

NCKU 112.1
Introduction to Analysis

Eric Liu

CONTENTS

Chapter 0

Basic Set Theory

0.1 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 0.1.1. (Atomic Formula) In our discussion in this section, there are precisely two atomic formulas:

$$x \in A \tag{1}$$

and

$$A = B \tag{2}$$

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 0.1.2. (Axiom of Extension)

$$\forall x(x \in A \longleftrightarrow x \in B) \longrightarrow A = B \quad (3)$$

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 0.1.3. (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 0.1.4. (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 0.1.5. (Axiom of Subset) Let P be a predicate. The axiom states

$$\forall A \exists B \forall x(x \in B \longleftrightarrow x \in A \wedge P(x)) \quad (4)$$

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 \emptyset is in our alphabet.

Axiom 0.1.6. (Axiom of Existence of Empty Set)

$$\exists \emptyset \forall y \neg (y \in \emptyset) \quad (5)$$

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 0.1.7. (Axiom of Pairing)

$$\forall x \forall y \exists A \forall v (v \in A \longleftrightarrow v = x \vee v = y) \quad (6)$$

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 0.1.8. (Axiom of Union)

$$\forall A \exists B \forall x (x \in B \longleftrightarrow \exists C (C \in A \wedge x \in C)) \quad (7)$$

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 0.1.9. (Axiom of Power Set)

$$\forall A \exists B \forall x (x \in B \longleftrightarrow \forall y (y \in x \longrightarrow y \in A)) \quad (8)$$

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 \emptyset 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 $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, but one can not construct the set of all natural numbers, that is, informally writing, $\mathbb{N} = \{0, 1, 2, 3, \dots\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \dots\}$.

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 0.1.10. (Axiom of Infinity)

$$\exists I (\emptyset \in I \wedge \forall y (y \in I \longrightarrow y \cup \{y\} \in I)) \quad (9)$$

Notice that every set interpreted as a natural number can be proved to belong to the set I in Axiom 0.1.10. 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 0.1.11. (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

$$0 \in U \text{ (First Peano axiom)} \quad (10)$$

$$a \in U \implies S(a) \in U \text{ (Second Peano axiom)} \quad (11)$$

$$\forall x \in U, S(x) \neq 0 \text{ (Third Peano axiom)} \quad (12)$$

$$S(a) = S(b) \implies a = b \text{ (Fourth Peano axiom)} \quad (13)$$

$$0 \in X \text{ and } \forall u \in U, u \in X \implies S(u) \in X \implies U \subseteq X \text{ (Induction axiom)} \quad (14)$$

Before reading the proof of the following theorem, notice that the set $\{y \in \bigcup x : \forall z(z \in x \implies y \in z)\}$ 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 0.1.12. (Existence of Model of Natural Numbers in ZF) We first define five predicate resembling the five Peano axioms.

$$P_1(x) := \emptyset \in x \quad (15)$$

$$P_2(x) := \forall a(a \in x \implies a \cup \{a\} \in x) \quad (16)$$

$$P_3(x) := \forall a(a \in x \implies a \cup \{a\} \neq \emptyset) \quad (17)$$

$$P_4(x) := \forall a \forall b(a \in x \wedge b \in x \wedge a \cup \{a\} = b \cup \{b\} \implies a = b) \quad (18)$$

$$P_5(x) := \forall a([\forall b(b \in a \implies b \in x) \wedge \emptyset \in a \wedge \forall b(b \in a \implies b \cup \{b\} \in a)] \implies x = a) \quad (19)$$

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

The theorem states:

$$\exists A(P_1(A) \wedge P_2(A) \wedge P_3(A) \wedge P_4(A) \wedge P_5(A)) \quad (20)$$

Proof. Let I be the set in Axiom 0.1.10. Define $\mathcal{L} := \{X \in \mathcal{P}(I) : \emptyset \in X \wedge \forall a(a \in X \longrightarrow a \cup \{a\} \in X)\}$. 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

$$P_1(Y) \wedge P_2(Y) \wedge P_3(Y) \wedge P_4(Y) \wedge P_5(Y) \quad (21)$$

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

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

We now prove $P_5(Y)$

Let $E \subseteq \bigcap \mathcal{L}$, $\emptyset \in E$ 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 $\{\emptyset\} \in S$, 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 $\{0, 1, 2, \dots; a, b\}$ 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 $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \dots$ 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 0.1.12 is relatively weak. This is not the case however. The fact is that, we don't really require anything extra of Theorem 0.1.12 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: $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \dots$ can each be formally proved to belong to $\bigcap \mathcal{L}$. We now introduce the last axiom.

Axiom 0.1.13. (Axiom of Foundation)

$$\forall X (X \neq \emptyset \longrightarrow \exists Y (Y \in X \wedge \forall z (z \in Y \longrightarrow \neg z \in X))) \quad (22)$$

In English, the axiom of foundation states that every nonempty set X has an element Y such that $X \cap Y = \emptyset$. 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 0.1.14. (No set belong to itself)

$$\forall x, \neg x \in x \quad (23)$$

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

Theorem 0.1.15. (No infinitely descending chain of sets) Informally, we can say that the following situation will never happen

$$\cdots \in x_3 \in x_2 \in x_1 \quad (24)$$

Formally, we say

$$\neg \exists X (\forall x (x \in X \longrightarrow \exists y (y \in X \wedge y \in x))) \quad (25)$$

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** . ■

0.2 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 first introduce axioms for ordered sets.

Axiom 0.2.1. (Axiom for Poset) Let \leq be a relation on set S . We say $\langle S, \leq \rangle$ is partially ordered if

$$\forall x \in S, x \leq x \text{ (Reflexive)} \quad (26)$$

$$\forall x, y, z \in S, x \leq y \text{ and } y \leq z \implies x \leq z \text{ (Transitive)} \quad (27)$$

$$\forall x, y \in S, x \leq y \text{ and } y \leq x \implies x = y \text{ (Antisymmetric)} \quad (28)$$

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

Axiom 0.2.2. (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

$$\forall x \in S, x \leq x \text{ (Reflexive)} \quad (29)$$

$$\forall x, y \in S, x \leq y \text{ or } y \leq x \text{ (Connected)} \quad (30)$$

$$\forall x, y, z \in S, x \leq y \text{ and } y \leq z \implies x \leq z \text{ (Transitive)} \quad (31)$$

$$\forall x, y \in S, x \leq y \text{ and } y \leq x \implies x = y \text{ (Antisymmetric)} \quad (32)$$

We now give the axioms for strictly ordered sets.

Axiom 0.2.3. (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

$$\forall x \in S, x \not< x \text{ (Irreflexive)} \quad (33)$$

$$\forall x, y \in S, x < y \implies y \not< x \text{ (Asymmetric)} \quad (34)$$

$$\forall x, y, z \in S, x < y \text{ and } y < z \implies x < z \text{ (Transitive)} \quad (35)$$

$$\forall x, y \in S, x \neq y \implies x < y \text{ or } y < x \text{ (Connected)} \quad (36)$$

$\langle S, < \rangle$ 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 0.2.4. (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 0.2.5. (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 0.2.6. (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:

$$\forall x \in S, x \not\in x \text{ (Irreflexive)} \quad (37)$$

$$\forall x, y \in S, x \in y \implies y \not\in x \text{ (Asymmetric)} \quad (38)$$

$$\forall x, y, z \in S, x \in y \text{ and } y \in z \implies x \in z \text{ (Transitive)} \quad (39)$$

$$\forall x, y \in S, x \neq y \implies x \in y \text{ or } y \in x \text{ (Connected)} \quad (40)$$

$$\forall V \subseteq S, \exists x \in V, \forall y \in V, x \in y \text{ (Well Order)} \quad (41)$$

$$\forall x \in S, x \subseteq S \text{ (Subset)} \quad (42)$$

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 antisymmetric is trivially satisfied by axiom of foundation.

Theorem 0.2.7. (Every Element of Ordinal is an Ordinal) If S be an ordinal, then $A \in S$ is an ordinal.

Proof. Notice that if $A = \emptyset$, then the proof is all trivial, so we only have to consider $A \neq \emptyset$. 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 0.2.8. (Ordinal Successor) If a is an ordinal, then $a \cup \{a\}$ is an ordinal.

Proof. Notice that $b \in a \cup \{a\} \implies b \in a \text{ or } b \in \{a\} \implies b \in a \text{ or } b = a$. 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, WOLOG, $y = a$, 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 is 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 0.2.9. (Basic Property of Ordinals) If a, b are ordinals, we have

$$a \in b \iff a \subset b \quad (43)$$

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 0.2.10. 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 0.2.11. (Trichotomy of Ordinals) If a, b are ordinals, we have

$$a \subset b \text{ or } a = b \text{ or } b \subset a \quad (44)$$

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

By Lemma 0.2.10, we know $a \cap b$ is an ordinal. Because $a \setminus b \neq \emptyset$, we know $a \cap b \subset a$, and then by Theorem 0.2.9 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 0.2.12. (Trichotomy of Ordinals) If a, b are ordinals, we have

$$a \in b \text{ or } a = b \text{ or } b \in a \quad (45)$$

Corollary 0.2.13. (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 0.2.12. ■

Theorem 0.2.14. (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 0.1.15 ■

Theorem 0.2.15. (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

$$\neg \exists A \forall x (x \in A \longleftrightarrow P(x)) \quad (46)$$

Proof. Assume Ω is the set of all ordinals. By Theorem 0.2.14, we see Ω satisfy at least irreflexive, asymmetric, transitive, connected and well ordering requirement of ordinal.

Let $x \in \Omega$, and $y \in x$. By Theorem 0.2.7, 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

$$0, 1, 2, 3, \dots; \omega, \omega + 1, \omega + 2, \dots \quad (47)$$

Theorem 0.2.16. (All Naturals are Ordinals) All element of $\bigcap \mathcal{L}$ are ordinals.

Proof. We prove by induction, which is the $P_5(\cap \mathcal{L})$ of Theorem 0.1.12.

Notice that \emptyset trivially satisfy all property of ordinals, which finish the proof for base step, and notice that by Theorem 0.2.8, the proof for induction step is also finished. ■

Theorem 0.2.17. (\mathbb{N} as a set is an ordinal) $\cap \mathcal{L}$ is an ordinal.

Proof. By Theorem 0.2.16, 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 $X := \{x \in \cap \mathcal{L} : x \subseteq \cap \mathcal{L}\}$. Base case: $\emptyset \in X$ is trivial. Let $a \in X$. Then we have $a \in \cap \mathcal{L}$ and $a \subseteq \cap \mathcal{L}$, so we have $a \cup \{a\} \subseteq \cap \mathcal{L}$ (done) ■

Theorem 0.2.18. (Transfinite Induction) Let X be a nonempty ordinal. For all $x \in X$, we define

$$S_x := \{z \in X : z \in x\} \quad (48)$$

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

$$\forall x \in X (S_x \subseteq Y \longrightarrow x \in Y) \quad (49)$$

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 $S_r = \emptyset \subseteq Y$. 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 . ■

0.3 Axiom of Choice and its equivalents

In this section, we aim to establish the equivalence between the Axiom of Choice and Zorn's Lemma. Additionally, we will explore and provide proofs for various applications of Zorn's Lemma. We'll begin by laying down foundational theorems in order theory, which will serve as our lemmas. Using these, we will demonstrate the progression from the Axiom of Choice to Zorn's Lemma. Interestingly, the reverse direction, proving from Zorn's Lemma to the Axiom of Choice, is considerably more straightforward and can be viewed as an "application" of Zorn's Lemma. Towards the end, we will delve into other notable applications of Zorn's Lemma.

Notice in this section, P is a non-empty partially ordered set (poset).

Definition 0.3.1. (Definition of Chain) We say $X \subseteq P$ is a chain if X is totally ordered.

Definition 0.3.2. (Definition of Well Ordered Set) Let $X \subseteq P$. We say X is well ordered only if X is totally ordered.

Definition 0.3.3. (Definition of Upper Bound) Let $X \subseteq P$. We say X is bounded above by t if

$$\forall x \in X, x \leq t \quad (50)$$

Definition 0.3.4. (Definition of Initial Segment) Let C be a chain of P . We say T is an initial segment of C if

$$T \subseteq C \text{ and } \forall v \in T, \forall u \in C, u \leq v \implies u \in T \quad (51)$$

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 0.3.5. An initial segment of a chain is also a chain

Lemma 0.3.6. 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 $V \neq \emptyset$. 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 0.3.7. If \mathcal{C} is a set of well ordered chains in P where

$$X, Y \in \mathcal{C} \implies X \text{ is an initial segment of } Y \text{ or } Y \text{ is an initial segment of } X \quad (52)$$

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

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

$$X, Y \in \mathcal{C} \implies X \subseteq Y \text{ or } Y \subseteq X \quad (53)$$

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 0.3.8. 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 0.3.7. 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 ■

While Lemma 0.3.7 might appear to be a potent assertion, implying that the union of a totally ordered (by inclusion) set of well-ordered subsets is itself well-ordered, this is far from the truth. For instance, consider the union of the sets $\{\{-1\}, \{-1, -2\}, \{-1, -2, -3\}, \dots\}$. This union is the set of negative integer, apparently not well-ordered, even though the set $\{\{-1\}, \{-1, -2\}, \dots\}$ is a set of well-ordered subsets of \mathbb{Z} when ordered by inclusion. This counterexample is even more compelling than our original premise.

Next, we'll move on to demonstrating the link from the Axiom of Choice to Zorn's Lemma. But before delving into that, let's first formulate the concept of a function within the ZF set theory.

Definition 0.3.9. (Construction of Function) In ZF, we interpret the set

$$\{\{a\}, \{a, b\}\} \quad (54)$$

as (a, b) and we interpret the set

$$\{(x, y) : x \in X, y \in Y\} \quad (55)$$

as a function from X to Y . Carefully notice we can construct predicate $Q(f)$ as below to tell if a set should be interpreted as a function.

$$\forall A(A \in f \longrightarrow \forall x(\{x\} \in A \longrightarrow \exists y H(A, x, y) \vee A = \{\{x\}\}) \wedge G(f, A, x)) \quad (56)$$

where predicate $H(A, x, y)$ is defined by

$$x \neq y \wedge \{x, y\} \in A \wedge \forall z(z \in A \longrightarrow z = \{x\} \vee z = \{x, y\}) \quad (57)$$

and predicate $G(f, A, x)$ is defined by

$$\forall B(B \in f \longrightarrow B = A \vee \{x\} \notin B) \quad (58)$$

Also, carefully notice we can construct predicate $Q_f(x, y)$

$$\exists A \in f(\{x\} \in A \wedge \{x, y\} \in A) \quad (59)$$

which if hold true should be interpreted as $f(x) = y$

Theorem 0.3.10. (Axiom of Choice Implies Zorn's Lemma) If

AC: For each set of nonempty sets Y , there exists a function $f : Y \rightarrow \bigcup Y$ such that $f(X) \in X$

Then

Zorn's Lemma: 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 h be a choice function on image of g . Then, we define $f = h \circ g$, so we have $\forall x \in C, x \leq f(C)$.

The above use the axiom of choice. We now define

$$\mathcal{C} := \{C \subseteq P : C \text{ is well ordered, } \min C = p, T \subset' C \longrightarrow f(T) = \min C \setminus T\} \quad (60)$$

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 0.3.7. Let $A, B \in \mathcal{C}$, and define

$$R := \bigcup \{T : T \subseteq' A \text{ and } T \subseteq' B\} \quad (61)$$

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

$$a < r' \implies \begin{cases} a < f(R) = \min A \setminus R \\ a < r \in R \subseteq' A \end{cases} \implies \begin{cases} a \notin A \setminus R \\ a \in R \end{cases} \implies a \in R \quad (62)$$

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

$$U := \bigcup \mathcal{C} \quad (63)$$

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

Notice that by Lemma 0.3.7, 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 0.3.8, 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 0.3.6, 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 $T \subset' U \cup \{f(U)\} \implies f(T) = \min U \cup \{f(U)\} \setminus T$.

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

$$f(U) \notin T \text{ and } T \subset' U \cup \{f(U)\} \quad (64)$$

$$\implies T \subseteq' U \quad (65)$$

$$\implies T = U \text{ or } T \subset' U \quad (66)$$

$$\implies f(T) = f(U) = \min U \cup \{f(U)\} \setminus U = \min U \cup \{f(U)\} \setminus T \quad (67)$$

$$\text{or } f(T) = \min U \setminus T = \min U \cup \{f(U)\} \setminus T \text{ (**done**)} \quad (68)$$

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

We will first delve into a straightforward application of Zorn's Lemma. Following that, we'll demonstrate how Zorn's Lemma inherently suggests the Axiom of Choice.

Theorem 0.3.11. (Zorn's Lemma Implies Every Vector Space Has a Basis) If

Zorn's Lemma: For every partially ordered set P such that every chain is bounded, there exists a maximal element.

then

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 $v \in V \setminus \text{span}(U)$. Then we see $U \subset U \cup \{v\}$ and see $U \cup \{v\} \in \mathcal{X}$ CaC (done) ■

In direct application of Zorn's Lemma, especially those involving chains ordered by inclusion, a common tactic is to use the union of a chain as an upper bound. Conceptually, this can be likened to computing the limit or supremum of a monotonically increasing sequence.

We now prove that Zorn's Lemma also implies axiom of choice, so we know Zorn's Lemma is equivalent to axiom of choice.

Theorem 0.3.12. (Zorn's Lemma Implies Axiom of Choice) If

Zorn's Lemma: For every partially ordered set P such that every chain is bounded above, there exists a maximal element

Then

AC: For each set of nonempty set Y , there exists a function $f : Y \rightarrow \bigcup Y$ such that $f(X) \in X$

Proof. Notice that we prove the theorem in the sense of ZF.

Let P be the set of all function f from subset of Y to $\bigcup Y$ such that $f(X) \in X$. In P , we define

$$g \leq h \text{ if } g \subseteq h \quad (69)$$

Let \mathcal{C} be a chain in P , we now prove $\bigcup \mathcal{C}$ is an upper bound of \mathcal{C} .

Let $A \in \bigcup \mathcal{C}$, so we know there exists $f \in \mathcal{C}$, such that $A \in f$, then we know A is indeed a set of the form $\{\{x\}, \{x, y\}\}$.

Let $B = \{\{x\}, \{x, y\}\} \in \bigcup \mathcal{C}$. Assume $\exists z, \exists C \in \mathcal{C}, C = \{\{x\}, \{x, z\}\} \in \bigcup \mathcal{C}$. Arbitrarily find B', C' such that $B \in B' \in \mathcal{C}$ and $C \in C' \in \mathcal{C}$. Then WOLG, let $B' \subseteq C'$, so we see $B \in C'$ then we see C' is not a function **CaC**. For now we have proven $\bigcup \mathcal{C} \in P$, then the fact that $\bigcup \mathcal{C}$ is an upper bound of \mathcal{C} trivially follows. (done)

Let f be a maximal element of P . Assume **Domain of f is a proper subset of Y** . Then we can arbitrarily select an element $X \in Y \setminus \text{Dom}(f)$, and arbitrarily select an element $x \in X$, then define $f' = f \cup \{\{X\}, \{X, x\}\}$ **CaC** ■

In the proof above, there are a couple of crucial observations to make:

Brevity of the Proof: The proof is notably more concise than proving that the Axiom of Choice (AC) implies Zorn's Lemma. This conciseness is a reason why some might prefer to take Zorn's Lemma as an axiomatic starting point, even though it's just an equivalent formulation of AC.

Choice from a Set: In the concluding section of the proof, we "select" an element from specific sets on two occasions. Beginners might assume that a choice function can only be defined if AC is assumed. However, this isn't the case for a finite collection of sets. One doesn't necessarily need the Axiom of Choice for this.

The real potency of the Axiom of Choice lies in its ability to address *situations that demand an infinite sequence of selections, where there's no predefined order or method to the selection process*.

In our proof, when we "select" an element, it's from a singular set. Consequently, the domain of this choice function is of cardinality 1. It's evident that the Axiom of Choice isn't necessary here.

To delve into why the Axiom of Choice isn't required for finite situations, consider attempting a formal construction of a choice function for the set $\{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$.

At this point, one may wonder, is it true that no construction of choice function on countable set of nonempty set require Axiom of Choice? This question in simple words want to ask: if a set X is countable, meaning I have already build a bijective function from \mathbb{N} to X , meaning I can well order X , and to select an element from one nonempty set is doable as above explained, why can't I recursively do this and have a choice function on X ? The answer is: To *recursively* do such, what we should do is to construct a function g that maps nonempty set to an element of itself, and recursively define the desired choice function by $f(x) = g(x)$, but one can see this fake "technique" is not only trivial, and more importantly, wrong, since without AC, we don't know if g exists, let alone using g to construct f .

Theorem 0.3.13. (Zorn's Lemma Implies Well Ordering Theorem) If

Zorn's Lemma: For every partially ordered set P such that every chain is bounded above, there exists a maximal element.

then

Well Ordering Principle: Every set X can be well ordered.

Proof. Define

$$P := \{(A_i, \leq_i) : A_i \subseteq X \text{ and } A_i \text{ is well ordered by } \leq_i\} \quad (70)$$

and we define a partial order \leq for P by

$$(A_i, \leq_i) \leq (A_j, \leq_j) \text{ if } A_i \subseteq' A_j \text{ with respect to } \leq_j \text{ and } x_i \leq_i y_i \longrightarrow x_i \leq_j y_i \quad (71)$$

Let \mathcal{C}_d be a chain of P . Define order \leq_d on $\bigcup \mathcal{C}_d$ by

$$x \leq_d y \text{ if } \exists A_i \in \mathcal{C}_d, x, y \in A_i \text{ and } x \leq_i y \quad (72)$$

we first prove \leq_d is well defined, which is just to prove that \leq_d is antisymmetric. Let $x \leq_d y$ and $y \leq_d x$, and let $x, y \in A_i \cup A_j$ such that $x \leq_i y$ and $y \leq_j x$. Because \mathcal{C} is a chain, WOLG, let $A_i \leq A_j$, we then have $x \leq_j y$ and $y \leq_j x$, so $x = y$ (done)

We now prove \leq_d is reflexive, connected and transitive.

For reflexive, notice $x \in \bigcup \mathcal{C}_d$ implies $\exists A_i \in \mathcal{C}_d, x \in A_i$, and the rest is trivial. For connected, for $x, y \in \bigcup \mathcal{C}_d$, let $x \in A_i \in \mathcal{C}_d$ and $y \in A_j \in \mathcal{C}_d$. Because \mathcal{C} is a chain, WOLG, we can let $A_i \leq A_j$, and then we see $x, y \in A_j$, so because A_j is totally ordered by \leq_j , we know x, y is comparable by \leq_d . For transitive, let $x \leq_d y \leq_d z$, and let $x, y \in A_i \in \mathcal{C}_d$ and $x \leq_i y$ and $y, z \in A_j \in \mathcal{C}_d$ and $y \leq_j z$. If $A_j \leq A_i$, we see $x \leq_i y \leq_i z$, then the proof is finished. If $A_i \leq A_j$, we see $x \leq_j y \leq_j z$ (done).

We now prove $\bigcup \mathcal{C}_d \in P$, which is just to prove $\bigcup \mathcal{C}_d$ is well ordered by \leq_d .

Let $V \subseteq \bigcup \mathcal{C}_d$, and arbitrarily pick $A_i \in \mathcal{C}_d$ such that $A_i \cap V$ is nonempty, because A_i is well ordered by \leq_i , we know there exists a minimal element v' of $A_i \cap V$ with respect to \leq_i . Let $v \in V$, if $v \in A_i \cap V$, then by definition of \leq_d , we see $v' \leq_d v$. If $v \in V \setminus A_i$, then for each $A_j \in \mathcal{C}$ containing v , we know $A_i \subset A_j$, and because \mathcal{C} is a chain, we know $A_i \leq A_j$. Assume $v \not\leq_j v'$, because $v' \in A_i \subseteq' A_j$, we can deduce $v' \in A_i$ CaC, so we have $v' <_j v$, which implies $v' <_d v$. We have proved v' is the smallest element of V (done)

We now prove $\bigcup \mathcal{C}_d$ is an upper bound of \mathcal{C}_d with respect to \leq . Arbitrarily pick $x_i, y_i \in A_i \in \mathcal{C}_d$, we immediately see $x_i \leq_i y_i \implies x_i \leq_d y_i$, so we only have to prove $A_i \subseteq' \bigcup \mathcal{C}_d$ with respect to \leq_d . Let $v_i \in A_i, v_j \in \bigcup \mathcal{C}_d$ and $v_j \leq_d v_i$. Arbitrarily pick A_j containing v_j and Assume $v_j \notin A_i$, so we have $A_i < A_j$. Because $v_j \leq_d v_i$ and $v_i, v_j \in A_j$, we know $v_j \leq_j v_i$. Then because $A_i \subseteq' A_j$ with respect to \leq_j , we have $v_j \in A_i$ CaC. We have proved $v_j \in A_i$ by rejecting the red assumption. Notice the premise. We have proved $A_i \subseteq' \bigcup \mathcal{C}_d$ with respect to \leq_d (done)

Let (A_k, \leq_k) be a maximal element of P . Assume $A_k \subset X$, we can arbitrarily pick $x \in X \setminus A_k$, and for all $a_k \in A_k$ define $a_k \leq_k x$. By Lemma 0.3.6, we see $(A_k \cup \{x\}, \leq_k)$ is well ordered, so $(A_k \cup \{x\}, \leq_k) \in P$. To prove $(A_k, \leq_k) < (A_k \cup \{x\}, \leq_k)$ is trivial. CaC ■

0.4 Cardinality and Cardinal

In this section, we first present a definition and a theorem as a primer rather than the primary focus. This serves as a foundation to instill confidence in the reader regarding the use of set theory language.

Definition 0.4.1. (Formulation of Property of Function) We use predicate $P(f)$

$$\forall A \forall B (A \in f \wedge B \in f \wedge \exists x (\{x\} \in A \wedge \{x\} \in B) \longrightarrow A = B) \quad (73)$$

to tell if f is injective, and we use predicate $P(f, Y)$

$$\forall y \exists A \exists B ((A = \{\{y\}\} \wedge A \in f) \vee (A \in f \wedge B \in A \wedge y \in B \wedge \exists x (x \in B \wedge x \neq y))) \quad (74)$$

Theorem 0.4.2. (Basic Property of Function) There exists a one-to-one function from A to B if and only if there exists an onto function from B to A

While it's straightforward to "prove" the theorem, formalization of such proof offers a robust assurance of its correctness. I encourage you to attempt formalization. With that, let's delve into the topic of cardinality, and introduce a highly useful theorem.

Definition 0.4.3. (Definition of Cardinality) We write

$$A \leq_c B \text{ or write } B \geq_c A \quad (75)$$

If there exists a one-to-one function from A to B , or, equivalently, if there exists an onto function from B to A . Also, we write

$$A =_c B \quad (76)$$

if there exists a bijective function from A to B (one-to-one correspondence between A, B)

Theorem 0.4.4. (Schroder-Berstein Theorem) Let A, B be sets. We have

$$A \leq_c B \text{ and } B \leq_c A \iff A =_c B \quad (77)$$

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

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

$$C_{k+1} := g(f(C_k)) \quad (78)$$

Also, Define

$$C := \bigcup_{k=0}^{\infty} C_k \quad (79)$$

Define

$$h(x) := \begin{cases} f(x) & x \in C \\ g^{-1}(x) & x \in A \setminus C \end{cases} \quad (80)$$

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

$$x \in A \setminus C \implies x \in A \setminus (A \setminus g(B)) \quad (81)$$

$$\iff x \in A \text{ and } x \notin A \setminus g(B) \quad (82)$$

$$\iff x \in A \text{ and } (x \notin A \text{ or } x \in g(B)) \quad (83)$$

$$\iff (x \in A \text{ and } x \notin A) \text{ or } (x \in A \text{ and } x \in g(B)) \quad (84)$$

$$\iff x \in A \text{ and } x \in g(B) \quad (85)$$

$$\iff x \in g(B) \text{ (done)} \quad (86)$$

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

$$h(A) = f(C) \cup g^{-1}(A \setminus C) \text{ and } h(A) \subseteq B \quad (87)$$

So we just prove

$$B \subseteq f(C) \cup g^{-1}(A \setminus C) \quad (88)$$

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

$$x \notin g^{-1}(A \setminus C) \quad (89)$$

$$\iff g(x) \notin A \setminus C \quad (90)$$

$$\iff g(x) \in C \quad (91)$$

$$\iff \exists n \in \mathbb{N}, g(x) \in C_n \quad (92)$$

$$\iff \exists n \in \mathbb{N}, \exists u \in C_{n-1}, g(x) = g(f(u)) \quad (93)$$

$$\iff \exists n \in \mathbb{N}, \exists u \in C_{n-1}, x = f(u) \quad (94)$$

$$\iff x \in f(C) \text{ (done)} \quad (95)$$

Schroder-Berstein Theorem in simple word is just that the relation \leq_c between sets is antisymmetric. One may intuitively think such "simple" consequence should be easy to prove but in fact isn't.

That the relation \leq_c is reflexive and transitive is trivial to prove. What being said, this does not implies \leq_c is a "total order", since we haven't proved it is connected.

Theorem 0.4.5. (Cardinality of All Set are Comparable if Given AC) If given AC, for all sets A, B , we have

$$A \leq_c B \text{ or } B \leq_c A \quad (96)$$

Proof. Let P be the set of one-to-one function from subset of A to B . We define order relation on P by

$$f \leq g \text{ if } \forall x \in \text{Dom}(f), f(x) = g(x) \text{ and } \text{Dom}(f) \subseteq \text{Dom}(g) \quad (97)$$

To prove this do define a order relation is trivial. By Zorn's Lemma, we can pick a maximal element of f . If Domain of f is not A and range of f is not B , then we can define a larger function by arbitrarily picking $x \in A \setminus \text{Dom}(f)$ and $y \in B \setminus \text{Im}(f)$ and defining

$$f'(z) = \begin{cases} f(z) & \text{if } z \in \text{Dom}(f) \\ y & \text{if } z = x \end{cases}. \text{ It is trivial to prove } f < f'. \quad \blacksquare$$

So far, we have proved \leq_c is a total order on the class of all sets. Now, we give a necessary and sufficient condition of one set having strictly less cardinality than that of the other.

Definition 0.4.6. (Definition of Cardinality) Let A, B be sets. If $A \geq_c B$ and there does not exist a one-to-one correspondence between A and B , which we denote $A \neq_c B$, then we write

$$A >_c B \text{ or write } B <_c A \quad (98)$$

Theorem 0.4.7. (Basic Property of Cardinality) For all sets A, B , we have

$$A >_c B \iff \text{no function from } A \text{ to } B \text{ is one-to-one} \quad (99)$$

$$A >_c B \iff \text{no function from } B \text{ to } A \text{ is onto} \quad (100)$$

Proof. The fact that an one-to-one function from A to B exists if and only if an onto function from B to A exists is proved in the beginning of the section, so we only have to prove the first iff and the proof is finished. From left to right, Assume **there exists an one-to-one function from A to B** . Because $A >_c B \implies A \geq_c B \implies$ there exists another one-to-one function from B to A , so by Schroder-Berstein Theorem, we have $A =_c B$ **CaC**.

From right to left, Assume **the negate of $A >_c B$** . Notice the negate of $A >_c B$ is no function from B to A is one-to-one or $A =_c B$. If $A =_c B$, then there exists a bijective, so an one-to-one function from A to B . If no function from B to A is one-to-one, and no function from A to B is one-to-one, then the cardinality of A, B is incomparable **CaC** ■

Notice that we haven't define the cardinality in a "constructive" manner. We defined cardinality by defining relation between sets. We now define cardinal, not cardinality, in a constructive manner compatible to our old definition.

Definition 0.4.8. (Von Neumann Cardinal Assignment) We now assign each set X to a cardinal

$$|X| = \min\{\sigma \in \text{Ord} : \text{and } \sigma =_c X\} \quad (101)$$

And we say an ordinal is a cardinal if the ordinal if an caridnality of some sets.

Theorem 0.4.9. (Every Set Can be Assigned to a Cardinal if Given AC, Needed to be formalized) For arbitrary set X , there exists an ordinal of same cardinality. More precisely, every well order set (some set can not be well ordered without AC) can be assigned a cardinal.

Proof. By AC, we can well order X . Define function f that map ordinal to element in X by

$$f(0) := \min X \text{ and } f(\alpha) := \min X \setminus \{f(\beta) : \beta < \alpha\} \quad (102)$$

Assume $\forall \alpha \in \text{Ord}, f[\alpha] \subset X$. Then we know $\text{Ord} = \text{Dom}(f)$, and then we can formally construct a inverse function f^{-1} , and formally construct the set that is the image of f^{-1} , which we know is Ord . **CaC** ■

Theorem 0.4.10. (Compatibility between Von Neumann Cardinal and Ordinal) For all set A, B , we have

$$A <_c B \implies |A| \in |B| \text{ and } A =_c B \implies |A| = |B| \text{ and } n \in \bigcap \mathcal{L} \implies |n| = n \quad (103)$$

where infimum is with respect to strict well order \in on the class of ordinals.

Proof. Assume $|B| \in |A|$. Notice that trivially we can prove the relation $=_c$ on the class of all sets is symmetric, so we have $A <_c B$, since if $A =_c B$, we should by definition have $|B| = |A|$. Notice $|B| \in |A| \implies |B| \subset |A|$, so we have an one-to-one function from $|B|$ to $|A|$ by mapping each element of B to itself, then we have $|B| \leq_c |A|$, notice that by definition we have $B =_c |B| \leq_c |A| =_c A$. Then we can construct function by composition to show $B \leq_c A$ **CaC**.

We prove by induction. Base case is simple: notice $\emptyset =_c \emptyset$ implies $\emptyset \in \{\sigma \in \text{Ord} : \sigma =_c \emptyset\}$, so we have $|\emptyset| = \emptyset$. For induction case, let $k \in \bigcap \mathcal{L}$ and $|k| = k$, we wish to

prove $k \cup \{k\} = \min\{\sigma \in \text{Ord} : \sigma =_c k \cup \{k\}\}$. Trivial, we have $k \cup \{k\} \in \{\sigma \in \text{Ord} : \sigma =_c k \cup \{k\}\}$. We first prove k is the greatest ordinal in $k \cup \{k\}$. Let $m \in k \cup \{k\}$. Assume $k \in m$. Then by definition of ordinal we have $k \subseteq m$, and because $k \in m$, we then have $k \cup \{k\} \subseteq m$. On the other side, notice by definition of ordinal we have $m \subseteq k \cup \{k\}$, so in summary we have $m = k \cup \{k\}$ CaC (done). So, we see if $k \cup \{k\}$ is not the smallest element of $\{\sigma : \sigma =_c k \cup \{k\}\}$, then k belong to $\{\sigma : \sigma =_c k \cup \{k\}\}$. We now prove $k \neq_c k \cup \{k\}$ by induction. Base case is trivial: $\emptyset \neq_c \{\emptyset\}$. We now prove $k \neq_c k \cup \{k\} \longrightarrow k \cup \{k\} \neq_c k \cup \{k\} \cup \{k \cup \{k\}\}$. Assume there exists a bijective function f from $k \cup \{k\}$ to $k \cup \{k\} \cup \{k \cup \{k\}\}$. Then we can define a bijective function $g(x) : k \rightarrow k \cup \{k\}$ by
$$\begin{cases} f(k) & \text{if } x = f^{-1}(\{k \cup \{k\}\}) \\ f(x) & \text{if } x \neq f^{-1}(\{k \cup \{k\}\}) \end{cases} \text{ CaC (done)} \quad \blacksquare$$

The theorem discussed above is fundamental to the theory of ordinals and cardinals. However, it is infrequently employed in the discussions of real analysis. This is because, using the elementary methods of set description, the sets under consideration often possess cardinalities equivalent to either \mathbb{N} , \mathbb{R} , or are finite.

We now present the standard definition of *infinite* in the language of set theory, a definition widely accepted in most analysis texts. Subsequently, we introduce the concept of *Dedekind-infinite*. This definition harnesses our intuition: that shifting every term in an infinite sequence is permissible. It's important to note that the traditional definition of *infinite* and that of *Dedekind-infinite* are not always equivalent. However, if the Axiom of Choice (AC) holds true, then these definitions become equivalent. Since most people in the mathematical community accept AC, the two definitions are often treated as equivalent in practice.

Definition 0.4.11. (Definition of Infinite) We say a set A is finite if there exists an ordinal σ in $\bigcap \mathcal{L}$ such that $|A| = \sigma$, and we say A is infinite if A is not finite.

Definition 0.4.12. (Definition of Dedekind Infinite) Let A be a set. We say A is Dedekind-infinite if there exists a proper subset of A that have the same cardinality as A , and we say A is Dedekind-finite if A is not Dedekind-infinite.

Theorem 0.4.13. (Equivalence of Infinite and Dedekind Infinite under AC) Under AC, a set A is infinite if and only if A is Dedekind-infinite.

Proof. We first prove A is finite $\implies A$ is Dedekind-finite by induction. Base case is trivial, since \emptyset doesn't even have proper subsets. For induction case, let $k \in \bigcap \mathcal{L}$ be Dedekind-finite. Assume $\exists m \subset k \cup \{k\}, m =_c k \cup \{k\}$. If m contain k , then we have $m \setminus \{k\} \subseteq k$ and $m \setminus \{k\} =_c k$, where we use the same technique in proof of Theorem 0.4.10. So we know $m \subseteq k$. Then we have $k \cup \{k\} = m \leq_k k$ CaC (done).

We now prove A is infinite $\implies A$ is Dedekind-infinite. Define function

$$f : \mathbb{N} \rightarrow \mathcal{P}(\mathcal{P}(A)), n \mapsto \{X : \mathcal{P}(A) : |X| = n\} \quad (104)$$

We first prove for all naturals n , we have $f(n)$ is nonempty by induction.

$f(1)$ is nonempty since we can arbitrarily select an element and use axiom of pairing to construct a singleton subset. If $f(k)$ is nonempty, we can arbitrarily select an element X from $f(k)$ and to see if $A = X$ then A is finite, so we can arbitrarily select an element from $A \setminus X$ and attach to X to show $f(k+1)$ is nonempty. (done)

We now construct a bijective from A to a proper subset missing one element by constructing a countable subset U of A and map each element in U to its successor.

By AC, we know there exists a function $g(n)$ such that $\forall n \in \mathbb{N}, g(n) \in f(n)$. Let $U = \bigcup g[\mathbb{N}]$. Define $h : U \rightarrow \mathbb{N}$ by mapping elements in $g(n)$ to elements in $\{x \in \mathbb{N} : \frac{(n-1)n}{2} < x \leq \frac{n(n+1)}{2}\}$. h is onto, since $U = \bigcup g[\mathbb{N}]$, and because $|g(n)| = n = |x \in \mathbb{N} : \frac{(n-1)n}{2} < x \leq \frac{n(n+1)}{2}|$, we know h is one-to-one. Now we define function $B : A \rightarrow A \setminus \{h^{-1}(1)\}$, where y is the element in U that was mapped to 1 by

$$B(x) = \begin{cases} x & \text{if } x \notin U \\ h^{-1}(h(x) + 1) & \text{if } x \in U \end{cases} \quad (105)$$

■

Corollary 0.4.14. (\mathbb{N} is the "smallest" infinite set) Every infinite set have either same or greater cardinality than \mathbb{N}

Proof. Consider the set U in proof above.

