem that Lemma 2.3.7 is extremely powerful as it "seems" to suggest, the union of a well ordered by inclusion set of well ordered set is well ordered. But this is not the case as one can see that $\bigcup\{\{-1\},\{-1,-2\},\{-1,-2,-3\},\ldots\}$ is not well ordered.

We now prove from Axiom of Choice to Zorn's Lemma.

Theorem 2.3.10. (Axiom of Choice Implies Zorn's Lemma) If for arbitrary set of nonempty set Y and arbitrary function $g : X \to Y$, there exists a choice function $f : X \to \bigcup Y$ such that $f(x) \in g(x)$, then for every partially ordered set P such that every chain is bounded, there exists a maximal element.

Proof. Because the set of upper bound of chain is always nonempty, we can let g be a function that map each chain C in P to the set of upper bounds of C, and let f be the choice function of g, that is, $f(x) \in g(x)$. Additionally, let $f(\varnothing) := p$.

The above use the axiom of choice. We now define

$$\mathcal{C} := \{C \subseteq P : C \text{ is well ordered}, \min C = p, T \subset' C \longrightarrow f(T) = \min C \setminus T\} \quad (2.54)$$

We see $\{p\} \in \mathcal{C}$, so $\mathcal{C}$ is nonempty. We first prove $\langle \mathcal{C}, \subseteq' \rangle$ is totally ordered.

Reflexive part, transitive part and antisymmetric part are implied by the definition of $\subseteq'$, as shown in the proof of Lemma 2.3.7. Let $A, B \in \mathcal{C}$, and define

$$R := \bigcup\{T : T \subseteq' A \text{ and } T \subseteq' B\} \quad (2.55)$$

Let $x \in R$ and $y \in A$. Then we find T such that $x \in T \subseteq' A$. Observe if $y < x$, then $y \in T \subseteq R$. This prove $R \subseteq' A$. Similarly, we can prove $R \subseteq' B$. Then by the definition of R, we know R is greatest common initial segment of A and B.

Assume $R \subset' A$ and $R \subset' B$. Then we see $f(R) = \min A \setminus R = \min B \setminus R$. Let $a \in A$ and $r' \in R \cup \{f(R)\}$. Observe

$$a < r' \implies \begin{cases} a < f(R) = \min A \setminus R \\ a < r \in R \subseteq' A \end{cases} \implies \begin{cases} a \notin A \setminus R \\ a \in R \end{cases} \implies a \in R \quad (2.56)$$

So we have $R \cup \{f(R)\} \subseteq' A$. Similarly we have $R \cup \{f(R)\} \subseteq' B$. Because $f(R) = \min A \setminus R \notin R$, so we know $R \subset R \cup \{f(R)\} \subseteq R$ by definition of R CaC

Because $R \subseteq' A$ and $R \subseteq' B$. WOLG, we can let $R = A$, and we see $A = R \subseteq' B$ (done)

Let

$$U := \bigcup \mathcal{C} \tag{2.57}$$

We now prove $U \in \mathcal{C}$.

Notice that by Lemma 2.3.7, we know U is a well ordered chain. Because $x \in U \implies \exists D, x \in D \in \mathcal{C} \implies p \leq x$, so p is also the minimal element of U. Then now we only have to prove $T \subset' U \longrightarrow f(T) = \min U \setminus T$.

Arbitrarily find $T \subset' U$, and arbitrarily pick $s \in U \setminus T$. By definition of U, we know there exists $S \in \mathcal{C}$ such that $s \in S$.

We wish to prove $S \subseteq' U$. Arbitrarily pick $v \in U$ less than s and let $v \in V \in \mathcal{C}$. By violet part, we know $S \subseteq' V$ or $V \subseteq' S$. If $V \subseteq' S$, then we prove $v < s \implies v \in V \subseteq S$. If $S \subseteq' V$, then $v < s \implies v \in S$. Notice both of them mean $S \subseteq' U$, since v is arbitrarily picked from U (done) .

Because $T \subseteq' U$ and $S \subseteq' U$, by Lemma 2.3.8, we know either $S \subseteq' T$ or $T \subseteq' S$. Notice $s \in S$ and $s \in U \setminus T$, so we know $T \subset' S$. Then because $S \in \mathcal{C}$, we know $f(T) = \min S \setminus T$. Notice that it can be trivially proved $S \subseteq' U \implies S \setminus T \subseteq' U \setminus T$, so in fact we have $f(T) = \min S \setminus T = \min U \setminus T$ (done)

Assume U does not contain a maximal element of P. We now prove $U \cup \{f(U)\} \in \mathcal{C}$

Notice that by Lemma 2.3.6, we know $U \cup \{f(U)\}$ is a well ordered chain, and because $f(U)$ is an upper bound of U, we know $p = \min U \cup \{f(U)\}$, so we only have to prove $T \subset' U \cup \{f(U)\} \longrightarrow f(T) = \min U \cup \{f(U)\} \setminus T$.

Because $f(U)$ is an upper bound of U, we see $f(U) = \max U \cup \{f(U)\}$, and because U contain no maximal element, we know $f(U) \notin U$, that is, $U \subset U \cup \{f(U)\}$.

Notice we have $f(U) \notin T$, since $f(U) \in T \implies T = U \cup \{f(U)\}$. Then observe

$$\begin{aligned}
& f(U) \notin T \text{ and } T \subset' U \cup \{f(U)\} && (2.58)\\
\implies & T \subseteq' U && (2.59)\\
\implies & T = U \text{ or } T \subset' U && (2.60)\\
\implies & f(T) = f(U) = \min U \cup \{f(U)\} \setminus U = \min U \cup \{f(U)\} \setminus T && (2.61)\\
& \text{or } f(T) = \min U \setminus T = \min U \cup \{f(U)\} \setminus T \text{ (done)} && (2.62)
\end{aligned}$$

Then we see $U \cup \{f(U)\} \in \mathcal{C} \implies U \cup \{f(U)\} \subseteq \bigcup \mathcal{C} = U$ CaC ■

We now prove a simple application of Zorn's Lemma, showcasing something intuitive that in fact rely on Zorn's Lemma.

Theorem 2.3.11. (Zorn's Lemma Implies Every Vector Space Has a Basis) Every vector space has a basis.

Proof. Let V be a vector space and let $\mathcal{X}$ be the set of all linearly independent set of V. Order $\mathcal{X}$ by inclusion, we now prove every chain $\mathcal{C}$ in $\mathcal{X}$ is bounded above by $\bigcup \mathcal{C}$.

We first prove $\bigcup \mathcal{C}$ is linearly independent. Arbitrarily pick a finite subset S of $\bigcup \mathcal{C}$, and for each vector $s_i \in S$, pick a linearly independent set $S_i \in \mathcal{C}$ such that $s_i \in S_i$. Let S_n be the maximal element of $\{S_i\}$ by inclusion. Then we have $S \subseteq S_n$, so S is linearly independent (done)

By definition, we have $C \in \mathcal{C} \implies C \subseteq \bigcup \mathcal{C}$ (done) .

Then by Zorn's Lemma, we know there exist a maximal element W of $\mathcal{X}$. We now prove U spans V.

Assume there exists $v \in V \setminus U$. Then we see $U \subset U \cup \{v\}$ and see $U \cup \{v\} \in \mathcal{X}$ CaC ■

Theorem 2.3.12. (Zorn's Lemma Implies Axiom of Choice) If for every partially ordered set P such that every chain is bounded above, there exists a maximal element, then for each set of nonempty set Y, there exists a function $f : Y \to \bigcup Y$ such that $f(X) \in X$

2.4 Unfinished Notes for Cardinal

In this section, I first have to assert that there are a lot of abuse of notation. For people who are not familiar with set theory, just note that if everything have to be "down to the bottom" rigorous, instead of writing integer $0, 1, 2, 3$, we have to either use the notation $\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$, or explain the notation $0, 1, 2, 3$ in detail unrelated to our topic.

Definition 2.4.1. (Definition of Cardinality) Two sets A, B are said to *have the same cardinality* or *have the same cardinal number* or *of the same cardinality* if there exists a bijective function from A to B. If A, B have the same cardinality, then we write

$$|A| = |B| \tag{2.63}$$

Theorem 2.4.2. (Basic Property of Cardinality) For all sets A, B, we have

$$|A| = |A| \text{ (Reflexive)} \tag{2.64}$$

$$|A| = |B| \implies |B| = |A| \text{ (Symmetric)} \tag{2.65}$$

$$|A| = |B| \text{ and } |B| = |C| \implies |A| = |C| \text{ (Transitive)} \tag{2.66}$$

Proof. Please note that so far in our definition and notation, when we say $|A| = |B|$, we does not mean $|A| = x = |B|$ for some mathematical object x. When we say $|A| = |B|$, we merely mean that there exists a bijective function from A to B.