■

Corollary 0.4.15. (Another Definition of Infinity) A set is infinite if and only if it has cardinality equal to or greater than \mathbb{N}

Now, we will demonstrate a few theorems specific to countable cardinalities. While these can be generalized to encompass all cardinalities using more advanced set theory tools, our final theorem will be applicable to all cardinalities without exception.

Definition 0.4.16. (Definition of Countable) We say a set is countable if the set has cardinality less than or equal to \mathbb{N}

Theorem 0.4.17. (Finite Cartesian Product of Countable Set is Countable) For all naturals n , we have $|\mathbb{N}^n| = |\mathbb{N}|$

Proof. Because we can map $p \mapsto (p, 1, \dots, 1)$, we know $|\mathbb{N}| \leq |\mathbb{N}^n|$. Because we know $|X_m := \{(p_1, \dots, p_n) \in \mathbb{N}^n : \sum p_i = m\}| = H_{m-n}^n$, we can map each element in X_m one-to-one into $(\sum_{i=3}^{m-1} H_{i-n}^n, \sum_{i=3}^m H_{i-n}^n]$, so we know $|\mathbb{N}^n| \leq |\mathbb{N}|$ ■

Theorem 0.4.18. (Countable Union of Countable Set is Countable) Let $S = \{E_n : n \in \mathbb{N}\}$ be a countable set of countable sets. We have $\bigcup S$ is countable.

Proof. For each $n \in \mathbb{N}$, because E_n is countable, we can define bijective $f_n : \mathbb{N} \rightarrow E_n$. Define $g : \mathbb{N} \rightarrow \bigcup S$ by

$$g(x) = f_{m-(h-2)}(h) \text{ where } m = \max\{n \in \mathbb{N} : \frac{(n+1)n}{2} < x\} \text{ and } h = x - \frac{(m+1)m}{2} \quad (106)$$

We now prove g is onto.

Let $x \in \bigcup S$, we know $\exists p \in \mathbb{N}, x \in E_p$, so we know $\exists p, q \in \mathbb{N}, x = f_p(q)$. One can verify $g(q + \frac{(p+q-2)^2 + (p+q-2)}{2}) = f_p(q) = x$ where $m = p + q - 2$ and $n = q$ (done) ■

Theorem 0.4.19. (Every Nonempty Set Has Cardinality Strictly Less than Its Power Set) If A is nonempty, then $|A| < |\mathcal{P}(A)|$.

Proof. Let f be a function from A to $\mathcal{P}(A)$. Define $T = \{a \in A : a \notin f(a)\}$. Observe $b \in T \implies b \notin f(b) \implies T \neq f(b)$ and observe $b \notin T \implies b \in f(b) \implies T \neq f(b)$, so $T \in \mathcal{P}(A) \setminus f[A]$. ■

0.5 Ordinal Arithmetic

0.6 Cardinal Arithmetic

0.7 Set Theoretic Limit

Definition 0.7.1. (Partial Order on Set) For sets A, B , we say

$$A \leq B \quad (107)$$

if

$$A \subseteq B \quad (108)$$

The class of sets form a partial order.

Theorem 0.7.2. (Upper and Lower Bound) Let σ be a collection of sets

$$\sigma = \{A_i : i \in I\} \quad (109)$$

The collection of lower bounds of σ is exactly

$$\mathcal{P}(\bigcap \sigma) \quad (110)$$

and

$$A \text{ is an upper bound of } \sigma \iff \bigcup \sigma \subseteq A \quad (111)$$

Definition 0.7.3. (Infimum and Supremum For Set) For collection σ of sets, we have

$$\inf \sigma = \bigcap \sigma \text{ and } \sup \sigma = \bigcup \sigma \quad (112)$$

Theorem 0.7.4. (Algebraic Interpretation of Notation for Infimum and Supremum)

$$\inf \sigma \text{ is the greatest set smaller than all sets in } \sigma \quad (113)$$

$$\sup \sigma \text{ is the smallest set greater than all sets in } \sigma \quad (114)$$

Above introduce set-theoretic infimum and supremum

Definition 0.7.5. (Lim-Sup for Sets) Let $\{A_n\}$ be a sequence of sets. We define

$$\liminf_{n \rightarrow \infty} A_n := \bigcup_{n \geq 1} \bigcap_{j \geq n} A_j \text{ and } \limsup_{n \rightarrow \infty} A_n := \bigcap_{n \geq 1} \bigcup_{j \geq n} A_j \quad (115)$$

Theorem 0.7.6. (Meaning Behind Lim-Sup) We have

$$x \in \liminf_{n \rightarrow \infty} A_n \iff \exists n, \forall j \geq n, x \in A_j \quad (116)$$

$$x \in \limsup_{n \rightarrow \infty} A_n \iff \forall n, \exists j \geq n, x \in A_j \quad (117)$$

Theorem 0.7.7. (Meaning Behind Lim-Sup)

$$\liminf_{n \rightarrow \infty} A_n = \bigcup_{n \geq 1} \inf\{A_j\}_{j=n}^{\infty} = \sup\{\inf\{A_j\}_{j=n}^{\infty}\}_{n=1}^{\infty} \quad (118)$$

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{n \geq 1} \sup\{A_j\}_{j=n}^{\infty} = \inf\{\sup\{A_j\}_{j=n}^{\infty}\}_{n=1}^{\infty} \quad (119)$$

Corollary 0.7.8. (Verification of Theorem 0.7.6) We have

$$x \in \liminf_{n \rightarrow \infty} A_n \iff \exists n, \forall j \geq n, x \in A_j \quad (120)$$

$$x \in \limsup_{n \rightarrow \infty} A_n \iff \forall n, \exists j \geq n, x \in A_j \quad (121)$$

And

$$X \subseteq \liminf_{n \rightarrow \infty} A_n \iff \exists n, \forall j \geq n, X \subseteq A_j \quad (122)$$

$$(123)$$

Proof. Suppose

$$x \in \liminf_{n \rightarrow \infty} A_n = \sup\{\inf\{A_j\}_{j=n}^{\infty}\}_{n=1}^{\infty} \quad (124)$$

By definition of supremum, we know $x \in \inf\{A_j\}_{j=n}^{\infty}$ for some n , then the theorem follows.

Suppose

$$x \in \limsup_{n \rightarrow \infty} A_n = \inf\{\sup\{A_j\}_{j=n}^{\infty}\}_{n=1}^{\infty} \quad (125)$$

By definition of infimum, we know $x \in \sup\{A_j\}_{j=n}^{\infty}$ for all n , then the theorem follows. ■**Corollary 0.7.9. (Manipulation of Lim-Sup)**

$$\inf\{A_n\} \subseteq \liminf_{n \rightarrow \infty} A_n \subseteq \limsup_{n \rightarrow \infty} A_n \subseteq \sup\{A_n\} \quad (126)$$

Proof. The first and the third inequality is easy to see. One can verify them using a notation that is more intuitive in this case

$$\inf\{A_n\} \in \{\inf\{A_j : j \geq n\} : n \geq 1\} \implies \inf\{A_n\} \subseteq \sup\{\inf\{A_j : j \geq n\} : n \geq 1\} \quad (127)$$

$$\sup\{A_n\} \in \{\sup\{A_j : j \geq n\} : n \geq 1\} \implies \sup\{A_n\} \subseteq \inf\{\sup\{A_j : j \geq n\} : n \geq 1\} \quad (128)$$

For the second inequality, we can use Theorem 0.7.6. Let $x \in \liminf_{n \rightarrow \infty} A_n$. We know there exists a threshold m such that $x \in A_j$ for every $j \geq m$. Then for each n , we can pick a number j greater than $\max\{m, n\}$, and observe $x \in A_j$. ■

Discuss when the sequence is monotonic

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

$$\max X \in X \text{ and } \forall x \in X, x \leq \max X \quad (1.1)$$

$$\min X \in X \text{ and } \forall x \in X, \min X \leq x \quad (1.2)$$

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

$$p^2 < 2 \implies p < q \text{ and } q^2 < 2 \quad (1.3)$$

$$p^2 > 2 \implies q < p \text{ and } 2 < q^2 \quad (1.4)$$

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 .

$$q - p = \frac{2 - p^2}{a} \quad (1.5)$$

$$q^2 - 2 = \frac{p^2 - 2}{b} \quad (1.6)$$

Now we solve for a, b

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

$$\iff \left[\frac{2 - p^2}{a} + p \right]^2 = \frac{p^2 - 2}{b} + 2 \quad (1.8)$$

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

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

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

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

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

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

$$\forall x \in S, x \sim x \quad (1.14)$$

$$\forall x, y \in S, x \sim y \text{ or } y \sim x \quad (1.15)$$

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

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

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

$$\exists a \in S, \forall b \in E, a \geq b \quad (1.18)$$

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

$$\forall b \in E, c \leq b \quad (1.19)$$

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

$$\sup E := \min A \quad (1.20)$$

and

$$\inf E := \max B \quad (1.21)$$

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

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

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 0.1.3, 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

$$E \subseteq S \text{ and } E \neq \emptyset \implies \sup E \text{ exists} \quad (1.23)$$

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

$$y < z \implies x + y < x + z \quad (1.24)$$

$$x > 0 \text{ and } y > 0 \implies xy > 0 \quad (1.25)$$

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

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

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

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

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

$$\begin{cases} a^x < a^y & \iff a > 1 \\ a^x = a^y & \iff a = 1 \\ a^x > a^y & \iff 0 < a < 1 \end{cases} \quad (1.28)$$

Proof. By Theorem 0.2.4, observe $1 < a \iff 1 < a < a^2 \iff 1 < a < a^2 < a^3 \iff \dots \iff 1 < a < \dots < a^{y-x} \iff a^x < a^y$. If $a = 1$, we know $a^x = 1 = a^y$. Again by Theorem 0.2.4, observe $0 < a < 1 \iff a^2 < a < 1 \iff \dots \iff a^{y-x} < \dots < 1 \iff a^y < a^x$. ■

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

$$\begin{cases} b^z < c^z & \iff 0 < z \\ b^z = c^z & \iff 0 = z \\ b^z > c^z & \iff 0 > z \end{cases} \quad (1.29)$$

Proof. If $z > 0$, then by Theorem 0.2.4, observe $0 < b < c \implies 0 < b^2 < bc < c^2 \implies b^3 < b^2c < bc^2 < c^3 \implies \dots \implies b^z < c^z$. 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 0.2.1 and Theorem 0.2.7. The latter result is a direct consequence of Theorem 0.2.3 and Theorem 0.2.5 ■

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

$$x^{p+q} = x^p x^q \quad (1.30)$$

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

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

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

$$x^p y^p = (xy)^p \quad (1.32)$$

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 $S = \emptyset$ 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, $S = \emptyset$ 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 natural 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 Theorem 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 $(0 < m^n < y \implies m < b)$. In the followings, the domain of a and b are only positives.

First we construct a . By Theorem 0.2.9, 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 0.2.10, 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 0.2.9, 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

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

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 $E \iff (m^n < y \implies m < x - k) \iff (m \geq x - k \implies m^n \geq y) \iff (m \geq x - k \implies x^n - m^n \leq x^n - y)$.

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 $a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b + \dots + b^{n-1})$, 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 $m \geq x - k \implies x^n - m^n \leq x^n - y$, 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

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

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

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

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

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

Proof. Observe

$$x = (a^m)^{\frac{1}{n}} \quad (1.37)$$

$$\iff x^n = a^m \quad (1.38)$$

$$\iff (x^n)^q = (a^m)^q \quad (1.39)$$

$$\iff x^{nq} = a^{mq} = a^{np} \text{ (because } n, m, q \in \mathbb{Z}) \quad (1.40)$$

$$\iff (x^q)^n = (a^p)^n \text{ (again, because } n, p, q \in \mathbb{Z}) \quad (1.41)$$

$$\iff x^q = a^p \text{ (by Theorem 1.3.6)} \quad (1.42)$$

$$\iff x = (a^p)^{\frac{1}{q}} \text{ (by Theorem 1.3.6)} \quad (1.43)$$

$$\iff (a^m)^{\frac{1}{n}} = x = (a^p)^{\frac{1}{q}} \quad (1.44)$$

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

$$\begin{cases} a^x < a^y & \iff a > 1 \\ a^x = a^y & \iff a = 1 \\ a^x > a^y & \iff 0 < a < 1 \end{cases} \quad (1.45)$$

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

$$a^x < a^y \implies a^{\frac{q}{p}} < a^{\frac{n}{m}} \quad (1.46)$$

$$\implies a^q = (a^{\frac{q}{p}})^p < (a^{\frac{n}{m}})^p \text{ (by Theorem 0.2.10)} \quad (1.47)$$

$$\implies a^{mq} = (a^q)^m < ((a^{\frac{n}{m}})^p)^m = ((a^{\frac{n}{m}})^m)^p = (a^n)^p = a^{np} \quad (1.48)$$

$$\implies a > 1 \text{ (by Theorem 0.2.9)} \quad (1.49)$$

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

$$\begin{cases} b^z < c^z & \iff 0 < z \\ b^z = c^z & \iff 0 = z \\ b^z > c^z & \iff 0 > z \end{cases} \quad (1.50)$$

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 0.2.10, 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

$$a^{r+s} = a^r a^s \quad (1.51)$$

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

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

$$= a^{np+mq} \quad (1.53)$$

Then observe

$$(a^r a^s)^{nq} = (a^{\frac{p}{q}} a^{\frac{m}{n}})^{nq} \quad (1.54)$$

$$= (a^{\frac{p}{q}})^{nq} (a^{\frac{m}{n}})^{nq} \text{ (by Theorem 0.2.14)} \quad (1.55)$$

$$= ((a^{\frac{p}{q}})^q)^n ((a^{\frac{m}{n}})^n)^q \quad (1.56)$$

$$= (a^p)^n (a^m)^q \quad (1.57)$$

$$= a^{np+mq} = (a^{r+s})^{nq} \text{ (Theorem 0.2.12 and Theorem 0.2.13)} \quad (1.58)$$

Then by Theorem 0.3.6, 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

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

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

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

$$= a^{mp} \quad (1.61)$$

Then observe

$$((a^r)^s)^{nq} = (((a^{\frac{p}{q}})^{\frac{m}{n}})^n)^q \quad (1.62)$$

$$= ((a^{\frac{p}{q}})^m)^q \quad (1.63)$$

$$= (a^{\frac{p}{q}})^{mq} \quad (1.64)$$

$$= ((a^{\frac{p}{q}})^q)^m \quad (1.65)$$

$$= (a^p)^m = a^{mp} = (a^{rs})^{nq} \quad (1.66)$$

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

$$a^r b^r = (ab)^r \quad (1.67)$$

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

$$\begin{cases} b^x < b^y & \iff b > 1 \\ b^x = b^y & \iff b = 1 \\ b^x > b^y & \iff 0 < b < 1 \end{cases} \quad (1.68)$$

Proof. By Theorem 0.3.5, 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 a 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

$$\sup SU = \sup S \sup U \quad (1.69)$$

and have

$$\inf SU = \inf S \inf U \quad (1.70)$$

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 a lower bound of SU . Now we show $\sup S \sup U$ is the *greatest* lower bound of SU .

Assume **there is a 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

$$a^{r+s} = a^r a^s \quad (1.71)$$

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 $\phi : S \rightarrow \mathbb{F}$ is an *order homomorphism*, if for all $a, b \in S$, we have

$$a < b \implies \phi(a) < \phi(b) \quad (1.72)$$

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 $\phi : S \rightarrow U$ is an order isomorphism. Let $US_\phi = \{s\phi(s) : s \in S\}$. We have

$$\sup S \sup U = \sup US_\phi \quad (1.73)$$

and have

$$\inf S \inf U = \inf US_\phi \quad (1.74)$$

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 $0 \leq s \leq \sup S$ and $0 \leq \phi(s) \leq \sup U$, so we can deduce $s\phi(s) \leq \sup S \sup U$. 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 $s \geq \inf S$ and $\phi(s) \geq \inf U$, so we can deduce $s\phi(s) \geq \inf S \inf U$. 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

$$b^z c^z = (bc)^z \quad (1.75)$$

Proof. From now, we will use $C(z)$ to denote $\{c^u : u \in \mathbb{Q}, u \leq z\}$. Define $\phi : B(z) \rightarrow C(z)$ 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 $\{(bc)^u : u \in \mathbb{Q}, u \leq z\} = \{b^t \phi(b^t) : t \in \mathbb{Q}, t \leq z\} = C(z)B(z)_\phi$. Then by Lemma 1.4.8, our proof is finished. We now show $(b^{-1})^z b^z = 1$.

WOLG, let $b < 1$. Denote $H := \{b^t : t \in \mathbb{Q}, t \leq z\}$, $E := \{b^{-t} : t \in \mathbb{Q}, t \leq z\}$, 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 $t \leq z$ 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 $t \leq z$ 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

$$\begin{cases} b^z < c^z & \iff 0 < z \\ b^z = c^z & \iff 0 = z \\ b^z > c^z & \iff 0 > z \end{cases} \quad (1.76)$$

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 a 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 $c - 1 = d^m - 1 = (d - 1)(d^{m-1} + d^{m-2} + \dots + 1) \geq (d - 1)m = (c^{\frac{1}{m}} - 1)m$ (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 $c - 1 = d^m - 1 = (d - 1)(d^{m-1} + d^{m-2} + \dots + 1) \leq (d - 1)m = (c^{\frac{1}{m}} - 1)m$ (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 $t = \frac{y}{b^x}$ 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 $b^{x-\frac{1}{n}} = \frac{b^x}{b^{\frac{1}{n}}} < \frac{b^x}{t} = y$. 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

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

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

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

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

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

So, we have

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

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

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

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 0.3.9, $(a^r)^s = ((a^{rs})^{\frac{1}{s}})^s = a^{rs}$ (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 field \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 .

$$(c, d) := \{\{c\}, \{c, d\}\} \quad (1.84)$$

$$(b, c, d) := (b, (c, d)) = \{\{b\}, \{b, (c, d)\}\} = \{\{b\}, \{b, \{\{c\}, \{c, d\}\}\}\} \quad (1.85)$$

$$(a, b, c, d) := (a, (b, c, d)) = \dots \quad (1.86)$$

Let $\{v_1, \dots, v_n\}$ be a basis for V , and define $\phi : \{v_1, \dots, v_n\} \rightarrow \mathbb{F}^n$ by $c_1v_1 + \dots + c_nv_n \mapsto (c_1, \dots, c_n)$. Notice that ϕ can be 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, \dots, 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 a 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

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

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 $\mathbf{x} \in \mathbb{R}^n$

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

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 $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ called Euclidean inner product defined by

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

and equipped with 2-norm

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

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

$$|\mathbf{x}||\mathbf{y}| \geq |\mathbf{x} \cdot \mathbf{y}| \quad (1.91)$$

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

$$f(t) = \sum (x_i t - y_i)^2 \quad (1.92)$$

$$= \sum x_i^2 t^2 - 2x_i y_i t + y_i^2 \quad (1.93)$$

$$= t^2 \sum x_i^2 - 2t \sum x_i y_i + \sum y_i^2 \quad (1.94)$$

$$= |\mathbf{x}|^2 t^2 - 2(\mathbf{x} \cdot \mathbf{y})t + |\mathbf{y}|^2 \quad (1.95)$$

$f(t)$ is a 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

$$\mathbf{x}\mathbf{y} = (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1) \quad (1.96)$$

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

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

$$\mathbf{y}\mathbf{x} = (y_1x_1 - y_2x_2, y_1x_2 + y_2x_1) = (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1) = \mathbf{x}\mathbf{y} \quad (1.97)$$

$$(\mathbf{x}\mathbf{y})\mathbf{z} = (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1)(z_1, z_2) \quad (1.98)$$

$$= (x_1y_1z_1 - x_2y_2z_1 - x_1y_2z_2 - x_2y_1z_2, x_1y_1z_2 - x_2y_2z_2 + x_1y_2z_1 + x_2y_1z_1) \quad (1.99)$$

$$= (x_1(y_1z_1 - y_2z_2) - x_2(y_2z_1 + y_1z_2), x_1(y_1z_2 + y_2z_2) + x_2(y_1z_1 - y_2z_2)) \quad (1.100)$$

$$= (x_1, x_2)(y_1z_1 - y_2z_2, y_1z_2 + y_2z_1) = \mathbf{x}(\mathbf{y}\mathbf{z}) \quad (1.101)$$

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

$$\mathbf{x}\left(\frac{x_1}{x_1^2 + x_2^2}, \frac{-x_2}{x_1^2 + x_2^2}\right) = \left(\frac{x_1^2 + x_2^2}{x_1^2 + x_2^2}, \frac{-x_1x_2 + x_1x_2}{x_1^2 + x_2^2}\right) \quad (1.103)$$

$$= (1, 0) \quad (1.104)$$

$$\mathbf{x}(\mathbf{y} + \mathbf{z}) = (x_1, x_2)(y_1 + z_1, y_2 + z_2) \quad (1.105)$$

$$= (x_1y_1 + x_1z_1 - x_2y_2 - x_2z_2, x_2y_1 + x_2z_1 + x_1y_2 + x_1z_2) \quad (1.106)$$

$$= (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1) + (x_1z_1 - x_2z_2, x_2z_1 + x_1z_2) \quad (1.107)$$

$$= \mathbf{x}\mathbf{y} + \mathbf{x}\mathbf{z} \quad (1.108)$$

■

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, $\mathbb{R} \not\subset \mathbb{C}$. 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 $\text{Re}(a + bi) := a$ and $\text{Im}(a + bi) := b$. 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

$$\overline{(\bar{z})} = z \quad (1.109)$$

$$\overline{z + w} = \bar{z} + \bar{w} \quad (1.110)$$

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

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

$$z + \bar{z} = 2\text{Re}(z) \text{ and } z - \bar{z} = 2i\text{Im}(z) \quad (1.113)$$

$$|\bar{z}| = |z| \quad (1.114)$$

$$|zw| = |z||w| \quad (1.115)$$

$$|z^n| = |z|^n \quad (1.116)$$

Proof. We prove only the following. Let $z = z_1 + z_2i$ and $w = w_1 + w_2i$. We have

$$\overline{(\bar{z})(\bar{w})} = \overline{(z_1 - z_2i)(w_1 - w_2i)} \quad (1.117)$$

$$= \overline{z_1w_1 - z_2w_2 - (z_2w_1 + z_1w_2)i} \quad (1.118)$$

$$= z_1w_1 - z_2w_2 + (z_2w_1 + z_1w_2)i \quad (1.119)$$

$$= zw \quad (1.120)$$

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

$$|zw| = |z_1w_1 - z_2w_2 + (z_2w_1 + z_1w_2)i| \quad (1.122)$$

$$= [(z_1w_1 - z_2w_2)^2 + (z_2w_1 + z_1w_2)^2]^{\frac{1}{2}} \quad (1.123)$$

$$= [(z_1w_1)^2 + (z_2w_2)^2 - 2z_1z_2w_1w_2 + (z_2w_1)^2 + (z_1w_2)^2 + 2z_1z_2w_1w_2]^{\frac{1}{2}} \quad (1.124)$$

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

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

■

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, $p : V \rightarrow \mathbb{R}$ is a function that satisfy

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

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

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

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

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

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

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

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

and observe

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

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

Observe

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

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

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

$$= |\lambda| |\mathbf{x}| \quad (1.138)$$

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 $\|\mathbf{x}\| := \max(|x_1|, |x_2|, \dots, |x_n|)$. 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 $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$ is a function that satisfy the following three axioms:

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

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

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

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

$$\mathbf{x} \cdot \mathbf{y} = \sum x_i y_i = \sum y_i x_i = \mathbf{y} \cdot \mathbf{x} = \overline{\mathbf{y} \cdot \mathbf{x}} \quad (1.142)$$

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

$$\mathbf{x} \cdot \mathbf{x} = \sum x_i^2 \geq 0 \quad (1.144)$$

■

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

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

Lemma 1.6.9. (Arising from inner product \rightarrow 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

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle \quad (1.146)$$

$$= \langle \mathbf{x}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{x} - \mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle \quad (1.147)$$

$$= 2\langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{x} \rangle \quad (1.148)$$

$$= 2\langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{y}, \mathbf{y} \rangle = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) \quad (1.149)$$

■

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

$$\|\mathbf{x} + \mathbf{y}\|^2 > \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| \quad (1.150)$$

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

$$\|\mathbf{x} - \mathbf{y}\|^2 > \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| \quad (1.151)$$

$$= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| \quad (1.152)$$

So we have

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

This **CaC** to Lemma 1.6.8.

Observe

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

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 $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{C}^n$ and $\mathbf{w} = (w_1, \dots, w_n) \in \mathbb{C}^n$

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

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

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

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

$$\sum |Bv_j - Cw_j|^2 = \sum (Bv_j - Cw_j) \overline{(Bv_j - Cw_j)} \quad (1.157)$$

$$= \sum (Bv_j - Cw_j) (B\overline{v_j} - \overline{Cw_j}) \quad (1.158)$$

$$= \sum B^2 |v_j|^2 - BC\overline{v_j}w_j - B\overline{C}v_j\overline{w_j} + C\overline{C}|w_j|^2 \quad (1.159)$$

$$= B^2A - BC\overline{C} - B\overline{C}C + C\overline{C}B \quad (1.160)$$

$$= B^2A - BC\overline{C} \quad (1.161)$$

$$= B(AB - |C|^2) \quad (1.162)$$

Because $B > 0$, then we can deduce

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

So we can deduce

$$\sum |v_j|^2 \sum |w_j|^2 \geq |\sum v_j \overline{w_j}|^2 \quad (1.164)$$

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

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

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

$$\|B\mathbf{v} - C\mathbf{w}\|^2 = \langle B\mathbf{v} - C\mathbf{w}, B\mathbf{v} - C\mathbf{w} \rangle \quad (1.166)$$

$$= \langle B\mathbf{v}, B\mathbf{v} \rangle - \langle C\mathbf{w}, B\mathbf{v} \rangle - \langle B\mathbf{v}, C\mathbf{w} \rangle + \langle C\mathbf{w}, C\mathbf{w} \rangle \quad (1.167)$$

$$= B^2 \langle \mathbf{v}, \mathbf{v} \rangle - CB \langle \mathbf{w}, \mathbf{v} \rangle - B\overline{C} \langle \mathbf{v}, \mathbf{w} \rangle + C\overline{C} \langle \mathbf{w}, \mathbf{w} \rangle \quad (1.168)$$

$$= B^2A - CB\overline{C} - B\overline{C}C + C\overline{C}B \quad (1.169)$$

$$= B(AB - |C|^2) \quad (1.170)$$

Because $B > 0$, we can deduce

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

So we can deduce

$$\|\mathbf{v}\|^2\|\mathbf{w}\|^2 \geq |\langle \mathbf{v}, \mathbf{w} \rangle|^2 \quad (1.172)$$

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

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\| \quad (1.173)$$

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

$$\|\mathbf{v} + \mathbf{w}\|^2 = \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle \quad (1.174)$$

$$= \|\mathbf{v}\|^2 + \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle + \|\mathbf{w}\|^2 \quad (1.175)$$

$$= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + \langle \mathbf{v}, \mathbf{w} \rangle + \overline{\langle \mathbf{v}, \mathbf{w} \rangle} \quad (1.176)$$

$$= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\operatorname{Re}(\langle \mathbf{v}, \mathbf{w} \rangle) \quad (1.177)$$

$$\leq \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2|\langle \mathbf{v}, \mathbf{w} \rangle| \quad (1.178)$$

$$\leq \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2(\|\mathbf{v}\|)(\|\mathbf{w}\|) \quad (1.179)$$

$$= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2 \quad (1.180)$$

■

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

$$\inf A = -\sup(-A) \quad (1.181)$$

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

$$|\sum z_j| \leq \sum |z_j| \quad (1.182)$$

Question 8

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

$$||x| - |y|| \leq |x - y| \quad (1.183)$$

Question 9

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

$$|1 + z|^2 + |1 - z|^2 \quad (1.184)$$

Question 10

Let $\mathbf{x} \in \mathbb{R}^k$ and $\mathbf{x} \neq \mathbf{0}$. When $k = 1$, prove that there does not exist $\mathbf{y} \in \mathbb{R}^k$, 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 $\mathbf{z} \in \mathbb{R}^k$ 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 0.2.1, we present the two ordered field axioms:

$$y < z \longrightarrow x + y < x + z \quad (1.185)$$

$$x > 0 \text{ and } y > 0 \longrightarrow xy > 0 \quad (1.186)$$

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 $y > 0 \rightarrow xy > 0$, 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

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

Question 15

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

Chapter 2

Basic Topology

2.1 Different but Equivalent Ways to Define A Topology

This section illustrates three methods to define a topology. However, in practice, the approach using open sets is predominantly employed.

Definition 2.1.1. (Definition of Topology, via Open Set) Given a set X , which called a topological space, if we say a family \mathfrak{D} of subsets of X is a topology on X , then we mean

$$X, \emptyset \in \mathfrak{D} \quad (2.1)$$

$$A, B \in \mathfrak{D} \implies A \cap B \in \mathfrak{D} \quad (2.2)$$

$$\mathfrak{A} \subseteq \mathfrak{D} \implies \bigcup \mathfrak{A} \in \mathfrak{D} \quad (2.3)$$

If a subset of X belong to \mathfrak{D} , we say it is an open set, and if a subset of X is an complement of an open set, we say it is an closed set.

The method above is the usual way to define a topology. Next, we'll look at some properties of closed sets. Keep in mind that the related properties for open sets are basic axioms

Theorem 2.1.2. (Property of Closed Sets) Let \mathfrak{F} be the family of precisely all closed sets in a topological space. We have

$$X, \emptyset \in \mathfrak{F} \quad (2.4)$$

$$A, B \in \mathfrak{F} \implies A \cup B \in \mathfrak{F} \quad (2.5)$$

$$\mathfrak{B} \subseteq \mathfrak{F} \implies \bigcap \mathfrak{B} \in \mathfrak{F} \quad (2.6)$$

Proof.

$$X = \emptyset^c \text{ and } \emptyset = X^c \in \mathfrak{F} \quad (2.7)$$

$$A, B \in \mathfrak{F} \implies A^c, B^c \in \mathfrak{D} \implies A^c \cap B^c = (A \cup B)^c \in \mathfrak{D} \implies A \cup B \in \mathfrak{F} \quad (2.8)$$

$$\mathfrak{B} \subseteq \mathfrak{F} \implies \{F^c : F \in \mathfrak{B}\} \subseteq \mathfrak{D} \implies \bigcup_{F \in \mathfrak{B}} F^c \in \mathfrak{D} \implies \bigcap \mathfrak{B} = (\bigcup_{F \in \mathfrak{B}} F^c)^c \in \mathfrak{F} \quad (2.9)$$

■

Theorem 2.1.3. (Equivalent Ways to Define The Same Topology, Part 1) Define f on $\mathcal{P}(\mathcal{P}(X))$ by

$$f(\mathfrak{A}) = \{A^c : A \in \mathfrak{A}\} \quad (2.10)$$

If \mathfrak{D} is a topology on X , then $f(\mathfrak{D})$ are the family of precisely all closed sets on X . Also, we have

$$f = f^{-1} \quad (2.11)$$

This mean if we are given a family of closed sets, the family of closed sets given by the topology induced by our original family is our original family.

Proof. Trivial. ■

Let's move on to the axioms for the neighborhood system. After covering that, we'll demonstrate that these axioms and those for open sets aren't just interconnected – they're essentially the same thing. Notice the proof below uses Zorn's Lemma.

Definition 2.1.4. (Definition of Neighborhood Function) We say a function $\mathcal{N} : X \rightarrow \mathcal{P}(\mathcal{P}(X))$ is neighborhood function if

$$\forall x \in X, \mathcal{N}(x) \neq \emptyset \text{ (Existence)} \quad (2.12)$$

$$N \in \mathcal{N}(x) \implies x \in N \text{ (Around)} \quad (2.13)$$

$$\exists N \in \mathcal{N}(x), N \subseteq M \implies M \in \mathcal{N}(x) \text{ (Super set)} \quad (2.14)$$

$$N, M \in \mathcal{N}(x) \implies N \cap M \in \mathcal{N}(x) \text{ (Intersection)} \quad (2.15)$$

$$\forall N \in \mathcal{N}(x), \exists M \in \mathcal{N}(x), M \subseteq N \text{ and } \forall y \in M, N \in \mathcal{N}(y) \quad (2.16)$$

The last property reads: every neighborhood N around arbitrary point x contain a neighborhood M around x such that N is a neighborhood of each points of M .

Theorem 2.1.5. (Equivalent Ways to Define The Same Topology, Part 2) There exists a one-to-one correspondence between the class of all topology on X and the class of all neighborhood function on X .

The correspondence can be written in the form of

$$(g(\mathfrak{D}))(x) = \{N \subseteq X : \exists O \in \mathfrak{D}, x \in O \subseteq N\} \quad (2.17)$$

The inverse of g will be proved to be

$$g^{-1}(\mathcal{N}) = \{O : \forall x \in O, O \in \mathcal{N}(x)\} \quad (2.18)$$

Proof. We first prove g do map topology to neighborhood function.

$g(\mathfrak{D})(x)$ is always nonempty because

$$X \in \mathfrak{D} \implies X \in (g(\mathfrak{D}))(x) \quad (2.19)$$

By definition of g , every defined neighborhood around x do contain x . A super set of $N \in (g(\mathfrak{D}))(x)$ belong to $(g(\mathfrak{D}))(x)$, since an open set containing x contained by N is also contained by the super set.

Let $N, M \in (g(\mathfrak{D}))(x)$ and $x \in O_N \subseteq N, x \in O_M \subseteq M$. We have

$$x \in O_N \cap O_M \subseteq N \cap M \quad (2.20)$$

For each neighborhood N around x , the open set containing x contained by the neighborhood is also a neighborhood around x , since it contain itself and it is an open set. Then we see N is a neighborhood around each point of the open set, since each point is contained by the open set contained by N (done)

We now prove g^{-1} do map neighborhood function to topology.

Proving $\emptyset \in g^{-1}(\mathcal{N})$ is trivial. We have $X \in g^{-1}(\mathcal{N})$ because $\mathcal{N}(x)$ is nonempty for all x and X is a super set of any neighborhood.

Let $A, B \in g^{-1}(\mathcal{N})$. Arbitrarily pick x in $A \cap B$, we know by definition of g^{-1} , A is a neighborhood of x and so is B . Then by definition of neighborhood function, we know $A \cap B \in \mathcal{N}(x)$. Because x is arbitrary picked from $A \cap B$, we have proved $A \cap B \in g^{-1}(\mathcal{N})$.

Let $\mathfrak{A} \subseteq g^{-1}(\mathcal{N})$. We deduce

$$x \in \bigcup \mathfrak{A} \implies \exists O \in \mathfrak{A}, x \in O \in \mathfrak{A} \subseteq g^{-1}(\mathcal{N}) \implies O \in \mathcal{N}(x) \implies O \subseteq \bigcup \mathfrak{A} \in \mathcal{N}(x) \quad (2.21)$$

By definition of $g^{-1}(\mathcal{N})$, the above deduction shows $\bigcup \mathfrak{A} \in g^{-1}(\mathcal{N})$ (done)

Lastly, we prove g^{-1} is the inverse of g . To prove such, we have to first prove $g^{-1}(g(\mathfrak{D})) = \mathfrak{D}$.

Arbitrarily pick $Z \in \mathfrak{D}$. Deduce

$$x \in Z \implies \exists Z \in \mathfrak{D}, x \in Z \subseteq Z \implies Z \in (g(\mathfrak{D}))(x) \quad (2.22)$$

By definition of g^{-1} , our deduction give us $Z \in g^{-1}(g(\mathfrak{D}))$. Because Z is arbitrarily picked from \mathfrak{D} , we have $\mathfrak{D} \subseteq g^{-1}(g(\mathfrak{D}))$.

Arbitrarily pick $Y \in g^{-1}(g(\mathfrak{D}))$. Define $\mathfrak{S} := \{O \in \mathfrak{D} : O \subseteq Y\}$. By definition of g^{-1} and g , we have

$$y \in Y \implies Y \in (g(\mathfrak{D}))(y) \implies \exists O \in \mathfrak{D}, y \in O \subseteq Y \implies y \in \bigcup \mathfrak{S} \quad (2.23)$$

The last implication hold true because $O \in \mathfrak{S}$ by definition of \mathfrak{S} . We have proved $Y \subseteq \bigcup \mathfrak{S}$. Also, trivially by definition of \mathfrak{S} , we have $\bigcup \mathfrak{S} \subseteq Y$. Then we have $Y = \bigcup \mathfrak{S} \in \mathfrak{D}$. Because Y is arbitrarily picked from $g^{-1}(g(\mathfrak{D}))$, we have proved $g^{-1}(g(\mathfrak{D})) \subseteq \mathfrak{D}$ (done)

We now prove $g(g^{-1}(\mathcal{N})) = \mathcal{N}$. More precisely, we want to prove $\forall x, g(g^{-1}(\mathcal{N}))(x) = \mathcal{N}(x)$

Let $x \in X$. Observe

$$Y \in g(g^{-1}(\mathcal{N}))(x) \implies \exists O \in g^{-1}(\mathcal{N}), x \in O \subseteq Y \quad (2.24)$$

By definition of g^{-1} , we see $O \in \mathcal{N}(x)$. Then by super set property, we know $Y \in \mathcal{N}(x)$. We have proved $g(g^{-1}(\mathcal{N}))(x) \subseteq \mathcal{N}(x)$.

Arbitrarily pick $N \in \mathcal{N}(x)$. Let

$$U := \{y \in N : N \in \mathcal{N}(y)\} \quad (2.25)$$

Deduce

$$y \in U \implies N \in \mathcal{N}(y) \quad (2.26)$$

$$\implies \exists M \in \mathcal{N}(y), M \subseteq N, \forall z \in M, N \in \mathcal{N}(z) \quad (2.27)$$

$$\implies \forall z \in M, z \in N \text{ and } N \in \mathcal{N}(z) \quad (2.28)$$

$$\implies \forall z \in M, z \in U \implies M \subseteq U \implies U \in \mathcal{N}(y) \quad (2.29)$$

We have proved

$$\forall y \in U, U \in \mathcal{N}(y) \quad (2.30)$$

which means

$$U \in g^{-1}(\mathcal{N}) \quad (2.31)$$

By definition of U , we know

$$U \subseteq N \quad (2.32)$$

The fact N is chosen from $\mathcal{N}(x)$ let us know $x \in U \subseteq N$, so we can deduce

$$N \in g(g^{-1}(\mathcal{N}))(x) \quad (2.33)$$

Because N is arbitrarily picked from $\mathcal{N}(x)$, we have $\mathcal{N}(x) \subseteq g(g^{-1}(\mathcal{N}))(x)$ (done) (done) ■

We now give the formal definition of neighborhood while a topology is given, and translate Theorem ?? into language without neighborhood function.

Definition 2.1.6. (Definition of Neighborhood) We say a set E is a neighborhood around a point x , if E contains an open set containing x .

Corollary 2.1.7. (Property of Neighborhood, Part 1)

- (a) For all points x , there exists a neighborhood around x .
- (b) The super set of a neighborhood around x is a neighborhood around x .
- (c) The intersection of two neighborhood around x is a neighborhood around x .
- (d) For each neighborhood N around x , there exists a smaller neighborhood M around x such that N is a neighborhood around all points in M .

Theorem 2.1.8. (Property of Neighborhood, Part 2)

$$E \text{ is open} \iff E \text{ is a neighborhood around all points in } E \quad (2.34)$$

Proof. From left to right, it trivially follows from our definition of neighborhood.

From right to left, we know every point x in E is contained by an open set contained by E . By axiom, the union of such open sets is an open set and contained by E . Also, every point in E belong to an open set contained by E , so belong to the union. We have proved E is the union. ■

2.2 Closure and Interior Operator

This section won't go into any detailed general topology. It merely work as a comparison for the next section.

Definition 2.2.1. (Definition of Closure and Interior Operator) We define the closure \overline{E} of E to be smallest closed set containing E

$$\overline{E} := \bigcap \{F \in \mathfrak{F} : E \subseteq F\} \quad (2.35)$$

and define the interior E° of E to be the largest open set contained by E

$$E^\circ := \bigcup \{O \in \mathfrak{D} : O \subseteq E\} \quad (2.36)$$

Theorem 2.2.2. (Basic Property of Closure and Interior Operator) We have

$$\overline{E} = E \iff E \in \mathfrak{F} \text{ and } E^\circ = E \iff E \in \mathfrak{D} \quad (2.37)$$

$$\overline{(\overline{E})} = \overline{E} \text{ and } (E^\circ)^\circ = E^\circ \quad (2.38)$$

$$S \subseteq T \implies S^\circ \subseteq T^\circ \text{ and } \overline{S} \subseteq \overline{T} \quad (2.39)$$

Theorem 2.2.3. (Basic Property of Closure and Interior Operator) We have

$$\overline{E \cup F} = \overline{E} \cup \overline{F} \text{ and } (E \cap F)^\circ = E^\circ \cap F^\circ \quad (2.40)$$

Now, we define interior point and limit point without referencing interior and closure. We'll also demonstrate that the interior, described as the largest open set within E , can in turn define interior points.

Aside from the last section, this is another example in general topology where definitions can be expressed in several ways, yet they remain equivalent.

Definition 2.2.4. (Definition of Interior Point) A point p is an interior point of E if

$$E \in \mathcal{N}(p) \quad (2.41)$$

Definition 2.2.5. (Definition of Limit and Isolated Point) A point $p \in X$ is a limit point of a subset E of X , if

$$\forall O \in \mathfrak{D} : p \in O, \exists u \neq p, u \in O \cap E \quad (2.42)$$

or equivalently

$$\forall O \in \mathfrak{D} : p \in O, O \cap E \setminus \{p\} \neq \emptyset \quad (2.43)$$

or equivalently

$$\forall N \in \mathcal{N}(p), N \cap E \setminus \{p\} \neq \emptyset \quad (2.44)$$

We denote the set of limit point of E by E' .

Theorem 2.2.6. (Interior Contain Exactly All Interior Points) We have

$$E^\circ = \{p : E \in \mathcal{N}(p)\} \quad (2.45)$$

Proof. Arbitrarily pick p from $\{p : E \in \mathcal{N}(p)\}$. By definition, we have

$$\exists O \in \mathfrak{D}, p \in O \subseteq E \text{ where } O \subseteq E^\circ \quad (2.46)$$

Then we have $p \in E^\circ$. Because p is arbitrarily picked from $\{p : E \in \mathcal{N}(p)\}$, we now have $\{p : E \in \mathcal{N}(p)\} \subseteq E^\circ$. Arbitrarily pick p' from E° . We deduce

$$p' \in E^\circ \subseteq E \implies E \in \mathcal{N}(p') \implies p' \in \{p : E \in \mathcal{N}(p)\} \quad (2.47)$$

Because p' is arbitrarily picked from E° , we have $E^\circ \subseteq \{p : E \in \mathcal{N}(p)\}$ ■

Theorem 2.2.7. (Closure Contain Exactly All Limit Points of E and Points in E) We have

$$\overline{E} = E' \cup E \quad (2.48)$$

Proof. Arbitrarily pick $p \in (E' \cup E)^c$. Because p is neither a limit point nor in E , we know there exists an open set O containing p and disjoint with E . We know O doesn't contain a limit point of E , otherwise O intersect with E . We have constructed an open set containing p in $(E' \cup E)^c$, so we have proved $(E' \cup E)^c$ is open. Then we know $E' \cup E$ is closed.

Let F be a closed set containing E . Arbitrarily pick $p \in F^c$. Clearly p is not in E , and we can see p is not a limit point of E , otherwise the open set F^c would have intersect with E . We have proved $F^c \subseteq (E' \cup E)^c$. In other words, we have proved $E' \cup E$ is the smallest closed set containing E . ■

Corollary 2.2.8. (Every Point in Closure And Not in the Set is a Limit Point)

$$\overline{E} \setminus E \subseteq E' \quad (2.49)$$

Above illustrate two different but equivalent way to define closure and interior. One must be familiar with both of them.

We now give some useful facts that shall later be compared with.

Theorem 2.2.9. (Point in Closure if and only if No Neighborhood is Disjoint)
We have

$$\overline{E} = \{p : \forall M \in \mathcal{N}(p), M \cap E \neq \emptyset\} \quad (2.50)$$

Proof. There are two possibilities for $p \in \overline{E}$

$$p \in E \text{ or } p \in E' \quad (2.51)$$

There are two possibilities for point have no disjoint neighborhood

$$p \in E \text{ or } p \notin E \quad (2.52)$$

Notice each first possibilities are identical, and each second possibilities are equivalent. ■

Theorem 2.2.10. (Point is Limit Point if and only if In the Closure of Set Excluding The Point) We have

$$E' = \{p : p \in \overline{E \setminus \{p\}}\} \quad (2.53)$$

Proof. Notice each limit point $p \in E'$ is still a limit point of $E \setminus \{p\}$, and notice each limit point of E is a limit point of $E \cup \{p\}$. ■

2.3 Three Non-Equivalent Compactness and Separable

In this section, we will first give two equivalent definition of base and prove they are equivalent.

Definition 2.3.1. (Definition of Base) We say a sub-collection \mathcal{O} of topology τ on X is a base if every open neighborhood O around x contain an open set U in \mathcal{O} containing U . That is

$$\forall x \in X, \forall O \in \tau : x \in O, \exists U \in \mathcal{O}, x \in U \subseteq O \quad (2.54)$$

We call open sets in a base basic open sets.

Theorem 2.3.2. (Equivalent Definition of Base) A sub-collection of topology τ is a base if and only if every open set can be expressed as a union of a set of basic open sets. That is, the necessary and sufficient condition for \mathcal{O} to be a base is

$$\forall O \in \tau, \exists \mathcal{U} \subseteq \mathcal{O}, O = \bigcup \mathcal{U} \quad (2.55)$$

Proof. From left to right, for each point $x \in O$, we know there exists $U_x \in \mathcal{O}$ such that $x \in U_x \subseteq O$. Collect U_x for each $x \in O$ as \mathcal{U} , and we are partially done.

From right to left, let O be an open neighborhood around x . If there exists no $U \in \mathcal{O}$ such that $x \in U \subseteq O$, then every union of sub-collection $\mathcal{U} \subseteq \mathcal{O}$ that contain x contain some points not in O , which is impossible. ■

Then, we will define separable, and give a sufficient condition of separable.

Definition 2.3.3. (Definition of Separable) We say X is separable if

$$X \text{ has a countable dense subset} \quad (2.56)$$

Theorem 2.3.4. (Base and Dense Set)

$$\text{A set intersecting with all basic open set is dense} \quad (2.57)$$

Proof. For each point p not in A and for each open neighborhood N_p around p , there exists an basic open set O_p contained by N_p . The fact O_p intersect A implies N_p intersect with A . We have proved every point p not in A is a limit point of A . ■

Corollary 2.3.5. (Countable Base Implies Separable)

$$X \text{ has a countable base} \implies X \text{ is separable} \quad (2.58)$$

Proof. Let \mathcal{O} be a countable base for X . For each $O \in \mathcal{O}$, we can collect a point $x_O \in O$, and have $A := \{x_O : O \in \mathcal{O}\}$. Because \mathcal{O} is countable, we know A is countable, and by Theorem ??, we know A is dense. ■

The above let us draw a Venn diagram, where existence of countable base is a small circle in the big circle separable.

Now, we give definition of three different compactness.

Definition 2.3.6. (Definition of Open Cover) We say a collection $\{G_\alpha\}$ of open sets is an open cover of E if

$$E \subseteq \bigcup \{G_\alpha\} \quad (2.59)$$

Definition 2.3.7. (Definition of Compact Space) We say a topological space X is compact if

$$\text{every open cover for } X \text{ has a finite open sub-cover} \quad (2.60)$$

Definition 2.3.8. (Definition of Countably Compact Space) We say a topological space X is countably compact if

$$\text{every countable open cover for } X \text{ has a finite open sub-cover} \quad (2.61)$$

Definition 2.3.9. (Definition of Limit Point Compact Space) We say a topological space X is limit point compact if

$$\text{every infinite subset of } X \text{ has a limit point} \quad (2.62)$$

It is clear that compact implies countably compact, as every countable cover in compact space has a finite sub-cover. We now explore the relation of the three compactness.

Theorem 2.3.10. (Compact Implies Countably Compact)

$$X \text{ is compact} \implies X \text{ is countably compact} \quad (2.63)$$

Theorem 2.3.11. (Countably Compact Implies Limit Point Compact)

$$X \text{ is countably compact} \implies X \text{ is limit point compact} \quad (2.64)$$

Proof. Arbitrarily pick an infinite A . We know A is either countable or uncountable. We first prove *if A is countably infinite, then A has a limit point*.

Assume *A has no limit point*. For each $a \in A$, because a is not a limit point of A , we know there exists an open neighborhood O_a around a such that $O_a \cap A = \{a\}$. Notice that A has no limit point let us deduce

$$A' = \emptyset \implies \overline{A} = A \implies A^c \text{ is open} \quad (2.65)$$

For each $a \in A$, collect one O_a into the collection $\mathcal{O} = \{O_a : a \in A\}$.

Then, because A is countable, we see $\mathcal{O} \cup \{A^c\}$ is an countable cover for X , and every finite open sub-cover contain only finite amount of point of A when A is infinite. **CaC (done)**

If A is uncountable, then A contain a countable set B . We know B has a limit point p , and we know p is also a limit point for A . ■

Corollary 2.3.12. (Compact Implies Limit Point Compact) If X is compact, then X is limit point compact.

Lastly, we give a theorem involving countable base.

Theorem 2.3.13. (Existence of Countable Base Implies Existence of Countable Sub-Cover)

$$X \text{ has a countable base} \implies \text{every open cover for } X \text{ has a countable sub-cover} \quad (2.66)$$

Proof. Let \mathcal{O} be a countable base for X and let \mathcal{G} be an open cover of X . We wish to find an countable open sub-cover of \mathcal{G} .

Because \mathcal{O} is a base, for each $G \in \mathcal{G}$, there exists basic open set $O \in \mathcal{O}$ contained by G . Then because \mathcal{G} is an open cover, for each $p \in X$, there exists G_p, O_p such that

$$p \in O_p \subseteq G_p \quad (2.67)$$

Collect all G_p from above so we have a sub-cover \mathcal{G}' . Because \mathcal{O} is countable, we know \mathcal{G}' is also countable. ■

Now, if we want to draw a conclusion for this section by drawing a Venn diagram, we should first draw compact inside countably compact inside limit point compact.

Chapter 3

Important Notions in Metric Space

3.1 Sequence

Definition 3.1.1. (Definition of Metric) We say a real-valued function $d : X^2 \rightarrow \mathbb{R}$ is a metric if

$$d(x, x) = 0 \quad (3.1)$$

$$x \neq y \longrightarrow d(x, y) > 0 \text{ (Positive Definiteness)} \quad (3.2)$$

$$d(x, y) = d(y, x) \text{ (Commutative)} \quad (3.3)$$

$$d(x, z) \leq d(x, y) + d(y, z) \text{ (Triangle Inequality)} \quad (3.4)$$

Definition 3.1.2. (Definition of Sequence) In this chapter, by a sequence in X , we mean a function from \mathbb{N} to X

Above is specification.

Definition 3.1.3. (Definition of Convergence of Reals) Let $\{a_n\}$ be a sequence of reals. By $\{a_n\}$ converges to a

$$\lim_{n \rightarrow \infty} a_n = a \quad (3.5)$$

We mean for each $\epsilon \in \mathbb{R}^+$, there exists great enough $N_\epsilon \in \mathbb{N}$ such that for each $n \in \mathbb{N}$

$$n > N_\epsilon \implies |a_n - a| < \epsilon \quad (3.6)$$

Definition 3.1.4. (Definition of Convergence) Let $\{a_n\}$ be a sequence in metric space X . By $\{a_n\}$ converge to a

$$\lim_{n \rightarrow \infty} a_n = a \quad (3.7)$$

We mean for each $\epsilon \in \mathbb{R}^+$, there exists great enough $N_\epsilon \in \mathbb{N}$ such that for each $n \in \mathbb{N}$

$$n > N_\epsilon \implies d(a_n, a) < \epsilon \quad (3.8)$$

Above is definition.

Theorem 3.1.5. (Other Definition for Convergence) Let $\{a_n\}$ be a sequence in metric space X . We have

$$\lim_{n \rightarrow \infty} a_n = a \iff \lim_{n \rightarrow \infty} d(a_n, a) = 0 \quad (3.9)$$

$$\iff (\forall \epsilon, \exists N_\epsilon, n \geq N_\epsilon \iff d(a_n, a) < \epsilon) \quad (3.10)$$

$$\iff (\forall \epsilon, \exists N_\epsilon, a_n \in B_\epsilon(a) \iff n \geq N_\epsilon) \quad (3.11)$$

Theorem 3.1.6. (Limit of Sequence is Unique) Let $\{a_n\}$ be a sequence in metric space X . We have

$$\lim_{n \rightarrow \infty} a_n = a \text{ and } \lim_{n \rightarrow \infty} a_n = a' \implies a' = a \quad (3.12)$$

Proof. Assume not unique and use $\epsilon = \frac{d(a', a)}{2}$ to cause a contradiction. ■

Below discuss sub-sequence.

Definition 3.1.7. (Definition of Sub-Sequence) We use

$$\{a_{n_k}\} \quad (3.13)$$

To denote sub-sequence of $\{a_n\}$, so $n_k \in \mathbb{N}$ must satisfy

$$n_k > \max\{n_1, \dots, n_{k-1}\} \text{ and } n_k \geq k \quad (3.14)$$

Theorem 3.1.8. (Sequence Converge if and only if All Sub-sequence Converge)

$$\text{Sequence } \{a_n\} \text{ converge to } a \iff \text{all sub-sequence } \{a_{n_k}\} \text{ converge to } a \quad (3.15)$$

Proof. From right to left is trivial. From left to right, notice $k_n \geq n$. ■

Theorem 3.1.9. (Existence of a Convenient sub-Sequence)

$$\{a_n\} \text{ converge to } a \quad (3.16)$$

$$\implies \text{there exists a sub-sequence } \{a_{n_k}\} \text{ such that } \forall k \in \mathbb{N}, d(a_{n_k}, a) < \frac{1}{k} \quad (3.17)$$

Proof. The construction is done by recursion of selecting a point in $B_{\frac{1}{k+1}}(a) \cap \{a_n\}$ after a_{n_k} . Such point must exists, otherwise all point after a_{n_k} is of distance to a greater than $\frac{1}{k+1}$. ■

Below introduce Cauchy.

Definition 3.1.10. (Definition of Cauchy Sequence) We say

$$\{a_n\} \text{ is a Cauchy sequence} \quad (3.18)$$

If for each $\epsilon \in \mathbb{R}^+$, there exists great enough $N_\epsilon \in \mathbb{N}$ such that for each $n, k \in \mathbb{N}$

$$n, k > N_\epsilon \implies d(a_n, a_k) < \epsilon \quad (3.19)$$

Theorem 3.1.11. (Every Convergent Sequence is Cauchy)

$$\{a_n\} \text{ converge} \implies \{a_n\} \text{ is Cauchy} \quad (3.20)$$

Proof. Let $\{a_n\}$ converge to a . We know for each ϵ there exists $N_{\frac{\epsilon}{2}}$ such that $n > N_{\frac{\epsilon}{2}} \implies d(a_n, a) < \frac{\epsilon}{2}$. Observe $m, k > N_{\frac{\epsilon}{2}} \implies d(a_m, a_k) \leq d(a_m, a) + d(a, a_k) < \epsilon$ ■

Lastly, we close this section with example of uncompleted metric space.

Theorem 3.1.12. (Not Every Cauchy Sequence Converge)

$$\{a_n\} \text{ is Cauchy NOT} \implies \{a_n\} \text{ converge} \quad (3.21)$$

Proof. Let $X = \mathbb{R} \setminus \{p\}$ be standard. Observe $\{a_n = p + \frac{1}{n}\}$ is in X and for each $\epsilon \in \mathbb{R}^+$, we have

$$m \geq k > \frac{1}{\epsilon} \implies d(a_m, a_k) = a_k - a_m < \frac{1}{k} - \frac{1}{m} < \frac{1}{k} < \epsilon \quad (3.22)$$

Let $q \in X$. We know $q \neq p$. Let $\epsilon = \frac{|p-q|}{2}$. Let $k > \frac{1}{\epsilon}$. Observe

$$\frac{1}{k} < \frac{|p-q|}{2} \implies \frac{q-p}{2} < \frac{1}{k} < \frac{p-q}{2} \quad (3.23)$$

And observe

$$d(a_k, q) = |p + \frac{1}{k} - q| > p - q + \frac{1}{k} > \frac{p-q}{2} \quad (3.24)$$

$$> q - p - \frac{1}{k} > \frac{q-p}{2} \quad (3.25)$$

So we have

$$d(a_k, q) > \frac{|p-q|}{2} = \epsilon \quad (3.26)$$

We have found an open ball around q that contain only finite amount of $\{a_n\}$, showing $\{a_n\}$ does not converge to q when q arbitrarily picked from X . ■

3.2 Closed

Definition 3.2.1. (Definition of Closure and Closed in Metric Space) Let E be a subset of (X, d) . By closure \overline{E} , we mean the set containing precisely all possible points to which the sequence in E may converge. We say E is closed if every convergent sequence in E converge to some point in E , in other words:

$$\overline{E} = E \quad (3.27)$$

Theorem 3.2.2. (Another Definition of Closure)

$$\overline{E} \text{ is the smallest closed set containing } E \quad (3.28)$$

Proof. To see $E \subseteq \overline{E}$ is simple, for each $x \in E$, the trivial sequence $\{x_n = x\}$ is a sequence in E that converge to x .

We first prove \overline{E} is closed. Let $\{x_n\}$ be a convergent sequence in \overline{E} , and let

$$\lim_{n \rightarrow \infty} x_n = x \quad (3.29)$$

We wish to prove $x \in \overline{E}$. If $\{x_n\}$ contain infinite terms in E , then we know $\{x_n\}$ has a sub-sequence $\{x_{n_k}\}$ in E which converge to x by Theorem ???. By definition of closure, we know $x \in \overline{E}$ and we are done.

Now, we only have to consider when $\{x_n\}$ contain finite terms in E , which tell us $\{x_n\}$ contain infinite terms (a sub-sequence that converge to x) $\{x_{n_k}\}$ in $\overline{E} \setminus E$.

We wish to find a sequence in E that converge to x , then we know x is in \overline{E} . By definition of closure, we know each term x_{n_k} of the sub-sequence $\{x_{n_k}\}$ is a limit of a sequence $\{x_m^k\}$ in E . Then for each x_{n_k} , we can pick a point $x_{m_k}^k$ in E being $d(x, x_{n_k})$ closed to x_{n_k}

$$x_{m_k}^k \in E \text{ and } d(x_{m_k}^k, x_{n_k}) < d(x, x_{n_k}) \quad (3.30)$$

This also give us

$$d(x_{m_k}^k, x) \leq d(x_{m_k}^k, x_{n_k}) + d(x_{n_k}, x) < 2d(x, x_{n_k}) \quad (3.31)$$

Now, we wish to prove

$$\lim_{k \rightarrow \infty} x_{m_k}^k = x \quad (3.32)$$

Because

$$\lim_{k \rightarrow \infty} x_{n_k} = x \quad (3.33)$$

For each small ϵ , we know there exists N such that

$$k > N_\epsilon \implies d(x, x_{n_k}) < \frac{\epsilon}{2} \quad (3.34)$$

Then we have

$$k > N_\epsilon \implies d(x_{m_k}^k, x) < 2d(x, x_{n_k}) < \epsilon \text{ (done)} \quad (3.35)$$

We now prove \overline{E} is the smallest closed set containing E .

Let K be a closed set containing E . We wish to show all $x \in \overline{E}$ is also in K . By definition, we know x is a limit of a sequence $\{x_n\}$ in $E \subseteq K$, then because K is closed, we know x is also in K (done) ■

Theorem 3.2.3. (Manipulation of Closure) We have

$$\overline{\overline{X}} = \overline{X} \text{ and } \overline{\overline{\emptyset}} = \overline{\emptyset} \quad (3.36)$$

$$\bigcup_{n=1}^{\infty} \overline{A_n} \subseteq \overline{\bigcup_{n=1}^{\infty} A_n} \text{ and } \overline{\bigcap_{n=1}^{\infty} A_n} \subseteq \bigcap_{n=1}^{\infty} \overline{A_n} \quad (3.37)$$

where the \subseteq may be proper. Notice that ∞ does not need to be countable.

Proof. $\overline{\overline{X}} = \overline{X}$ and $\overline{\overline{\emptyset}} = \overline{\emptyset}$ are clear. We now prove $\bigcup_{n=1}^{\infty} \overline{A_n} \subseteq \overline{\bigcup_{n=1}^{\infty} A_n}$

Let $x \in \bigcup_{n=1}^{\infty} \overline{A_n}$. We know x is a limit of some sequence $\{x_k\}$ in some A_i . Notice that the sequence $\{x_k\}$ is also in $\bigcup_{n=1}^{\infty} A_n$, so we also have $x \in \overline{\bigcup_{n=1}^{\infty} A_n}$. (done)

For a case where \subseteq is proper, well order \mathbb{Q} into $\{a_n\}_{n \in \mathbb{N}}$. We know $\overline{\{a_n\}} = \{a_n\}$, so we have

$$\bigcup_{n=1}^{\infty} \overline{A_n} = \mathbb{Q} \quad (3.38)$$

We also have

$$\overline{\bigcup_{n=1}^{\infty} A_n} = \overline{\mathbb{Q}} = \mathbb{R} \quad (3.39)$$

For a rigorous proof of $\overline{\mathbb{Q}} = \mathbb{R}$, notice for each $x \in \mathbb{R}$, the supremum of the set $\{q \in \mathbb{Q} : q < x\}$ is x .

We now prove $\overline{\bigcap_{n=1}^{\infty} A_n} \subseteq \bigcap_{n=1}^{\infty} \overline{A_n}$

Let $x \in \overline{\bigcap_{n=1}^{\infty} A_n}$. We know there is a sequence $\{x_n\}$ in $\bigcap_{n=1}^{\infty} A_n$ that converge to x . Notice that $\{x_n\}$ is a sequence in A_i for all i , so by definition of closure, the limit x is in $\overline{A_i}$ for all

i. (done)

For a case where \subseteq is proper, see the below proof for Theorem ??.

If the above are said to be a "general" statements of how closure act on infinite union and intersection. The below shows a stronger property of closure, where one and only one of the inclusion become equality.

Theorem 3.2.4. (Manipulation of Closure) We have

$$\overline{E \cup F} = \overline{E} \cup \overline{F} \text{ and } \overline{E \cap F} \subseteq \overline{E} \cap \overline{F} \quad (3.40)$$

Proof. We first prove $\overline{E \cup F} = \overline{E} \cup \overline{F}$.

$\overline{E \cup F} \subseteq \overline{E} \cup \overline{F}$ is clear, since a convergent sequence in E is a convergent sequence in $E \cup F$. To see $\overline{E} \cup \overline{F} \subseteq \overline{E \cup F}$, let x be a limit of a sequence $\{x_n\}$ in $E \cup F$. We know $\{x_n\}$ has infinite terms (a sub-sequence converging to x) $\{x_{n_k}\}$ either in E , in F or in both. Then by definition of closure, we know x is either in \overline{E} , \overline{F} or both. (done)

We now prove $\overline{E \cap F} \subseteq \overline{E} \cap \overline{F}$.

Notice that if a sequence $\{x_n\}$ is in $E \cap F$, then the sequence $\{x_n\}$ is both in E and in F . Then we know the limit is both in \overline{E} and \overline{F} . (done)

For a case where \subseteq is proper, observe

$$E = (0, 1) \text{ and } F = (1, 2) \quad (3.41)$$

We have

$$\overline{E \cap F} = \overline{\emptyset} = \emptyset \subset \{1\} = [0, 1] \cap [1, 2] = \overline{E} \cap \overline{F} \quad (3.42)$$

Theorem 3.2.5. (Manipulation of Closed Sets) Let \mathfrak{F} be the family of closed sets.

We have

$$\emptyset \text{ and } X \text{ are closed.} \quad (3.43)$$

$$\text{Intersection of closed sets are closed.} \quad (3.44)$$

$$\text{Union of finite closed sets are closed, but union of infinite closed set may not be closed.} \quad (3.45)$$

Proof. In Theorem ??, we have proved \emptyset and X are closed. Let $\{A_\lambda\}$ be an infinite collection of closed sets. Because A_λ are closed, by Theorem ??, we have

$$\bigcap_{\lambda \in \Lambda} A_\lambda \subseteq \bigcap_{\lambda \in \Lambda} \overline{A_\lambda} = \bigcap_{\lambda \in \Lambda} A_\lambda \quad (3.46)$$

And notice the other side is trivial.

Finite union of closed sets is closed is an immediate consequence of Theorem ?? by induction. To see $\bigcup_{\lambda \in \Lambda} A_\lambda$ need not always be closed. Notice that if $\bigcup_{\lambda \in \Lambda} A_\lambda$ are closed, then we have

$$\overline{\bigcup_{\lambda \in \Lambda} A_\lambda} = \bigcup_{\lambda \in \Lambda} A_\lambda = \bigcup_{\lambda \in \Lambda} \overline{A_\lambda} \quad (3.47)$$

but Theorem ?? and Theorem ?? tell us the equality between $\overline{\bigcup_{\lambda \in \Lambda} A_\lambda} = \bigcup_{\lambda \in \Lambda} \overline{A_\lambda}$ need not hold true. For a constructive counter-example, let $A_n = [\frac{1}{n}, 1]$, and observe

$$\bigcup_{n=1}^{\infty} A_n = (0, 1] \notin \mathfrak{F} \quad (3.48)$$

Notice if we let $A_n = (\frac{1}{n}, 1]$, we have a counter-example for $\bigcup_{\lambda \in \Lambda} \overline{A_\lambda} = \overline{\bigcup_{\lambda \in \Lambda} A_\lambda}$

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} (\frac{1}{n}, 1] = (0, 1] \subset [0, 1] = \overline{(0, 1]} = \overline{\bigcup_{n=1}^{\infty} (\frac{1}{n}, 1]} = \overline{\bigcup_{n=1}^{\infty} A_n} \quad (3.49)$$

We now introduce an important idea hidden behind our above proofs for each theorems.

Definition 3.2.6. (Definition of Limit Point) We say x is a limit point of E if there is a sequence $\{x_n\}$ in E such that

$$\lim_{n \rightarrow \infty} x_n = x \text{ and } \forall n, x_n \neq x \quad (3.50)$$

We denote the set of limit points of E by E' . We say a point in E that isn't a limit point is an isolated point. Also, we say a set is perfect if $E' = E$.

Theorem 3.2.7. (Property of Limit Points) We have

$$\overline{E} = E' \cup E \quad (3.51)$$

Proof. The fact $E \cup E' \subseteq \overline{E}$ is clear, as we notice for each x in E , the sequence $\{x_n = x\}$ is a sequence in E that converges to x . Let $x \in \overline{E} \setminus E$. We wish to show $x \in E'$. Let $\{x_n\}$ be a sequence in E that converge to x . Because $\{x_n\}$ is a sequence in E and x is not in E , we know $\{x_n\}$ must not contain x , and we are done. ■

Corollary 3.2.8. (Property of Limit Points) We have

$$E \text{ is closed} \iff E' \subseteq E \quad (3.52)$$

Proof. If E is closed, then $E' \cup E = \overline{E} = E$ tell us $E' \subseteq E$. If $E' \subseteq E$, then $\overline{E} = E' \cup E = E$ ■

Theorem 3.2.9. (Property of Limit Points)

$$p \in E' \iff \forall \epsilon, B_\epsilon(p) \cap E \text{ is infinite} \quad (3.53)$$

Proof. From left to right, let $\{x_n\}$ be a sequence in E that converge to p and does not contain p . For all ϵ , we know there exists N_ϵ such that $n > N_\epsilon \implies x_n \in B_\epsilon(p)$. Assume **the terms after N_ϵ is finitely distinct**. We can let $\epsilon = \min\{x_k : k > N_\epsilon\}$ and **CaC**

From right to left, we wish to construct a sequence $\{x_n\}$ in E that converge to p . For all n , by premise we know there exists a point in $B_{\frac{1}{n}}(p) \cap E$ that isn't p (because $B_\epsilon(p) \cap E$ is infinite). Select that point as x_n . Then the sequence $\{x_n\}$ does converge to p and does not contain p . ■

Theorem 3.2.10. (Property of Limit Points) We have

$$(E \cap F)' \subseteq E' \cap F' \quad (3.54)$$

$$E' \cup F' = (E \cup F)' \quad (3.55)$$

Where the \subseteq may be proper. We also have

$$E \subseteq F \implies E' \subseteq F' \quad (3.56)$$

Where the converse may not hold true. We also have

$$E' \text{ is closed.} \quad (3.57)$$

Proof. For the first two statements, use the fact that a sequence in $E \cap F$ is in E and a sequence in E is in $E \cup F$, and that a sequence in $E \cup F$ must has a sub-sequence in E or F .

For an example of \subseteq being proper, let $E = (0, 1)$ and $F = (1, 2)$. We have

$$(E \cap F)' = \emptyset \subset \{1\} = E' \cap F' \quad (3.58)$$

For the third statements, use the fact a sequence in E is in F . To see the converse may not hold true, let $E = (0, 1) \cup \{2\}$ and $F = (0, 1)$

We now prove **E' is closed**. Let $\{x_n\}_{n \in \mathbb{N}}$ be a convergent sequence in E' . Let $x = \lim_{n \rightarrow \infty} x_n$. We wish to prove $x \in E'$. Notice that we can construct a sub-sequence $\{x_{n_m}\}_{m \in \mathbb{N}}$ such that for all m , we have

$$d(x_{n_m}, x) < \frac{1}{m} \quad (3.59)$$

Remember $\{x_{n_m}\}$ converge to x , and $\{x_{n_m}\} \subseteq \{x_n\}$ is in E' . From now, we use $\{x_u\}$ to denote $\{x_{n_m}\}$ where $u = m$.

Because $\{x_u\}$ is in E' , so if $x \in \{x_u\}$, our proof is trivially done. We only have to consider when $x \notin \{x_u\}$.

Because $\{x_u\}_{u \in \mathbb{N}}$ is in E' , we know for each $u \in \mathbb{N}$, there exists a sequence $\{x_k^u\}_{k \in \mathbb{N}}$ in E that converges to x_u . Using the infinite amount of sequence $\{\{x_k^u\}_{k \in \mathbb{N}} : u \in \mathbb{N}\}$ in E , we wish to construct a sequence in E that converge to x and doesn't contain x .

Fix u . Because

$$\lim_{k \rightarrow \infty} x_k^u = x_u \quad (3.60)$$

We know there exists $N_{\frac{1}{u}}$ such that

$$k > N_{\frac{1}{u}} \implies d(x_k^u, x_u) < \frac{1}{u} \quad (3.61)$$

Notice that because the limit is unique and we are under the consideration $x \notin \{x_n\}$, it is impossible to happen that

$$\forall k > N_{\frac{1}{u}}, x_k^u = x \quad (3.62)$$

Otherwise we can see $\{x_k^n\}$ converge to x , not x_n .

Then, we know there exists some k_u such that $d(x_{k_u}^u, x_u) < \frac{1}{u}$ and $x_{k_u}^u < \frac{1}{u}$. Collect all such $x_{k_u}^u$ for all u , and we have a sequence in E that doesn't contain x . We now prove $\{x_{k_u}^u\}_{u \in \mathbb{N}}$ converge to x .

Fix ϵ . Recall how we construct $\{x_u\}$. We have

$$\forall u \in \mathbb{N}, d(x_u, x) < \frac{1}{u} \quad (3.63)$$

Observe

$$u > \frac{2}{\epsilon} \implies d(x_{k_u}^u, x) \leq d(x_{k_u}^u, x_u) + d(x_u, x) < \frac{1}{u} + \frac{1}{u} < \frac{2}{u} < \epsilon \text{ (done)} \quad (3.64)$$

■

Lastly, we verify the idea of subspace topology. (need to formulate for limit point and closure)

Theorem 3.2.11. (Subspace Topology) Let Y be a subspace of (X, d) , let $E \subseteq X$, and let $p \in Y$. We have

$$p \text{ is a limit point of } E \cap Y \text{ in } Y \implies p \text{ is a limit point of } E \text{ in } X \quad (3.65)$$

$$\text{The closure of } E \cap Y \text{ in } Y \text{ is a subset of the closure of } E \text{ in } X \quad (3.66)$$

$$E \text{ is closed in } X \implies E \cap Y \text{ is closed in } Y \quad (3.67)$$

where the converse of the first and the third statement may not hold true, and the subset relation in the second statement may be proper.

Proof. For the first statement, use the fact a sequence is in $E \cap Y$ is in E .

For a non-trivial example of of the converse in first statement mat not hold true, let $X = \mathbb{R}$, let $Y = \{1\} \cup [2, 3]$ and let $E = (0, 3)$. We see 1 and 0 are both limit points of E in X , and neither of them is a limit point of $E \cap Y$ in Y , as $1 \in Y$ is an isolated point and 0 is not even in Y .

We now prove the second statement. Let p be in the closure of $E \cap Y$, and let $\{x_n\}$ be a sequence in $E \cap Y$ that converge to p . Notice that the sequence $\{x_n\}$ is in E and we are done.

We now prove the last statement. From left to right, let $\{x_n\}$ be a sequence in $E \cap Y$ that converge to some p in Y . We wish to prove p is in $E \cap Y$. Notice that $\{x_n\}$ is also in E , so we know p is in E . Then because p is in Y , we know $p \in E \cap Y$.

For an non-trivial example of the converse in last statement may not hold true. Let $X = \mathbb{R}$, let $Y = (0, 2)$ and $E = [1, 2)$. Clearly E is not closed, but $E \cap Y = [1, 2)$ is closed in Y (One can even observe that E in Y is even "informally" unbounded).

■

3.3 Open

Definition 3.3.1. (Definition of Interior Points and Open) Let E be a subset of (X, d) . By an interior point p of E , we mean a point such every sequence $\{x_n\}$ converge to p must eventually be in E :

$$\exists N, \forall n > N, x_n \in E \quad (3.68)$$

We call the set of interior points of E interior, and denote interior by E° . We say E is open if every sequence that converge to a point in E must eventually be in E , in other words, every point of E is an interior point of E :

$$E^\circ = E \quad (3.69)$$

Theorem 3.3.2. (Another Definition of Interior)

$$E^\circ \text{ is the greatest open set in } E \quad (3.70)$$

Proof. We first prove $E^\circ \subseteq E$.

Let p be an interior point of E . Verify for the trivial sequence $\{x_n = p\}_{n \in \mathbb{N}}$. (done)

We now prove E° is open.

Let x be a point in E° , and let $\{x_n\}$ be a sequence converge to x . We wish to show $\{x_n\}$ must eventually be in E° . Assume $\{x_n\}$ doesn't end in E° . That is, there exists a sub-sequence $\{x_u\}$ of $\{x_n\}$ in $E \setminus E^\circ$.

Because $\{x_u\}$ is a sub-sequence of $\{x_n\}$, we know $\{x_u\}$ converge to x . Because of such, by Theorem ??, we know there exists a sub-sequence $\{x_r\}$ of $\{x_u\}$ such that

$$\forall r \in \mathbb{N}, d(x_r, x) < \frac{1}{r} \quad (3.71)$$

Fix r . Notice that $x_r \in \{x_u\}$, so we know x_r is not in E° . Then we know there exists a sequence $\{x_m^r\}$ converge to x_r that does not end in E .

Then by Theorem ??, we know there exists a sub-sequence $\{x_k^r\} \subseteq \{x_m^r\}$ such that

$$\forall k \in \mathbb{N}, d(x_k^r, x_r) < \frac{1}{k} \quad (3.72)$$

Now for each r , notice that $k = r$ implies

$$d(x_k^r, x) < d(x_k^r, x_r) + d(x_r, x) < \frac{1}{k} + \frac{1}{r} = \frac{2}{r} \quad (3.73)$$

Collect all such x_r^r as a sequence $\{x_r^r\}$. Notice that every term x_r^r is in the sequence $\{x_m^r\}$ which are not in E , so we know the sequence is not in E . Observe

$$\forall \epsilon, r > \frac{2}{\epsilon} \implies d(x_r^r, x) < \frac{2}{r} < \epsilon \quad (3.74)$$

We have constructed a sequence $\{x_r^r\}$ that converge to x but completely not in E **CaC** to $x \in E^\circ$. (done)

Lastly, we prove E° is the greatest open set contained by E . Assume **There exists an open set $O \subseteq E$ such that $O \not\subseteq E^\circ$** . Let $x \in O \setminus E^\circ$. Because $x \notin E^\circ$, we know there exists a sequence $\{x_n\}$ converge to x and have a sub-sequence $\{x_k\} \subseteq \{x_n\}$ not in E . Then because $x \in O$, we know the sub-sequence $\{x_k\}$ must eventually end in $O \subseteq E$ **CaC**. (done) ■

Theorem 3.3.3. (Manipulation of Interior) We have

$$X = X^\circ \text{ and } \emptyset^\circ = \emptyset \quad (3.75)$$

$$\bigcup_{n=1}^{\infty} A_n^\circ \subseteq \left(\bigcup_{n=1}^{\infty} A_n\right)^\circ \quad (3.76)$$

$$\left(\bigcap_{n=1}^{\infty} A_n\right)^\circ \subseteq \bigcap_{n=1}^{\infty} A_n^\circ \quad (3.77)$$

where the \subseteq can be proper.