Define $f : A \to A$ by $x \mapsto x$, and we can use this function to prove the Reflexive part. For the rest two parts, use inverse and composition of functions. ■

Notice that at this point, the concept of cardinality is defined completely based on bijective function, which does not rely on the elements of sets.

We now give a notation related to cardinality, and introduce Schroder-Berstein Theorem, which plays an important role in our proofs for properties of cardinality.

Definition 2.4.3. (Definition of Cardinality) Given two sets A, B, if there exists an one-to-one function from A to B, we write

$$|A| \leq |B| \text{ or write } |B| \geq |A| \tag{2.67}$$

Theorem 2.4.4. (Basic Property of Cardinality) Let A, B be sets. We have

$$A \subseteq B \implies |A| \leq |B| \tag{2.68}$$

Proof. Define $f : A \to B$ by $x \mapsto x$ and we are done. ■

Theorem 2.4.5. (Schroder-Berstein Theorem) Let A, B be sets. We have

$$|A| \leq |B| \text{ and } |B| \leq |A| \iff |A| = |B| \tag{2.69}$$

Proof. Let $f : A \to B$, $g : B \to A$ be two one-to-one function.

Define $C_0 := A \setminus g(B)$, and for all nonnegative integer k, define

$$C_{k+1} := g(f(C_k)) \tag{2.70}$$

Also, Define

$$C := \bigcup_{k=0}^{\infty} C_k \tag{2.71}$$

Define

$$h(x) := \begin{cases} f(x) & x \in C \\ g^{-1}(x) & x \in A \setminus C \end{cases} \tag{2.72}$$

We first prove h is defined every where on A.

Because $C_0 \subseteq C$, we know that $A \setminus C \subseteq A \setminus C_0 = A \setminus (A \setminus g(B))$. Then we can deduce

$$x \in A \setminus C \implies x \in A \setminus (A \setminus g(B)) \tag{2.73}$$
$$\iff x \in A \text{ and } x \notin A \setminus g(B) \tag{2.74}$$
$$\iff x \in A \text{ and } (x \notin A \text{ or } x \in g(B)) \tag{2.75}$$
$$\iff (x \in A \text{ and } x \notin A) \text{ or } (x \in A \text{ and } x \in g(B)) \tag{2.76}$$
$$\iff x \in A \text{ and } x \in g(B) \tag{2.77}$$
$$\iff x \in g(B) \text{ (done)} \tag{2.78}$$

We now prove h is one-to-one.

Assume there exists $x \neq y \in A$ such that $h(x) = h(y)$. We know x, y are not both in C, otherwise f is not one-to-one. We also know x, y are not both in $A \setminus C$, otherwise g is not a function. WOLG, let $x \in C$ and $y \in A \setminus C$, so we have $f(x) = g^{-1}(y)$. Then, we have $y = g(f(x))$. Because $x \in C$, we know there exists a nonnegative integer n such that $x \in C_n$. Then we see $y \in g(f(C_n)) = C_{n+1}$ CaC (done)

Lastly we prove h is onto.

We wish to prove $B = h(A)$, where

$$h(A) = f(C) \cup g^{-1}(A \setminus C) \text{ and } h(A) \subseteq B \tag{2.79}$$

So we just prove

$$B \subseteq f(C) \cup g^{-1}(A \setminus C) \tag{2.80}$$

Let $x \in B$. If all x is in $g^{-1}(A \setminus C)$, then the proof is over. If not, keep in mind that $g(x) \notin C_0$, since $g(x) \in C_0 = A \setminus g(B)$ implies $g(x) \notin g(B)$ and observe

$$x \notin g^{-1}(A \setminus C) \tag{2.81}$$
$$\iff g(x) \notin A \setminus C \tag{2.82}$$
$$\iff g(x) \in C \tag{2.83}$$
$$\iff \exists n \in \mathbb{N}, g(x) \in C_n \tag{2.84}$$
$$\iff \exists n \in \mathbb{N}, \exists u \in C_{n-1}, g(x) = g(f(u)) \tag{2.85}$$
$$\iff \exists n \in \mathbb{N}, \exists u \in C_{n-1}, x = f(u) \tag{2.86}$$
$$\iff x \in f(C) \text{ (done)} \tag{2.87}$$

■

We now give another notation related to cardinality.