Proof. The first statement is trivial. For the second and the third, use the fact a sequence that ends in some A_n will end in $\bigcup A_n$ and the fact a sequence that ends in $\bigcap A_n$ will end in each A_n .

For an example of the second statement may be proper, well order all numbers in $(0, 1)$, and let A_n contain exactly one number. Then

$$\bigcup_{n=1}^{\infty} A_n^\circ = \bigcup_{n=1}^{\infty} \emptyset = \emptyset \subset (0, 1) = \left(\bigcup_{n=1}^{\infty} A_n\right)^\circ \quad (3.78)$$

For an example of the third statement may be proper, let $A_n := (1 - \frac{1}{n}, 1 + \frac{1}{n})$. Then

$$\left(\bigcap_{n=1}^{\infty} A_n\right)^\circ = \{1\}^\circ = \emptyset \subset \{1\} = \bigcap_{n=1}^{\infty} A_n^\circ \quad (3.79)$$

■

Theorem 3.3.4. (Manipulation of Interior) We have

$$E^\circ \cup F^\circ \subseteq (E \cup F)^\circ \quad (3.80)$$

$$(E \cap F)^\circ = E^\circ \cap F^\circ \quad (3.81)$$

And if $E \subseteq F$, we have

$$E^\circ \subseteq F^\circ \quad (3.82)$$

Proof. For the first statement, use the fact a sequence that ends in E is a sequence that ends in $E \cup F$. Notice that a sequence end in $E \cup F$ need not end in neither E nor F .

To see the \subseteq may be proper, let $E = [0, 1]$ and $F = [1, 2]$. We have

$$E^\circ \cup F^\circ = (0, 1) \cup (1, 2) \neq (0, 2) = [0, 2]^\circ = (E \cup F)^\circ \quad (3.83)$$

For $(E \cap F)^\circ \subseteq E^\circ \cap F^\circ$, use the fact a sequence that ends in $E \cap F$ will end in both E and F .

For $E^\circ \cap F^\circ \subseteq (E \cap F)^\circ$, use the fact a sequence that end in both E and F will end in $E \cap F$. In more rigor, let N_E satisfy

$$n > N_E \implies x_n \in E \quad (3.84)$$

and let N_F satisfy

$$n > N_F \implies x_n \in F \quad (3.85)$$

Then we see

$$n > \max\{N_E, N_F\} \implies x_n \in E \cap F \quad (3.86)$$

For the last statement, use the fact a sequence that ends in E is a sequence that ends in F ■

Theorem 3.3.5. (Manipulation of Open Sets) We have

$$X \text{ and } \emptyset \text{ are open.} \quad (3.87)$$

$$\text{Union of open sets are open.} \quad (3.88)$$

$$\text{Finite intersection of open sets are open} \quad (3.89)$$

$$\text{Infinite intersection of open sets may not be open.} \quad (3.90)$$

Proof. In Theorem ??, we have proved X and \emptyset are open. Let $\{A_\lambda : \lambda \in \Lambda\}$ be a collection of open sets. By Theorem ??, we know

$$\bigcup_{\lambda \in \Lambda} A_\lambda = \bigcup_{\lambda \in \Lambda} A_\lambda^\circ \subseteq \left(\bigcup_{\lambda \in \Lambda} A_\lambda\right)^\circ \quad (3.91)$$

The other side follows from Theorem ??.

Finite intersection of open sets are open follows from Theorem ??:

$$(E \cap F)^\circ = E^\circ \cap F^\circ = E \cap F \quad (3.92)$$

For an example of infinite intersection of open sets may not be open, let $A_n = (1 - \frac{1}{n}, 1 + \frac{1}{n})$. We have $\bigcap_{n=1}^{\infty} A_n = \{1\}$ ■

We now introduce some interaction between interior and closure operator.

Theorem 3.3.6. (Interaction of Interior and Closure) We have

$$(E^\circ)^c = \overline{E^c} \quad (3.93)$$

Proof. $x \in \overline{E^c}$ means there exists a sequence in E^c converge to x . That sequence has no term in E , so clearly does not end in E . This show $x \notin E^\circ$, which means $x \in (E^\circ)^c$.

$x \notin E^\circ$ means there exists a sequence $\{x_n\}$ that converges to x and does not end in E . We know $\{x_n\}$ has a sub-sequence $\{x_k\}$ not in E that converges to x . The sequence $\{x_k\}$ is in E^c , and converge to x , so we know $x \in \overline{E^c}$ ■

Theorem 3.3.7. (Non-relation)

$$\overline{E^\circ} \subseteq \overline{E} \text{ and } E^\circ \subseteq (\overline{E})^\circ \quad (3.94)$$

Are true, but the \subseteq can be proper. Also, $\overline{E^\circ}$ and $(\overline{E})^\circ$ are incomparable.

Proof. For $\overline{E^\circ} \subseteq \overline{E}$ and $E^\circ \subseteq (\overline{E})^\circ$, let $E = \mathbb{Q}$. We have

$$\overline{E^\circ} = \overline{\emptyset} = \emptyset \subset \mathbb{R} = \overline{\mathbb{Q}} = \overline{E} \quad (3.95)$$

And we have

$$E^\circ = \emptyset \subset \mathbb{R} = (\mathbb{R})^\circ = (\overline{\mathbb{Q}})^\circ = (\overline{E})^\circ \quad (3.96)$$

For $\overline{E^\circ} \not\subseteq (\overline{E})^\circ$, let $E = [0, 1]$. We have

$$\overline{E^\circ} = [0, 1] \not\subseteq (0, 1) = (\overline{E})^\circ \quad (3.97)$$

For $(\overline{E})^\circ \not\subseteq \overline{E^\circ}$, let $E = \mathbb{Q}$. We have

$$(\overline{E})^\circ = \mathbb{R} \not\subseteq \emptyset = \overline{E^\circ} \quad (3.98)$$

We now shed light on subspace topology of sequential space.

Theorem 3.3.8. (Subspace Topology) Let Y be a subspace of (X, d) , let $E \subseteq X$, and let $p \in Y$. We have

$$p \text{ is an interior point of } E \text{ in } X \implies p \text{ is an interior point of } E \cap Y \text{ in } Y \quad (3.99)$$

$$Y \cap E^\circ \text{ in } X \text{ is a subset of the interior of } E \cap Y \text{ in } Y \quad (3.100)$$

$$E \text{ is open in } X \implies E \cap Y \text{ is open in } Y \quad (3.101)$$

where the converse may not hold true.

Proof. We first prove the first statement. Let $\{x_n\}$ be a sequence in Y that converge to p . Because p is an interior point of E in X , and $\{x_n\}$ is in X , as $Y \subseteq X$, we know there exists N such that

$$n > N \implies x_n \in E \quad (3.102)$$

Notice $x_n \in Y$, and we are done.

For a nontrivial example of the converse of the first statement may not hold true, let $E = (0, 2)$, let $Y = \{1\} \cup (2, 3)$. We see 1 is an interior point of E in \mathbb{R} , but 1 isn't an interior point of $\{1\} = E \cap Y$ in Y .

The second and the third statement follows from the first statement. ■

We now show that metric space is sequential.

Theorem 3.3.9. (Open Balls are Sequentially Open)

$$B_r(p) \text{ is open} \quad (3.103)$$

Proof. Let $x \in B_r(p)$. We wish to show every sequence $\{x_n\}$ that converge to x ends in $B_r(p)$. Let $u = r - d(p, x)$. Notice that $B_u(x) \subseteq B_r(p)$ and that $\{x_n\}$ ends in $B_u(x)$, and we are done. ■

Theorem 3.3.10. (Metric Spaces are Sequential)

$$p \text{ is an interior point of } E \quad (3.104)$$

$$\iff \text{there exists an small enough open ball } B_r(p) \text{ contained by } E \quad (3.105)$$

Proof. We first prove from left to right. Assume **every open ball $B_r(p)$ is too big to be contained by E** . That is:

$$\forall \epsilon, B_\epsilon(p) \setminus E \neq \emptyset \quad (3.106)$$

Then for each n , we let $x_n \in B_{\frac{1}{n}}(p) \setminus E$. Then $\{x_n\}$ is a sequence outside of E that converge to p **CaC** to $p \in E^\circ$.

We now prove from right to left. Let $\{x_n\}$ be a sequence converge to p . We know there exists N such that

$$n > N \implies x_n \in B_r(p) \subseteq E \quad (3.107)$$

This finish our proof. ■

Corollary 3.3.11. (Metric Spaces are Sequential)

$$E \text{ is open} \quad (3.108)$$

$$\iff \text{for each point } p \text{ in } E \text{ there exists an small enough open ball } B_r(p) \text{ contained by } E \quad (3.109)$$

Lastly, the most important.

Theorem 3.3.12. (Topology Axiom)

$$E \text{ is open} \iff E^c \text{ is closed} \quad (3.110)$$

Proof. We first prove from left to right. Let $\{x_n\}$ be a sequence in E^c that converge to x . We wish to show $x \in E^c$. Assume $x \in E$. Because E is open, we know $\{x_n\}$ eventually ends in E CaC .

We now prove from right to left. Let $x \in E$, and let $\{x_n\}$ be a sequence converge to x . We wish to show $\{x_n\}$ eventually ends in E . Because $x \notin E^c$, and E^c is closed, we know $\{x_n\}$ has no sub-sequence in E^c , otherwise we can see that sub-sequence in E^c converge to $x \notin E^c$.

Then we know there exists only finite terms in E^c . Let N be the maximal index of all finite terms in E^c . We see

$$n > N \implies x_n \notin E^c \implies x_n \in E \quad (3.111)$$

■

3.4 Connected

Definition 3.4.1. (Definition of Separated) We say two sets A, B are separated if

$$A \cap \overline{B} = \emptyset = \overline{A} \cap B \quad (3.112)$$

Theorem 3.4.2. (Subsets of Separated Sets are Separated) Let $A \subseteq N$ and $B \subseteq M$.

$$N, M \text{ are separated} \implies A, B \text{ are separated} \quad (3.113)$$

Proof. Notice

$$\overline{A} \subseteq \overline{N} \subseteq M^c \subseteq B^c \quad (3.114)$$

And

$$\overline{B} \subseteq \overline{M} \subseteq N^c \subseteq A^c \quad (3.115)$$

■

Theorem 3.4.3. (Subspace Topology) Let Y be a subspace of (X, d) .

$$A, B \text{ are separated} \implies A \cap Y \text{ and } B \cap Y \text{ are separated in } Y \quad (3.116)$$

The converse may not hold true.

Proof. In Theorem ??, we have proved the closure of $A \cap Y$ in Y is a subset of the closure of A in X . Then because A, B are separated, we know the closure of $A \cap Y$ in Y is disjoint to B , thus disjoint to $B \cap Y$. The other side is similar.

To see the converse may not hold true, let $Y = (0, 1) \cup (2, 3)$, and let $A = (0, 2)$ and $B = (1, 3)$. ■

Theorem 3.4.4. (Equivalent Definition of Connected) The following are equivalent

$$E \text{ is not a union of two nonempty separated sets.} \quad (3.117)$$

$$\iff E \text{ has no proper nontrivial clopen subset in subspace topology for } E. \quad (3.118)$$

Proof. We prove

$$E \text{ has a proper non-trivial clopen subset in subspace topology for } E \quad (3.119)$$

$$\iff E \text{ is a union of two non-empty separated sets} \quad (3.120)$$

From **left to right**, let $M \subseteq E$ be clopen in subspace topology of E . We wish to show M and $E \setminus M$ are separated, and we are done. Assume **M and $E \setminus M$ are not separated**. WOLG, let $x \in M' \cap (E \setminus M)$.

Because $x \in E \setminus M$ and $E \setminus M$ is open in subspace topology of E , we know there exists open ball $B_r(x)$ such that $B_r(x) \cap E \subseteq E \setminus M$. Because $x \in M'$, we know $B_r(x)$ contain infinite points in M **CaC** to $B_r(x) \cap E \subseteq E \setminus M$. **(done)**

We now prove from **right to left**. Let A and B be separated and $E = A \cup B$. We wish to show A is clopen in subspace topology of E . Notice we have

$$A = \overline{A} \cup E \quad (3.121)$$

so we know A is closed in subspace topology of E . Because we have

$$A \subseteq \overline{B}^c \quad (3.122)$$

So we have

$$A = \overline{B}^c \cap E \quad \text{(done)} \quad (3.123)$$

■

Theorem 3.4.5. (Union of Connected Sets that have Nonempty Intersection is Connected) Let \mathcal{F} be a class of connected sets. We have

$$\bigcap \mathcal{F} \neq \emptyset \implies \bigcup \mathcal{F} \text{ is connected} \quad (3.124)$$

Proof. Assume $\bigcup \mathcal{F}$ is not connected. Let

$$\bigcup \mathcal{F} = A \dot{\cup} B \text{ and } A \neq \emptyset \neq B \quad (3.125)$$

And let A, B be relatively open to $\bigcup \mathcal{F}$. We know $\bigcap \mathcal{F}$ must intersect with either A , or B , or both.

WOLG, let

$$A \cap \bigcap \mathcal{F} \neq \emptyset \quad (3.126)$$

Because B is non-empty and $B \subseteq \bigcup \mathcal{F}$, we know B must intersect with some $F_n \in \mathcal{F}$. Notice that because $A \cap \bigcap \mathcal{F} \neq \emptyset$, we have $A \cap F_n \neq \emptyset$. Then by Theorem ??, we see $A \cap F_n$ and $B \cap F_n$ are both relatively open to F_n , while $F_n = (A \cap F_n) \dot{\cup} (B \cap F_n)$ **CaC** ■

3.5 Metric Space: Open

Definition 3.5.1. (Definition of Open Ball) Let (X, d) be a metric space, we say the set

$$B_r(x) = \{y \in X : d(x, y) < r\} \quad (3.127)$$

is the open ball of radius r around x

Theorem 3.5.2. (Basic Property of Open Ball) If $r \leq h$, then $B_r(x) \subseteq B_h(x)$.

Give clear notice that although we use the word "open" ball, we have not proved "open" ball is open. In fact, we have not even defined what would be open.

Notice in \mathbb{R}^1 , the open ball $B_r(x)$ is just the interval $(x - r, x + r)$, and the interval (a, b) is the open ball $B_{\frac{b-a}{2}}(\frac{a+b}{2})$. Also notice $B_0(x) \neq (x, x) = \emptyset$.

The following first define interior points, then define open set from definition of interior points.

Definition 3.5.3. (Definition of Interior Points and Open) A point p is an interior point of E if

$$\exists r \in \mathbb{R}^+, B_r(x) \subseteq E \quad (3.128)$$

The set of interior points of E is called interior E° of E . Also, we say E is open if

$$E^\circ = E \quad (3.129)$$

Theorem 3.5.4. (Basic Property of Interior) For all set E , we have $E^\circ \subseteq E$

Now, we prove an "open" ball is indeed open. The proof of such should of course be completely rigorous, but to come up with a proof rely more than rigor. One can and should try to use geometric intuition to form a proof, and try to rigorously formulate the proof to verify whether the idea also hold true in all metric spaces.

Notice Theorem ?? will prove that our definition of open do satisfy the axioms of open sets. In other words, every metric induce a topology.

Theorem 3.5.5. (Open Ball is Open) Open ball is open.

Proof. For each point $p \in B_r(x)$ with distance to the center r' , we can see the ball of radius $r - r'$ centering p is contained by $B_r(x)$. ■

Theorem 3.5.6. (Compatibility of Definitions of Open) Let \mathfrak{O} be the family of open set, by our definition of open in metric space X , we have

$$\emptyset, X \in \mathfrak{O} \quad (3.130)$$

$$A, B \in \mathfrak{O} \implies A \cap B \in \mathfrak{O} \quad (3.131)$$

$$\mathfrak{A} \subseteq \mathfrak{O} \implies \bigcup \mathfrak{A} \in \mathfrak{O} \quad (3.132)$$

Proof. Notice $\emptyset^\circ = \emptyset$. Every open ball centering any point is clearly contained by X .

For any point in the intersection of two open sets, we by definition have two open balls where each of them is in A or B . Pick the smaller one and we are done.

For each point p in an union of open sets, we know there exists an open set O in the union containing p . By definition we have an open ball centering p contained by O . This open ball is contained by the union. ■

Now, we present some example of open sets in different metric spaces. Give clear notice that some open set in a sub-metric space may not be open in the larger metric space.

Theorem 3.5.7. (Examples of Open Sets in Euclidean Metric Space) The followings are open if not specified to be not open

$$(-\infty, a) \text{ in } \mathbb{R} \quad (3.133)$$

$$(-\infty, a] \text{ is not open in } \mathbb{R} \quad (3.134)$$

$$(a, b) \text{ in } \mathbb{R} \quad (3.135)$$

$$[a, b) \text{ is not open in } \mathbb{R} \quad (3.136)$$

$$(a, b) \times \{c\} \text{ is not open in } \mathbb{R}^2 \quad (3.137)$$

$$(a, b) \times (c, d) \text{ in } \mathbb{R}^2 \quad (3.138)$$

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \neq 1\} \text{ (}\mathbb{R}^2 \text{ minus a sphere)} \quad (3.139)$$

$$\mathbb{R}^2 \text{ in } \mathbb{R}^2 \quad (3.140)$$

$$\mathbb{Q} \text{ in } \mathbb{Q} \quad (3.141)$$

$$\mathbb{Q} \text{ is not open in } \mathbb{R} \quad (3.142)$$

$$(1, 2) \cap \mathbb{Q} \text{ in } \mathbb{Q} \quad (3.143)$$

$$[1, 2] \cap \mathbb{Q} \text{ is not open in } \mathbb{Q} \quad (3.144)$$

$$(1, 2) \cap \mathbb{Q} \text{ is not open in } \mathbb{R} \quad (3.145)$$

Theorem 3.5.8. (Example: Discrete Metric) Discrete metric is defined as

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases} \quad (3.146)$$

every set in discrete metric space is open.

Proof. Every ball is either the whole space or the center, depending on if the radius is greater than 1. Pick any ball of radius smaller than or equal to 1 and we are done. ■

Theorem 3.5.9. (Subspace Topology) Let (X, d) be a metric space and let $(Y, d_Y) \subseteq (X, d)$, where d_Y is the restriction of d . Let τ be the standard topology for X . The standard topology τ_Y for (Y, d_Y) is exactly

$$\tau_Y = \{O \cap Y : O \in \tau\} \quad (3.147)$$

Proof. In this proof, in metric space (Y, d^Y) , we will denote the open ball centering p of radius r as $B_r^Y(p)$. In symbol

$$B_r^Y(p) := B_r(p) \cap Y \quad (3.148)$$

Arbitrarily pick O from τ . We wish to prove $O \cap Y$ is open in Y . Because O is open, for each $p \in O \cap Y$, we know there exists an open ball $B_r(p)$ contained by O . Then because $B_r^Y(p) \subseteq B_r(p)$ and $B_r^Y(p) \subseteq Y$, we have

$$B_r^Y(p) \subseteq B_r(p) \cap Y \subseteq O \cap Y \text{ (done)} \quad (3.149)$$

Arbitrarily pick O_Y from τ_Y . We wish to prove

$$\exists O \in \tau, O_Y = O \cap Y \quad (3.150)$$

For each point $p \in O_Y$, because O_Y is open in Y , we know there exists an radius r_p small enough so that the open ball $B_{r_p}^Y(p)$ is contained by O_Y . Collect all such r_p , and let

$$O := \bigcup_{p \in O_Y} B_{r_p}(p) \quad (3.151)$$

We know O is open in X , we wish to prove

$$O_Y = O \cap Y \quad (3.152)$$

Recall the definition of O . It is clear that $O_Y \subseteq O \cap Y$. Let $x \in O \cap Y$. Because $x \in O$, we know there exists $q \in O_Y$ such that

$$x \in B_{r_q}(q) \quad (3.153)$$

then because $x \in Y$, we know

$$x \in B_{r_q}(q) \cap Y = B_{r_q}^Y(q) \subseteq O_Y \text{ (done) (done)} \quad (3.154)$$

■

Theorem 3.5.10. (General Sufficient Condition for Open)

$$E \text{ is finite or } E^c \text{ is finite} \implies E \text{ is open} \quad (3.155)$$

Proof. Verify for $B_{\min\{d(p,y):y \in E^c\}}(p)$ ■

3.6 Metric Space: Closed and Limit Point

Definition 3.6.1. (Definition of Limit Points and Closed) A point p is a limit point of E if

$$\forall r \in \mathbb{R}^+, \exists q \neq p, q \in B_r(p) \cap E \quad (3.156)$$

or equivalently

$$\forall r \in \mathbb{R}^+, B_r(p) \cap E \setminus \{p\} \neq \emptyset \quad (3.157)$$

The set of limit points of E is denoted E' . We say E is closed if

$$E' \subseteq E \quad (3.158)$$

Theorem 3.6.2. (Property of Limit Points) Let $E \subseteq F$. We have

$$E' \subseteq F' \quad (3.159)$$

Proof. If p is a limit point of E , then we know every open ball centering p non-trivially intersect with E , which implies non-trivially intersecting with F . ■

Theorem 3.6.3. (Property of Limit Points) If $p \in E'$, then every open neighborhood around p non-trivially intersect with E .

Proof. Notice that an open neighborhood around p must contain an open ball centering p . ■

Notice Theorem ?? is written in language of general topology. Now, we prove that our definition of closed and limit point in metric space is compatible to those in general topology. Then we can apply the result we have in general topology into metric space.

Theorem 3.6.4. (Compatibility of Definition of Limit Point) A point p is a limit point of E in metric space (X, d) if and only if p is a limit point of E in topology induced by (X, d)

Proof. From left to right, notice every open neighborhood around p contain an open ball centering p . From right to left, notice every open ball is an open neighborhood. ■

Theorem 3.6.5. (Compatibility of Definition of Open and Closed) Let $\mathfrak{O}, \mathfrak{F}$ respectively be the family of open sets and family of closed set under our metric definition. We have

$$E \in \mathfrak{O} \iff E^c \in \mathfrak{F} \quad (3.160)$$

Proof. From left to right, let p be a limit point of E^c . If p is not in E^c , we know there exists an open ball centering p contained in E . Meanwhile, because p is a limit point of E^c , we know every open ball centering p contain a point $q \neq p$ in E^c . Then we can deduce p must

be in E^c .

From right to left, let $p \in E$. Because p is not a limit point of E^c , we know there exists an open ball centering p disjoint E^c , that is, contained by E . ■

Corollary 3.6.6. (Axioms for Closed Sets) Let \mathfrak{F} be the family of closed sets under our metric definition. We have

$$X, \emptyset \in \mathfrak{F} \quad (3.161)$$

$$A, B \in \mathfrak{F} \implies A \cup B \in \mathfrak{F} \quad (3.162)$$

$$\mathfrak{B} \subseteq \mathfrak{F} \implies \bigcap \mathfrak{B} \in \mathfrak{F} \quad (3.163)$$

Notice Corollary ?? is proved using language of general topology in the first section of this chapter. I personally haven't tried to prove Corollary ?? using limit point definition, but I think this won't be difficult.

Again, we now give some examples of closed set. One can see that the proof for some sets are closed if following straight from the limit point definition is extremely long: One have to first find the set of limit points, by checking every points whether for all real number a property stand, and then check if the set of limit points is a subset of E . A possible optimal way is proving the complement is open, since proving something is open is more.

Theorem 3.6.7. (Example of Closed Sets in Euclidean Metric Space) The followings are closed set if not specified.

$$[a, b] \text{ in } \mathbb{R} \quad (3.164)$$

$$[a, b) \text{ is not closed in } \mathbb{R} \quad (3.165)$$

$$[a, \infty) \text{ in } \mathbb{R} \quad (3.166)$$

$$(a, \infty) \text{ is not closed in } \mathbb{R} \quad (3.167)$$

$$[a, b] \times [c, d] \text{ in } \mathbb{R}^2 \quad (3.168)$$

$$\{c\} \text{ in } \mathbb{R} \quad (3.169)$$

$$[a, b] \times [c, c] \text{ in } \mathbb{R}^2 \quad (3.170)$$

$$[a, b) \times [c, d] \text{ is not closed in } \mathbb{R}^2 \quad (3.171)$$

$$[a, b) \times [c, c] \text{ is not closed in } \mathbb{R}^2 \quad (3.172)$$

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \text{ A circle} \quad (3.173)$$

$$[0, 1] \cap \mathbb{Q} \text{ is closed in } \mathbb{Q} \text{ but not in } \mathbb{R} \quad (3.174)$$

$$\text{finite set in } \mathbb{R}^n \quad (3.175)$$

$$\mathbb{Z} \text{ in } \mathbb{R} \quad (3.176)$$

$$\mathbb{R}^2 \text{ in } \mathbb{R}^2 \quad (3.177)$$

Theorem 3.6.8. (Example: Discrete Metric) Every set in discrete metric space is closed.

Theorem 3.6.9. (Subspace Topology) Let $Y \subseteq X$. If p is a limit point of $E \cap Y$ in Y , then p is a limit point of E in X .

Corollary 3.6.10. (Subspace Topology) Let (X, d) be a metric space and let $(Y, d_Y) \subseteq (X, d)$, where d_Y is the restriction of d . Let τ^c be the family of closed set with respect to X , and let τ_Y^c be the family of closed set with respect to Y . We have

$$\tau_Y^c = \{F \cap Y : F \in \tau^c\} \quad (3.178)$$

Proof. If $p \in Y$ is a limit point of $F \cap Y$ in Y , then p is a limit point of F , which implies $p \in F$. We have proved $\{F \cap Y : F \in \tau^c\} \subseteq \tau_Y^c$. We now prove

$$\tau_Y^c \subseteq \{F \cap Y : F \in \tau^c\} \quad (3.179)$$

For each $F_Y \in \tau_Y^c$, we know $Y \setminus F_Y$ is open with respect to Y . Then we know there exists open $O \in \tau$ such that $O \cap Y = Y \setminus F_Y$. Then we have $O^c \cap Y = F_Y$ (done) ■

Theorem 3.6.11. (General Sufficient Condition for Closed) If a set E is finite, it is closed.

3.7 Metric Space: Cauchy Sequence

We first introduce sequence and convergence.

Definition 3.7.1. (Definition of Convergence in Metric Space) Let X be a metric space. We say $\{a_n\}$ converge to a

$$\lim_{n \rightarrow \infty} a_n = a \quad (3.180)$$

If

$$\forall \epsilon \in \mathbb{R}^+, \exists N_\epsilon \in \mathbb{N}, n > N_\epsilon \implies d(a_n, a) < \epsilon \quad (3.181)$$

We now introduce two properties concerning limit and convergence.

Theorem 3.7.2. (In Metric Space, Sequence Converge to At Most One Point) Let $\{a_n\}$ be a sequence in (X, d) . We have

$$\lim_{n \rightarrow \infty} a_n = a \text{ and } \lim_{n \rightarrow \infty} a_n = a' \implies a' = a \quad (3.182)$$

Proof. Consider when $\epsilon = \frac{d(a, a')}{2}$ ■

Theorem 3.7.3. (Limit Point is Limit of Some Sequence) Let $E \subseteq X$ and let p be a limit point of E . There exists a sequence $\{p_n\}$ in E such that

$$\lim_{n \rightarrow \infty} p_n = p \quad (3.183)$$

Proof. For each $k \in \mathbb{N}$, we know there exists point $q \in E$ in open ball $B_{\frac{1}{k}}(p)$. For each $k \in \mathbb{N}$, let such q be p_k . ■

Theorem 3.7.4. (Basic Property for Convergence Sequence) Let X be a metric space. We have

$$\lim_{n \rightarrow \infty} a_n = a \iff \text{every open ball } B_r(a) \text{ contain all but finitely many term of } \{a_n\} \quad (3.184)$$

Proof. From left to right is trivial. From right to left, let k be the greatest natural such that $a_k \in \{a_n\} \cap B_r(a)$. Use $N = k + 1$ and we are done. ■

We now introduce two equivalent definition of bounded.

Definition 3.7.5. (Definition of Bounded) We say a set E is bounded if there exists some open ball $B_r(p)$ that contain E .

Theorem 3.7.6. (Equivalent Definition of Bounded)

$$E \text{ is bounded} \iff \exists r \in \mathbb{R}^+, \forall p, q \in E, d(p, q) < r \quad (3.185)$$

Proof. From left to right, notice that every two points in $B_r(p)$ is of distance smaller than $2r$. From right to left, verify $B_r(p)$. ■

Theorem 3.7.7. (Converge Implies Bounded)

$$\lim_{n \rightarrow \infty} a_n = a \implies \{a_n\} \text{ is bounded} \quad (3.186)$$

Proof. Fix r . Because $\{a_n\}$ converge to a , there exists $N_\epsilon \in \mathbb{N}$ such that $n > N_r \implies a_n \in B_r(a)$. We know the ball $B_r(a)$ contain all but finitely many points of $\{a_n\}$. Then we know there exists a point in $\{a_n\}$ that is most far from a . Denote the most far point from a by a_k . Observe that $B_{d(a,a_k)+1}(a)$ contain $\{a_n\}$. ■

We now introduce Cauchy Sequence

Definition 3.7.8. (Definition of Cauchy Sequence) We say $\{a_n\}$ is a Cauchy sequence if

$$\forall \epsilon \in \mathbb{R}^+, \exists N_\epsilon \in \mathbb{N}, m, k > N_\epsilon \implies d(m, k) < \epsilon \quad (3.187)$$

Theorem 3.7.9. (Every Convergent Sequence is Cauchy) We have

$$\{a_n\} \text{ is convergent} \implies \{a_n\} \text{ is Cauchy} \quad (3.188)$$

Proof. Let $\{a_n\}$ converge to a . Notice every two point is ϵ -close in $B_{\frac{\epsilon}{2}}(a)$. ■

3.8 Metric Space: Continuous Function

Definition 3.8.1. (Limit of a Function) Let $f : X \rightarrow Y$

Definition 3.8.2. (Definition of Continuous at a Point) We say $f : X \rightarrow Y$ is continuous at a if

$$f(a) \text{ is defined and } \forall \epsilon \in \mathbb{R}^+, \exists \delta \in \mathbb{R}^+, x \in B_\delta(a) \implies d(f(x), f(a)) < \epsilon \quad (3.189)$$

3.9 Metric Space: Interior, Closure and Limit Points

Definition 3.9.1. (Definition of Closure) We define the closure \overline{E} of E to be

$$\overline{E} = E' \cup E \quad (3.190)$$

Theorem 3.9.2. (Property of Limit Points) If p is not in \overline{E} , then there exists $B_r(p)$ disjoint E .

Proof. Because p is not in E' , we know there exists an open ball $B(p)$ not intersecting with E at everywhere except possibly p . We know p is not in E , so we know $B(p)$ is disjoint with E . ■

Theorem 3.9.3. (Property of Closure and Interior) Let $E \subseteq F$. We have

$$E^\circ \subseteq F^\circ \text{ and } \overline{E} \subseteq \overline{F} \quad (3.191)$$

Proof. For $E^\circ \subseteq F^\circ$, notice an open ball contained by E is also contained by F . For $\overline{E} \subseteq \overline{F}$, notice $E' \subseteq F'$, by Theorem ?? ■

Theorem 3.9.4. (Property (definition) of Closure and Interior)

$$\overline{E} \text{ is the smallest closed set containing } E \quad (3.192)$$

$$E^\circ \text{ is the largest open subset of } E \quad (3.193)$$

Proof. We first prove \overline{E} is closed.

Let p be a limit point of \overline{E} . We know every open ball $B_r(p)$ contain a limit point of E or a point in E . If $B_r(p)$ contain a limit point q of E , we know the open ball $B_{r-d(q,p)}(q) \subseteq B_r(p)$ contain a point in E . (done) .

Let F be a closed set containing E . We can deduce

$$E' \subseteq F' \subseteq F \implies \overline{E} \subseteq \overline{F} \quad (3.194)$$

Let p be an interior point of E . We know there exists an open ball $B_r(p)$ contained in E . If $B_r(p)$ contain a point q not in interior, then we see $B_{r-d(q,p)}(q) \subseteq B_r(p)$ contain point not in E , which is impossible.

Let O be an open set contained by E , if O contain any point q not in interior, then there exists an open ball $B_r(q)$ contained in O and containing a point not in E , which is impossible. ■

Corollary 3.9.5. (Property of Closure)

$$E \text{ is closed} \iff \overline{E} = E \quad (3.195)$$

$$E \text{ is open} \iff E^\circ = E \quad (3.196)$$

Also, notice Theorem ?? have been proved in language of general topology, and here we give a simple proof in language of metric.

Lastly, we show some result in metric topology.

Theorem 3.9.6. (A Point is Limit Point if and only if All Open Ball Contain Infinite Point of E)

Proof. From right to left, notice all open $B_r(p)$ contain infinite points of E implies each open ball $B_r(p)$ contain some point in E that isn't p .

From left to right, let p be a limit point of E . If an open ball $B_r(p)$ contain only finite amount of point in E , we can let $r' = \min\{d(p, q) : q \in B_r(p) \cap E\}$ and verify $B_{r'}(p)$ contain no point in E , which is impossible. ■

Corollary 3.9.7. (Finite Points Set In Metric Topology has no Limit Points)**Corollary 3.9.8. (Finite Points Set In Metric Topology is Always Clopen)**

Proof. For closed, we can prove the complement is open or use above corollary. For open, verify for $B_{\min\{d(x, y) : y \in O\}}(x)$. ■

Theorem 3.9.9. (Counter Example of Above Statement in General Topology)

Let $X = \{a, b, c\}$ and define $\mathfrak{D} := \{\emptyset, X\}$. We have

$$\{a, b\}' = X \quad (3.197)$$

Not only does $\{a, b\}$ has limit point, also the neighborhood X around limit point contain only finite points. Notice $\{a, b\}$ is neither open nor closed.

3.10 Metric Space: Compact, Countably Compact and Limit Point Compact

There are three tools we need to prove the MEGA TFAE.

Theorem 3.10.1. (In Metric Space, Existence of Countable Base is Equivalent to Separable) Let (X, d) be a metric space.

$$X \text{ has a countable base} \iff X \text{ is separable} \quad (3.198)$$

Proof. Because by Corollary ??, X is separable if X has a countable base, we only have to prove separable implies existence of countable base. Let X be separable. Then we have a countable dense subset E of X . We now prove

$$\mathcal{O} := \{B_r(p) : p \in E \text{ and } r \in \mathbb{Q}\} \quad (3.199)$$

is a countable base.

Notice $\mathcal{O} = \bigcup \{\{B_r(p) : r \in \mathbb{Q}\} : p \in E\}$ is a countable union of countable sets, so \mathcal{O} is countable. We only have to prove \mathcal{O} is a base.

For each open neighborhood O around $x \in X$, by definition of open in metric space, we know there exists an open ball $B_r(x)$ contained by O . Because E is dense, we know x either in E or in E' . If x is in E , we can pick a positive rational n smaller than r , so we have

$$x \in B_n(x) \subseteq B_r(x) \subseteq O \text{ and } B_n(x) \in \mathcal{O} \quad (3.200)$$

and we are done. If $x \in E'$, we know there exists a point q in $E \cap B_{\frac{r}{2}}(x)$. Then we can pick a rational n between $d(q, x)$ and $\frac{r}{2}$. Now, observe

$$x \in B_n(q) \subseteq B_r(x) \subseteq O \text{ and } B_n(q) \in \mathcal{O} \quad (3.201)$$

and we are done. (done) ■

Lemma 3.10.2. (Existence of Intuitively Small Enough Open Ball) Let $p \in B_r(q)$. There exists an open ball $B_{r'}(p)$ small enough to be contained by $B_r(q)$ and not containing q

Proof. If $d(p, q) \leq \frac{r}{2}$, we can let $r' = d(p, q)$. Clearly $B_{r'}(p)$ doesn't contain q , and for each $x \in B_{r'}(p)$ we have

$$d(x, q) \leq d(x, p) + d(p, q) \leq r' + \frac{r}{2} \leq r \quad (3.202)$$

If $d(p, q) \geq \frac{r}{2}$, we can let $r' = r - d(p, q)$. Clearly $B_{r'}(p) \subseteq B_r(q)$, and we can deduce

$$2d(p, q) \geq r \implies d(p, q) \leq r - d(p, q) = r' \quad (3.203)$$

so we have $q \in B_{r'}(p)$. In short, we use the open ball $B_{\min\{d(p, q), r - d(p, q)\}}(p)$ ■

Theorem 3.10.3. (In Metric Space, Limit Point Compact Implies Separable)

Let K be a metric space.

$$K \text{ is limit point compact} \implies K \text{ is separable} \quad (3.204)$$

Proof. Let $\delta > 0$ and let $a_{\delta,1} \in K$. We now prove there exists a finite set $A_\delta = \{a_{\delta,1}, a_{\delta,2}, \dots, a_{\delta,n}\}$ such that $K \subseteq \bigcup \{B_\delta(a_{\delta,i}) : a_{\delta,i} \in A_\delta\}$.

To prove such, we construct A_δ by adding element $a_{\delta,i}$ such that $\forall j < i, d(a_{\delta,j}, a_{\delta,i}) \geq \delta$ to $\{a_{\delta,1}\}$ until impossible. Assume such finite set A_δ can not be constructed using this method. In other words, we assume our method of adding points can go on infinitely. Then we have set $A^\mathbb{N} = \{a_{\delta,i} : i \in \mathbb{N}\}$ such that

$$\forall a_{\delta,i}, a_{\delta,j} \in A^\mathbb{N}, d(a_{\delta,i}, a_{\delta,j}) \geq \delta \quad (3.205)$$

Arbitrarily pick $p \in K$, we wish to find an open ball $B_r(p)$ small enough so that

$$B_r(p) \cap A^\mathbb{N} \subseteq \{p\} \quad (3.206)$$

Let

$$E = \bigcup \{B_\delta(a_{\delta,i}) : a_{\delta,i} \in A^\mathbb{N}\} \quad (3.207)$$

Notice that each ball $B_\delta(a_{\delta,i})$ must not contain any other $a_{\delta,j}$.

If p is in E , we know p is contained by some ball $B_\delta(a_{\delta,i})$. If $p = a_{\delta,i}$ then $B_\delta(p)$ is small enough. If $p \neq a_{\delta,i}$, we wish to find an open ball $B_r(p)$ small enough so that it neither contain $a_{\delta,i}$ nor any other $a_{\delta,j}$. Notice that if the ball $B_r(p)$ is contained by $B_\delta(a_{\delta,i})$, then $B_r(p)$ contain no $a_{\delta,j}$. By Lemma ??, we have found an open ball small enough to satisfy our need.

If p is not in E , clearly $B_\delta(p)$ is small enough.

We have proved $A^\mathbb{N}$ has no limit point CaC to K is limit point compact. (done)

Let A_δ be the finite set from violet part. Consider the collection

$$A = \bigcup \{A_1, A_{\frac{1}{2}}, A_{\frac{1}{3}}, \dots\} \quad (3.208)$$

Because A is a countable union of finite sets, we know A is countable. We now prove A is dense.

Arbitrarily pick $p \in K \setminus A$. For each open ball $B_r(p)$, there exists $\frac{1}{n}$ small enough to be smaller than r to give us $p \in B_{\frac{1}{n}}(a)$ for some $a \in A_{\frac{1}{n}}$, and give us $a \in B_{\frac{1}{n}}(p) \subseteq B_r(p)$ (done) ■

Theorem 3.10.4. (In Metric Space, Limit Point Compact Implies Compact)

Proof. Let \mathcal{G} be an open cover for K . By Theorem ??, we know K is separable. Then by Theorem ??, we know K has a countable base. Then by Theorem ??, we know \mathcal{G} contain a countable sub-cover

$$\{G_i\}_{i \in \mathbb{N}} \quad (3.209)$$

We now prove there exists $m \in \mathbb{N}$ such that $\{G_i\}_{i=1}^m$ cover K .

For each $n \in \mathbb{N}$, denote

$$E_n = \bigcup \{G_i\}_{i=1}^n \quad (3.210)$$

Obviously we have

$$E_1 \subseteq E_2 \subseteq E_3 \subseteq \dots \quad (3.211)$$

We know E_n converge to K , that is

$$\bigcup \{E_i\}_{i=1}^\infty = K \quad (3.212)$$

Assume $\forall m \in \mathbb{N}, E_m \neq K$. Then we can construct an infinite sequence

$$\{x_i : x_i \in E_i^c\}_{i=1}^\infty \quad (3.213)$$

Because K is limit point compact, we know $\{x_i\}_{i=1}^\infty$ has a limit point p .

Because $\{G_i\}_{i=1}^\infty$ is an open cover, we know $p \in G_n$ for some n . Then we know there exist open ball $B_r(p)$ small enough to be contained by G_n . Notice $B_r(p)$ is contained by G_n implies $B_r(p)$ is contained by E_n , so we know $B_r(p)$ at best contain all x_j where $j \leq n$ **CaC** to Theorem ?? (done) ■

We now prove the MEGA TFAE.

Theorem 3.10.5. (In Metric Space, Limit Point Compact is Equivalent to Compact) For metric space K , the followings three are equivalent

- (a) K is a limit point compact space
- (b) K is a compact space
- (c) K is a countably compact space

Above three implies the following two equivalent statement (\mathbb{R} is an example the above statements are stronger)

- (1) K is separable

(2) K has a countable base

and all above implies

(i) every open cover for K has a countable sub-cover.

Proof. In Theorem ??, we prove that if K has a count base then K is separable by collecting one point from each basic open set, and in Theorem ??, we construct a countable base by collecting all open balls of rational radius around each point in the countable dense subset.

In Theorem ??, we prove if K is limit point compact then K is separable. We proved such by covering K with open ball of radius $\frac{1}{n}$ for each $n \in \mathbb{N}$, and argue the set of centers A is countable by proving to cover K always require only finite amount of point (which is easy to see if one know K is bounded, but our method here is to show if require infinite centers, then the set of infinite centers has no limit point, which is also easy to see if one use geometric intuition), and argue that each point p that isn't in A is a limit point of A by arguing for each distance r , there exists an $a \in A$ closed enough to be closer to p than r .

One may observe, at the end of the proof for Theorem ??, we also construct a countable base for K .

In Theorem ??, we prove if K has a countable base then every open cover for K has a countable sub-cover by filtering the open cover with countable base.

Notice that we have proved that if K is limit point compact, then K is separable, has a countable base and every open cover for K has a countable sub-cover.

Then, in Theorem ??, we proved that if K is limit point compact, the countable sub-cover $\{G_i\}_{i \in \mathbb{N}}$ for K must has a finite sub-cover $\{G_i\}_{i=1}^m$, otherwise we can construct an infinite set $\{x_i\}_{i=1}^{\infty}$ by picking a point from area that isn't covered by $\{G_i\}_{i=1}^n$ for each natural n , and show that the limit point of $\{x_i\}_{i=1}^{\infty}$, being in some G_n , must contain at most of n amount of points of $\{x_i\}_{i=1}^{\infty}$, which is impossible.

It is clear that K is compact implies K is countably compact. In Theorem ??, we proved if K is countably compact, then K is limit point compact by showing if no point is a limit point of a countably infinite set A , then for each point $a \in A$, we can find an open neighborhood around a small enough to contain only one point in A , i.e. a itself, and the collection of such small enough open neighborhood and A^c form an open cover that has no finite sub-cover, as any finite sub-cover contain only finite amount of point in A , and

observe every uncountable set must contain some countable set thus also have a limit point.

In summary

$$(2) \implies (1) \text{ by Corollary ??} \quad (3.214)$$

$$(1) \implies (2) \text{ by Theorem ??} \quad (3.215)$$

$$(a) \implies (1) \text{ and } (2) \text{ by Theorem ??} \quad (3.216)$$

$$(2) \implies (i) \text{ by Theorem ??} \quad (3.217)$$

$$(a) \implies (b) \text{ by Theorem ??, which use (i)} \quad (3.218)$$

$$(b) \implies (c) \text{ by Theorem ??, which is trivial} \quad (3.219)$$

$$(c) \implies (a) \text{ by Theorem ??} \quad (3.220)$$



Above is a very important result. One must know how to prove such.

3.11 Metric Space: Some Properties Implied by Compact

Notice there is a difference between compact space and compact set

Definition 3.11.1. (Definition of Compact Subset) Let X be a metric space, we say $K \subseteq X$ is a compact subset if

$$\text{every open cover for } K \text{ has a finite sub-cover} \quad (3.221)$$

Theorem 3.11.2. (Compact Subset is itself a Compact Space)

$$\text{A compact subset regarded as a metric space is compact} \quad (3.222)$$

Proof. Let K be a compact subset, and regard K as a metric space. Let $\{G_\lambda : \lambda \in \Lambda\}$ be an open cover for K with respect to K . We know there exists a collection of open set $\{G_i : i \in I\}$ such that

$$\{G_\lambda : \lambda \in \Lambda\} = \{G_i \cap K : i \in I\} \quad (3.223)$$

Because $\{G_\lambda\}$ is an open cover for K with respect to K , we know $\{G_i\}$ is an open cover for K . Then because K is compact, there exists a finite sub-cover $\{G_j\}$. Notice $\{G_j \cap K\}$ is a finite sub-cover for K with respect to K . ■

Theorem 3.11.3. (Compact Sets Are Closed) Let X be a metric space, and let $K \subseteq X$

$$K \text{ is compact} \implies K \text{ is closed} \quad (3.224)$$

Proof. We prove K^c is open. Let $p \in K^c$. For each $k \in K$, collect the open balls $B_{\frac{d(k,p)}{2}}(k)$ as \mathcal{O} , so we have an open cover \mathcal{O} for K . Because K is a compact subset, we know \mathcal{O} has a finite sub-cover

$$\mathcal{O}' := \{B_{\frac{d(k_i,p)}{2}}(k_i)\}_{i=1}^m \quad (3.225)$$

Let $r = \min\{\frac{d(k_i,p)}{2} : 1 \leq i \leq m\}$. We only have to prove

$$\text{the open neighborhood } B_r(p) \text{ is disjoint to } \bigcup \mathcal{O}'. \quad (3.226)$$

This isn't difficult to prove if one have geometric intuition. Assume $q \in B_r(p) \cap \bigcup \mathcal{O}'$. Because $q \in \bigcup \mathcal{O}'$, we know there exists $B_{\frac{d(k_i,p)}{2}}(k_i)$ that contain q , so we have

$$d(k_i, q) \leq \frac{d(k_i, p)}{2} \quad (3.227)$$

Then we have

$$d(p, q) + d(q, k_i) < r + \frac{d(k_i, p)}{2} \leq d(k_i, p) \text{ CaC (done)} \quad (3.228)$$

Definition 3.11.4. (Definition of Bounded) Let X be a metric space, we say $E \subseteq X$ is bounded if there exists an open ball $B_r(p)$ centering some p in E , that contained E

Theorem 3.11.5. (Compact Sets are Bounded) Let X be a metric space and $K \subseteq X$. We have

$$K \text{ is compact} \implies K \text{ is bounded} \quad (3.229)$$

Proof. Assume K is unbounded. Arbitrarily pick q from K . We know for each $n \in \mathbb{N}$, the open ball $B_n(q)$ doesn't cover K

$$K \setminus B_n(q) \neq \emptyset \quad (3.230)$$

Then we can construct a sequence $\{x_n\}_{n \in \mathbb{N}}$ by arbitrarily picking x_n from $K \setminus B_n(q)$ for each $n \in \mathbb{N}$. Because K is compact implies K is limit point compact, we know $\{x_n\}_{n \in \mathbb{N}}$ has a limit point p in K . Let $r = d(p, q)$, and let m be the smallest natural greater than r . Because $B_{m-r}(p)$ is contained by $B_m(q)$, we know $B_{m-r}(p)$ contain at most first m points in $\{x_n\}_{n=1}^\infty$ CaC to every open ball centering limit point contain infinite point. ■

Corollary 3.11.6. (Compact Sets are Closed and Bounded)

$$K \text{ is compact} \implies K \text{ is closed and bounded} \quad (3.231)$$

We close this section with a method to generate new compact sets. ■

Theorem 3.11.7. (Closed Subsets of Compact Sets Are Compact)

$$K \text{ is compact and } F \subseteq K \text{ is closed} \implies F \text{ is compact} \quad (3.232)$$

Proof. Arbitrarily pick an open cover \mathcal{G} for F . We wish to prove there is a finite sub-cover for F . Because F is closed, we know F^c is open, so we know $\mathcal{G} \cup \{F^c\}$ is also an open cover, and moreover, open cover for X and K . Then because K is compact, we know $\mathcal{G} \cup \{F^c\}$ has a finite sub-cover \mathcal{G}' for K , and apparently, for F . If $\mathcal{G}' \subseteq \mathcal{G}$, then we are done. If $F^c \subseteq \mathcal{G}'$, then we can see $\mathcal{G}' \setminus \{F^c\}$ is a finite sub-cover of \mathcal{G} for F . ■

Corollary 3.11.8. Let F be closed and K be compact. We have $F \cap K$ is compact.

Proof. The fact that K is compact tell us K is closed. Then we know $F \cap K$ is closed and is a subset to K ■

3.12 Euclidean Metric

Definition 3.12.1. (Definition of k-cell) In $(\mathbb{R}^k, d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|)$, we say I is a k -cell, when there exists k intervals $\{[a_1, b_1], \dots, [a_k, b_k]\}$ such that

$$I = \prod_{n=1}^k [a_n, b_n] \quad (3.233)$$

Lemma 3.12.2. (Shrinking k-cell Sequence Convergence) Let $\{I_n\}_{n=1}^\infty$ be a sequence of k -cell such that

$$\forall n \in \mathbb{N}, I_{n+1} \subseteq I_n \quad (3.234)$$

Then we have

$$\bigcap \{I_n\}_{n=1}^\infty \neq \emptyset \quad (3.235)$$

Proof. For each $n \in \mathbb{N}$, denote the first component of I_n by $[a_n, b_n]$. We know $\{a_n\}_{n=1}^\infty$ is an monotonically increasing sequence and bounded above by every b_n . Notice that the supremum of a set A is exactly the infimum of the set of upper bounds of A . Applying this fact into our situation, we know for each n , $\sup\{a_n\}_{n=1}^\infty$ is greater than or equal to a_n and smaller than or equal to b_n . Then we know $\sup\{a_n\}_{n=1}^\infty$ must be contained by every $[a_n, b_n]$. In other words,

$$\sup\{a_k\}_{k=1}^\infty \in \bigcap_{n=1}^\infty [a_n, b_n] \quad (3.236)$$

Use the same argument in every other component to get a point in $\bigcap \{I_n\}_{n=1}^\infty$. ■

Lemma 3.12.3. (Maximum Distance of Two Points In k-cell is the length of Diagonal Line) Let $I = [a_1, b_1] \times \dots \times [a_k, b_k]$. Let $\mathbf{a} = (a_1, \dots, a_k)$ and $\mathbf{b} = (b_1, \dots, b_k)$. The maximum distance of two points in k -cell are

$$|\mathbf{b} - \mathbf{a}| = \left(\sum_{i=1}^k (b_i - a_i)^2 \right)^{\frac{1}{2}} \quad (3.237)$$

Proof. Let $\mathbf{x}, \mathbf{y} \in I$. For each $n \in \mathbb{N}$ smaller then or equal to k , because $x_n, y_n \in [a_n, b_n]$, we have $|x_n - y_n| \leq b_n - a_n$. Then we can deduce

$$|\mathbf{x} - \mathbf{y}| = \left(\sum_{i=1}^k |x_i - y_i|^2 \right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^k (b_i - a_i)^2 \right)^{\frac{1}{2}} = |\mathbf{b} - \mathbf{a}| \quad (3.238)$$

■

Theorem 3.12.4. (k-Cell is Compact)

Every k-cell I is compact. (3.239)

Proof. Let I_0 be $[a_1, b_1] \times \cdots \times [a_k, b_k]$. Assume $\{G_\lambda : \lambda \in \Lambda\}$ is an open cover for I_0 that has no finite sub-cover. For each $n \in \mathbb{N} : 1 \leq n \leq k$, let

$$c_n = \frac{a_n + b_n}{2} \quad (3.240)$$

Now, we can break I_0 into 2^k amount of k -cell of same size, where the n -th component of each of them is either $[a_n, c_n]$ or $[c_n, b_n]$. Notice that at least one of them can not be covered by finite amount of G_λ , since if all of them can be covered by a finite subset of $\{G_\lambda\}$ the union of finite sub-covers for each small k -cell is a finite sub-cover for I_0 , where we have assumed such finite sub-cover does not exist.

Denote the smaller k -cell that can not be covered by finite amount of G_λ by I_1 . Notice that $\{G_\lambda : \lambda \in \Lambda\}$ is an open cover for I_1 , and there exists no finite sub-cover of $\{G_\lambda : \lambda \in \Lambda\}$ for I_1 , just like I_0 .

Let $I_1 = [x_1, y_1] \times \cdots \times [x_k, y_k]$. For each $n \in \mathbb{N} : 1 \leq n \leq k$, let

$$z_n = \frac{x_n + y_n}{2} \quad (3.241)$$

Again, break I_1 into 2^k amount of k -cell of even smaller size using z_n . And, again, argue that one of the even smaller k -cell has can not be covered by a finite subset of $\{G_\lambda\}$.

Let I_2 be one of the even smaller k -cell that can not be covered by finite amount of G_λ , and use the same technique to construct a even even smaller k -cell I_3 , which can not be covered by finite amount of G_λ . This procedure can go on infinitely, so we have a sequence

$$\cdots \subseteq I_3 \subseteq I_2 \subseteq I_1 \subseteq I_0 \quad (3.242)$$

Denote this sequence $\{I_n\}_{n=1}^\infty$. By Lemma ??, we know

$$\bigcap \{I_n\}_{n=0}^\infty \neq \emptyset \quad (3.243)$$

Then, we can let

$$\mathbf{x} \in \bigcap \{I_n\}_{n=0}^\infty \quad (3.244)$$

Because for each non-negative integer n ,

$$\mathbf{x} \in I_n \text{ and } \{G_\lambda : \lambda \in \Lambda\} \text{ is an open cover for } I_n \quad (3.245)$$

We know there exists G_λ that contain \mathbf{x} . By definition of open, we know there exists an open ball $B_r(\mathbf{x})$ small enough to be contained by G_λ

$$\mathbf{x} \in B_r(\mathbf{x}) \subseteq G_\lambda \quad (3.246)$$

Now, we wish to pick a even smaller $I_m \in \{I_n\}_{n=0}^\infty$ contained by $B_r(\mathbf{x})$, so we have

$$\mathbf{x} \in I_m \subseteq B_r(\mathbf{x}) \subseteq G_\lambda \quad (3.247)$$

to show $\{G_\lambda\}$ is a cardinality-one sub-cover for I_m , which is impossible as we construct $\{I_n\}_{n=0}^\infty$ to satisfy that each of them can not be covered by finite subset of $\{G_\lambda : \lambda \in \Lambda\}$.

Because of the way we break each k-cell to smaller k-cell is to cut the k-cell right in half, it is simple to use geometric intuition to see such small enough I_m should exist.

Let

$$\delta = \left(\sum_{j=1}^k (b_j - a_j)^2 \right)^{\frac{1}{2}} = |\mathbf{b} - \mathbf{a}| \quad (3.248)$$

We can use mathematical induction to rigorously prove that the length of diagonal line for each I_n is exactly $2^{-n}\delta$.

Now, using Lemma ??, it is easy to verify when $m > \log_2 \frac{\delta}{r}$, the k-cell I_m is small enough to be contained by $B_r(\mathbf{x})$ **CaC** . ■

Corollary 3.12.5. (Heine-Borel Theorem) In $(\mathbb{R}^k, d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|)$, we have

$$K \text{ is closed and bounded } \iff K \text{ is compact} \quad (3.249)$$

Proof. We have proved from right to left in Theorem ?? and Theorem ??. To prove from left to right, just observe K is bounded implies that K is contained by some k-cell. Then because closed subset of compact set is compact, we can deduce K is compact. ■

Chapter 4

Numerical Sequence and Series

4.1 Convergence

Definition 4.1.1. (Definition of Sequentially Compact) Let $E \subseteq X$ be a topological space. We say E is sequentially compact if

$$\text{for each sequence } \{a_n\} \text{ there exists a converging sub-sequence } \{a'_n\} \quad (4.1)$$

Notice that in general topology, a sequence can converge to multiple point, the trivial topology of multiple points is an example.	Notice for N
--	----------------

Theorem 4.1.2. (In Metric Space, Limit Point Compact Implies Sequentially Compact)

$$K \text{ is limit point compact} \implies K \text{ is sequentially compact} \quad (4.2)$$

Proof. Let $\{p_n\}$ be a sequence in K . If the range of $\{p_n\}$ is finite, then there exists a point p_m appears infinitely many time in the sequence, otherwise the sequence end at a finite term. Then, we can pick the sub-sequence that contain only p_m and we are done.

If the range of $\{p_n\}$ is infinite, then because K is limit point compact, we know the range of $\{p_n : n \in \mathbb{N}\}$ has a limit point in K . Denote this point by p' , and pick the first term $p'_1 = p_{k_1}$ of our sub-sequence by arbitrarily picking p_{k_1} in $B_1(p)$. Now, we wish to find points in $B_{\frac{1}{2}}(p) \cap \{p_n : k_1 < n \in \mathbb{N}\}$. Notice the set

$$B_{\frac{1}{2}}(p) \cap \{p_n : n \in \mathbb{N}\} \quad (4.3)$$

is infinite, so it is impossible

$$B_{\frac{1}{2}}(p) \cap \{p_n : n \in \mathbb{N}\} \subseteq \{p_n : p \leq k_1 \in \mathbb{N}\} \quad (4.4)$$

then we know

$$B_{\frac{1}{2}}(p) \cap \{p_n : n > k_1 \in \mathbb{N}\} \neq \emptyset \quad (4.5)$$

So we can let $p'_2 = p_{k_2} \in B_{\frac{1}{2}}(p)$. By axiom of choice, we can do this infinite amount of time and have a sub-sequence $\{p'_n\}$ that converge to p . ■

Theorem 4.1.3. (In Euclidean Space, Bounded Implies Existence of Convergent Sub-sequence) In \mathbb{R}^n , we have

$$\{p_n\} \text{ is bounded} \implies \{p_n\} \text{ has a convergent sub-sequence} \quad (4.6)$$

Proof. In Euclidean space, if $\{p_n\}$ is bounded, we can find a k-cell I that contain $\{p_n\}$. We know k-cell I is compact, so we know it is sequentially compact. Then we know $\{p_n\}$ has a convergent sub-sequence in I . ■

4.2 Cauchy Sequence

Definition 4.2.1. (Definition of Cauchy Sequence) Let X be a metric space. We say $\{p_n\}$ is a Cauchy sequence if

$$\forall \epsilon \in \mathbb{R}^+, \exists N \in \mathbb{N}, \forall k, m > N, d(p_m, p_k) < \epsilon \quad (4.7)$$

Definition 4.2.2. (Definition of Diameter) Let $E \subseteq X$. We say the diameter of E is

$$\text{diam } E = \sup\{d(p, q) : p, q \in E\} \quad (4.8)$$

Theorem 4.2.3. (Convergent Implies Cauchy)

$$\{p_n\} \text{ converge} \implies \{p_n\} \text{ is a Cauchy sequence} \quad (4.9)$$

Proof. Let $\{p_n\}$ converge to p . Fix ϵ , we know there exists $N_{\frac{\epsilon}{2}}$ such that

$$k, m > N_{\frac{\epsilon}{2}} \implies d(p_k, p), d(p_m, p) < \frac{\epsilon}{2} \implies d(p_k, p_m) < d(p_k, p) + d(p_m, p) = \epsilon \quad (4.10)$$

■

Theorem 4.2.4. (Diameter of Closure Equals to Original Diameter) Let $E \subseteq X$. We have

$$\text{diam } \overline{E} = \text{diam } E \quad (4.11)$$

Proof. We wish to prove

$$\sup\{d(p, q) : p, q \in \overline{E}\} = \sup\{d(p, q) : p, q \in E\} \quad (4.12)$$

Because $E \subseteq \overline{E}$, we have

$$\sup\{d(p, q) : p, q \in \overline{E}\} \geq \sup\{d(p, q) : p, q \in E\} \quad (4.13)$$

Assume $\sup\{d(p, q) : p, q \in \overline{E}\} > \sup\{d(p, q) : p, q \in E\}$. We know there exists $p', q' \in \overline{E}$ such that

$$d(p', q') > \text{diam } E \quad (4.14)$$

Clearly, p', q' can not both be in E . Let $p' \in \overline{E} \setminus E$. We know

$$p' \in E' \quad (4.15)$$

If $q' \in E$, denote q' by q . We see

$$\forall p \in E, d(p', q) \leq d(p, p') + d(p, q) \quad (4.16)$$

where $d(p, p') > 0$ can be arbitrarily small. So if

$$d(p, q) < \text{diam } E \quad (4.17)$$

We have

$$d(p', q) \leq d(p, p') + d(p, q) < \text{diam } E \quad (4.18)$$

Then we have

$$\forall p \in E, d(p, q) = \text{diam } E \quad (4.19)$$

Which give us

$$E = \{q\} \text{ CaC} \quad (4.20)$$

Let $q' \notin E$. We know

$$q' \in E' \quad (4.21)$$

We see

$$\forall p, q \in E, d(p', q') \leq d(p', p) + d(p, q) + d(q, q') \quad (4.22)$$

where $d(p', p)$ and $d(q, q')$ can be arbitrarily small.

Let $d(p, q)$ satisfy

$$d(p, q) < \text{diam } E \quad (4.23)$$

Let $d(p', p)$ and $d(q, q')$ small enough to have

$$d(p', q') \leq d(p', p) + d(p, q) + d(q, q') < \text{diam } E \quad (4.24)$$

■

Theorem 4.2.5. (Sufficient Condition for Infinite Intersection of Compact Sets being Non-empty) Let $\{K_\lambda : \lambda \in \Lambda\}$ be an infinite collection of compact sets.

$$\text{every finite sub-collection of } \{K_\lambda : \lambda \in \Lambda\} \text{ is non-empty} \implies \bigcap_{\lambda \in \Lambda} K_\lambda \text{ is non-empty} \quad (4.25)$$

Proof. Assume $\bigcap_{\lambda \in \Lambda} K_\lambda = \emptyset$. Fix K_α . Because

$$K_\alpha \cap \left(\bigcap_{\lambda \in \Lambda, \lambda \neq \alpha} K_\lambda \right) = \emptyset \quad (4.26)$$

We know no point of K_α belong to every K_λ , otherwise that point is in $K_\alpha \cap (\bigcap_{\lambda \in \Lambda, \lambda \neq \alpha} K_\lambda)$.

Then we see

$$\{K_\lambda^c : \lambda \in \Lambda\} \quad (4.27)$$

Is an open cover for K_α .

Because K_α is compact, there exists a finite sub-cover

$$\{K_1^c, \dots, K_m^c\} \quad (4.28)$$

Notice that for all $x \in K_\alpha$, we have

$$\exists n : 1 \leq n \leq m, x \in K_n^c \quad (4.29)$$

Where $x \in K_n^c \implies x \notin K_n$. This tell us

$$\text{no points in } K_\alpha \text{ belong to all } K_1, \dots, K_m \quad (4.30)$$

Meaning

$$K_\alpha \cap \bigcap \{K_1, \dots, K_m\} = \emptyset \quad \text{CaC} \quad (4.31)$$

■

Corollary 4.2.6. (Single Point) Let $\{K_n\}$ be a sequence of compact sets such that

$$\forall n, K_{n+1} \subseteq K_n \text{ and } \lim_{n \rightarrow \infty} \text{diam } K_n = 0 \quad (4.32)$$

We have

$$\bigcap_{n=1}^{\infty} K_n \text{ consists of exact one point} \quad (4.33)$$

Proof. Notice if $\bigcap_{n=1}^{\infty} K_n$ has more than one point, we can pick two different points in $\bigcap_{n=1}^{\infty} K_n$ to show

$$\text{diam } \bigcap_{n=1}^{\infty} K_n > 0 \quad (4.34)$$

However, because for all m , we have

$$\bigcap_{n=1}^{\infty} K_n \subseteq K_m \quad (4.35)$$

We know

$$\forall m, 0 < \text{diam } \bigcap_{n=1}^{\infty} K_n \leq \text{diam } K_m \quad (4.36)$$

CaC to

$$\lim_{n \rightarrow \infty} \text{diam } K_n = 0 \quad (4.37)$$

■

Definition 4.2.7. (Definition of Completed) We say a subset X of a metric space is complete if

$$\text{every Cauchy sequence in } X \text{ converge to some point in } X \quad (4.38)$$

Theorem 4.2.8. (Compact Implies Complete)

$$K \text{ is compact} \implies K \text{ is complete} \quad (4.39)$$

Proof. Let $\{p_n\}$ be a Cauchy sequence in K . For each n let

$$E_n := \{p_k : k > n\} \quad (4.40)$$

We first prove

$$\lim_{n \rightarrow \infty} \text{diam } \overline{E_n} = 0 \quad (4.41)$$

Fix ϵ . We know there exists N such that

$$m, k > N \implies d(p_m, p_k) < \epsilon \quad (4.42)$$

This implies

$$\epsilon \geq \text{diam } E_N = \text{diam } \overline{E_N} \text{ (done)} \quad (4.43)$$

Notice that

$$\overline{E_n} \subseteq K \text{ and } \overline{E_n} \text{ is closed} \implies \overline{E_n} \text{ is compact} \quad (4.44)$$

Then by Corollary ??, we know there exists a unique p such that

$$\forall n, p \in \overline{E_n} \quad (4.45)$$

Now, we wish to prove

$$\lim_{n \rightarrow \infty} p_n = p \quad (4.46)$$

Fix ϵ . Because

$$\lim_{n \rightarrow \infty} \text{diam } \overline{E_n} = 0 \quad (4.47)$$

We know there exists N such that

$$\forall n > N, \text{diam } \overline{E_n} < \epsilon \quad (4.48)$$

Let $N_1 > N$. We have

$$\text{diam } \overline{E_{N_1}} < \epsilon \quad (4.49)$$

Notice

$$E_{N_1} = \{p_{N_1+1}, p_{N_1+2}, \dots\} \subseteq \overline{E_{N_1}} \quad (4.50)$$

This implies

$$\forall n > N_1, d(p_n, p) \leq \text{diam } \overline{E_n} < \epsilon \text{ (done)} \quad (4.51)$$

■

Corollary 4.2.9. (\mathbb{R}^n is complete)

$$\mathbb{R}^n \text{ is complete} \quad (4.52)$$

Proof. Let $\{p_n\}$ be a Cauchy sequence in \mathbb{R}^n . We first prove $\{p_n\}$ is bounded.

Again, let

$$E_n := \{p_k : k > n\} \quad (4.53)$$

Because

$$\lim_{n \rightarrow \infty} \text{diam } \overline{E_n} = 0 \quad (4.54)$$

We know there exists n such that

$$\text{diam } E_n = \text{diam } \overline{E_n} \in \mathbb{R}^+ \quad (4.55)$$

This implies $\{p_n\}$ is bounded. (done)

Pick a k -cell that contain $\{p_n\}$. We know k -cell is compact. By Theorem ??, we see $\{p_n\}$ must converge to some point in that k -cell. ■

Corollary 4.2.10. (Cauchy Criterion for Series)

$$\sum a_n \text{ converge} \iff \forall \epsilon, \exists N, \forall m > u > N, \left| \sum_{n=u}^m a_n \right| \leq \epsilon \quad (4.56)$$

Proof. Let

$$s_n = \sum_{k=1}^n a_k \quad (4.57)$$

We see

$$|s_m - s_u| = \left| \sum_{n=u}^m a_n \right| \quad (4.58)$$

Because \mathbb{R} is complete, we know

$$s_n \text{ converge} \iff s_n \text{ is Cauchy} \quad (4.59)$$

Notice s_n is Cauchy means

$$\forall \epsilon, \exists N, \forall m, u > N, |s_m - s_u| < \epsilon \quad (4.60)$$

■

4.3 Sub-sequential Limit

Definition 4.3.1. (Definition of Sub-sequential Limit) We say

$$x \text{ is a sub-sequential limit of } \{x_n\} \quad (4.61)$$

$$\iff \text{there exists a sub-sequence } \{x_{f(n)}\} \text{ such that } \lim_{n \rightarrow \infty} x_{f(n)} = x \quad (4.62)$$

Above is specification. We first give definition of lim-sup.

Definition 4.3.2. (Definition of Limit Superior) For real sequence $\{x_n\}$, we define

$$\liminf_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \inf\{x_k : k > n\} \quad (4.63)$$

And

$$\limsup_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \sup\{x_k : k > n\} \quad (4.64)$$

Notice that the sequence $\{\inf\{x_k : k > n\}\}_{n \in \mathbb{N}}$ monotonically increase, and the sequence $\{\sup\{x_k : k > n\}\}_{n \in \mathbb{N}}$ monotonically decrease.

Theorem 4.3.3. (Main Property of Limit Superior)

$$\Lambda > \limsup_{n \rightarrow \infty} x_n \implies \text{There exists only finite terms greater than } \Lambda \quad (4.65)$$

$$\text{There exists infinite terms greater than } \Lambda \implies \Lambda \leq \limsup_{n \rightarrow \infty} x_n \quad (4.66)$$

Proof. We first prove

$$\Lambda > \limsup_{n \rightarrow \infty} x_n \implies \text{There exists only finite terms greater than } \Lambda \quad (4.67)$$

Because $\Lambda > \limsup_{n \rightarrow \infty} x_n$, we know there exists N such that

$$\Lambda > \sup\{x_n : n > N\} \quad (4.68)$$

In other words,

$$\exists N, \forall n > N, x_n \leq \sup\{x_n : n > N\} < \Lambda \quad (4.69)$$

Then we know the amount of terms greater than Λ is at most N (done) .

We now prove

$$\text{There exists infinite terms greater than } \Lambda \implies \Lambda \leq \limsup_{n \rightarrow \infty} x_n \quad (4.70)$$

■

Theorem 4.3.4. (Equivalent Definition of Limit Inferior) Let E be the set of subsequential limit of $\{x_n\}$. We have

$$\liminf_{n \rightarrow \infty} x_n = \inf E \quad (4.71)$$

even in the context of extended real numbers.

Proof. Notice that the sequence

$$\{\inf\{x_k : k > n\}\}_{n \in \mathbb{N}} \quad (4.72)$$

monotonically increase.

We first prove

$$\liminf_{n \rightarrow \infty} x_n = -\infty \implies \inf E = -\infty \quad (4.73)$$

Because $\{\inf\{x_k : k > n\}\}_{n \in \mathbb{N}}$ monotonically increase, we know

$$\inf\{x_n\} = -\infty \quad (4.74)$$

For each n , let

$$R_n = -n \quad (4.75)$$

Fix n . We know

$$\forall N, \exists u_n > N, x_{u_n} < R_n \quad (4.76)$$

Otherwise, say $\exists N, \forall u > N, x_u \geq -R$, we will have $\inf\{x_k\}_{k \in \mathbb{N}} \geq \min\{R_n, x_1, \dots, x_N\} \in \mathbb{R}$.

Then for each n , we know there exists u_n greater than $\max\{u_1, \dots, u_{n-1}\}$, such that

$$x_{u_n} < R_n = -n \quad (4.77)$$

Let $f(n) = u_n$. We see

$$\lim_{n \rightarrow \infty} x_{f(n)} = -\infty \quad (4.78)$$

which implies $\inf E = -\infty$ (done)

We now prove

$$\liminf_{n \rightarrow \infty} x_n = L \implies \inf E = L \quad (4.79)$$

Because $\{\inf\{x_k : k > n\}\}$ is a monotonically increasing sequence. For each n , we know

$$\inf\{x_k : k > n\} \leq L \quad (4.80)$$

Then for each n , we can let $f(n)$ be a number greater than $\max\{f(1), \dots, f(n-1), n\}$ such that $x_{f(n)} \leq L$. Then by squeeze Theorem, we know

$$\lim_{n \rightarrow \infty} x_{f(n)} = L \quad (4.81)$$

We have proved $\inf E \leq L$. Assume $\inf E < L$. Let $M = \frac{\inf E + L}{2}$. We have

$$\inf E < M < L \quad (4.82)$$

Then there exists a sub-sequence $\{x_{f(n)}\}$ such that

$$\lim_{n \rightarrow \infty} x_{f(n)} = M \quad (4.83)$$

Then we know there exists N such that

$$n > N \implies x_{f(n)} < \frac{L + M}{2} < L \quad (4.84)$$

Then we see for all n

$$\inf\{x_k : k > n\} < \frac{L + M}{2} < L \text{ CaC (done)} \quad (4.85)$$

Lastly we prove

$$\liminf_{n \rightarrow \infty} x_n = \infty \implies \inf E = \infty \quad (4.86)$$

Let $\{x_{f(n)}\}$ be a sub-sequence. We wish to prove

$$\lim_{n \rightarrow \infty} x_{f(n)} = \infty \quad (4.87)$$

Notice that we have

$$\liminf_{n \rightarrow \infty} x_n = \infty \quad (4.88)$$

$$\iff \forall R \in \mathbb{R}^+, \exists N, n > N \implies \inf\{x_k : k > n\} > R \quad (4.89)$$

$$\iff \forall R \in \mathbb{R}^+, \exists N, \inf\{x_k : k > N\} > R \quad (4.90)$$

$$\implies \forall R \in \mathbb{R}^+, \exists N, k > N \implies x_k > R \quad (4.91)$$

Then for all R , we know there must exist N_n such that $f(N_n) > N$, then we see

$$n > N_n \implies f(n) > f(N_n) > N \implies x_{f(n)} > R \text{ (done)} \quad (4.92)$$

■

Corollary 4.3.5. (Equivalent Definition of Limit Superior) Let E be the set of sub-sequential limit of $\{x_n\}$. We have

$$\limsup_{n \rightarrow \infty} x_n = \sup E \quad (4.93)$$

Theorem 4.3.6. (Sub-sequential Limit of a Sequence in Metric Space form a Closed Set) Let E be the set of all sub-sequential limit of $\{x_n\}$. We have

$$E \text{ is closed} \quad (4.94)$$

Proof. Let $q \in E'$. We wish to prove $q \in E$. In other words, we wish to find a sub-sequence $\{x_{f(n)}\}$ of $\{x_n\}$ that converge to q .

Because $q \in E'$, we know q is a limit of some sequence $\{p_n\}$ in E such that $q \notin \{p_n\}$.

Let the sequence $\{p_n\}$ satisfy

$$d(q, p_n) < \frac{1}{n} \quad (4.95)$$

Because for each k , the point p_k is a sub-sequential limit of $\{x_n\}$. We know for each k , there exists a sub-sequence $\{x_{f_k(n)}\} \subseteq \{x_n\}$ such that for each k , we have

$$\lim_{n \rightarrow \infty} x_{f_k(n)} = p_k \quad (4.96)$$

We now construct a sub-sequence $\{x_{f(n)}\}$ that converge to q .