Definition 2.4.6. (Definition of Cardinality) Let A, B be sets. If $|A| \geq |B|$ and there does not exist a bijective function from A to B, which we denote $|A| \neq |B|$, then we write

$$|A| > |B| \text{ or write } |B| < |A| \tag{2.88}$$

Theorem 2.4.7. (Basic Property of Cardinality) Let A, B be two sets. We have

$$|A| > |B| \iff \text{no function from } A \text{ to } B \text{ is one-to-one} \tag{2.89}$$

$$|A| > |B| \iff \text{no function from } B \text{ to } A \text{ is onto} \tag{2.90}$$

Proof. $(\longrightarrow)$

$|A| > |B|$ implies that there exists an one-to-one function from B to A. If there exists a one-to-one function from A to B, then by Schroder-Berstein Theorem 2.1.5, we have a contradiction. If there exists an onto function $g : B \to A$, then with g^{-1} and Schroder-Berstein Theorem 2.1.5, we have a contradiction.

$(\longleftarrow)$

We prove the negate of $|A| > |B|$ is $|A| \leq |B|$ ■

Definition 2.4.8. (Definition of finite and infinite) For all nonnegative integer n, let

$$J_n = \{m \in \mathbb{N}, 0 < m \leq n\} \tag{2.91}$$

Let A be a set. If for some nonnegative integer n, we have $|A| = |J_n|$, then we say A is *finite*. A is *infinite* if A is not finite.

Theorem 2.4.9. (Basic Property of Finite Set) Let A, B be two finite sets, and let $|A| = |B|$. Let f be a function from A to B. We have

$$f \text{ is one-to-one} \iff f \text{ is bijective} \iff f \text{ is onto} \tag{2.92}$$

Theorem 2.4.10. (A coincidence ?) If we define naturals by $0 = \varnothing$ and $x+1 = x \cup \{x\}$, then for all nonnegative integer n, we have $|n| = J_n$

Theorem 2.4.11. ($\mathbb{N}$ is infinite) $\mathbb{N}$ is infinite.

Proof. Assume $\mathbb{N}$ is finite. Define $f : \mathbb{N} \to \mathbb{N}$ by $x \mapsto x + 1$. f is one-to-one implies that f is onto, but $\forall x \in \mathbb{N}, f(x) \neq 1$ CaC ■

Theorem 2.4.12. ($\mathbb{N}$ is the "smallest" infinite set) Every infinite set A have either same or greater cardinality than $\mathbb{N}$

Proof. ■

Theorem 2.4.13. (Equivalent to AC) Let A, B be sets. We have

$$|A| \leq |B| \text{ or } |A| \geq |B| \tag{2.93}$$

Proof. ■

Theorem 2.4.14. (Equivalent to AC) Let A, B be sets. We have

$$|A| > |B| \iff |A| \not\leq |B| \tag{2.94}$$

Proof. ■

2.5 Metric Space

Definition 2.5.1. (Definition of Metric Space) We say $(X, d : X^2 \to \mathbb{R})$ is a metric space if for all $x, y, z \in X$, we have

$$d(x, y) \geq 0 \text{ (Nonnegativity)} \tag{2.95}$$

$$x \neq y \longrightarrow d(x, y) = 0 \text{ (Positive Definitness)} \tag{2.96}$$

$$d(x, y) = d(y, x) \text{ (Commutative)} \tag{2.97}$$

$$d(x, z) \leq d(x, y) + d(y, z) \text{ (Triangle Inequality)} \tag{2.98}$$

Theorem 2.5.2. (Examples of Metric Space)

$$\mathbb{R}^n, d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| \tag{2.99}$$

$$\mathbb{R}^n, d(\mathbf{x}, \mathbf{y}) = \sum |x_i - y_i| \tag{2.100}$$

Proof. The first example is an Euclidean space. We notice $|\mathbf{x} - \mathbf{y}| = \sqrt{\sum (x_i - y_i)^2} \geq 0$ and $|\mathbf{x} - \mathbf{y}| = 0 \implies \sqrt{\sum (x_i - y_i)^2} = 0 \implies \sum (x_i - y_i)^2 = 0 \implies \forall i, x_i = y_i \implies \mathbf{x} = \mathbf{y}$. Notice $|\mathbf{x} - \mathbf{y}| = \sqrt{\sum}$ ■