We wish that for each n , we have

$$d(x_{f(n)}, q) < \frac{2}{n} \quad (4.97)$$

Notice that because

$$\lim_{n \rightarrow \infty} x_{f_k(n)} = p_k \quad (4.98)$$

For all n , there exists N_n such that

$$u > N_n \implies d(x_{f_n(u)}, p_n) < \frac{1}{n} \implies d(x_{f_n(u)}, q) \leq d(x_{f_n(u)}, p_n) + d(p_n, q) < \frac{2}{n} \quad (4.99)$$

Then we can let $f(n) > \max\{N_n, f(1), \dots, f(n-1)\}$ and we are done. (done) ■

Theorem 4.3.7. (Properties of Limit Inferior and Superior) Let $\{x_n\}$ be a real sequence. We have

$$\liminf_{n \rightarrow \infty} x_n = \inf E \in E \text{ and } \limsup_{n \rightarrow \infty} x_n = \sup E \in E \quad (4.100)$$

And have

$$L < \liminf_{n \rightarrow \infty} x_n \implies (\exists N, n > N \implies x_n > L) \quad (4.101)$$

Proof. The first statement hold true because E is closed. Let $L < \liminf_{n \rightarrow \infty} x_n$. We know there exists N such that

$$L < \inf\{x_k : k > N\} \quad (4.102)$$

This tell us

$$\forall k > N, x_k > L \quad (4.103)$$

This finish the proof. ■

Theorem 4.3.8. (Useful Theorem) If there exists N such that

$$\forall n > N, x_n \leq y_n \quad (4.104)$$

Then

$$\limsup_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} y_n \quad (4.105)$$

And

$$\liminf_{n \rightarrow \infty} x_n \leq \liminf_{n \rightarrow \infty} y_n \quad (4.106)$$

Proof. Assume $\limsup_{n \rightarrow \infty} x_n > \limsup_{n \rightarrow \infty} y_n$. Let β satisfy

$$\limsup_{n \rightarrow \infty} x_n > \beta > \limsup_{n \rightarrow \infty} y_n \quad (4.107)$$

We know there exists N_1 such that

$$\forall n > N_1, y_n < \beta \quad (4.108)$$

Let $\{x_{f(n)}\}$ be a sub-sequence converge to $\limsup_{n \rightarrow \infty} x_n$. Then we know there exists N_2 such that

$$\forall n > N_2, x_{f(n)} > \beta \quad (4.109)$$

Then we can select $n > \max\{N_1, N_2\}$, and because $f(n) > f(N_1) \geq N_1$, we have

$$y_{f(n)} < \beta < x_{f(n)} \text{ CaC} \quad (4.110)$$

The other side is similar. ■

4.4 Special Sequences

Theorem 4.4.1. (Sequence Arithmetic) Let $\{x_n\}, \{y_n\}$ be two real sequences that respectively converge to x and y . We have

$$\lim_{n \rightarrow \infty} x_n + y_n = x + y \quad (4.111)$$

$$\lim_{n \rightarrow \infty} cx_n = cx \text{ and } \lim_{n \rightarrow \infty} c + x_n = c + x \quad (4.112)$$

$$\lim_{n \rightarrow \infty} x_n y_n = xy \quad (4.113)$$

$$\lim_{n \rightarrow \infty} x_n^{-1} = x^{-1} \quad (4.114)$$

Proof. Fix ϵ . We know there exists $N_{(x, \frac{\epsilon}{2})}$ and $N_{(y, \frac{\epsilon}{2})}$ such that

$$n > N_{(x, \frac{\epsilon}{2})} \implies |x_n - x| < \frac{\epsilon}{2} \text{ and } n > N_{(y, \frac{\epsilon}{2})} \implies |y_n - y| < \frac{\epsilon}{2} \quad (4.115)$$

Then we have

$$n > \max\{N_{(x, \frac{\epsilon}{2})}, N_{(y, \frac{\epsilon}{2})}\} \implies |x_n + y_n - (x + y)| \leq |x_n - x| + |y_n - y| < \epsilon \quad (4.116)$$

We know there exists $N_{(x, \frac{\epsilon}{|c|})}$ such that

$$n > N_{(x, \frac{\epsilon}{|c|})} \implies |x_n - x| < \frac{\epsilon}{|c|} \implies |cx_n - cx| = |c||x_n - x| < \epsilon \quad (4.117)$$

We know there exists N_ϵ such that

$$n > N_{(x, \epsilon)} \implies |c + x_n - c + x| = |x_n - x| < \epsilon \quad (4.118)$$

Notice that

$$|x_n y_n - xy| = |(x_n - x)(y_n - y) + x(y_n - y) + y(x_n - x)| \quad (4.119)$$

$$\leq |x_n - x||y_n - y| + |x||y_n - y| + |y||x_n - x| \quad (4.120)$$

We know there exists $N_{(x, \sqrt{\frac{\epsilon}{3}})}$ and $N_{(y, \sqrt{\frac{\epsilon}{3}})}$ and $N_{(x, \frac{\epsilon}{3|y|})}$ and $N_{(y, \frac{\epsilon}{3|x|})}$ such that if

$$n > \max\{N_{(x, \sqrt{\frac{\epsilon}{3}})}, N_{(y, \sqrt{\frac{\epsilon}{3}})}, N_{(x, \frac{\epsilon}{3|y|})}, N_{(y, \frac{\epsilon}{3|x|})}\} \quad (4.121)$$

Then we have

$$|x_n y_n - xy| \leq |x_n - x||y_n - y| + |x||y_n - y| + |y||x_n - x| \quad (4.122)$$

$$\leq \sqrt{\frac{\epsilon}{3}} \sqrt{\frac{\epsilon}{3}} + |x| \frac{\epsilon}{3|y|} + |y| \frac{\epsilon}{3|x|} = \epsilon \quad (4.123)$$

Notice that

$$\left| \frac{1}{x_n} - \frac{1}{x} \right| = \left| \frac{x - x_n}{xx_n} \right| = \frac{1}{|x|} \frac{1}{|x_n|} |x - x_n| \quad (4.124)$$

We know there exists $N_{(x, \frac{|x|}{2})}$ such that

$$n > N_{(x, \frac{|x|}{2})} \implies |x| - |x_n| \leq |x - x_n| < \frac{|x|}{2} \implies \frac{|x|}{2} < |x_n| \implies \frac{1}{|x|} \frac{1}{|x_n|} < \frac{2}{|x^2|} \quad (4.125)$$

We also know there exists $N_{(x, \frac{|x^2|\epsilon}{2})}$ such that

$$n > \max\{N_{(x, \frac{|x|}{2})}, N_{(x, \frac{|x^2|\epsilon}{2})}\} \implies \left| \frac{1}{x_n} - \frac{1}{x} \right| = \frac{1}{|x|} \frac{1}{|x_n|} |x - x_n| < \frac{2}{|x^2|} \frac{|x^2|\epsilon}{2} = \epsilon \quad (4.126)$$

■

Theorem 4.4.2. (Special Sequence) Let $p > 0$. We have

$$\lim_{n \rightarrow \infty} \frac{1}{n^p} = 0 \quad (4.127)$$

$$(4.128)$$

Proof. Fix ϵ . We have

$$\lim_{n \rightarrow \infty} \frac{1}{n^p} = \lim_{n \rightarrow \infty} \left(\frac{1}{n}\right)^p = \left(\lim_{n \rightarrow \infty} \frac{1}{n}\right)^p \quad (4.129)$$

We only have to prove

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0 \quad (4.130)$$

Let $N = \frac{1}{\epsilon}$ and we have

$$n > \frac{1}{\epsilon} \implies \left| \frac{1}{n} - 0 \right| = \frac{1}{n} < \epsilon \quad (4.131)$$

■

Theorem 4.4.3. (Special Sequence) Let $p > 0$. We have

$$\lim_{n \rightarrow \infty} \sqrt[n]{p} = 1 \quad (4.132)$$

Proof. Let

$$u_n = \sqrt[n]{p} - 1 \quad (4.133)$$

Notice that

$$\lim_{n \rightarrow \infty} u_n = 0 \implies \lim_{n \rightarrow \infty} \sqrt[n]{p} - 1 = 0 \implies \lim_{n \rightarrow \infty} \sqrt[n]{p} = 1 \quad (4.134)$$

So we only have to prove

$$\lim_{n \rightarrow \infty} u_n = 0 \quad (4.135)$$

Also, notice

$$1 + \binom{n}{1}u_n + \binom{n}{2}u_n^2 + \cdots + \binom{n}{n}u_n^n = (u_n + 1)^n = (\sqrt[n]{p} - 1 + 1)^n = p \quad (4.136)$$

We first prove

$$p > 1 \implies \lim_{n \rightarrow \infty} u_n = 0 \quad (4.137)$$

We have

$$p > 1 \implies \sqrt[n]{p} > 1 \implies u_n > 0 \implies \binom{n}{2}u_n^2 + \cdots + \binom{n}{n}u_n^n > 0 \quad (4.138)$$

Then because

$$1 + nu_n + \binom{n}{2}u_n^2 + \cdots + \binom{n}{n}u_n^n = p \quad (4.139)$$

We have

$$1 + nu_n < p \quad (4.140)$$

Which give us

$$0 < u_n < \frac{p-1}{n} \quad (4.141)$$

Then by Squeeze Theorem, we have

$$0 \leq \lim_{n \rightarrow \infty} u_n \leq \lim_{n \rightarrow \infty} \frac{p-1}{n} = 0 \text{ (done)} \quad (4.142)$$

Notice that if $0 < p < 1$, then by what we have just proved, we know

$$\lim_{n \rightarrow \infty} p^{-\frac{1}{n}} = 1 \quad (4.143)$$

Then we know

$$\lim_{n \rightarrow \infty} p^{\frac{1}{n}} = 1 \quad (4.144)$$

■

Theorem 4.4.4. (Special Sequence)

$$\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1 \quad (4.145)$$

Proof. Let

$$e_n = \sqrt[n]{n} - 1 > 0 \quad (4.146)$$

We wish to prove

$$\lim_{n \rightarrow \infty} e_n = 0 \quad (4.147)$$

Notice

$$n = (e_n + 1)^n = 1 + \binom{n}{1}e_n + \binom{n}{2}e_n^2 + \cdots + \binom{n}{n}e_n^n \quad (4.148)$$

So we have

$$\frac{n(n-1)}{2}e_n^2 = \binom{n}{2}e_n^2 < 1 + \binom{n}{1}e_n + \binom{n}{2}e_n^2 + \cdots + \binom{n}{n}e_n^n = n \quad (4.149)$$

Then we have

$$e_n^2 < \frac{2}{n-1} \quad (4.150)$$

So we have

$$0 < e_n < \sqrt{\frac{2}{n-1}} \quad (4.151)$$

By Squeeze Theorem, we have

$$0 \leq \lim_{n \rightarrow \infty} e_n \leq \lim_{n \rightarrow \infty} \sqrt{\frac{2}{n-1}} = \sqrt{2} \lim_{n \rightarrow \infty} \frac{1}{(n-1)^{\frac{1}{2}}} = 0 \quad (4.152)$$

■

Theorem 4.4.5. (Special Sequence) Let $\alpha \in \mathbb{R}$ and $q > 1$

$$\lim_{n \rightarrow \infty} \frac{n^\alpha}{q^n} = 0 \quad (4.153)$$

Proof. Let

$$q = 1 + p \quad (4.154)$$

So we have

$$q^n = (1 + p)^n = 1 + np + \binom{n}{2}p^2 + \cdots + \binom{n}{n}p^n \quad (4.155)$$

Let $k \in \mathbb{N}$ satisfy

$$k > \alpha^+ \quad (4.156)$$

Then for n greater than k , because $q > 1 \implies p > 0$, we have

$$q^n = 1 + np + \binom{n}{2}p^2 + \cdots + \binom{n}{n}p^n > \binom{n}{k}p^k \quad (4.157)$$

If we require n to be greater than $2k$, we see

$$n > 2k \implies n > 2k - 2 \implies \frac{n}{2} > k - 1 \implies n - k + 1 > \frac{n}{2} \quad (4.158)$$

Then we have

$$\binom{n}{k}p^k = \frac{n(n-1)\cdots(n-k+1)}{k!}p^k = n(n-1)\cdots(n-k+1)\frac{p^k}{k!} > \left(\frac{n}{2}\right)^k \frac{p^k}{k!} \quad (4.159)$$

Notice that we have

$$q^n > \binom{n}{k}p^k > \left(\frac{n}{2}\right)^k \frac{p^k}{k!} \quad (4.160)$$

Then we have

$$\frac{n^\alpha}{q^n} < \frac{n^\alpha 2^k}{n^k p^k} k! = n^{\alpha-k} \left(\frac{2}{p}\right)^k k! \quad (4.161)$$

Because $k > \alpha^+$, we know $u = k - \alpha > 0$. Then we have

$$0 \leq \lim_{n \rightarrow \infty} \frac{n^\alpha}{q^n} \leq \left(\frac{2}{p}\right)^k k! \lim_{n \rightarrow \infty} n^{-u} = 0 \quad (4.162)$$

■

Theorem 4.4.6. (Special Sequence) Let $|x| < 1$. We have

$$\lim_{n \rightarrow \infty} x^n = 0 \quad (4.163)$$

Proof. Theorem ?? tell us $\forall \alpha \in \mathbb{R}, \forall q > 1$ we have

$$\lim_{n \rightarrow \infty} \frac{n^\alpha}{q^n} = 0 \quad (4.164)$$

If we let $\alpha = 0$, we have

$$\lim_{n \rightarrow \infty} \frac{1}{q^n} = 0 \quad (4.165)$$

Notice

$$|x| < 1 \implies \left|\frac{1}{x}\right| > 1 \quad (4.166)$$

If x is positive, we immediately have

$$\lim_{n \rightarrow \infty} x^n = \lim_{n \rightarrow \infty} \frac{1}{\left(\frac{1}{x}\right)^n} = 0 \quad (4.167)$$

■

4.5 Comparison Test, Root Test and Ratio Test

Theorem 4.5.1. (Geometric Series) Let x be non-negative

$$\sum_{n=0}^{\infty} x^n \text{ converge} \iff x < 1 \quad (4.168)$$

Proof. Let

$$s_n := \sum_{k=1}^n x^k \quad (4.169)$$

Notice

$$(1-x)s_n = \sum_{k=1}^n x^k - \sum_{k=1}^n x^{k+1} = x - x^{n+1} \quad (4.170)$$

This give us

$$s_n = \frac{x - x^{n+1}}{1-x} = \frac{x}{1-x}(1 - x^{n+1}) \quad (4.171)$$

Then we have

$$x < 1 \implies \lim_{n \rightarrow \infty} x^{n+1} = 0 \implies \lim_{n \rightarrow \infty} s_n = \frac{x}{1-x} \quad (4.172)$$

And have

$$x > 1 \implies \lim_{n \rightarrow \infty} x^{n+1} = \infty \implies \lim_{n \rightarrow \infty} s_n = \frac{x}{1-x}(-\infty) \quad (4.173)$$

■

Definition 4.5.2. (Monotonic) We say a real sequence $\{x_n\}$ is monotonic if $\{x_n\}$ monotonically increase or monotonically decrease.

We say $\{x_n\}$ monotonically increase if

$$\forall n, x_{n+1} \geq x_n \quad (4.174)$$

We say $\{x_n\}$ monotonically decrease if

$$\forall n, x_{n+1} \leq x_n \quad (4.175)$$

Theorem 4.5.3. (Monotonic Sequence Converge if and only if Bounded) Let $\{x_n\}$ be a monotonic sequence. We have

$$\{x_n\} \text{ converge} \iff \{x_n\} \text{ is bounded} \quad (4.176)$$

Proof. Let $\{x_n\}$ monotonically increase. We know $\{x_n\}$ is bounded below by x_1 . We first prove

$$\{x_n\} \text{ is bounded above} \implies \{x_n\} \text{ converge} \quad (4.177)$$

Notice that

$$\{x_n\} \text{ is bounded above} \implies \sup\{x_n\} \in \mathbb{R} \quad (4.178)$$

Let $x := \sup\{x_n\}$. We wish to prove

$$\lim_{n \rightarrow \infty} x_n = x \quad (4.179)$$

Fix ϵ . We know there exists some N such that

$$x_N > x - \epsilon \quad (4.180)$$

Notice that

$$\forall m > u > N, x_m, x_u \in (x - \epsilon, x) \text{ and } |x_m - x_u| < \epsilon \quad (4.181)$$

We have proved $\{x_n\}$ is Cauchy. (done)

We now prove

$$\{x_n\} \text{ converge} \implies \{x_n\} \text{ is bounded} \quad (4.182)$$

Let $\epsilon \in \mathbb{R}$. We know there exists N such that

$$\forall n > N, x_n < x + \epsilon \quad (4.183)$$

Then we see $\{x_n\}$ is bounded above by $\max\{x_1, \dots, x_N, x + \epsilon\}$. (done)

Let $\{x_n\}$ monotonically decrease. We know $\{-x_n\}$ increase. If $\{x_n\}$ is bounded, then $\{-x_n\}$ is bounded, which implies $\{-x_n\}$ converges and $\{x_n\}$ converges.

If $\{x_n\}$ converge, then $\{-x_n\}$ converge, which implies $\{-x_n\}$ is bounded and $\{x_n\}$ is bounded. ■

Theorem 4.5.4. (Comparison Test)

$$\sum c_n \text{ converge and } \exists N, \forall n > N, |a_n| \leq c_n \implies \sum a_n \text{ converge} \quad (4.184)$$

Proof. By Cauchy criterion, we know

$$\forall \epsilon, \exists N_0, \forall m > u > N_0, \sum_{k=u}^m c_k < \epsilon \quad (4.185)$$

Fix ϵ , let $N_1 = \max\{N_0, N\}$. We have

$$\forall m > u > N_1, \sum_{k=u}^m a_k < \sum_{k=u}^m |a_k| \leq \sum_{k=u}^m c_k < \epsilon \quad (4.186)$$

And we are done, because we have just prove the series $\sum a_n$ satisfy Cauchy Criterion. ■

Corollary 4.5.5. (Comparison Test)

$$\sum d_n \text{ diverge and } \exists N, \forall n > N, a_n \geq d_n \geq 0 \implies \sum a_n \text{ diverge} \quad (4.187)$$

Proof. If $\sum a_n$ converge, then $\sum d_n$ converge by Theorem ?? ■

Theorem 4.5.6. (Root Test) Given $\sum a_n$. Let

$$\alpha := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \quad (4.188)$$

Then

$$(a) \alpha < 1 \implies \sum a_n \text{ converge}$$

$$(b) \alpha > 1 \implies \sum a_n \text{ diverge}$$

Proof. Let

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \alpha < \beta < 1 \quad (4.189)$$

We know there exists N such that

$$\forall n > N, \sqrt[n]{|a_n|} < \beta < 1 \quad (4.190)$$

Then we have

$$\forall n > N, a_n \leq |a_n| < \beta^n \quad (4.191)$$

Because $\beta < 1$, we know

$$\sum_{n=1}^{\infty} \beta^n = \frac{\beta}{1 - \beta} \quad (4.192)$$

Then by Comparison Test, we know $\sum a_n$ converge. If

$$\alpha = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1 \quad (4.193)$$

Let $\{a_{f(n)}\}$ be a sub-sequence such that

$$\lim_{n \rightarrow \infty} \sqrt[f(n)]{|a_{f(n)}|} = \alpha > 1 \quad (4.194)$$

Then we see there exists $N_{\alpha-1}$ such that

$$\forall n > N_{\alpha-1}, \sqrt[n]{|a_{f(n)}|} > 1 \text{ and } |a_{f(n)}| > 1 \quad (4.195)$$

Notice that $\sum_{n=1}^{\infty} a_n$ converge implies $\lim_{n \rightarrow \infty} a_n = 0$ **CaC** ■

Theorem 4.5.7. (Ratio Test) Given series $\sum a_n$. We have

$$(a) \limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1 \implies \sum a_n \text{ converge}$$

$$(b) \liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1 \implies \sum a_n \text{ diverge}$$

Proof. We first prove

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1 \implies \sum a_n \text{ converge} \quad (4.196)$$

Let

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1 \quad (4.197)$$

Then we can pick β such that

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < \beta < 1 \quad (4.198)$$

We know there exists N that satisfy

$$\forall n \geq N, \left| \frac{a_{n+1}}{a_n} \right| < \beta < 1 \quad (4.199)$$

This give us

$$\forall n > N, |a_n| < \beta^{n-N} |a_N| \quad (4.200)$$

In other words,

$$\forall k, |a_{N+k}| < \beta^k |a_N| \quad (4.201)$$

Then because

$$\sum_{k=1}^{\infty} \beta^k |a_N| = \frac{|a_N| \beta}{1 - \beta} \quad (4.202)$$

By comparison test we know the series

$$\sum_{k=1}^{\infty} a_{N+k} \text{ converges} \quad (4.203)$$

This tell us

$$\sum a_n = \sum_{k=1}^N a_k + \sum_{k=1}^{\infty} a_{N+k} \text{ converge } (\text{done}) \quad (4.204)$$

We now prove

$$\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1 \implies \sum a_n \text{ diverge} \quad (4.205)$$

Assume $\sum a_n$ converge. We have

$$\lim_{n \rightarrow \infty} a_n = 0 \quad (4.206)$$

By

$$\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1 \quad (4.207)$$

We know there exists N such that

$$\forall n \geq N, \left| \frac{a_{n+1}}{a_n} \right| > 1 \quad (4.208)$$

This tell us

$$\forall n \geq N, |a_n| > |a_N| \quad (4.209)$$

Let $\epsilon = |a_N|$, clearly

$$\forall N_1, \exists n > \max\{N_1, N\}, |a_n| > |a_N| = \epsilon \text{ CaC} \quad (4.210)$$

■

Theorem 4.5.8. (Root Test is Stronger than Ratio Test) For positive sequence $\{a_n\}$, we have

$$\liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \leq \limsup_{n \rightarrow \infty} \sqrt[n]{a_n} \leq \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \quad (4.211)$$

Proof. Clearly,

$$\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \leq \limsup_{n \rightarrow \infty} \sqrt[n]{a_n} \quad (4.212)$$

We first prove

$$\liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq \liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \quad (4.213)$$

Let

$$\alpha = \liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \quad (4.214)$$

If $\alpha = -\infty$, the theorem hold true trivially. We only have to consider when $\alpha > -\infty$. Arbitrarily pick positive β smaller than α :

$$\beta < \alpha = \liminf_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \quad (4.215)$$

Then we know there exists N such that

$$\forall n \geq N, \frac{a_{n+1}}{a_n} > \beta \quad (4.216)$$

This implies

$$\forall k, a_{N+k} > \beta^k a_N \quad (4.217)$$

Then for all $n > N$, we have

$$\sqrt[n]{a_n} > \sqrt[n]{\beta^{n-N} a_N} = \beta \sqrt[n]{\beta^{-N} a_N} \quad (4.218)$$

Because

$$\lim_{n \rightarrow \infty} \beta \sqrt[n]{\beta^{-N} a_N} = \beta \quad (4.219)$$

We see

$$\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \geq \beta \quad (4.220)$$

Then because $\beta < \alpha$ can be arbitrarily close to α , we see

$$\liminf_{n \rightarrow \infty} \sqrt[n]{a_n} \geq \alpha \text{ (done)} \quad (4.221)$$

■

Chapter 5

Vector Analysis

5.1 Important Inequalities

Theorem 5.1.1. (Young's Inequality) If $a, b \geq 0$ and $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad (5.1)$$

Proof.

$$ab = \exp(\ln a + \ln b) \quad (5.2)$$

$$= \exp\left[\left(\frac{1}{p} \ln a^p\right) + \left(\frac{1}{q} \ln b^q\right)\right] \quad (5.3)$$

$$\leq \frac{1}{p} \exp(\ln a^p) + \frac{1}{q} \exp(\ln b^q) \quad (5.4)$$

$$= \frac{a^p}{p} + \frac{b^q}{q} \quad (5.5)$$

The inequality above hold true because \exp is concave upward, and $\frac{1}{p} + \frac{1}{q} = 1$. ■

Theorem 5.1.2. (Definition of p -norm) Recall in Chapter 1, we define absolute value on \mathbb{R} and \mathbb{C} . With that definition, we have: for all $0 < p \in \mathbb{R}$, the definition

$$\|\mathbf{v}\|_p := \left(\sum |v_i|^p\right)^{\frac{1}{p}} \quad (5.6)$$

almost form a norm.

Proof. Observe

$$\|\lambda \mathbf{v}\|_p = \left(\sum |\lambda v_i|^p \right)^{\frac{1}{p}} \quad (5.7)$$

$$= \left[\sum (|\lambda| |v_i|)^p \right]^{\frac{1}{p}} \quad (5.8)$$

$$= \left[\sum (|\lambda|)^p (|v_i|)^p \right]^{\frac{1}{p}} \quad (5.9)$$

$$= [(|\lambda|)^p \sum (|v_i|)^p]^{\frac{1}{p}} \quad (5.10)$$

$$= [(|\lambda|)^p]^{\frac{1}{p}} \left(\sum |v_i|^p \right)^{\frac{1}{p}} \quad (5.11)$$

$$= |\lambda| (\|\mathbf{v}\|_p) \quad (5.12)$$

Nonnegative and positive definiteness is trivial. Triangle Inequality *is left to proved*. ■

Theorem 5.1.3. (Holder's Inequality in \mathbb{C}^n) Let $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, $\frac{1}{p} + \frac{1}{q} = 1$. Define \mathbf{xy} by $\mathbf{xy}_i = \mathbf{x}_i \mathbf{y}_i$. We have

$$\|\mathbf{xy}\|_1 \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q \quad (5.13)$$

Proof. Normalize

$$\hat{\mathbf{x}} := \frac{\mathbf{x}}{\|\mathbf{x}\|_p} \text{ and } \hat{\mathbf{y}} := \frac{\mathbf{y}}{\|\mathbf{y}\|_q} \quad (5.14)$$

So we only wish to prove

$$\|\hat{\mathbf{x}} \hat{\mathbf{y}}\|_1 \leq 1 \quad (5.15)$$

Notice because $(\sum |\hat{x}_i|^p)^{\frac{1}{p}} = \|\hat{\mathbf{x}}\|_p = \frac{\|\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = 1$, and similarly vice versa for \mathbf{y} , we have

$$\sum |\hat{x}_i|^p = 1 = \sum |\hat{y}_i|^q \quad (5.16)$$

So by Young's inequality, we have

$$\|\hat{\mathbf{x}} \hat{\mathbf{y}}\|_1 = \sum |\hat{x}_i \hat{y}_i| = \sum |\hat{x}_i| |\hat{y}_i| \leq \sum \frac{|\hat{x}_i|^p}{p} + \frac{|\hat{y}_i|^q}{q} = \frac{1}{p} + \frac{1}{q} = 1 \quad (5.17)$$

■

Corollary 5.1.4. (Minkowskis Inequality in \mathbb{C}^n and $p > 1$) Let $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, and let $p \in \mathbb{R}^+$ and $p \neq 1$. We have triangle inequality for all supposed p -norm

$$p > 1 \implies \|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p \quad (5.18)$$

Proof. Let $q = \frac{p}{p-1}$, so we have

$$\frac{1}{p} + \frac{1}{q} = \frac{p}{p} = 1 \quad (5.19)$$

Then we deduce

$$\|\mathbf{x} + \mathbf{y}\|_p^p = \sum |x_i + y_i|^p \quad (5.20)$$

$$= \sum |(x_i + y_i)^p| \quad (5.21)$$

$$= \sum |(x_i + y_i)(x_i + y_i)^{p-1}| \quad (5.22)$$

$$\leq \sum |x_i(x_i + y_i)^{p-1}| + |y_i(x_i + y_i)^{p-1}| \quad (5.23)$$

$$= \|\mathbf{x}(\mathbf{x} + \mathbf{y})^{p-1}\|_1 + \|\mathbf{y}(\mathbf{x} + \mathbf{y})^{p-1}\|_1 \text{ (Operation is component wise)} \quad (5.24)$$

$$\leq \|\mathbf{x}\|_p \|(\mathbf{x} + \mathbf{y})^{p-1}\|_q + \|\mathbf{y}\|_p \|(\mathbf{x} + \mathbf{y})^{p-1}\|_q \text{ (Holder's Inequality)} \quad (5.25)$$

$$= \|\mathbf{x}\|_p \left[\sum (|(x_i + y_i)^{p-1}|)^q \right]^{\frac{1}{q}} + \|\mathbf{y}\|_p \left[\sum (|(x_i + y_i)^{p-1}|)^q \right]^{\frac{1}{q}} \quad (5.26)$$

$$= \|\mathbf{x}\|_p \left[\sum (|x_i + y_i|^{p-1})^q \right]^{\frac{1}{q}} + \|\mathbf{y}\|_p \left[\sum (|x_i + y_i|^{p-1})^q \right]^{\frac{1}{q}} \quad (5.27)$$

$$= \|\mathbf{x}\|_p \left(\sum |x_i + y_i|^p \right)^{\frac{1}{q}} + \|\mathbf{y}\|_p \left(\sum |x_i + y_i|^p \right)^{\frac{1}{q}} \text{ (because } q(p-1) = p \text{)} \quad (5.28)$$

$$= (\|\mathbf{x}\|_p + \|\mathbf{y}\|_p) \|\mathbf{x} + \mathbf{y}\|_p^{\frac{p}{q}} \quad (5.29)$$

So

$$\|\mathbf{x} + \mathbf{y}\|_p = \|\mathbf{x} + \mathbf{y}\|_p^{p - \frac{p}{q}} \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p \quad (5.30)$$

■

Definition 5.1.5. (Definition of p -norm) On \mathbb{C}^n , if and only if $p > 1$, we say the p -norm is

$$\|\mathbf{v}\|_p := \left(\sum |v_i|^p \right)^{\frac{1}{p}} \quad (5.31)$$

5.2 Definition of Differential-able on Euclidean Space

Definition 5.2.1. (Little-o Notation) If we write $o(g(x)) = f(x)$, we mean

$$\lim_{g(x) \rightarrow 0} \frac{f(x)}{g(x)} = 0 \quad (5.32)$$

Definition 5.2.2. (Differentiable on \mathbb{R}) We say a real-valued function $f(x)$ on \mathbb{R} is differentiable on a if

$$\exists L \in \mathbb{R}, \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = L \quad (5.33)$$

Notice L should be viewed as slope (not a number on real number line), or more abstract, a linear transformation from \mathbb{R}^1 to \mathbb{R}^1 .

Definition 5.2.3. (A General Definition of Differentiable of Function from Normed \mathbb{R}^n to \mathbb{R}^m) Let $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$. We say \mathbf{f} is differentiable at \mathbf{a} if there exists a linear transformation $\mathbf{J} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - \mathbf{J}(\mathbf{h})\|_{\mathbb{R}^m}}{\|\mathbf{h}\|_{\mathbb{R}^n}} = 0 \quad (5.34)$$

There is no restriction on which norm to use on \mathbb{R}^n and \mathbb{R}^m , but the metric implicitly used in the limit symbol is induced by the norm in the denominator. That is

$$\forall \epsilon, \exists \delta, \forall \mathbf{h}, \|\mathbf{h}\|_{\mathbb{R}^n} < \delta \longrightarrow \frac{\|\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - \mathbf{J}(\mathbf{h})\|_{\mathbb{R}^m}}{\|\mathbf{h}\|_{\mathbb{R}^n}} < \epsilon \quad (5.35)$$

We also say \mathbf{J} is a differential or a derivative of \mathbf{f} at \mathbf{a} .

Definition 5.2.4. (Tangent Plane) Let $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be differentiable at \mathbf{a} . Let \mathbf{J} be a differential of \mathbf{f} at \mathbf{a} . We say

$$\{(\mathbf{x}, \mathbf{J}(\mathbf{x})) \in \mathbb{R}^{n+m} : \mathbf{x} \in \mathbb{R}^n\} \text{ where } \mathbf{J}(\mathbf{a}) = \mathbf{f}(\mathbf{a}) \quad (5.36)$$

is the tangent plane of \mathbf{f} at \mathbf{a} .

Definition 5.2.5. (Definition of Gradient) Let $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. The gradient of f is defined as

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right) \quad (5.37)$$

Notice that we can verify Gradient is Jacobian, and jacobian is always a derivative.

Definition 5.2.6. (Jacobian Matrix) For function $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, The Jacobian matrix $\mathbf{J}_{\mathbf{f}}$ is

$$\mathbf{J}_{\mathbf{f}} := \begin{bmatrix} \nabla^T f_1 \\ \vdots \\ \nabla^T f_m \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \quad (5.38)$$

The notation is

$$D(\mathbf{f}) = \mathbf{J} = \mathbf{J}_{\mathbf{f}} = \nabla \mathbf{f} = \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_n)} \quad (5.39)$$

Theorem 5.2.7. (If Differentiable, Then Jacobian give a Differential) Let $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be differentiable at \mathbf{a} . Define

$$\mathbf{J} : \mathbb{R}^n \rightarrow \mathbb{R}_m, \mathbf{x} \mapsto \mathbf{J}_f(\mathbf{a}) \cdot \mathbf{x} \quad (5.40)$$

We have

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \mathbf{J}(\mathbf{x} - \mathbf{a})\|_{\mathbb{R}^m}}{\|\mathbf{x} - \mathbf{a}\|_{\mathbb{R}^n}} = 0 \quad (5.41)$$

5.3 Example Exercises for Differentiable

Theorem 5.3.1. (Example of Differentiable Function)

$$f(x, y) = x^2 + 2y \text{ is differentiable every where} \quad (5.42)$$

Proof. Notice

$$\mathbf{J} = [2x \ 2] \quad (5.43)$$

For each point (a, b) , we let $\mathbf{H} := \mathbf{J}(a, b) = [2a \ 2]$, and we have

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \mathbf{H}(\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|_{\mathbb{R}^n}} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{(x - a)^2}{\sqrt{(x - a)^2 + (y - b)^2}} \quad (5.44)$$

Notice that we have

$$\lim_{x \rightarrow a} \frac{(x - a)^2}{\sqrt{(x - a)^2}} = 0 \quad (5.45)$$

In ϵ, δ language, we have

$$\forall \epsilon, \exists \delta, \forall x, |x - a| < \delta \implies \frac{(x - a)^2}{\sqrt{(x - a)^2}} < \epsilon \quad (5.46)$$

For all ϵ , we then can pick the δ above so we have

$$\sqrt{(x - a)^2 + (y - b)^2} < \delta \implies |x - a| < \delta \implies \frac{(x - a)^2}{\sqrt{(x - a)^2 + (y - b)^2}} \leq \frac{(x - a)^2}{\sqrt{(x - a)^2}} < \epsilon \quad (5.47)$$

This prove equation 4.40 approach to 0 ■

Theorem 5.3.2. (Example of Differentiable Function) Let $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$

$$\text{If } \mathbf{f} \text{ is linear, then } \mathbf{f} \text{ is differentiable everywhere} \quad (5.48)$$

Proof. Let $\mathbf{H} = \mathbf{f}$, then we have

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \mathbf{H}(\mathbf{x} - \mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x} - \mathbf{a} - (\mathbf{x} - \mathbf{a}))\|}{\|\mathbf{x} - \mathbf{a}\|} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{0})\|}{\|\mathbf{x} - \mathbf{a}\|} = 0 \quad (5.49)$$

■

Theorem 5.3.3. (Example of Jacobian Exists, But Non-Differentiable)

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{if } (x, y) \neq \mathbf{0} \\ 0 & \text{if } (x, y) = \mathbf{0} \end{cases} \text{ exists jacobian but non-differentiable at } \mathbf{0} \quad (5.50)$$

Proof. Notice $\mathbf{J}(0,0) = \begin{bmatrix} 0 & 0 \end{bmatrix}$. Now we prove f is non-differentiable. Observe

$$\lim_{(x,y) \rightarrow (0,0)} \frac{f(x,y) - f(0,0) - \mathbf{H}(0,0)}{\sqrt{x^2 + y^2}} = \frac{f(x,y)}{\sqrt{x^2 + y^2}} = \frac{xy}{(x^2 + y^2)^{\frac{3}{2}}} \quad (5.51)$$

We know

$$\lim_{x \rightarrow 0} \frac{x^2}{(2x^2)^{\frac{3}{2}}} = \infty \quad (5.52)$$

In ϵ, δ language, we have

$$\forall \epsilon, \exists \delta, x = \delta \implies \frac{x^2}{(2x^2)^{\frac{3}{2}}} > \epsilon \quad (5.53)$$

Then we have

$$\forall \epsilon, \exists \delta, (x,y) = \left(\frac{\delta}{\sqrt{2}}, \frac{\delta}{\sqrt{2}}\right) \implies \frac{xy}{(x^2 + y^2)^{\frac{3}{2}}} = \frac{\delta^2}{2(\delta^2)^{\frac{3}{2}}} > \epsilon \quad (5.54)$$

which prove equation 4.47 does not approach to 0. ■

Theorem 5.3.4. (Chain Rule) Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and let $g : \mathbb{R}^k \rightarrow \mathbb{R}^n$. We have

$$D_{\mathbf{a}}(f \circ g) = D_{g(\mathbf{a})}(f) \circ D_{\mathbf{a}}(g) \quad (5.55)$$

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^1$, we can write

$$h(\mathbf{x}) := f \circ g(\mathbf{x}) = f(g_1(\mathbf{x}), \dots, g_k(\mathbf{x})) \quad (5.56)$$

and we have

$$\frac{\partial h}{\partial x_i} = \sum_{1 \leq j \leq n} \frac{\partial f}{\partial g_j} \frac{\partial g_j}{\partial x_i} \quad (5.57)$$

Theorem 5.3.5. (Example of Chain Rule) Let $\delta : (0,1) \rightarrow \mathbb{R}^n$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable. Define

$$F(t) := f(\delta(t)) \quad (5.58)$$

We have

$$L_{F'(t)} = D_{\delta(t)}(f) \circ D_t(\delta) \quad (5.59)$$

Proof. Notice that $D_t(F) = L_{F'(t)}$, and we are done. ■

Theorem 5.3.6. (Example of Chain Rule) Let $f(u,v,w) = u^2v + wv^2$ and $\mathbf{g}(x,y) = (xy, \sin x, e^x)$. Let $h = f \circ \mathbf{g}$. We have

$$\frac{\partial h}{\partial x} = 2xy^2 \sin x + (x^2y^2 + 2e^x \sin x) \cos x + e^x \sin^2 x \quad (5.60)$$

Proof. We have

$$\frac{\partial h}{\partial x} = \frac{\partial h}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial h}{\partial v} \frac{\partial v}{\partial x} + \frac{\partial h}{\partial w} \frac{\partial w}{\partial x} \quad (5.61)$$

$$= (2uv)(y) + (u^2 + 2wv)(\cos x) + (v^2)(e^x) \quad (5.62)$$

$$= 2xy(\sin x)y + (x^2y^2 + 2e^x \sin x) \cos x + (\sin^2 x)e^x \quad (5.63)$$

$$= 2xy^2 \sin x + (x^2y^2 + 2e^x \sin x) \cos x + e^x \sin^2 x \quad (5.64)$$

■

Theorem 5.3.7. (Mean Value Theorem) Let $\mathbf{f} : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$. Suppose \mathbf{f} is differentiable on U and $\mathbf{x}, \mathbf{y} \in U$. For all $i \in \{1, \dots, n\}$, there exists $0 \leq c_i \leq 1$ such that

$$f_i(\mathbf{y}) - f_i(\mathbf{x}) = (Df_i)(\mathbf{x} + c_i(\mathbf{y} - \mathbf{x})) \quad (5.65)$$

5.4 Directional Derivative

Definition 5.4.1. (Definition of Directional Derivative) Let f be real valued and defined on a neighborhood of $\mathbf{x}_0 \in \mathbb{R}^n$ and let $\mathbf{v} \in \mathbb{R}^n$ with $\|\mathbf{v}\| = 1$. We define

$$(D_{\mathbf{v}}f)(\mathbf{x}_0) := \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v})}{t} \quad (5.66)$$

Theorem 5.4.2. (Directional Derivative) Let U be open and $f : U \rightarrow \mathbb{R}^n$ be differentiable at x_0 . We have

$$(D_{\mathbf{v}}f)(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v} = \mathbf{J}_f(\mathbf{x}_0) \cdot \mathbf{v} \quad (5.67)$$

Proof. Outline: in the limit of differentiable equation, let $x = x_0 + tv$ and see what happen when $t \rightarrow 0$ ■

Theorem 5.4.3. (Example) Let

$$f(x, y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{if } (x, y) \neq 0 \\ 0 & \text{if } (x, y) = (0, 0) \end{cases} \quad (5.68)$$

We have $f(x, y)$ is non-differentiable at v

Proof. Observe

$$D_v f(0) = \lim_{t \rightarrow 0} \frac{f(tv_1, tv_2) - f(0, 0)}{t} \quad (5.69)$$

$$= \lim_{t \rightarrow 0} \frac{v_1^3}{v_1^2 + v_2^2} = v_1^3 \neq Df(0)(v) \quad (5.70)$$

■

5.5 Gradient

Definition 5.5.1. (Level Set) Let f be a scalar function. A level set is a set of points

$$L(c) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = c\} \quad (5.71)$$

Theorem 5.5.2. (Gradient and Tangent Plane) Let f be a scalar function. Vector $\nabla f(\mathbf{p})$ is normal to the level set $L(f(\mathbf{p}))$.

Theorem 5.5.3. (Example) Let $S = \{(x, y, z) : x^2 + y^2 + z^2 = 3\}$. Find a vector normal to S at $(1, 1, 1)$

Proof. Define

$$f(x, y, z) := x^2 + y^2 + z^2 \quad (5.72)$$

We have

$$S = L(f(1, 1, 1)) \quad (5.73)$$

Compute

$$\nabla f(1, 1, 1) = (4, 4, 4) \quad (5.74)$$

■

Theorem 5.5.4. (TEST Example) Let $S = \{(x, y, z) : x^2 - y^2 + xyz = 1\}$. Find the tangent plane of S at $(1, 0, 1)$

Proof. Define

$$f(x, y, z) = x^2 - y^2 + xyz \quad (5.75)$$

We have

$$S = L(f(1, 0, 1)) \quad (5.76)$$

Compute

$$\nabla f(1, 0, 1) = (2, 1, 0) \quad (5.77)$$

The tangent plane is

$$2x + y = 2 \quad (5.78)$$

■

5.6 Integral

Theorem 5.6.1. (Integral) Let

$$\forall x \in S, g_1(x) \leq g_2(x) \quad (5.79)$$

Let

$$A = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : \forall x \in S, g_1(x) \leq y \leq g_2(x)\} \quad (5.80)$$

Let $f : A \rightarrow \mathbb{R}$ be continuous, we have

$$\int_A f(x, y) d(x, y) = \int_S \left(\int_{g_1(y)}^{g_2(y)} f(x, y) dy \right) dx \quad (5.81)$$

This enable us to calculate integral by iteration.

Theorem 5.6.2. (Example) Let

$$A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, x \leq y \leq 1\} \quad (5.82)$$

Find

$$\int_A xy dA \quad (5.83)$$

Proof. Compute

$$\int_A xy dA = \int_0^1 \int_0^y xy dx dy \quad (5.84)$$

$$= \int_0^1 \frac{y^3}{2} dy \quad (5.85)$$

$$= \frac{1}{8} \quad (5.86)$$

■

Theorem 5.6.3. (Example) Let

$$A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, \sqrt{x} \leq y \leq 1\} \quad (5.87)$$

Find

$$\int_A e^{y^3} dA \quad (5.88)$$

Proof. Compute

$$\int_A e^{y^3} dA = \int_0^1 \int_0^{y^2} e^{y^3} dx dy \quad (5.89)$$

$$= \int_0^1 y^2 e^{y^3} dy \quad (5.90)$$

$$= \frac{1}{3}(e - 1) \quad (5.91)$$

■

Theorem 5.6.4. (Example) Let

$$A = \{(x_1, x_2, x_3) : x_1, x_2, x_3 \geq 0, x_1 + x_2 + x_3 \leq 1\} \quad (5.92)$$

Find

$$\int_A (x_1 + x_2 + x_3)^2 dx_1 dx_2 dx_3 \quad (5.93)$$

Proof.

$$\int_S (x_1 + x_2 + x_3)^2 dx_1 dx_2 dx_3 = \int_0^1 \int_0^{1-x_3} \int_0^{1-x_3-x_2} (x_1 + x_2 + x_3)^2 dx_1 dx_2 dx_3 \quad (5.94)$$

$$= \int_0^1 \int_0^{1-x_3} \frac{1}{3} - \frac{(x_2 + x_3)^3}{3} dx_2 dx_3 \quad (5.95)$$

$$= \int_0^1 \frac{x_3^4}{12} - \frac{x_3}{3} + \frac{1}{4} dx_3 \quad (5.96)$$

$$= \frac{1}{10} \quad (5.97)$$

The answer is $\frac{4}{5}$

■

Theorem 5.6.5. (Example) Let $V(n)$ be the volume of unit n -ball. Compute $V(n)$

Proof. We first write $V(n)$ as

$$V(n) = \int_{-1}^1 \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{n-1}^2}} dx_n dx_{n-1} \cdots dx_1 \quad (5.98)$$

Then we can observe

$$V(n) = V(n-1) \int_{-1}^1 (1-x_1^2)^{\frac{n-1}{2}} dx_1 \quad (5.99)$$

$$= V(n-1) \int_0^1 (1-u)^{n-1} \frac{u^{-\frac{1}{2}}}{2} du (\because u = x_1^2) \quad (5.100)$$

■

Chapter 6

Advanced Calculus HW

6.1 HW1

Question 16

1. Prove that the following statements are equivalent: for a given sequence $\{x_n\}$,
- (a) for every $0 < \epsilon \in \mathbb{Q}$, there exists $N \in \mathbb{N}$ such that $|x_n - x| < \epsilon$ whenever $n \geq N$.
 - (b) for every $0 < \epsilon \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that $|x_n - x| < \epsilon$ whenever $n \geq N$.

Proof. From (b) to (a), just observe $\mathbb{Q} \subseteq \mathbb{R}$ and we are done. Now we prove from (a) to (b).

Because \mathbb{Q} is dense in \mathbb{R} , for all $\epsilon \in \mathbb{R}^+$, we can pick $\epsilon' \in \mathbb{Q}$ such that $0 < \epsilon' < \epsilon$. By (a), we know there exists $N \in \mathbb{N}$ such that $|x_n - x| < \epsilon' < \epsilon$ whenever $n \geq N$. This finish the proof. ■

Question 17

2. Let $\{x_n\}_{n=1}^{\infty}$ be a monotone increasing sequence such that

$$x_{n+1} - x_n \leq \frac{1}{n}.$$

Determine whether the sequence converges. (If yes, prove it; if not, disprove it or give a counterexample.)

Proof. No, consider p -series. The sequence $\{x_n\}_{n=1}^{\infty}$ defined by $x_i := \sum_{j=1}^i \frac{1}{j}$ is monotone increasing, satisfying the desired property, and from our knowledge, diverge. ■

Question 18

Let $M_{n \times m}$ be the collection of all $n \times m$ matrices with real entries. Define a function $\|\cdot\| : M_{n \times m} \rightarrow \mathbb{R}$ by

$$\|A\| = \sup \left\{ \frac{\|Ax\|_2}{\|x\|_2} : x \in \mathbb{R}^m, x \neq 0 \right\},$$

where we recall that $\|\cdot\|_2$ is the 2-norm on Euclidean space given by

$$\|Ax\|_2 = \left(\sum_{i=1}^k x_i^2 \right)^{1/2} \quad \text{if } x \in \mathbb{R}^k.$$

Show that:

1. $\|A\| = \sup \{\|Ax\|_2 : x \in \mathbb{R}^m, \|x\|_2 = 1\} = \inf \{M \in \mathbb{R} : \|Ax\|_2 \leq M\|x\|_2 \forall x \in \mathbb{R}^m\}.$
2. $\|Ax\|_2 \leq \|A\|\|x\|_2$ for all $x \in \mathbb{R}^m$.
3. $\|\cdot\|$ defines a norm on $M_{n \times m}$.

Proof. In this proof, we use $|\cdot|$ to denote $\|\cdot\|_2$, and if we write x without specification, it belong to \mathbb{R}^m

We first show $\|A\| = \sup\{|Ax| : |x| = 1\}$

Assume $\sup\{\frac{|Ax|}{|x|} : x \neq 0\} = \|A\| > \sup\{|Ax| : |x| = 1\}$. Then we know $\sup\{|Ax| : |x| = 1\}$ is not an upper bound of $\{\frac{|Ax|}{|x|} : x \neq 0\}$, so we know there exists $x \in \mathbb{R}^m$ such that $\frac{|Ax|}{|x|} > \sup\{|Ay| : |y| = 1\}$.

Define $\hat{x} := \frac{x}{|x|}$. We have $\frac{|Ax|}{|x|} = |\frac{Ax}{|x|}| = |A\hat{x}| \leq \sup\{|Ay| : |y| = 1\}$, since $|\hat{x}| = |\frac{x}{|x|}| = \frac{|x|}{|x|} = 1 \implies |A\hat{x}| \in \{|Ay| : |y| = 1\}$. This **CaC**.

Assume $\sup\{\frac{|Ax|}{|x|} : x \neq 0\} = \|A\| < \sup\{|Ax| : |x| = 1\}$. Then we know $\sup\{\frac{|Ax|}{|x|} : x \neq 0\}$ is not an upper bound of $\sup\{|Ax| : |x| = 1\}$, so we know there exists $\hat{x} \in \mathbb{R}^m : |\hat{x}| = 1$ such that $|A\hat{x}| > \sup\{\frac{|Ay|}{|y|} : y \neq 0\}$.

We see $|A\hat{x}| > \sup\{\frac{|Ay|}{|y|} : y \neq 0\} \geq \frac{|A\hat{x}|}{|\hat{x}|} = \frac{|A\hat{x}|}{1} = |A\hat{x}|$ **CaC** (done)

Observe $\inf \{M \in \mathbb{R} : \|Ax\|_2 \leq M\|x\|_2 \forall x \in \mathbb{R}^m\} = \inf\{c \in \mathbb{R} : \forall x \neq 0, c \geq \frac{|Ax|}{|x|}\}$, since

$$\forall M, |A0| \leq M|0|.$$

Observe that $\{c \in \mathbb{R} : \forall x \neq 0, c \geq \frac{|Ax|}{|x|}\}$ is the set of upper bound of $\{\frac{|Ax|}{|x|} : |x| \neq 0\}$, so $\inf\{c \in \mathbb{R} : \forall x, c \geq \frac{|Ax|}{|x|}\} = \|A\| = \sup\{\frac{|Ax|}{|x|} : |x| \neq 0\}$. ■

Proof. In this proof, we use $|\cdot|$ to denote $\|\cdot\|_2$, and if we write x without specification, it belong to \mathbb{R}^m

If $x = 0$, then we trivially have $|Ax| = |0| = 0 \leq \|A\||x| = 0$, so from now, we only have to consider $x \neq 0$.

If $x \neq 0$, we have $|Ax| \leq \|A\||x| \iff \frac{|Ax|}{|x|} \leq \|A\| = \sup\{\frac{|Ax|}{|x|} : x \neq 0\}$, trivially true. ■

Proof. In this proof, we use $|\cdot|$ to denote $\|\cdot\|_2$, and if we write x without specification, it belong to \mathbb{R}^m

For non-negativity, observe $\forall x \neq 0, \frac{|Ax|}{|x|} \geq 0 \implies \|A\| = \sup\{\frac{|Ax|}{|x|} : x \in \mathbb{R}^m, x \neq 0\} \geq 0$.

For definite-positive, observe $A = 0 \implies \forall x \neq 0, \frac{|Ax|}{|x|} = \frac{|0|}{|x|} = 0 \implies \|A\| = 0$. Also, if $A \neq 0$, we can pick a column, say p -th column, that contain non-zero entry. We see the vector $e \in \mathbb{R}^m$ where the only non-zero entry is the p -th entry being 1 satisfy $|Ae| > 0$, thus $\frac{|Ae|}{|e|} > 0$. Because $e \neq 0$, we see $\|A\| = \sup\{\frac{|Ax|}{|x|} : x \in \mathbb{R}^m, x \neq 0\} \geq \frac{|Ae|}{|e|} > 0$.

For absolute-homogenity, let $c \in \mathbb{R}$ and $A \in M_{n \times m}$. We wish to prove $\|cA\| = \sup\{\frac{|cAx|}{|x|} : x \neq 0\} = |c| \sup\{\frac{|Ax|}{|x|} : x \neq 0\} = |c|\|A\|$. Notice that $\frac{|cAx|}{|x|} = \frac{|c||Ax|}{|x|}$, so we only have to prove the more general statement : $c > 0 \implies c \sup X = \sup\{cx : x \in X\}$. Notice that $\forall x \in X, c \sup X \geq cx$, so we have $c \sup X \geq \sup\{cx : x \in X\}$. If $c \sup X$ is not the smallest upper bound, we see there exists cx such that $c \sup X < cx$, and we see $\sup X < x$, causing a contradiction, so we do have $c \sup X = \sup\{cx : x \in X\}$ (done)

For triangle-inequality, first observe $\frac{|(A+B)x|}{|x|} = \frac{|Ax+Bx|}{|x|} \leq \frac{|Ax|+|Bx|}{|x|} = \frac{|Ax|}{|x|} + \frac{|Bx|}{|x|}$. Assume $\|A+B\| > \|A\| + \|B\|$.

Because $\|A\| + \|B\|$ is not an upper bound of $\{\frac{|(A+B)x|}{|x|} : x \neq 0\}$, we know there exists x' such that $\frac{|(A+B)x'|}{|x'|} > \|A\| + \|B\|$. Further, by definition of $\|A\|, \|B\|$, we have

$$\frac{|(A+B)x'|}{|x'|} > \frac{|Ax'|}{|x'|} + \frac{|Bx'|}{|x'|} \quad \text{CaC} \tag{6.1}$$

Question 19

Suppose that S_1, S_2, \dots, S_n are sets in \mathbb{R} and

$$S = \bigcup_{i=1}^n S_i.$$

Define $B_i = \sup S_i$ for $i = 1, \dots, n$.

1. Show that $\sup S = \max\{B_1, B_2, \dots, B_n\}$.
2. If S is the union of an infinite collection of S_i , find the relation between $\sup S$ and B_i .

Proof. Let $\sup S_j = B_j = \max\{B_1, \dots, B_n\}$. We first show $\sup S_j$ is an upper bound of S .

By definition, we have $\forall x \in S_j, x \leq \sup S_j$ and have $\forall i \neq j, \forall x \in S_i, x \leq \sup S_i \leq \sup S_j$, so we have $\forall x \in S, \exists k \in \{1, \dots, n\}, x \in S_k \implies x \leq \sup S_k \leq \sup S_j$ (done)

We now show $\sup S_j$ is the least upper bound of S .

Assume there exists an upper bound of S smaller than $\sup S_j$. Denote that upper bound y . Because y is smaller than $\sup S_j$, we know y is not an upper bound of S_j , so we know there is a number $z \in S_j$ greater than y . Observe that the fact y is an upper bound of S implies y is greater than or equal to $z \in S_j \subseteq S$ CaC (done) ■

Proof. We prove $\sup S = \sup\{B_i\}$.

Notice $\sup S$ is an upper bound of S_i , so we have $\forall i, \sup S \geq \sup S_i = B_i$. This means $\sup S$ is an upper bound of $\{B_i\}$. We have proved $\sup S \geq \sup\{B_i\}$. Assume $\sup S > \sup\{B_i\}$. Then because $\sup\{B_i\}$ is not an upper bound of S , we know there exists $s \in S$ such that $s > \sup\{B_i\}$. But because $S = \bigcup\{S_i\}$, we know $\exists S_j, s \in S_j$, which give us $s \leq \sup S_j = B_j \leq \sup\{B_i\}$ CaC (done) ■

Question 20

Let A be a non-empty set of \mathbb{R} which is bounded below. Define the set $-A$ by

$$-A \equiv \{-x \in \mathbb{R} : x \in A\}.$$

Prove that

$$\inf(A) = -\sup(-A).$$

Proof. Observe $\forall x \in -A, \sup(-A) \geq x \implies \forall a \in A, -\sup(-A) \leq a$. So $-\sup(-A)$ is an lower bound of A . Assume $-\sup(-A)$ is not the greatest lower bound of A (greatest lower bound exists because bounded below and completeness). Let $b > -\sup(-A)$ be another lower bound of A . We have $-b < \sup(-A)$, so we know $-b$ is not an upper bound of $-A$, then we know $\exists x \in -A, -b < x$. Then we know $\exists a \in A, -b < -a$, which implies $\exists a \in A, b > a$, but b is an lower bound of A CaC ■

Question 21

1. Let A, B be non-empty subsets of \mathbb{R} . Define $A + B$ as

$$A + B = \{x + y : x \in A, y \in B\}.$$

Justify if the following statements are true or false by providing a proof for the true statements and giving a counter-example for the false ones.

- (a) $\sup(A + B) = \sup A + \sup B$.
- (b) $\inf(A + B) = \inf A + \inf B$.
- (c) $\sup(A \cap B) \leq \min\{\sup A, \sup B\}$.
- (d) $\sup(A \cap B) = \min\{\sup A, \sup B\}$.
- (e) $\sup(A \cup B) \geq \max\{\sup A, \sup B\}$.
- (f) $\sup(A \cup B) = \max\{\sup A, \sup B\}$.

Proof. We prove $\sup(A + B) = \sup A + \sup B$. For all $a + b \in A + B$, we by definition have $a \leq \sup A, b \leq \sup B$, so we have $a + b \leq \sup A + \sup B$. This prove $\sup A + \sup B$ is an upper bound of $A + B$. Assume there exists an upper bound x of $A + B$ smaller than $\sup A + \sup B$. We have $x - \sup B < \sup A$, so we know $x - \sup B$ is not an upper bound of A . Then we know $\exists a' \in A, x - \sup B < a'$. This implies $x - a' < \sup B$, so we know $x - a'$ is not an upper bound of B . Then we know there exists $b' \in B$ such that $x - a' < b'$. This implies $x < a' + b' \in A + B$ CaC to x is an upper bound of $A + B$ ■

Proof. We prove $\inf(A + B) = \inf A + \inf B$. For all $a + b \in A + B$, we by definition have $\inf A \leq a, \inf B \leq b$, so we have $\inf A + \inf B \leq a + b$. This prove $\inf A + \inf B$ is an lower bound of $A + B$. Assume there exists an lower bound x of $A + B$ greater than $\inf A + \inf B$. We have $x - \inf A > \inf B$, so we know $x - \inf A$ is not an lower bound of B . Then we

know $\exists b' \in B, x - \inf A > b'$. This implies $x - b' > \inf A$, so we know $x - b'$ is not a lower bound of A . Then we know $\exists a' \in A, x - b' > a'$. So we know $x < a' + b' \in A + B$ **CaC** to x is a lower bound of $A + B$ ■

Proof. We prove $\sup(A \cap B) \leq \min\{\sup A, \sup B\}$.

WOLG, let $\sup A \leq \sup B$. By definition $x \in A \cap B \implies x \in A \implies x \leq \sup A$, so we know $\sup A$ is an upper bound of $A \cap B$. This implies $\sup A \cap B \leq \sup(A \cap B)$ ■

Proof. We show $\sup(A \cap B) = \min\{\sup A, \sup B\}$ is not always correct. Let $A = [0, 2]$ and $B = [0, 1] \cup [3, 4]$. We have $\sup(A \cap B = [0, 1]) = 1 \neq \min\{\sup A = 2, \sup B = 4\}$ ■

Proof. $\sup A \cup B$ is an upper bound of both A and B , so $\sup A \cup B > \sup A$ and $\sup A \cup B > \sup B$ ■

Proof. We prove $\sup(A \cup B) = \max\{\sup A, \sup B\}$. WOLG, let $\sup A \geq \sup B$. Assume **$\sup B \leq \sup A < \sup(A \cup B)$** . Let x be a number between $\sup A$ and $\sup A \cup B$ (x exists since it can be $\frac{\sup A + \sup(A \cup B)}{2}$). Because $x < \sup(A \cup B)$, we know x is not an upper bound of $A \cup B$. By definition, we know there exists $z \in A \cup B$ such that $x < z$. We know either $z \in A$ or $z \in B$, but we see $z \in A \implies z \leq \sup A < x$ and we see $z \in B \implies z \leq \sup B \leq \sup A < x$ **CaC** ■

Question 22

7. Let $S \subseteq \mathbb{R}$ be bounded below and non-empty. Show that

$$\inf S = \sup\{x \in \mathbb{R} : x \text{ is a lower bound for } S\}.$$

Proof. Denote $B = \{x \in \mathbb{R} : x \text{ is a lower bound for } S\}$. Assume **$\inf S > \sup B$** . Let $\inf S > x > \sup B$. Notice $\inf S > x$ implies there is a lower bound b of S greater than x . Observe $b > x$ and $b \in B \implies x$ is not an upper bound of B **CaC** $x > \sup B$.

Assume **$\inf S < \sup B$** . Let $\inf S < x < \sup B$. Notice $\inf S < x$ implies there exists $s' \in S$ such that $s' < x$, and notice $x > \sup B$ implies there exists $b' \in B$ such that $b' > x$. We see $b' > x > s'$ while b' is a lower bound of S **CaC** ■

Question 23

8. Let f be a continuous function on \mathbb{R} and D is a dense subset in \mathbb{R} . Prove that:

$$1. \sup_{x \in D} f(x) = \sup_{x \in \mathbb{R}} f(x).$$

2. There exists a sequence $\{x_n\}_{n=1}^{\infty}$ in D such that

$$\lim_{n \rightarrow \infty} f(x_n) = \sup_{x \in \mathbb{R}} f(x).$$

Proof. Because $D \subseteq \mathbb{R}$, we must have $\sup\{f(x) : x \in D\} \leq \sup\{f(x) : x \in \mathbb{R}\}$, since any upper bound of the latter will be one of the former.

Assume $\sup\{f(x) : x \in D\} < \sup\{f(x) : x \in \mathbb{R}\}$. Because $\sup\{f(x) : x \in D\}$ is not an upper bound of $\{f(x) : x \in \mathbb{R}\}$, we know there exists $x' \in \mathbb{R}$ such that

$$f(x') > \sup\{f(x) : x \in D\} \quad (6.2)$$

We first construct a sequence $\{x_i\}_{i=1}^{\infty}$ in D such that

$$\lim_{i \rightarrow \infty} x_i = x' \text{ which reads } \forall \epsilon, \exists N, n > N \longrightarrow |x_n - x'| < \epsilon \quad (6.3)$$

By Axiom of Choice and the fact that D is dense in \mathbb{R} , we can pick $x_i \in (x' - \frac{1}{i}, x' + \frac{1}{i})$, so we have

$$\forall i, |x_i - x'| < \frac{1}{i} \quad (6.4)$$

For all ϵ , we can pick a natural $N > \frac{1}{\epsilon}$, so we have

$$n > N \implies |x_n - x'| < \frac{1}{n} < \frac{1}{N} < \frac{1}{\epsilon} \text{ (done)} \quad (6.5)$$

We now prove

$$\lim_{i \rightarrow \infty} f(x_i) = f(x') \text{ which reads } \forall \epsilon, \exists N, n > N \longrightarrow |f(x_n) - f(x')| < \epsilon \quad (6.6)$$

Because f is continuous, we have

$$\forall \epsilon, \exists \delta, \forall u \in \mathbb{R}, |u - x'| < \delta \implies |f(u) - f(x')| < \epsilon \quad (6.7)$$

Then for all ϵ , we can first let δ satisfy $|u - x'| < \delta \implies |f(u) - f(x')| < \epsilon$. Then by the **violet fact** we can pick N such that

$$n > N \implies |x_n - x'| < \delta \implies |f(x_n) - f(x')| < \epsilon \text{ (done)} \quad (6.8)$$

Let $H = \sup\{f(x) : x \in D\}$. Now we prove

$$\lim_{i \rightarrow \infty} f(x_i) \leq H \quad (6.9)$$

Assume $f(x') = \lim_{i \rightarrow \infty} f(x_i) > H$. We know there exists N such that

$$n > N \longrightarrow |f(x_n) - f(x')| < |H - f(x')| = f(x') - H \quad (6.10)$$

The last equality hold true due to the premise equation (5.2). Notice

$$|f(x_n) - f(x')| < f(x') - H \implies H - f(x') < f(x_n) - f(x') \implies f(x_n) > H \quad (6.11)$$

so in fact we know there exists N such that

$$n > N \longrightarrow f(x_n) > H = \sup\{f(x) : x \in D\} \text{ CaC (done)} \quad (6.12)$$

Now, using all our proven facts, we have

$$\sup\{f(x) : x \in D\} < f(x') = \lim_{i \rightarrow \infty} f(x_i) \leq H = \sup\{f(x) : x \in D\} \text{ CaC} \quad (6.13)$$

where the first inequality follows from premise equation (5.2) ■

Proof. Let $H = \sup\{f(x) : x \in D\}$. We first prove

$$\forall i \in \mathbb{N}, \{f(x) : x \in D \text{ and } H - f(x) < \frac{1}{i}\} \neq \emptyset \quad (6.14)$$

Assume there exists some $n \in \mathbb{N}$ such that the set is empty. We then have

$$\forall x \in D, H \geq f(x) + \frac{1}{n} \quad (6.15)$$

So we have

$$\forall x \in D, H - \frac{1}{2n} \geq f(x) + \frac{1}{2n} > f(x) \quad (6.16)$$

Then we see $H - \frac{1}{2n}$ is an upper bound of $\{f(x) : x \in D\}$ smaller than H CaC (done)

By Axiom of Choice, we can construct a sequence $\{x_i\}_{i=1}^{\infty}$ by picking $x_i : f(x_i) \in \{f(x) : x \in D \text{ and } H - f(x) < \frac{1}{i}\}$. Then we have

$$\forall \epsilon, n > \left[\frac{1}{\epsilon}\right] + 1 \implies n > \frac{1}{\epsilon} \implies |f(x_n) - H| < \frac{1}{n} < \epsilon \quad (6.17)$$

This written in limit sign is

$$\lim_{i \rightarrow \infty} f(x_i) = H = \sup\{f(x) : x \in D\} = \sup\{f(x) : x \in \mathbb{R}\} \quad (6.18)$$

■

6.2 HW2

Question 24

Let $A \subseteq \mathbb{R}^n$ be an open set and $B \subset \mathbb{R}^n$ be any set. Then the set

$$A + B \equiv \{a + b : a \in A \text{ and } b \in B\}$$

is open.

Proof. Notice

$$A + B = \bigcup \{\{a + b : a \in A\} : b \in B\} \quad (6.19)$$

We only have to prove for all $b \in B$, the set $A + b := \{a + b : a \in A\}$ is open.

Fix b . Arbitrarily pick $a + b \in A + b$. Because A is open, we know there exists $r \in \mathbb{R}^+$ such that

$$B_r(a) \subseteq A \quad (6.20)$$

Let

$$B_r(a) + b := \{x + b : x \in B_r(a)\} \quad (6.21)$$

We now prove

$$B_r(a) + b = B_r(a + b) \quad (6.22)$$

Arbitrarily pick $x + b \in B_r(a) + b$. We have

$$|(x + b) - (a + b)| = |x - a| < r \quad (6.23)$$

We have proved $B_r(a) + b \subseteq B_r(a + b)$. Arbitrarily pick $y \in B_r(a + b)$. Let $z = y - b$. We have

$$y = z + b \text{ and } |z - a| = |y - (a + b)| < r \quad (6.24)$$

The latter tell us $z \in B_r(a)$, so we have

$$y = z + b \in B_r(a) + b \quad (6.25)$$

Because y is arbitrarily picked from $B_r(a + b)$, we have proved $B_r(a + b) \subseteq B_r(a) + b$ (done)

Notice that r is selected to satisfy

$$B_r(a) \subseteq A \quad (6.26)$$

and it is clear that

$$B_r(a) + b \subseteq A + b \quad (6.27)$$

So we have

$$B_r(a + b) = B_r(a) + b \subseteq A + b \quad (6.28)$$

Notice $a + b$ is arbitrarily picked from $A + b$. Our proof is done (done) ■

Question 25

Show that every open set in \mathbb{R} is the union of at most countable collection of disjoint open intervals.

Proof. Because \mathbb{Q} is countable and dense in \mathbb{R} , so in Theorem ??, we have proved the set

$$\mathcal{O} := \{B_r(p) : p \in \mathbb{Q} \text{ and } r \in \mathbb{Q}^+\} \quad (6.29)$$

is a countable base. Notice that every open ball $B_r(p)$ can be expressed as $(p - r, p + r)$, so we know \mathcal{O} is a countable collection of open interval.

Let A be an open set. We wish to prove *A can be expressed as union of a countable collection of disjoint open interval.*

Because \mathcal{O} is a base, for each $a \in A$, we can find $B_{r_a}(p_a)$, such that

$$a \in B_{r_a}(p_a) \subseteq A \quad (6.30)$$

Collect all such open ball for each points in A and call this collection \mathcal{O}' .

Define a relation \sim on \mathcal{O}' by

$$(a, b) \sim (c, d) \text{ if } \exists S \in \mathcal{P}(\mathcal{O}') \text{ such that} \quad (6.31)$$

$$\bigcup S \text{ is an open interval and } (a, b) \subseteq \bigcup S \text{ and } (c, d) \subseteq \bigcup S \quad (6.32)$$

We wish to prove *\sim is an equivalence relation.*

To show $(a_1, b_1) \sim (a_1, b_1)$, use $S = \{(a_1, b_1)\}$. To show $(a_1, b_1) \sim (a_2, b_2) \implies (a_2, b_2) \sim (a_1, b_1)$, use the same S . It is left to prove

$$(a_1, b_1) \sim (a_2, b_2) \text{ and } (a_2, b_2) \sim (a_3, b_3) \implies (a_1, b_1) \sim (a_3, b_3) \quad (6.33)$$

Let $S_1 \in \mathcal{P}(\mathcal{O}')$ be from $(a_1, b_1) \sim (a_2, b_2)$, and let $S_2 \in \mathcal{P}(\mathcal{O}')$ be from $(a_2, b_2) \sim (a_3, b_3)$.

Define

$$S_3 := S_1 \cup S_2 \quad (6.34)$$

Because $(a_2, b_2) \subseteq \bigcup S_1 \cap \bigcup S_2$, we know $\bigcup S_1 \cap \bigcup S_2$ are not disjoint. Then because union of two intersecting open interval is again an open interval, we know $\bigcup S_3$ is an open interval.

Deduce

$$(a_1, b_1) \subseteq \bigcup_{170} S_1 \subseteq \bigcup S_3 \quad (6.35)$$

and deduce

$$(a_3, b_3) \subseteq \bigcup S_2 \subseteq \bigcup S_3 \text{ (done)} \quad (6.36)$$

Let E be the collection of union of each equivalent class of \sim on \mathcal{O}' . Because \mathcal{O}' is countable, we know E is countable. Every element of \mathcal{O}' is a subset of A by definition, so we know $\bigcup E = \bigcup \mathcal{O}' \subseteq A$. Every element of A is in some $O \in \mathcal{O}'$ by definition, so we have $A \subseteq \bigcup \mathcal{O}' = \bigcup E$. It is only left to prove **each member in E is an open interval** and **each two members of E are disjoint**.

We first prove **disjoint**.

Let $B, C \in E$ and let B', C' be equivalent class that satisfy

$$B = \bigcup B' \text{ and } C = \bigcup C' \quad (6.37)$$

Assume **B, C is not disjoint**, say, **$x \in B \cap C$** , we have

$$\exists (a, b) \in B', x \in (a, b) \subseteq B \quad (6.38)$$

and have

$$\exists (c, d) \in C', x \in (c, d) \subseteq C \quad (6.39)$$

We then can see (a, b) and (c, d) intersect, then we can use $S = \{(a, b), (c, d)\}$ to show $(a, b) \sim (c, d)$ **CaC (done)**.

We now prove **open interval**.

Let $B \in E$, and let B' be equivalent class that satisfy $B = \bigcup B'$. We wish to prove $B = (\inf B, \sup B)$ where negative and positive infinity is taken into account.

Because B' contain only open sets, we know B is open. Then we know $\sup B$ and $\inf B$ if exists, is not in B . To see such, just observe every open ball centering $\sup B$ has a number greater than all numbers in B , and similarly for $\inf B$ vice versa.

We have proved $B \subseteq (\inf B, \sup B)$.

We only have to prove every point in $(\inf B, \sup B)$ is in B , where infimum and supremum can be negative or positive infinity.

Assume **$\exists x \in (\inf B, \sup B), x \notin B$** . We have

$$\exists (y_1, z_1) \in B', (y_1, z_1) \subseteq (\inf B, x) \text{ and } \exists (y_2, z_2) \in B', (y_2, z_2) \subseteq (x, \sup B) \quad (6.40)$$

Because $(y_1, z_1) \sim (y_2, z_2)$ there exists $S \in \mathcal{O}'$ such that $(y_1, z_2) \subseteq \bigcup S$ and $\bigcup S$ is an open interval.

If S contain any element (v, w) not in B' , we can use S to show $(v, w) \sim (y_2, z_2)$, causing a contradiction. We have proved S is a subset of B' .

Now, Because $x \in \bigcup S$, we know there must exists some interval in $S \subseteq B'$ containing x CaC to $x \notin B$ (done) (done) ■

Question 26

Let $A \subseteq B \subseteq \mathbb{R}$. Suppose that A is a dense subset of B .

1. Prove that $B \subseteq \overline{A}$.
2. If B is closed, determine whether $B = \overline{A}$.

Proof. A is a dense subset of B means that every point $b \in B \setminus A$ is a limit point of A in the scope of B . Notice that b is a limit point of A in the scope of B also means b is a limit point in the scope of \mathbb{R} . Then we have proved $B \setminus A \subseteq A'$. It follows

$$B = A \cup (B \setminus A) \subseteq A \cup A' = \overline{A} \quad (6.41)$$

If B is closed, we have

$$\overline{A} \subseteq \overline{B} = B \quad (6.42)$$

Then because $B \subseteq \overline{A}$, we have $B = \overline{A}$ ■

Definition 0.1. A metric space X is *sequentially compact* if every sequence of points in X has a convergent sub-sequence converging to a point in X .

Question 27

Let A and B be subsets of a metric space (M, d) and denote $\text{cl}(A) = \overline{A}$. Show that

1. $\text{cl}(\text{cl}(A)) = \text{cl}(A)$.
2. $\text{cl}(A \cup B) = \text{cl}(A) \cup \text{cl}(B)$.
3. $\text{cl}(A \cap B) \subseteq \text{cl}(A) \cap \text{cl}(B)$. Find an example such that $\text{cl}(A \cap B) \subsetneq \text{cl}(A) \cap \text{cl}(B)$.

Proof. Notice that closure is the smallest closed set containing the set.

Because \overline{A} is a closed set, we know the smallest closed set containing \overline{A} is \overline{A} itself.

Deduce

$$A \subseteq A \cup B \implies \overline{A} \subseteq \overline{A \cup B} \quad (6.43)$$

and deduce

$$B \subseteq A \cup B \implies \overline{B} \subseteq \overline{A \cup B} \quad (6.44)$$

so deduce

$$\overline{A} \cup \overline{B} \subseteq \overline{A \cup B} \quad (6.45)$$

Notice that $\overline{A} \cup \overline{B}$ is a closed set containing both A and B , thus containing $A \cup B$, so because $\overline{A \cup B}$ by definition is the smallest closed set containing $A \cup B$, we have

$$\overline{A \cup B} \subseteq \overline{A} \cup \overline{B} \quad (6.46)$$

Notice that \overline{A} and \overline{B} are both closed set containing $A \cap B$. Then because $\overline{A \cap B}$ is the smallest closed set containing $A \cap B$. We have

$$\overline{A \cap B} \subseteq \overline{A} \text{ and } \overline{A \cap B} \subseteq \overline{B} \quad (6.47)$$

Then have

$$\overline{A \cap B} \subseteq \overline{A} \cap \overline{B} \quad (6.48)$$

Let $A = (0, 1)$ and $B = (1, 2)$. We have

$$\overline{A \cap B} = \overline{\emptyset} = \emptyset \text{ and } \overline{A} \cap \overline{B} = [0, 1] \cap [1, 2] = \{1\} \quad (6.49)$$

■

Question 28

Let K be a sequentially compact set in a metric space (M, d) and let $F \subseteq K$ be closed. Prove that F is sequentially compact.

Proof. Let $\{x_i\}_{i=1}^{\infty}$ be a sequence in F

$$\{x_i\}_{i=1}^{\infty} \subseteq F \subseteq K \quad (6.50)$$

Because K is sequentially compact, we know there exists a sub-sequence $\{x_{n_i}\}_{i=1}^{\infty}$ converge to some point a

$$\lim_{i \rightarrow \infty} \{x_{n_i}\} = a \quad (6.51)$$

We only have to prove a is in F .

Assume $a \notin F$. Because F is closed, we know F^c is open. Then $a \in F^c$ tell us there exists an open ball $B_r(a)$ contain no point in F , thus disjoint to $\{x_{n_i}\}$ **CaC**

■

Definition 0.2. The discrete metric d on a set X is defined by

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y, \\ 0 & \text{if } x = y, \end{cases}$$

for any $x, y \in X$. In this case (X, d) is called a discrete metric space.

Question 29

Let (M, d) be a metric space with discrete metric d . Prove that every compact set in M is finite.

Proof. Notice we have

$$\forall p \in M, B_1(p) = \{p\} \quad (6.52)$$

So we know no set in M has a limit point. Let $K \subseteq M$ be compact. By Theorem ??, we know K is limit point compact. Assume K is infinite. Then K has a limit point x . ■

Question 30

Let $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ be a convergent sequence with the limit x . Prove that the set $\{x_n : n \in \mathbb{N}\} \cup \{x\}$ is compact.

Proof. Let

$$E \subseteq \{x_n : n \in \mathbb{N}\} \cup \{x\} \quad (6.53)$$

If E is infinite, we can regard $E \setminus \{x\}$ as a sub-sequence of $\{x_n\}$. Then because $\{x_n\}_{n=1}^{\infty}$ converge to x , we know $E \setminus \{x\}$ converge to x . Then we know x is a limit point of $E \setminus \{x\}$ thus a limit point of E . We have proved every infinite subset of $\{x_n : n \in \mathbb{N}\} \cup \{x\}$ has a limit point x in $\{x_n : n \in \mathbb{N}\} \cup \{x\}$. By Theorem ??, $\{x_n : n \in \mathbb{N}\} \cup \{x\}$ is compact. ■

Question 31

1. Let A and B be two subsets of a metric space (M, d) . The distance between A and B is defined by

$$d(A, B) \equiv \inf\{d(x, y) : x \in A, y \in B\}.$$

- (a) Give an example of two disjoint, nonempty, closed sets A and B in \mathbb{R}^n for which $d(A, B) = 0$.
- (b) Let A, B be nonempty sets in \mathbb{R}^n with A closed and B compact. Show that there are points $a \in A$ and $b \in B$ such that $d(a, b) = |a - b|$. Deduce that $d(A, B)$ is positive if such A, B are disjoint.

Proof. Let $n = 1$. Let

$$A = \{a_k = k - 1 + \sum_{i=1}^{k-1} 2^{-i-1} : k \in \mathbb{N}\} \quad (6.54)$$

Let

$$B = \{b_k = k - \sum_{i=1}^{k-1} 2^{-i-1} : u \in \mathbb{N}\} \quad (6.55)$$

Notice the pattern

$$a_1 = 0 < \frac{1}{2} < 1 = b_1 \quad (6.56)$$

$$a_2 = 1 + \frac{1}{4} < 1 + \frac{1}{2} < 2 - \frac{1}{4} = b_2 \quad (6.57)$$

$$a_3 = 2 + \frac{1}{4} + \frac{1}{8} < 2 + \frac{1}{2} < 3 - \frac{1}{4} - \frac{1}{8} = b_3 \quad (6.58)$$

$$\vdots \quad (6.59)$$

and compute

$$b_k - a_k = 2^{1-k} \quad (6.60)$$

First observe for each non-negative integer m the interval $[m, m+1]$ contain exactly one point from A and contain exactly one point from B . To see A and B are disjoint, observe that because A, B contain only non-negative numbers, so every point is in $[m, m+1]$ for some non-negative integer m , and we know each interval only contain a_{m+1} from A and b_{m+1} from B where $b_{m+1} - a_{m+1} = 2^{-m}$.

We now show A, B are both closed. For negative $-x \in \mathbb{R}$, observe $B_x(-x)$ intersect with neither A nor B . We know $B_1(0)$ does not intersect with B and $0 \in A$. For positive x not in A , let $m = \lfloor x \rfloor$. We know point closest to x is either a_m, a_{m+1} or a_{m+2} . Then we can let $r = \min\{|x - a_m|, |x - a_{m+1}|, |x - a_{m+2}|\}$, and see $B_r(x)$ does not intersect with A . For positive x not in B , let $m = \lfloor x \rfloor$. We know point closest to x is either b_m, b_{m+1} or b_{m+2} . Then we can let $r = \min\{|x - b_m|, |x - b_{m+1}|, |x - b_{m+2}|\}$, and see $B_r(x)$ does not intersect with B .

Lastly, to prove $d(A, B) = 0$ is to prove for each positive real r , there exists a pair of a, b such that $|a - b| < r$. Observe

$$b_k - a_k = 2^{1-k} < r \iff 1 - k < \log_2 r \iff k > 1 - \log_2 r \quad (6.61)$$

And we are done, by picking great enough k .

Now we do (b).

Because

$$\{d(a, b) : a \in A, b \in B\} = \bigcup \{\{d(a, b) : a \in A\} : b \in B\} \quad (6.62)$$

so we have

$$d(A, B) = \inf \{d(a, b) : a \in A, b \in B\} \quad (6.63)$$

$$= \inf \bigcup \{\{d(a, b) : a \in A\} : b \in B\} \quad (6.64)$$

$$= \inf \{\inf \{d(a, b) : a \in A\} : b \in B\} \quad (6.65)$$

$$= \inf \{d(A, \{b\}) : b \in B\} \quad (6.66)$$

where the third equality hold true by the fourth question of HW1.

We first prove if $d(A, B) = 0$, then A and B intersect.

If B is finite, then $0 = d(A, B) = \min \{d(A, \{b\}) : b \in B\}$, so we know there exists $b \in B$ such that $0 = d(A, \{b\}) = \inf \{d(a, b) : a \in A\}$. Then we can deduce every open ball $B_r(b)$ contain a point $a \in A$, since there exists a such that $d(a, b) < r$. We have proved b is a limit point of A , then because A is closed, we know $b \in A$.

If B is infinite, then $0 = \inf \{d(A, \{b\}) : b \in B\}$. For each n , we then know there exists b_n such that $d(A, \{b_n\}) < \frac{1}{n}$.

Construct an infinite sequence $\{b_n\}_{n \in \mathbb{N}}$ by picking b_n such that $d(A, \{b_n\}) < \frac{1}{n}$. We can see our method for picking b_n will give us an infinite sequence as every b satisfy $d(A, \{b\}) > 0$. By Theorem ??, we know B is limit point compact, so we know $\{b_n\}_{n \in \mathbb{N}}$ has a limit point in B .

Denote the limit point for $\{b_n\}_{n \in \mathbb{N}}$ as p . We now show $p \in A$. Assume $p \in A^c$. Because A is closed, we know there exists an open ball $B_r(p)$ disjoint with A . Then we know $d(A, \{p\}) \geq r$.

Notice that if we fix $b' \in B$. We have triangle inequality

$$d(A, \{p\}) \leq d(A, \{b'\}) + d(b', p) \quad (6.67)$$

Because

$$d(A, \{b'\}) + d(b', p) = \inf \{d(a, b') : a \in A\} + d(b', p) = \inf \{d(a, b') + d(b', p) : a \in A\} \quad (6.68)$$

where for each a , we have

$$d(A, \{p\}) \leq d(a, p) \leq d(a, b') + d(b', p) \quad (6.69)$$

Then from

$$\forall b \in B, r \leq d(A, \{p\}) \leq d(A, \{b\}) + d(\{b\}, \{p\}) \quad (6.70)$$

we have

$$\forall b \in B [d(A, \{b\}) < \frac{r}{2} \implies d(\{b\}, \{p\}) > \frac{r}{2}] \quad (6.71)$$

Then we see if $n > \frac{2}{r}$, then b_n is not contained by the open ball $B_{\frac{r}{2}}(p)$. In other words, the open ball $B_{\frac{r}{2}}(p)$ that center a limit point p contain at most $\lfloor \frac{2}{r} \rfloor$ amount of point in $\{b_n\}_{n \in \mathbb{N}}$
CaC (done)

Lastly, we prove there exists $a \in A, b \in B$ such that $d(A, B) = d(a, b)$. If A and B are not disjoint, the fact $\exists a \in A, \exists b \in B, d(A, B) = |a - b|$ is clear by picking $a = b$. We only have to consider when A and B are disjoint.

Let $d(A, B) = u$. We know $u > 0$, because of the result we have proved.

We have

$$\inf\{d(A, \{b\}) : b \in B\} = d(A, B) = u \quad (6.72)$$

At this stage, we wish to prove $\exists b \in B$ such that $d(A, \{b\}) = u$. Assume **there does not exist such b**

We know, for each $n \in \mathbb{N}$, there exists a point b_n such that

$$d(A, \{b_n\}) < u + \frac{1}{n} \quad (6.73)$$

Because $\forall b \in B, d(A, \{b\}) > u$, we know if we collect one b_n that satisfy $d(A, \{b_n\}) < u + \frac{1}{n}$ for each $n \in \mathbb{N}$, the sequence $\{b_n\}_{n \in \mathbb{N}}$ is infinite.

Then by Theorem ??, we know there exists a limit point $p \in B$ for $\{b_n\}_n \in \mathbb{N}$.

We know $d(A, \{p\}) > u$ by our assumption.

Let $m = d(A, \{p\}) > u$. Then from

$$\forall b \in B, m = d(A, \{p\}) < d(A, \{b\}) + d(b, p) \text{ This was proved at } * \quad (6.74)$$

we have

$$\forall b \in B, d(A, \{b\}) < u + \frac{1}{n} \implies d(b, p) > m - u - \frac{1}{n} \quad (6.75)$$

Let k be a natural such that $u + \frac{1}{k} < m$. For each natural t greater than k , we have

$$d(b_t, p) > m - u - \frac{1}{t} > m - u - \frac{1}{k} \quad (6.76)$$

In other words, the open ball $B_{m-u-\frac{1}{k}}(p)$ centering a limit point contain at most k amount of point in $\{b_n\}_{n \in \mathbb{N}}$ **CaC** (done)

Let $b \in B$ satisfy $d(A, \{b\}) = u$. At the final stage, we wish to prove $\exists a \in A, d(a, b) = u$.

Assume $\forall a \in A, d(a, b) > u$. We know, for each $n \in \mathbb{N}$, there exists a point a_n such that

$$d(a_n, b) < u + \frac{1}{n} \quad (6.77)$$

Because $\forall a \in A, d(a, b) > u$, we know if we collect one a_n that satisfy $d(a_n, b) < u + \frac{1}{n}$ for each $n \in \mathbb{N}$, the sequence $\{a_n\}_{n \in \mathbb{N}}$ is infinite.

Notice that the sequence $\{a_n\}_{n \in \mathbb{N}}$ is also bounded as one can check $B_{2(u+1)}(a_1)$ contain all $\{a_n\}_{n \in \mathbb{N}}$, using b as a "pivot".

Then because we are in \mathbb{R}^n and $\{a_n\}_{n \in \mathbb{N}}$ is closed and bounded, we know $\{a_n\}_{n \in \mathbb{N}}$ is compact, then by Theorem ??, we know $\{a_n\}_{n \in \mathbb{N}}$ has a limit point in itself.

Let a_k be a limit point of $\{a_n\}_{n \in \mathbb{N}}$.

Let $m = d(a_k, b) > u$. Observe that

$$\forall a_n, m = d(a_k, b) < d(b, a_n) + d(a_n, a_k) \quad (6.78)$$

give us

$$\forall a_n, d(a_n, b) < u + \frac{1}{n} \implies d(a_n, a_k) > m - u - \frac{1}{n} \quad (6.79)$$

Let $s \in \mathbb{N}$ be great enough so that $m - u - \frac{1}{s} > 0$. Then we see for each natural w greater than s , we have

$$d(a_w, b) < u + \frac{1}{w} < u + \frac{1}{s} \quad (6.80)$$

so we have

$$d(a_w, a_k) > m - u - \frac{1}{s} \quad (6.81)$$

Then we see the open ball $B_{m-u-\frac{1}{s}}(a_k)$ contain at most s amount of point in $\{a_n\}_{n \in \mathbb{N}}$ **CaC** (done) ■

Question 32

9. Let (M, d) be a metric space.

- (a) Show that the union of a finite number of compact subsets of M is compact.
- (b) Show that the intersection of an arbitrary collection of compact subsets of M is compact.

Proof. Let $\mathcal{K} = \{K_1, \dots, K_n\}$ be a finite collection of compact subset of M . We wish to show $\bigcup \mathcal{K}$ is compact. Let \mathcal{G} be an open cover for \mathcal{K} . For each $m \in \mathbb{N} : 1 \leq m \leq n$, because $K_m \subseteq \bigcup \mathcal{K}$, we know \mathcal{G} is also an open cover for K_m . Then, we can pick a finite sub-cover $\mathcal{G}_m \subseteq \mathcal{G}$ for K_m . Collect all such finite sub-cover $\mathcal{G}_1, \dots, \mathcal{G}_m$. The union $\bigcup \{\mathcal{G}_1, \dots, \mathcal{G}_m\}$ is a finite open cover for $\bigcup \mathcal{K}$.

Notice that a compact set must be closed, and arbitrary collection of closed sets are closed. Then, we know the intersection of an arbitrary collection of compact set must be a closed set and a subset of every compact set in the collection. Then because closed subset of compact set is compact, we know the intersection is compact. ■

Question 33

10. A metric space (M, d) is said to be *separable* if there is a countable subset A which is dense in M . Show that every compact set is separable.

Proof. See the proof in Theorem ??, there is a summary where you can find the numbering of rigorous proof for each step. Some proof for corollary is omitted because they are immediate consequences of theorem before.

Because I use a program written by artificial intelligence to update my reference in latex code whenever I edit my latex code, some reference may be directed to a wrong theorem, but all the tools are there in Chapter 2. ■

6.3 HW3

Question 34

1. Let S_1 and S_2 be two nonempty subsets in a metric space with $S_1 \cap \overline{S_2} = \overline{S_1} \cap S_2 = \emptyset$. If $A \subseteq S_1 \cup S_2$ is a connected set, then either $A \subseteq S_1$ or $A \subseteq S_2$.

Proof. Assume **we have** $A \not\subseteq S_1$ and $A \not\subseteq S_2$. Then we know $A \cap S_1$ and $A \cap S_2$ are both non-empty. We know

$$A = (A \cap S_1) \cup (A \cap S_2) \quad (6.82)$$

We wish to show

$$A \cap S_1 \text{ and } A \cap S_2 \text{ are separated.} \quad (6.83)$$

Notice

$$\overline{A \cap S_1} \subseteq \overline{S_1} \quad (6.84)$$

Then because $\overline{S_1}$ and S_2 are disjoint, we know $\overline{A \cap S_1}$ and S_2 are disjoint. Then because $A \cap S_2 \subseteq S_2$, we know $\overline{A \cap S_1}$ and $A \cap S_2$ are disjoint. Similarly, notice

$$\overline{A \cap S_2} \subseteq \overline{S_2} \quad (6.85)$$

Then because $\overline{S_2}$ and S_1 are disjoint, we know $\overline{A \cap S_2}$ and S_1 are disjoint. Then because $A \cap S_1 \subseteq S_1$, we know $\overline{A \cap S_2}$ and $A \cap S_1$ are disjoint.

We have proved $A \cap S_1$ and $A \cap S_2$ are separated, which **CaC** to A is connected. ■

Theorem 6.3.1. (Subspace Topology) Let Y be a subspace of (X, d) , let $E \subseteq X$, and let $p \in Y$. We have

$$p \text{ is a limit point of } E \cap Y \text{ in } Y \implies p \text{ is a limit point of } E \text{ in } X \quad (6.86)$$

$$\text{The closure of } E \cap Y \text{ in } Y \text{ is a subset of the closure of } E \text{ in } X \quad (6.87)$$

$$E \text{ is closed in } X \implies E \cap Y \text{ is closed in } Y \quad (6.88)$$

where the converse of the first and the third statement may not hold true, and the subset relation in the second statement may be proper.

Proof. For the first statement, use the fact a sequence in $E \cap Y$ is in E .

For a non-trivial example of the converse in first statement mat not hold true, let $X = \mathbb{R}$, let $Y = \{1\} \cup [2, 3]$ and let $E = (0, 3)$. We see 1 and 0 are both limit points of E in X , and neither of them is a limit point of $E \cap Y$ in Y , as $1 \in Y$ is an isolated point and 0 is not even in Y .

We now prove the second statement. Let p be in the closure of $E \cap Y$, and let $\{x_n\}$ be a sequence in $E \cap Y$ that converge to p . Notice that the sequence $\{x_n\}$ is in E and we are done.

We now prove the last statement. From left to right, let $\{x_n\}$ be a sequence in $E \cap Y$ that converge to some p in Y . We wish to prove p is in $E \cap Y$. Notice that $\{x_n\}$ is also in E , so we know p is in E . Then because p is in Y , we know $p \in E \cap Y$.

For an non-trivial example of the converse in last statement may not hold true. Let $X = \mathbb{R}$, let $Y = (0, 2)$ and $E = [1, 2)$. Clearly E is not closed, but $E \cap Y = [1, 2)$ is closed in Y (One can even observe that E in Y is even "informally" unbounded).

■

Theorem 6.3.2. (Connected) If A is disconnected, then A can be partitioned into two non-empty disjoint relatively open subsets.

Proof. Because A is disconnected, we know $A = E \cup F$ for some pair E, F of separated sets. We wish to prove E, F are both relatively open to A .

Because $\overline{E} \cap F = \emptyset$, we know the closure of E in subspace topology of A is disjoint to F . Then with respect to A , we can see $\overline{E} \subseteq F^c = E$, so we know E is relatively closed to A .

Similarly, because $\overline{F} \cap E = \emptyset$, we know the closure of F in subspace topology of A is disjoint to E . Then with respect to A , we can see $\overline{F} \subseteq E^c = F$, so we know F is relatively closed to A .

Then because E and F are both relatively closed to A and $A = E \cup F$ where E, F are disjoint, we know E and F are both relatively open to A .

■

Question 35

2. If A_1 and A_2 are two nonempty and connected sets with $A_1 \cap A_2 \neq \emptyset$. Prove or disprove that
- (a) $A_1 \cap A_2$ is connected
 - (b) $A_1 \cup A_2$ is connected

Proof. Let $A_1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ and let $A_2 = \{(0, y) \in \mathbb{R}^2 : y \in \mathbb{R}\}$.

We see $A_1 \cap A_2 = \{(0, 1), (0, -1)\}$. We see $\{(0, 1)\}$ and $\{(0, -1)\}$ are separated, because they are close.

We now prove A_2 is connected.

For (b), see next Theorem. ■

Question 36

3. Let $\{A_k\}_{k=1}^{\infty}$ be a family of connected subsets of M , and suppose that A is a connected subset of M such that $A_k \cap A \neq \emptyset$ for all $k \in \mathbb{N}$. Show that the union $(\bigcup_{k \in \mathbb{N}} A_k) \cup A$ is also connected.

Proof. Assume $\bigcup_{k \in \mathbb{N}} A_k \cup A$ is not connected. Let $\bigcup_{k \in \mathbb{N}} A_k \cup A$ be partitioned into two non-empty E, F relatively open to $\bigcup_{k \in \mathbb{N}} A_k \cup A$.

If E, F both intersect with A , observe that A can be partitioned into two non-empty $E \cap A$ and $F \cap A$, which are relatively open to A , causing a contradiction to A is connected.

Then, we only have to consider when only one of E, F intersect with A . WOLG, let E intersect with A .

We know F must intersect with some A_n . From last question, we know $A_n \cup A$ is connected. Notice that $A_n \cup A$ can be partitioned into two non-empty $E \cap (A \cup A_n)$ and $F \cap (A \cup A_n)$, and they are relatively open to $A \cup A_n$ CaC to $A_n \cup A$ is connected. ■

Question 37

4. Let $\{a_k\}_{k=1}^{\infty}$ be a sequence, and define $s_n = \frac{1}{n} \sum_{k=1}^n a_k$. Prove or disprove that

(a) If a_k converges, then s_n converges.

(b) If s_n converges, then a_k converges.

(c) Let $t = \frac{(2n-1)a_1 + (2n-3)a_2 + \dots + 3a_{n-1} + a_n}{n^2}$. Assume a_k converges to a . Does t_n also converge to a ?

Proof. First notice

$$|s_n - L| = \left| \frac{\sum_{k=1}^n a_k}{n} - L \right| = \left| \frac{(\sum_{k=1}^n a_k) - nL}{n} \right| \quad (6.89)$$

$$= \left| \frac{\sum_{k=1}^n (a_k - L)}{n} \right| \quad (6.90)$$

$$= \frac{1}{n} \left| \sum_{k=1}^n (a_k - L) \right| \quad (6.91)$$

$$\leq \frac{1}{n} \sum_{k=1}^n |a_k - L| \quad (6.92)$$

We prove $\lim_{k \rightarrow \infty} a_k = L \implies \lim_{k \rightarrow \infty} s_k = L$.

Arbitrarily pick $R \in \mathbb{R}^+$. We wish to find N such that

$$k > N \implies |s_k - L| < R \quad (6.93)$$

Because $\lim_{n \rightarrow \infty} a_n = L$, we know there exists N_0 such that

$$n > N_0 \iff |a_n - L| < \frac{R}{2} \quad (6.94)$$

Let

$$H = \sum_{k=1}^{N_0} |a_k - L| \text{ and } m = \frac{2}{R}(H - N_0 R) \quad (6.95)$$

We wish to prove

$$n > N_0 + m \implies |s_n - L| < R \quad (6.96)$$

Let $n = N_0 + u$ where $u > m$. Observe

$$|s_n - L| \leq \frac{1}{n} \sum_{k=1}^n |a_k - L| = \frac{\sum_{k=1}^n |a_k - L|}{N_0 + u} \quad (6.97)$$

$$= \frac{\sum_{k=1}^{N_0} |a_k - L| + \sum_{k=N_0+1}^{N_0+u} |a_k - L|}{N_0 + u} \quad (6.98)$$

$$\leq \frac{H + u \frac{R}{2}}{N_0 + u} \quad (6.99)$$

and Observe

$$u > m = \frac{2}{R}(H - N_0 R) \implies \frac{Ru}{2} > H - N_0 R \quad (6.100)$$

$$\implies N_0 R + Ru > H + \frac{Ru}{2} \quad (6.101)$$

$$\implies R > \frac{H + \frac{Ru}{2}}{N_0 + u} \quad (6.102)$$

In other words, we have

$$n > N_0 + m \implies u > m \implies |s_n - L| \leq \frac{H + \frac{Ru}{2}}{N_0 + u} < R \text{ (done)} \quad (6.103)$$

For (b), we raise an counter-example. Let

$$a_n = \begin{cases} 1 & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases} \quad (6.104)$$

To see $\{a_n\}$ does not converge to any number, consider the sub-sequences

$$\{a_{n_k}\} \text{ where } n_k = 2k - 1 \text{ and } \{a_{n_u}\} \text{ where } n_u = 2u \quad (6.105)$$

We have

$$\forall k, a_{n_k} = 1 \text{ and } \forall u, a_{n_u} = 0 \quad (6.106)$$

Then we see

$$\lim_{k \rightarrow \infty} a_{n_k} = 1 \text{ and } \lim_{u \rightarrow \infty} a_{n_u} = 0 \quad (6.107)$$

If $\{a_n\}$ converge, then these two sub-sequence should converge to the same number.

We wish to show

$$\lim_{n \rightarrow \infty} s_n = \frac{1}{2} \quad (6.108)$$

With simple logical computation, we have

$$s_n = \begin{cases} \frac{n-1}{2n} & \text{if } n \text{ is odd} \\ \frac{1}{2} & \text{if } n \text{ is even} \end{cases} \quad (6.109)$$

Notice

$$\frac{n-1}{2n} = \frac{1}{2} - \frac{1}{2n} \quad (6.110)$$

Fix $\epsilon \in \mathbb{R}^+$. We see that

$$n > \frac{1}{2\epsilon} \implies |s_n - \frac{1}{2}| = \begin{cases} \frac{1}{2n} & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases} < \epsilon \quad (6.111)$$

Lastly, we prove

$$\lim_{n \rightarrow \infty} a_n = a \implies \lim_{n \rightarrow \infty} t_n = a \quad (6.112)$$

First notice

$$|t_n - a| = \left| \frac{(2n-1)a_1 + (2n-3)a_2 + \cdots + a_n}{n^2} - a \right| \quad (6.113)$$

$$= \left| \frac{1}{n^2} [(2n-1)(a_1 - a) + (2n-3)(a_2 - a) + \cdots + (a_n - a)] \right| \quad (6.114)$$

$$= \frac{1}{n^2} \left| \sum_{k=1}^n (2n-2k+1)(a_k - a) \right| \quad (6.115)$$

$$\leq \frac{1}{n^2} \sum_{k=1}^n |(2n-2k+1)(a_k - a)| \quad (6.116)$$

$$= \frac{1}{n^2} \sum_{k=1}^n (2n-2k+1)|a_k - a| \quad (6.117)$$

$$\leq \frac{1}{n^2} \sum_{k=1}^n (2n-1)|a_k - a| = \frac{2n-1}{n^2} \sum_{k=1}^n |a_k - a| \quad (6.118)$$

Arbitrarily pick $\epsilon \in \mathbb{R}^+$. Because $\lim_{n \rightarrow \infty} a_n = a$. We know there exists N_0 such that

$$n > N_0 \implies |a_n - a| < \frac{\epsilon}{4} \implies \left(\frac{2n-1}{n}\right)|a_n - a| = \left(2 - \frac{1}{n}\right)|a_n - a| < 2|a_n - a| < \frac{\epsilon}{2} \quad (6.119)$$

Let

$$H = \sum_{k=1}^{N_0} |a_k - a| \quad (6.120)$$

We have

$$|t_n - a| \leq \frac{2n-1}{n^2} \sum_{k=1}^n |a_k - a| = \frac{2n-1}{n^2} \sum_{k=1}^{N_0} |a_k - a| + \frac{2n-1}{n^2} \sum_{k=N_0+1}^n |a_k - a| \quad (6.121)$$

$$= \frac{2n-1}{n^2} H + \frac{2n-1}{n^2} \sum_{k=N_0+1}^n |a_k - a| \quad (6.122)$$

$$= \frac{2n-1}{n^2} H + \frac{1}{n} \sum_{k=N_0+1}^n \frac{2n-1}{n} |a_k - a| \quad (6.123)$$

$$\leq \frac{2n-1}{n^2} H + \frac{1}{n} (n - N_0) \frac{\epsilon}{2} \quad (6.124)$$

Then if we let

$$X_n := \frac{2n-1}{n^2} H + \frac{(n - N_0)\epsilon}{2n} \quad (6.125)$$

For all $n > N_0$, we have

$$|t_n - a| < X_n \quad (6.126)$$

Notice

$$X_n = \frac{2H}{n} - \frac{H}{n^2} + \frac{\epsilon}{2} - \frac{N_0\epsilon}{2n} \quad (6.127)$$

So we have

$$\lim_{n \rightarrow \infty} X_n = \frac{\epsilon}{2} \quad (6.128)$$

Then we know there exists some N_1 such that

$$n > N_1 \implies |X_n - \frac{\epsilon}{2}| < \frac{\epsilon}{2} \implies |t_n - a| \leq X_n < \epsilon \text{ (done)} \quad (6.129)$$

■

Question 38

5. If $a_k > 0$ for all $k \in \mathbb{N}$, prove that

$$\liminf_{k \rightarrow \infty} \frac{a_{k+1}}{a_k} \leq \liminf_{k \rightarrow \infty} \sqrt[k]{a_k} \leq \limsup_{k \rightarrow \infty} \sqrt[k]{a_k} \leq \limsup_{k \rightarrow \infty} \frac{a_{k+1}}{a_k}$$

Moreover, find a sequence $\{a_k\}_{k=1}^{\infty}$ such that

$$\limsup_{k \rightarrow \infty} \sqrt[k]{a_k} < \limsup_{k \rightarrow \infty} \frac{a_{k+1}}{a_k}$$

Proof. In book Theorem 3.37. ■

Question 39

6. If $s_1 = \sqrt{2}$, and

$$s_{n+1} = \sqrt{2 + \sqrt{s_n}}, \quad n = 1, 2, 3, \quad (6.130)$$

prove that s_n converge and bounded above by 2

Proof. We first prove $\{s_n\}$ increase monotonically by induction.

Base case:

$$s_1^2 = 2 < 2 + \sqrt{2 + \sqrt{\sqrt{2}}} = s_2^2 \implies s_1 < s_2 \quad (6.131)$$

Induction case: Let $s_k < s_{k+1}$. We wish to prove $s_{k+1} < s_{k+2}$. Because $s_{k+1} = \sqrt{2 + \sqrt{s_k}}$, we have

$$s_k \leq s_{k+1} = \sqrt{2 + \sqrt{s_k}} \quad (6.132)$$

Then we can deduce

$$s_k \leq \sqrt{2 + \sqrt{s_k}} \quad (6.133)$$

$$\implies \sqrt{s_k} \leq \sqrt{\sqrt{2 + \sqrt{s_k}}} \quad (6.134)$$

$$\implies 2 + \sqrt{s_k} \leq 2 + \sqrt{\sqrt{2 + \sqrt{s_k}}} \quad (6.135)$$

$$\implies \sqrt{2 + \sqrt{s_k}} \leq \sqrt{2 + \sqrt{\sqrt{2 + \sqrt{s_k}}}} \quad (6.136)$$

$$\implies s_{k+1} = \sqrt{2 + \sqrt{s_k}} \leq \sqrt{2 + \sqrt{s_{k+1}}} = s_{k+2} \text{ (done)} \quad (6.137)$$

We now prove $\{s_n\}$ is bounded above by 2 by induction.

Base case: $s_1 = \sqrt{2} < 2$.

Induction case: Let $s_k \leq 2$. We wish to prove $s_{k+1} \leq 2$. Observe

$$s_k \leq 2 \implies s_k \leq 4 \quad (6.138)$$

$$\implies \sqrt{s_k} \leq \sqrt{4} = 2 \quad (6.139)$$

$$\implies 2 + \sqrt{s_k} \leq 4 \quad (6.140)$$

$$\implies s_{k+1} = \sqrt{2 + \sqrt{s_k}} \leq \sqrt{4} = 2 \text{ (done)} \quad (6.141)$$

The fact that $\{s_n\}$ monotonically increase and bounded above tell us $\{s_n\}$ converge. ■

Question 40

7. Suppose $a_n > 0$ and $s_n = \sum_{k=1}^n a_k$. If s_n diverges, prove or disprove that $t_n = \sum_{k=1}^n \frac{a_k}{1+a_k}$ diverges. What can be said about

$$S_n = \sum_{k=1}^n \frac{a_k}{1 + ka_k} \quad (6.142)$$

$$T_n = \sum_{k=1}^n \frac{a_k}{1 + k^2 a_k} \quad (6.143)$$

$$\text{If } s_n = \sum_{k=1}^n a_k \text{ converge, does } J_n = \sum_{k=1}^n ka_k \text{ converge} \quad (6.144)$$

Proof. We prove

$$t_n \text{ converge} \implies s_n \text{ converge} \quad (6.145)$$

Notice that

$$a_n = \frac{a_n + a_n^2}{1 + a_n} = \frac{a_n}{1 + a_n} + \frac{a_n^2}{1 + a_n} \quad (6.146)$$

So we have

$$s_n = t_n + \sum_{k=1}^n \frac{a_n^2}{1 + a_n} \quad (6.147)$$

Because t_n converge, above tell us we only have to prove $\sum_{k=1}^n \frac{a_n^2}{1+a_n}$ converge.

Because t_n converge, we know

$$\lim_{n \rightarrow \infty} \frac{a_n}{1 + a_n} = 0 \quad (6.148)$$

Assume $\lim_{n \rightarrow \infty} a_n \neq 0$. Then there exists ϵ such that

$$\forall N \in \mathbb{N}, \exists n > N, a_n > \epsilon \quad (6.149)$$

Notice

$$a_n > \epsilon \implies \frac{a_n}{1 + a_n} = 1 - \frac{1}{1 + a_n} > 1 - \frac{1}{1 + \epsilon} \implies \left| \frac{a_n}{1 + a_n} - 0 \right| > 1 - \frac{1}{1 + \epsilon} \quad (6.150)$$

In other words, there exists a sub-sequence $\{a_{f(n)}\}$ such that

$$\lim_{n \rightarrow \infty} \frac{a_{f(n)}}{1 + a_{f(n)}} \neq 0 \text{ CaC} \quad (6.151)$$

We have proved $\lim_{n \rightarrow \infty} a_n = 0$. Then we know there exists N_1 such that

$$\forall n > N_1, a_n < 1 \quad (6.152)$$

In other words,

$$\forall n > N_1, \frac{a_n^2}{1 + a_n} < \frac{a_n}{1 + a_n} \quad (6.153)$$

By comparison test, our proof is done (done) .

We show

$$\text{It is possible } s_n \text{ diverge and } S_n \text{ converge.} \quad (6.154)$$

Let

$$a_k = \begin{cases} 1 & \text{if } \exists u \in \mathbb{N}, k = u^2 \\ \frac{1}{k^2} & \text{otherwise} \end{cases} \quad (6.155)$$

Clearly, $\lim_{n \rightarrow \infty} s_n = \infty$. Yet, we have

$$\lim_{n \rightarrow \infty} S_n = \sum_{k=1}^{\infty} \frac{a_k}{1 + ka_k} = \sum_{u=1}^{\infty} \frac{a_{u^2}}{1 + u^2 a_{u^2}} + \sum_{k \in \mathbb{N} \setminus \{r^2: r \in \mathbb{N}\}} \frac{a_k}{1 + ka_k} \quad (6.156)$$

$$= \sum_{u=1}^{\infty} \frac{1}{1 + u^2} + \sum_{k \in \mathbb{N} \setminus \{r^2: r \in \mathbb{N}\}} \frac{\frac{1}{k^2}}{1 + \frac{1}{k}} \quad (6.157)$$

$$= \sum_{u=1}^{\infty} \frac{1}{1 + u^2} + \sum_{k \in \mathbb{N} \setminus \{r^2: r \in \mathbb{N}\}} \frac{1}{k(k+1)} \quad (6.158)$$

Notice that

$$\frac{1}{1 + u^2} < \frac{1}{u^2} \text{ and } \frac{1}{k(k+1)} < \frac{1}{k^2} \quad (6.159)$$

Then by comparison test, we know

$$\text{both } \sum_{u=1}^{\infty} \frac{1}{1 + u^2} \text{ and } \sum_{k \in \mathbb{N} \setminus \{r^2: r \in \mathbb{N}\}} \frac{1}{k(k+1)} \text{ converge} \quad (6.160)$$

Then we know

$$S_n \text{ also converge (done)} \quad (6.161)$$

Notice that for all n

$$k \geq 1 \implies 1 + k^2 a_k > 1 + ka_k \implies \frac{a_k}{1 + k^2 a_k} < \frac{a_k}{1 + ka_k} \implies T_n < S_n \quad (6.162)$$

Then because the term is non-negative, by comparison test, we know the example above also satisfy T_n converge while s_n diverge.

Notice that if we let $a_k = \frac{1}{k^2}$, then s_n converge and $J_n = \sum_{k=1}^n \frac{1}{k}$ diverge.

■

Question 41

8. Assume $A \subset \mathbb{R}$ is compact and let $a \in A$. Suppose $\{a_n\}$ is a sequence in A such that every convergent sub-sequence of $\{a_n\}$ converges to a .

1. Does the sequence $\{a_n\}$ also converge to a ?
2. Without the assumption that A is compact, does the sequence $\{a_n\}$ converge to a ?

Proof. Assume a_n does not converge to a . We know there exists ϵ such that there exists a sub-sequence $\{a_{n_k}\}_{k \in \mathbb{N}}$

$$\forall k \in \mathbb{N}, a_{n_k} \notin B_\epsilon(a) \quad (6.163)$$

Because A is (sequentially) compact, we know there must exist a sub-sequence $\{a_{n_{k_u}}\}_{u \in \mathbb{N}}$ such that

$$\{a_{n_{k_u}}\} \text{ converge} \quad (6.164)$$

Notice that

$$\{a_{n_{k_u}}\} \text{ is a sub-sequence of } \{a_n\} \quad (6.165)$$

So we know

$$\lim_{u \rightarrow \infty} a_{n_{k_u}} = a \text{ CaC to } \forall n \in \mathbb{N}, a_{n_k} \neq B_\epsilon(a) \quad (6.166)$$

Let $A = \mathbb{Q}$, and let

$$a_n = \begin{cases} 0 & \text{if } n \text{ is odd} \\ n & \text{if } n \text{ is even} \end{cases} \quad (6.167)$$

and we see every convergent sub-sequence converge to 0 but a_n itself does not converge to 0 ■

Question 42

9. Suppose that $a_k \neq 0$ for large k and that

$$p = \lim_{k \rightarrow \infty} \frac{\ln\left(\frac{1}{|a_k|}\right)}{\ln(k)}$$

exists as an extended real number.

- (a) If $p > 1$, then $\sum_{k=1}^{\infty} a_k$ converges absolutely.
- (b) If $p < 1$, then $\sum_{k=1}^{\infty} a_k$ diverge

Proof. Let $p > \alpha > 1$. Then we know there exists N such that

$$\forall n > N, \frac{\ln(\frac{1}{|a_n|})}{\ln n} > \alpha > 1 \quad (6.168)$$

This give us

$$\forall n > N, \ln(\frac{1}{|a_n|}) > \alpha \ln n \quad (6.169)$$

This give us

$$\forall n > N, \frac{\ln(\frac{1}{|a_n|})}{\alpha} > \ln n \quad (6.170)$$

This give us

$$\forall n > N, \ln(|a_n|^{-\frac{1}{\alpha}}) > \ln n \quad (6.171)$$

This give us

$$\forall n > N, |a_n|^{\frac{-1}{\alpha}} > n \quad (6.172)$$

This give us

$$\forall n > N, |a_n| < n^{-\alpha} < n^{-1} \quad (6.173)$$

By comparison test, we are done.

Let $p < \beta < 1$. Then we know there exists N such that

$$\forall n > N, \frac{\ln(\frac{1}{|a_n|})}{\ln n} < \beta < 1 \quad (6.174)$$

This give us

$$\forall n > N, \frac{\ln(|a_n|^{-1})}{\beta} < \ln n \quad (6.175)$$

This give us

$$\forall n > N, \ln(|a_n|^{\frac{-1}{\beta}}) < \ln n \quad (6.176)$$

This give us

$$\forall n > N, |a_n|^{\frac{-1}{\beta}} < n \quad (6.177)$$

This give us

$$\forall n > N, |a_n| > n^{-\beta} > n^{-1} \quad (6.178)$$

By comparison test, we are done. ■

Question 43

10. Suppose that $f : \mathbb{R} \rightarrow (0, \infty)$ is differentiable, that $f(x) \rightarrow 0$ as $x \rightarrow \infty$, and that

$$\alpha = \lim_{x \rightarrow \infty} \frac{xf'(x)}{f(x)}$$

exists. If $\alpha < -1$, prove that

$$\sum_{k=1}^{\infty} f(k)$$

converges.

Proof. Let β satisfy

$$\alpha < \beta < -1 \quad (6.179)$$

Because

$$\lim_{x \rightarrow \infty} \frac{xf'(x)}{f(x)} = \alpha \quad (6.180)$$

We know there exists R such that

$$\forall x > R, \frac{xf'(x)}{f(x)} \leq \beta \quad (6.181)$$

Then we have

$$\forall x > R, \frac{d}{dx} \ln f(x) = \frac{f'(x)}{f(x)} \leq \frac{\beta}{x} \quad (6.182)$$

This tell us

$$\forall y > R, \int_R^y \frac{d}{dx} \ln f(x) dx \leq \int_R^y \frac{\beta}{x} dx \quad (6.183)$$

Which means

$$\forall y > R, \ln\left(\frac{f(y)}{f(R)}\right) = \ln f(y) - \ln f(R) \leq \beta(\ln(y) - \ln(R)) = \ln\left(\frac{y}{R}\right)^\beta \quad (6.184)$$

This shows

$$\forall y > R, \frac{f(y)}{f(R)} \leq \left(\frac{y}{R}\right)^\beta \quad (6.185)$$

It means

$$\forall y > R, f(y) \leq \frac{f(R)}{R^\beta} y^\beta \quad (6.186)$$

Notice $\beta < -1$, we know the series

$$\sum_{n>R} \frac{f(R)}{R^\beta} n^\beta = \frac{f(R)}{R^\beta} \sum_{n>R} n^\beta \text{ converge} \quad (6.187)$$

Then because $f(y)$ is always positive (given by the question), we know $\sum_{n=1}^{\infty} f(n)$ converge by comparison test. ■

Lemma 6.3.3. (Bernoulli Inequality) Let $r \geq 1$ and $x \geq -1$. We have

$$(1+x)^r \geq 1+rx \quad (6.188)$$

(6.189)

Question 44

11. Suppose that $\{a_n\}$ is a sequence of nonzero real numbers and that

$$p = \lim_{k \rightarrow \infty} k \left(1 - \left| \frac{a_{k+1}}{a_k} \right| \right)$$

exists as an extended real number. Prove that

$$\sum_{k=1}^{\infty} |a_k|$$

converges absolutely when $p > 1$.

Proof. Because ■